# Pacific Journal of Mathematics

Volume 328 No. 2 February 2024

**PACIFIC JOURNAL OF MATHEMATICS**

Founded in 1951 by E. F. Beckenbach (1906–1982) and F. Wolf (1904–1989)

msp.org/pjm



See inside back cover or msp.org/pjm for submission instructions.

The subscription price for 2024 is US $645/year for the electronic version, and $875/year for print and electronic.
Subscriptions, requests for back issues and changes of subscriber address should be sent to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163, U.S.A. The Pacific Journal of Mathematics is indexed by Mathematical Reviews, Zentralblatt MATH, PASCAL CNRS Index, Referativnyi Zhurnal, Current Mathematical Publications and Web of Knowledge (Science Citation Index).

The Pacific Journal of Mathematics (ISSN 1945-5844 electronic, 0030-8730 printed) at the University of California, c/o Department of Mathematics, 798 Evans Hall #3840, Berkeley, CA 94720-3840, is published twelve times a year. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices. POSTMASTER: send address changes to Pacific Journal of Mathematics, P.O. Box 4163, Berkeley, CA 94704-0163.

PJM peer review and production are managed by EditFlow® from Mathematical Sciences Publishers.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

PACIFIC JOURNAL OF MATHEMATICS
Vol. 328, No. 2, 2024
https://doi.org/10.2140/pjm.2024.328.193

# HANKEL OPERATORS ON $L^p(\mathbb{R}_+)$ AND THEIR $p$-COMPLETELY BOUNDED MULTIPLIERS

LORIS ARNOLD, CHRISTIAN LE MERDY AND SAFOURA ZADEH

**We show that for any $1 < p < \infty$, the space $\mathrm{Hank}_p(\mathbb{R}_+) \subseteq B(L^p(\mathbb{R}_+))$ of all Hankel operators on $L^p(\mathbb{R}_+)$ is equal to the $w^*$-closure of the linear span of the operators $\theta_u : L^p(\mathbb{R}_+) \to L^p(\mathbb{R}_+)$ defined by $\theta_u f = f(u - \cdot)$, for $u > 0$. We deduce that $\mathrm{Hank}_p(\mathbb{R}_+)$ is the dual space of $A_p(\mathbb{R}_+)$, a half-line analogue of the Figà-Talamanca–Herz algebra $A_p(\mathbb{R})$. Then we show that a function $m : \mathbb{R}_+^* \to \mathbb{C}$ is the symbol of a $p$-completely bounded multiplier $\mathrm{Hank}_p(\mathbb{R}_+) \to \mathrm{Hank}_p(\mathbb{R}_+)$ if and only if there exist $\alpha \in L^\infty(\mathbb{R}_+; L^p(\Omega))$ and $\beta \in L^\infty(\mathbb{R}_+; L^{p'}(\Omega))$ such that $m(s+t) = \langle \alpha(s), \beta(t) \rangle$ for a.e. $(s,t) \in \mathbb{R}_+^{*2}$. We also give analogues of these results in the (easier) discrete case.**

## 1. Introduction

For any $u > 0$ and for any function $f : \mathbb{R}_+ \to \mathbb{C}$, let $\tau_u f : \mathbb{R}_+ \to \mathbb{C}$ be the shifted function defined by $\tau_u f = f(\,\cdot - u)$. Let $1 < p,\, p' < \infty$ be two conjugate indices. We say that a bounded operator $T : L^p(\mathbb{R}_+) \to L^p(\mathbb{R}_+)$ is Hankelian if $\langle T\tau_u f, g\rangle = \langle Tf, \tau_u g\rangle$ for all $f \in L^p(\mathbb{R}_+)$ and $g \in L^{p'}(\mathbb{R}_+)$. Let $B(L^p(\mathbb{R}_+))$ denote the Banach space of all bounded operators on $L^p(\mathbb{R}_+)$. The main object of this paper is the subspace $\mathrm{Hank}_p(\mathbb{R}_+) \subseteq B(L^p(\mathbb{R}_+))$ of all Hankel operators on $L^p(\mathbb{R}_+)$.

The case $p = 2$ has received a lot of attention; see [Nikolski 2002; 2020; Peller 2003; Yafaev 2015; 2017a; 2017b]. The most important result in this case is that $\mathrm{Hank}_2(\mathbb{R}_+)$ is isometrically isomorphic to the quotient space $L^\infty(\mathbb{R})/H^\infty(\mathbb{R})$, where $H^\infty(\mathbb{R}) \subset L^\infty(\mathbb{R})$ is the classical Hardy space of essentially bounded functions whose Fourier transform has support in $\mathbb{R}_+$ (see [Nikolski 2020, Section IV.5.3] or [Peller 2003, Theorem I.8.1]). This result is the real line analogue of Nehari's classical theorem describing Hankel operators on $\ell^2$ (see [Nikolski 2020, Theorem II.2.2.4], [Peller 2003, Theorem I.1.1] or [Power 1982, Theorem 1.3]). An equivalent formulation of the above result is that

$$\mathrm{Hank}_2(\mathbb{R}_+) \simeq H^1(\mathbb{R})^*, \tag{1}$$

where $H^1(\mathbb{R}) \subseteq L^1(\mathbb{R})$ is the Hardy space of all integrable functions whose Fourier transform vanishes on $\mathbb{R}_-$.

*MSC2020:* primary 47B35; secondary 46L07.
*Keywords:* $p$-complete boundedness, multipliers, Hankel operators.

The first main result of this paper is that for any $1 < p < \infty$, the Banach space $\mathrm{Hank}_p(\mathbb{R}_+)$ coincides with $\overline{\mathrm{Span}}^{w^*}\{\theta_u : u > 0\} \subset B(L^p(\mathbb{R}_+))$, where, for any $u > 0$, $\theta_u : L^p(\mathbb{R}_+) \to L^p(\mathbb{R}_+)$ is the Hankel operator defined by $\theta_u f = f(u - \cdot\,)$. As a consequence, we show that

$$\mathrm{Hank}_p(\mathbb{R}_+) \simeq A_p(\mathbb{R}_+)^*, \tag{2}$$

where $A_p(\mathbb{R}_+)$ is a half-line analogue of the Figà-Talamanca–Herz algebra $A_p(\mathbb{R})$ (see, e.g., [Derighetti 2011, Chapter 3]). We will see in Remark 4.2(a) that $A_2(\mathbb{R}_+) \simeq H^1(\mathbb{R})$. Thus, the duality result (2), established in Theorem 4.1, is an $L^p$-version of (1).

By a multiplier of $\mathrm{Hank}_p(\mathbb{R}_+)$, we mean a $w^*$-continuous operator

$$T : \mathrm{Hank}_p(\mathbb{R}_+) \to \mathrm{Hank}_p(\mathbb{R}_+)$$

such that $T(\theta_u) = m(u)\theta_u$ for all $u > 0$, for some function $m : \mathbb{R}_+^* \to \mathbb{C}$. In this case, we set $T = T_m$ and it turns out that $m$ is necessarily bounded and continuous, see Lemma 4.5. The second main result of this paper is a characterization of $p$-completely bounded multipliers $T_m$. We refer to Section 2 for some background on $p$-complete boundedness, whose definition goes back to [Pisier 1990] (see also [Daws 2010; Le Merdy 1996; Pisier 2001]). We prove in Theorem 4.6 that $T_m : \mathrm{Hank}_p(\mathbb{R}_+) \to \mathrm{Hank}_p(\mathbb{R}_+)$ is a $p$-completely bounded multiplier if and only if there exist a measure space $(\Omega, \mu)$ and two essentially bounded measurable functions $\alpha : \mathbb{R}_+ \to L^p(\Omega)$ and $\beta : \mathbb{R}_+ \to L^{p'}(\Omega)$ such that $m(s+t) = \langle \alpha(s), \beta(t)\rangle$ for almost every $(s, t) \in \mathbb{R}_+^{*2}$. This is a generalisation of [Arnold et al. 2022, Theorem 3.1]. Indeed, the result in [Arnold et al. 2022] provides a characterization of $S^1$-bounded multipliers on $H^1(\mathbb{R})$. Using (1), this yields a characterization of completely bounded multipliers on $\mathrm{Hank}_2(\mathbb{R}_+)$, which is nothing but the case $p = 2$ of Theorem 4.6. See Remark 4.7 for more on this.

Let us briefly explain the plan of the paper. Section 2 contains some preliminary results. Section 3 is devoted to $\mathrm{Hank}_p(\mathbb{N}) \subset B(\ell^p)$, the space of Hankel operators on $\ell^p$. We establish analogues of the aforementioned results in the discrete setting. Results for $\mathrm{Hank}_p(\mathbb{N})$ are easier than those concerning $\mathrm{Hank}_p(\mathbb{R}_+)$ and Section 3 can be considered as a warm up. The main results are stated and proved in Section 4.

## 2. Preliminaries

All our Banach spaces are complex ones. For any Banach spaces $X$, $Z$, we let $B(X, Z)$ denote the Banach space of all bounded operators from $X$ into $Z$ and we write $B(X)$ instead of $B(X, X)$ when $Z = X$. For any $x \in X$ and $x^* \in X^*$, the duality action $x^*(x)$ is denoted by $\langle x^*, x\rangle_{X^*,X}$, or simply by $\langle x^*, x\rangle$ if there is no risk of confusion.

We start with duality on tensor products. Let $X$, $Y$ be Banach spaces. Let $X \widehat{\otimes} Y$ denote their projective tensor product [Diestel and Uhl 1977, Section VIII.1]. We will use the classical isometric identification

$$(X \widehat{\otimes} Y)^* \simeq B(X, Y^*) \tag{3}$$

provided, e.g., by [Diestel and Uhl 1977, Corollary VIII.2.2]. More precisely, for any $\xi \in (X \widehat{\otimes} Y)^*$, there exists a necessarily unique $R_\xi \in B(X, Y^*)$ such that $\xi(x \otimes y) = \langle R_\xi(x), y\rangle$ for all $x \in X$ and $y \in Y$. Moreover $\|R_\xi\| = \|\xi\|$ and the mapping $\xi \mapsto R_\xi$ is onto.

**Lemma 2.1.** *Let $A \subset X$ and $B \subset Y$ such that* Span$\{A\}$ *is dense in $X$ and* Span$\{B\}$ *is dense in $Y$. Assume that $(R_\iota)_\iota$ is a bounded net of $B(X, Y^*)$. Then $R_\iota$ converges to some $R \in B(X, Y^*)$ in the $w^*$-topology if and only if $\langle R_\iota(x), y\rangle \to \langle R(x), y\rangle$ for all $x \in A$ and $y \in B$.*

*Proof.* Assume the latter property. Since the algebraic tensor product $X \otimes Y$ is dense in $X \widehat{\otimes} Y$, it implies that $\langle R_\iota, z\rangle \to \langle R, z\rangle$, for all $z$ belonging to a dense subspace of $X \widehat{\otimes} Y$. Next, the boundedness of $(R_\iota)_\iota$ implies that $\langle R_\iota, z\rangle \to \langle R, z\rangle$, for all $z$ belonging to $X \widehat{\otimes} Y$. The equivalence follows. □

We will use the above duality principles in the case when $X = Y^*$ is an $L^p$-space $L^p(\Omega)$, for some index $1 < p < \infty$.

We now give a brief background on $p$-completely bounded maps, following [Pisier 1990] (see also [Daws 2010; Le Merdy 1996; Pisier 2001]). Let $1 < p < \infty$ and let $SQ_p$ denote the collection of quotients of subspaces of $L^p$-spaces, where we identify spaces which are isometrically isomorphic. Let $E$ be an $SQ_p$-space. Let $n \geq 1$ be an integer and let $[T_{ij}]_{1\leq i,j\leq n} \in M_n \otimes B(E)$ be an $n \times n$ matrix with entries $T_{ij}$ in $B(E)$. We equip $M_n \otimes B(E)$ with the norm defined by

$$\|[T_{ij}]\| = \sup\Big\{\Big(\sum_{i=1}^n \Big\|\sum_{j=1}^n T_{ij}(x_j)\Big\|^p\Big)^{\frac{1}{p}} : x_1, \dots, x_n \in E,\ \sum_{i=1}^n \|x_i\|^p \leq 1\Big\}. \tag{4}$$

If $S \subset B(E)$ is any subspace, then we let $M_n(S)$ denote $M_n \otimes S$ equipped with the induced norm.

Let $S_1$ and $S_2$ be subspaces of $B(E_1)$ and $B(E_2)$, respectively, for some $SQ_p$-spaces $E_1$ and $E_2$. Let $w : S_1 \to S_2$ be a linear map. For any $n \geq 1$, let $w_n : M_n(S_1) \to M_n(S_2)$ be defined by $w_n([T_{ij}]) = [w(T_{ij})]$, for any $[T_{ij}]_{1\leq i,j\leq n} \in M_n(S_1)$. By definition, $w$ is called $p$-completely bounded if the maps $w_n$ are uniformly bounded. In this case, the *$p$-cb* norm of $w$ is defined by $\|w\|_{p-cb} = \sup_n \|w_n\|$. We further say that $w$ is $p$-completely contractive if $\|w\|_{p-cb} \leq 1$ and that $w$ is a $p$-complete isometry if $w_n$ is an isometry for all $n \geq 1$. Note that the case $p = 2$ corresponds to the classical notion of completely bounded maps (see, e.g., [Paulsen 2002; Pisier 2001]).

We recall the following factorisation theorem of Pisier (see [Le Merdy 1996, Theorem 1.4; Pisier 1990, Theorem 2.1]), which extends Wittstock's factorisation theorem [Paulsen 2002, Theorem 8.4].

**Theorem 2.2.** *Let $(\Omega_1, \mu_1)$ and $(\Omega_2, \mu_2)$ be measure spaces and let $1 < p < \infty$. Let $S \subseteq B(L^p(\Omega_1))$ be a unital subalgebra. Let $w : S \to B(L^p(\Omega_2))$ be a linear map and let $C \geq 0$ be a constant. The following assertions are equivalent.*

(i) *The map $w$ is $p$-completely bounded and $\|w\|_{p-cb} \leq C$.*

(ii) *There exist an $SQ_p$-space $E$, a unital, nondegenerate $p$-completely contractive homomorphism $\pi : S \to B(E)$ as well as operators $V : L^p(\Omega_2) \to E$ and $W : E \to L^p(\Omega_2)$ such that $\|V\|\|W\| \leq C$ and for any $x \in S$, $w(x) = W\pi(x)V$.*

**Remark 2.3.** Let $1 < p < \infty$ and let $p'$ be its conjugate index. Let $E$ be an $SQ_p$-space. Then by assumption, there exist a measure space $(\Omega, \mu)$ and two closed subspaces $E_2 \subseteq E_1 \subseteq L^p(\Omega)$ such that $E = E_1/E_2$. Then $E_1^\perp \subseteq E_2^\perp \subseteq L^{p'}(\Omega)$ and we have an isometric identification

$$E^* \simeq \frac{E_2^\perp}{E_1^\perp}, \tag{5}$$

by the classical duality between subspaces and quotients of Banach spaces. More explicitly, let $f \in E_1$ and let $g \in E_2^\perp$. Let $\dot{f} \in E$ denote the class of $f$ modulo $E_2$ and let $\dot{g} \in E^*$ denote the element associated to the class of $g$ modulo $E_1^\perp$ through the identification (5). Then we have

$$\langle \dot{g}, \dot{f} \rangle_{E^*, E} = \langle g, f \rangle_{L^{p'}, L^p}. \tag{6}$$

We now turn to Bochner spaces. Let $(\Sigma, \nu)$ be a measure space and let $X$ be a Banach space. For any $1 \leq p \leq \infty$, we let $L^p(\Sigma; X)$ denote the space of all measurable functions $\phi : \Sigma \to X$ (defined up to almost everywhere zero functions) such that the norm function $t \mapsto \|\phi(t)\|$ belongs to $L^p(\Sigma)$. This is a Banach space for the norm $\|\phi\|_p$, defined as the $L^p(\Sigma)$-norm of $\|\phi(\,\cdot\,)\|$ (see, e.g., [Diestel and Uhl 1977, Chapters I and II]).

Assume that $p$ is finite and note that in this case, $L^p(\Sigma) \otimes X$ is dense in $L^p(\Sigma; X)$. Let $p'$ be the conjugate index of $p$. For all $\phi \in L^p(\Sigma; X)$ and $\psi \in L^{p'}(\Sigma; X^*)$, the function $t \mapsto \langle \psi(t), \phi(t) \rangle_{X^*, X}$ belongs to $L^1(\Sigma)$ and the resulting duality paring $\langle \psi, \phi \rangle := \int_\Omega \langle \psi(t), \phi(t) \rangle \, d\nu(t)$ extends to an isometric embedding $L^{p'}(\Sigma; X^*) \hookrightarrow L^p(\Sigma; X)^*$. Furthermore, this embedding is onto if $X$ is reflexive, that is,

$$L^{p'}(\Sigma; X^*) \simeq L^p(\Sigma; X)^* \quad \text{if } X \text{ is reflexive}. \tag{7}$$

We refer to [Diestel and Uhl 1977, Corollary III.2.13 and Section IV.1] for these results and complements.

Let $(\Sigma, \nu)$ and $(\Omega, \mu)$ be two measure spaces. Then we have an isometric identification

$$L^p(\Sigma; L^p(\Omega)) \simeq L^p(\Sigma \times \Omega),$$

from which it follows that for any $T \in B(L^p(\Sigma))$, the tensor extension

$$T \otimes I_{L^p(\Omega)} : L^p(\Sigma) \otimes L^p(\Omega) \to L^p(\Sigma) \otimes L^p(\Omega)$$

extends to a bounded operator $T \overline{\otimes} I_{L^p(\Omega)}$ on $L^p(\Sigma \times \Omega)$, whose norm is equal to the norm of $T$. The following is elementary.

**Lemma 2.4.** *The mapping $\pi : B(L^p(\Sigma)) \to B(L^p(\Sigma \times \Omega))$ defined by $\pi(T) = T \overline{\otimes} I_{L^p(\Omega)}$ is a $p$-complete isometry.*

*Proof.* Let $n \geq 1$ and let $J_n = \{1, \dots, n\}$. It follows from (4) that $M_n(B(L^p(\Sigma))) = B(\ell_n^p(L^p(\Sigma)))$ and hence $M_n(B(L^p(\Sigma))) = B(L^p(J_n \times \Sigma))$ isometrically. Likewise, we have $M_n(B(L^p(\Sigma \times \Omega))) = B(L^p(J_n \times \Sigma \times \Omega))$ isometrically. Through these identifications,

$$[T_{ij} \overline{\otimes} I_{L^p(\Omega)}] = [T_{ij}] \overline{\otimes} I_{L^p(\Omega)},$$

for all $[T_{ij}]_{1 \leq i,j \leq n}$ in $M_n(B(L^p(\Sigma)))$. The result follows at once. □

We finally state an important result concerning Schur products on $B(\ell_I^p)$-spaces. Let $I$ be an index set, let $1 < p < \infty$ and let $\ell_I^p$ denote the discrete $L^p$-space over $I$. Let $(e_t)_{t \in I}$ be its canonical basis. To any $T \in B(\ell_I^p)$, we associate a matrix of complex numbers, $[a_{st}]_{s,t \in I}$, defined by $a_{st} = \langle T(e_t), e_s \rangle$, for all $s, t \in I$. Following [Pisier 2001, Chapter 5], we say that a bounded family $\{\varphi(s,t)\}_{(s,t) \in I^2}$ of complex numbers is a bounded Schur multiplier on $B(\ell_I^p)$ if for all $T \in B(\ell_I^p)$, with matrix $[a_{st}]_{s,t \in I}$, the matrix $[\varphi(s,t) a_{st}]_{s,t \in I}$ represents an element of $B(\ell_I^p)$. In this case, the mapping $[a_{st}] \to [\varphi(s,t) a_{st}]$ is a bounded operator from $B(\ell_I^p)$ into itself. We note that $\{\varphi(s,t)\}_{(s,t) \in I^2}$ is a bounded Schur multiplier with norm $\leq C$ if and only if for all $n \geq 1$, all $[a_{ij}]_{1 \leq i,j \leq n}$ in $M_n$ and all $t_1, \dots, t_n, s_1, \dots, s_n$ in $I$, we have

$$\|[\varphi(s_i, t_j) a_{ij}]\|_{B(\ell_n^p)} \leq C \|[a_{ij}]\|_{B(\ell_n^p)}. \tag{8}$$

In the sequel, we apply the above definitions to the case when $I = \mathbb{R}_+^*$.

**Theorem 2.5.** *Let $\varphi : \mathbb{R}_+^{*2} \to \mathbb{C}$ be a continuous bounded function. Let $1 < p, p' < \infty$ be conjugate indices and let $C \geq 0$ be a constant. The following assertions are equivalent.*

(i) *The family $\{\varphi(s,t)\}_{(s,t) \in \mathbb{R}_+^{*2}}$ is a bounded Schur multiplier on $B(\ell_{\mathbb{R}_+^*}^p)$, with norm $\leq C$.*

(ii) *There exist a measure space $(\Omega, \mu)$ and two functions $\alpha \in L^\infty(\mathbb{R}_+; L^p(\Omega))$ and $\beta \in L^\infty(\mathbb{R}_+; L^{p'}(\Omega))$ such that $\|\alpha\|_\infty \|\beta\|_\infty \leq C$ and $\varphi(s,t) = \langle \alpha(s), \beta(t) \rangle_{L^p, L^{p'}}$ for almost every $(s,t) \in \mathbb{R}_+^{*2}$.*

*Proof.* According to [Coine 2018, Section 4.1], (ii) is equivalent to the fact that as an element of $L^\infty(\mathbb{R}_+^2)$,

(ii') $\varphi$ is a bounded Schur multiplier on $B(L^p(\mathbb{R}_+))$.

It further follows from [Herz 1974, Lemmas 1 and 2] that since $\varphi$ is continuous, (ii') is equivalent to (i). The result follows. □

## 3. Hankel operators on $\ell^p$ and their multipliers

In this section we work on the sequence spaces $\ell^p = \ell^p_{\mathbb{N}}$, where $\mathbb{N} = \{0, 1, \dots\}$. For any $1 < p < \infty$, we let $(e_n)_{\geq 0}$ denote the classical basis of $\ell^p$. For any $T \in B(\ell^p)$, the associated matrix $[t_{ij}]_{i,j\geq 0}$ is given by $t_{ij} = \langle T(e_j), e_i\rangle$, for all $i, j \geq 0$.

Let $\mathrm{Hank}_p(\mathbb{N}) \subseteq B(\ell^p)$ be the subspace of all $T \in B(\ell^p)$ whose matrix is Hankelian, i.e., has the form $[c_{i+j}]_{i,j\geq 0}$ for some sequence $(c_k)_{k\geq 0}$ of complex numbers.

Let $p'$ be the conjugate index of $p$ and regard $\ell^p \otimes \ell^{p'} \subset B(\ell^p)$ in the usual way. We set

$$\gamma_k = \sum_{i+j=k} e_i \otimes e_j$$

for any $k \geq 0$. Then each $\gamma_k$ belongs to $\mathrm{Hank}_p(\mathbb{N})$, and $\|\gamma_k\| = 1$. Indeed, the matrix of $\gamma_k$ is $[c_{i+j}]_{i,j\geq 0}$ with $c_k = 1$ and $c_l = 0$ for all $l \neq k$.

**Lemma 3.1.** *For any* $1 < p < \infty$, *the space* $\mathrm{Hank}_p(\mathbb{N})$ *is the* $w^*$*-closure of the linear span of* $\{\gamma_k : k \geq 0\}$.

*Proof.* It is plain that $\mathrm{Hank}_p(\mathbb{N})$ is a $w^*$-closed subspace of $B(\ell^p)$, hence one inclusion is straightforward.

To check the other one, consider $T \in \mathrm{Hank}_p(\mathbb{N})$. By the definition of this space, there is a sequence $(c_k)_{k\geq 0}$ of $\mathbb{C}$ such that

$$\langle T(e_j), e_i\rangle = c_{i+j}, \quad \text{for all } i, j \geq 0.$$

For any $n \geq 1$, let $K_n$ be the Fejér kernel defined by

$$K_n(t) = \sum_{k=-n}^{n} \Big(1 - \frac{|k|}{n}\Big) e^{int}, \quad t \in \mathbb{R}.$$

Then let $T_n \in B(\ell^p)$ be the finite rank operator whose matrix is $[\widehat{K_n}(i+j)c_{i+j}]_{i,j\geq 0}$. Note that

$$T_n = \sum_{k=0}^{n} \Big(1 - \frac{|k|}{n}\Big) c_k\, \gamma_k \in \mathrm{Span}\{\gamma_k : k \geq 0\}.$$

We show that $\|T_n\| \leq \|T\|$. To see this, let $\alpha = (\alpha_j)_{j\geq 0} \in \ell^p$ and $(\beta_m)_{m\geq 0} \in \ell^{p'}$. We have that

$$\begin{aligned}\langle T_n(\alpha), \beta\rangle &= \sum_{m,j\geq 0} \widehat{K_n}(m+j)c_{m+j}\alpha_j\beta_m \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(t) \sum_{m,j\geq 0} c_{m+j}\alpha_j\beta_m e^{-i(m+j)t}\, dt.\end{aligned}$$

Since $K_n \geq 0$, we deduce

$$|\langle T_n(\alpha), \beta\rangle| \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(t) \Big|\sum_{m,j\geq 0} c_{m+j}\alpha_j\beta_m e^{-i(m+j)t}\Big|\, dt.$$

Now for all $t \in [-\pi, \pi]$, we have

$$\begin{aligned}\Big|\sum_{m,j\geq 0} c_{m+j}\alpha_j\beta_m e^{-i(m+j)t}\Big| &= \Big|\sum_{m,j\geq 0} c_{m+j}(e^{-ijt}\alpha_j)(e^{-imt}\beta_m)\Big| \\ &= \Big|\Big\langle T((e^{-ijt}\alpha_j)_{j\geq 0}), (e^{-imt}\beta_m)_{m\geq 0}\Big\rangle\Big| \\ &\leq \|T\|\Big(\sum_{j\geq 0} |e^{-ijt}\alpha_j|^p\Big)^{\frac{1}{p}}\Big(\sum_{m\geq 0} |e^{-imt}\beta_m|^{p'}\Big)^{\frac{1}{p'}} \\ &\leq \|T\|\|\alpha\|_p\|\beta\|_{p'},\end{aligned}$$

Since

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} K_n(t)dt = 1,$$

we therefore obtain that $|\langle T_n(\alpha), \beta\rangle| \leq \|T\|\|\alpha\|_p\|\beta\|_{p'}$. This proves that $\|T_n\| \leq \|T\|$, as requested.

For all $i, j \geq 0$,

$$\langle T_n(e_j), e_i\rangle = \widehat{K_n}(i+j)\,\langle Te_j, e_i\rangle \to \langle Te_j, e_i\rangle,$$

when $n \to \infty$. Hence $T_n \to T$ in the $w^*$-topology, by Lemma 2.1. Consequently, $T$ belongs to the $w^*$-closure of $\mathrm{Span}\{\gamma_k : k \geq 0\}$. □

**Remark 3.2.** (a) Nehari's celebrated theorem (see, e.g., [Nikolski 2020, Theorem II.2.2.4], [Peller 2003, Theorem I.1.1] or [Power 1982, Theorem 1.3]) asserts that

$$\mathrm{Hank}_2(\mathbb{N}) \simeq \frac{L^\infty(\mathbb{T})}{H^\infty(\mathbb{T})}. \tag{9}$$

Here $\mathbb{T}$ stands for the unit circle of $\mathbb{C}$ and $H^\infty(\mathbb{T}) \subset L^\infty(\mathbb{T})$ is the Hardy space of functions whose negative Fourier coefficients vanish. The isometric isomorphism $J : L^\infty(\mathbb{T})/H^\infty(\mathbb{T}) \to \mathrm{Hank}_2(\mathbb{N})$ providing (9) is defined as follows. Given any $F \in L^\infty(\mathbb{T})$, let $\dot{F}$ denote its class modulo $H^\infty(\mathbb{T})$. Then $J(\dot{F})$ is the operator whose matrix is equal to $[\widehat{F}(-i-j-1)]_{i,j\geq 0}$.

(b) We remark that $\mathrm{Hank}_p(\mathbb{N}) \subseteq \mathrm{Hank}_2(\mathbb{N})$. To see this, note that if $T \in \mathrm{Hank}_p(\mathbb{N})$, then because of the symmetry in its matrix representation due to being a Hankelian matrix, $T$ has the same matrix representation as $T^*$, and therefore $T$ extends to a bounded operator on $\ell^{p'}$. By interpolation, $T$ extends to a bounded operator on $\ell^2$, which is represented by the same matrix as $T$. Hence, $T$ belongs to $\mathrm{Hank}_2(\mathbb{N})$.

However for $1 < p \neq 2 < \infty$, there is no description of $\mathrm{Hank}_p(\mathbb{N})$ similar to Nehari's theorem.

(c) The definition of $\mathrm{Hank}_p(\mathbb{N})$ readily extends to the case $p = 1$ isometrically:

$$\mathrm{Hank}_1(\mathbb{N}) \simeq \ell^1.$$

Indeed, let $J_1 : \ell^1 \to \mathrm{Hank}_1(\mathbb{N})$ be defined by

$$J_1(c) = \sum_{k=0}^{\infty} c_k \gamma_k, \quad c = (c_k)_{k\geq 0} \in \ell^1.$$

Next, let $J_2 : \mathrm{Hank}_1(\mathbb{N}) \to \ell^1$ be defined by $J_2(T) = T(e_0)$. Then $J_1$, $J_2$ are contractions and it is easy to check that they are inverse to each other. Hence $J_1$ is an isometric isomorphism.

We say that a sequence $m = (m_k)_{k\geq 0}$ in $\mathbb{C}$ is the symbol of a multiplier on $\mathrm{Hank}_p(\mathbb{N})$ if there is a $w^*$-continuous operator $T_m : \mathrm{Hank}_p(\mathbb{N}) \to \mathrm{Hank}_p(\mathbb{N})$ such that

$$T_m(\gamma_k) = m_k \gamma_k, \quad k \geq 0.$$

Such an operator is uniquely defined. In this case, $m \in \ell^\infty$ and $\|m\|_\infty \leq \|T_m\|$.

The following is a simple extension of [Pisier 2001, Theorems 6.1 and 6.2].

**Theorem 3.3.** *Let $1 < p < \infty$, let $C \geq 0$ be a constant and let $m = (m_k)_{k\geq 0}$ be a sequence in $\mathbb{C}$. The following assertions are equivalent.*

(i) *$m$ is the symbol of a $p$-completely bounded multiplier on* $\mathrm{Hank}_p(\mathbb{N})$, *and*

$$\|T_m : \mathrm{Hank}_p(\mathbb{N}) \to \mathrm{Hank}_p(\mathbb{N})\|_{p-cb} \leq C.$$

(ii) *There exist a measure space $(\Omega, \mu)$, and bounded sequences $(\alpha_i)_{i\geq 0}$ in $L^p(\Omega)$ and $(\beta_j)_{j\geq 0}$ in $L^{p'}(\Omega)$ such that $m_{i+j} = \langle \alpha_i, \beta_j \rangle$, for every $i, j \geq 0$, and*

$$\sup_{i\geq 0} \|\alpha_i\|_p \sup_{j\geq 0} \|\beta_j\|_{p'} \leq C.$$

*Proof.* By homogeneity, we may assume that $C = 1$ throughout this proof.

Assume (i). Let $\kappa : \ell^p_{\mathbb{Z}} \to \ell^p_{\mathbb{Z}}$ be defined by $\kappa((a_k)_{k\in\mathbb{Z}}) = (a_{-k})_{k\in\mathbb{Z}}$, let $J : \ell^p_{\mathbb{N}} \to \ell^p_{\mathbb{Z}}$ be the canonical embedding and let $Q : \ell^p_{\mathbb{Z}} \to \ell^p_{\mathbb{N}}$ be the canonical projection. Define $q : B(\ell^p_{\mathbb{Z}}) \to B(\ell^p_{\mathbb{N}})$ by $q(T) = Q\kappa T J$. According to the easy implication (ii) $\Longrightarrow$ (i) of Theorem 2.2, the mapping $q$ is $p$-completely contractive. We note that if $[t_{i,j}]_{(i,j)\in\mathbb{Z}^2}$ is the matrix of some $T \in B(\ell^p_{\mathbb{Z}})$, then the matrix of $q(T)$ is equal to $[t_{-i,j}]_{(i,j)\in\mathbb{N}^2}$.

Let $\mathcal{M}_p(\mathbb{Z}) \subseteq B(\ell^p_{\mathbb{Z}})$ be the space of all bounded Fourier multipliers on $\ell^p_{\mathbb{Z}}$; this is a unital subalgebra. Let $T \in \mathcal{M}_p(\mathbb{Z})$ and let $\phi \in L^\infty(\mathbb{T})$ denote its symbol. Then the matrix of $T$ is equal to $[\widehat{\phi}(i-j)]_{(i,j)\in\mathbb{Z}^2}$, hence the matrix of $q(T)$ is equal to $[\widehat{\phi}(-i-j)]_{(i,j)\in\mathbb{N}^2}$. Hence, $q(T)$ is Hankelian. We can therefore consider the restriction map

$$q_{|\mathcal{M}_p(\mathbb{Z})} : \mathcal{M}_p(\mathbb{Z}) \to \mathrm{Hank}_p(\mathbb{N}).$$

Let $\mathtt{s} : \ell^p_{\mathbb{Z}} \to \ell^p_{\mathbb{Z}}$ be the shift operator defined by $\mathtt{s}(e_j) = e_{j+1}$, for all $j \in \mathbb{Z}$. We observe (left to the reader) that

$$q(\mathtt{s}^{-k}) = \gamma_k, \quad k \in \mathbb{N}. \tag{10}$$

We assume that $T_m : \mathrm{Hank}_p(\mathbb{N}) \to \mathrm{Hank}_p(\mathbb{N})$ is $p$-completely contractive. Consider $w : \mathcal{M}_p(\mathbb{Z}) \to \mathrm{Hank}_p(\mathbb{N}) \subseteq B(\ell^p)$ defined by $w := T_m \circ q_{|\mathcal{M}_p(\mathbb{Z})}$. Then $w$ is $p$-completely contractive. Applying Theorem 2.2 to $w$, we obtain an $SQ_p$-space $E$, a contractive homomorphism $\pi : \mathcal{M}_p(\mathbb{Z}) \to B(E)$ and contractive maps $V : \ell^p_{\mathbb{N}} \to E$ and $W : E \to \ell^p_{\mathbb{N}}$ such that

$$w(T) = W\pi(T)V, \quad T \in \mathcal{M}_p(\mathbb{Z}). \tag{11}$$

Let $i, j \geq 0$. By (10), we have

$$w(\mathtt{s}^{-(i+j)}) = T_m(q(\mathtt{s}^{-(i+j)})) = T_m(\gamma_{i+j}) = m_{i+j}\gamma_{i+j},$$

hence $\langle w(\mathtt{s}^{-(i+j)})e_i, e_j\rangle = m_{i+j}$. Consequently, from (11), we obtain that

$$m_{i+j} = \langle \pi(\mathtt{s}^{-(i+j)})V(e_i), W^*(e_j)\rangle_{E,E^*}.$$

The mapping $\pi$ is multiplicative, hence this implies that

$$m_{i+j} = \langle \pi(\mathtt{s}^{-i})V(e_i), \pi(\mathtt{s}^{-j})^* W^*(e_j)\rangle_{E,E^*}.$$

Set $x_i := \pi(\mathtt{s}^{-i})V(e_i) \in E$ and $y_j := \pi(\mathtt{s}^{-j})^*W^*(e_j) \in E^*$. Then, for all $i, j \geq 0$ we have $\|x_i\| \leq 1$, $\|y_j\| \leq 1$ and $m_{i+j} = \langle x_i, y_j\rangle_{E,E^*}$.

Let us now apply Remark 2.3. As in the latter, consider a measure space $(\Omega, \mu)$ and closed subspaces $E_2 \subset E_1 \subset L^p(\Omega)$ such that $E = E_1/E_2$. Recall (5). For any $i \geq 0$, pick $\alpha_i \in E_1$ such that $\|\alpha_i\|_p = \|x_i\|$ and $\dot{\alpha_i} = x_i$. Likewise, for any $j \geq 0$, pick $\beta_j \in E_2^\perp$ such that $\|\beta_j\|_{p'} = \|y_j\|$ and $\dot{\beta_j} = y_j$. Then for all $i, j \geq 0$, we both have $\|\alpha_i\|_p \leq 1$, $\|\beta_j\|_{p'} \leq 1$ and $m_{i+j} = \langle \alpha_i, \beta_j\rangle_{L^p, L^{p'}}$. This proves (ii).

Conversely, assume (ii). By [Pisier 2001, Corollary 8.2], the family $\{m_{i+j}\}_{(i,j)\in\mathbb{N}^2}$ induces a $p$-completely contractive Schur multiplier on $B(\ell^p)$. It is clear that the restriction of this Schur multiplier maps $\mathrm{Hank}_p(\mathbb{N})$ into itself. More precisely, it maps $\gamma_k$ to $m_k\gamma_k$ for all $k \geq 0$. Hence $m$ is the symbol of a $p$-completely contractive multiplier on $\mathrm{Hank}_p(\mathbb{N})$. □

## 4. Hankel operators on $L^p(\mathbb{R}_+)$

Throughout we let $1 < p < \infty$ and we let $p'$ denote its conjugate index. For any $u > 0$, we set $\tau_u f := f(\,\cdot - u)$, for all $f \in L^1(\mathbb{R}) + L^\infty(\mathbb{R})$. Let

$$\mathrm{Hank}_p(\mathbb{R}_+) \subseteq B(L^p(\mathbb{R}_+))$$

be the space of Hankelian operators on $L^p(\mathbb{R}_+)$, consisting of all bounded operators $T : L^p(\mathbb{R}_+) \to L^p(\mathbb{R}_+)$ such that

$$\langle T\tau_u f, g\rangle = \langle Tf, \tau_u g\rangle,$$

for all $f \in L^p(\mathbb{R}_+)$, $g \in L^{p'}(\mathbb{R}_+)$ and $u > 0$.

For any $u > 0$, let $\theta_u : L^p(\mathbb{R}_+) \to L^p(\mathbb{R}_+)$ be defined by $\theta_u f = f(u - \cdot\,)$. Note that $\theta_u$ is a Hankelian operator on $L^p(\mathbb{R}_+)$. Indeed, for all $f \in L^p(\mathbb{R}_+)$, $g \in L^{p'}(\mathbb{R}_+)$ and $v > 0$, we have

$$\langle \theta_u \tau_v f, g\rangle = \int_v^u f(u - s) g(s - v)\, ds = \langle \theta_u f, \tau_v g\rangle$$

if $v < u$, and $\langle \theta_u \tau_v f, g\rangle = \langle \theta_u f, \tau_v g\rangle = 0$ if $v \geq u$. The operators $\theta_u$ are the continuous counterparts of the operators $\gamma_k$ from Section 3. From this point of view, part (1) of Theorem 4.1 below is an analogue of Lemma 3.1. However its proof is more delicate.

We introduce a new space $A_p(\mathbb{R}_+) \subseteq C_0(\mathbb{R}_+)$ by

$$A_p(\mathbb{R}_+) := \Big\{ F = \sum_{n=1}^{\infty} f_n * g_n : f_n \in L^p(\mathbb{R}_+),\ g_n \in L^{p'}(\mathbb{R}_+) \text{ and } \sum_{n=1}^{\infty} \|f_n\|_p \|g_n\|_{p'} < \infty \Big\},$$

and we equip it with the norm

$$\|F\|_{A_p} = \inf\Big\{ \sum_{n=1}^{\infty} \|f_n\|_p \|g_n\|_{p'} \Big\}, \tag{12}$$

where the infimum runs over all possible representations of $F$ as above. The space $A_p(\mathbb{R}_+)$ is a half-line analogue of the classical Figà-Talamanca–Herz algebra $A_p(\mathbb{R})$; see, e.g., [Derighetti 2011]. The classical arguments showing that the latter is a Banach space show as well that (12) is a norm on $A_p(\mathbb{R}_+)$ and that $A_p(\mathbb{R}_+)$ is a Banach space.

It follows from the above definitions that there exists a (necessarily unique) contractive map

$$Q_p : L^p(\mathbb{R}_+) \mathbin{\widehat{\otimes}} L^{p'}(\mathbb{R}_+) \to A_p(\mathbb{R}_+)$$

such that $Q_p(f \otimes g) = f * g$, for all $f \in L^p(\mathbb{R}_+)$ and $g \in L^{p'}(\mathbb{R}_+)$. Moreover $Q_p$ is a quotient map. Hence the adjoint

$$Q_p^* : A_p(\mathbb{R}_+)^* \to B(L^p(\mathbb{R}_+))$$

of $Q_p$ is an isometry. This yields an isometric identification $A_p(\mathbb{R}_+)^* \simeq \ker(Q_p)^\perp$ $(= \operatorname{ran}(Q_p^*))$.

We observe that

$$\ker(Q_p)^\perp = \overline{\operatorname{Span}}^{w^*}\{\theta_u : u > 0\}. \tag{13}$$

To prove this, we note that

$$\langle \theta_u, f \otimes g\rangle = \langle \theta_u(f), g\rangle = (f * g)(u), \tag{14}$$

for all $f \in L^p(\mathbb{R})$, $g \in L^{p'}(\mathbb{R}_+)$ and $u > 0$. Hence,

$$\Big\langle \theta_u, \sum_{n=1}^\infty f_n \otimes g_n\Big\rangle = \Big(\sum_{n=1}^\infty f_n * g_n\Big)(u)$$

for all sequences $(f_n)_n$ in $L^p(\mathbb{R}_+)$ and $(g_n)_n$ in $L^{p'}(\mathbb{R}_+)$ such that

$$\sum_{n=1}^\infty \|f_n\|_p \|g_n\|_{p'} < \infty,$$

and all $u > 0$. This implies that $\operatorname{Span}\{\theta_u : u > 0\}_\perp = \ker(Q_p)$, and (13) follows.

**Theorem 4.1.** (1) *The space* $\operatorname{Hank}_p(\mathbb{R}_+)$ *is equal to the* $w^*$*-closure of the linear span of* $\{\theta_u : u > 0\}$.

(2) *We have an isometric identification*

$$\operatorname{Hank}_p(\mathbb{R}_+) \simeq A_p(\mathbb{R}_+)^*.$$

*Proof.* Part (2) follows from part (1) and the discussion preceding the statement of Theorem 4.1. For any $f \in L^p(\mathbb{R}_+)$, $g \in L^{p'}(\mathbb{R}_+)$ and $u > 0$, the functionals $T \mapsto \langle T\tau_u f, g\rangle$ and $T \mapsto \langle Tf, \tau_u g\rangle$ are $w^*$-continuous on $B(L^p(\mathbb{R}_+))$. Consequently, $\operatorname{Hank}_p(\mathbb{R}_+)$ is $w^*$-closed. Hence $\operatorname{Hank}_p(\mathbb{R}_+)$ contains the $w^*$-closure of $\operatorname{Span}\{\theta_u : u > 0\}$. To prove the reverse inclusion, it suffices to show, by (13), that

$$\operatorname{Hank}_p(\mathbb{R}_+) \subset \ker(Q_p)^\perp.$$

We will use a double approximation process. First, let $k, l$ in $C_c(\mathbb{R})$, the space of continuous functions with compact support. To any $T \in B(L^p(\mathbb{R}_+))$, we associate $T_{k,l} \in B(L^p(\mathbb{R}_+))$ defined by

$$\langle T_{k,l}(f), g\rangle = \int_{\mathbb{R}} \langle T(\tau_u k \cdot f), \tau_{-u} l \cdot g\rangle\, du, \quad f \in L^p(\mathbb{R}_+),\ g \in L^{p'}(\mathbb{R}_+).$$

We note that

$$\begin{aligned}\int_{\mathbb{R}} \big|\langle T(\tau_u k \cdot f), \tau_{-u} l \cdot g\rangle\big|\, du &\le \|T\|_p \Big(\int_{\mathbb{R}} \|\tau_u k f\|_p^p\, du\Big)^{\frac{1}{p}} \Big(\int_{\mathbb{R}} \|\tau_{-u} l g\|_{p'}^{p'}\, du\Big)^{\frac{1}{p'}}\\ &= \|T\|_p \|f\|_p \|g\|_{p'} \|k\|_p \|l\|_{p'}.\end{aligned}$$

Thus, $T_{k,l}$ is well-defined and $\|T_{k,l}\| \leq \|T\|\|k\|_p\|l\|_{p'}$. We are going to show that

$$T \in \mathrm{Hank}_p(\mathbb{R}_+) \implies T_{k,l} \in \ker(Q_p)^{\perp}. \tag{15}$$

Let $\alpha \in C_c(\mathbb{R}_+)^+$ such that $\|\alpha\|_1 = 1$. Let $R_\alpha \in B(L^p(\mathbb{R}_+))$ be defined by

$$R_\alpha(f) = \alpha * f, \quad f \in L^p(\mathbb{R}_+).$$

We show that $(TR_\alpha)_{k,l}$ belongs to $\ker(Q_p)^\perp$ if $T \in \mathrm{Hank}_p(\mathbb{R}_+)$, and we use these auxiliary operators to establish (15).

We fix some $T \in \mathrm{Hank}_p(\mathbb{R}_+)$. Let $z \in \ker(Q_p)$. Since $C_c(\mathbb{R}_+)$ is both dense in $L^p(\mathbb{R}_+)$ and $L^{p'}(\mathbb{R}_+)$, it follows, e.g., from [Derighetti 2011, Chapter 3, Proposition 6] that there exist sequences $(f_n)_{n\geq 1}$ and $(g_n)_{n\geq 1}$ in $C_c(\mathbb{R}_+)$ such that $\sum_{n=1}^{\infty} \|f_n\|_p\|g_n\|_{p'} < \infty$ and $z = \sum_{n=1}^{\infty} f_n \otimes g_n$. Since $z \in \ker(Q_p)$, we have $\sum_{n=1}^{\infty} f_n * g_n = 0$, pointwise.

We write $R_\alpha f = \int_{\mathbb{R}_+} f(s)\tau_s\alpha\, ds$ as a Bochner integral, for all $f \in C_c(\mathbb{R}_+)$. A simple application of Fubini's theorem leads to

$$k * l \cdot f_n * g_n = \int_{\mathbb{R}}\int_{\mathbb{R}_+} (\tau_u k \cdot f_n)(s)\tau_s(\tau_{-u} l \cdot g_n)\, ds du,$$

for all $n \geq 1$. We deduce that

$$\begin{aligned}
\sum_{n=1}^{\infty} \langle (TR_\alpha)_{k,l}(f_n), g_n\rangle &= \sum_{n=1}^{\infty} \int_{\mathbb{R}} \langle TR_\alpha(\tau_u k \cdot f_n), \tau_{-u} l \cdot g_n\rangle\, du \\
&= \sum_{n=1}^{\infty} \int_{\mathbb{R}} \langle T((\tau_u k \cdot f_n) * \alpha), \tau_{-u} l \cdot g_n\rangle\, du \\
&= \sum_{n=1}^{\infty} \int_{\mathbb{R}}\int_{\mathbb{R}_+} (\tau_u k \cdot f_n)(s)\langle T(\tau_s\alpha), \tau_{-u} l \cdot g_n\rangle\, ds du \\
&= \sum_{n=1}^{\infty} \int_{\mathbb{R}}\int_{\mathbb{R}_+} \langle T(\alpha), (\tau_u k \cdot f_n)(s)\tau_s(\tau_{-u} l \cdot g_n)\rangle\, ds du \\
&= \sum_{n=1}^{\infty} \langle T(\alpha), k * l \cdot f_n * g_n\rangle \\
&= \Big\langle T(\alpha), k * l \cdot \sum_{n=1}^{\infty} f_n * g_n \Big\rangle = 0.
\end{aligned}$$

This shows that $(TR_\alpha)_{k,l}$ belongs to $\ker(Q_p)^\perp$.

For $z$, $f_n$, $g_n$ as above, write

$$\sum_{n=1}^{\infty} \langle T_{k,l}(f_n), g_n\rangle = \sum_{n=1}^{\infty} \langle T_{k,l}(f_n), g_n\rangle - \sum_{n=1}^{\infty} \langle (TR_\alpha)_{k,l}(f_n), g_n\rangle.$$

Then we have

$$
\begin{aligned}
&\Big|\sum_{n=1}^{\infty}\langle T_{k,l}(f_n),\, g_n\rangle\Big| \\
&\qquad\le \sum_{n=1}^{\infty}\int_{\mathbb{R}}\big|\langle T(\tau_u k\cdot f_n-(\tau_u k\cdot f_n)*\alpha),\, \tau_{-u}l\cdot g_n\rangle\big|\,du \\
&\qquad\le \sum_{n=1}^{\infty}\|T\|\Big(\int_{\mathbb{R}}\|\tau_u k\cdot f_n-(\tau_u k\cdot f_n)*\alpha\|_p^p\,du\Big)^{\frac{1}{p}}\Big(\int_{\mathbb{R}}\|\tau_{-u}l\cdot g_n\|_{p'}^{p'}\,du\Big)^{\frac{1}{p'}} \\
&\qquad\le \|T\|\|l\|_{p'}\sum_{n=1}^{\infty}\|g_n\|_{p'}\Big(\int_{\mathbb{R}}\|\tau_u k\cdot f_n-\Big(\tau_u k\cdot f_n\Big)*\alpha\|_p^p\,du\Big)^{\frac{1}{p}}.
\end{aligned}
$$

Recall that by assumption, $\alpha\ge 0$ and $\int_{\mathbb{R}_+}\alpha(s)ds=1$. Then we deduce from above that

$$
\begin{aligned}
&\Big|\sum_{n=1}^{\infty}\langle T_{k,l}(f_n),\, g_n\rangle\Big| \\
&\qquad\le \|T\|\|l\|_{p'}\sum_{n=1}^{\infty}\|g_n\|_{p'}\Big(\int_{\mathbb{R}}\Big\|\int_{\mathbb{R}_+}\alpha(s)(\tau_u k\cdot f_n-\tau_s(\tau_u k\cdot f_n))ds\Big\|_p^p\,du\Big)^{\frac{1}{p}} \\
&\qquad\le \|T\|\|l\|_{p'}\sum_{n=1}^{\infty}\|g_n\|_{p'}\Big(\int_{\mathbb{R}}\int_{\mathbb{R}_+}\alpha(s)\big\|\tau_u k\cdot f_n-\tau_s(\tau_u k\cdot f_n)\big\|_p^p\,dsdu\Big)^{\frac{1}{p}}.
\end{aligned}
$$

The integral in the right-hand side satisfies

$$
\begin{aligned}
&\Big(\int_{\mathbb{R}}\int_{\mathbb{R}_+}\alpha(s)\big\|\tau_u k\cdot f_n-\tau_s(\tau_u k\cdot f_n)\big\|_p^p\,dsdu\Big)^{\frac{1}{p}} \\
&\ \le\Big(\int_{\mathbb{R}}\int_{\mathbb{R}_+}\alpha(s)\big\|\tau_u k\cdot f_n-\tau_{s+u}k\cdot f_n\big\|_p^p dsdu\Big)^{\frac{1}{p}} \\
&\qquad+\Big(\int_{\mathbb{R}}\int_{\mathbb{R}_+}\alpha(s)\big\|\tau_{s+u}k\cdot f_n-\tau_s(\tau_u k\cdot f_n)\big\|_p^p\,dsdu\Big)^{\frac{1}{p}} \\
&\ \le\Big(\int_{\mathbb{R}}\int_{\mathbb{R}_+}\alpha(s)\big\|\tau_u((k-\tau_s k)\cdot f_n)\big\|_p^p\,dsdu\Big)^{\frac{1}{p}} \\
&\qquad+\Big(\int_{\mathbb{R}}\int_{\mathbb{R}_+}\alpha(s)\big\|\tau_{s+u}k\cdot(f_n-\tau_s f_n)\big\|_p^p dsdu\Big)^{\frac{1}{p}} \\
&\ \le\sup_{s\in\mathrm{supp}(\alpha)}\Big(\int_{\mathbb{R}}\big\|\tau_u(k-\tau_s k)\cdot f_n\big\|_p^p\,du\Big)^{\frac{1}{p}}+\sup_{s\in\mathrm{supp}(\alpha)}\Big(\int_{\mathbb{R}}\big\|\tau_{s+u}k\cdot(f_n-\tau_s f_n)\big\|_p^p\,du\Big)^{\frac{1}{p}} \\
&\ =\sup_{s\in\mathrm{supp}(\alpha)}\|k-\tau_s k\|_p\|f_n\|_p+\sup_{s\in\mathrm{supp}(\alpha)}\|k\|_p\|f_n-\tau_s f_n\|_p.
\end{aligned}
$$

Hence we obtain that

$$
\begin{aligned}
&\Big|\sum_{n=1}^{\infty}\langle T_{k,l}(f_n),\, g_n\rangle\Big| \\
&\ \le\|T\|\|l\|_{p'}\sum_{n=1}^{\infty}\|g_n\|_{p'}\big(\sup_{s\in\mathrm{supp}(\alpha)}\|k-\tau_s k\|_p\|f_n\|_p+\sup_{s\in\mathrm{supp}(\alpha)}\|k\|_p\|f_n-\tau_s f_n\|_p\big).
\end{aligned}
$$

Given $\epsilon > 0$, choose $M$ such that

$$\sum_{n=M+1}^{\infty} \|f_n\|_p \|g_n\|_{p'} < \epsilon.$$

We may find $s_0 > 0$ such that for all $s \in (0, s_0)$ and all $1 \le n \le M$, we have that

$$\|k - \tau_s k\|_p \le \frac{\epsilon \|k\|_p}{\sum_{n=1}^{\infty} \|f_n\|_p \|g_n\|_{p'}} \quad \text{and} \quad \|f_n - \tau_s f_n\|_p \le \frac{\epsilon}{M \|g_n\|_{p'}}.$$

We may now choose $\alpha$ so that $\mathrm{supp}(\alpha) \subseteq (0, t_0)$. Then we obtain from above that

$$\begin{aligned}
&\Big| \sum_{n=1}^{\infty} \langle T_{k,l}(f_n), g_n \rangle \Big| \\
&\quad \le \|T\| \|l\|_{p'} \Big( \epsilon \|k\|_p + \sum_{n=1}^{M} \|g_n\|_{p'} \cdot \sup_{s \in \mathrm{supp}(\alpha)} \|k\|_p \|f_n - \tau_s f_n\|_p \\
&\qquad\qquad\qquad\qquad + \sum_{n=M+1}^{\infty} \|g_n\|_{p'} \cdot \sup_{s \in \mathrm{supp}(\alpha)} \|k\|_p \|f_n - \tau_s f_n\|_p \Big) \\
&\quad \le \|T\| \|l\|_{p'} \Big( 2\epsilon \|k\|_p + \sum_{n=M+1}^{\infty} 2 \|k\|_p \|g_n\|_{p'} \|f_n\|_p \Big) \\
&\quad \le 4\epsilon \|T\| \|l\|_{p'} \|k\|_p.
\end{aligned}$$

Since $\epsilon$ was arbitrary, this shows that $\sum_{n=1}^{\infty} \langle T_{k,l}(f_n), g_n \rangle = 0$. Since $z = \sum_{n=1}^{\infty} f_n \otimes g_n$ was an arbitrary element of $\ker(Q_p)$, we obtain (15).

Next, we construct a sequence $(T_{k_n,l_n})_n$ which tends to $T$ in the $w^*$-topology of $B(L^p(\mathbb{R}_+))$. In the sequel, we assume that $k, l$ in $C_c(\mathbb{R})$ are such that

$$\|k\|_p = 1, \quad \|l\|_{p'} = 1 \quad \text{and} \quad \int_{\mathbb{R}} k(-s) l(s)\, ds = 1. \tag{16}$$

Consider any $f, g \in C_c(\mathbb{R}_+)$. We have

$$\begin{aligned}
&\big| \langle T(f), g \rangle - \langle T_{k,l}(f), g \rangle \big| \\
&\quad = \Big| \int_{\mathbb{R}} \langle T(k(-s) f), l(s) g \rangle - \langle T(\tau_s k \cdot f), \tau_{-s} l \cdot g \rangle \, ds \Big| \\
&\quad \le \int_{\mathbb{R}} \big| \langle T((k(-s) - \tau_s k) f), l(s) g \rangle \big| \, ds + \int_{\mathbb{R}} \big| \langle T(\tau_s k \cdot f), (l(s) - \tau_{-s} l) g \rangle \big| \, ds \\
&\quad \le \|T\| \Big( \int_{\mathbb{R}} \big\| (k(-s) - \tau_s k) f \big\|_p^p ds \Big)^{\frac{1}{p}} \Big( \int_{\mathbb{R}} \| l(s) g \|_{p'}^{p'} \, ds \Big)^{\frac{1}{p'}} \\
&\qquad\qquad + \|T\| \Big( \int_{\mathbb{R}} \| \tau_s k \cdot f \|_p^p ds \Big)^{\frac{1}{p}} \Big( \int_{\mathbb{R}} \| (l(s) - \tau_{-s} l) g \|_{p'}^{p'} \, ds \Big)^{\frac{1}{p'}} \\
&\quad \le \|T\| \|g\|_{p'} \Big( \int_{\mathbb{R}_+} |f(t)|^p \| \tau_t \check{k} - \check{k} \|_p^p \, dt \Big)^{\frac{1}{p}} \\
&\qquad\qquad + \|T\| \|f\|_p \Big( \int_{\mathbb{R}_+} |g(t)|^{p'} \| \tau_{-t} l - l \|_{p'}^{p'} \, dt \Big)^{\frac{1}{p'}}.
\end{aligned}$$

Here $\check{k}$ denotes the function $s \mapsto k(-s)$.

Now for $n \geq 1$, set

$$k_n := \frac{\chi_{[-n,n]}}{(2n)^{\frac{1}{p}}} \quad \text{and} \quad l_n := \frac{\chi_{[-n,n]}}{(2n)^{\frac{1}{p'}}},$$

where $\chi_{[-n,n]}$ is the indicator function of the interval $[-n, n]$. Then $\|k_n\|_p = \|l_n\|_{p'} = 1$ and $\int_{\mathbb{R}} k_n(-s) l_n(s)\, ds = 1$ as in (16). Let $K = \mathrm{supp}(f) \cup \mathrm{supp}(g)$ and let $r = \sup(K)$. Note that $\check{k}_n = k_n$ and that we have

$$\sup_{t \in K} \|\tau_t k_n - k_n\|_p \leq \left(\frac{r}{n}\right)^{\frac{1}{p}} \quad \text{and} \quad \sup_{t \in K} \|\tau_{-t} l_n - l_n\|_{p'} \leq \left(\frac{r}{n}\right)^{\frac{1}{p'}}.$$

Therefore,

$$\left|\langle T(f), g\rangle - \langle T_{k_n,l_n}(f), g\rangle\right| \leq \frac{2r}{n} \|T\| \|f\|_p \|g\|_{p'},$$

hence $\langle T_{k_n,l_n}(f), g\rangle \xrightarrow[n\to\infty]{} \langle T(f), g\rangle$. Since $\|T_{k_n,l_n}\| \leq \|T\|$ for all $n \geq 1$, this implies, by Lemma 2.1, that $T_{k_n,l_n} \to T$ in the $w^*$-topology of $B(L^p(\mathbb{R}_+))$. Consequently, $T \in \ker(Q_p)^{\perp}$ as expected. $\square$

**Remark 4.2.** (a) For any $1 \leq p \leq \infty$, let $H^p(\mathbb{R}) \subset L^p(\mathbb{R})$ be the subspace of all $f \in L^p(\mathbb{R})$ whose Fourier transform has support in $\mathbb{R}_+$. Recall the factorisation property

$$H^1(\mathbb{R}) = H^2(\mathbb{R}) \times H^2(\mathbb{R}).$$

More precisely, the product $h_1 h_2 \in H^1(\mathbb{R})$ and $\|h_1 h_2\|_1 \leq \|h_1\|_2 \|h_2\|_2$ for all $h_1, h_2 \in H^2(\mathbb{R})$ and conversely, for all $h \in H^1(\mathbb{R})$, there exist $h_1, h_2 \in H^2(\mathbb{R})$ such that $h = h_1 h_2$ and $\|h\|_1 = \|h_1\|_2 \|h_2\|_2$.

Recall that by definition,

$$A_2(\mathbb{R}_+) = \left\{\sum f_n * g_n : f_n, g_n \in L^2(\mathbb{R}_+), \sum \|f_n\|_2 \|g_n\|_2 < \infty\right\}.$$

It therefore follows from the above factorisation property and the identification of $L^2(\mathbb{R}_+)$ with $H^2(\mathbb{R})$ via the Fourier transform that

$$A_2(\mathbb{R}_+) = \{\hat{h} : h \in H^1(\mathbb{R})\},$$

with $\|\hat{h}\|_{A_2(\mathbb{R}_+)} = \|h\|_{H^1(\mathbb{R})}$. Therefore, we have an isometric identification

$$A_2(\mathbb{R}_+) \cong H^1(\mathbb{R}).$$

Since $H^1(\mathbb{R})^{\perp} = H^{\infty}(\mathbb{R})$, we have

$$H_1(\mathbb{R})^* \cong \frac{L^{\infty}(\mathbb{R})}{H^{\infty}(\mathbb{R})}.$$

Applying Theorem 4.1(2), we recover the well-known fact (see [Nikolski 2020, Section IV.5.3] or [Peller 2003, Theorem I.8.1]) that

$$\mathrm{Hank}_2(\mathbb{R}_+) \cong \frac{L^\infty(\mathbb{R})}{H^\infty(\mathbb{R})}.$$

(b) Notice that $\mathrm{Hank}_p(\mathbb{R}_+) \subseteq \mathrm{Hank}_2(\mathbb{R}_+)$. Indeed, suppose that $T \in \mathrm{Hank}_p(\mathbb{R}_+)$ and note that the adjoint mapping $T^* \in B(L^{p'}(\mathbb{R}_+))$ coincides with $T$ on $L^p(\mathbb{R}_+) \cap L^{p'}(\mathbb{R}_+)$. To see this, take $f, g \in L^p(\mathbb{R}_+) \cap L^{p'}(\mathbb{R}_+)$ and observe that $f \otimes g - g \otimes f$ belongs to $\ker(Q_p)$. This implies that $\langle T(f), g\rangle = \langle T(g), f\rangle$. Therefore, $T$ coincides with $T^*$ on $L^p(\mathbb{R}_+) \cap L^{p'}(\mathbb{R}_+)$. It then follows by interpolation that $T$ extends to a bounded operator on $L^2(\mathbb{R}_+)$, say $\widetilde{T}$. Since $T$ and $\widetilde{T}$ coincide on $L^p(\mathbb{R}_+) \cap L^2(\mathbb{R}_+)$ and $T$ is Hankelian, it follows from the definition of Hankel operators that $\widetilde{T}$ is also a Hankel operator and hence belongs to $\mathrm{Hank}_2(\mathbb{R}_+)$.

(c) The definition of $\mathrm{Hank}_p(\mathbb{R}_+)$ extends to the case $p = 1$. In analogy with Remark 3.2(c), we have an isometric identification

$$\mathrm{Hank}_1(\mathbb{R}_+) \simeq M(\mathbb{R}_+^*),$$

where $M(\mathbb{R}_+^*)$ denotes the space of all bounded Borel measures on $\mathbb{R}_+^*$. To establish this, we first note that for all $f \in L^1(\mathbb{R}_+)$, the function $u \mapsto \theta_u(f)$ is bounded and continuous from $\mathbb{R}_+^*$ into $L^1(\mathbb{R}_+)$. Hence for all $\nu \in M(\mathbb{R}_+^*)$, we may define $H_\nu \in B(L^1(\mathbb{R}_+))$ by

$$H_\nu(f) = \int_{\mathbb{R}_+^*} \theta_u(f)\, d\nu(u), \qquad f \in L^1(\mathbb{R}_+). \tag{17}$$

It is clear that $H_\nu$ is Hankelian. It follows from (14) that

$$\big\langle H_\nu(f), g\big\rangle = \int_{\mathbb{R}_+^*} (f * g)(u)\, d\nu(u), \qquad f \in L^1(\mathbb{R}_+),\ g \in L^\infty(\mathbb{R}_+).$$

We note that the mapping $\nu \mapsto H_\nu$ is a one-to-one contraction from $M(\mathbb{R}_+^*)$ into $\mathrm{Hank}_1(\mathbb{R}_+)$. We shall now prove that this mapping is an onto isometry.

We use the isometric identification $M(\mathbb{R}_+^*) \simeq C_0(\mathbb{R}_+^*)^*$ provided by the Riesz theorem and we regard $L^1(\mathbb{R}_+) \subseteq M(\mathbb{R}_+^*)$ in the obvious way. Let $T \in \mathrm{Hank}_1(\mathbb{R}_+)$. We observe that for all $h, f \in L^1(\mathbb{R}_+)$ and all $g \in C_0(\mathbb{R}_+^*)$, we have

$$\langle T(h * f), g\rangle = \langle T(h), f * g\rangle \tag{18}$$

Indeed, write $h*f = \int_0^\infty f(s)\tau_s h\, ds$. This implies that $T(h*f) = \int_0^\infty f(s) T(\tau_s h)\, ds$, hence

$$\langle T(h * f), g\rangle = \int_0^\infty f(s)\langle T\tau_s h, g\rangle\, ds = \int_0^\infty f(s)\langle Th, \tau_s g\rangle\, ds = \langle T(h), f * g\rangle.$$

Let $(h_n)_{n\geq 1}$ be a norm one approximate unit of $L^1(\mathbb{R}_+)$. Then $(T(h_n))_{n\geq 1}$ is a bounded sequence of $L^1(\mathbb{R}_+)$. Hence it admits a cluster point $\nu \in M(\mathbb{R}_+^*)$

in the $w^*$-topology of $M(\mathbb{R}_+^*)$. Thus, for all $g \in C_0(\mathbb{R}_+^*)$, the complex number $\int_{\mathbb{R}_+^*} g(u)\, d\nu(u)$ is a cluster point of the sequence $(\langle T(h_n), g\rangle)_{n\geq 1}$. Furthermore, we have $\|\nu\| \leq \|T\|$. Let $f \in L^1(\mathbb{R}_+)$ and let $g \in C_0(\mathbb{R}_+^*)$. Since $h_n * f \to f$ in $L^1(\mathbb{R}_+)$, we have that $\langle T(h_n * f), g\rangle \to \langle T(f), g\rangle$. By (18), we may write $\langle T(h_n * f), g\rangle = \langle T(h_n), f * g\rangle$. We deduce that

$$\langle T(f), g\rangle = \int_{\mathbb{R}_+^*} (f * g)(u)\, d\nu(u).$$

This implies that $T = H_\nu$, see (17), which concludes the proof.

**Definition 4.3.** We say that a function $m : \mathbb{R}_+^* \to \mathbb{C}$ is the symbol of a multiplier on $\mathrm{Hank}_p(\mathbb{R}_+)$ if there exist a $w^*$-continuous operator $T_m : \mathrm{Hank}_p(\mathbb{R}_+) \to \mathrm{Hank}_p(\mathbb{R}_+)$ such that for every $u > 0$, $T_m(\theta_u) = m(u)\theta_u$. (Note that such an operator $T_m$ is necessarily unique.)

**Remark 4.4.** Suppose that $T_m : \mathrm{Hank}_p(\mathbb{R}_+) \to \mathrm{Hank}_p(\mathbb{R}_+)$ is a multiplier as defined above. Using Theorem 4.1(2), let $S_m : A_p(\mathbb{R}_+) \to A_p(\mathbb{R}_+)$ be the operator such that $S_m^* = T_m$. For $f \in L^p(\mathbb{R}_+)$ and $g \in L^{p'}(\mathbb{R}_+)$, we have, by (14),

$$\begin{aligned}
[S_m(f * g)](u) &= \langle \theta_u, S_m(f * g)\rangle \\
&= \langle T_m(\theta_u), f * g\rangle \\
&= m(u)\langle \theta_u, f * g\rangle \\
&= m(u)(f * g)(u).
\end{aligned}$$

We deduce that $S_m(F) = m \cdot F$, for every $F \in A_p(\mathbb{R}_+)$.

Conversely, if $m : \mathbb{R}_+^* \to \mathbb{C}$ is such that $S_m : A_p(\mathbb{R}_+) \to A_p(\mathbb{R}_+)$ given by $S_m(F) = m \cdot F$ is well-defined and bounded, then $S_m^*$ is a multiplier on $\mathrm{Hank}_p(\mathbb{R}_+)$.

**Lemma 4.5.** *If $m : \mathbb{R}_+^* \to \mathbb{C}$ is the symbol of a multiplier on* $\mathrm{Hank}_p(\mathbb{R}_+)$, *then $m$ is continuous and bounded.*

*Proof.* For all $u > 0$, we have $m(u)\theta_u = T_m(\theta_u)$, hence $|m(u)| \leq \|T_m\|$. Thus, $m$ is bounded. For any $a > 0$, let $\chi_{(0,a)}$ be the indicator function of the interval $(0, a)$. Then $m \cdot \chi_{(0,a)} * \chi_{(0,a)}$ belongs to $A_p(\mathbb{R}_+)$, hence to $C_b(\mathbb{R}_+^*)$, by Remark 4.4. Since $\chi_{(0,a)} * \chi_{(0,a)} > 0$ on $(0, 2a)$, it follows that $m$ is continuous on $(0, 2a)$. Thus, $m$ is continuous on $\mathbb{R}_+^*$. □

**Theorem 4.6.** *Let $1 < p < \infty$, let $C \geq 0$ be a constant and let $m : \mathbb{R}_+^* \to \mathbb{C}$ be a function. The following assertions are equivalent.*

(i) *$m$ is the symbol of a $p$-completely bounded multiplier on* $\mathrm{Hank}_p(\mathbb{R}_+)$, *and*

$$\|T_m : \mathrm{Hank}_p(\mathbb{R}_+) \to \mathrm{Hank}_p(\mathbb{R}_+)\|_{p-cb} \leq C.$$

(ii) *$m$ is continuous and there exist a measure space $(\Omega, \mu)$ and two functions $\alpha \in L^\infty(\mathbb{R}_+; L^p(\Omega))$ and $\beta \in L^\infty(\mathbb{R}_+; L^{p'}(\Omega))$ such that $\|\alpha\|_\infty \|\beta\|_\infty \le C$ and $m(s+t) = \langle \alpha(s), \beta(t)\rangle$, for almost every $(s,t) \in \mathbb{R}_+^{*2}$.*

*Proof.* By homogeneity, we may assume that $C = 1$ throughout this proof.

Assume (i). The continuity of $m$ follows from Lemma 4.5. Let $T_m : \mathrm{Hank}_p(\mathbb{R}_+) \to \mathrm{Hank}_p(\mathbb{R}_+)$ be the $p$-completely contractive multiplier associated with $m$. Let $\kappa : L^p(\mathbb{R}) \to L^p(\mathbb{R})$ be defined by $(\kappa f)(t) = f(-t)$, for all $f \in L^p(\mathbb{R})$. Let $J : L^p(\mathbb{R}_+) \to L^p(\mathbb{R})$ be the canonical embedding and let $Q : L^p(\mathbb{R}) \to L^p(\mathbb{R}_+)$ be the canonical projection defined by $Qf = f_{|\mathbb{R}_+}$. Let $q : B(L^p(\mathbb{R})) \to B(L^p(\mathbb{R}_+))$ be given by $q(T) = Q\kappa T J$, for all $T \in B(L^p(\mathbb{R}))$. Applying the easy implication (ii) $\Longrightarrow$ (i) of Theorem 2.2 we obtain that $q$ is $p$-completely contractive.

Let $\mathcal{M}_p(\mathbb{R}) \subseteq B(L^p(\mathbb{R}))$ denote the subalgebra of bounded Fourier multipliers. Let us show that if $T \in \mathcal{M}_p(\mathbb{R})$, then $q(T) \in \mathrm{Hank}_p(\mathbb{R}_+)$. For any $s \in \mathbb{R}$, recall $\tau_s \in B(L^p(\mathbb{R}))$ given by $\tau_s(f) = f(\,\cdot - s)$. Note that $\tau_s \in \mathcal{M}_p(\mathbb{R})$ and that $\mathcal{M}_p(\mathbb{R}) = \overline{\mathrm{Span}}^{w^*}\{\tau_s : s \in \mathbb{R}\}$. For all $f \in L^p(\mathbb{R}_+)$, we have

$$q(\tau_s)f = Q\tau(f(\,\cdot - s)) = Q(f(-(\,\cdot + s))) = \{t \in \mathbb{R}_+ \mapsto f(-t-s)\}.$$

Hence, if $s \ge 0$, then $q(\tau_s) = 0$ and if $s < 0$, then $q(\tau_s) = \theta_{-s}$. It is plain that $q$ is $w^*$-continuous. Since $\mathrm{Hank}_p(\mathbb{R}_+)$ is $w^*$-closed, we deduce that $q$ maps $\mathcal{M}_p(\mathbb{R})$ into $\mathrm{Hank}_p(\mathbb{R}_+)$.

Consider the mapping

$$q_0 := q_{|\mathcal{M}_p(\mathbb{R})} : \mathcal{M}_p(\mathbb{R}) \to \mathrm{Hank}_p(\mathbb{R}_+)$$

and set

$$\Gamma := T_m \circ q_0 : \mathcal{M}_p(\mathbb{R}) \to B(L^p(\mathbb{R}_+)).$$

It follows from above that

$$\Gamma(\tau_{-s}) = m(s)\theta_s, \quad s > 0. \tag{19}$$

Since $q$ is $p$-completely contractive, $\Gamma$ is also $p$-completely contractive. Applying Theorem 2.2 to $\Gamma$, we obtain the existence of an $SQ_p$-space $E$, a unital $p$-completely contractive, nondegenerate homomorphism $\pi : \mathcal{M}_p(\mathbb{R}) \to B(E)$ as well as operators $V : L^p(\mathbb{R}_+) \to E$ and $W : E \to L^p(\mathbb{R}_+)$ such that $\|V\|\|W\| \le 1$ and for every $x \in \mathcal{M}_p(\mathbb{R})$, $\Gamma(x) = W\pi(x)V$.

Let $c : L^1(\mathbb{R}) \to \mathcal{M}_p(\mathbb{R})$ be defined by $[c(g)](f) = g * f$, for all $g \in L^1(\mathbb{R})$ and $f \in L^p(\mathbb{R})$. Let $\lambda : L^1(\mathbb{R}) \to B(E)$ be given by $\lambda = \pi \circ c$. Then $\lambda$ is a contractive, nondegenerate homomorphism. By [de Pagter and Ricker 2008, Remark 2.5], there exists $\sigma : \mathbb{R} \to B(E)$, a bounded strongly continuous representation such that for all $g \in L^1(\mathbb{R})$, $\lambda(g) = \int_{\mathbb{R}} g(t)\sigma(t)\,dt$ (defined in the strong sense). Let us show that

$$\Gamma(\tau_{-s}) = W\sigma(-s)V, \quad s > 0. \tag{20}$$

Let $\eta \in L^1(\mathbb{R})_+$ be such that $\int_{\mathbb{R}} \eta(t)\,dt = 1$. For any $r > 0$, let $\eta_r(t) = r\eta(rt)$. Since $\sigma : \mathbb{R} \to B(E)$ is strongly continuous, the function $t \mapsto \langle \sigma(t)x, x^* \rangle$ is continuous and we have

$$\int_{\mathbb{R}} \eta_r(-s-t)\langle \sigma(t)x, x^* \rangle \,dt \xrightarrow[r\to\infty]{} \langle \sigma(-s)x, x^* \rangle, \tag{21}$$

for all $x \in E$ and $x^* \in E^*$. Since the left-hand side in (21) is equal to

$$\langle \pi(c(\eta_r(-s-\cdot)))x, x^* \rangle,$$

we obtain, by Lemma 2.1, that $\pi(c(\eta_r(-s-\cdot))) \to \sigma(-s)$ in the $w^*$-topology of $B(E)$. This implies that $W\pi(c(\eta_r(-s-\cdot)))V \to W\sigma(-s)V$ in the $w^*$-topology of $B(L^p(\mathbb{R}_+))$. We next show that $W\pi(c(\eta_r(-s-\cdot)))V \to \Gamma(\tau_{-s})$ in the $w^*$-topology of $B(L^p(\mathbb{R}_+))$, which will complete the proof of (20). Since

$$W\pi(c(\eta_r(-s-\cdot)))V = \Gamma(c(\eta_r(-s-\cdot)))$$

and $\Gamma$ is $w^*$-continuous, it suffices to show that $c(\eta_r(-s-\cdot)) \to \tau_{-s}$ in the $w^*$-topology of $B(L^p(\mathbb{R}))$. To see this, let $f \in L^p(\mathbb{R})$ and $g \in L^{p'}(\mathbb{R})$. We have that

$$\begin{aligned}\langle c(\eta_r(-s-\cdot))f, g \rangle &= \langle \eta_r(-s-\cdot) * f, g \rangle \\ &= \langle \delta_{-s} * \eta_r * f, g \rangle \\ &\to \langle \delta_{-s} * f, g \rangle = \langle \tau_{-s} f, g \rangle.\end{aligned}$$

By Lemma 2.1 again, this proves that $c(\eta_r(-s-\cdot)) \to \tau_{-s}$ in the $w^*$-topology, as expected.

Given any $\epsilon > 0$, let $m_\epsilon : \mathbb{R}_+^* \to \mathbb{C}$ be defined by

$$m_\epsilon(t) = m(t+\epsilon), \quad t > 0.$$

Let $f \in L^p(\mathbb{R}_+)$ be given by $f = \epsilon^{-\frac{1}{p}} \chi_{(0,\epsilon)}$ and let $g \in L^{p'}(\mathbb{R}_+)$ be given by $g = \epsilon^{-\frac{1}{p'}} \chi_{(0,\epsilon)}$. For any $s, t > 0$, set

$$\alpha_\epsilon(s) := \sigma\big(-s - \tfrac{\epsilon}{2}\big)V(\tau_s f) \quad \text{and} \quad \beta_\epsilon(t) := \sigma\big(-t - \tfrac{\epsilon}{2}\big)^* W^*(\tau_t g).$$

Since $\sigma$ is strongly continuous, $\alpha_\epsilon$ and $\beta_\epsilon$ are continuous. By (19) and (20), we have that

$$\begin{aligned}\langle \alpha_\epsilon(s), \beta_\epsilon(t) \rangle_{E,E^*} &= \big\langle \sigma\big(-s - \tfrac{\epsilon}{2}\big)V(\tau_s f), \sigma\big(-t - \tfrac{\epsilon}{2}\big)^* W^*(\tau_t g) \big\rangle \\ &= \langle W\sigma(-s-t-\epsilon)V(\tau_s f), \tau_t g \rangle \\ &= \langle (\Gamma(\tau_{-s-t-\epsilon}))(\tau_s f), \tau_t g \rangle \\ &= m(s+t+\epsilon)\langle \theta_{s+t+\epsilon}(\tau_s f), \tau_t g \rangle \\ &= m_\epsilon(s+t)\langle \epsilon^{-1/p}\chi_{(t,t+\epsilon)}, \epsilon^{-1/p'}\chi_{(t,t+\epsilon)} \rangle \\ &= m_\epsilon(s+t),\end{aligned}$$

for all $s, t > 0$. Moreover, $\|\alpha_\epsilon(s)\| \leq \|V\|$ and $\|\beta_\epsilon(t)\| \leq \|W\|$ for all $t, s > 0$. Since $\alpha_\epsilon$ and $\beta_\epsilon$ are continuous, this implies that $\alpha_\epsilon \in L^\infty(\mathbb{R}_+; E)$, $\beta_\epsilon \in L^\infty(\mathbb{R}_+; E^*)$ and $\|\alpha_\epsilon\|_\infty \|\beta_\epsilon\|_\infty \leq \|V\| \|W\| \leq 1$.

We now show that the $SQ_p$-space $E$ can be replaced by an $L^p$-space in the above factorisation property of $m_\epsilon$. Following Remark 2.3, assume that $E = E_1/E_2$, with $E_2 \subseteq E_1 \subseteq L^p(\Omega)$, and for all $f \in E_1$, let $\dot{f} \in E$ denote the class of $f$. Recall (5) and for all $g \in E_2^\perp$, let $\dot{g} \in E^*$ denote the class of $g$. Since $E$ is a quotient of $E_1$, we have an isometric embedding $E^* \subseteq E_1^*$. More precisely,

$$E^* = \frac{E_2^\perp}{E_1^\perp} \hookrightarrow \frac{L^{p'}(\Omega)}{E_1^\perp} = E_1^*.$$

This induces an isometric embedding

$$L^1(\mathbb{R}_+; E^*) \subseteq L^1(\mathbb{R}_+; E_1^*).$$

Since $E^*$ and $E_1^*$ are reflexive, we may apply the identifications

$$L^1(\mathbb{R}_+; E^*)^* \simeq L^\infty(\mathbb{R}_+; E) \quad \text{and} \quad L^1(\mathbb{R}_+; E_1^*)^* \simeq L^\infty(\mathbb{R}_+; E_1)$$

provided by (7). By the Hahn–Banach theorem, we deduce the existence of $\widetilde{\alpha_\epsilon} \in L^\infty(\mathbb{R}_+; E_1)$ such that $\|\widetilde{\alpha_\epsilon}\|_\infty = \|\alpha_\epsilon\|_\infty$ and the functional $L^1(\mathbb{R}_+; E_1^*) \to \mathbb{C}$ induced by $\widetilde{\alpha_\epsilon}$ extends the functional $L^1(\mathbb{R}_+; E^*) \to \mathbb{C}$ induced by $\alpha_\epsilon$. It is easy to check that the latter means that $\dot{\widetilde{\alpha_\epsilon}}(s) = \alpha_\epsilon(s)$ almost everywhere on $\mathbb{R}_+$. Likewise, there exist $\widetilde{\beta_\epsilon} \in L^\infty(\mathbb{R}_+; E_2^\perp)$ such that $\|\widetilde{\beta_\epsilon}\|_\infty = \|\beta_\epsilon\|_\infty$ and $\dot{\widetilde{\beta_\epsilon}}(t) = \beta_\epsilon(t)$ almost everywhere on $\mathbb{R}_+$. Regard $\widetilde{\alpha}_\epsilon$ as an element of $L^\infty(\mathbb{R}_+, L^p(\Omega))$ and $\widetilde{\beta}_\epsilon$ as an element of $L^\infty(\mathbb{R}_+, L^{p'}(\Omega))$. By (6), we then have

$$\langle \alpha_\epsilon(s), \beta_\epsilon(t) \rangle_{E, E^*} = \langle \widetilde{\alpha}_\epsilon(s), \widetilde{\beta}_\epsilon(t) \rangle_{L^p, L^{p'}},$$

for almost every $(s, t) \in \mathbb{R}_+^{*2}$.

We therefore obtain that $m_\epsilon : \mathbb{R}_+^* \to \mathbb{C}$ satisfies condition (ii) of the theorem (with $C = 1$).

Define $\varphi : \mathbb{R}_+^{*2} \to \mathbb{C}$ by $\varphi(s, t) = m(s + t)$. Likewise, for any $\epsilon > 0$, define $\varphi_\epsilon : \mathbb{R}_+^{*2} \to \mathbb{C}$ by $\varphi(s, t) = m_\epsilon(s + t)$. Since $m$ is continuous, the functions $\varphi$ and $\varphi_\epsilon$ are continuous. It follows from above that for all $\epsilon > 0$, $\varphi_\epsilon$ satisfies condition (ii) in Theorem 2.5, with $C = 1$. The latter theorem therefore implies that the family $\{\varphi_\epsilon(s, t)\}_{(s,t) \in \mathbb{R}_+^{*2}}$ is a bounded Schur multiplier on $B(\ell^p_{\mathbb{R}_+^*})$, with norm less than one. Thus for all $[a_{ij}]_{1 \leq i, j \leq n}$ in $M_n$ and for all $t_1, \ldots, t_n, s_1, \ldots, s_n$ in $\mathbb{R}_+^*$, we have $\|[\varphi_\epsilon(s_i, t_j) a_{ij}]\|_{B(\ell^p_n)} \leq \|[a_{ij}]\|_{B(\ell^p_n)}$. Since $m$ is continuous, $\varphi_\epsilon \to \varphi$ pointwise when $\epsilon \to 0$. We deduce that $\varphi$ satisfies (8) with $C = 1$ for all $[a_{ij}]_{1 \leq i, j \leq n}$ in $M_n$ and all $t_1, \ldots, t_n, s_1, \ldots, s_n$ in $\mathbb{R}_+^*$. Consequently, the family $\{\varphi(s, t)\}_{(s,t) \in \mathbb{R}_+^{*2}}$ is a bounded Schur multiplier on $B(\ell^p_{\mathbb{R}_+^*})$, with norm less than one. Applying the implication (i) $\Longrightarrow$ (ii) in Theorem 2.5, we deduce the assertion (ii) of Theorem 4.6.

Conversely, assume (ii). Following Lemma 2.4, let

$$\pi : B(L^p(\mathbb{R}_+)) \to B(L^p(\mathbb{R}_+ \times \Omega))$$

be the $p$-completely isometric homomorphism defined by $\pi(T) = T \overline{\otimes} I_{L^p(\Omega)}$. This map is $w^*$-continuous. Indeed, let $(T_\iota)_\iota$ be a bounded net of $B(L^p(\mathbb{R}_+))$ converging to some $T \in B(L^p(\mathbb{R}_+))$ in the $w^*$-topology. For any $f \in L^p(\mathbb{R}_+)$, $g \in L^{p'}(\mathbb{R}_+)$, $\varphi \in L^p(\Omega)$ and $\psi \in L^{p'}(\Omega)$, we have

$$\big\langle \pi(T_\iota), (f \otimes \varphi) \otimes (g \otimes \psi)\big\rangle = \langle T_\iota f, g\rangle_{L^p(\mathbb{R}_+), L^{p'}(\mathbb{R}_+)} \langle \varphi, \psi\rangle_{L^p(\Omega), L^{p'}(\Omega)},$$

where the duality pairing in the left-hand side refers to the identification

$$\big(L^p(\mathbb{R}_+ \times \Omega) \widehat{\otimes} L^{p'}(\mathbb{R}_+ \times \Omega)\big)^* \simeq B(L^p(\mathbb{R}_+ \times \Omega)).$$

Since $\langle T_\iota f, g\rangle \to \langle T f, g\rangle$, we deduce that

$$\langle \pi(T_\iota), (f \otimes \varphi) \otimes (g \otimes \psi)\rangle \to \langle \pi(T), (f \otimes \varphi) \otimes (g \otimes \psi)\rangle.$$

Since $L^p(\mathbb{R}_+) \otimes L^p(\Omega)$ and $L^{p'}(\mathbb{R}_+) \otimes L^{p'}(\Omega)$ are dense in $L^p(\mathbb{R}_+ \times \Omega)$ and $L^{p'}(\mathbb{R}_+ \times \Omega)$, respectively, we deduce that $\pi(T_\iota) \to \pi(T)$ in the $w^*$-topology, by Lemma 2.1. This proves that $\pi$ is $w^*$-continuous.

Let $V : L^p(\mathbb{R}_+) \to L^p(\mathbb{R}_+; L^p(\Omega)) \simeq L^p(\mathbb{R}_+ \times \Omega)$ be defined by

$$V(f) = f\alpha, \quad f \in L^p(\mathbb{R}_+).$$

This is a well-defined contraction. Likewise we define a contraction

$$W : L^p(\mathbb{R}_+ \times \Omega) \to L^p(\mathbb{R}_+)$$

by setting

$$W^*(g) = g\beta, \quad g \in L^{p'}(\mathbb{R}_+).$$

It follows from above and from the implication (ii) $\Longrightarrow$ (i) of Theorem 2.2 that the mapping

$$w : B(L^p(\mathbb{R}_+)) \to B(L^p(\mathbb{R}_+)), \quad w(T) = W\pi(T)V$$

is a $w^*$-continuous $p$-complete contraction.

We claim that for all $u > 0$, we have

$$w(\theta_u) = m(u)\theta_u. \tag{22}$$

To prove this, consider $f \in L^p(\mathbb{R}_+)$ and $g \in L^{p'}(\mathbb{R}_+)$. For all $u > 0$, we have

$$\langle w(\theta_u) f, g\rangle = \langle \pi(\theta_u) V(f), W^*(g)\rangle = \langle \pi(\theta_u)(f\alpha), (g\beta)\rangle.$$

By the definitions of $\pi$ and $\theta_u$, we have $\pi(\theta_u)(f\alpha) = (f\alpha)(u - \cdot\,)$. Consequently,

$$\langle w(\theta_u) f, g\rangle = \int_0^u f(u-t) g(t) \langle \alpha(u-t), \beta(t)\rangle \, dt, \quad u > 0.$$

Let $h \in L^1(\mathbb{R}_+)$ be an auxiliary function. Then using Fubini's theorem and setting $s = u - t$ in due place, we obtain that

$$\begin{aligned}\int_0^\infty \langle w(\theta_u) f, g\rangle h(u)\,du &= \int_0^\infty \int_t^\infty h(u) f(u-t) g(t) \langle \alpha(u-t), \beta(t)\rangle\, dudt \\ &= \int_0^\infty \int_0^\infty h(s+t) f(s) g(t) \langle \alpha(s), \beta(t)\rangle\, dsdt.\end{aligned}$$

Applying the a.e. equality $m(s+t) = \langle \alpha(s), \beta(t)\rangle$ and reversing this computation, we deduce that

$$\int_0^\infty \langle w(\theta_u) f, g\rangle h(u)\,du = \int_0^\infty m(u)(f * g)(u) h(u)\,du.$$

Since $h$ is arbitrary, this implies that $\langle w(\theta_u) f, g\rangle = m(u)(f * g)(u)$ for a.e. $u > 0$. Equivalently, $\langle w(\theta_u) f, g\rangle = m(u)\langle \theta_u f, g\rangle$ for a.e. $u > 0$. It is plain that $u \mapsto \theta_u$ is $w^*$-continuous on $B(L^p(\mathbb{R}_+))$. Since $w$ is $w^*$-continuous, the function $u \mapsto \langle w(\theta_u) f, g\rangle$ is continuous as well. Since $m$ is assumed continuous, we deduce that $\langle w(\theta_u) f, g\rangle = m(u)\langle \theta_u f, g\rangle$ for all $u > 0$. This yields (22), for all $u > 0$.

By part (1) of Theorem 4.1 and the $w^*$-continuity of $w$, the identity (22) implies that $\mathrm{Hank}_p(\mathbb{R}_+)$ is an invariant subspace of $w$. Further the restriction of $w$ to $\mathrm{Hank}_p(\mathbb{R}_+)$ is the multiplier associated to $m$. The assertion (i) follows. ☐

**Remark 4.7.** We proved in [Arnold et al. 2022, Theorem 3.1] that a continuous function $m : \mathbb{R}_+^* \to \mathbb{C}$ is the symbol of an $S^1$-bounded Fourier multiplier on $H^1(\mathbb{R})$, with $S^1$-bounded norm $\leq C$, if and only if there exist a Hilbert space $\mathcal{H}$ and two functions $\alpha, \beta \in L^\infty(\mathbb{R}_+; \mathcal{H})$ such that $\|\alpha\|_\infty \|\beta\|_\infty \leq C$ and $m(s+t) = \langle \alpha(t), \beta(s)\rangle_{\mathcal{H}}$ for almost every $(s,t) \in \mathbb{R}_+^{*2}$. It turns out that using (1), a mapping $S : H^1(\mathbb{R}) \to H^1(\mathbb{R})$ is an $S^1$-bounded Fourier multiplier with $S^1$-bounded norm $\leq C$ if and only if $S^* : \mathrm{Hank}_2(\mathbb{R}_+) \to \mathrm{Hank}_2(\mathbb{R}_+)$ is a completely bounded multiplier with completely bounded norm $\leq C$. See [Arnold et al. 2022, Remark 3.4] for more on this. Thus the statement in [Arnold et al. 2022, Theorem 3.1] is equivalent to the case $p = 2$ of Theorem 4.6. In this regard, Theorem 4.6 can be regarded as a $p$-analogue of [Arnold et al. 2022, Theorem 3.1].

**Remark 4.8.** Let $f \in L^p(\mathbb{R}_+)$ and $g \in L^{p'}(\mathbb{R}_+)$. For any $s, t > 0$, we may write

$$(f * g)(s+t) = \int_{\mathbb{R}} f(s+r) g(t-r)\,dr.$$

Equivalently,

$$(f * g)(s+t) = \langle \tau_{-s} f, \tau_t \check{g}\rangle_{L^p(\mathbb{R}_+), L^{p'}(\mathbb{R}_+)}.$$

According to the implication (ii) $\Longrightarrow$ (i) of Theorem 4.6 and Remark 4.4, $f * g$ is therefore a pointwise multiplier of $A_p(\mathbb{R}_+)$, with norm less than or equal to $\|f\|_p \|g\|_{p'}$. We deduce that every $F \in A_p(\mathbb{R}_+)$ is a pointwise multiplier of $A_p(\mathbb{R}_+)$, with norm less than or equal to $\|F\|_{A_p}$. This means that $A_p(\mathbb{R}_+)$ is a Banach algebra for the pointwise product.

## Acknowledgements

Le Merdy was supported by the ANR project *Noncommutative analysis on groups and quantum groups* (No. ANR-19-CE40-0002). Zadeh was supported by Projet I-SITE MultiStructure *Harmonic Analysis of noncommutative Fourier and Schur multipliers over operator spaces*. The authors gratefully thank the Heilbronn Institute for Mathematical Research and the UKRI/EPSRC Additional Funding Program for Mathematical Sciences for the financial support through Heilbronn Grant R102688-101.

## References

[Arnold et al. 2022] L. Arnold, C. L. Merdy, and S. Zadeh, "$S^1$-bounded Fourier multipliers on $H^1(\mathbb{R})$ and functional calculus for semigroups", 2022. arXiv 2203.16829

[Coine 2018] C. Coine, "Schur multipliers on $\mathcal{B}(L^p, L^q)$", *J. Operator Theory* **79**:2 (2018), 301–326. MR Zbl

[Daws 2010] M. Daws, "$p$-operator spaces and Figà-Talamanca–Herz algebras", *J. Operator Theory* **63**:1 (2010), 47–83. MR Zbl

[Derighetti 2011] A. Derighetti, *Convolution operators on groups*, Lecture Notes of the Unione Matematica Italiana **11**, Springer, 2011. MR Zbl

[Diestel and Uhl 1977] J. Diestel and J. J. Uhl, Jr., *Vector measures*, Mathematical Surveys **15**, American Mathematical Society, Providence, RI, 1977. MR Zbl

[Herz 1974] C. Herz, "Une généralisation de la notion de transformée de Fourier-Stieltjes", *Ann. Inst. Fourier* (*Grenoble*) **24**:3 (1974), xiii, 145–157. MR Zbl

[Le Merdy 1996] C. Le Merdy, "Factorization of $p$-completely bounded multilinear maps", *Pacific J. Math.* **172**:1 (1996), 187–213. MR Zbl

[Nikolski 2002] N. K. Nikolski, *Operators*, *functions*, *and systems*: *an easy reading*, *I*: *Hardy*, *Hankel*, *and Toeplitz*, vol. 1, Mathematical Surveys and Monographs **92**, American Mathematical Society, Providence, RI, 2002. MR Zbl

[Nikolski 2020] N. K. Nikolski, *Toeplitz matrices and operators*, Cambridge Studies in Advanced Mathematics **182**, Cambridge University Press, 2020. MR Zbl

[de Pagter and Ricker 2008] B. de Pagter and W. J. Ricker, "$R$-bounded representations of $L^1(G)$", *Positivity* **12**:1 (2008), 151–166. MR

[Paulsen 2002] V. Paulsen, *Completely bounded maps and operator algebras*, Cambridge Studies in Advanced Mathematics **78**, Cambridge University Press, 2002. MR Zbl

[Peller 2003] V. V. Peller, *Hankel operators and their applications*, Springer, 2003. MR Zbl

[Pisier 1990] G. Pisier, "Completely bounded maps between sets of Banach space operators", *Indiana Univ. Math. J.* **39**:1 (1990), 249–277. MR Zbl

[Pisier 2001] G. Pisier, *Similarity problems and completely bounded maps*, Lecture Notes in Mathematics **1618**, Springer, 2001. MR Zbl

[Power 1982] S. C. Power, *Hankel operators on Hilbert space*, Research Notes in Mathematics **64**, Pitman, Boston, 1982. MR Zbl

[Yafaev 2015] D. R. Yafaev, "On finite rank Hankel operators", *J. Funct. Anal.* **268**:7 (2015), 1808–1839. MR Zbl

[Yafaev 2017a] D. R. Yafaev, "Hankel and Toeplitz operators: continuous and discrete representations", *Opuscula Math.* **37**:1 (2017), 189–218. MR Zbl

[Yafaev 2017b] D. R. Yafaev, "Quasi-diagonalization of Hankel operators", *J. Anal. Math.* **133** (2017), 133–182. MR Zbl

Received February 3, 2023.

LORIS ARNOLD
INSTITUTE OF MATHEMATICS
POLISH ACADEMY OF SCIENCES
WARSAW
POLAND
larnold@impan.pl

CHRISTIAN LE MERDY
LABORATOIRE DE MATHÉMATIQUES DE BESANÇON, UMR CNRS 6623
UNIVERSITÉ DE FRANCHE-COMTÉ
BESANÇON
FRANCE
clemerdy@univ-fcomte.fr

SAFOURA ZADEH
LABORATOIRE DE MATHÉMATIQUES BLAISE PASCAL, UMR 6620
UNIVERSITÉ CLERMONT AUVERGNE
FRANCE
and
SCHOOL OF MATHEMATICS
UNIVERSITY OF BRISTOL
BRISTOL
UNITED KINGDOM
and
INSTITUT DES HAUTES ÉTUDES SCIENTIFIQUES
BURES-SUR-YVETTE
FRANCE
jsafoora@gmail.com

PACIFIC JOURNAL OF MATHEMATICS
Vol. 328, No. 2, 2024
https://doi.org/10.2140/pjm.2024.328.217

# STABLE FUNCTORIAL EQUIVALENCE OF BLOCKS

SERGE BOUC AND DENİZ YILMAZ

**Let $k$ be an algebraically closed field of characteristic $p > 0$, let $R$ be a commutative ring and let $\mathbb{F}$ be an algebraically closed field of characteristic 0. We introduce the category $\overline{\mathcal{F}^{\Delta}_{Rpp_k}}$ of stable diagonal $p$-permutation functors over $R$. We prove that the category $\overline{\mathcal{F}^{\Delta}_{\mathbb{F}pp_k}}$ is semisimple and give a parametrization of its simple objects in terms of the simple diagonal $p$-permutation functors.**

**We also introduce the notion of a stable functorial equivalence over $R$ between blocks of finite groups. We prove that if $G$ is a finite group and if $b$ is a block idempotent of $kG$ with an abelian defect group $D$ and Frobenius inertial quotient $E$, then there exists a stable functorial equivalence over $\mathbb{F}$ between the pairs $(G, b)$ and $(D \rtimes E, 1)$.**

## 1. Introduction

Various notions of equivalences between blocks of finite groups have been studied such as splendid Morita equivalence, splendid Rickard equivalence, $p$-permutation equivalence, isotypies and perfect isometries [Broué 1990; Boltje and Xu 2008; Boltje and Perepelitsky 2020]. These equivalences are related to prominent conjectures in modular representation theory such as Broué's abelian defect group conjecture [Linckelmann 2018, Conjecture 9.7.6], Puig's finiteness conjecture [Linckelmann 2018, Conjecture 6.4.2] and Donovan's conjecture [Linckelmann 2018, Conjecture 6.1.9].

In [Bouc and Yılmaz 2022] we introduced another equivalence of blocks, namely *functorial equivalence*, using the notion of diagonal $p$-permutation functors: Let $k$ be an algebraically closed field of characteristic $p > 0$, let $\mathbb{F}$ be an algebraically closed field of characteristic 0 and let $R$ be a commutative ring. We denote by $Rpp_k^{\Delta}$ the category whose objects are finite groups and for finite groups $G$ and $H$ whose morphisms from $H$ to $G$ are the Grothendieck group $RT^{\Delta}(G, H)$ of *diagonal* $p$-permutation $(kG, kH)$-bimodules. An $R$-linear functor from $Rpp_k^{\Delta}$ to ${}_R\mathsf{Mod}$ is called a *diagonal $p$-permutation functor*. To each pair $(G, b)$ of a finite group $G$

*MSC2020:* 16S34, 20C20, 20J15.

*Keywords:* block, diagonal $p$-permutation functors, functorial equivalence, Frobenius inertial quotient.

and a block idempotent $b$ of $kG$, we associate a canonical diagonal $p$-permutation functor over $R$, denoted by $RT^{\Delta}_{G,b}$. If $(H, c)$ is another such pair, we say that $(G, b)$ and $(H, c)$ are *functorially equivalent over* $R$ if the functors $RT^{\Delta}_{G,b}$ and $RT^{\Delta}_{H,c}$ are isomorphic.

In [Bouc and Yılmaz 2022] we proved that the category of diagonal $p$-permutation functors over $\mathbb{F}$ is semisimple, parametrized simple functors and provided three equivalent descriptions of the decomposition of the functor $\mathbb{F}T^{\Delta}_{G,b}$ in terms of the simple functors [Bouc and Yılmaz 2022, Corollary 6.15 and Theorem 8.22]. We proved that the number of isomorphism classes of simple modules, the number of ordinary irreducible characters, and the defect groups are preserved under functorial equivalences over $\mathbb{F}$ [Bouc and Yılmaz 2022, Theorem 10.5]. Moreover we proved that for a given finite $p$-group $D$, there are only finitely many pairs $(G, b)$, where $G$ is a finite group and $b$ is a block idempotent of $kG$, up to functorial equivalence over $\mathbb{F}$ [Bouc and Yılmaz 2022, Theorem 10.6] and we provided a sufficient condition for two blocks to be functorially equivalent over $\mathbb{F}$ in the situation of Broué's abelian defect group conjecture [Bouc and Yılmaz 2022, Theorem 11.1].

In this paper, we introduce the notion of *stable* diagonal $p$-permutation functors and *stable* functorial equivalences. We denote by $\overline{Rpp^{\Delta}_k}$ the quotient category of $Rpp^{\Delta}_k$ by the morphisms that factor through the trivial group. A *stable diagonal $p$-permutation functor over* $R$ is an $R$-linear functor from $\overline{Rpp^{\Delta}_k}$ to ${}_R\mathsf{Mod}$, or equivalently, a diagonal $p$-permutation functor which vanishes at the trivial group. In particular, the simple diagonal $p$-permutation functors $S_{L,u,V}$ with $L \neq 1$ are (simple) stable diagonal $p$-permutation functors. Our first main result is the following.

**Theorem 1.1.** *The category $\overline{\mathcal{F}^{\Delta}_{\mathbb{F}pp_k}}$ of stable diagonal $p$-permutation functors over $\mathbb{F}$ is semisimple. The simple stable diagonal $p$-permutation functors are precisely the simple diagonal $p$-permutation functors $S_{L,u,V}$ with $L \neq 1$.*

Given a finite group $G$ and a block idempotent $b$ of $kG$, we define a stable diagonal $p$-permutation functor $\overline{RT^{\Delta}_{G,b}}$ similar to $RT^{\Delta}_{G,b}$; see Definition 4.1. Note that $\overline{RT^{\Delta}_{G,b}}$ is the zero functor if and only if $b$ has defect 0. We say that two pairs $(G, b)$ and $(H, c)$ are *stably functorially equivalent over* $R$ if the functors $\overline{RT^{\Delta}_{G,b}}$ and $\overline{RT^{\Delta}_{H,c}}$ are isomorphic. For a block algebra $kGb$, let $k(kGb)$ and $l(kGb)$ denote the number of irreducible ordinary characters and the number of irreducible Brauer characters of $b$, respectively.

**Theorem 1.2.** *Let $b$ be a block idempotent of $kG$ and let $c$ be a block idempotent of $kH$.*

(i) *The pairs $(G, b)$ and $(H, c)$ are stably functorially equivalent over $\mathbb{F}$ if and only if the multiplicities of $S_{L,u,V}$ in $\mathbb{F}T^{\Delta}_{G,b}$ and $\mathbb{F}T^{\Delta}_{H,c}$ are the same for any simple diagonal $p$-permutation functor $S_{L,u,V}$ with $L \neq 1$. In this case, $(G, b)$ and $(H, c)$ are functorially equivalent over $\mathbb{F}$ if and only if $l(kGb) = l(kHc)$.*

(ii) *If the pairs $(G, b)$ and $(H, c)$ are stably functorially equivalent over $\mathbb{F}$, then $b$ and $c$ have isomorphic defect groups and one has*

$$k(kGb) - l(kGb) = k(kHc) - l(kHc).$$

We also consider the blocks with abelian defect groups and Frobenius inertial quotient.

**Theorem 1.3.** *Let $G$ be a finite group, $b$ a block idempotent of $kG$ with a nontrivial abelian defect group $D$. Let $E = N_G(D, e_D)/C_G(D)$ denote the inertial quotient of $b$. Suppose that $E$ acts freely on $D \setminus \{1\}$. Then*:

(i) *There exists a functorial equivalence over $\mathbb{F}$ between $(G, b)$ and $(D \rtimes E, 1)$ if and only if $l(kGb) = l(k(D \rtimes E))$.*

(ii) *Suppose that $E$ is abelian. Then there exists a functorial equivalence over $\mathbb{F}$ between $(G, b)$ and $(D \rtimes E, 1)$ if and only if $(G, b)$ and $(D \rtimes E, 1)$ are $p$-permutation equivalent.*

In Section 2 we recall diagonal $p$-permutation functors and functorial equivalences between blocks. In Section 3 we introduce the category of stable diagonal $p$-permutation functors and prove Theorem 1.1. In Section 4 we introduce the notion of stable functorial equivalences between blocks and prove Theorem 1.2. Finally, in Section 5 we prove Theorem 1.3.

## 2. Preliminaries

(a) Let $(P, s)$ be a pair, where $P$ is a $p$-group and $s$ is a generator of a $p'$-group acting on $P$. We write $P\langle s\rangle := P \rtimes \langle s\rangle$ for the corresponding semidirect product. We say that two pairs $(P, s)$ and $(Q, t)$ are *isomorphic* and write $(P, s) \cong (Q, t)$, if there is a group isomorphism $f : P\langle s\rangle \to Q\langle t\rangle$ that sends $s$ to a conjugate of $t$. We set $\mathrm{Aut}(P, s)$ to be the group of the automorphisms of the pair $(P, s)$ and $\mathrm{Out}(P, s) = \mathrm{Aut}(P, s)/\mathrm{Inn}(P\langle s\rangle)$. Recall from [Bouc and Yılmaz 2020] that a pair $(P, s)$ is called a *$D^\Delta$-pair*, if $C_{\langle s\rangle}(P) = 1$.

(b) Let $G$ and $H$ be finite groups. We denote by $T(G)$ the Grothendieck group of $p$-permutation $kG$-modules and by $T^\Delta(G, H)$ the Grothendieck group of $p$-permutation $(kG, kH)$-bimodules whose indecomposable direct summands have twisted diagonal vertices. Let $Rpp_k^\Delta$ denote the following category:

- objects: finite groups.
- $\mathrm{Mor}_{Rpp_k^\Delta}(G, H) = R \otimes_{\mathbb{Z}} T^\Delta(H, G) = RT^\Delta(H, G)$.
- composition is induced from the tensor product of bimodules.
- $\mathrm{Id}_G = [kG]$.

An $R$-linear functor from $Rpp_k^\Delta$ to ${}_R\mathsf{Mod}$ is called a *diagonal p-permutation functor* over $R$. Together with natural transformations, diagonal $p$-permutation functors form an abelian category $\mathcal{F}_{Rpp_k}^\Delta$.

(c) Recall from [Bouc and Yılmaz 2022] that the category $\mathcal{F}_{\mathbb{F}pp_k}^\Delta$ is semisimple. Moreover, the simple diagonal $p$-permutation functors, up to isomorphism, are parametrized by the isomorphism classes of triples $(L, u, V)$, where $(L, u)$ is a $D^\Delta$-pair, and $V$ is a simple $\mathbb{F}\mathrm{Out}(L, u)$-module (see [loc. cit., Sections 6 and 7] for more details on simple functors).

(d) Let $G$ be a finite group and $b$ a block idempotent of $kG$. Recall from [loc. cit.] that the block diagonal $p$-permutation functor $RT_{G,b}^\Delta$ is defined as

$$RT_{G,b}^\Delta : Rpp_k^\Delta \to {}_R\mathsf{Mod}, \quad H \mapsto RT^\Delta(H, G) \otimes_{kG} kGb.$$

See [loc. cit., Section 8] for the decomposition of $\mathbb{F}T_{G,b}^\Delta$ in terms of the simple functors $S_{L,u,V}$.

(e) Let $b$ be a block idempotent of $kG$ and let $c$ be a block idempotent of $kH$. We say that the pairs $(G, b)$ and $(H, c)$ are *functorially equivalent* over $R$, if the corresponding diagonal $p$-permutation functors $RT_{G,b}^\Delta$ and $RT_{H,c}^\Delta$ are isomorphic in $\mathcal{F}_{Rpp_k}^\Delta$ [loc. cit., Definition 10.1]. By [loc. cit., Lemma 10.2] the pairs $(G, b)$ and $(H, c)$ are functorially equivalent over $R$ if and only if there exists $\omega \in bRT^\Delta(G, H)c$ and $\sigma \in cRT^\Delta(H, G)b$ such that

$$\omega \cdot_H \sigma = [kGb] \quad \text{in} \quad bRT^\Delta(G, G)b \quad \text{and} \quad \sigma \cdot_G \omega = [kHc] \quad \text{in} \quad cRT^\Delta(H, H)c.$$

## 3. Stable diagonal $p$-permutation functors

In this section we introduce the category of stable diagonal $p$-permutation functors.

For a finite group $G$, let $P(G)$ denote the subgroup of $T(G)$ generated by the indecomposable projective $kG$-modules. Let also $\overline{T(G)}$ denote the quotient group $T(G)/P(G)$. For $X \in T(G)$, we denote by $\overline{X}$ the image of $X$ in $\overline{T(G)}$. If $H$ is another finite group, we define $P(G, H)$ and $\overline{T^\Delta(G, H)}$ similarly.

**Lemma 3.1.** *For finite groups $G$ and $H$ one has $P(G, H) = T^\Delta(G, 1) \circ T^\Delta(1, H)$.*

*Proof.* This follows from the fact that the projective indecomposable $k(G \times H)$-modules are of the form $P \otimes_k Q$, where $P$ and $Q$ are projective indecomposable $kG$ and $kH$-modules, respectively. □

**Definition 3.2.** Let $\overline{Rpp_k^\Delta}$ denote the following category:

- objects: finite groups.
- $\mathrm{Mor}_{\overline{Rpp_k^\Delta}}(G, H) = R \otimes_{\mathbb{Z}} \overline{T^\Delta(H, G)} = \overline{RT^\Delta(H, G)}$.
- composition is induced from the tensor product of bimodules.
- $\mathrm{Id}_G = \overline{[kG]}$.

**Definition 3.3.** An $R$-linear functor $\overline{Rpp_k^\Delta} \to {}_R\mathsf{Mod}$ is called a *stable diagonal $p$-permutation functor* over $R$. Together with natural transformations, stable diagonal $p$-permutation functors form an abelian category $\overline{\mathcal{F}_{Rpp_k}^\Delta}$.

**Remark 3.4.** The functor

$$\Gamma : \overline{\mathcal{F}_{Rpp_k}^\Delta} \to \mathcal{F}_{Rpp_k}^\Delta$$

obtained by composition with the projection $Rpp_k^\Delta \to \overline{Rpp_k^\Delta}$ gives a description of $\overline{\mathcal{F}_{Rpp_k}^\Delta}$ as a full subcategory of $\mathcal{F}_{Rpp_k}^\Delta$. Moreover, $\Gamma$ has a left adjoint $\Sigma$, constructed as follows: If $F$ is a diagonal $p$-permutation functor over $R$ and $G$ is a finite group, set

$$\bar{F}(G) := F(G)/RT^\Delta(G, \mathbf{1})F(\mathbf{1}).$$

Then $\bar{F}$ is a diagonal $p$-permutation functor, equal to the quotient of $F$ by the subfunctor generated by $F(\mathbf{1})$. Obviously, $\bar{F}$ vanishes at the trivial group, so it is a stable diagonal $p$-permutation functor. The functor $\Sigma : F \mapsto \bar{F}$ is a left adjoint to the above functor $\Gamma$. In particular, $\overline{\mathcal{F}_{Rpp_k}^\Delta}$ is a reflective subcategory of $\mathcal{F}_{Rpp_k}^\Delta$.

Let $G$ be a finite group. Recall that by [Bouc and Yılmaz 2022, Corollary 8.23(i)], the multiplicity of the simple diagonal $p$-permutation functor $S_{1,1,\mathbb{F}}$ in the representable functor $\mathbb{F}T^\Delta(-, G)$ is equal to the number $l(kG)$ of the isomorphism classes of simple $kG$-modules. Let $\mathcal{I}(-, G)$ denote the sum of simple subfunctors of $\mathbb{F}T^\Delta(-, G)$ isomorphic to $S_{1,1,\mathbb{F}}$. Let also $\mathbb{F}\mathrm{Proj}(-, G)$ denote the subfunctor of $\mathbb{F}T^\Delta(-, G)$ sending a finite group $H$ to $\mathbb{F}\mathrm{Proj}(H, G)$.

**Lemma 3.5.** *The subfunctors $\mathcal{I}(-, G)$ and $\mathbb{F}\mathrm{Proj}(-, G)$ of the representable functor $\mathbb{F}T^\Delta(-, G)$ are equal.*

*Proof.* For finite groups $G$ and $H$, the number of isomorphism classes of projective indecomposable $k(G \times H)$-modules, or equivalently, the number of isomorphism classes of simple $k(G \times H)$-modules is equal to the number of conjugacy classes of $p'$-elements of $G \times H$. Hence the $\mathbb{F}$-dimension of the evaluation $\mathbb{F}\mathrm{Proj}(H, G)$ is equal to

$$l(k(G \times H)) = l(kG)l(kH)$$

which is equal to the $\mathbb{F}$-dimension of $l(kG)S_{1,1,\mathbb{F}}(H)$ by [Bouc and Yılmaz 2022, Corollary 8.23(i)], and hence to the $\mathbb{F}$-dimension of $\mathcal{I}(H, G)$.

Note that $\mathbb{F}\mathrm{Proj}(-, G)$ is equal to the functor

$$\mathbb{F}T^\Delta(-, 1) \circ \mathbb{F}T^\Delta(1, G).$$

Moreover $S_{L,u,V}(1) = 0$ for $L \neq 1$, and hence $\mathbb{F}T^\Delta(1, G) = \mathcal{I}(1, G)$. Therefore,

$$\mathbb{F}\mathrm{Proj}(-, G) = \mathbb{F}T^\Delta(-, 1) \circ \mathbb{F}T^\Delta(1, G) = \mathbb{F}T^\Delta(-, 1) \circ \mathcal{I}(1, G) \subseteq \mathcal{I}(-, G).$$

Since the $\mathbb{F}$-dimensions of $\mathbb{F}\mathrm{Proj}(H, G)$ and $\mathcal{I}(H, G)$ are the same for any finite group $H$, it follows that $\mathbb{F}\mathrm{Proj}(-, G) = \mathcal{I}(-, G)$. □

*Proof of Theorem 1.1.* For a finite group $G$, the representable diagonal $p$-permutation functor $\mathbb{F}T^{\Delta}(-,G)$ decomposes as a direct sum of simple functors $S_{L,u,V}$, and hence we have

$$\mathbb{F}T^{\Delta}(-,G)\cong\mathcal{I}(-,G)\bigoplus_{\substack{(L,u,V)\\ L\neq 1}} S_{L,u,V}^{m_{L,u,V}},$$

for some nonnegative integers $m_{L,u,V}$, where $(L,u,V)$ runs over a set of isomorphism classes of $D^{\Delta}$-pairs $(L,u)$ with $L\neq 1$, and simple $\mathbb{F}\mathrm{Out}(L,u)$-modules $V$. By Lemma 3.5, the representable stable diagonal $p$-permutation functor $\overline{\mathbb{F}T^{\Delta}(-,G)}$ is isomorphic to the direct sum

$$\bigoplus_{\substack{(L,u,V)\\ L\neq 1}} S_{L,u,V}^{m_{L,u,V}}$$

of simple diagonal $p$-permutation functors, and each of these simple functors is a simple stable diagonal $p$-permutation functor. Since the functor category $\overline{\mathcal{F}^{\Delta}_{\mathbb{F}pp_k}}$ is generated by the representable functors the result follows. □

## 4. Stable functorial equivalences

Let $G$ and $H$ be finite groups.

**Definition 4.1.** Let $b$ a block idempotent of $kG$. The stable diagonal $p$-permutation functor $\overline{RT^{\Delta}_{G,b}}$ is defined as

$$\overline{RT^{\Delta}_{G,b}}:\overline{Rpp^{\Delta}_k}\to {}_R\mathsf{Mod},\qquad H\mapsto \overline{RT^{\Delta}(H,G)\otimes_{kG}kGb}.$$

See Section 2(d) for the definition of $RT^{\Delta}_{G,b}$ and note that $\overline{RT^{\Delta}_{G,b}}=\Sigma(RT^{\Delta}_{G,b})$.

**Definition 4.2.** Let $b$ be a block idempotent of $kG$ and let $c$ be a block idempotent of $kH$. We say that the pairs $(G,b)$ and $(H,c)$ are *stably functorially equivalent over* $R$, if their corresponding stable diagonal $p$-permutation functors $\overline{RT^{\Delta}_{G,b}}$ and $\overline{RT^{\Delta}_{H,c}}$ are isomorphic in $\overline{\mathcal{F}^{\Delta}_{Rpp_k}}$.

**Lemma 4.3.** *Let $b$ be a block idempotent of $kG$ and let $c$ be a block idempotent of $kH$.*

(a) *The pairs $(G,b)$ and $(H,c)$ are stably functorially equivalent over $R$ if and only if there exists $\omega\in bRT^{\Delta}(G,H)c$ and $\sigma\in cRT^{\Delta}(H,G)b$ such that*

$$\omega\cdot_H\sigma=[kGb]+[P] \text{ in } bRT^{\Delta}(G,G)b \text{ and } \sigma\cdot_G\omega=[kHc]+[Q] \text{ in } cRT^{\Delta}(H,H)c$$

*for some $P\in R\mathrm{Proj}(kGb,kGb)$ and $Q\in R\mathrm{Proj}(kHc,kHc)$.*

(b) *If the pairs $(G,b)$ and $(H,c)$ are functorially equivalent over $R$, then they are also stably functorially equivalent over $R$.*

*Proof.* By the Yoneda lemma, the pairs $(G, b)$ and $(H, c)$ are stably functorially equivalent over $R$ if and only if there exists $\bar{\omega} \in \overline{bRT^\Delta(G, H)c}$ and $\bar{\sigma} \in \overline{cRT^\Delta(H, G)b}$ such that

$$\bar{\omega} \cdot_H \bar{\sigma} = \overline{[kGb]} \quad \text{in} \quad \overline{bRT^\Delta(G, G)b} \quad \text{and} \quad \bar{\sigma} \cdot_G \bar{\omega} = \overline{[kHc]} \quad \text{in} \quad \overline{cRT^\Delta(H, H)c}.$$

Hence (a) follows and (b) is clear. □

*Proof of Theorem 1.2.* (i) The first statement follows from Theorem 1.1 and the second statement follows since by [Bouc and Yılmaz 2022, Corollary 8.23(i)] the multiplicity of the simple functor $S_{1,1,\mathbb{F}}$ in $\mathbb{F}T^\Delta_{G,b}$ is equal to $l(kGb)$.

(ii) The first statement follows from the proof of [loc. cit., Theorem 10.5(iii)] and the second statement follows from the proof of [loc. cit., Theorem 10.5(ii)]. □

## 5. Blocks with Frobenius inertial quotient

(a) Recall the assumptions of Theorem 1.3: Let $G$ be a finite group, $b$ a block idempotent of $kG$ with a nontrivial abelian defect group $D$. Let $(D, e_D)$ be a maximal $b$-Brauer pair and let $E = N_G(D, e_D)/C_G(D)$ denote the inertial quotient of $b$. Suppose that $E$ acts freely on $D\backslash\{1\}$. This condition is equivalent to requiring $D \rtimes E$ be a Frobenius group. Let $\mathcal{F}_b$ be the fusion system of $b$ with respect to $(D, e_D)$. Then $\mathcal{F}_b$ is equal to the fusion system $\mathcal{F}_{D\rtimes E}(D)$ on $D$ determined by $D \rtimes E$.

Under these assumptions, we know from [Linckelmann 2018, Theorem 10.5.1] that there is a stable equivalence of Morita type between $(G, b)$ and $(D \rtimes E, 1)$, induced by a bimodule with an endopermutation source. Since $D$ is abelian, by [Rickard 1996, Theorem 7.2] this bimodule admits an endosplit $p$-permutation resolution, which yields a stable $p$-permutation equivalence between $(G, b)$ and $(D \rtimes E, 1)$; see, for instance, [Linckelmann 2018, Remark 9.11.7]. In particular $(G, b)$ and $(D \rtimes E, 1)$ are stably functorially equivalent.

Hereafter we give an alternative proof of the existence of this stable functorial equivalence, which relies only on Theorem 1.1. Together with Theorem 1.2, this now implies Part (i) of Theorem 1.3. Part (ii) of Theorem 1.3 follows from Part (i) and [Linckelmann 2018, Theorem 10.5.10].

(b) Let $S_{L,u,V}$ be a simple diagonal $p$-permutation functor such that $L$ is nontrivial and isomorphic to a subgroup of $D$. Recall that by [Bouc and Yılmaz 2022, Theorem 8.22] the multiplicity of $S_{L,u,V}$ in $\mathbb{F}T^\Delta_{G,b}$ is equal to the $\mathbb{F}$-dimension of

$$\bigoplus_{(P,e_P)\in[\mathcal{F}_b]} \; \bigoplus_{\pi\in[N_G(P,e_P)\backslash\mathcal{P}_{(P,e_P)}(L,u)/\mathrm{Aut}(L,u)]} \mathbb{F}\mathrm{Proj}(ke_PC_G(P), u) \otimes_{\mathrm{Aut}(L,u)_{\overline{(P,e_P,\pi)}}} V,$$

where $[\mathcal{F}_b]$ denotes a set of isomorphism classes of objects in $\mathcal{F}_b$, $\mathcal{P}_{(P,e_P)}$ is the set of group isomorphisms $\pi : L \to P$ with $\pi i_u \pi^{-1} \in \mathrm{Aut}_{\mathcal{F}_b}(P, e_P)$, and $\mathrm{Aut}(L, u)_{\overline{(P,e_P,\pi)}}$ is the stabilizer in $\mathrm{Aut}(L, u)$ of the $G$-orbit of $(P, e_P, \pi)$. Since $b$

is a block with Frobenius inertial quotient, the block $kC_G(P)e_P$ is nilpotent for every nontrivial subgroup $P$ of $D$; see, for instance, [Linckelmann 2018, Theorem 10.5.2]. Therefore, we have $l(ke_PC_G(P)) = 1$, and hence the multiplicity formula reduces to

$$\bigoplus_{(P,e_P)\in[\mathcal{F}_b]} \bigoplus_{\pi\in[N_G(P,e_P)\backslash\mathcal{P}_{(P,e_P)}(L,u)/\mathrm{Aut}(L,u)]} V^{\mathrm{Out}(L,u)_{\overline{(P,e_P,\pi)}}}.$$

Let $\mathcal{Q}_{D\rtimes E,p}$ denote the set of pairs $(P,s)$ of $p$-subgroups $P$ of $D\rtimes E$ and $p'$-elements $s$ of $N_{D\rtimes E}(P)$. Let also $[\mathcal{Q}_{D\rtimes E,p}]$ denote a set of representatives of $D\rtimes E$-orbits on $\mathcal{Q}_{D\rtimes E,p}$ under the conjugation map. Recall from [Bouc and Yılmaz 2022, Corollary 7.4] that the multiplicity of $S_{L,u,V}$ in $\mathbb{F}T^{\Delta}_{D\rtimes E}$ is equal to the $\mathbb{F}$-dimension of

$$\bigoplus_{\substack{(P,s)\in[\mathcal{Q}_{D\rtimes E,p}]\\ (\tilde{P},\tilde{s})\cong(L,u)}} V^{N_{D\rtimes E}(P,s)},$$

where for a pair $(P,s)\in\mathcal{Q}_{D\rtimes E,p}$ with $(\tilde{P},\tilde{s})\cong(L,u)$, we fix an isomorphism $\phi_{P,s}: L\to P$ with $\phi_{P,s}({}^u l) = {}^s\phi_{P,s}(l)$ for all $l\in L$ and we view $V$ as an $\mathbb{F}N_{D\rtimes E}(P,s)$-module via the group homomorphism

$$N_G(P,s)\to \mathrm{Out}(L,u) \tag{1}$$

that sends $g\in N_G(P,s)$ to the image of $\phi_{P,s}^{-1}\circ i_g\circ\phi_{P,s}$ in $\mathrm{Out}(L,u)$.

(c) Let $\mathcal{P}_b(G,L,u)$ denote the set of triples $(P,e,\pi)$ where $(P,e)\in\mathcal{F}_b$ and $\pi\in\mathcal{P}_{(P,e_P)}(L,u)$. Let also $\mathcal{Q}_{D\rtimes E,p}(L,u)$ denote the set of pairs $(P,s)$ in $\mathcal{Q}_{D\rtimes E,p}$ with the property that $(\tilde{P},\tilde{s})\cong(L,u)$.

If $(P,e,\pi)\in\mathcal{P}_b(G,L,u)$, then $\pi i_u\pi^{-1}\in\mathrm{Aut}_{\mathcal{F}_b}(P,e_P)$ by definition and since $\mathcal{F}_b$ is equal to $\mathcal{F}_{D\rtimes E}(D)$, it follows that there exists a $p'$-element $s$ of $N_{D\rtimes E}(P)$ with $\pi i_u\pi^{-1} = i_s$. This implies by [Bouc and Yılmaz 2022, Lemma 3.3] that $(\tilde{P},\tilde{s})\cong(L,u)$ and therefore we have a map

$$\Psi:\mathcal{P}_b(G,L,u)\to\mathcal{Q}_{D\rtimes E,p}(L,u),\quad (P,e,\pi)\mapsto(P,s).$$

**Lemma 5.1.** *The map $\Psi$ induces a bijection*

$$\overline{\Psi}:[G\backslash\mathcal{P}_b(G,L,u)/\mathrm{Aut}(L,u)]\to[\mathcal{Q}_{D\rtimes E,p}(L,u)].$$

*Proof.* First we show that the map $\overline{\Psi}$ is well-defined. Let $(P,e,\pi)$ and $(Q,f,\rho)$ be two elements in $\mathcal{P}_b(G,L,u)$ that lie in the same $G\times\mathrm{Aut}(L,u)$-orbit. We need to show that $\overline{\Psi}(P,e,\pi) = \overline{\Psi}(Q,f,\rho)$. Write $\Psi(P,e,\pi) = (P,s)$ and $\Psi(Q,f,\rho) = (Q,t)$. Let $g\in G$ and $\varphi\in\mathrm{Aut}(L,u)$ such that

$$g\cdot(P,e,\pi)\cdot\varphi = (Q,f,\rho).$$

Then $(P,e)$ and $(Q,f)$ lie in the same isomorphism class in $[\mathcal{F}_b]$ and hence $P$ and $Q$ are $D\rtimes E$-conjugate since $\mathcal{F}_b = \mathcal{F}_{D\rtimes E}(D)$. Thus, there exists $h\in D\rtimes E$

with $i_g = i_h : P \to Q$. Hence $\rho = i_g \pi \varphi = i_h \pi \varphi : L \to Q$. Since $\varphi \in \mathrm{Aut}(L, u)$, one has $\varphi \circ i_u = i_u \circ \varphi$. Therefore,

$$i_t = \rho i_u \rho^{-1} = i_h \pi \varphi i_u \varphi^{-1} \pi^{-1} i_{h^{-1}} = i_h \pi i_u \pi^{-1} i_{h^{-1}} = i_h i_s i_{h^{-1}} = i_{hsh^{-1}}.$$

This shows that $(Q, t) = h \cdot (P, s)$ and hence the map $\overline{\Psi}$ is well-defined.

Now we show that $\overline{\Psi}$ is surjective. Let $(P, s) \in \mathcal{Q}_{D \rtimes E, p}(L, u)$. Since $(\tilde{P}, \tilde{s}) \cong (L, u)$, again by [Bouc and Yılmaz 2022, Lemma 3.3], there exists $\pi : L \to P$ such that $\pi i_u = i_s \pi$, i.e., $\pi i_u \pi^{-1} = i_s : P \to P$. Since $\mathcal{F}_{D \rtimes E}(D) = \mathcal{F}_b$, it follows that there exists $g \in N_G(P, e)$ with $i_s = i_g$, and hence $(P, e, \pi) \in \mathcal{P}_b(G, L, u)$ with $\overline{\Psi}(P, e, \pi) = (P, s)$. Thus, $\overline{\Psi}$ is surjective.

Finally, we show that $\overline{\Psi}$ is injective. Let $(P, e, \pi), (Q, f, \rho) \in \mathcal{P}_b(G, L, u)$ be elements with $\overline{\Psi}(P, e, \pi) = \overline{\Psi}(Q, f, \rho)$. Write $(P, s) = \Psi(P, e, \pi)$ and $(Q, t) = \Psi(Q, f, \rho)$. Then there exists $h \in D \rtimes E$ such that

$$h \cdot (P, s) = (Q, t).$$

Again, there exists $g \in G$ such that $i_g = i_h : P \to Q$. Define

$$\varphi := \pi^{-1} \circ i_g^{-1} \circ \rho : L \to L.$$

One has

$$\begin{aligned} \varphi \circ i_u &= \pi^{-1} \circ i_g^{-1} \circ \rho \circ i_u = \pi^{-1} \circ i_g^{-1} \circ i_t \circ \rho = \pi^{-1} \circ i_g^{-1} \circ i_g \circ i_s \circ i_g^{-1} \circ \rho \\ &= \pi^{-1} \circ i_s \circ i_g^{-1} \circ \rho = i_u \circ \pi^{-1} \circ i_g^{-1} \circ \rho = i_u \circ \varphi \end{aligned}$$

which shows that $\varphi \in \mathrm{Aut}(L, u)$. Moreover, one has

$$g \cdot (P, e, \pi) \cdot \varphi = (Q, f, \rho)$$

and so the map $\overline{\Psi}$ is injective. □

**Lemma 5.2.** *Let* $(P, e, \pi) \in [G \backslash \mathcal{P}_b(G, L, u)/\mathrm{Aut}(L, u)]$ *and* $(P, s) = \overline{\Psi}(P, e, \pi) \in [\mathcal{Q}_{D \rtimes E, p}]$. *Then the image of* $N_{D \rtimes E}(P, s)$ *in* $\mathrm{Out}(L, u)$ *is equal to* $\mathrm{Out}(L, u)_{\overline{(P,e,\pi)}}$.

*Proof.* We have $\pi i_u \pi^{-1} = i_s$ and hence the image of $N_{D \rtimes E}(P, s)$ is given by

$$N_{D \rtimes E}(P, s) \to \mathrm{Out}(L, u), \quad h \mapsto \overline{\pi^{-1} \circ i_h \circ \pi}$$

Note that since ${}^h s = s$, we have $i_h i_s = i_s i_h$, i.e., $i_h \pi i_u \pi^{-1} = \pi i_u \pi^{-1} i_h$. Therefore the image is

$$\begin{aligned} &\{\overline{\pi^{-1} i_h \pi} \mid h \in D \rtimes E, i_h : P \to P, {}^h s = s\} \\ &\qquad = \{\overline{\pi^{-1} i_g \pi} \mid g \in N_G(P, e), i_g \pi i_u \pi^{-1} = \pi i_u \pi^{-1} i_g\} \\ &\qquad = \{\overline{\pi^{-1} i_g \pi} \in \mathrm{Out}(L, u) \mid \pi^{-1} i_g \pi = i_g, g \in N_G(P, e)\} \\ &\qquad = \mathrm{Out}(L, u)_{\overline{(P,e,\pi)}} \end{aligned}$$

as was to be shown. □

*Proof of Theorem 1.3.* We need to show that for any $L \neq 1$, the multiplicities of a simple diagonal $p$-permutation functor $S_{L,u,V}$ in $\mathbb{F}T^{\Delta}_{G,b}$ and in $\mathbb{F}T^{\Delta}_{D\rtimes E}$ are equal. But this follows from Lemmas 5.1 and 5.2. Now Part (i) follows from Theorem 1.2(i), and Part (ii) follows from [Linckelmann 2018, Theorem 10.5.10]. □

## Acknowledgement

The second author was supported by the BAGEP Award of the Science Academy.

## References

[Boltje and Perepelitsky 2020] R. Boltje and P. Perepelitsky, "$p$-permutation equivalences between blocks of group algebras", 2020. arXiv 2007.09253

[Boltje and Xu 2008] R. Boltje and B. Xu, "On $p$-permutation equivalences: between Rickard equivalences and isotypies", *Trans. Amer. Math. Soc.* **360**:10 (2008), 5067–5087. MR Zbl

[Bouc and Yılmaz 2020] S. Bouc and D. Yılmaz, "Diagonal $p$-permutation functors", *J. Algebra* **556** (2020), 1036–1056. MR Zbl

[Bouc and Yılmaz 2022] S. Bouc and D. Yılmaz, "Diagonal $p$-permutation functors, semisimplicity, and functorial equivalence of blocks", *Adv. Math.* **411** (2022), art. id. 108799. MR Zbl

[Broué 1990] M. Broué, "Isométries parfaites, types de blocs, catégories dérivées", pp. 61–92 in *Représentations linéaires des groupes finis*, Astérisque **181-182**, 1990. MR Zbl

[Linckelmann 2018] M. Linckelmann, *The block theory of finite group algebras*, *II*, Cambridge University Press, 2018. MR Zbl

[Rickard 1996] J. Rickard, "Splendid equivalences: derived categories and permutation modules", *Proc. London Math. Soc.* (3) **72**:2 (1996), 331–358. MR Zbl

Received November 5, 2023.

SERGE BOUC
CNRS-LAMFA
UNIVERSITÉ DE PICARDIE
AMIENS
FRANCE
serge.bouc@u-picardie.fr

DENİZ YILMAZ
DEPARTMENT OF MATHEMATICS
BILKENT UNIVERSITY
ANKARA
TURKEY
d.yilmaz@bilkent.edu.tr

PACIFIC JOURNAL OF MATHEMATICS
Vol. 328, No. 2, 2024
https://doi.org/10.2140/pjm.2024.328.227

# LORENTZ–SHIMOGAKI AND ARAZY–CWIKEL THEOREMS REVISITED

LÉONARD CADILHAC, FEDOR SUKOCHEV AND DMITRIY ZANIN

**By using the space $L_0$ of finitely supported functions as a left endpoint on the interpolation scale of $L_p$-spaces, we present a new approach to the Lorentz–Shimogaki and Arazy–Cwikel theorems which covers the whole range of $p, q \in (0, \infty]$. In particular, we show that for $0 \leq p < q < r < s \leq \infty$,**

$$\mathrm{Int}(L_q, L_r) = \mathrm{Int}(L_p, L_r) \cap \mathrm{Int}(L_q, L_s)$$

**if the underlying space is $(0, \alpha)$, $\alpha \in (0, \infty]$ equipped with the Lebesgue measure. As a byproduct of our result, we solve a conjecture of Levitina, Sukochev and Zanin (2020).**

## 1. Introduction

Descriptions of interpolation spaces for couples of $L_p$-spaces for $1 \leq p \leq \infty$ were extensively researched from the 1960s to the 80s, providing satisfying answers to most problems that were considered relevant at the time.

However, new questions arising from noncommutative analysis recently highlighted some gaps in our knowledge of this subject, especially for the case of $p < 1$ of quasi-Banach spaces. In this paper, we revisit some important results of the literature [1; 21; 27], generalizing them and thus filling some of the holes that were revealed in the theory. In particular, we answer a question asked by Levitina, Sukochev and Zanin in [20] and already partially studied in [11] regarding the interpolation theory of sequence spaces (see Theorem 1.2). Besides this new result, this paper introduces a general approach that covers the range of all $0 \leq p \leq \infty$ and is self-contained. It emphasizes the use of the space $L_0$ of all finitely supported measurable functions. As far as the authors know this space rarely appears in interpolation theory (however, see [2; 16; 24] and [15]). We provide evidence that $L_0$ is a suitable "left endpoint" on the interpolation scale of $L_p$-spaces, despite its possessing an atypical structure, that of a normed abelian group.

A function space $E$ is an *interpolation space* for the couple $(L_p, L_q)$ if any linear operator $T$ bounded on $L_p$ and $L_q$ is also bounded on $E$ (see Definition 2.4). This

*MSC2020:* 43A15, 46B70, 46M35.
*Keywords:* interpolation, symmetric function spaces, $L_p$-spaces, majorization.

notion provides a way of transferring inequalities well known in $L_p$-spaces to more exotic ones. To both understand the range of applicability of this technique and be able to check whether it applies to a given function space $E$, we are interested in simple descriptions of interpolation spaces for the couple $(L_p, L_q)$.

This problem has a long history starting with the Calderón–Mityagin theorem (see [9; 22]) on the couple $(L_1, L_\infty)$ and followed by Lorentz and Shimogaki's [21] results on the couples $(L_1, L_q)$ and $(L_p, L_\infty)$ with $1 \le p, q \le \infty$. A remarkable result of Arazy and Cwikel then states that a space $E$ is an interpolation space for the couple $(L_p, L_q)$, $1 < p < q < \infty$ if and only if it is an interpolation space for the couples $(L_1, L_q)$ and $(L_p, L_\infty)$.

Describing interpolation spaces often comes down to understanding certain orders. In fact, at a very fundamental level, being an interpolation space can be understood as a monotonicity property. Indeed, given two compatible quasi-Banach spaces $A, B$, denote by $\mathbb{C}(A, B)$ the set of operators $A + B \to A + B$ that restrict to contractions on $A$ and $B$. Consider the following order on $A + B$:

$$f \le_{p,q} g \Leftrightarrow \exists T \in \mathbb{C}(A, B), \quad T(g) = f.$$

With this definition in mind, $E$ is an interpolation space for the couple $(A, B)$ if and only if

$$\forall f \in E, \forall g \in L_p + L_q, \quad g \le_{p,q} f \Rightarrow g \in E.$$

In fact, the fundamental theorem of Calderón and Mityagin precisely describes the order $\le_{L_1,L_\infty}$ (from now on denoted by $\prec\prec_{\mathrm{hd}}$). It states that for $f, g \in L_1 + L_\infty$,

$$g \prec\prec_{\mathrm{hd}} f \Leftrightarrow g \le_{L_1,L_\infty} f \Leftrightarrow \forall t > 0, \quad \int_0^t \mu(s, g)\, ds \le \int_0^t \mu(s, f)\, ds,$$

where $\mu(g) : t \to \mu(t, g)$ denotes the right-continuous decreasing rearrangement of $g$. We will call this order *head majorization*. Moreover, if $f, g \in L_1$ and $\|f\|_1 = \|g\|_1$, then we write $g \prec\prec_{\mathrm{hd}} f$. Variants of this order allow to describe interpolation spaces for any couple $(L_p, L_\infty)$, $p \in (0, \infty)$ (see [21] for the Banach range and [8] for $p < 1$).

Another phenomenon, this time specific to the study of interpolation theory of $L_p$-spaces, is that to guarantee that a space $E$ is an interpolation space for the couple $(L_p, L_q)$, $p < q$ it is natural to impose two conditions: one which will impose that $E$ is "on the right of $L_p$" and one that will impose that $E$ is "on the left of $L_q$". An example of such a result is the above-mentioned Arazy–Cwikel theorem but one can think also of convexity/concavity conditions or Boyd indices (see [18] for an overview and [8, Theorem 1.4]).

In this spirit, the natural counterpart of head majorization is *tail majorization* defined on $L_0 + L_1$ by

$$g \prec\prec_{\text{tl}} f \Leftrightarrow \forall t > 0, \quad \int_t^\infty \mu(s, g)\, ds \le \int_t^\infty \mu(s, f)\, ds.$$

We'll show later that this order is in fact equivalent to $\le_{L_0,L_1}$. Moreover, if $f, g \in L_1$ and $\|f\|_1 = \|g\|_1$, then we write $g \prec_{\text{tl}} f$. Remark that $g \prec_{\text{tl}} f$ if and only if $f \prec_{\text{hd}} g$. Note that tail majorization coincides with the weak supermajorization of [13].

Let us now state our main theorem. Let $\mathcal{X}$ be the linear space of all measurable functions. If not specified otherwise, the *underlying measure space* we are working on is $(0, \infty)$ equipped with the Lebesgue measure $m$. We obtain:

**Theorem 1.1.** *Let $E \subset \mathcal{X}$ be a quasi-Banach function space (a priori, not necessarily symmetric). Let $p, q \in (0, \infty)$ such that $p < q$. Then:*

(a) *$E$ is an interpolation space for the couple $(L_p, L_\infty)$ if and only if there exists $c_{p,E} > 0$ such that for any $f \in E$ and $g \in L_p + L_\infty$,*

$$|g|^p \prec\prec_{\text{hd}} |f|^p \Rightarrow g \in E \quad \textit{and} \quad \|g\|_E \le c_{p,E}\|f\|_E.$$

(b) *$E$ is an interpolation space for the couple $(L_0, L_q)$ if and only if there exists $c_{q,E} > 0$ such that for any $f \in E$ and $g \in L_0 + L_q$,*

$$|g|^q \prec\prec_{\text{tl}} |f|^q \Rightarrow g \in E \quad \textit{and} \quad \|g\|_E \le c_{q,E}\|f\|_E.$$

(c) *$E$ is an interpolation space for the couples $(L_0, L_q)$ and $(L_p, L_\infty)$ if and only if it is an interpolation space for the couple $(L_p, L_q)$.*

This extends the results of Lorentz–Shimogaki and Arazy–Cwikel to the quasi-Banach setting and contributes to the two first questions asked by Arazy in [12, p. 232] in the particular case of $L_p$-spaces for $0 < p < \infty$. As mentioned before, our approach places $L_0$ as a left endpoint on the interpolation scale of $L_p$-spaces, in sharp contrast to earlier results which focused mostly on Banach spaces and had $L_1$ playing this part. An advantage of our approach is that it naturally encompasses every symmetric quasi-Banach space since they are all interpolation spaces for the couple $(L_0, L_\infty)$ (see [2; 16]). On the contrary, there exist some symmetric Banach spaces which are *not* interpolation spaces for the couple $(L_1, L_\infty)$ (see [26]). This led to some difficulties which were customarily circumvented with the help of various technical conditions such as the Fatou property (as appears, e.g., in [4]).

Compared to [8] where the first author investigates similar characterizations, the novelty of this theorem is statement (b) that deals with the space $L_0$. A deeper advantage of our new approach is that it no longer relies on Sparr's $K$-monotonicity result [27] for couples of $L_p$-spaces which was instrumental in [8].

Indeed, our strategy in this paper is different from the techniques used in [1; 2; 3; 4; 8; 9; 10; 11; 12; 16; 18; 21; 22; 27] and is based on partition lemmas, which were originally developed in a deep paper due to Braverman and Mekler [7] devoted to the study of the symmetric Banach function spaces $E$ such that the set $\{f \in E : f \prec\prec_{\mathrm{hd}} g\}$ coincides with the closure of the convex hull of its extreme points.

The approach of Braverman and Mekler was subsequently revised and redeveloped in [28] and precisely this revision constitutes the core of our approach in this paper.

We restate partition lemmas based on [28, Proposition 19] in Section 4. These lemmas allow us to restrict head and tail majorizations to very simple situations and reduce the problem to functions taking at most two values. Then, we deduce interpolation results from those structural lemmas.

Note that this scheme of proof is quite direct and in particular, does not involve at any point duality related arguments which apply only to Banach spaces [21] or more generally to $L$-convex quasi-Banach spaces [17; 25].

In Section 6, we pursue the same type of investigation, but in the setting of sequence spaces. The nondiffuse aspect of the underlying measure generates substantial technical difficulties. In particular, we require a new partition lemma which is not as efficient as those in Section 4 (compare Lemmas 6.2 and 4.6). This deficiency has been first pointed out to the authors by Cwikel. However, we are still able to resolve the conjecture of [20] (in the affirmative) by combining Lemma 6.1 with a Boyd-type argument which we borrow from Montgomery and Smith [23]. In particular, we substantially strengthen the results in [11]. Here is the precise statement:

**Theorem 1.2.** *Let $E \subset \ell_\infty$ be a quasi-Banach sequence space and $q \geq 1$. The following conditions are equivalent*:

(a) *There exists $p < q$ such that $E$ is an interpolation space for the couple $(\ell^p, \ell^q)$.*

(b) *There exists $c > 0$ such that for any $u \in E$ and $v \in \ell_\infty$,*

$$|v|^q \prec\prec_{\mathrm{tl}} |u|^q \Rightarrow v \in E \quad \textit{and} \quad \|v\|_E \leq c\|u\|_E.$$

(c) *For any $u \in E$ and $v \in \ell_\infty$,*

$$|v|^q \prec\prec_{\mathrm{tl}} |u|^q \Rightarrow v \in E.$$

In this section, we freely use results of Cwikel [10] and Cadilhac [8] to avoid repeating too many similar arguments.

Note that Theorem 1.2 was since proved independently in [5] where a deeper analysis of the interpolation scale of sequence spaces $\ell_p,\ 0 \leq p \leq \infty$ is presented. In particular, it is shown in [5] that for any $0 < p < q < \infty,\ E$ is an interpolation space for $(\ell_p, \ell_q)$ if and only if it is an interpolation space for $(\ell_0, \ell_q)$ and $(\ell_p, \ell_\infty)$, thus providing a counterpart to our main theorem in the sequence setting.

## 2. Preliminaries

***Interpolation spaces.*** The reader is referred to [6] for more details on interpolation theory and to [19] for an introduction to symmetric spaces. In the remainder of this section, $p$ and $q$ will denote two nonnegative reals such that $p \le q$.

Let $(\Omega, m)$ be any measure space (in particular the following definitions apply to $\mathbb{N}$ equipped with the counting measure, i.e., sequence spaces). As previously mentioned, $L_0(\Omega) \subset \mathcal{X}(\Omega)$ denotes the set of functions whose supports have finite measures, it is naturally equipped with the group norm

$$\|f\|_0 = m(\text{supp } f), \quad f \in L_0(\Omega).$$

The "norm" of a linear operator $T : L_0(\Omega) \to L_0(\Omega)$, is defined as

$$\|T\|_{L_0 \to L_0} = \sup_{f \in L_0} \frac{m(\text{supp}(Tf))}{m(\text{supp}(f))}.$$

**Definition 2.1.** A linear space $E \subset \mathcal{X}(\Omega)$ becomes a quasi-Banach function space when equipped with a complete quasinorm $\|\cdot\|_E$ such that:

- If $f \in E$ and $g \in \mathcal{X}(\Omega)$ are such that $|g| \le |f|$, then $g \in E$ and $\|g\|_E \le \|f\|_E$.

**Definition 2.2.** A quasi-Banach function space $E \subset \mathcal{X}(\Omega)$ is called symmetric if

- $f \in E$ and $g \in \mathcal{X}(\Omega)$ are such that $\mu(f) = \mu(g)$, then $g \in E$ and $\|g\|_E = \|f\|_E$.

**Definition 2.3** (bounded operator on a couple of quasi-Banach function spaces). Let $X$ and $Y$ be quasi-Banach function spaces. We say that a linear operator $T$ is bounded on $(X, Y)$ if $T$ is defined from $X + Y$ to $X + Y$ and restricts to a bounded operator from $X$ to $X$ and from $Y$ to $Y$. Set

$$\|T\|_{(X,Y)\to(X,Y)} = \max(\|T\|_{X\to X}, \|T\|_{Y\to Y}).$$

Let us recall the precise abstract definition of an interpolation space (see [6; 19]).

**Definition 2.4** (interpolation space between function spaces). Let $X$, $Y$ and $Z$ be either quasi-Banach function spaces on $\Omega$ or $L_0(\Omega)$. We say that $Z$ is an interpolation space for the couple $(X, Y)$ if $X \cap Y \subset Z \subset X + Y$ and any bounded operator on $(X, Y)$ restricts to a bounded operator on $Z$. Denote by $\text{Int}(X, Y)$ the set of interpolation spaces for the couple $(X, Y)$.

For quasi-Banach spaces, the above definition is equivalent to a seemingly stronger quantitative property.

**Proposition 2.5.** *Let $X$, $Y$, $Z$ be quasi-Banach function spaces. If $Z$ is an interpolation space for the couple $(X, Y)$, then there exists a constant $c(X, Y, Z) > 0$ such that for any bounded operator $T$ on $(X, Y)$,*

$$\|T\|_{Z\to Z} \le c(X, Y, Z) \cdot \|T\|_{(X,Y)\to(X,Y)}.$$

*The best possible value of* $c(X, Y, Z)$ *is called interpolation constant of* $Z$ *with respect to the couple* $(X, Y)$.

*Proof.* In [19, Lemma I.4.3], the assertion is proved for Banach spaces. The argument for quasi-Banach spaces is identical (because it relies on the closed graph theorem, which holds for $F$-spaces, and hence for quasi-Banach spaces). □

***K-functional and E-functional.*** In the remainder of the subsection, $X$, $Y$ and $Z$ will denote function spaces which are either quasi-Banach, or $L_0$.

**Definition 2.6.** Let $f \in X + Y$ and $t > 0$. Define

$$K_t(f, X, Y) := \inf_{g+h=f} \|g\|_X + t\|h\|_Y \quad \text{and} \quad E_t(f, X, Y) := \inf_{\|g\|_X \le t} \|f - g\|_Y.$$

These two notions are closely related to one another (see [24]) and the $K$-functional in particular plays a major role in the study of general interpolation spaces. Note that

$$K_t(f, L_1, L_\infty) = \int_0^t \mu(s, f)\, ds \quad \text{and} \quad E_t(f, L_0, L_1) = \int_t^\infty \mu(s, f)\, ds.$$

Thus the head and tail majorizations we consider can be in fact expressed in terms of $K$ and $E$ functionals. We say that $Z$ is $K$*-monotone* with respect to the couple $(X, Y)$ if $X \cap Y \subset Z \subset X + Y$ and for any $f \in Z,\ g \in X + Y$,

$$\forall t > 0, \quad K_t(g, X, Y) \le K_t(f, X, Y) \Rightarrow g \in Z.$$

Similarly, $Z$ is $E$*-monotone* with respect to the couple $(X, Y)$ if $X \cap Y \subset Z \subset X + Y$ and for any $f \in Z,\ g \in X + Y$,

$$\forall t > 0, \quad E_t(g, X, Y) \le E_t(f, X, Y) \Rightarrow g \in Z.$$

**Remark 2.7.** It is clear from the definitions that if $Z$ is either $E$-monotone or $K$-monotone for the couple $(X, Y)$ then $Z$ is an interpolation space for $(X, Y)$.

***Symmetry of interpolation spaces.*** In this subsection, we show that a quasi-Banach interpolation space for a couple of symmetric spaces can always be renormed into a symmetric space. Note that similar results can be found in the literature, see, for example, [19, Theorem 2.1].

As usual, we will use the term *measure preserving* for a measurable map $\omega$ between measure spaces $(\Omega_1, \mathcal{A}_1, m_1)$ and $(\Omega_2, \mathcal{A}_2, m_2)$ verifying,

$$\forall A \in \mathcal{A}_1, \qquad \omega(A) \in \mathcal{A}_2 \quad \text{and} \quad m_2(\omega(A)) = m_1(A).$$

**Lemma 2.8.** *Assume that* $\Omega$ *is* $(0, 1)$, $(0, \infty)$ *or* $\mathbb{N}$. *Let* $0 \le f, g \in L_0(\Omega) + L_\infty(\Omega)$ *and let* $\varepsilon > 0$. *Assume that* $\mu(f) = \mu(g)$. *There exists a measure preserving map* $\omega : \mathrm{supp}(g) \to \mathrm{supp}(f)$ *such that* $(1 + \varepsilon)(f \circ \omega) \ge g$.

*Proof.* **Case 1**. Suppose first that $\mu(\infty, f) = \mu(\infty, g) = 0$.

Define, for any $n \in \mathbb{Z}$,

$$F_n = \{t : (1+\varepsilon)^n < f(t) \leq (1+\varepsilon)^{n+1}\}, \quad G_n = \{t : (1+\varepsilon)^n < g \leq (1+\varepsilon)^{n+1}\}.$$

By assumption, $m(F_n) = m(G_n)$ for every $n \in \mathbb{Z}$. Let $\omega_n : G_n \to F_n$ be an arbitrary measure preserving bijection.

Define the measure preserving map $\omega : \mathrm{supp}(g) \to \mathrm{supp}(f)$ by concatenating $\omega_n : G_n \to F_n,\ n \in \mathbb{Z}$. For every $t \in G_n$, we have

$$f(\omega(t)) \geq (1+\varepsilon)^n, \quad g \leq (1+\varepsilon)^{n+1}.$$

Thus,

$$(1+\varepsilon) f(\omega(t)) \geq g, \quad t \in \mathrm{supp}(g).$$

This completes the proof of Case 1.

**Case 2**. Let $\delta$ such that $(1+\delta)^2 = (1+\varepsilon)$. Let $a = \mu(\infty, f) = \mu(\infty, g) > 0$. Define, for any $n \geq 1$,

$$F_n = \{t : a(1+\delta)^n < f(t) \leq a(1+\delta)^{n+1}\}, \quad G_n = \{t : a(1+\delta)^n < g \leq a(1+\delta)^{n+1}\}$$

and

$$F_0 = \{t : (1+\delta)^{-1} a \leq f(t) \leq (1+\delta)\, a\}, \quad G_0 = \{t : 0 < g \leq (1+\delta)\, a\}.$$

By assumption, for any $n \geq 1,\ m(G_n) = m(F_n)$ and $m(G_0) = m(F_0) = \infty$. For any $n \geq 0$, choose a measure preserving bijection $\omega_n$ from $G_n$ to $F_n$.

Define the measure preserving map $\omega : \mathrm{supp}(g) \to \mathrm{supp}(f)$ by concatenating the $\omega_n$'s. For any $n \geq 0$ and any $t \in G_n$,

$$f(\omega(t)) \geq a(1+\delta)^{n-1}, \quad g \leq a(1+\delta)^{n+1}.$$

Thus,

$$(1+\delta)^2 f(\omega(t)) = (1+\varepsilon) f(\omega(t)) \geq g, \quad t \in \mathrm{supp}(g). \qquad \square$$

**Lemma 2.9.** *Assume that* $\Omega$ *is* $(0, 1)$, $(0, \infty)$ *or* $\mathbb{N}$. *Let* $E, A, B \subset (L_0 + L_\infty)(\Omega)$ *be quasi-Banach function spaces. Assume that* $A$ *and* $B$ *are symmetric and that* $E$ *is an interpolation space for the couple* $(A, B)$. *Then* $E$ *admits an equivalent symmetric quasinorm.*

*Proof.* Let $f \in E$ and $g \in L_0 + L_\infty$. Assume that $\mu(g) \leq \mu(f)$. By Lemma 2.8, there exists a map $\omega : \mathrm{supp}(g) \to \mathrm{supp}(f)$ such that for any $t \in \mathrm{supp}(g)$,

$$2|f \circ \omega(t)| \geq |g(t)|.$$

Define, for any $h \in \mathcal{X}(\Omega)$,

$$T(h) := \begin{cases} \frac{g}{f \circ \omega} h \circ \omega & \text{on } \mathrm{supp}(g), \\ 0, & \text{elsewhere.} \end{cases}$$

Since $\omega$ is measure preserving, $T$ is bounded on $A$ and $B$ of norm less than 2. Let $c_E$ be the interpolation constant of $E$ for the couple $(A, B)$ (as in Proposition 2.5). We know that $Tf = g \in E$ and

$$\|g\| \leq 2c_E\|f\|. \tag{2-1}$$

Define, for any $f \in E$,

$$\|f\|_{E'} = \inf_{\mu(g)\geq\mu(f)} \|g\|_E.$$

By (2-1), $\|f\|_{E'} \leq \|f\|_E \leq 2c_E\|f\|_{E'}$ and $(E, \|\cdot\|_{E'})$ is a symmetric space. □

**Remark 2.10.** It is not difficult to see that if the underlying measure space $\Omega$ contains both a continuous part and atoms, then Lemma 2.9 is no longer true for $A = L_p(\Omega)$, $B = L_q(\Omega)$ and $p < 1$. However, one can observe that if $A$ and $B$ are *fully symmetric* (i.e., interpolation spaces between $L_1(\Omega)$ and $L_\infty(\Omega)$), Lemma 2.9 remains valid for any $\Omega$. This is reminiscent of the conditions required in [27, Section 4].

## 3. Interpolation for the couple $(L_0, L_q)$

In this section, $\Omega = (0, \infty)$ (for brevity, we omit $\Omega$ in the notations). We investigate some basic properties of the interpolation couple $(L_0, L_q)$. First, we provide a statement analogous to Proposition 2.5 and applicable to $L_0$.

Since the closed graph theorem does not apply to $L_0$ (it is not an $F$-space), our proof uses concrete constructions that rely on the structure of the underlying measure space.

For any $f \in \mathcal{X}$, denote by $M_f$ the multiplication operator $g \mapsto f \cdot g$.

**Theorem 3.1.** *Let $E$ be a quasi-Banach function space and $q \in (0, \infty]$. Assume that $E$ is an interpolation space for the couple $(L_0, L_q)$. Then, there exists a constant $c$ such that for any contraction $T$ on $(L_0, L_q)$, $\|T\|_{E\to E} \leq c$.*

*Proof.* Let $(A_n)_{n\geq1}$ be a partition of $(0, \infty)$ such that $m(A_n) = \infty$ for every $n \geq 1$.

Let $\gamma_n : A_n \to A_n^c$ be a measure preserving bijective transform. Set

$$(U_n x)(t) = \begin{cases} x(\gamma_n(t)), & t \in A_n, \\ 0, & t \in A_n^c, \end{cases} \qquad (V_n x)(t) = \begin{cases} x(\gamma_n^{-1}(t)), & t \in A_n^c, \\ 0, & t \in A_n. \end{cases}$$

Obviously, $U_n$ and $V_n$ are bounded operators on the couple $(L_0, L_q)$. By assumption, $U_n, V_n : E \to E$ are bounded mappings.

Note that

$$V_n U_n = M_{\chi_{A_n^c}}, \quad n \geq 1.$$

Let us argue by contradiction. For any $n \geq 1$, choose an operator $T_n$ which is a contraction on $(L_0, L_q)$ and such that

$$\|T_n\|_{E\to E} \geq 4^n \cdot \max\{\|U_n\|_{E\to E}, \|V_n\|_{E\to E}, 1\}^2. \tag{3-1}$$

It is immediate that

$$\begin{aligned} T_n &= M_{\chi_{A_n}} T_n M_{\chi_{A_n}} + M_{\chi_{A_n^c}} T_n M_{\chi_{A_n}} + M_{\chi_{A_n}} T_n M_{\chi_{A_n^c}} + M_{\chi_{A_n^c}} T_n M_{\chi_{A_n^c}} \\ &= T_{1,n} + V_n T_{2,n} + T_{3,n} U_n + V_n T_{4,n} U_n, \end{aligned}$$

where

$$T_{1,n} = M_{\chi_{A_n}} T_n M_{\chi_{A_n}}, \quad T_{2,n} = U_n T_n M_{\chi_{A_n}}, \quad T_{3,n} = M_{\chi_{A_n}} T_n V_n, \quad T_{4,n} = U_n T_n V_n.$$

By quasitriangle inequality, we have

$$\|T_n\|_{E\to E} \leq C_E^2 \cdot \left(\sum_{k=1}^{4} \|T_{k,n}\|_{E\to E}\right) \cdot \max\{\|U_n\|_{E\to E}, \|V_n\|_{E\to E}.1\}^2.$$

Let $k_n \in \{1, 2, 3, 4\}$ be such that

$$\|T_{k_n,n}\|_{E\to E} = \max_{1\leq k\leq 4} \|T_{k,n}\|_{E\to E}.$$

We, therefore, have

$$\|T_n\|_{E\to E} \leq 4C_E^2 \|T_{k_n,n}\|_{E\to E} \cdot \max\{\|U_n\|_{E\to E}, \|V_n\|_{E\to E}.1\}^2. \tag{3-2}$$

Set $S_n = T_{k_n,n}$. Note that $\|S_n\|_{L_0\to L_0} \leq 1$ and $\|S_n\|_{L_q\to L_q} \leq 1$. A combination of (3-1) and (3-2) yields

$$\|S_n\|_{E\to E} \geq 4^{n-1} C_E^{-2}, \quad n \geq 1.$$

Note that $S_n = M_{\chi_{A_n}} S_n M_{\chi_{A_n}}$. Set

$$S = \sum_{n\geq 1} S_n.$$

Since the $S_n$'s are in direct sum, we have

$$\|S\|_{L_0\to L_0} = \sup_{n\geq 1} \|S_n\|_{L_0\to L_0} \leq 1 \quad \text{and} \quad \|S\|_{L_q\to L_q} = \sup_{n\geq 1} \|S_n\|_{L_q\to L_q} \leq 1.$$

Moreover, $E$ is an interpolation space for the couple $(L_0, L_q)$, it follows that $S : E \to E$ is bounded.

For any $n \geq 1$, choose $f_n \in E$ such that $\|f_n\|_E \leq 1$ and $\|S_n f_n\|_E \geq 4^{n-2} C_E^{-2}$. Recall that $S_n = S_n M_{\chi_{A_n}}$. Hence, we may assume without loss of generality that $f_n$ is supported on $A_n$. Thus, $S(f_n) = S_n(f_n)$ and

$$\|S(f_n)\|_E = \|S_n(f_n)\|_E \geq 4^{n-2} C_E^{-2}.$$

This contradicts the boundedness of $S$. □

**Remark 3.2.** Theorem 3.1 above remains true for other underlying measure spaces:

- *For sequence spaces*. Indeed, in the proof of Theorem 3.1, we only use properties of the underlying measure space in the first sentence, namely when we consider a partition of $(0, \infty)$ into countably many sets, each of them isomorphic to $(0, \infty)$. Since a partition satisfying the same property exists for $\mathbb{Z}_+$, Theorem 3.1 remains true for interpolation spaces between $\ell_0$ and $\ell_q$.
- *For* $(0, 1)$. The same general idea applies in this case but some modifications have to be made because the maps $\gamma_n$ introduced in the proof cannot be assumed to be measure-preserving. The details are left to the reader.

**Lemma 3.3.** *Let $E, Y \subset L_0 + L_\infty$ be quasi-Banach function spaces. Assume that $Y$ is symmetric and that $E$ is an interpolation space for the couple $(L_0, Y)$. Then $E$ admits an equivalent symmetric quasinorm.*

*Proof.* The argument follows that in Lemma 2.9 mutatis mutandi. □

The following assertion is a special case of Theorem 1.1 and an important ingredient in the proof of the latter theorem.

**Corollary 3.4.** *Let $X$ be a quasi-Banach function space and $q \in (0, \infty)$. Assume that $L_0 \cap L_q \subset X \subset L_0 + L_q$ and that for any $f \in E$ and $g \in L_0 + L_q$,*

$$|g|^q \prec\prec_{\mathrm{tl}} |f|^q \Rightarrow g \in E.$$

*Then $E$ is an interpolation space for the couple $(L_0, L_q)$.*

*Proof.* It is clear that the condition on $X$ is equivalent to $E$-monotonicity with respect to the couple $(L_0, L_q)$ so by Remark 2.7, $E$ is an interpolation space for the couple $(L_0, L_q)$. □

Corollary 3.4 applies in particular to $L_p$-spaces, $p \leq q$. We decided to add a more precise statement and to provide a direct proof of the latter.

**Corollary 3.5.** *Let $p, q \in (0, \infty)$ such that $p < q$. Then, $L_p$ is an interpolation space for the couple $(L_0, L_q)$. More precisely, if $T$ is a contraction on $(L_0, L_q)$, then $T$ is a contraction on $L_p$.*

*Proof.* Let us first consider characteristic functions. Let $E$ be a set with finite measure. Since $T$ is a contraction on $L_0$, the measure of the support of $T(\chi_E)$ is less than $m(E)$. So by Hölder's inequality, setting $r = (p^{-1} - q^{-1})^{-1}$, we have

$$\|T(\chi_E)\|_p \leq \|T(\chi_E)\|_q \cdot m(E)^{1/r} \leq \|\chi_E\|_q \cdot m(E)^{1/r} = \|\chi_E\|_p.$$

First, consider the case $p \leq 1$. Let $f \in L_p$ be a step function, i.e.,

$$f = \sum_{i \in \mathbb{N}} a_i \chi_{E_i},$$

where $a_i \in \mathbb{C}$ and the sets $E_i$ are disjoint sets with finite measure. By the $p$-triangular inequality we write

$$\|Tf\|_p^p \le \sum_{i\in\mathbb{N}} |a_i|^p \, \|T(\chi_{E_i})\|_p^p \le \sum_{i\in\mathbb{N}} |a_i|^p \, \|\chi_{E_i}\|_p^p = \|f\|_p^p.$$

Since $T : L_p \to L_p$ is bounded by Corollary 3.4 and since step functions are dense in $L_p$, it follows that $T : L_p \to L_p$ is a contraction (for $p \le 1$).

Now consider the case $p > 1$. Since $p < q$, it follows that $q > 1$. By the preceding paragraph, $T : L_1 \to L_1$ is a contraction. By complex interpolation, $T : L_p \to L_p$ is also contraction. □

## 4. Construction of contractions on $(L_0, L_q)$ and $(L_p, L_\infty)$

Let $p, q \in (0, \infty)$. In this section, we are interested in the following question. Given functions $f$ and $g$ in $L_0 + L_q$ (resp. $L_p + L_\infty$), does there exist a bicontraction $T$ on $(L_0, L_q)$ (resp. $(L_p, L_\infty)$) such that $T(f) = g$? We show that such an operator exists provided that $|g|^q \prec\prec_{\mathrm{tl}} |f|^q$ (resp. $|g|^p \prec\prec_{\mathrm{hd}} |f|^p$. This directly implies a necessary condition for a symmetric space to be an interpolation space for the couple $(L_0, L_q)$ (resp. $(L_p, L_\infty)$) which will be exploited in the next section.

Our method of proof is very direct. We construct the bicontraction $T$ as direct sums of very simple operators. This is made possible by two partition lemmas that enable us to understand the orders $\prec\prec_{\mathrm{tl}}$ and $\prec\prec_{\mathrm{hd}}$ as direct sums of simple situations.

***Partition lemmas.*** We state our first lemma without proof since it essentially repeats that of Proposition 19 in [28].

**Lemma 4.1.** *Let $f, g \in L_1$ be positive decreasing step functions. Assume that $g \prec_{\mathrm{hd}} f$. There exists a sequence of intervals $\{I_k, J_k\}_{k\ge0}$ of $(0, \infty)$ such that:*

(i) *For every $k \ge 0$, $I_k$ and $J_k$ are disjoint intervals of finite length.*

(ii) *$(I_k \cup J_k) \cap (I_l \cup J_l) = \varnothing$ for $k \ne l$.*

(iii) *$f$ and $g$ are constant on $I_k$ and on $J_k$.*

(iv) *$g|_{I_k\cup J_k} \prec_{\mathrm{hd}} f|_{I_k\cup J_k}$ for every $k \ge 0$.*

(v) *$g \le f$ on the complement of $\bigcup_{k\ge0} I_k \cup J_k$.*

*If furthermore $f, g \in L_1$ and $g \prec_{\mathrm{hd}} f$ then $f = g$ on the complement of $\bigcup_{k\ge0} I_k \cup J_k$.*

**Scholium 4.2.** *Let $f, g \in \mathcal{X}$ be positive decreasing functions. Let $\Delta \subset (0, \infty)$ be an arbitrary measurable set.*

(i) *If $f, g \in L_1 + L_\infty$ are such that*

$$\int_{[0,t]\cap\Delta} g \le \int_{[0,t]\cap\Delta} f, \quad t > 0,$$

*then $g\chi_\Delta \prec\prec_{\mathrm{hd}} f\chi_\Delta$.*

(ii) *If $f, g \in L_0 + L_1$ are such that*

$$\int_{(t,\infty)\cap\Delta} g \le \int_{(t,\infty)\cap\Delta} f, \quad t > 0,$$

*then $g\chi_\Delta \prec\prec_{\mathrm{tl}} f\chi_\Delta$.*

The second partition lemma deals with describing the order $\prec\prec_{\mathrm{tl}}$ in terms of $\prec_{\mathrm{tl}}$ and $\le$.

**Lemma 4.3.** *Let $f, g \in L_0 + L_1$ be such that $f = \mu(f)$, $g = \mu(g)$ and $g \prec\prec_{\mathrm{tl}} f$. There exists a collection $\{\Delta_k\}_{k\ge 0}$ of pairwise disjoint sets such that:*

(i) *$f|_{\Delta_k} \prec_{\mathrm{hd}} g|_{\Delta_k}$ for every $k \ge 0$.*

(ii) *$g \le f$ on the complement of $\bigcup_{k\ge 0} \Delta_k$.*

*Proof.* Consider the set $\{g > f\}$. Similarly to the previous proof, connected components of the set $\{g > f\}$ are intervals (closed or not) not reduced to points. Let us enumerate these intervals as $(a_k, b_k)$, $k \ge 0$.

We have

$$\int_t^\infty (f-g)_+ - \int_t^\infty (f-g)_- = \int_t^\infty (f-g) \ge 0.$$

Let

$$H(t) = \sup\biggl\{u : \int_u^\infty (f-g)_+ = \int_t^\infty (f-g)_-\biggr\}.$$

Obviously, $H$ is a monotone function, $H(t) \ge t$ for all $t > 0$ and

$$\int_{H(t)}^\infty (f-g)_+ = \int_t^\infty (f-g)_-.$$

Set

$$\Delta_k = (a_k, b_k) \cup \bigl((H(a_k), H(b_k)) \cap \{g \le f\}\bigr).$$

Note that

$$\int_{b_k}^\infty (f-g)_+ = \int_{a_k}^\infty (f-g)_+ \ge \int_{a_k}^\infty (f-g)_-$$

and therefore, $H(a_k) \ge b_k$.

We claim that $\Delta_k \cap \Delta_l = \varnothing$ for $k \ne l$. Indeed, let $a_k < b_k \le a_l < b_l$. We have $H(a_k) \le H(b_k) \le H(a_l) \le H(b_l)$. Thus, $(H(a_k), H(b_k)) \cap (H(a_l), H(b_l)) = \varnothing$. We now have

$$\Delta_k \cap \Delta_l = \bigl((\Delta_k \cap \{f < g\}) \cap (\Delta_l \cap \{f < g\})\bigr) \cup \bigl((\Delta_k \cap \{f \ge g\}) \cap (\Delta_l \cap \{f \ge g\})\bigr).$$

Obviously,

$$(\Delta_k \cap \{f < g\}) \cap (\Delta_l \cap \{f < g\}) = (a_k, b_k) \cap (a_l, b_l) = \varnothing,$$

$$(\Delta_k \cap \{f \ge g\}) \cap (\Delta_l \cap \{f \ge g\}) = (H(a_k), H(b_k)) \cap (H(a_l), H(b_l)) \cap \{f \ge g\} = \varnothing.$$

This proves the claim.

We now claim that

$$\int_{(t,\infty)\cap\Delta_k}(f-g)\geq 0.$$

If $t\geq b_k$, then taking into account that $H(a_k)\geq b_k$, we infer that

$$(t,\infty)\cap\Delta_k\subset\{f\geq g\}$$

and the claim follows immediately. If $t\in(a_k,b_k)$, then

$$\begin{aligned}\int_{(t,\infty)\cap\Delta_k}(f-g)&=\int_{(H(a_k),H(b_k))}(f-g)_+-\int_t^{b_k}(f-g)_-\\&\geq\int_{(H(a_k),H(b_k))}(f-g)_+-\int_{a_k}^{b_k}(f-g)_-=0.\end{aligned}$$

This proves the claim.

It follows from the claim and Scholium 4.2 that $g\chi_{\Delta_k}\prec\prec_{\mathrm{tl}}f\chi_{\Delta_k}$. Since

$$\int_{\Delta_k}g=\int_{\Delta_k}f,$$

it follows that $g\chi_{\Delta_k}\prec_{\mathrm{tl}}f\chi_{\Delta_k}$, which immediately implies the first assertion.

By construction, $(a_k,b_k)\subset\Delta_k$. Thus,

$$\{g>f\}=\bigcup_{k\geq 0}(a_k,b_k)\subset\bigcup_{k\geq 0}\Delta_k.$$

The second assertion is now obvious. $\square$

***Construction of operators.*** We repeat the same structure as in the previous subsection, proving four lemmas, each one dealing with a certain order: $\prec_{\mathrm{hd}}$, $\prec_{\mathrm{tl}}$, $\prec\prec_{\mathrm{hd}}$, and finally $\prec\prec_{\mathrm{tl}}$.

**Lemma 4.4.** *Let $p\in(0,\infty)$. Let $f,g\in L_p(0,\infty)$, assume that $|g|^p\prec_{\mathrm{hd}}|f|^p$, $f=\mu(f)$ and $g=\mu(g)$. There exists a linear operator $T:\mathcal{X}(0,\infty)\to\mathcal{X}(0,\infty)$ such that $g=T(f)$ and*

$$\|T\|_{L_p\to L_p}\leq 2\cdot 3^{1/p},\quad \|T\|_{L_\infty\to L_\infty}\leq 2\cdot 2^{1/p}.$$

*Proof.* **Step 1**. First, let us assume that $f$ and $g$ are step functions.

Apply Lemma 4.1 to the functions $f^p$ and $g^p$ and let $I_k$ and $J_k$ be as in Lemma 4.1. Without loss of generality, the interval $I_k$ is located to the left of the interval $J_k$.

For every $k\geq 0$, let's define the mapping $S_k:\mathcal{X}(I_k\cup J_k)\to\mathcal{X}(I_k\cup J_k)$ as below. The construction of this mapping will depend on whether $f^p|_{J_k}\leq\frac{1}{2}g^p|_{J_k}$ or $f^p|_{J_k}>\frac{1}{2}g^p|_{J_k}$.

If $f^p|_{J_k} \le \frac{1}{2} g^p|_{J_k}$, then

$$\begin{aligned} g^p|_{J_k} \cdot m(J_k) &\le g^p|_{I_k} \cdot m(I_k) + g^p|_{J_k} \cdot m(J_k) \\ &= f^p|_{I_k} \cdot m(I_k) + f^p|_{J_k} \cdot m(J_k) \le f^p|_{I_k} \cdot m(I_k) + \tfrac{1}{2} g^p|_{J_k} \cdot m(J_k). \end{aligned}$$

Therefore,

$$g^p|_{J_k} \cdot m(J_k) \le 2 f^p|_{I_k} \cdot m(I_k).$$

Let $l_k$ be a linear bijection from $J_k$ to $I_k$. We set

$$S_k\, x = \frac{g|_{I_k}}{f|_{I_k}} \cdot x\chi_{I_k} + \frac{g|_{J_k}}{f|_{I_k}} \cdot (x \circ l_k)\, \chi_{J_k}.$$

Clearly, $S_k$ is a contraction in the uniform norm.

Let $x \in L_p$. We have

$$\begin{aligned} \|S_k\, x\|_p^p &\le \frac{g^p|_{I_k}}{f^p|_{I_k}} \cdot \|x\chi_{I_k}\|_p^p + \frac{g^p|_{J_k}}{f^p|_{I_k}} \cdot \|(x \circ l_k)\, \chi_{J_k}\|_p^p \\ &\le \frac{g^p|_{I_k}}{f^p|_{I_k}} \cdot \|x\|_p^p + \frac{g^p|_{J_k}}{f^p|_{I_k}} \cdot \frac{m(J_k)}{m(I_k)} \cdot \|x\|_p^p \le 3\|x\|_p^p. \end{aligned}$$

Also, we have

$$\|S_k\, x\|_\infty \le \|x\|_\infty.$$

If $f^p|_{J_k} > \frac{1}{2} g^p|_{J_k}$, then we set $S_k = M_{gf^{-1}}$. Clearly, $\|S_k\, x\|_\infty \le 2^{1/p}\|x\|_\infty$ and $\|S_k\, x\|_p \le 2^{1/p}\|x\|_p$.

We define $S : \mathcal{X} \to \mathcal{X}$ by

$$S = \bigoplus_{k \ge 0} S_k.$$

Remark that $\|S\|_{r\to r} = \sup_{k\ge 0}\|S_k\|_{L_r \to L_r}$ for any $r \in [0, \infty]$. Hence,

$$\|S\|_{L_p \to L_p} \le 3^{1/p}, \quad \|S\|_{L_\infty \to L_\infty} \le 2^{1/p}.$$

It remains only to note that $Sf = g$.

**Step 2**. Now, only assume that $f$ and $g$ are positive and nonincreasing. Define, for any $n \in \mathbb{Z}$,

$$a_n = \sup\{t \in (0, \infty) : f(t) \ge 2^{n/2}\} \quad \text{and} \quad b_n = \sup\{t \in (0, \infty) : g(t) \ge 2^{n/2}\}.$$

Let $\mathcal{A}$ be the $\sigma$-algebra generated by the intervals $(a_n, a_{n+1})$ and $(b_n, b_{n+1})$. Define

$$f_0^p = \mathbb{E}[f^p \mid \mathcal{A}] \quad \text{and} \quad g_0^p = \mathbb{E}[g^p \mid \mathcal{A}].$$

Note that $f_0$ and $g_0$ are step functions such that

$$g_0^p \prec_{\mathrm{hd}} f_0^p, \quad 2^{-1/2} f \le f_0 \le 2^{1/2} f \quad \text{and} \quad 2^{-1/2} g \le g_0 \le 2^{1/2} g.$$

Apply Step 1 to $f_0$ and $g_0$ to obtain an operator $S$ and set

$$T = M_{ff_0^{-1}} \circ S \circ M_{g_0 g^{-1}}.$$

Clearly, $Tf = g$ and

$$\|T\|_{L_p \to L_p} \leq 2\|S\|_{L_p \to L_p} \leq 2 \cdot 2^{1/p}, \quad \|T\|_{L_\infty \to L_\infty} \leq 2\|S\|_{L_\infty \to L_\infty} \leq 2^2 \cdot 2^{1/p}. \ \square$$

**Lemma 4.5.** *Let $f, g \in L_q(0,\infty)$ be positive nonincreasing functions such that $g^q \prec_{\mathrm{tl}} f^q$. Let $d > 1$. There exists a linear operator $T : \mathcal{X}(0,\infty) \to \mathcal{X}(0,\infty)$ such that $g = T(f)$ and*

$$\|T\|_{L_0 \to L_0} \leq 4, \quad \|T\|_{L_q \to L_q} \leq 2 \cdot 3^{1/q}.$$

*Proof.* Following Step 2 of Lemma 4.4, we are reduced to dealing with step functions.

Apply Lemma 4.1 to the functions $g^q$ and $f^q$ and let $(I_k)_{k\geq 1}$ and $(J_k)_{k\geq 1}$ be as in Lemma 4.1. Without loss of generality, the intervals $I_k$ is located to the left of the intervals $J_k$.

Let $k \geq 1$. Define the mappings $S_k : \mathcal{X}(I_k \cup J_k) \to \mathcal{X}(I_k \cup J_k)$ as below.

Note that since $f^q|_{I_k \cup J_k} \prec_{\mathrm{hd}} g^q|_{I_k \cup J_k}$, we have

$$g|_{I_k} \geq f|_{I_k} \geq f|_{J_k} \geq g|_{J_k}.$$

The construction of $S_k$ will depend on whether we have $\|f\chi_{I_k}\|_q \leq \|f\chi_{J_k}\|_q$ or $\|f\chi_{I_k}\|_q > \|f\chi_{J_k}\|_q$.

If $\|f\chi_{I_k}\|_q \leq \|f\chi_{J_k}\|_q$, then $m(I_k) \leq m(J_k)$ and

$$g_0^q|_{I_k} \cdot m(I_k) \leq 2 f^q|_{J_k} \cdot m(J_k).$$

Let $l_k : I_k \to J_k$ be a linear bijection. We set

$$S_k\, x = \frac{g|_{I_k}}{f|_{J_k}} (x \circ l_k)\, \chi_{I_k} + \frac{g}{f} x \chi_{J_k}.$$

Let $x \in L_q$. We have

$$\|S_k\, x\|_0 \leq \|x\chi_{J_k}\|_0 + \|(x \circ l_k)\, \chi_{I_k}\|_0 \leq \left(1 + \frac{m(I_k)}{m(J_k)}\right)\|x\|_0.$$

Thus, $\|S_k\|_{L_0 \to L_0} \leq 2$. We have

$$\begin{aligned}\|S_k\, x\|_q^q &= \frac{g^q|_{I_k}}{f^q|_{J_k}} \|(x \circ l_k)\, \chi_{I_k}\|_q^q + \frac{g^q|_{J_k}}{f^q|_{J_k}} \|x\chi_{J_k}\|_q^q \\ &\leq \frac{g^q|_{I_k}}{f^q|_{J_k}} \cdot \frac{m(I_k)}{m(J_k)} \cdot \|x\|_q^q + \frac{g^q|_{J_k}}{f^q|_{J_k}} \|x\|_q^q \leq 3\|x\|_q^q.\end{aligned}$$

If $\|f\chi_{I_k}\|_q > \|f\chi_{J_k}\|_q$, then

$$g^q|_{I_k} \cdot m(I_k) \leq 2 f^q|_{I_k} \cdot m(I_k)$$

and therefore, $g_0 \le 2^{1/q} f_0$. We now set $S_k = M_{gf^{-1}}$. Obviously, $\|S_k\|_{L_0 \to L_0} \le 1$ and $\|S_k\|_{L_q \to L_q} \le 2^{1/q}$.

We set

$$S = \bigoplus_{k \ge 0} S_k.$$

Since $S : \mathcal{X}(0, \infty) \to \mathcal{X}(0, \infty)$ is defined as a direct sum:

$$\|S\|_{L_0 \to L_0} = \sup_{k \ge 1} \|S_k\|_{L_0 \to L_0} \le 2, \quad \|S\|_{L_q \to L_q} = \sup_{k \ge 1} \|S_k\|_{L_q \to L_q} \le 3^{1/q},$$

it remains only to note that $Sf = g$. □

**Lemma 4.6.** *Let $f, g \in (L_p + L_\infty)(0, \infty)$ be such that $|g|^p \prec\prec_{\mathrm{hd}} |f|^p$, $f = \mu(f)$ and $g = \mu(g)$. There exists a linear operator $T : \mathcal{X} \to \mathcal{X}$ such that $g = T(f)$ and*

$$\|T\|_{L_p \to L_p} \le 2 \cdot 3^{1/p}, \quad \|T\|_{L_\infty \to L_\infty} \le 2 \cdot 2^{1/p}.$$

*Proof.* Let $(\Delta_k)_{k \ge 0}$ be as in Lemma 4.1 and let $\Delta_\infty$ be the complement of $\bigcup_{k \ge 0} \Delta_k$. By Lemma 4.4, there exists $T_k : \mathcal{X}(\Delta_k) \to \mathcal{X}(\Delta_k)$ such that $T_k(f) = g$ on $\Delta_k$ and

$$\|T_k\|_{L_p(\mathcal{X}_k) \to L_p(\mathcal{X}_k)} \le 2 \cdot 9^{1/p}, \quad \|T_k\|_{L_\infty(\mathcal{X}_k) \to L_\infty(\mathcal{X}_k)} \le 2 \cdot 4^{1/p}.$$

Set $T_\infty = M_{g/f}$ on $\mathcal{X}(\Delta_\infty)$. We now set

$$T = T_\infty \bigoplus \Big( \bigoplus_{k \ge 0} T_k \Big).$$

Obviously, $Tf = g$ on $(0, \infty)$ and

$$\|T\|_{L_p \to L_p} \le 2 \cdot 9^{1/p}, \quad \|T\|_{L_\infty \to L_\infty} \le 2 \cdot 4^{1/p}.$$ □

**Lemma 4.7.** *Let $f, g \in (L_0 + L_q)(0, \infty)$ be such that $|g|^q \prec\prec_{\mathrm{tl}} |f|^q$, $f = \mu(f)$ and $g = \mu(g)$. There exists a linear operator $T : \mathcal{X} \to \mathcal{X}$ such that $g = T(f)$ and*

$$\|T\|_{L_0 \to L_0} \le 4, \quad \|T\|_{L_q \to L_q} \le 2 \cdot 3^{1/q}.$$

*Proof.* Without loss of generality, $g = \mu(g)$ and $f = \mu(f)$. Let $(\Delta_k)_{k \ge 0}$ be as in Lemma 4.3 and let $\Delta_\infty$ be the complement of $\bigcup_{k \ge 0} \Delta_k$. By Lemma 4.5, there exists $T_k : \mathcal{X}(\Delta_k) \to \mathcal{X}(\Delta_k)$ such that $T_k(f) = g$ on $\Delta_k$ and

$$\|T_k\|_{L_0(\mathcal{X}_k) \to L_0(\mathcal{X}_k)} \le 8, \quad \|T_k\|_{L_q(\mathcal{X}_k) \to L_q(\mathcal{X}_k)} \le 2 \cdot 9^{1/q}.$$

Set $T_\infty = M_{g/f}$ on $\mathcal{X}(\Delta_\infty)$. We now set

$$T = T_\infty \bigoplus \Big( \bigoplus_{k \ge 0} T_k \Big).$$

Obviously, $Tf = g$ on $(0, \infty)$ and

$$\|T\|_{L_0 \to L_0} \le 4, \quad \|T\|_{L_q \to L_q} \le 2 \cdot 3^{1/q}.$$ □

## 5. Interpolation spaces for the couple $(L_p, L_q)$

In this section, we obtain characterizations of interpolation spaces for the couple $(L_p, L_q)$, in terms of the majorization notions studied earlier. The necessity of the condition we consider is a direct consequence of the constructions explained in the previous section.

**Theorem 5.1.** *Let $0 \le p < q \le \infty$. Let $E$ be a quasi-Banach function space such that $E \in \mathrm{Int}(L_p, L_q)$. There exist $c_{p,E}$ and $c_{q,E}$ in $\mathbb{R}_{>0}$ such that*:

(i) *Suppose $p \neq 0$. For any $f \in E$ and $g \in L_p + L_\infty$ if $|g|^p \prec\prec_{\mathrm{hd}} |f|^p$, then $g \in E$ and $\|g\|_E \le c_{p,E}\|f\|_E$.*

(ii) *Suppose $q \neq \infty$. For any $f \in E$ and $g \in L_0 + L_q$ if $|g|^q \prec\prec_{\mathrm{tl}} |f|^q$, then $g \in E$ and $\|g\|_E \le c_{q,E}\|f\|_E$.*

*Proof.* By Lemma 2.9 (for $p > 0$) or Lemma 3.3 (for $p = 0$), we may assume without loss of generality that $E$ is a symmetric function space.

Assume that $p \neq 0$. Let $f \in E$ and let $g \in L_p + L_\infty$ be such that $|g|^p \prec\prec_{\mathrm{hd}} |f|^p$. Since $E$ is symmetric, we may assume without loss of generality that $f = \mu(f)$ and $g = \mu(g)$. By Lemma 4.6, there exists an operator $T$ such that $T(f) = g$ and

$$\|T\|_{(L_p, L_\infty) \to (L_p, L_\infty)} \le 2 \cdot 3^{1/p}.$$

Recall that $L_q$ is an interpolation space for the couple $(L_p, L_\infty)$ (one can take, for example, real or complex interpolation method). Let $c_{p,q}$ be the interpolation constant of $L_q$ for the couple $(L_p, L_\infty)$. We have

$$\|T\|_{L_q \to L_q} \le c_{p,q} \cdot 2 \cdot 3^{1/p}.$$

Let $c_E$ be the interpolation constant of $E$ for $(L_p, L_q)$ (see Proposition 2.5). Then,

$$\|T\|_{E \to E} \le c_E \cdot \max\{1, c_{p,q}\} \cdot 2 \cdot 3^{1/p}.$$

Thus,

$$\|g\|_E \le \|T\|_{E \to E}\|f\|_E \le c_E \cdot \max\{1, c_{p,q}\} \cdot 2 \cdot 3^{1/p}\|f\|_E.$$

This proves the first assertion. The proof of the second one follows mutatis mutandi using Corollary 3.5 instead of complex interpolation and (for $p = 0$) Theorem 3.1 instead of Proposition 2.5. ☐

**Lemma 5.2.** *Assume that $0 < p < q < \infty$. Let $f, g \in L_p + L_q$ such that $f = \mu(f)$ and $g = \mu(g)$. Suppose that at every $t > 0$, one of the following inequalities holds*:

$$\int_0^t g^p \, ds \le \int_0^t f^p \, ds \quad \text{or} \quad \int_t^\infty g^q \, ds \le \int_t^\infty f^q \, ds.$$

*Then, there exist $g_1, g_2 \in (L_p + L_q)^+$ which satisfies $g = g_1 + g_2$, $g_1^p \prec\prec_{\mathrm{hd}} f^p$ and $g_2^q \prec\prec_{\mathrm{tl}} f^q$.*

*Proof.* Set

$$A=\left\{t>0:\int_0^t g(s)^p\,ds\le\int_0^t f(s)^p\,ds\right\},$$
$$B=\left\{t>0:\int_t^\infty g(s)^q\,ds\le\int_t^\infty f(s)^q\,ds\right\}.$$

Let

$$u_+(t)=\inf\{s\in A:s\ge t\},\quad u_-(t)=\sup\{s\in A:s\le t\},$$
$$v_+(t)=\inf\{s\in B:s\ge t\},\quad v_-(t)=\sup\{s\in B:s\le t\}.$$

Note that $f^p\chi_{(u_-(t),u_+(t))}\prec_{\mathrm{hd}} g^p\chi_{(u_-(t),u_+(t))}$ for $t\notin A$ and, therefore,

$$g(u_+(t)-0)\le f(u_+(t)-0).$$

Set $h_1(t)=g(u_+(t)-0),\ t>0$. By definition, $u_+(t)\ge t$ for all $t>0$. Since $g$ is decreasing, it follows that $h_1\le g$. Set $h_2=g\chi_B$. Since $u_+(t)=t$ for $t\in A$, it follows that $h_1=g$ on $A$. Thus, $h_1+h_2\ge g\chi_A+g\chi_B\ge g$.

We claim that

$$\int_0^t\mu(s,h_1)^p\,ds\le\int_0^t f(s)^p\,ds.$$

Indeed, for $t\in A$, we have

$$\int_0^t h_1(s)^p\,ds\le\int_0^t g(s)^p\,ds\le\int_0^t f(s)^p\,ds.$$

For $t\notin A$, we have $h_1(s)=g(u_+(t))$ for all $s\in(u_-(t),u_+(t))$. Thus,

$$\int_0^t h_1(s)^p\,ds=\int_0^{u_-(t)}h_1(s)^p\,ds+\int_{u_-(t)}^t h_1(s)^p\,ds$$
$$\le\int_0^{u_-(t)}g(s)^p\,ds+\int_{u_-(t)}^t g(u_+(t))^p\,ds.$$

Since

$$\int_0^{u_-(t)}g(s)^p\,ds=\int_0^{u_-(t)}f(s)^p\,ds,\quad g(u_+(t))\le f(u_+(t)),$$

it follows that

$$\int_0^t h_1(s)^p\,ds\le\int_0^{u_-(t)}f(s)^p\,ds+\int_{u_-(t)}^t f(u_+(t))^p\,ds\le\int_0^t f(s)^p\,ds.$$

Since $h_1=\mu(h_1)$, the claim follows.

We claim that

$$\int_t^\infty\mu(s,h_2)^q\,ds\le\int_t^\infty f(s)^q\,ds.$$

For $t \in B$, we have

$$\int_t^\infty h_2(s)^q\, ds \le \int_t^\infty g(s)^q\, ds \le \int_t^\infty f(s)^q\, ds.$$

For $t \notin B$, we have

$$\int_t^\infty h_2(s)^q\, ds = \int_{v_+(t)}^\infty h_2(s)^q\, ds \le \int_{v_+(t)}^\infty g(s)^q\, ds = \int_{v_+(t)}^\infty f(s)^q\, ds \le \int_t^\infty f(s)^q\, ds.$$

In either case, we have

$$\int_t^\infty \mu(s, h_2)^q\, ds \le \int_t^\infty h_2(s)^q\, ds \le \int_t^\infty f(s)^q\, ds.$$

This proves the claim.

Setting

$$g_1 = \frac{h_1}{h_1 + h_2} g, \quad g_2 = \frac{h_2}{h_1 + h_2} g,$$

we complete the proof. $\square$

**Theorem 5.3.** *Let $0 \le p < q \le \infty$ with either $p \ne 0$ or $q \ne \infty$. Let $E$ be a quasi-Banach function space. Assume that there exist $c_{p,E}$ and $c_{q,E}$ in $\mathbb{R}_{>0}$ such that*:

(i) *Suppose $p \ne 0$. For any $f \in E$ and $g \in L_0 + L_p$, if $|g|^p \prec\prec_{\mathrm{hd}} |f|^p$, then $g \in E$ and $\|g\|_E \le c_{p,E} \|f\|_E$.*

(ii) *Suppose $q \ne \infty$. For any $f \in E$ and $g \in L_0 + L_q$, if $|g|^q \prec\prec_{\mathrm{tl}} |f|^q$, then $g \in E$ and $\|g\|_E \le c_{q,E} \|f\|_E$.*

*Then $E$ belongs to* $\mathrm{Int}(L_p, L_q)$.

*Proof.* Assume that $p \ne 0$. Let us show that the first condition implies $E \subset L_p + L_\infty$. Indeed, assume the contrary and choose $f \in E$ such that $\mu(f)\, \chi_{(0,1)} \notin L_p$. Let

$$f_n = \min\left\{\mu\left(\frac{1}{n}, f\right), \mu(f)\, \chi_{(0,1)}\right\}, \quad n \ge 1.$$

Obviously, $\|f_n\|_E \le \|\mu(f)\, \chi_{(0,1)}\|_E \le \|f\|_E$. On contrary, $\|f_n\|_p^p \chi_{(0,1)} \prec\prec_{\mathrm{hd}} f_n^p$. By the first condition on $E$, we have $\|f_n\|_p \|\chi_{(0,1)}\|_E \le c_{p,E} \|f\|_E$. However, we have $\|f_n\|_p \uparrow \|\mu(f)\, \chi_{(0,1)}\|_p = \infty$. This contradiction shows that our initial assumption was incorrect. Thus, $E \subset L_p + L_\infty$.

A similar argument shows that the second condition implies $E \subset L_0 + L_q$. Thus, a combination of both conditions implies $E \subset L_p + L_q$.

Let $T$ be a contraction on $(L_p, L_q)$ and $f \in E$. To conclude the proof, it suffices to show that $Tf$ belongs to $E$. First, note that

$$K(t, Tf, L_p, L_q) \le K(t, f, L_p, L_q).$$

Assume that $p > 0$ and $q < \infty$. Let $\alpha^{-1} = \frac{1}{p} - \frac{1}{q}$. By the Holmstedt formula for the $K$-functional (see [14]), there exists a constant $c_{p,q} > 0$ such that for any $t \in \mathbb{R}_{>0}$,

$$\left(\int_0^{t^\alpha} \mu(s, Tf)^p \, ds\right)^{1/p} + t\left(\int_{t^\alpha}^\infty \mu(s, Tf)^q \, ds\right)^{1/q} \\ \le c_{p,q}\left(\left(\int_0^{t^\alpha} \mu(s, f)^p \, ds\right)^{1/p} + t\left(\int_{t^\alpha}^\infty \mu(s, f)^q \, ds\right)^{1/q}\right).$$

Hence, for any given $t > 0$, we have either

$$\int_0^{t^\alpha} \mu(s, Tf)^p \, ds \le \int_0^{t^\alpha} \mu(s, c_{p,q} f)^p \, ds$$

or

$$\int_{t^\alpha}^\infty \mu(s, Tf)^q \, ds \le \int_{t^\alpha}^\infty \mu(s, c_{p,q} f)^q \, ds.$$

By Lemma 5.2, one can write

$$\mu(Tf) = g_1 + g_2, \quad g_1^p \prec\prec_{\mathrm{hd}} (c_{p,q}\mu(f))^p, \quad g_2 \prec\prec_{\mathrm{tl}} (c_{p,q}\mu(f))^q. \tag{5-1}$$

By assumption, we have

$$\|g_1\|_E \le c_{p,E} \|f\|_E, \quad \|g_2\|_E \le c_{q,E} \|f\|_E.$$

By triangle inequality, we have

$$\|Tf\|_E \le c_{p,q,E} \|f\|_E.$$

Assume now that $p > 0$ and $q = \infty$. This case is simpler since by the Holmstedt formula (see [14]), there exists $c_p \in \mathbb{R}_{>0}$ such that for any $t \in \mathbb{R}_{>0}$,

$$\left(\int_0^{t^p} \mu(s, Tf)^p \, ds\right)^{1/p} \le c_p \left(\int_0^{t^p} \mu(s, f)^p \, ds\right)^{1/p}.$$

This means that $|Tf|^p \prec\prec_{\mathrm{hd}} |c_p f|^p$ so by assumption (1), $Tf$ belongs to $E$ and

$$\|Tf\|_E \le c_p \, c_{p,E} \|f\|_E.$$

The case of $p = 0$ and $q < \infty$ is given by Corollary 3.4. □

Theorem 1.1 claimed in the introduction compiles some results of this section.

*Proof of Theorem 1.1.* Assertion (a) is obtained by combining Theorems 5.1 and 5.3 with $q = \infty$.

Assertion (b) is derived similarly from Theorems 5.1 and 5.3 by applying them with $p = 0$.

Finally, using assertions (a) and (b) of Theorem 1.1, Theorems 5.1 and 5.3, for $0 < p < q < \infty$, one obtains Theorem 1.1(c). □

**Remark 5.4.** In the spirit of Corollary 3.4, we could have used a nonquantitative condition to deal with the case of $q = \infty$ in Theorem 5.3. Let $E$ be a quasi-Banach function space and $p, q \in (0, \infty)$. This means that the following two conditions are equivalent:

(i) For any $f \in E,\ g \in L_p + L_\infty$,

$$|g|^p \prec\prec_{\mathrm{hd}} |f|^p \Rightarrow g \in E.$$

(ii) There exists $c > 0$ such that for any $f \in E,\ g \in L_p + L_\infty$,

$$|g|^p \prec\prec_{\mathrm{hd}} |f|^p \Rightarrow g \in E \quad \text{and} \quad \|g\|_E \leq c\|f\|_E.$$

Similarly, the following two conditions are equivalent:

(i) For any $f \in E,\ g \in L_0 + L_q$,

$$|g|^q \prec\prec_{\mathrm{tl}} |f|^q \Rightarrow g \in E.$$

(ii) There exists $c > 0$ such that for any $f \in E,\ g \in L_0 + L_q$,

$$|g|^q \prec\prec_{\mathrm{tl}} |f|^q \Rightarrow g \in E \quad \text{and} \quad \|g\|_E \leq c\|f\|_E.$$

## 6. Interpolation spaces for couples of $\ell^p$-spaces

In this section, we show that our approach to the Lorentz–Shimogaki and Arazy–Cwikel theorems also applies to sequence spaces. We follow a structure similar to the previous sections, proving partition lemmas, then constructing bounded operators on couples $(\ell^p, \ell^q)$ with suitable properties to finally conclude on the interpolation spaces of the couple $(\ell^p, \ell^q)$. Additional arguments involving Boyd indices will be required to prove Theorem 1.2.

We identify sequences with bounded functions on $(0, \infty)$ which are almost constant on intervals of the form $(k, k+1),\ k \in \mathbb{Z}^+$ by

$$i : \ell^\infty \to L_\infty, \quad (u_k)_{k\in\mathbb{Z}^+} \mapsto \sum_{k=0}^{\infty} u_k\, \mathbf{1}_{(k,k+1)}.$$

***An interpolation theorem for the couple $(\ell^p, \ell^q)$.*** We start with a partition lemma playing, for sequence spaces, the role of Lemma 4.3.

**Lemma 6.1.** *Let $a = (a_n)_{n\in\mathbb{Z}^+},\ b = (b_n)_{n\in\mathbb{Z}^+}$ be two positive decreasing sequences such that $b \prec\prec_{\mathrm{tl}} a$. There exists a sequence $(\Delta_n)_{n\in\mathbb{Z}^+}$ of subsets of $\mathbb{Z}^+$ such that*:

(i) *For every $k \in \mathbb{Z}_+$, we have $|\{n \in \mathbb{Z}^+ : k \in \Delta_n\}| \leq 3$.*

(ii) $\sum_{k\in\Delta_n} a_k \geq b_n$ *for any $n \in \mathbb{Z}^+$.*

*Proof.* Define $I = \{n \in \mathbb{Z}^+ : b_n > a_n\}$. For $n \notin I$, set $\Delta_n = \{n\}$.

For any $n \in \mathbb{Z}^+$, define

$$i_n = \sup\left\{ i : \sum_{k=i}^{\infty} a_k \geq \sum_{k=n}^{\infty} b_k \right\}$$

and for $n \in I$, let $\Delta_n = \{i_n, \dots, i_{n+1}\}$.

From the definition of $i_n$, we have

$$\sum_{k \geq i_n} a_k \geq \sum_{k \geq n} b_k.$$

From the definition of $i_{n+1}$, we have

$$\sum_{k > i_{n+1}} a_k < \sum_{k \geq n+1} b_k.$$

Taking the difference of these inequalities, we infer that

$$\sum_{k \in \Delta_n} a_k \geq b_n.$$

This proves the second condition.

Note that since $b \prec\prec_{\mathrm{tl}} a$, we write $i_n \geq n$ for any $n \in \mathbb{Z}^+$. Hence, if $n \in I$, then $b_n > a_n \geq a_{i_{n+1}}$. Furthermore, by definition of $i_{n+1}$, we have

$$\sum_{k > i_{n+1}} a_k < \sum_{k=n+1}^{\infty} b_k, \quad \text{so} \sum_{k=i_{n+1}}^{\infty} a_k < \sum_{k=n}^{\infty} b_k.$$

Hence, by definition of $i_n$, we have $i_{n+1} > i_n$ for $n \in I$.

Let us now check the first condition. Suppose there exist distinct numbers $n_1, n_2, n_3 \in I$ such that $k \in \Delta_{n_1}, \Delta_{n_2}, \Delta_{n_3}$. Without loss of generality, $n_1 < n_2 < n_3$. Since $k \in \Delta_{n_1}$, it follows that $k \leq i_{n_1+1} \leq i_{n_2}$. Since $k \in \Delta_{n_3}$, it follows that $k \geq i_{n_3} \geq i_{n_2+1}$. Hence, $i_{n_2+1} \leq k \leq i_{n_2}$ and, therefore, $i_{n_2+1} = i_{n_2}$. Since $n_2 \in I$, it follows $i_{n_2+1} > i_{n_2}$. This contradiction shows that $|\{n \in I : k \in \Delta_n\}| \leq 2$. By definition, $k$ also belongs to at most one set $\Delta_n$, $n \notin I$. Consequently,

$$|\{n \in \mathbb{Z}^+ : k \in \Delta_n\}| \leq 3. \qquad \square$$

From the partition lemma, we deduce an operator lemma similar to Lemma 4.6. It extends Proposition 2 in [3], which is established there for the special case $p = 1$ by a completely different method.

**Lemma 6.2.** *Let $p \geq 1$. Let $a, b \in \ell^p$ such that $|b|^p \prec\prec_{\mathrm{tl}} |a|^p$. Then there exists an operator $T : \ell^p \to \ell^p$ such that*:

(i) $T(a) = b$.

(ii) $\|T\|_{p \to p} \leq 3^{1/p}$ *and* $\|T\|_{0 \to 0} \leq 3$.

*Proof.* We can assume that both sequences are nonnegative and decreasing. Apply Lemma 6.1 to $|a|^p$ and $|b|^p$. For every $n \in \mathbb{Z}^+$, choose a linear form $\varphi_n$ on $\ell^p$ of norm less than 1, supported on $\Delta_n$ and such that $\varphi_n(a) = \varphi_n(a\mathbf{1}_{\Delta_n}) = b_n$. Define

$$T : x \in \ell^p \mapsto (\varphi_n(x))_{n\in\mathbb{Z}^+}.$$

By construction, $T(a) = b$. Let us check the norm estimates. Let $x \in \ell^p$, then

$$\begin{aligned}\|T(x)\|_p^p = \sum_{n\in\mathbb{Z}^+} |\varphi_n(x)|^p &= \sum_{n\in\mathbb{Z}^+} |\varphi_n(x\mathbf{1}_{\Delta_n)})|^p \\ &\leq \sum_{n\in\mathbb{Z}^+}\sum_{k\in\Delta_n} |x_k|^p = \sum_{k\in\mathbb{Z}^+} |\{n : k \in \Delta_n\}|\,|x_k|^p \leq 3\|x\|_p^p.\end{aligned}$$

The second estimate is clear, using once again the fact that an integer $k$ belongs to at most three $\Delta_n$'s. □

The following remarks were communicated to the authors by Cwikel and Nilsson.

**Remark 6.3.** A bounded linear operator on $\ell^p$, $p \leq 1$ extends automatically to a bounded linear operator on $\ell^1$.

*Proof.* Indeed, let $(e_n)_{n\in\mathbb{Z}^+}$ be the canonical basis of $\ell^\infty$. Let $T$ be a contraction on $\ell^p$, $p < 1$. Then by Hölder's inequality $\|T(e_n)\|_1 \leq \|T(e_n)\|_p \leq \|T\|_{p\to p}$. By the triangle inequality, for any finite sequence $a = (a_n)_{n\in\mathbb{Z}^+}$,

$$\|T(a_n)\|_1 \leq \sum_{n\in\mathbb{Z}^+} |a_n|\|T(e_n)\|_1 \leq \|a\|_1\|T\|_{p\to p}.$$

Hence, $T$ extends to a contraction on $\ell^1$. □

**Remark 6.4.** The condition $p \geq 1$ in Lemma 6.2 is necessary.

*Proof.* Let us show that Lemma 6.2 cannot be true for $p < 1$. Assume by contradiction that there exists $c > 0$ such that for any finite sequences $a$ and $b$ in $\ell^p$ such that $|b|^p \prec\prec_{\mathrm{tl}} |a|^p$ there exists $T$ with $\|T\|_{p\to p} \leq c$ and $T(a) = b$. By Remark 6.3 above, we also have $\|T\|_{1\to 1} \leq c$. In particular, $\|b\|_1 \leq c\|a\|_1$. By considering $b = e_1$ and $a = \frac{1}{N^{1/p}}\sum_{i=1}^N e_i$ for $N$ large enough, one obtains a contradiction. □

We will not prove a sequence version of Lemma 4.5 to avoid the repetition of too many similar arguments. Fortunately, the expected result already appears in the literature, see [10, Theorem 3].

**Lemma 6.5.** *Let $p > 0$. Let $a, b \in \ell^\infty$ such that $|b|^p \prec\prec_{\mathrm{hd}} |a|^p$. Then there exists an operator $T : \ell^p \to \ell^p$ such that*:

(i) $T(a) = b$.

(ii) $\|T\|_{p\to p} \leq 8^{1/p}$ *and* $\|T\|_{\infty\to\infty} \leq 2^{1/p}$.

We conclude this subsection with a new interpolation theorem.

**Theorem 6.6.** *Let $p < q \in (0, \infty]$ such that $q \geq 1$. Let $E$ be a quasi-Banach sequence space. Then $E$ belongs to $Int(\ell^p, \ell^q)$ if and only if there exists $c_{p,E}$ and $c_{q,E}$ in $\mathbb{R}_{>0}$ such that*:

(i) *For any $u \in E$ and $v \in \ell^\infty$, if $|v|^p \prec\prec_{\text{hd}} |u|^p$, then $v \in E$ and $\|v\|_E \leq c_{p,E}\|u\|_E$.*

(ii) *Suppose $q < \infty$. For any $u \in E$ and $v \in \ell^\infty$, if $|v|^q \prec\prec_{\text{tl}} |u|^q$, then $v \in E$ and $\|v\|_E \leq c_{q,E}\|u\|_E$.*

*Proof.* The proof of the "only if" implication is identical to the proof of Theorem 5.1 using Lemmas 6.2 and 6.5 instead of Lemmas 4.5 and 4.4. The "if" implication is given by [8, Theorem 4.7]. □

***Upper Boyd index.*** Let us now recall the definition of the upper Boyd index, in the case of sequence spaces. For any $n \in \mathbb{N}$ define the dilation operator

$$D_n \colon \ell^\infty \to \ell^\infty, \quad (u_k)_{k\in\mathbb{Z}^+} \mapsto (u_{\lfloor k/n \rfloor})_{k\in\mathbb{Z}^+}.$$

Let $E$ be a symmetric function space. Define the Boyd index associated to $E$ by

$$\beta_E = \lim_{k\to\infty} \frac{\log\|D_k\|_{E\to E}}{\log k}.$$

Note that since $E$ is a quasi-Banach space, $\beta_E < \infty$.

In the next proposition, we relate the upper Boyd index to an interpolation property. We follow [23, Theorem 2].

**Proposition 6.7.** *Assume that $E$ is a quasi-Banach symmetric sequence space. Let $p < 1/\beta_E$. There exists a constant $C$ such that for any $u \in E$ and $v \in \ell^\infty$, satisfying $|v|^p \prec\prec_{\text{hd}} |u|^p$, we have $v \in E$ and $\|v\|_E \leq C\|u\|_E$.*

Define the map $V : \ell_\infty \to \ell_\infty$ by setting

$$Vu = \sum_{n=0}^{\infty} 2^{-n} D_{2^n} u$$

and the map $C : \ell_\infty \to \ell_\infty$ by

$$(Cu)(n) = \frac{1}{n+1} \sum_{i=0}^{n} u_n.$$

**Lemma 6.8.** *If $p < 1/\beta_E$, then*

$$\|(V(u^p))^{1/p}\|_E \leq c_{p,E}\|u\|_E, \quad 0 \leq u \in E.$$

*Proof.* Let $E_p$ be the $p$-concavification of $E$, that is,

$$E_p = \{f : |f|^{1/p} \in E\}, \quad \|f\|_{E_p} = \||f|^{1/p}\|_E^p.$$

Obviously, $E_p$ is a quasi-Banach space. Apply the Aoki–Rolewicz theorem to the space $E_p$ and fix $q = q_{p,E} > 0$ such that

$$\Big\|\sum_{n\geq 0} x_n\Big\|_{E_p}^q \leq C_{p,E}\sum_{n\geq 0}\|x_n\|_{E_p}^q.$$

For every $u \in E$, we have

$$\begin{aligned}\|(V(u^p))^{1/p}\|_E^{qp} = \|V(u^p)\|_{E_p}^q &= \Big\|\sum_{n=0}^{\infty}\frac{1}{2^n}(D_{2^n}u)^p\Big\|_{E_p}^q\\ &\leq C_{p,E}\sum_{n=0}^{\infty}\Big\|\frac{1}{2^n}(D_{2^n}u)^p\Big\|_{E_p}^q\\ &= C_{p,E}\sum_{n=0}^{\infty}2^{-nq}\|D_{2^n}u\|_E^{qp}.\end{aligned}$$

Let $r \in (p, \beta_E^{-1})$. By the definition of $\beta_E$, there exists $c_{p,E} > 0$ such that $\|D_n\|_{E\to E} \leq c_{p,E}\, n^{1/r}$ for any $n \in \mathbb{N}$. Therefore,

$$\begin{aligned}\|(V(u^p))^{1/p}\|_E^{qp} &\leq C_{p,E}\cdot c_{p,E}^q\cdot\sum_{n=0}^{\infty}2^{-nq}\,2^{nqp/r}\|u\|_E^{qp}\\ &= C_{p,E}\cdot c_{p,E}^q\cdot\frac{2^q}{2^q-2^{qp/r}}\cdot\|u\|_E^{qp}.\end{aligned}$$

□

**Lemma 6.9.** *If $x = \mu(x)$, then $Cx \leq 3Vx$ for every $x \in \ell_\infty$.*

*Proof.* Let $k \geq 0$. Since $x$ is decreasing, it follows that

$$(Cx)(2^k-1) = \frac{1}{2^k}\Big(x(0) + \sum_{i=0}^{k-1}\sum_{j=2^i}^{2^{i+1}-1}x(j)\Big) \leq \frac{1}{2^k}\Big(x(0)+\sum_{i=0}^{k-1}2^i\,x(2^i)\Big).$$

On the other hand, we have

$$\begin{aligned}(Vx)(2^{k+1}-1) &= \sum_{n\geq 0}2^{-n}\,x\Big(\Big\lfloor\frac{2^{k+1}-1}{2^n}\Big\rfloor\Big)\\ &= \sum_{n=0}^{k}2^{-n}x(2^{k+1-n}-1) + \sum_{n=k+1}^{\infty}2^{-n}\,x(0)\\ &= \frac{1}{2^k}\Big(x(0)+\sum_{i=0}^{k}2^i\,x(2^{i+1}-1)\Big).\end{aligned}$$

Again using the fact that $x$ is decreasing, we obtain

$$\sum_{i=0}^{k-1} 2^i x(2^i) = x(1) + \sum_{i=0}^{k-2} 2^{i+1} x(2^{i+1}) \\ \leq x(1) + 2\sum_{i=0}^{k-2} 2^i x(2^{i+1}-1) \leq 3\sum_{i=0}^{k} 2^i x(2^{i+1}-1).$$

Combining the three previous inequalities, we have just shown that for any $k \geq 0$,

$$(Cx)(2^k-1) \leq 3(Vx)(2^{k+1}-1).$$

Now let $n \geq 0$ and choose $k$ such that $n \in [2^k-1, 2^{k+1}-1]$. Since $Cx$ and $Vx$ are decreasing, we have

$$(Cx)(n) \leq (Cx)(2^k-1) \leq 3(Vx)(2^{k+1}-1) \leq 3(Vx)(n). \qquad \square$$

*Proof of Proposition 6.7.* Without loss of generality, $u = \mu(u)$ and $v = \mu(v)$. Since $v^p \prec\prec_{\mathrm{hd}} u^p$, it follows that

$$|v|^p \leq C(|v|^p) \leq C(|u|^p) \leq 3V(|u|^p),$$

where we used Lemma 6.9 to obtain the last inequality. By Lemma 6.8, we have

$$\|v\|_E \leq 3^{1/p} \|(V(|u|^p))^{1/p}\|_E \leq c_{p,E}\|u\|_p. \qquad \square$$

We are now ready to deliver a complete resolution of the conjecture stated by Levitina et al. in [20].

*Proof of Theorem 1.2.* Let $E$ be a quasi-Banach sequence space. Let $q \geq 1$. Recall that Theorem 1.2 states that the following two conditions are equivalent:

(a) There exists $p < q$ such that $E$ is an interpolation space for the couple $(\ell^p, \ell^q)$.

(b) There exists $c > 0$ such that for any $u \in E$ and $|v|^q \prec\prec_{\mathrm{tl}} |u|^q$, then $v \in E$ and $\|v\|_E \leq c\|u\|_E$.

Note that by Lemma 2.9, we may assume that $E$ is a symmetric space.

(a) $\Rightarrow$ (b). This is immediate by Theorem 6.6.

(b) $\Rightarrow$ (a). Let $p < 1/\beta_E$. By Proposition 6.7, for any sequence $u \in E$ and $v \in \ell^\infty$, if $|v|^p \prec\prec_{\mathrm{hd}} |u|^p$, $v \in E$ and $\|v\|_E \leq c_{p,E}\|u\|_E$. Applying Theorem 6.6 for indices $p$ and $q$, we obtain that $E$ belongs to $\mathrm{Int}(\ell^p, \ell^q)$. $\square$

## References

[1] J. Arazy and M. Cwikel, "A new characterization of the interpolation spaces between $L^p$ and $L^q$", *Math. Scand.* **55**:2 (1984), 253–270. MR Zbl

[2] S. V. Astashkin, “Interpolation of operators in quasinormed groups of measurable functions”, *Sibirsk. Mat. Zh.* **35**:6 (1994), 1215–1222. MR Zbl

[3] S. V. Astashkin, “A version of Calderón–Mityagin theorem for the class of rearrangement invariant groups”, *Nonlinear Anal.* **200** (2020), art. id. 112063. MR Zbl

[4] S. V. Astashkin and L. Maligranda, “Interpolation between $L_1$ and $L_p$, $1 < p < \infty$”, *Proc. Amer. Math. Soc.* **132**:10 (2004), 2929–2938. MR Zbl

[5] S. V. Astashkin, M. Cwikel, and P. G. Nilsson, “Arazy–Cwikel and Calderón–Mityagin type properties of the couples $(\ell^p, \ell^q)$, $0 \leq p < q \leq \infty$”, *Ann. Mat. Pura Appl.* (4) **202**:4 (2023), 1643–1678. MR Zbl

[6] J. Bergh and J. Löfström, *Interpolation spaces*: *an introduction*, Grundl. Math. Wissen. **223**, Springer, Berlin, 1976. MR Zbl

[7] M. Š. Braverman and A. A. Mekler, “The Hardy–Littlewood property for symmetric spaces”, *Sibirsk. Mat. Ž.* **18**:3 (1977), 522–540, 717. MR Zbl

[8] L. Cadilhac, “Majorization, interpolation and noncommutative Khinchin inequalities”, *Stud. Math.* **258**:1 (2021), 1–26. Zbl

[9] A.-P. Calderón, “Spaces between $L^1$ and $L^\infty$ and the theorem of Marcinkiewicz”, *Studia Math.* **26** (1966), 273–299. MR Zbl

[10] M. Cwikel, “Monotonicity properties of interpolation spaces, II”, *Ark. Mat.* **19**:1 (1981), 123–136. MR Zbl

[11] M. Cwikel and P. G. Nilsson, “An alternative characterization of normed interpolation spaces between $\ell^1$ and $\ell^q$”, preprint, 2017. Zbl arXiv 1711.04809

[12] M. Cwikel and J. Peetre (editors), *Interpolation spaces and allied topics in analysis*, Lecture Notes in Mathematics **1070**, Springer, Berlin, 1984. MR Zbl

[13] P. G. Dodds, T. K. Dodds, and F. A. Sukochev, “On $p$-convexity and $q$-concavity in non-commutative symmetric spaces”, *Integral Equations Operator Theory* **78**:1 (2014), 91–114. MR Zbl

[14] T. Holmstedt, “Interpolation of quasi-normed spaces”, *Math. Scand.* **26** (1970), 177–199. MR Zbl

[15] J. Huang and F. Sukochev, “Interpolation between $L_0(\mathcal{M}, \tau)$ and $L_\infty(\mathcal{M}, \tau)$”, *Math. Z.* **293**:3-4 (2019), 1657–1672. MR Zbl

[16] H. Hudzik and L. Maligranda, “An interpolation theorem in symmetric function $F$-spaces”, *Proc. Amer. Math. Soc.* **110**:1 (1990), 89–96. MR Zbl

[17] N. J. Kalton, “Convexity conditions for nonlocally convex lattices”, *Glasgow Math. J.* **25**:2 (1984), 141–152. MR Zbl

[18] N. Kalton and S. Montgomery-Smith, “Interpolation of Banach spaces”, pp. 1131–1175 in *Handbook of the geometry of Banach spaces*, vol. 2, North-Holland, 2003. MR Zbl

[19] S. G. Kreĭn, Y. I. Petunīn, and E. M. Semënov, *Interpolation of linear operators*, Translations of Mathematical Monographs **54**, American Mathematical Society, Providence, RI, 1982. MR Zbl

[20] G. Levitina, F. Sukochev, and D. Zanin, “Cwikel estimates revisited”, *Proc. Lond. Math. Soc.* (3) **120**:2 (2020), 265–304. MR Zbl

[21] G. G. Lorentz and T. Shimogaki, “Interpolation theorems for the pairs of spaces $(L^p, L^\infty)$ and $(L^1, L^q)$”, *Trans. Amer. Math. Soc.* **159** (1971), 207–221. MR Zbl

[22] B. S. Mityagin, “An interpolation theorem for modular spaces”, *Mat. Sb.* (*N.S.*) **66(108)**:4 (1965), 473–482. MR Zbl

[23] S. J. Montgomery-Smith, "The Hardy operator and Boyd indices", pp. 359–364 in *Interaction between functional analysis, harmonic analysis, and probability* (Columbia, MO, 1994), Lecture Notes in Pure and Appl. Math. **175**, Dekker, New York, 1996. MR Zbl

[24] J. Peetre and G. Sparr, "Interpolation of normed abelian groups", *Ann. Mat. Pura Appl.* (4) **92** (1972), 217–262. MR Zbl

[25] N. Popa, "Rearrangement invariant $p$-spaces on [0, 1], $0 < p < 1$, containing a complemented subspace isomorphic to $l_2$", *Rev. Roumaine Math. Pures Appl.* **28**:2 (1983), 153–160. MR Zbl

[26] A. A. Sedaev, E. M. Semenov, and F. A. Sukochev, "Fully symmetric function spaces without an equivalent Fatou norm", *Positivity* **19**:3 (2015), 419–437. MR Zbl

[27] G. Sparr, "Interpolation of weighted $L_p$-spaces", *Studia Math.* **62**:3 (1978), 229–271. MR Zbl

[28] F. Sukochev and D. Zanin, "Orbits in symmetric spaces", *J. Funct. Anal.* **257**:1 (2009), 194–218. MR Zbl

Received December 15, 2021. Revised March 6, 2024.

LÉONARD CADILHAC
SORBONNE UNIVERSITÉ
INSTITUT DE MATHÉMATIQUES DE JUSSIEU
PARIS
FRANCE
cadilhac@imj-prg.fr

FEDOR SUKOCHEV
SCHOOL OF MATHEMATICS AND STATISTICS
UNIVERSITY OF NEW SOUTH WALES
KENSINGTON
AUSTRALIA
f.sukochev@unsw.edu.au

DMITRIY ZANIN
SCHOOL OF MATHEMATICS AND STATISTICS
UNIVERSITY OF NEW SOUTH WALES
KENSINGTON
AUSTRALIA
d.zanin@unsw.edu.au

PACIFIC JOURNAL OF MATHEMATICS
Vol. 328, No. 2, 2024
https://doi.org/10.2140/pjm.2024.328.255

# FINITE AXIOMATIZABILITY OF THE RANK AND THE DIMENSION OF A PRO-$\pi$ GROUP

MARTINA CONTE AND BENJAMIN KLOPSCH

*In memory of Avinoam Mann*

**The Prüfer rank $\mathrm{rk}(G)$ of a profinite group $G$ is the supremum, across all open subgroups $H$ of $G$, of the minimal number of generators $\mathrm{d}(H)$. It is known that, for any given prime $p$, a profinite group $G$ admits the structure of a $p$-adic analytic group if and only if $G$ is virtually a pro-$p$ group of finite rank. The dimension $\dim G$ of a $p$-adic analytic profinite group $G$ is the analytic dimension of $G$ as a $p$-adic manifold; it is known that $\dim G$ coincides with the rank $\mathrm{rk}(U)$ of any uniformly powerful open pro-$p$ subgroup $U$ of $G$.**

**Let $\pi$ be a finite set of primes, let $r \in \mathbb{N}$ and let $\mathrm{r} = (r_p)_{p\in\pi}$, $\mathrm{d} = (d_p)_{p\in\pi}$ be tuples in $\{0, 1, \dots, r\}$. We show that there is a single sentence $\sigma_{\pi,r,\mathrm{r},\mathrm{d}}$ in the first-order language of groups such that for every pro-$\pi$ group $G$ the following are equivalent: (i) $\sigma_{\pi,r,\mathrm{r},\mathrm{d}}$ holds true in the group $G$, that is, $G \models \sigma_{\pi,r,\mathrm{r},\mathrm{d}}$; (ii) $G$ has rank $r$ and, for each $p \in \pi$, the Sylow pro-$p$ subgroups of $G$ have rank $r_p$ and dimension $d_p$.**

**Loosely speaking, this shows that, for a pro-$\pi$ group $G$ of bounded rank, the precise rank of $G$ as well as the ranks and dimensions of the Sylow subgroups of $G$ can be recognized by a single sentence in the basic first-order language of groups.**

## 1. Introduction

Nies, Segal and Tent [Nies et al. 2021] carried out an investigation of the model-theoretic concept of finite axiomatizability in the context of profinite groups. For instance, a profinite group $G$ is finitely axiomatizable within a class $\mathcal{C}$ of profinite groups, with respect to the first-order language $\mathcal{L}_{\mathrm{gp}}$ of groups, if there is a sentence $\psi_{G,\mathcal{C}}$ in $\mathcal{L}_{\mathrm{gp}}$ such that the following holds: a profinite group $H$ in $\mathcal{C}$ is isomorphic to $G$ if and only if $\psi_{G,\mathcal{C}}$ holds true in $H$, in symbols $H \models \psi_{G,\mathcal{C}}$. More generally,

The research was conducted in the framework of the DFG-funded research training group “GRK 2240: Algebro-Geometric Methods in Algebra, Arithmetic and Topology”.

*MSC2020:* primary 20E18; secondary 03C98, 20A15, 20D15, 20D20, 22E20.

*Keywords:* pro-$\pi$ group, $p$-adic analytic group, finite nilpotent group, Prüfer rank, dimension, finite axiomatizability.

one takes interest in whether specific properties or invariants of profinite groups, again within a given class $\mathcal{C}$, can be detected uniformly by a single sentence in $\mathcal{L}_{\mathrm{gp}}$.

Our main interest is in finitely generated profinite groups. Nikolov and Segal [2007] established that such groups are strongly complete; loosely speaking, this means that the topology of a finitely generated profinite group is already predetermined by the abstract group structure. Jarden and Lubotzky [2008] used Nikolov and Segal's finite width results for certain words to prove that every finitely generated profinite group is "first-order rigid", i.e., determined up to isomorphism by its first-order theory, within the class of profinite groups. By restricting to *finite* axiomatizability, we probe for more delicate first-order properties within suitable classes of finitely generated profinite groups.

In this paper we focus on the class of profinite groups of finite Prüfer rank, from now on "rank" for short. This invariant is connected to, but not to be confused with, the minimal number of generators: the rank of a profinite group $G$ is defined as

$$\mathrm{rk}(G) = \sup\{\mathrm{d}(H) \mid H \leq_{\mathrm{o}} G\} = \sup\{\mathrm{d}(H) \mid H \leq_{\mathrm{c}} G\},$$

where $\mathrm{d}(H)$ denotes the minimal number of generators of a topological group $H$ and, as indicated, $H$ runs over all open or all closed subgroups of $G$. It is not difficult to see that the rank of $G$ is the supremum of the ranks of its finite continuous quotients, i.e., $\mathrm{rk}(G) = \sup\{\mathrm{rk}(G/N) \mid N \trianglelefteq_{\mathrm{o}} G\}$. The rank plays a central role in the structure theory of $p$-adic Lie groups. It is known that, for any given prime $p$, a profinite group $G$ admits the structure of a $p$-adic analytic group if and only if $G$ is virtually a pro-$p$ group of finite rank. The dimension $\dim G$ of a $p$-adic analytic profinite group $G$ is the analytic dimension of $G$ as a $p$-adic manifold; in fact, $\dim G \leq \mathrm{rk}(G)$ and $\dim G$ coincides with the rank $\mathrm{rk}(U)$ of any uniformly powerful open pro-$p$ subgroup $U$ of $G$. Further details and related results about $p$-adic analytic pro-$p$ groups can be found in [Dixon et al. 1999]; the concise introduction [Klopsch 2011] summarizes key aspects of the theory.

Loosely speaking, our aim is to show that, for every finite set of primes $\pi$, the precise rank $r$ as well as the ranks $\mathbf{r} = (r_p)_{p\in\pi}$ and dimensions $\mathbf{d} = (d_p)_{p\in\pi}$ of the Sylow pro-$p$ subgroups of any pro-$\pi$ group $G$ of finite rank can be recognized by a single sentence $\sigma_{\pi,r,\mathbf{r},\mathbf{d}}$ in the first-order language of groups $\mathcal{L}_{\mathrm{gp}}$. The starting point for our investigation is Proposition 5.1 in [Nies et al. 2021] which states: Given $r \in \mathbb{N}$, there is an $\mathcal{L}_{\mathrm{gp}}$-sentence $\rho_{p,r}$ such that for every pro-$p$ group $G$, the following implications hold:

$$\mathrm{rk}(G) \leq r \quad \Longrightarrow \quad G \models \rho_{p,r} \quad \Longrightarrow \quad \mathrm{rk}(G) \leq r(2+\log_2(r)).$$

Our first theorem both strengthens and generalizes this result. The $p$-rank $\mathrm{rk}_p(G)$ of a profinite group $G$ is the common rank of all Sylow pro-$p$ subgroups of $G$. A sentence $\phi$ in $\mathcal{L}_{\mathrm{gp}}$ is called an $\exists\forall\exists$-sentence if it results from a quantifier-free formula $\phi_0$ by means of a sequence of existential, universal and existential quantifications (in this order), rendering the free variables of $\phi_0$ to be bound in $\phi$; compare with Example 3.1.

**Theorem 1.1.** *Let $\pi$ be a finite set of primes. Let $r \in \mathbb{N}$ and let $\mathbf{r} = (r_p)_{p\in\pi}$ be a tuple in $\{0, 1, \ldots, r\}$. Then there exists an $\exists\forall\exists$-sentence $\varrho_{\pi,r,\mathbf{r}}$ in $\mathcal{L}_{\mathrm{gp}}$ such that, for every pro-$\pi$ group $G$, the following are equivalent*:

(i) $\mathrm{rk}(G) = r$, *and* $\mathrm{rk}_p(G) = r_p$ *for every* $p \in \pi$.

(ii) $\varrho_{\pi,r,\mathbf{r}}$ *holds in $G$, i.e., $G \models \varrho_{\pi,r,\mathbf{r}}$.*

It is no coincidence that the sentences $\varrho_{\pi,r,\mathbf{r}}$ which we manufacture to prove the theorem depend on the given set of primes $\pi$. A standard ultraproduct construction reveals that, for every *infinite* set of primes $\tilde{\pi}$ and $r \in \mathbb{N}$, there is no $\mathcal{L}_{\mathrm{gp}}$-sentence $\vartheta_{\tilde{\pi},r}$ which could identify, uniformly across $p \in \tilde{\pi}$, among pro-$p$ groups $G$ those with rank $\mathrm{rk}(G) = r$; see Proposition 3.3.

In addition to Theorem 1.1 we establish a corresponding theorem which concerns the dimensions of the Sylow subgroups of a profinite group of finite rank.

**Theorem 1.2.** *Let $\pi$ be a finite set of primes. Let $r \in \mathbb{N}$ and let $\mathbf{d} = (d_p)_{p\in\pi}$ be a tuple in $\{0, 1, \ldots, r\}$. Then there exists an $\exists\forall\exists$-sentence $\tau_{\pi,r,\mathbf{d}}$ in $\mathcal{L}_{\mathrm{gp}}$ such that, for every pro-$\pi$ group $G$ with $\mathrm{rk}(G) = r$, the following are equivalent*:

(i) *For every $p \in \pi$, the Sylow pro-$p$ subgroups of $G$ have dimension $d_p$.*

(ii) $\tau_{\pi,r,\mathbf{d}}$ *holds in $G$, i.e., $G \models \tau_{\pi,r,\mathbf{d}}$.*

In combination, the two theorems provide the first-order sentences $\sigma_{\pi,r,\mathbf{r},\mathbf{d}}$ with the properties promised above. It is remarkable that such sentences exist in the basic language $\mathcal{L}_{\mathrm{gp}}$ of groups. In connection with $p$-adic analytic profinite groups, it is often necessary to employ suitably expanded languages in order to capture part of the topological or analytic structure; compare with [Macpherson and Tent 2016]. We do not need to enlarge the language at all. Moreover, the complexity of $\sigma_{\pi,r,\mathbf{r},\mathbf{d}}$ remains within three alternations of $\exists$- and $\forall$-quantifiers, even though the sentences that we manufacture depend strongly on the given set of primes $\pi$.

As we will show, the proofs of Theorems 1.1 and 1.2 reduce, in a certain sense, to the simpler setting of pronilpotent pro-$\pi$ groups, termed $\mathcal{C}_\pi$-groups by [Nies et al. 2021, Section 5]. We recall that, even in the pronilpotent case, Sylow subgroups are not in general definable and there is no standard reduction to pro-$p$ groups; this can be seen from relative quantifier elimination results (down to positive primitive formulas) for modules over rings; see [Prest 1988, Sections 2.4 and 2.$\mathbb{Z}$]. Part of our task is to develop appropriate tools to by-pass this obstacle.

Key to our approach for proving Theorems 1.1 and 1.2 are purely group-theoretic considerations leading to Theorem 2.1 and its corollary, about profinite groups which are virtually pronilpotent and of finite rank. Specialising to the setting of finite nilpotent groups to ease the exposition at this point, we can formulate the central insight as follows.

**Theorem 1.3.** *Let $G$ be a finite nilpotent group of rank $r = \mathrm{rk}(G)$. Then*

$$\mathrm{rk}(G) = \mathrm{rk}\big(G/\Phi^{j(r)}(G)\big) \quad \text{for } j(r) = 2r + \lceil \log_2(r) \rceil + 2,$$

*where $\Phi^{j(r)}(G)$ denotes the $j(r)$-th iterated Frattini subgroup of $G$.*

It is an open problem to identify, if at all possible, even smaller canonical quotients which witness the full rank of a finite nilpotent group.

Following a suggestion of González-Sánchez, we derive from a result of Héthelyi and Lévai [2003] a new description of the dimension of a finitely generated powerful pro-$p$ group; this is useful for establishing Theorem 1.2, but also of independent interest.

**Theorem 1.4.** *Let $G$ be a finitely generated powerful pro-$p$ group with torsion subgroup $T$, and let $\Omega_{\{1\}}(G) = \{g \in G \mid g^p = 1\}$ denote the set of all elements of order 1 or $p$ in $G$. Then*

$$\dim(G) = \mathrm{d}(G) - \log_p |\Omega_{\{1\}}(G)| = \mathrm{d}(G) - \mathrm{d}(T).$$

With a view toward possible future investigations, we add a final comment and a question. Naturally one wonders whether "being of finite rank" per se can be captured by a suitable first-order sentence. Results of Feferman and Vaught [1959] imply that, even for a fixed prime $p$, there is no set $\mathcal{T}_p$ of $\mathcal{L}_{\mathrm{gp}}$-sentences (and in particular no single sentence) which identifies among the collection of all pro-$p$ groups those that possess finite rank. Indeed, the class of pro-$p$ groups of finite rank is closed under taking finite cartesian products, but an infinite cartesian product of nontrivial pro-$p$ groups of finite rank is not even finitely generated. Therefore [Feferman and Vaught 1959, Corollary 6.7] shows that no $\mathcal{T}_p$ with the desired property exists. However, a modified question suggests itself. Given $d \geq 2$, is there a set $\mathcal{T}_{p,d}$ of $\mathcal{L}_{\mathrm{gp}}$-sentences (possibly a single sentence) such that the following holds for pro-$p$ groups $G$ with $\mathrm{d}(G) \leq d$: the group $G$ has finite rank if and only if $G$ satisfies $\mathcal{T}_{p,d}$?

**Remark.** Our proofs for Theorems 1.1 and 1.2 involve results of Lucchini [1997] and an observation of Mazurov [1994] which currently rely on the classification of finite simple groups. However, in suitable circumstances, e.g., if we restrict attention to prosoluble groups, the required ingredients are known to hold without use of the classification; compare with [Lucchini 1989, Section 5]. If $2 \notin \pi$, the Odd Order Theorem guarantees that all pro-$\pi$ groups are prosoluble.

**Organization and Notation.** In Section 2 we prove Theorem 2.1 and its corollary, which specialize to Theorem 1.3. In Example 2.3 we discuss limitations of our strategy; Proposition 3.3 shows that Theorem 2.1 does not generalize to groups involving infinitely many primes. In Section 3 we establish Theorem 1.1. In Section 4 we prove Theorem 1.4 and deduce Theorem 1.2.

Our notation is mostly standard and in line with current practice. For instance, $\mathrm{Z}(G)$ denotes the centre of a group $G$, and $C_n$ denotes a cyclic group of order $n$. The meaning of possibly less familiar terms, such as $\Phi(G)$ for the Frattini subgroup and $\Phi_p(G)$ for the $p$-Frattini subgroup of a group $G$, are explained at their first occurrence. We deal exclusively with profinite groups. Accordingly, notions such as the Frattini subgroup, the commutator subgroup or the subgroup generated by a given set are tacitly understood in the topological sense: in each case we mean the topological closure of the corresponding abstract subgroup. Basic model-theoretic concepts which are employed without further reference are covered by standard texts such as [Hodges 1993].

## 2. Detecting the rank in bounded quotients

Every compact $p$-adic analytic group $G$ has finite rank and contains an open normal powerful pro-$p$ subgroup $F$. Since $F$ is a pro-$p$ group, its Frattini subgroup $\Phi(F)$ coincides with $[F, F]F^p$ and $F/\Phi(F)$ is elementary abelian. Since $F$ is powerful, we know that $\mathrm{rk}(F) = \mathrm{d}(F) = \mathrm{rk}(F/\Phi(F))$; see [Dixon et al. 1999, Theorem 3.8]. Furthermore, the iterated Frattini series $\Phi^j(F)$, $j \in \mathbb{N}$, of $F$ coincides with both the lower $p$-series and the iterated $p$-power series of $F$. It provides a base of neighbourhoods for 1 in $G$ consisting of open normal subgroups. Consequently, the rank of $G$ is given by

$$\mathrm{rk}(G) = \sup\{\mathrm{rk}(G/\Phi^j(F)) \mid j \in \mathbb{N}\} = \max\{\mathrm{rk}(G/\Phi^j(F)) \mid j \in \mathbb{N}\};$$

in other words, $\mathrm{rk}(G)$ is the terminal value of the nondecreasing, eventually constant sequence $\mathrm{rk}(G/\Phi^j(F))$, $j \in \mathbb{N}$.

It is natural to look for an upper bound for the smallest $j \in \mathbb{N}$ such that $\mathrm{rk}(G) = \mathrm{rk}(G/\Phi^j(F))$, a bound that is, as far as possible, independent of $p$ and any special features of the pair $F \le G$. Based on our current knowledge, the strongest possible outcome could be that $\mathrm{rk}(G) = \mathrm{rk}(G/\Phi(F))$ holds without any exceptions. More modestly, one can ask for weaker bounds, possibly contingent on additional information regarding $\mathrm{rk}(G)$.

We establish a result of the latter kind, which applies more generally to profinite groups $G$ of finite rank that admit a pronilpotent open normal subgroup $F$. We recall that the $p$-rank $\mathrm{rk}_p(G)$ of a profinite group $G$ is simply the rank $\mathrm{rk}(P)$ of a Sylow pro-$p$ subgroup $P$ of $G$. Furthermore, we write $\Phi_p(G) = [G, G]G^p$ for the $p$-Frattini subgroup of $G$; the $p$-Frattini quotient $G/\Phi_p(G)$ is the largest elementary abelian pro-$p$ quotient of the profinite group $G$.

**Theorem 2.1.** *Let $R \in \mathbb{N}$. Suppose that the profinite group $G$ has an open normal subgroup $F \trianglelefteq_{\mathrm{o}} G$ which is pronilpotent and such that each Sylow subgroup of $F$ is powerful.*

(i) *For every prime $p$ such that $\mathrm{rk}_p(G) \le R$, the $p$-rank satisfies*

$$\mathrm{rk}_p(G) = \mathrm{rk}_p\big(G/\Phi^{2R+1}(F)\big).$$

(ii) *If $\mathrm{rk}(G) \le R$, then*

$$\mathrm{rk}(G) = \mathrm{rk}\big(G/\Phi^{2R+1}(F)\big).$$

*Proof.* It is convenient to write $F_i = \Phi^i(F)$ for $i \in \mathbb{N}$.

(i) Let $p$ be a prime such that $r_p = \mathrm{rk}_p(G) \le R$. We show that $r_p = \mathrm{rk}_p(G/F_{2R+1})$. Since $F$ is pronilpotent, its Hall pro-$p'$ subgroup $P'$ is normal in $G$; compare with [Ribes and Zalesskii 2010, Section 2.3]. Working modulo $P'$, we may assume without loss of generality that $F$ is a powerful pro-$p$ group. In this situation $G$ is virtually a pro-$p$ group. Clearly, we have $r_p \ge \mathrm{rk}_p(G/F_{2R+1})$. For a contradiction, we assume that $r_p > \mathrm{rk}_p(G/F_{2R+1})$. Choose a pro-$p$ subgroup $H \le_{\mathrm{o}} G$ of minimal index among the open pro-$p$ subgroups of $G$ with $\mathrm{d}(H) = r_p$. In particular, this means that $\mathrm{d}(H) > \mathrm{d}(HF_{2R+1}/F_{2R+1})$.

The sequence $\mathrm{d}(HF_j/F_j)$, $j \in \mathbb{N}$, is nondecreasing and eventually constant, with final constant value $\mathrm{d}(H)$. Since $\mathrm{d}(H) = r_p < 2R + 1$, we conclude that $\mathrm{d}(HF_j/F_j)$, $j \in \mathbb{N}$, cannot be strictly increasing until it becomes constant. Hence there exists $j = j(H) \in \mathbb{N}$ such that

$$\begin{aligned} \text{(2-1)} \qquad \mathrm{d}(HF_j/F_j) = \mathrm{d}(HF_{j+1}/F_{j+1}) &< \mathrm{d}(HF_{j+2}/F_{j+2}) \\ &< \cdots < \mathrm{d}(HF_{j+k+1}/F_{j+k+1}) = \mathrm{d}(H) \end{aligned}$$

for suitable $k = k(H)$ with $1 \le k \le r_p \le R$. In particular, this set-up implies that $j + k + 1 > 2R + 1$, hence $j > R$ and $2j \ge j + R + 1 \ge j + k + 1$. Consequently, we see that $[F_j, F_j] \subseteq F_{2j} \subseteq F_{j+k+1}$ and there is no harm in assuming that

$$[F_j, F_j] = F_{2j} = 1.$$

This reduction renders $G$ finite, with abelian normal $p$-subgroups

$$A = F_j \quad \text{and} \quad B = F_{j+1} = \Phi(F_j) = A^p.$$

We set $l = \mathrm{d}(H/(H \cap B)) = \mathrm{d}(HB/B) < \mathrm{d}(H) = r_p$ and choose generators $y_1, \ldots, y_l$ for $H$ modulo $H \cap B$ so that

$$L = \langle y_1, \ldots, y_l \rangle \le H$$

satisfies $LB = HB$. Put $m = \mathrm{d}(H) - l = r_p - l \ge 1$. A collection of elements generates $H$ if and only if it generates the Frattini quotient $H/\Phi(H)$; the latter is elementary abelian, because $H$ is a $p$-group. Thus the minimal generating set $y_1, \ldots, y_l$ modulo $H \cap B$ can be supplemented to a minimal generating set for $H$:

there are $b_1, \dots, b_m \in B$ such that

$$H = \langle y_1, \dots, y_l, b_1, \dots, b_m \rangle \quad \text{with } \mathrm{d}(H) = r_p = l + m.$$

We put $M = \langle b_1, \dots, b_m \rangle^H \trianglelefteq H$ so that $H = LM$.

Choose $a_1, \dots, a_m \in A$ with $b_i = a_i^p$ for $1 \le i \le m$ and set

$$\widetilde{H} = \langle y_1, \dots, y_l, a_1, \dots, a_m \rangle \le G.$$

We claim that $\widetilde{H}$ is a $p$-subgroup of $G$ such that

$$|G : \widetilde{H}| < |G : H| \quad \text{and} \quad \mathrm{d}(\widetilde{H}) = r_p, \tag{2-2}$$

which yields the required contradiction.

Clearly, $\widetilde{H} \le HA$ is a $p$-group and $H \subseteq \widetilde{H}$. Moreover, we see that $HA = \widetilde{H}A = LA$. We may assume without loss of generality that $G = LA$. In this situation $G$ is a $p$-group; furthermore, $L \cap A \trianglelefteq G$ is normal. By construction, compare with (2-1), we have $\mathrm{d}(L/(L \cap A)) = \mathrm{d}(HA/A) = \mathrm{d}(HB/B) = l = \mathrm{d}(L)$. Thus $L \cap A \subseteq \Phi(L) \subseteq \Phi(H)$ and there is no harm in assuming $L \cap A = 1$. This gives

$$G = L \ltimes A, \quad H = L \ltimes M \quad \text{and} \quad \widetilde{H} = L \ltimes \widetilde{M} \quad \text{for } \widetilde{M} = \langle a_1, \dots, a_m \rangle^{\widetilde{H}}.$$

We supplement $y_1, \dots, y_l$ to a minimal generating set $y_1, \dots, y_l, \tilde{a}_1, \dots, \tilde{a}_n$ for the $p$-group $\widetilde{H}$, for suitable $n \in \{0, 1, \dots, m\}$ and $\tilde{a}_1, \dots, \tilde{a}_n \in \widetilde{M}$. The $p$-power map $g \mapsto g^p$ induces a surjective $L$-invariant homomorphism $\alpha \colon \widetilde{M} \to M$ between finite abelian $p$-groups. This implies $|\widetilde{M}| > |M|$ and thus $|G : \widetilde{H}| < |G : H|$. Furthermore, using the identity map on $L$ in combination with $\alpha$, we obtain a surjective homomorphism from $\widetilde{H} = L \ltimes \widetilde{M}$ onto $L \ltimes M = H$. This shows that $r_p = \mathrm{d}(H) \le \mathrm{d}(\widetilde{H}) \le r_p$ and hence $\mathrm{d}(\widetilde{H}) = r_p$, which completes the proof of (2-2).

(ii) Now suppose that $\mathrm{rk}(G) \le R$. Clearly, the maximal local rank

$$\mathrm{mlr}(G) = \max\bigl(\{\mathrm{rk}_p(G) \mid p \text{ prime}\}\bigr)$$

is at most $\mathrm{rk}(G)$. Conversely, Lucchini [1997, Theorem 3 and Corollary 4] established that

$$\mathrm{rk}(G) \le \mathrm{mlr}(G) + 1,$$

with equality if and only if there are

- an odd prime $p$ such that $r_p = \mathrm{rk}_p(G) = \mathrm{mlr}(G)$ and
- an open subgroup $H \le_{\mathrm{o}} G$ and $N \trianglelefteq_{\mathrm{o}} H$ such that

  $$H/\Phi_p(N) \cong H/N \ltimes N/\Phi_p(N) \cong C_q \ltimes C_p^{\mathrm{mlr}(G)},$$

  where $H/N \cong C_q$ is cyclic of prime order $q \mid (p-1)$, the $p$-Frattini quotient $N/\Phi_p(N) \cong C_p^{\mathrm{mlr}(G)}$ is elementary abelian of rank $\mathrm{mlr}(G)$, and $H/N$ acts via conjugation faithfully on $N/\Phi_p(N)$ by power automorphisms (i.e., by nonzero homotheties if we regard $N/\Phi_p(N)$ as an $\mathbb{F}_p$-vector space).

For short let us refer, somewhat effusively, within this proof to such a pair $(H, N)$ as a "runaway couple" for $G$ with respect to $p$.

By (i), we have $\mathrm{mlr}(G) = \mathrm{mlr}(G/F_{2R+1})$, and hence it suffices to show: if $G$ admits a runaway couple, then so does $G/F_{2R+1}$, in fact, with respect to the same prime. Suppose that $(H, N)$ is a runaway couple for $G$ with respect to an odd prime $p$ so that $H/\Phi_p(N) \cong C_q \ltimes C_p^{r_p}$ as detailed above, with the additional property that $|G : H|$ is as small as possible. Assume for a contradiction that $G/F_{2R+1}$ does not admit a runaway couple.

As in the proof of (i) there is no harm in factoring out the Hall pro-$p'$ subgroup $P'$ of $F$, because $H \cap F \subseteq N$ and $H \cap P' \subseteq \Phi_p(N)$. Consequently we may as well assume that $F \trianglelefteq_{\mathrm{o}} G$ is a powerful pro-$p$ group, which makes $G$ virtually a pro-$p$ group.

As in the proof of (i), the sequence

$$\mathrm{d}\big(H/\big((H \cap F_j)\Phi_p(N)\big)\big) = \mathrm{d}\big(HF_j/\Phi_p(N)F_j\big), \quad j \in \mathbb{N},$$

is nondecreasing and eventually constant, with final constant value

$$\mathrm{d}(H/\Phi_p(N)) = \mathrm{d}(H) = r_p + 1 < 2R + 1.$$

We use the same arguments as before to conclude that there exists $j = j(H)$ such that the analogue of (2-1) for $H/\Phi_p(N)$ holds and we reduce to the situation where $[F_j, F_j] = F_{2j} = 1$. This reduction renders $G$ finite, with abelian normal $p$-subgroups

$$A = F_j \quad \text{and} \quad B = F_{j+1} = \Phi(F_j) = A^p;$$

furthermore, we have

$$l = \mathrm{d}\big(N/\big((H \cap A)\Phi_p(N)\big)\big) = \mathrm{d}\big(N/\big((H \cap B)\Phi_p(N)\big)\big) < \mathrm{d}(N/\Phi_p(N)) = r_p. \tag{2-3}$$

It suffices to produce a runaway couple $(\widetilde{H}, \widetilde{N})$ for the group $HA$ with respect to $p$ such that $|HA : \widetilde{H}| < |HA : H|$; thus we may assume that

$$G = HA.$$

This reduction allows us to conclude that $\Phi_p(N) \cap A \trianglelefteq G$ and there is no harm in assuming $\Phi_p(N) \cap A = 1$. Likewise $M = H \cap A \trianglelefteq G$, and reduction modulo $\Phi_p(N)$ induces an embedding of $M \le N$ into the elementary abelian group $N/\Phi_p(N) \cong C_p^{r_p}$. Using (2-3), we conclude that

$$M = H \cap A = H \cap B = \langle b_1, \ldots, b_m \rangle \cong C_p^m \quad \text{for } m = r_p - l \ge 1.$$

The normal subgroup $M\Phi_p(N) \trianglelefteq H$ decomposes as a direct product $M \times \Phi_p(N)$. Recall that $H/\Phi_p(N) \cong C_q \ltimes C_p^{r_p}$, with the action given by power automorphisms.

We build a minimal generating set $x, y_1, \dots, y_l, b_1, \dots, b_m$ for $H$ modulo $\Phi_p(N)$ by choosing

$$x \in H \smallsetminus N \quad \text{and} \quad y_1, \dots, y_l \in N$$

which supplement $b_1, \dots, b_m$ suitably. We set

$$L_1 = \langle x, y_1, \dots, y_l \rangle \le H \quad \text{and} \quad L = L_1 \Phi_p(N) \le H.$$

In this situation $H = LM$ and we claim that $L \cap M = 1$ so that

$$H = L \ltimes M.$$

Indeed, our construction yields that the intersection in $H/\Phi_p(N) \cong C_q \ltimes C_p^{l+m}$ of the subgroups

$$L/\Phi_p(N) = \langle \bar{x} \rangle \ltimes \langle \bar{y_1}, \dots, \bar{y_l} \rangle \cong C_q \ltimes C_p^l \quad \text{and} \quad M\Phi_p(N)/\Phi_p(N) \cong M \cong C_p^m$$

is trivial. This gives $L \cap M \subseteq \Phi_p(N)$ and consequently $L \cap M \subseteq \Phi_p(N) \cap M = 1$.

Put $\widetilde{M} = \{a \in A \mid a^p \in M\} \trianglelefteq G$. Recall that $M = H \cap B$ and $B = A^p$. The $p$-power map constitutes a surjective $G$-equivariant homomorphism $\widetilde{M} \to M$ whose kernel $K \trianglelefteq G$, say, includes $M$. From $L \cap M = 1$ we conclude that $LK \cap \widetilde{M} = (L \cap \widetilde{M})K \subseteq K$. Moreover, we have $L \cap K \subseteq H \cap A = M$ and thus $L \cap K \subseteq L \cap M = 1$.

These considerations show that the group $\widetilde{H} = L\widetilde{M}$ maps onto

$$\widetilde{H}/K \cong LK/K \ltimes \widetilde{M}/K \cong L \ltimes M = H,$$

and hence onto $C_q \ltimes C_p^{r_p}$. Thus $\widetilde{H}$ gives rise to a runaway couple for $G$, with respect to the prime $p$, just as $H$ does. To conclude the proof we observe that $|K| \ge |M| \ge p$ implies $|\widetilde{H}| > |\widetilde{H}|/|K| = |H|$ and hence $|G : \widetilde{H}| < |G : H|$. $\square$

The following corollary yields in particular Theorem 1.3 about finite nilpotent groups, which was showcased in the introduction for its succinctness.

**Corollary 2.2.** *Let $R \in \mathbb{N}$. Suppose that the profinite group $G$ has an open normal subgroup $F \trianglelefteq_{\mathrm{o}} G$ which is pronilpotent.*

(i) *If $\mathrm{rk}_p(G) \le R$ for some prime $p$, then*

$$\mathrm{rk}_p(G) = \mathrm{rk}_p\big(G/\Phi^{2R+\lceil \log_2(R) \rceil+2}(F)\big).$$

(ii) *If $\mathrm{rk}(G) \le R$, then*

$$\mathrm{rk}(G) = \mathrm{rk}\big(G/\Phi^{2R+\lceil \log_2(R) \rceil+2}(F)\big).$$

*Proof.* As in the proof of Theorem 2.1, one reduces to the case in which $F$ is a pro-$p$ group for a single prime $p$. From $\mathrm{rk}(F) \le R$ it follows that $\Phi^{\lceil \log_2(R) \rceil+1}(F) \trianglelefteq_{\mathrm{o}} G$ is powerful; compare with [Dixon et al. 1999, Chapter 2, Exercise 6]. Thus we can apply Theorem 2.1 to $\Phi^{\lceil \log_2(R) \rceil+1}(F)$ in place of $F$. $\square$

The following example puts the basic idea behind the proof of Theorem 2.1 into perspective. It indicates that one would need to take a different approach or at least make more careful choices in order to eliminate the dependency on the parameter $R$. Indeed, the example yields, for $p > n \geq 2$, a pro-$p$ group $G$, a powerful open normal subgroup $F \trianglelefteq_{\mathrm{o}} G$ and an open subgroup $H \leq_{\mathrm{o}} G$ such that $\mathrm{d}(H) = \mathrm{rk}(G)$ but $\mathrm{d}(\widetilde{H}\Phi^n(F)/\Phi^n(F)) < \mathrm{rk}(G)$ for all $\widetilde{H} \leq_{\mathrm{o}} G$ with $\widetilde{H} \supseteq H$.

**Example 2.3.** Let $n \in \mathbb{N}$ and consider the metabelian pro-$p$ group

$$G = C \ltimes A, \qquad \text{where } C = \langle c \rangle \cong \mathbb{Z}_p, \quad A = \langle a_1, \dots, a_n \rangle \cong \mathbb{Z}_p^n$$

and the action of $C$ on $A$ is given by

$$a_i^c = a_i a_{i+1} \quad \text{for } 1 \leq i < n, \qquad \text{and} \qquad a_n^c = a_n.$$

Here $\mathbb{Z}_p$ denotes the additive group of the $p$-adic integers, viz. the infinite procyclic pro-$p$ group. Then $G = \langle c, a_1 \rangle$ is 2-generated, nilpotent of class $n$ and has rank $\mathrm{rk}(G) = n + 1$. For instance,

$$H = \langle c, a_1^{p^{n-1}}, a_2^{p^{n-2}}, \dots, a_{n-1}^p, a_n \rangle \leq_{\mathrm{o}} G$$

requires $n + 1$ generators.

Suppose that $p > n \geq 2$. Then $F = \langle c^p \rangle \ltimes A \trianglelefteq_{\mathrm{o}} G$ is powerful, and $\Phi^j(F) = \langle c^{p^j} \rangle \ltimes A^{p^{j-1}}$ for $j \in \mathbb{N}$. Thus any subgroup $\widetilde{H} \leq_{\mathrm{o}} G$ with $\widetilde{H}F = HF = \langle c \rangle F$ and $\mathrm{d}(\widetilde{H}) = \mathrm{d}(\widetilde{H}\Phi^n(F)/\Phi^n(F))$ requires less than $\mathrm{d}(H) = n + 1$ generators, but nevertheless $\mathrm{rk}(G) = \mathrm{rk}(G/\Phi(F))$. The group

$$K = \langle c^p, a_1, \dots, a_n \rangle,$$

which is unrelated to $H$, requires $n + 1$ generators, even modulo $\Phi(F)$.

## 3. Finite axiomatizability of the rank

In this section we establish Theorem 1.1. We begin with a basic example which illustrates the concept of an $\exists\forall\exists$-sentence in $\mathcal{L}_{\mathrm{gp}}$ and related constructions which we use frequently; compare with [Nies et al. 2021, Sections 2 and 5]. Despite its simplicity, the example is a key building block in later proofs, where we need to control the quantifier complexity of more involved first-order formulae.

**Example 3.1.** Let $G$ be a profinite group and let $N \subseteq G$. Suppose that $N$ is definable in $G$; this means that there is an $\mathcal{L}_{\mathrm{gp}}$-formula $\varphi(x)$, with a single free variable $x$, such that $N = \{g \in G \mid \varphi(g)\}$.

Let $B = \{b_1, \dots, b_n\}$ be a finite group of order $n$, with multiplication "table"

$$b_i b_j = b_{m(i,j)}$$

encoded by a suitable function $m\colon \{1, \dots, n\} \times \{1, \dots, n\} \to \{1, \dots, n\}$.

Then the sentence

$$\exists a_1, \dots a_n \, \forall x, y, z : \varphi(1) \wedge \big((\varphi(x) \wedge \varphi(y)) \to \varphi(x^{-1}y)\big) \wedge \big(\varphi(x) \to \varphi(y^{-1}xy)\big) \\ \wedge \Big( \bigwedge_{1 \le i < j \le n} \neg\varphi(a_i^{-1}a_j)\Big) \wedge \Big( \bigvee_{1 \le i \le n} \varphi(a_i^{-1}y)\Big) \wedge \Big( \bigwedge_{1 \le i, j \le n} \varphi(a_{m(i,j)}^{-1}a_i a_j)\Big)$$

can be used to express that $N \trianglelefteq G$ and $G/N \cong B$. The quantifier complexity of this sentence is the same as the quantifier complexity of $\varphi$ increased by $\exists\forall$. In particular, if $N \subseteq_{\mathrm{c}} G$ is $\exists$-definable as a closed set, i.e., definable by means of an $\exists$-formula which implicitly ensures that $N$ is topologically closed, we obtain an $\exists\forall\exists$-sentence to express that $N \trianglelefteq_{\mathrm{c}} G$ and $G/N \cong B$.

For instance, if we know or suspect that the commutator word has a certain finite width in $G$, we may consider the $\exists$-definable set

$$N = \big\{[x_1, y_1] \cdots [x_r, y_r] \mid x_1, y_1, \dots, x_r, y_r \in G\big\} \subseteq_{\mathrm{c}} G,$$

for a given parameter $r \in \mathbb{N}$, and formulate an $\exists\forall\exists$-sentence in $\mathcal{L}_{\mathrm{gp}}$ which expresses that, indeed, $N$ is equal to the entire commutator subgroup $[G, G]$ and that the abelianization $G/[G, G]$ is isomorphic to a given finite group.

Sometimes we want to express, by means of an $\mathcal{L}_{\mathrm{gp}}$-sentence, extra features of a definable subgroup $H \le_{\mathrm{c}} G$. This process typically involves quantification over elements of $H$ rather than $G$ which, in general, may increase the quantifier complexity of the resulting sentences. However, if $H = \{g \in G \mid \varphi(g)\}$ is $\exists$-definable, where $\varphi(x)$ takes the form $\exists \underline{z} : \varphi_0(x, \underline{z})$ with $\varphi_0$ quantifier-free in free variables $x$ and $z_1, \dots, z_m$, say, then $H$ is "quantifier-neutral" in the following sense. First-order assertions about $H$ can be translated into assertions of the same quantifier complexity about $G$, simply by expressing universal quantification over elements of $H$ as $\forall x, \underline{z} : (\varphi_0(x, \underline{z}) \to \cdots)$ and existential quantification over elements of $H$ as $\exists x, \underline{z} : (\varphi_0(x, \underline{z}) \wedge \cdots)$. $\diamond$

It is convenient to establish the assertions of Theorem 1.1 first for pronilpotent groups before considering the general situation.

**Proposition 3.2.** *Let $\pi$ be a finite set of primes, let $r \in \mathbb{N}$ and let $\mathbf{r} = (r_p)_{p \in \pi}$ be a tuple in $\{0, 1, \dots, r\}$. Then there exists an $\exists\forall\exists$-sentence $\omega_{\pi,r,\mathbf{r}}$ in $\mathcal{L}_{\mathrm{gp}}$ such that, for every pronilpotent pro-$\pi$ group $H$, the following are equivalent*:

(i) $\mathrm{rk}(H) = r$, *and* $\mathrm{rk}_p(H) = r_p$ *for every* $p \in \pi$.

(ii) $\omega_{\pi,r,\mathbf{r}}$ *holds in* $H$, *i.e.*, $H \models \omega_{\pi,r,\mathbf{r}}$.

*Proof.* We set $k = |\pi|$, write $\pi = \{p_1, \dots, p_k\}$ and put $q = q(\pi) = p_1 \cdots p_k$. As $H$ is pronilpotent, it is the direct product $H = \prod_{i=1}^{k} H_i$ of its Sylow pro-$p_i$ subgroups $H_i$. We set $m = m(r) = \lceil \log_2(r) \rceil + 1$.

Similar to Example 3.1, there is an $\exists\forall\exists$-sentence $\beta_1$ in $\mathcal{L}_{\mathrm{gp}}$ to express that there are elements $a_1, \dots, a_r$ in $H$ such that every element $h \in H$ can be written as $h = \prod_{j=1}^{r} a_j^{e_j} b$, for suitable choices for $e_j \in \{0, 1, \dots, q-1\}$ and

$$b \in B(H) = \Bigl\{[x_1, y_1] \cdots [x_r, y_r] z^q \mid x_1, y_1, \dots, x_r, y_r, z \in H\Bigr\} \subseteq_{\mathrm{c}} \Phi(H).$$

We recall that $\mathrm{d}(H) = \mathrm{d}(H/\Phi(H))$ and that $\mathrm{d}(H) \le r$ implies $B(H) = \Phi(H)$; see [Dixon et al. 1999, Lemma 1.23]. Thus $\beta_1$ holds for $H$ if and only if $\mathrm{d}(H) \le r$. Moreover, in this case $\Phi(H) = B(H)$ is $\exists$-definable in $H$ and hence quantifier-neutral in the sense of Example 3.1. By recursion, there is an $\exists\forall\exists$-sentence $\beta_{m+1}$ such that $\beta_{m+1}$ holds for $H$ if and only if

$$\mathrm{rk}(\Phi^j(H)/\Phi^{j+1}(H)) \le r \quad \text{for } 0 \le j \le m; \tag{3-1}$$

in this case the subgroup $F = \Phi^m(H)$ is $\exists$-definable in $H$ and hence quantifier-neutral, moreover it satisfies $\mathrm{d}(F) \le r$. Furthermore, there is an $\forall\exists$-sentence $\gamma$ which expresses that every Sylow subgroup of $F$ is powerful, viz. that $F$ is semipowerful in the terminology introduced in [Nies et al. 2021, Section 5]. Indeed, by [Dixon et al. 1999, Proposition 2.6], it suffices to express that every commutator $[x, y]$ of elements $x, y \in F$ is a $(2q)$-th power $z^{2q}$ of a suitable $z \in F$.

Once $F$ is $r$-generated and semipowerful, we know that $\mathrm{rk}(F) \le r$. If, in addition, the rank bounds specified in (3-1) hold, we deduce that $\mathrm{rk}(H/F) \le mr$ and hence $\mathrm{rk}(H) \le R$ for $R = (m+1)r$. Furthermore, the group

$$\Phi^{2R+1}(F) = \Bigl\{x^{q^{2R+1}} \mid x \in F\Bigr\}$$

is $\exists$-definable in $H$ and hence quantifier-neutral; in particular, $H/\Phi^{2R+1}(F)$ is interpretable in $H$. Finally, $|H/\Phi^{2R+1}(F)|$ is bounded by $q^{(2R+m+1)r}$ and there is an $\exists\forall\exists$-sentence $\theta$ which expresses that $H/\Phi^{2R+1}(F)$ is one of the finitely many finite $\pi$-groups of suitable order which has rank $r$ and whose $p$-ranks are in agreement with the prescribed $\mathbf{r}$; compare with Example 3.1.

With the backing of Theorem 2.1, we form the conjunction of the sentences $\beta_{m+1}, \gamma, \theta$ to arrive at an $\exists\forall\exists$-sentence $\omega_{\pi,r,\mathbf{r}}$ with the desired property. □

*Proof of Theorem 1.1.* We analyse the structure of a pro-$\pi$ group $G$ of rank $\mathrm{rk}(G) = r$ to build step-by-step a first-order sentence $\eta_{\pi,r}$ that is satisfied by any such group $G$. Following that we check that, conversely, every pro-$\pi$ group satisfying $\eta_{\pi,r}$ has rank at most $2r$. Applying Theorem 2.1, we extend $\eta_{\pi,r}$ to a sentence $\varrho_{\pi,r,\mathbf{r}}$ which pins down precisely the rank as being $r$ and the ranks of the Sylow subgroups as being given by $\mathbf{r}$.

Our discussion involves upper bounds for certain integer parameters that depend on $\pi$ and $r$, but not on the specific group $G$ used in our discussion; for short, we say that such parameters are $(\pi, r)$-bounded. The proof proceeds in four steps along the following plan of action. In Step 1 we produce a pronilpotent open normal subgroup $K \trianglelefteq_o G$ of $(\pi, r)$-bounded index. This is used in Step 2 to describe an $\exists$-definable pronilpotent open normal subgroup $H \trianglelefteq_o G$ of $(\pi, r)$-bounded index. In Step 3 we show that the fact that $H$ is pronilpotent can be expressed by an $\exists\forall\exists$-sentence. This uses a simple but effective trick: we would like to express that $H$ is a direct product of its Sylow subgroups, but in general the latter fail to be definable; to overcome this problem we work modulo the centre $\mathrm{Z}(H)$ which is sufficient for our purposes. In Step 4 we use the tools that we already prepared in Example 3.1 and in Proposition 3.2 to conclude the argument.

*Step 1.* The classification of finite simple groups implies that, up to isomorphism, there are only finitely many finite simple $\pi$-groups; see [Mazurov 1994, Remark following Lemma 2]. A fortiori there is a finite set

$$\mathcal{S} = \mathcal{S}_{\pi,r}$$

of representatives for the isomorphism classes of finite simple $\pi$-groups $S$ such that $\mathrm{rk}(S) \leq r$. Consequently, the cardinality of the set

$$\Psi = \Psi_{G,\pi,r} = \{\psi \mid \psi\colon G \to \mathrm{Aut}(S^l) \text{ a homomorphism for } S \in \mathcal{S} \text{ and } 0 \leq l \leq r\}$$

is $(\pi, r)$-bounded, because $G$ can be generated by at most $r$ elements and any homomorphism between groups is determined by its effect on a chosen set of generators. From this we observe that the index of

$$K = K_{G,\pi,r} = \bigcap_{\psi \in \Psi} \ker \psi \trianglelefteq_o G$$

in $G$ is $(\pi, r)$-bounded. Thus there exists $f(\pi, r) \in \mathbb{N}$, depending on $\pi$ and $r$, but not on the specific group $G$, such that $|G\colon K|$ divides $f(\pi, r)$.

We claim that $K$ is pronilpotent. For this it suffices to show that $K/(K \cap L)$ is nilpotent for each $L \trianglelefteq_o G$. Let $L \trianglelefteq_o G$. By pulling back a chief series for the finite group $G/L$ to $G$, we obtain a normal series

$$L = G_{n+1} \trianglelefteq G_n \trianglelefteq \cdots \trianglelefteq G_1 = G$$

of finite length $n$ such that, for each $i \in \{1, \ldots, n\}$, the group $G_i/G_{i+1}$ is a minimal normal subgroup of $G/G_{i+1}$ and thus isomorphic to $S_i^{m(i)}$ for suitable choices of $S_i \in \mathcal{S}$ and $m(i) \in \mathbb{N}$. Since each of the groups $S_i^{m(i)}$ contains an elementary abelian $p$-subgroup of rank $m(i)$, for primes $p$ dividing $|S_i|$, we obtain $m(i) \leq \mathrm{rk}(S_i^{m(i)}) \leq \mathrm{rk}(G) = r$ for all $i \in \{1, \ldots, n\}$. Intersecting with $K$, we obtain a series

$$K \cap L = K \cap G_{n+1} \trianglelefteq K \cap G_n \trianglelefteq \ldots \trianglelefteq K \cap G_1 = K \tag{3-2}$$

consisting of $G$-invariant subgroups with factors $(K \cap G_i)/(K \cap G_{i+1}) \cong S_i^{l(i)}$ satisfying $0 \leq l(i) \leq m(i) \leq r$, for $i \in \{1, \ldots, n\}$. By construction, $K$ acts trivially on each of these factors so that $[K \cap G_i, K] \subseteq K \cap G_{i+1}$ for $i \in \{1, \dots, n\}$. Thus (3-2) constitutes a central series for $K/(K \cap L)$, and $K/(K \cap L)$ is nilpotent (of class at most $n$).

*Step* 2. Next we consider the group

$$H = G^{f(\pi,r)} = \langle g^{f(\pi,r)} \mid g \in G \rangle \trianglelefteq_{\mathrm{o}} G \quad \text{with } H \subseteq K;$$

the index $|G : H|$ is $(\pi, r)$-bounded, by the positive solution to the restricted Burnside problem. In fact, we do not require the general result, but a rather special case, which is easy to establish. Indeed, assume for the moment that the pro-$\pi$ group $G$ of rank $r$ is finite of exponent $f(\pi, r)$. We need to show that $|G|$ is $(\pi, r)$-bounded. In Step 1 we established that $G$ has a nilpotent normal subgroup $K$ of $(\pi, r)$-bounded index. Thus there is no harm in assuming that $G = K$. Furthermore, $K$ is a direct product of its Sylow $p$-subgroups, where $p$ ranges over the finite set $\pi$. Hence we may even assume that $G$ is a $p$-group of rank at most $r$, for some $p \in \pi$, and that $f(\pi, r)$ is a $p$-power, $p^e$ say. In this situation, $G$ contains a powerful normal subgroup of $(p, r)$-bounded index (see [Dixon et al. 1999, Theorem 2.13]), and we may assume that $G$ itself is powerful. The $p$-power series of a powerful $p$-group coincides with its lower $p$-series, and we obtain the bound $|G| \leq p^{re}$.

Next we observe that the verbal subgroup $H$ is an $\exists$-definable subgroup of $G$ and hence quantifier-neutral, in the sense discussed in Example 3.1. Indeed, by [Nikolov and Segal 2011, Theorem 1], every element of $H$ can be written as a product of a $(\pi, r)$-bounded number of $f(\pi, r)$-th powers. But again we only require the bound in a rather special case which is much easier to handle. Indeed, descending without loss of generality to a subgroup of $(\pi, r)$-bounded index, as above, it suffices to recall that in a powerful pro-$p$ group every product of $p^e$-th powers is itself a $p^e$-th power; see [Dixon et al. 1999, Corollary 3.5].

*Step* 3. Since $K$ is pronilpotent, so is $H$. In the situation at hand, this fact can be expressed by an $\exists\forall\exists$-sentence. Indeed, $H$ is pronilpotent if and only if $H/\mathrm{Z}(H)$ is pronilpotent. Hence it suffices to express the assertion that $H/\mathrm{Z}(H)$ is pronilpotent. Clearly, $\mathrm{Z}(H)$ is $\forall$-definable in $H$ and hence in $G$. We set $k = |\pi|$ and write $\pi = \{p_1, \dots, p_k\}$. As $H$ is pronilpotent, $H = \prod_{i=1}^{k} H_i$ is the direct product of its Sylow pro-$p_i$ subgroups $H_i$ and $\mathrm{Z}(H) = \prod_{i=1}^{k} \mathrm{Z}(H_i)$ so that $H/\mathrm{Z}(H) \cong \prod_{i=1}^{k} H_i/\mathrm{Z}(H_i)$. From

$$C_i = \mathrm{C}_H(H_i) = \prod_{j=1}^{i-1} H_j \times \mathrm{Z}(H_i) \times \prod_{j=i+1}^{k} H_j, \quad \text{for } i \in \{1, \dots, k\},$$

we deduce that

$$D_i = \bigcap \{C_j \mid 1 \leq j \leq k \text{ and } j \neq i\} = \prod_{j=1}^{i-1} \mathrm{Z}(H_j) \times H_i \times \prod_{j=i+1}^{k} \mathrm{Z}(H_j)$$

and thus

$$D_i/\mathrm{Z}(H) \cong H_i/\mathrm{Z}(H_i), \quad \text{for } i \in \{1, \dots, k\}.$$

As $\mathrm{rk}(G) \le r$, there exist, for each $i \in \{1, \dots, k\}$, elements $x_{i,1}, \dots, x_{i,r} \in H_i$ such that $H_i = \langle x_{i,1}, \dots, x_{i,r} \rangle$ and thus

$$C_i = \mathrm{C}_H(\{x_{i,1}, \dots, x_{i,r}\}).$$

Subject to the $kr$ parameters $x_{1,1}, \dots, x_{k,r}$, this makes $\mathrm{Z}(H) = \bigcap_{i=1}^{k} C_i$ and each of the groups $D_i$ quantifier-free definable, by suitable centralizer conditions; moreover $Q_i = D_i / \mathrm{Z}(H)$ becomes interpretable in $H$, for $1 \le i \le k$.

We conclude that it suffices to express in an $\forall\exists$-sentence, subject to the $(\pi, r)$-bounded number of parameters $x_{s,t}$, that

(a) $\bigcap_{i=1}^{k} C_i = \mathrm{Z}(H)$, hence $\mathrm{Z}(H) \subseteq D_i$, for $i \in \{1, \dots k\}$;

(b) $D_i / \mathrm{Z}(H)$ is a pro-$p_i$ group for $i \in \{1, \dots k\}$;

(c) $[D_i, D_j] \subseteq \mathrm{Z}(H)$ for $i, j \in \{1, \dots, k\}$ with $i \ne j$;

(d) $H = D_1 \cdot D_2 \cdot \ldots \cdot D_k$, where the right-hand side denotes the set of all products $y_1 \cdots y_k$ with factors $y_i \in D_i$ for $i \in \{1, \dots, k\}$;

for this implies that $H / \mathrm{Z}(H) = \prod_{i=1}^{k} D_i / \mathrm{Z}(H)$ is the direct product of its Sylow subgroups and thus pronilpotent. Turning the parameters $x_{s,t}$ into variables bound by an extra existential quantifier at the front, we arrive at an $\exists\forall\exists$-sentence without parameters which verifies that $H$ is pronilpotent.

Subject to the parameters $x_{s,t}$, the assertions in (a), (c) can be expressed by an $\forall$-sentence, and (d) can be achieved by means of an $\forall\exists$-sentence. The only tricky part occurs in (b) where we need to express that the group $Q_i = D_i / \mathrm{Z}(H)$ is a pro-$p_i$ group. Since we know a priori that $Q_i$ is a pro-$\pi$ group, this is achieved by demanding that every element of $Q_i$ is a $q_i$-th power, for $q_i = p_1 \cdots p_{i-1} p_{i+1} \cdots p_k$. This can be expressed by an $\forall\exists$-sentence at the level of $H$, because $\mathrm{Z}(H) = \bigcap_{i=1}^{k} C_i$ is quantifier-free definable subject to the parameters $x_{s,t}$.

*Step* 4. By Step 2, the group $G/H$ is interpretable in $G$ and finite of $(\pi, r)$-bounded order. There is an $\exists\forall\exists$-sentence that expresses that the factor group $G/H$ is among the finitely many finite groups of rank at most $r$ and exponent dividing $f(\pi, r)$; compare with Example 3.1. Using our results from Step 2, Step 3 and Proposition 3.2, we produce an $\exists\forall\exists$-sentence that expresses that the power word $x^{f(\pi,r)}$ has $(\pi, r)$-bounded width in $G$ and that $H = G^{f(\pi,r)}$ is pronilpotent of rank at most $r$.

The conjunction of these two sentences yields an $\exists\forall\exists$-sentence $\eta_{\pi,r}$ such that

- every pro-$\pi$ group $G$ of rank $\mathrm{rk}(G) = r$ satisfies $\eta_{\pi,r}$;
- conversely, if a pro-$\pi$ group $\widetilde{G}$ satisfies $\eta_{\pi,r}$, then $\widetilde{H} = \widetilde{G}^{f(\pi,r)} \trianglelefteq_{\mathrm{o}} \widetilde{G}$ is pronilpotent and both $\widetilde{H}$ and $\widetilde{G}/\widetilde{H}$ have rank at most $r$; in particular, this ensures that $\mathrm{rk}(\widetilde{G}) \le R$ for $R = 2r$.

We put $m = m(R) = \lceil \log_2(R) \rceil + 1$. As in the proof of Proposition 3.2 we see that $F = \Phi^{m(R)}(H) \trianglelefteq_o G$ is $\exists$-definable, hence quantifier-neutral, and semipowerful. Furthermore, $\Phi^{2R+1}(F)$ is $\exists$-definable, hence quantifier-neutral, and, by Theorem 2.1,

$$\mathrm{rk}(G) = \mathrm{rk}\big(G/\Phi^{2R+1}(F)\big) \quad \text{and} \quad \mathrm{rk}_p(G) = \mathrm{rk}_p\big(G/\Phi^{2R+1}(F)\big) \quad \text{for every } p \in \pi.$$

Just as in the proof of Proposition 3.2 we find an $\exists\forall\exists$-sentence which in conjunction with $\eta_{\pi,r}$ produces an $\exists\forall\exists$-sentence $\varrho_{\pi,r,\mathbf{r}}$ with the desired property. □

The next result complements Theorem 1.1. It illustrates that the rank of a pro-$p$ group cannot be detected by a first-order sentence uniformly across all primes $p$, even if the language $\mathcal{L}_{\mathrm{gp}}$ was to be enlarged by an extra function to be interpreted as the $p$-power map $x \mapsto x^p$ in pro-$p$ groups. (Note that regarding elementary abelian $p$-groups it is futile to enlarge the language in this way.) We sketch a proof for completeness; it relies on a standard ultraproduct construction and a well-known quantifier elimination result in model theory.

**Proposition 3.3.** *Let $\tilde{\pi}$ be an infinite set of primes and let $r \in \mathbb{N}$. Then there is no $\mathcal{L}_{\mathrm{gp}}$-sentence $\vartheta_{\tilde{\pi},r}$ such that, for every $p \in \tilde{\pi}$ and every finite elementary abelian $p$-group $G$, the following are equivalent*:

(i) $\mathrm{rk}(G) = r$.

(ii) *$\vartheta_{\tilde{\pi},r}$ holds in $G$, i.e., $G \models \vartheta_{\tilde{\pi},r}$.*

*Proof.* For a contradiction, assume that the $\mathcal{L}_{\mathrm{gp}}$-sentence $\vartheta = \vartheta_{\tilde{\pi},r}$ has the desired property. Then $C_p^r \models \vartheta$ and $C_p^{r+1} \models \neg\vartheta$ for all $p \in \tilde{\pi}$. We regard $C_p^r$ and $C_p^{r+1}$ as the additive groups of the vector spaces $\mathbb{F}_p^r$ and $\mathbb{F}_p^{r+1}$ over the prime field $\mathbb{F}_p$.

Let $\mathfrak{U}$ be a nonprincipal ultrafilter on the infinite index set $\tilde{\pi}$. By Łoś's theorem,

$$\mathcal{K} = \big(\textstyle\prod_{p \in \tilde{\pi}} \mathbb{F}_p\big)/\sim_{\mathfrak{U}}$$

is a field of characteristic 0, and

$$\mathcal{V} = \big(\textstyle\prod_{p \in \tilde{\pi}} \mathbb{F}_p^r\big)/\sim_{\mathfrak{U}} \quad \text{and} \quad \mathcal{W} = \big(\textstyle\prod_{p \in \tilde{\pi}} \mathbb{F}_p^{r+1}\big)/\sim_{\mathfrak{U}}$$

are nonzero $\mathcal{K}$-vector spaces. Let $\mathcal{L}_{\mathcal{K}\text{-vs}}$ denote the language of $\mathcal{K}$-vector spaces, which comprises the language of groups (for the additive group of vectors) and, for each scalar $c \in \mathcal{K}$, a 1-ary operation $f_c$ (to denote scalar multiplication by $c$). Clearly, the $\mathcal{L}_{\mathrm{gp}}$-sentence $\vartheta$ gives rise to an $\mathcal{L}_{\mathcal{K}\text{-vs}}$-sentence $\theta$, not involving scalar multiplication at all, such that by Łoś's theorem

$$\mathcal{V} \models \theta \quad \text{and} \quad \mathcal{W} \models \neg\theta,$$

in contradiction to the known fact that the infinite $\mathcal{K}$-vector spaces $\mathcal{V}$ and $\mathcal{W}$ have the same theory, due to quantifier elimination; see [Hodges 1993, Section 8.4]. □

## 4. Finite axiomatizability of the dimension

In this section we establish Theorems 1.4 and 1.2. We derive the former from a result of Héthelyi and Lévai [2003] about finite powerful $p$-groups; compare with [Wilson 2002; Fernández-Alcober 2007]. We recall from [Dixon et al. 1999, Theorem 4.20] that the elements of finite order in a finitely generated powerful pro-$p$ group form a powerful finite subgroup, its torsion subgroup.

*Proof of Theorem 1.4.* The torsion subgroup $T$ is finite and characteristic in $G$ so that $\mathrm{C}_G(T) \trianglelefteq_{\mathrm{o}} G$. We choose a uniformly powerful open normal subgroup $U \trianglelefteq_{\mathrm{o}} G$ such that $U \subseteq \mathrm{C}_G(T)$ and $U \subseteq \Phi(G)$. Since $U$ is torsion-free, this implies that

$$N = U \times T \trianglelefteq_{\mathrm{o}} G \quad \text{and} \quad \mathrm{d}(G) = \mathrm{d}(G/U).$$

We show below that there exists $k \in \mathbb{N}$ such that $U^{p^k} = \Phi^k(U) \trianglelefteq_{\mathrm{o}} G$ satisfies

$$\Omega_{\{1\}}(G/U^{p^k}) = \Omega_{\{1\}}(N/U^{p^k}). \tag{4-1}$$

Since $N/U^{p^k} \cong U/U^{p^k} \times T$ and because $U$ is uniformly powerful, $\Omega_{\{1\}}(N/U^{p^k})$ is in bijection with the cartesian product of sets

$$\Omega_{\{1\}}(U/U^{p^k}) \times \Omega_{\{1\}}(T) = U^{p^{k-1}}/U^{p^k} \times \Omega_{\{1\}}(G)$$

and furthermore $\log_p|U^{p^{k-1}}/U^{p^k}| = \mathrm{d}(U)$. Put $s(G) = \log_p|\Omega_{\{1\}}(G)|$. Stringing all pieces together, we see that the finite powerful $p$-group $P = G/U^{p^k}$ satisfies

$$\log_p|\Omega_{\{1\}}(P)| = \mathrm{d}(U) + s(G) = \dim(G) + s(G).$$

The theorem of Héthelyi and Lévai [2003] yields $\log_p|\Omega_{\{1\}}(P)| = \mathrm{d}(P)$ and $s(G) = \log_p|\Omega_{\{1\}}(T)| = \mathrm{d}(T)$ so that

$$\dim(G) = \log_p|\Omega_{\{1\}}(P)| - s(G) = \mathrm{d}(P) - s(G) = \mathrm{d}(G) - s(G) = \mathrm{d}(G) - \mathrm{d}(T).$$

It remains to establish (4-1). Since $U^{p^k}$, $k \in \mathbb{N}$, is a base for the neighbourhoods of 1 in $G$, it suffices to show that there exists an open normal subgroup $W \trianglelefteq_{\mathrm{o}} G$ such that for every $x \in G \smallsetminus N \subseteq_{\mathrm{c}} G$ we have $x^p \notin W$, or in other words $x^p \not\equiv_W 1$. From $T \subseteq N$ we see that $G \smallsetminus N$ does not contain any elements of finite order. Hence for every $x \in G \smallsetminus N$ there exists $W_x \trianglelefteq_{\mathrm{o}} G$ such that $x^p \not\equiv_{W_x} 1$, and consequently $y^p \not\equiv_{W_x} 1$ for all $y \in xW_x \subseteq_{\mathrm{o}} G$. Since $G \smallsetminus N$ is compact, it is covered by a finite union of such cosets $xW_x$, i.e., $G \smallsetminus N \subseteq \bigcup_{x \in X} xW_x$ with $|X| < \infty$. This implies that $W = \bigcap_{x \in X} W_x \trianglelefteq_{\mathrm{o}} G$ has the required property. □

*Proof of Theorem 1.2.* Let $p \in \pi$ and put $d = d_p$. It suffices to explain how one can build an $\exists\forall\exists$-sentence $\tau_{\pi,r,p,d}$ in $\mathcal{L}_{\mathrm{gp}}$ which expresses that a pro-$\pi$ group $G$ of rank $\mathrm{rk}(G) = r$ has Sylow pro-$p$ subgroup dimension $d$. As in the proof of Theorem 1.1 we work with a general pro-$\pi$ group $G$ with $\mathrm{rk}(G) = r$ to concoct $\tau_{\pi,r,p,d}$.

Using the same approach as in the proof of Theorem 1.1, we find an $\exists$-definable and hence quantifier-neutral subgroup $H \trianglelefteq_{\mathrm{o}} G$ that is pronilpotent and has $(\pi, r)$-bounded index in $G$; moreover the arrangement can be expressed by means of a suitable $\exists\forall\exists$-sentence. We put $m = m(r) = \lceil \log_2(r) \rceil + 1$. In the proof of Proposition 3.2 we saw that we can use an $\exists\forall\exists$-sentence to describe that $\Phi^m(H)$ is semipowerful and of $(\pi, r)$-bounded index in $H$; in parallel we can realize $\Phi^m(H)$ as an $\exists$-definable and hence quantifier-neutral subgroup. The Sylow subgroup dimensions do not change if we pass from $G$ to an open subgroup. Replacing $G$ by $\Phi^m(H)$, we may therefore assume without loss of generality that $G$ itself is pronilpotent and semipowerful.

As $G$ is pronilpotent, $G$ is the direct product of its powerful Sylow subgroups; let $G_p$ denote the Sylow pro-$p$ subgroup and $T_p$ its torsion subgroup. By Theorem 1.4 it suffices to produce an $\exists\forall\exists$-sentence which pins down within the finite range $\{0, 1, \ldots, r\}$ the invariants

$$\mathrm{d}(G_p) = \log_p |G_p : \Phi(G_p)| \quad \text{and} \quad \mathrm{d}(T_p) = \log_p |\Omega_{\{1\}}(G_p)|,$$

where $\Omega_{\{1\}}(G_p) = \{g \in G_p \mid g^p = 1\}$ is the set of all elements of order 1 or $p$. We observe that $G_p/\Phi(G_p) \cong G/\Phi_p(G)$ is essentially the $p$-Frattini quotient of $G$ and that $\Omega_{\{1\}}(G_p) = \{g \in G \mid g^p = 1\}$.

The Frattini quotient $G/\Phi(G)$ has $(\pi, r)$-bounded order and maps onto the $p$-Frattini quotient $G/\Phi_p(G)$. As in the proof of Proposition 3.2, the group $G/\Phi(G)$ is interpretable in $G$. There is an $\exists\forall\exists$-sentence which detects any prescribed isomorphism type of $G/\Phi(G)$ among a $(\pi, r)$-bounded number of possibilities; compare with Example 3.1. Forming a suitable disjunction, we can also detect the isomorphism type of the $p$-Frattini quotient $G/\Phi_p(G)$ and hence the minimal numbers of generators $\mathrm{d}(G_p)$.

Clearly, the closed subset $\{g \in G \mid g^p = 1\} \subseteq_{\mathrm{c}} G$ is quantifier-free definable in $G$. Moreover, its size equals $p^{\mathrm{d}(T_p)}$ and is thus at most $p^r$. We can easily identify by means of an $\exists\forall$-sentence its precise size and hence the invariant $\mathrm{d}(T_p)$. $\square$

## Acknowledgements

The results form part of the first author's doctoral thesis "On some algebraic properties of $p$-adic analytic pro-$p$ groups" (to become available via the university library depository at `https://docserv.uni-duesseldorf.de/`). We thank Jon González-Sánchez for drawing our attention to structural results about finite powerful $p$-groups which led to Theorem 1.4 and thus helped us to streamline our proof of Theorem 1.2. We thank Immanuel Halupczok for valuable conversations on model-theoretic background; in particular, these led us to formulate Proposition 3.3. The referees' feedback led to improvements in the presentation.

## References

[Dixon et al. 1999] J. D. Dixon, M. P. F. du Sautoy, A. Mann, and D. Segal, *Analytic pro-$p$ groups*, 2nd ed., Cambridge Studies in Advanced Math. **61**, Cambridge University Press, 1999. MR Zbl

[Feferman and Vaught 1959] S. Feferman and R. L. Vaught, "The first order properties of products of algebraic systems", *Fund. Math.* **47** (1959), 57–103. MR Zbl

[Fernández-Alcober 2007] G. A. Fernández-Alcober, "Omega subgroups of powerful $p$-groups", *Israel J. Math.* **162** (2007), 75–79. MR Zbl

[Héthelyi and Lévai 2003] L. Héthelyi and L. Lévai, "On elements of order $p$ in powerful $p$-groups", *J. Algebra* **270**:1 (2003), 1–6. MR Zbl

[Hodges 1993] W. Hodges, *Model theory*, Encyclopedia of Mathematics and its Applications **42**, Cambridge University Press, 1993. MR Zbl

[Jarden and Lubotzky 2008] M. Jarden and A. Lubotzky, "Elementary equivalence of profinite groups", *Bull. Lond. Math. Soc.* **40**:5 (2008), 887–896. MR Zbl

[Klopsch 2011] B. Klopsch, "An introduction to compact $p$-adic Lie groups", pp. 7–61 in *Lectures on profinite topics in group theory*, edited by D. Segal, London Math. Soc. Stud. Texts **77**, Cambridge University Press, 2011. MR Zbl

[Lucchini 1989] A. Lucchini, "A bound on the number of generators of a finite group", *Arch. Math.* (*Basel*) **53**:4 (1989), 313–317. MR Zbl

[Lucchini 1997] A. Lucchini, "A bound on the presentation rank of a finite group", *Bull. London Math. Soc.* **29**:4 (1997), 389–394. MR Zbl

[Macpherson and Tent 2016] D. Macpherson and K. Tent, "Profinite groups with NIP theory and $p$-adic analytic groups", *Bull. Lond. Math. Soc.* **48**:6 (2016), 1037–1049. MR Zbl

[Mazurov 1994] V. D. Mazurov, "On the set of orders of elements of a finite group", *Algebra i Logika* **33**:1 (1994), 81–89. MR Zbl

[Nies et al. 2021] A. Nies, D. Segal, and K. Tent, "Finite axiomatizability for profinite groups", *Proc. Lond. Math. Soc.* (3) **123**:6 (2021), 597–635. MR Zbl

[Nikolov and Segal 2007] N. Nikolov and D. Segal, "On finitely generated profinite groups, I: strong completeness and uniform bounds", *Ann. of Math.* (2) **165**:1 (2007), 171–238. MR Zbl

[Nikolov and Segal 2011] N. Nikolov and D. Segal, "Powers in finite groups", *Groups Geom. Dyn.* **5**:2 (2011), 501–507. MR Zbl

[Prest 1988] M. Prest, *Model theory and modules*, London Mathematical Society Lecture Note Series **130**, Cambridge University Press, 1988. MR Zbl

[Ribes and Zalesskii 2010] L. Ribes and P. Zalesskii, *Profinite groups*, 2nd ed., Ergebnisse der Math. (3) **40**, Springer, 2010. MR Zbl

[Wilson 2002] L. Wilson, "On the power structure of powerful $p$-groups", *J. Group Theory* **5**:2 (2002), 129–144. MR Zbl

Received April 14, 2023. Revised February 6, 2024.

MARTINA CONTE
conte@uni-duesseldorf.de

BENJAMIN KLOPSCH
klopsch@math.uni-duesseldorf.de

(*all authors*)
HEINRICH-HEINE-UNIVERSITÄT
DÜSSELDORF
GERMANY

PACIFIC JOURNAL OF MATHEMATICS
Vol. 328, No. 2, 2024
https://doi.org/10.2140/pjm.2024.328.275

# ELLIPTIC GENUS AND STRING COBORDISM AT DIMENSION 24

FEI HAN AND RUIZHI HUANG

**It is known that spin cobordism can be determined by Stiefel–Whitney numbers and index theoretic invariants, namely KO-theoretic Pontryagin numbers. We show that string cobordism at dimension 24 can be determined by elliptic genus, a higher index theoretic invariant. We also compute the image of 24-dimensional string cobordism under elliptic genus. Using our results, we show that under certain curvature conditions, a compact 24-dimensional string manifold must bound a string manifold.**

## 1. Introduction

Cobordism is a fundamental tool in geometry and topology. For the oriented cobordism ring $\Omega_*^{\mathrm{SO}}$, there are spin cobordism $\Omega_*^{\mathrm{Spin}}$ and string cobordism $\Omega_*^{\mathrm{String}}$ as refinements through the Whitehead tower

$$\cdots \to \mathrm{String} \to \mathrm{Spin} \to \mathrm{SO}.$$

It is a classical problem to classify cobordism classes in terms of characteristic numbers. Historically, Wall [1960] showed that two closed oriented manifolds are oriented cobordant if and only if they have the same Stiefel–Whitney numbers and Pontryagin numbers. Anderson, Brown and Peterson [Anderson et al. 1967] showed that two closed spin manifolds are spin cobordant if and only if they have the same Stiefel–Whitney numbers and KO-theoretic characteristic numbers.

The problem for string manifolds is much more complicated. To our best knowledge, it is unknown yet which set of characteristic numbers classifies string cobordism. It is expected that TMF-theoretic characteristic numbers will play a similar role for string cobordism as KO-theoretic characteristic numbers do for spin cobordism. Here TMF stands for the topological modular form developed by Hopkins and Miller [Hopkins 2002]. The Witten genus [1987; 1988] plays a similar role in TMF as the $\widehat{A}$-genus does in KO and is refined to be the $\sigma$-orientation from the Thom spectrum of string cobordism to the spectrum TMF [Hopkins 2002; Ando et al. 2001].

*MSC2020:* primary 57R20, 58J26; secondary 53C21, 57R90.
*Keywords:* string cobordism, elliptic genus, string manifold, nonpositive Ricci curvature, Witten genus.

In this paper we show that the elliptic genus [Ochanine 1987], a higher index theoretic invariant, determines 24-dimensional string cobordism. As elliptic genus is a twisted Witten genus [Witten 1987; 1988; Liu 1995a; 1995b], it can be viewed as sort of TMF-theoretic characteristic numbers. This coincides with the expectation of the role that TMF-theoretic characteristic numbers should play for string cobordism. In the paper, we also compute the image of 24-dimensional string cobordism under elliptic genus as well as give some application of our results in geometry. It is worthwhile to remark that 24 is a dimension of special interest for string geometry. For instance, in this dimension, one has (see [Hirzebruch et al. 1994, pp. 85–87])

$$W(M) = \widehat{A}(M)\bar{\Delta} + \widehat{A}(M, T)\Delta,$$

where $W(M)$ is the Witten genus of $M$, $\widehat{A}(M)$ is the A-hat genus and $\widehat{A}(M, T)$ is the tangent bundle twisted A-hat genus of $M$, $\bar{\Delta} = E_4^3 - 744 \cdot \Delta$ with $E_4$ being the Eisenstein series of weight 4 and $\Delta$ being the modular discriminant of weight 12. Hirzebruch raised his prize question in [Hirzebruch et al. 1994] that whether there exists a 24-dimensional compact string manifold $M$ such that $W(M) = \bar{\Delta}$ (or equivalently $\widehat{A}(M) = 1$, $\widehat{A}(M, T) = 0$) and the Monster group acts on $M$ as self-diffeomorphisms. The existence of such a manifold was confirmed by Mahowald and Hopkins [2002]. They determined the image of Witten genus at this dimension via TMF. Based on their work, we [Han and Huang 2022] realized the kernel of Witten genus at dimension 24 and determined an integral basis of $\Omega_{24}^{\mathrm{String}}$. As applications, various Rokhlin type divisibility theorems were proved there, which significantly extend earlier relevant results of Chen and Han [2015] and Chen, Han and Zhang [Chen et al. 2012]. Additionally, Milivojević [2021] used rational homotopy theory to give a weak form solution to the Hirzebruch's prize question. However, the part of the question concerning the Monster group action is still open.

The *elliptic genus*, which was first constructed by Ochanine [1987] and Landweber and Stong [1988], is a graded ring homomorphism

$$\phi : \Omega_*^{\mathrm{SO}} \to \mathbb{Z}\left[\tfrac{1}{2}\right][\delta, \varepsilon] \tag{1-1}$$

from the oriented cobordism ring to the graded polynomial ring $\mathbb{Z}\left[\tfrac{1}{2}\right][\delta, \varepsilon]$ with the degrees $|\delta| = 4$, $|\varepsilon| = 8$, such that the logarithm is given by the formal integral

$$g(z) = \int_0^z \frac{dt}{\sqrt{1 - 2\delta t^2 + \varepsilon t^4}}. \tag{1-2}$$

The background and the developments of the theory of elliptic genus can be found in [Landweber 1988b; Segal 1988; Kreck and Stolz 1993; Hirzebruch et al. 1994; Liu 1996a; Hopkins 2002; Witten 1988].

It is shown [Chudnovsky et al. 1988; Landweber 1988a] that the image of the elliptic genus is

$$\phi(\Omega_*^{\mathrm{SO}}) = \mathbb{Z}[\delta, 2\gamma, 2\gamma^2, \ldots, 2\gamma^{2^s}, \ldots], \tag{1-3}$$

where $\gamma = \frac{1}{4}(\delta^2 - \varepsilon)$; and when restricted to spin cobordism,

$$\phi(\Omega_*^{\mathrm{Spin}}) = \mathbb{Z}[16\delta, (8\delta)^2, \varepsilon]. \tag{1-4}$$

It follows that at dimension 24 the image is spanned over $\mathbb{Z}$ by

$$(8\delta)^6, \quad (8\delta)^4\varepsilon, \quad (8\delta)^2\varepsilon^2, \quad \varepsilon^3.$$

The map $\phi : \Omega_{24}^{\mathrm{Spin}} \to \mathbb{Z}[8\delta, \varepsilon]$ has nontrivial kernel. Actually $E - F \cdot B$ is in the kernel, where $E$ is the total space of a fiber bundle of compact and connected structure group with $F$ being spin manifold as fiber and $B$ being the base. This comes from the multiplicativity of elliptic genus [Ochanine 1988], which is equivalent to the Witten–Bott–Taubes–Liu rigidity [Bott and Taubes 1989; Taubes 1989; Liu 1996b].

Our main result is stated as follows.

**Theorem 1.** *The elliptic genus*

$$\phi : \Omega_{24}^{\mathrm{String}} \to \mathbb{Z}[8\delta, \varepsilon]$$

*is injective and its image is a subgroup of* $\mathbb{Z}[8\delta, \varepsilon]$ *spanned by*

$$(8\delta)^6, \qquad 24(8\delta)^4\varepsilon, \qquad (8\delta)^2\varepsilon^2, \qquad 8\varepsilon^3.$$

The theorem shows us the following picture:

$$\begin{aligned}\Omega_{24}^{\mathrm{String}} \cong \phi(\Omega_{24}^{\mathrm{String}}) &\cong \mathbb{Z} \oplus 24\mathbb{Z} \oplus \mathbb{Z} \oplus 8\mathbb{Z} \\ &\leq \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \cong \phi(\Omega_{24}^{\mathrm{Spin}}).\end{aligned}$$

In particular, it supports the expectation that TMF-theoretic characteristic numbers will play a similar role for string cobordism as KO-theoretic characteristic numbers do for spin cobordism.

The key to the proof of Theorem 1 is a result in [Han and Huang 2022], where we determine an integral basis of $\Omega_{24}^{\mathrm{String}}$, which consists of two explicitly constructed manifolds in the kernel of the Witten genus, and another two-manifolds constructed by Mahowald and Hopkins [2002] determining the image of the Witten genus. Then we can apply two concrete elliptic genera (2-9) to reduce the computations of the elliptic genus to those of classical twisted and untwisted genera on the generators of $\Omega_{24}^{\mathrm{String}}$. The details are carried out in Section 3.

Theorem 1 has interesting application in geometry. A closed manifold $M$ is called almost flat if for any $\varepsilon > 0$, there is a Riemannian metric $g_\varepsilon$ on $M$ such that the diameter $\operatorname{diam}(M, g_\varepsilon) \leq 1$ and $g_\varepsilon$ is $\varepsilon$-flat, i.e., for the sectional curvature $K_{g_\varepsilon}$, we have $|K_{g_\varepsilon}| < \varepsilon$. Given $n$, there is a positive number $\varepsilon_n > 0$ such that if an $n$-dimensional manifold admits an $\varepsilon_n$-flat metric with diameter $\leq 1$, then it is almost flat. The classical result of Gromov [1978] says that every almost flat manifold is finitely covered by a nilmanifold, and this was refined by Ruh [1982] by proving that an almost flat manifold is diffeomorphic to an infranilmanifold. It has been conjectured by Farrell and Zdravkovska [1983] and independently by Yau [1993] that every almost flat manifold is the boundary of a closed manifold. Davis and Fang [2016] showed that this conjecture holds under the assumption that the 2-Sylow subgroup of holonomy group is cyclic or generalized quaternionic. The general case of the conjecture remains open. Davis and Fang [2016] also pointed out that it is a difficult question whether every almost flat spin manifold (up to changing spin structures) bounds a spin manifold.

By Chern–Weil theory, it can be shown that the Pontryagin numbers of an oriented almost flat manifold $M$ all vanish [Davis and Fang 2016]. Since the elliptic genus is determined by Pontryagin numbers, one can see from Theorem 1 that every 24-dimensional almost flat string manifold bounds a string manifold.

In [Chen and Han 2024], vanishing results for elliptic genus were proven under almost nonpositive Ricci curvature condition. Theorem 1.3 there shows that given $n \in \mathbb{N}$ and positive number $\lambda$, there exists some $\varepsilon = \varepsilon(n, \lambda) > 0$ such that if a compact $4n$-dimensional spin Riemannian manifold $(M, g)$ satisfies $\operatorname{diam}(M, g) \leq 1$, $\operatorname{Ric}(g) \leq \varepsilon$, sectional curvature $\geq -\lambda$ and has infinite isometry group, then the elliptic genus of $M$ vanishes. Combining with Theorem 1, we obtain:

**Corollary 2.** *Given positive number $\lambda$, there exists some $\varepsilon = \varepsilon(\lambda) > 0$ such that if a compact 24-dimensional string Riemannian manifold $(M, g)$ satisfies* $\operatorname{diam}(M, g) \leq 1$, $\operatorname{Ric}(g) \leq \varepsilon$, *sectional curvature $\geq -\lambda$ and has infinite isometry group, then $M$ bounds a string manifold.* □

## 2. Preliminaries

In this section we collect some necessary knowledge of elliptic genus used in the sequel. Details can be found in [Hirzebruch et al. 1994; Liu 1992; 1995a].

Let $f$ be the formal inverse function of the logarithm $g$ in (1-2). Then $Y = f'$, $X = f$ solve the Jacobi quadrics

$$Y^2 = 1 - 2\delta \cdot X^2 + \varepsilon X^4. \tag{2-1}$$

For concrete values of $\delta$ and $\varepsilon$, a solution $f$ gives an elliptic genus with logarithm $g$. For instance, when $\delta = \varepsilon = 1$, $f(z) = \tanh z$ and $\phi$ reduces to the $L$-genus or the signature, and when $\delta = -\frac{1}{8}$, $\varepsilon = 0$, $f(z) = 2\sinh\frac{z}{2}$ and $\phi$ reduces to the $\widehat{A}$-genus.

Recall that the four Jacobi theta-functions (see [Chandrasekharan 1985]) defined by infinite multiplications are

$$\theta(v,\tau)=2q^{1/8}\sin(\pi v)\prod_{j=1}^{\infty}\big[(1-q^j)(1-e^{2\pi\sqrt{-1}v}q^j)(1-e^{-2\pi\sqrt{-1}v}q^j)\big],$$
$$\theta_1(v,\tau)=2q^{1/8}\cos(\pi v)\prod_{j=1}^{\infty}\big[(1-q^j)(1+e^{2\pi\sqrt{-1}v}q^j)(1+e^{-2\pi\sqrt{-1}v}q^j)\big],$$
$$\theta_2(v,\tau)=\prod_{j=1}^{\infty}\big[(1-q^j)(1-e^{2\pi\sqrt{-1}v}q^{j-1/2})(1-e^{-2\pi\sqrt{-1}v}q^{j-1/2})\big],$$
$$\theta_3(v,\tau)=\prod_{j=1}^{\infty}\big[(1-q^j)(1+e^{2\pi\sqrt{-1}v}q^{j-1/2})(1+e^{-2\pi\sqrt{-1}v}q^{j-1/2})\big],$$

where $q=e^{2\pi\sqrt{-1}\tau}$. They are holomorphic functions for $(v,\tau)\in\mathbb{C}\times\mathbb{H}$, where $\mathbb{C}$ is the complex plane and $\mathbb{H}$ is the upper half plane. Write $\theta_j=\theta_j(0,\tau),\ 1\le j\le 3$, and $\theta'(0,\tau)=\frac{\partial}{\partial v}\theta(v,\tau)\big|_{v=0}$.

When

$$\text{(2-2)}\qquad \begin{aligned}\delta=\delta_1(\tau)&=\frac{1}{8}(\theta_2^4+\theta_3^4)=\frac{1}{4}+6\sum_{n=1}^{\infty}\sum_{\substack{d|n\\ d\text{ odd}}}dq^n=\frac{1}{4}+6q+6q^2+\cdots,\\ \varepsilon=\varepsilon_1(\tau)&=\frac{1}{16}\theta_2^4\theta_3^4=\frac{1}{16}+\sum_{n=1}^{\infty}\sum_{d|n}(-1)^d d^3q^n=\frac{1}{16}-q+7q^2+\cdots,\end{aligned}$$

equation (2-1) has the solution

$$f_1(z,\tau)=2\pi\sqrt{-1}\frac{\theta(z,\tau)}{\theta'(0,\tau)}\frac{\theta_1(0,\tau)}{\theta_1(z,\tau)}.$$

Similarly, when

$$\text{(2-3)}\qquad \begin{aligned}\delta=\delta_2(\tau)&=-\frac{1}{8}(\theta_1^4+\theta_3^4)=-\frac{1}{8}-3\sum_{n=1}^{\infty}\sum_{\substack{d|n\\ d\text{ odd}}}dq^{n/2}=-\frac{1}{8}-3q^{1/2}-3q+\cdots,\\ \varepsilon=\varepsilon_2(\tau)&=\frac{1}{16}\theta_1^4\theta_3^4=\sum_{n=1}^{\infty}\sum_{\substack{d|n\\ n/d\text{ odd}}}d^3q^{n/2}=q^{1/2}+8q+\cdots,\end{aligned}$$

equation (2-1) has the solution

$$f_2(z,\tau)=2\pi\sqrt{-1}\frac{\theta(z,\tau)}{\theta'(0,\tau)}\frac{\theta_2(0,\tau)}{\theta_2(z,\tau)}.$$

Let $M$ be a $4k$-dimensional closed smooth oriented manifold. Let $\{\pm 2\pi\sqrt{-1}x_i,\ 1\le i\le 2k\}$ be the formal Chern roots of the complexification $T_{\mathbb{C}}M = TM\otimes\mathbb{C}$. Consider the two characteristic numbers

$$\begin{aligned} \mathrm{Ell}_1(M,\tau) &= 2^{2k}\left\langle \prod_{i=1}^{2k}\frac{2\pi\sqrt{-1}x_i}{f_1(x_i,\tau)},[M]\right\rangle \in \mathbb{Q}[\![q]\!],\\ \mathrm{Ell}_2(M,\tau) &= \left\langle \prod_{i=1}^{2k}\frac{2\pi\sqrt{-1}x_i}{f_2(x_i,\tau)},[M]\right\rangle \in \mathbb{Q}[\![q^{1/2}]\!]. \end{aligned} \tag{2-4}$$

$\mathrm{Ell}_1(M,\tau)$, $\mathrm{Ell}_2(M,\tau)$ can be written as signature and $\widehat{A}$-genus twisted by the Witten bundles. More precisely, let

$$\widehat{A}(M) = \prod_{i=1}^{2k}\frac{\pi\sqrt{-1}x_i}{\sinh\pi\sqrt{-1}x_i} \tag{2-5}$$

be the $\widehat{A}$-class and

$$\widehat{L}(M) = \prod_{i=1}^{2k}\frac{2\pi\sqrt{-1}x_i}{\tanh\pi\sqrt{-1}x_i} \tag{2-6}$$

the $\widehat{L}$-class. Let $E$ be a complex vector bundle on $M$. $\langle \widehat{L}(M)\mathrm{ch}E,[M]\rangle$ is equal to the index of the twisted signature operator $\mathrm{ind}(d_s\otimes E)=\mathrm{Sig}(M,E)$. When $M$ is spin, $\langle\widehat{A}(M)\mathrm{ch}E,[M]\rangle$ is equal to the index of the twisted Atiyah–Singer Dirac operator $\mathrm{ind}(D\otimes E)$. When twisted by bundles naturally constructed from the tangent bundle $TM$ of $M$, denote

$$\begin{aligned} \widehat{A}(M,T^i\otimes\Lambda^j\otimes S^k) &:= \widehat{A}(M,\otimes^i T_{\mathbb{C}}M\otimes\Lambda^j(T_{\mathbb{C}}M)\otimes S^k(T_{\mathbb{C}}M)),\\ \mathrm{Sig}(M,T^i\otimes\Lambda^j\otimes S^k) &:= \mathrm{Sig}(M,\otimes^i T_{\mathbb{C}}M\otimes\Lambda^j(T_{\mathbb{C}}M)\otimes S^k(T_{\mathbb{C}}M)), \end{aligned}$$

where $\Lambda^j(T_{\mathbb{C}}M)$ and $S^k(T_{\mathbb{C}}M)$ are the $j$-th exterior and $k$-th symmetric powers of $T_{\mathbb{C}}M$ respectively.

For any complex variable $t$, let

$$\Lambda_t(E) = \mathbb{C} + tE + t^2\Lambda^2(E)+\cdots,\qquad S_t(E)=\mathbb{C}+tE+t^2S^2(E)+\cdots$$

denote respectively the total exterior and symmetric powers of $E$, which live in $K(M)[\![t]\!]$. Denote by

$$\Theta_1(T_{\mathbb{C}}M) = \bigotimes_{n=1}^{\infty} S_{q^n}(T_{\mathbb{C}}M-\mathbb{C}^{4k})\otimes\bigotimes_{m=1}^{\infty}\Lambda_{q^m}(T_{\mathbb{C}}M-\mathbb{C}^{4k}), \tag{2-7}$$

$$\Theta_2(T_{\mathbb{C}}M) = \bigotimes_{n=1}^{\infty} S_{q^n}(T_{\mathbb{C}}M-\mathbb{C}^{4k})\otimes\bigotimes_{m=1}^{\infty}\Lambda_{-q^{m-(1/2)}}(T_{\mathbb{C}}M-\mathbb{C}^{4k}) \tag{2-8}$$

the Witten bundles, which are elements in $K(M)[[q^{1/2}]]$. Then one has

$$\begin{aligned}\mathrm{Ell}_1(M,\tau)&=\big\langle\widehat{L}(M)\mathrm{ch}(\Theta_1(T_{\mathbb{C}}M)),[M]\big\rangle,\\ \mathrm{Ell}_2(M,\tau)&=\big\langle\widehat{A}(M)\mathrm{ch}(\Theta_2(T_{\mathbb{C}}M)),[M]\big\rangle.\end{aligned}\tag{2-9}$$

## 3. Proof of Theorem 1

Let $M$ be a $4k$-dimensional closed smooth oriented manifold. By (1-3),

$$\phi(M)=a_0(M)\delta^6+a_1(M)\delta^4\varepsilon+a_2(M)\delta^2\varepsilon^2+a_3(M)\varepsilon^3,\tag{3-1}$$

where $a_i(M)\in\mathbb{Z}\big[\frac{1}{2}\big]$, $0\le i\le 3$. First we show that one can express the 4 Pontryagin numbers $a_i(M)$ $(0\le i\le 3)$ in terms of $\widehat{A}$-genus, signature and their twists by the tangent bundle.

**Proposition 3.1.** *Let $M$ be a $4k$-dimensional closed smooth oriented manifold. One has*

$$\begin{aligned}a_0(M)&=2^{18}\widehat{A}(M),\\ a_1(M)&=-2^{15}\cdot 3\cdot 5\widehat{A}(M)-2^{12}\widehat{A}(M,T),\\ a_2(M)&=2^{16}\cdot 3\widehat{A}(M)+2^{13}\widehat{A}(M,T)+\frac{1}{2^5}\,\mathrm{Sig}(M,T),\\ a_3(M)&=2^{15}\widehat{A}(M)-2^{12}\widehat{A}(M,T)-\frac{1}{2^5}\,\mathrm{Sig}(M,T)+\mathrm{Sig}(M).\end{aligned}$$

*Proof.* From the preliminary in Section 2, we see that

$$\mathrm{Ell}_1(M)=2^{12}\big(a_0(M)\delta_1^6+a_1(M)\delta_1^4\varepsilon_1+a_2(M)\delta_1^2\varepsilon_1^2+a_3(M)\varepsilon_1^3\big),\tag{3-2}$$

$$\mathrm{Ell}_2(M)=a_0(M)\delta_2^6+a_1(M)\delta_2^4\varepsilon_2+a_2(M)\delta_2^2\varepsilon_2^2+a_3(M)\varepsilon_2^3.\tag{3-3}$$

On the other hand, by (2-7), (2-8) and (2-9), it is not hard to compute that

$$\begin{aligned}\mathrm{Ell}_1(M,\tau)&=\mathrm{Sig}(M)+(2\,\mathrm{Sig}(M,T)-48\,\mathrm{Sig}(M))q+\cdots,\\ \mathrm{Ell}_2(M,\tau)&=\widehat{A}(M)-(\widehat{A}(M,T)-24\widehat{A}(M))q^{1/2}+\cdots.\end{aligned}\tag{3-4}$$

With the help of (2-2) and (2-3), we can compare (3-4) with (3-3) and (3-2). For instance, by modulo higher terms $(q^{1/2})^i$ with $i\ge 2$,

$$\begin{aligned}\mathrm{Ell}_2(M)&\equiv\frac{a_0(M)}{8^6}(-1-24q^{1/2})^6+\frac{a_1(M)}{8^4}(-1-24q^{1/2})^4(q^{1/2})\\ &\equiv\frac{a_0(M)}{2^{18}}+\left(\frac{3^2a_0(M)}{2^{14}}+\frac{a_1(M)}{2^{12}}\right)q^{1/2}.\end{aligned}$$

Combining the above formula with (3-4), we have

$$\frac{a_0(M)}{2^{18}}=\widehat{A}(M),\qquad \frac{3^2a_0(M)}{2^{14}}+\frac{a_1(M)}{2^{12}}=-\widehat{A}(M,T)+24\widehat{A}(M).\tag{3-5}$$

Similarly, by modulo higher terms $q^i$ with $i \geq 2$,

$$\begin{aligned}\mathrm{Ell}_1(M) &\equiv 2^{12}\biggl(\frac{a_0(M)}{8^6}(2+48q)^6 + \frac{a_1(M)}{8^4}(2+48q)^4\Bigl(\frac{1}{16}-q\Bigr)\\ &\qquad + \frac{a_2(M)}{8^2}(2+48q)^2\Bigl(\frac{1}{16}-q\Bigr)^2 + a_3(M)\Bigl(\frac{1}{16}-q\Bigr)^3\biggr)\\ &\equiv (a_0(M)+a_1(M)+a_2(M)+a_3(M))\\ &\qquad + (144a_0(M)+80a_1(M)+16a_2(M)-48a_3(M))q.\end{aligned}$$

Combining the above formula with (3-4), we have

$$\begin{aligned} a_0(M)+a_1(M)+a_2(M)+a_3(M) &= \mathrm{Sig}(M),\\ 144a_0(M)+80a_1(M)+16a_2(M)-48a_3(M) &= 2\,\mathrm{Sig}(M,T)-48\,\mathrm{Sig}(M).\end{aligned} \tag{3-6}$$

The equalities in (3-5) and (3-6) can be organized to result in a matrix equation

$$\begin{pmatrix} \frac{1}{2^{18}} & 0 & 0 & 0\\ \frac{3^2}{2^{14}} & \frac{1}{2^{12}} & 0 & 0\\ 1 & 1 & 1 & 1\\ 144 & 80 & 16 & -48\end{pmatrix}\cdot\begin{pmatrix} a_0(M)\\ a_1(M)\\ a_2(M)\\ a_3(M)\end{pmatrix} = \begin{pmatrix} \widehat{A}(M)\\ -\widehat{A}(M,T)+24\widehat{A}(M)\\ \mathrm{Sig}(M)\\ 2\,\mathrm{Sig}(M,T)-48\,\mathrm{Sig}(M)\end{pmatrix}. \tag{3-7}$$

We can solve $a_i(M)$ from (3-7), and then the proposition is proved. □

Now suppose $M$ is further a string manifold. With the string condition, we can rewrite the equalities of $a_i(M)$ in Proposition 3.1 in terms of a new family of (twisted) genera, which is helpful for proving Theorem 1.

**Proposition 3.2.** *Let $M$ be a* 24*-dimensional closed smooth string manifold. Then*

$$\begin{pmatrix} a_0(M)\\ a_1(M)\\ a_2(M)\\ a_3(M)\end{pmatrix} = \begin{pmatrix} 2^{18} & 0 & 0 & 0\\ -2^{15}\cdot 3\cdot 5 & -2^{15}\cdot 3 & 0 & 0\\ 2^8\cdot 3\cdot 331 & 2^9\cdot 3^5 & 2^6 & 0\\ -2^8\cdot 97 & -2^9\cdot 3\cdot 17 & -2^6 & 2^3\end{pmatrix}\cdot\begin{pmatrix} \widehat{A}(M)\\ \frac{1}{24}\widehat{A}(M,T)\\ \widehat{A}(M,\Lambda^2)\\ \frac{1}{8}\,\mathrm{Sig}(M)\end{pmatrix}. \tag{3-8}$$

*Proof.* Under the string condition, the twisted and untwisted genera in Proposition 3.1 possess intrinsic relations. Indeed, by combining modularity of the Witten genus and a modular form constructed in [Liu and Wang 2013], Chen and Han [2015] showed that, when $M$ is a 24-dimensional closed smooth string manifold, one has

$$\mathrm{Sig}(M,T) = 2^{11}\bigl(\widehat{A}(M,\Lambda^2)-47\widehat{A}(M,T)+900\widehat{A}(M)\bigr). \tag{3-9}$$

With (3-9) we can rewrite the equalities of $a_i(M)$ in Proposition 3.1 as displayed in this proposition. □

In [Han and Huang 2022] we determined an integral basis of $\Omega_{24}^{\mathrm{String}}$, which is crucial for the proof of Theorem 1.

**Theorem 3.3** [Han and Huang 2022, Theorem 1 and Corollary 3]. *The correspondence* $\kappa : \Omega_{24}^{\mathrm{String}} \to \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z} \oplus \mathbb{Z}$ *defined by*

$$\kappa(M) = \left( \widehat{A}(M), \frac{1}{24}\widehat{A}(M, T), \widehat{A}(M, \Lambda^2), \frac{1}{8}\operatorname{Sig}(M) \right)$$

*is an isomorphism of abelian groups. Moreover, there exists a basis* $\{M_i\}_{1 \le i \le 4}$ *of* $\Omega_{24}^{\mathrm{String}}$ *such that*

$$K := \begin{pmatrix} \kappa(M_1) \\ \kappa(M_2) \\ \kappa(M_3) \\ \kappa(M_4) \end{pmatrix}^{\tau} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 2^3 \cdot 3^3 \cdot 5 & 2^2 \cdot 3 \cdot 17 \cdot 1069 & -1 & 0 \\ 2^8 \cdot 3 \cdot 61 & 2^8 \cdot 5 \cdot 37 & 2^2 \cdot 7 & 1 \end{pmatrix}. \qquad \square$$

Now we are ready to prove Theorem 1.

*Proof of Theorem 1.* Suppose $M$ satisfies that $\phi(M) = 0$. Then by (3-1) $a_i(M) = 0$ for $0 \le i \le 3$. Notice that in (3-8) the coefficient matrix is invertible. Then by Proposition 3.2, the 4 index numbers $\widehat{A}(M)$, $\frac{1}{24}\widehat{A}(M, T)$, $\widehat{A}(M, \Lambda^2)$ and $\frac{1}{8}\operatorname{Sig}(M)$ vanish. This means that $\kappa(M) = (0, 0, 0, 0)$. Since by Theorem 3.3 $\kappa$ is an isomorphism, $[M] = 0 \in \Omega_{24}^{\mathrm{String}}$. Hence $\phi$ is injective.

To compute the image of the elliptic genus, we need to compute the elliptic genus of the generators $M_i$ $(1 \le i \le 4)$ in Theorem 3.3. By Proposition 3.2 and Theorem 3.3, they can be computed by the matrix multiplication

$$\begin{aligned} &\begin{pmatrix} 2^{18} & 0 & 0 & 0 \\ -2^{15} \cdot 3 \cdot 5 & -2^{15} \cdot 3 & 0 & 0 \\ 2^8 \cdot 3 \cdot 331 & 2^9 \cdot 3^5 & 2^6 & 0 \\ -2^8 \cdot 97 & -2^9 \cdot 3 \cdot 17 & -2^6 & 2^3 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 2^3 \cdot 3^3 \cdot 5 & 2^2 \cdot 3 \cdot 17 \cdot 1069 & -1 & 0 \\ 2^8 \cdot 3 \cdot 61 & 2^8 \cdot 5 \cdot 37 & 2^2 \cdot 7 & 1 \end{pmatrix} \\ &\qquad = \begin{pmatrix} 0 & 2^{18} & 0 & 0 \\ 2^{15} \cdot 3 & -2^{15} \cdot 3 \cdot 5 & 0 & 0 \\ -2^{11} \cdot 3^3 & 2^{11} \cdot 3^3 \cdot 257 & -2^6 & 0 \\ 2^{12} \cdot 3^4 & -2^{12} \cdot 3^4 \cdot 41 & 2^5 \cdot 3^2 & 2^3 \end{pmatrix} = (a_j(M_i))_{j \times i}. \end{aligned}$$

Combining (3-1), the above matrix gives the 4 generators of the image $\phi(\Omega_*^{\mathrm{String}})$ as

$$\begin{aligned} 2^3 \cdot \varepsilon^3 &= \phi(M_4), \\ -(8\delta)^2\varepsilon^2 + 2^5 \cdot 3^2 \cdot \varepsilon^3 &= \phi(M_3), \\ 2^3 \cdot 3 \cdot (8\delta)^4\varepsilon - 2^5 \cdot 3^3 \cdot (8\delta)^2\varepsilon^2 + 2^{12} \cdot 3^4 \cdot \varepsilon^3 &= \phi(M_1), \\ (8\delta)^6 - 2^3 \cdot 3 \cdot 5 \cdot (8\delta)^4\varepsilon + 2^5 \cdot 3^3 \cdot 257 \cdot (8\delta)^2\varepsilon^2 - 2^{12} \cdot 3^4 \cdot 41 \cdot \varepsilon^3 &= \phi(M_2). \end{aligned}$$

It follows that $\phi(\Omega_*^{\mathrm{String}})$ is generated by $2^3\cdot\varepsilon^3$, $(8\delta)^2\varepsilon^2$, $2^3\cdot 3\cdot(8\delta)^4\varepsilon$ and $(8\delta)^6$. This completes the proof of the theorem. $\square$


## Acknowledgements

Fei Han was partially supported by the grant AcRF R-146-000-263-114 from the National University of Singapore. He thanks Professors Kefeng Liu and Weiping Zhang for helpful discussions. Ruizhi Huang was supported in part by the National Natural Science Foundation of China (Grant nos. 12331003, 12288201 and 11801544), the National Key R&D Program of China (No. 2021YFA1002300), the Youth Innovation Promotion Association of Chinese Academy Sciences, and the "Chen Jingrun" Future Star Program of AMSS.


## References


[Anderson et al. 1967] D. W. Anderson, E. H. Brown, Jr., and F. P. Peterson, "The structure of the Spin cobordism ring", *Ann. of Math.* (2) **86** (1967), 271–298. MR Zbl

[Ando et al. 2001] M. Ando, M. J. Hopkins, and N. P. Strickland, "Elliptic spectra, the Witten genus and the theorem of the cube", *Invent. Math.* **146**:3 (2001), 595–687. MR Zbl

[Bott and Taubes 1989] R. Bott and C. Taubes, "On the rigidity theorems of Witten", *J. Amer. Math. Soc.* **2**:1 (1989), 137–186. MR Zbl

[Chandrasekharan 1985] K. Chandrasekharan, *Elliptic functions*, Grundl. Math. Wissen. **281**, Springer, 1985. MR Zbl

[Chen and Han 2015] Q. Chen and F. Han, "Mod 3 congruence and twisted signature of 24 dimensional string manifolds", *Trans. Amer. Math. Soc.* **367**:4 (2015), 2959–2977. MR Zbl

[Chen and Han 2024] X. Chen and F. Han, "New Bochner type theorems", *Math. Ann.* **388**:4 (2024), 3757–3783. Zbl

[Chen et al. 2012] Q. Chen, F. Han, and W. Zhang, "Divisibility of modified signature of string manifold in dimension 8, 16 and 24", unpublished note, 2012.

[Chudnovsky et al. 1988] D. V. Chudnovsky, G. V. Chudnovsky, P. S. Landweber, S. Ochanine, and R. E. Stong, "Integrality and divisibility of the elliptic genus", preprint, 1988.

[Davis and Fang 2016] J. F. Davis and F. Fang, "An almost flat manifold with a cyclic or quaternionic holonomy group bounds", *J. Differential Geom.* **103**:2 (2016), 289–296. MR Zbl

[Farrell and Zdravkovska 1983] F. T. Farrell and S. Zdravkovska, "Do almost flat manifolds bound?", *Michigan Math. J.* **30**:2 (1983), 199–208. MR Zbl

[Gromov 1978] M. Gromov, "Almost flat manifolds", *J. Differential Geometry* **13**:2 (1978), 231–241. MR

[Han and Huang 2022] F. Han and R. Huang, "On characteristic numbers of 24 dimensional string manifolds", *Math. Z.* **300**:3 (2022), 2309–2331. MR Zbl

[Hirzebruch et al. 1994] F. Hirzebruch, T. Berger, and R. Jung, *Manifolds and modular forms*, 2nd ed., Aspects of Mathematics **E20**, Friedrich Vieweg & Sohn, Braunschweig, 1994. MR Zbl

[Hopkins 2002] M. J. Hopkins, "Algebraic topology and modular forms", pp. 291–317 in *Proceedings of the International Congress of Mathematicians* (Beijing, 2002), vol. I, edited by T. Li, Higher Ed. Press, Beijing, 2002. MR Zbl

[Kreck and Stolz 1993] M. Kreck and S. Stolz, “$\mathbf{HP}^2$-bundles and elliptic homology”, *Acta Math.* **171**:2 (1993), 231–261. MR Zbl

[Landweber 1988a] P. S. Landweber, “Elliptic cohomology and modular forms”, pp. 55–68 in *Elliptic curves and modular forms in algebraic topology* (Princeton, NJ, 1986), edited by P. S. Landweber, Lecture Notes in Math. **1326**, Springer, 1988. MR Zbl

[Landweber 1988b] P. S. Landweber, “Elliptic genera: an introductory overview”, pp. 1–10 in *Elliptic curves and modular forms in algebraic topology* (Princeton, NJ, 1986), edited by P. S. Landweber, Lecture Notes in Math. **1326**, Springer, 1988. MR Zbl

[Landweber and Stong 1988] P. S. Landweber and R. E. Stong, “Circle actions on Spin manifolds and characteristic numbers”, *Topology* **27**:2 (1988), 145–161. MR Zbl

[Liu 1992] K. Liu, “On mod 2 and higher elliptic genera”, *Comm. Math. Phys.* **149**:1 (1992), 71–95. MR Zbl

[Liu 1995a] K. Liu, “Modular invariance and characteristic numbers”, *Comm. Math. Phys.* **174**:1 (1995), 29–42. MR Zbl

[Liu 1995b] K. Liu, “On modular invariance and rigidity theorems”, *J. Differential Geom.* **41**:2 (1995), 343–396. MR Zbl

[Liu 1996a] K. Liu, “Modular forms and topology”, pp. 237–262 in *Moonshine, the Monster, and related topics* (South Hadley, MA, 1994), edited by C. Dong and G. Mason, Contemp. Math. **193**, Amer. Math. Soc., Providence, RI, 1996. MR Zbl

[Liu 1996b] K. Liu, “On elliptic genera and theta-functions”, *Topology* **35**:3 (1996), 617–640. MR Zbl

[Liu and Wang 2013] K. Liu and Y. Wang, “A note on modular forms and generalized anomaly cancellation formulas”, *Sci. China Math.* **56**:1 (2013), 55–65. MR Zbl

[Mahowald and Hopkins 2002] M. Mahowald and M. Hopkins, “The structure of 24 dimensional manifolds having normal bundles which lift to $BO[8]$”, pp. 89–110 in *Recent progress in homotopy theory* (Baltimore, MD, 2000), edited by D. M. Davis et al., Contemp. Math. **293**, Amer. Math. Soc., Providence, RI, 2002. MR Zbl

[Milivojević 2021] A. Milivojević, “The weak form of Hirzebruch’s prize question via rational surgery”, preprint, 2021, https://www.math.uwaterloo.ca/~amilivoj/rationalHirzebruchmanifold.pdf.

[Ochanine 1987] S. Ochanine, “Sur les genres multiplicatifs définis par des intégrales elliptiques”, *Topology* **26**:2 (1987), 143–151. MR Zbl

[Ochanine 1988] S. Ochanine, “Genres elliptiques équivariants”, pp. 107–122 in *Elliptic curves and modular forms in algebraic topology* (Princeton, NJ, 1986), edited by P. S. Landweber, Lecture Notes in Math. **1326**, Springer, 1988. MR Zbl

[Ruh 1982] E. A. Ruh, “Almost flat manifolds”, *J. Differential Geometry* **17**:1 (1982), 1–14. MR Zbl

[Segal 1988] G. Segal, “Elliptic cohomology (after Landweber-Stong, Ochanine, Witten, and others)”, exposé 695, pp. 187–201 in *Séminaire Bourbaki*, 1987/88, Astérisque **161-162**, 1988. MR Zbl

[Taubes 1989] C. H. Taubes, “$S^1$ actions and elliptic genera”, *Comm. Math. Phys.* **122**:3 (1989), 455–526. MR Zbl

[Wall 1960] C. T. C. Wall, “Determination of the cobordism ring”, *Ann. of Math.* (2) **72** (1960), 292–311. MR Zbl

[Witten 1987] E. Witten, “Elliptic genera and quantum field theory”, *Comm. Math. Phys.* **109**:4 (1987), 525–536. MR Zbl

[Witten 1988] E. Witten, “The index of the Dirac operator in loop space”, pp. 161–181 in *Elliptic curves and modular forms in algebraic topology* (Princeton, NJ, 1986), edited by P. S. Landweber, Lecture Notes in Math. **1326**, Springer, 1988. MR Zbl

[Yau 1993] S.-T. Yau, "Open problems in geometry", pp. 1–28 in *Differential geometry*: *partial differential equations on manifolds* (Los Angeles, CA, 1990), edited by R. Greene and S. T. Yau, Proc. Sympos. Pure Math., Part 1 **54**, Amer. Math. Soc., Providence, RI, 1993. MR Zbl

Received February 28, 2023. Revised October 20, 2023.

FEI HAN
DEPARTMENT OF MATHEMATICS
NATIONAL UNIVERSITY OF SINGAPORE
SINGAPORE
mathanf@nus.edu.sg

RUIZHI HUANG
INSTITUTE OF MATHEMATICS AND SYSTEMS SCIENCE
CHINESE ACADEMY OF SCIENCES
BEIJING
CHINA
huangrz@amss.ac.cn

PACIFIC JOURNAL OF MATHEMATICS
Vol. 328, No. 2, 2024
https://doi.org/10.2140/pjm.2024.328.287

# THE DOMINATION MONOID IN HENSELIAN VALUED FIELDS

MARTIN HILS AND ROSARIO MENNUNI

**We study the domination monoid in various classes of structures arising from henselian valuations, including $\mathcal{RV}$-expansions of henselian valued fields of equicharacteristic 0 (and, more generally, of benign valued fields), $\mathfrak{p}$-adically closed fields, monotone D-henselian differential valued fields with many constants, regular ordered abelian groups, and pure short exact sequences of abelian structures. We obtain Ax–Kochen–Ershov-type reductions to suitable fully embedded families of sorts in quite general settings, and full computations in concrete ones.**

In their seminal work [17] on stable domination, Haskell, Hrushovski and Macpherson introduced the *domination monoid* $\widetilde{\mathrm{Inv}}(\mathfrak{U})$, and showed that in algebraically closed valued fields it decomposes as $\widetilde{\mathrm{Inv}}(\mathrm{k}(\mathfrak{U})) \times \widetilde{\mathrm{Inv}}(\Gamma(\mathfrak{U}))$, where k denotes the residue field, $\Gamma$ the value group, and $\mathfrak{U}$ a *monster* model, that is, a sufficiently saturated and strongly homogeneous model. (Strictly speaking, Haskell et al. [17] work with $\overline{\mathrm{Inv}}(\mathfrak{U})$, which is in general different, but coincides with $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ in their setting. See [21, Remark 2.1.14 and Theorem 5.2.22].) A similar result was proven in [12; 23] in the case of real closed fields with a convex valuation. This paper revolves around understanding $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ in more general classes of valued fields, and expansions thereof. A special case of our results is the following.

**Theorem A** (Corollary 6.19). *Let $T$ be the theory of a henselian valued field of equicharacteristic* 0, *or algebraically maximal Kaplansky*, *possibly enriched on* k *and* $\Gamma$. *If all* $\mathrm{k}^\times/(\mathrm{k}^\times)^n$ *are finite*, *then* $\widetilde{\mathrm{Inv}}(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}(\mathrm{k}(\mathfrak{U})) \times \widetilde{\mathrm{Inv}}(\Gamma(\mathfrak{U}))$.

More generally, we obtain a two-step reduction, first to leading term structures, and then, using technology on pure short exact sequences recently developed in [2], to k and $\Gamma$, albeit in a form which, in general, is (necessarily) slightly more

The authors were supported by the Deutsche Forschungsgemeinschaft (DFG) via HI 2004/1-1 (part of the ANR-DFG project GeoMod) and under Germany's Excellence Strategy EXC 2044-390685587, Mathematics Münster: Dynamics-Geometry-Structure. Mennuni was supported by the Italian research projects PRIN 2017: "Mathematical Logic: models, sets, computability" Prot. 2017NWTM8RPRIN and PRIN 2022: "Models, sets and classifications" Prot. 2022TECZJA and is a member of the INdAM research group GNSAGA..

*MSC2020:* primary 03C45; secondary 03C60, 03C64, 12J10, 12L12.

*Keywords:* model theory, invariant types, domination monoid, valued fields, ordered abelian groups.

involved. We also compute $\widetilde{\mathrm{Inv}}(\Gamma(\mathfrak{U}))$ when the theory of $\Gamma$ has an archimedean model, and prove several accessory statements.

Before stating our results in more detail, let us give an informal account of the context (see Section 1 for the precise definitions). The starting point is the space $S^{\mathrm{inv}}(\mathfrak{U})$ of *invariant types* over a monster model $\mathfrak{U}$: those which are invariant over a small subset. It is a dense subspace of $S(\mathfrak{U})$, whose points may be canonically extended to larger parameter sets. Such extensions allow to define the *tensor product*, or *Morley product*, obtaining a semigroup $(S^{\mathrm{inv}}(\mathfrak{U}), \otimes)$, in fact a monoid. The space $S^{\mathrm{inv}}(\mathfrak{U})$ also comes with a preorder $\geq_{\mathrm{D}}$, called *domination*: roughly, $p \geq_{\mathrm{D}} q$ means that $q$ is recoverable from $p$ plus a small amount of information. The quotient by the induced equivalence relation, *domination-equivalence* $\sim_{\mathrm{D}}$, is then a poset, denoted by $(\widetilde{\mathrm{Inv}}(\mathfrak{U}), \geq_{\mathrm{D}})$. If $\otimes$ respects $\geq_{\mathrm{D}}$, i.e., if $(S^{\mathrm{inv}}(\mathfrak{U}), \otimes, \geq_{\mathrm{D}})$ is a preordered semigroup, then $\sim_{\mathrm{D}}$ is a congruence with respect to $\otimes$ and we say that *the domination monoid is well defined*, and equip $(\widetilde{\mathrm{Inv}}(\mathfrak{U}), \geq_{\mathrm{D}})$ with the operation induced by $\otimes$. Compatibility of $\otimes$ and $\geq_{\mathrm{D}}$ in a given theory can be shown by using certain sufficient criteria, isolated in [22] and applied, e.g., in [24], or by finding a nice system of representatives for $\sim_{\mathrm{D}}$-classes (see Proposition 1.3). Nevertheless, in general, $\otimes$ may fail to respect $\geq_{\mathrm{D}}$ [22]. Hence, when dealing with $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ in a given structure, one needs to understand whether it is well defined as a monoid; and, when dealing with it in the abstract, the monoid structure cannot be taken for granted.

Recall that to a valued field K are associated certain abelian groups augmented by an absorbing element, fitting in a short exact sequence

$$1 \to (\mathrm{k}, \times) \to (\mathrm{K}, \times)/(1+\mathfrak{m}) \to \Gamma \cup \{\infty\} \to 0,$$

denoted by $\mathcal{RV}$. This sequence is interpretable in K, and this interpretation endows it with extra structure. The amount of induced structure clearly depends on whether K has extra structure itself, but at a bare minimum k will carry the language of fields and $\Gamma$ that of ordered abelian groups. By [4] (see also [20], or [14; 15] for a more modern treatment), henselian valued fields of residue characteristic 0 eliminate quantifiers relatively to $\mathcal{RV}$, and the latter is fully embedded with the structure described above. This holds resplendently, in the sense that it is still true after arbitrary expansions of $\mathcal{RV}$. The same holds in the algebraically maximal Kaplansky case, by [20] (see also [15]).[1] These are known after [30] as classes of *benign* valued fields and, in several contexts, they turn out to be particularly amenable to model-theoretic investigation. One of our main results says the context of domination is no exception.

**Theorem B** (Theorem 6.18). *In every $\mathcal{RV}$-expansion of a benign theory of valued fields there is an isomorphism of posets $\widetilde{\mathrm{Inv}}(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}(\mathcal{RV}(\mathfrak{U}))$. If $\otimes$ respects $\geq_{\mathrm{D}}$ in $\mathcal{RV}(\mathfrak{U})$, then $\otimes$ respects $\geq_{\mathrm{D}}$ in $\mathfrak{U}$, and the above is an isomorphism of monoids.*

[1] Note that these quantifier elimination results are already implicitly contained in [9].

Having reduced $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ to the short exact sequence $\mathcal{RV}$, the next step is to reduce it to its kernel k and quotient $\Gamma$. If we add an angular component map, the sequence $\mathcal{RV}$ splits and we obtain a product decomposition as in Theorem A (Remark 6.1). Without an angular component, a product decomposition is not always possible; yet, k and $\Gamma$ still exert a tight control on $\mathcal{RV}$. This behaviour is not peculiar of $\mathcal{RV}$: it holds in short exact sequences of abelian structures, provided they satisfy a purity assumption, using the relative quantifier elimination from [2]. For reasons to be clarified later (Remark 4.17), here it is natural to look at types in infinitely many variables, say $\kappa$, and hence at the corresponding analogue $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U})$ of $\widetilde{\mathrm{Inv}}(\mathfrak{U})$.

**Theorem C** (Corollary 4.9). *Let $\mathfrak{U}$ be a pure short exact sequence*

$$0 \to \mathcal{A} \to \mathcal{B} \to \mathcal{C} \to 0$$

*of $L$-abelian structures, where $\mathcal{A}$ and $\mathcal{C}$ may carry extra structure. Let $\kappa \geq |L|$ be a small cardinal. There is an expansion $\mathcal{A}_{\mathcal{F}}$ of $\mathcal{A}$ by imaginary sorts yielding an isomorphism of posets $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}_\kappa(\mathcal{A}_{\mathcal{F}}(\mathfrak{U})) \times \widetilde{\mathrm{Inv}}_\kappa(\mathcal{C}(\mathfrak{U}))$. If $\otimes$ respects $\geq_{\mathrm{D}}$ in both $\mathcal{A}_{\mathcal{F}}(\mathfrak{U})$ and $\mathcal{C}(\mathfrak{U})$, then $\otimes$ respects $\geq_{\mathrm{D}}$ in $\mathfrak{U}$, and the above is an isomorphism of monoids.*

In algebraically or real closed valued fields, the isomorphism

$$\widetilde{\mathrm{Inv}}(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}(\mathrm{k}(\mathfrak{U})) \times \widetilde{\mathrm{Inv}}(\Gamma(\mathfrak{U}))$$

is complemented by a computation of the factors, carried out in [17; 23]. In particular, if $\Gamma(\mathfrak{U})$ is divisible, then $\widetilde{\mathrm{Inv}}(\Gamma(\mathfrak{U}))$ is isomorphic to the upper semilattice of finite sets of invariant convex subgroups of $\Gamma(\mathfrak{U})$ (in the sense of Definition 3.16). A further contribution of this work is the computation of $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ in the next simplest class of theories of ordered abelian groups: those with an archimedean model, known as *regular*. Denote by $\mathrm{CS}^{\mathrm{inv}}(\mathfrak{U})$ the set of invariant convex subgroups of $\mathfrak{U}$, by $\mathscr{P}_{\leq\kappa}(\mathrm{CS}^{\mathrm{inv}}(\mathfrak{U}))$ the upper semilattice of its subsets of size at most $\kappa$, and by $\hat{\kappa}$ the ordered monoid of cardinals smaller or equal than $\kappa$ with cardinal sum.

**Theorem D** (Corollary 3.33). *Let $T$ be the theory of a regular ordered abelian group, $\kappa$ a small infinite cardinal, and $\mathbb{P}_T$ the set of primes $\mathfrak{p}$ such that $\mathfrak{U}/\mathfrak{p}\mathfrak{U}$ is infinite. Then $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}^{\mathrm{eq}})$ is well defined, and $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}^{\mathrm{eq}}) \cong \mathscr{P}_{\leq\kappa}(\mathrm{CS}^{\mathrm{inv}}(\mathfrak{U})) \times \prod_{\mathbb{P}_T} \hat{\kappa}$.*

Theorem D applies to Presburger arithmetic, the theory of $(\mathbb{Z}, +, <)$. Pairing this with a suitable generalisation of Theorem B, we obtain the following.

**Theorem E** (Corollary 7.7). *In the theory* $\mathrm{Th}(\mathbb{Q}_\mathfrak{p})$ *of $\mathfrak{p}$-adically closed fields, $\otimes$ respects $\geq_{\mathrm{D}}$, and* $\widetilde{\mathrm{Inv}}(\mathfrak{U}) \cong \mathscr{P}_{<\omega}\big(\mathrm{CS}^{\mathrm{inv}}(\Gamma(\mathfrak{U}))\big)$.

A similar statement (Corollary 7.5) holds for Witt vectors over $\mathbb{F}_\mathfrak{p}^{\mathrm{alg}}$. Finally, we move to monotone D-henselian differential valued fields with many constants.

While Theorem B does not generalise to this context (Remark 8.5), its analogue for $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U})$ does (Theorem 8.2). We fully compute $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U})$ in the model companion $\mathsf{VDF}_{\mathcal{EC}}$. Similar results hold for $\sigma$-henselian valued difference fields (Remark 8.6).

**Theorem F** (Theorem 8.4). *In* $\mathsf{VDF}_{\mathcal{EC}}$, *for every small infinite cardinal* $\kappa$, *the monoid* $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U})$ *is well defined*, *and we have isomorphisms*

$$\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}_\kappa(\mathrm{k}(\mathfrak{U})) \times \widetilde{\mathrm{Inv}}_\kappa(\Gamma(\mathfrak{U})) \cong \prod_{\delta(\mathfrak{U})}^{\leq\kappa} \hat{\kappa} \times \mathscr{P}_{\leq\kappa}\big(\mathrm{CS}^{\mathrm{inv}}(\Gamma(\mathfrak{U}))\big),$$

*where* $\delta(\mathfrak{U})$ *is a certain cardinal*, *and* $\prod_{\delta(\mathfrak{U})}^{\leq\kappa}\hat{\kappa}$ *denotes the submonoid of* $\prod_{\delta(\mathfrak{U})}\hat{\kappa}$ *consisting of* $\delta(\mathfrak{U})$-*sequences with support of size at most* $\kappa$.

The paper is structured as follows. In the first two sections we recall some preliminary notions and facts, and deal with some easy observations about orthogonality of invariant types. In Section 3 we prove Theorem D, while in Section 4 we study expanded pure short exact sequences of abelian structures, proving Theorem C. The results from these two sections are then combined in Section 5 to deal with the case of ordered abelian groups with finitely many definable convex subgroups. In Section 6 we prove Theorem B, and illustrate how it may be combined with Theorem C to obtain statements such as Theorem A. Section 7 deals with finitely ramified mixed characteristic henselian valued fields and includes a proof of Theorem E, and Section 8 deals with the differential case, proving Theorem F.

## 1. Preliminaries

***Notation and conventions.*** We adopt the conventions and notations of [23, Section 1.1] (e.g., we usually (and tacitly) fix a *monster model* $\mathfrak{U}$, and *definable* means $\mathfrak{U}$-definable), with the following additions and differences. The set of prime natural numbers is denoted by $\mathbb{P}$. Sorts are denoted by upright letters, as in $\mathrm{A}, \mathrm{K}, \mathrm{k}, \Gamma$, families of sorts by calligraphic letters such as $\mathcal{C}$, and $S_{\mathcal{C}^{<\omega}}(A)$ stands for the disjoint union of all spaces of types in finitely many variables, each with sort in $\mathcal{C}$. Terms may contain parameters, as in $t(x,d)$; we write $t(x)$ if they do not.

***Domination.*** We assume familiarity with invariant types, and recall some basic definitions and facts about domination. See [23, Section 1.2], [21, Section 2.1.2] and [22] for a more thorough treatment.

If $p(x), q(y) \in S(\mathfrak{U})$, let $S_{pq}(A)$ be the set of types over $A$ in variables $xy$ extending $(p(x) \restriction A) \cup (q(y) \restriction A)$. We say that $p(x) \in S(\mathfrak{U})$ *dominates* $q(y) \in S(\mathfrak{U})$, and write $p \geq_{\mathrm{D}} q$, if there are a small $A \subset^+ \mathfrak{U}$ and $r \in S_{pq}(A)$ such that $p(x) \cup r(x,y) \vdash q(y)$. We say that $p, q \in S(\mathfrak{U})$ are *domination-equivalent*, and write $p \sim_{\mathrm{D}} q$, if $p \geq_{\mathrm{D}} q$ and $q \geq_{\mathrm{D}} p$. We denote the domination-equivalence class of $p$ by $[\![p]\!]$. The *domination poset* $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ is the quotient of $S^{\mathrm{inv}}(\mathfrak{U})$ by $\sim_{\mathrm{D}}$, equipped with the partial order induced by $\geq_{\mathrm{D}}$, denoted by the same symbol. In other words, domination is the semiisolation counterpart to $\mathrm{F}^{\mathrm{s}}_{\kappa(\mathfrak{U})}$-isolation in the sense of [29, Chapter IV]; the two notions are distinct, see [24, Example 3.3].

We will be mostly concerned with domination on $S^{\mathrm{inv}}(\mathfrak{U})$. When describing a witness to $p \geq_{\mathrm{D}} q$, we write, e.g., "let $r$ contain $\varphi(x, y)$" with the meaning "let $r \in S_{pq}(A)$ contain $\varphi(x, y)$, for an $A$ such that $p, q \in S^{\mathrm{inv}}(\mathfrak{U}, A)$". By [22, Lemma 1.14], if $p_0, p_1 \in S^{\mathrm{inv}}(\mathfrak{U})$ and $p_0 \geq_{\mathrm{D}} p_1$, then $p_0 \otimes q \geq_{\mathrm{D}} p_1 \otimes q$. We say that $\otimes$ *respects* $\geq_{\mathrm{D}}$ if $q_0 \geq_{\mathrm{D}} q_1$ implies $p \otimes q_0 \geq_{\mathrm{D}} p \otimes q_1$. If this is the case, the *domination monoid* is the expansion of $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ by the operation induced by $\otimes$, also denoted by $\otimes$. If we say $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ *is well defined* (as a partially ordered monoid) we mean "$\otimes$ respects $\geq_{\mathrm{D}}$". As $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ is always well defined as a poset, this should cause no confusion.

Adding imaginary sorts to $\mathfrak{U}$ may result in an enlargement of $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ [22, Corollary 3.8]. Yet, if $T$ eliminates imaginaries, even just geometrically, then the natural embedding $\widetilde{\mathrm{Inv}}(\mathfrak{U}) \hookrightarrow \widetilde{\mathrm{Inv}}(\mathfrak{U}^{\mathrm{eq}})$ is easily seen to be an isomorphism. By [22, Proposition 1.23], domination witnessed by algebraicity is compatible with $\otimes$: if $p, q_0, q_1 \in S^{\mathrm{inv}}(\mathfrak{U})$ and, for $i < 2$, there are realisations $a_i \vDash q_i$ such that $a_1 \in \mathrm{acl}(\mathfrak{U}a_0)$, then for all invariant $p$ we have $p \otimes q_0 \geq_{\mathrm{D}} p \otimes q_1$. In particular, if $T$ has geometric elimination of imaginaries, then $\widetilde{\mathrm{Inv}}(\mathfrak{U}^{\mathrm{eq}})$ is well defined if and only if $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ is.

Frequently, we will equip a family of sorts, say $\mathcal{A} = \{\mathrm{A}_s \mid s \in S\}$, with the traces of some $\varnothing$-definable relations, and consider it as a standalone structure. We call $\mathcal{A}$ *fully embedded* if, for each $s_0, \dots, s_n \in S$, every subset of $(\mathrm{A}_{s_0} \times \cdots \times \mathrm{A}_{s_n})(\mathfrak{U})$ is definable in $\mathfrak{U}$ if and only if it is definable in $\mathcal{A}(\mathfrak{U})$. When talking of a fully embedded $\mathcal{A}$ in the abstract, as below, we assume a structure on $\mathcal{A}$ to be fixed.

**Fact 1.1** [21, Proposition 2.3.31]. Let $\mathcal{A}$ be a fully embedded family of sorts, and let $\iota : S_{\mathcal{A}^{<\omega}}(\mathcal{A}(\mathfrak{U})) \to S(\mathfrak{U})$ send a type of $\mathcal{A}(\mathfrak{U})$ to the unique type of $\mathfrak{U}$ it entails. The type $p$ is invariant if and only if $\iota(p)$ is. The map $\iota \restriction S^{\mathrm{inv}}(\mathcal{A}(\mathfrak{U}))$ is an injective $\otimes$-homomorphism inducing an embedding of posets $\tilde{\iota} : \widetilde{\mathrm{Inv}}(\mathcal{A}(\mathfrak{U})) \hookrightarrow \widetilde{\mathrm{Inv}}(\mathfrak{U})$ which, if $\otimes$ respects $\geq_{\mathrm{D}}$ in $\mathfrak{U}$ (hence also in $\mathcal{A}(\mathfrak{U})$), is also an embedding of monoids.

**Remark 1.2.** With the notation and assumptions from Fact 1.1, if $p$ is an invariant $\mathcal{A}(\mathfrak{U})$-type, $\mathfrak{U}_1 \succ \mathfrak{U}$, and $\mathcal{A}(\mathfrak{U}) \subseteq B \subseteq \mathcal{A}(\mathfrak{U}_1)$, then $(p \mid B) \vdash (\iota p \mid \mathfrak{U}B)$.

*Proof.* Suppose $\varphi(x, w, t) \in L(\varnothing)$, $d \in \mathfrak{U}$, $e \in B$, and $\iota p(x) \mid B \vdash \varphi(x, d, e)$. Since $x, t$ are $\mathcal{A}$-variables, and $d \in \mathfrak{U}$, full embeddedness yields an $L_{\mathcal{A}}(\mathcal{A}(\mathfrak{U}))$-formula $\psi(x, t)$ equivalent to $\varphi(x, d, t)$. So $\psi(x, e) \in p \mid B$ and we are done. $\square$

**Proposition 1.3.** *Assume for all $p \in S^{\mathrm{inv}}(\mathfrak{U})$ there is a tuple $\tau^p$ of definable functions with codomains in a fully embedded $\mathcal{A}$ such that $p \sim_{\mathrm{D}} \tau^p_* p$ and $p \otimes q \sim_{\mathrm{D}} \tau^p_* p \otimes \tau^q_* q$. If $\otimes$ respects $\geq_{\mathrm{D}}$ in $\mathcal{A}(\mathfrak{U})$, then $\otimes$ respects $\geq_{\mathrm{D}}$ in $\mathfrak{U}$.*

*Proof.* We need to show that if $q_0 \geq_{\mathrm{D}} q_1$ then $p \otimes q_0 \geq_{\mathrm{D}} p \otimes q_1$. By assumption, $p \otimes q_0 \sim_{\mathrm{D}} \tau_*^p p \otimes \tau_*^{q_0} q_0$ and $\tau_*^p p \otimes \tau_*^{q_1} q_1 \sim_{\mathrm{D}} p \otimes q_1$. Since $\otimes$ respects $\geq_{\mathrm{D}}$ in $\mathcal{A}(\mathfrak{U})$, we obtain $\tau_*^p p \otimes \tau_*^{q_0} q_0 \geq_{\mathrm{D}} \tau_*^p p \otimes \tau_*^{q_1} q_1$, and we are done. $\square$

Note that a map $\tau$ as above induces an inverse of $\tilde{\iota}$.

***A word on $*$-types.*** We will deal with types in a small infinite number of variables, also known in the literature as $*$-types. We define $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U})$ as the quotient of $S_{<\kappa^+}(\mathfrak{U})$ by $\sim_{\mathrm{D}}$. Note that, by padding with realised coordinates and permuting variables, every $\sim_{\mathrm{D}}$-class has a representative with variables indexed by $\kappa$. We leave to the reader easy tasks such as defining the $\alpha$-th power $p^{(\alpha)}$, for $\alpha$ an ordinal, or such as convincing themselves that basic statements such as Fact 1.1 generalise.

Nevertheless, it is not clear if well-definedness of $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ implies well-definedness of $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U})$ (the converse is easy): for instance, at least a priori, one could have a situation where the finitary $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ is well defined, but there are a 1-type $q_0$ and a $\kappa$-type $q_1$ such that $q_0 \geq_{\mathrm{D}} q_1$ but, for some $p$, we have $p \otimes q_0 \not\geq_{\mathrm{D}} p \otimes q_1$. In the rest of the paper we will say, e.g., "$\otimes$ respects $\geq_{\mathrm{D}}$" with the understanding that, whenever $*$-types are involved, this is to be read as "$\otimes$ respects $\geq_{\mathrm{D}}$ on $*$-types".

**Question 1.4.** If $\otimes$ respects $\geq_{\mathrm{D}}$ on finitary types, does $\otimes$ respect $\geq_{\mathrm{D}}$ on $*$-types?

## 2. Orthogonality

**Definition 2.1.** We say that $p, q \in S(A)$ are *weakly orthogonal*, and write $p \perp^{\mathrm{w}} q$, if $p(x) \cup q(y)$ implies a complete $xy$-type over $A$. We say that $p, q \in S^{\mathrm{inv}}(\mathfrak{U})$ are *orthogonal*, and write $p \perp q$, if $(p \mid B) \perp^{\mathrm{w}} (q \mid B)$ for every $B \supseteq \mathfrak{U}$. Two definable sets $\varphi$, $\psi$ are *orthogonal* if for every $n, m \in \omega$, every $p \in S_{\varphi^n}(\mathfrak{U})$ and $q \in S_{\psi^m}(\mathfrak{U})$, we have $p \perp^{\mathrm{w}} q$. Two families of sorts $\mathcal{A}, \mathcal{C}$ are *orthogonal* if every cartesian product of sorts in $\mathcal{A}$ is orthogonal to every cartesian product of sorts in $\mathcal{C}$.

It is easily seen that if $p, q \in S^{\mathrm{inv}}(\mathfrak{U}, M)$ are weakly orthogonal and $\mathfrak{U}_1 \succ \mathfrak{U}$ is $|M|^+$-saturated and $|M|^+$-strongly homogeneous, then $(p \mid \mathfrak{U}_1) \perp^{\mathrm{w}} (q \mid \mathfrak{U}_1)$. This can fail for arbitrary $B \supseteq \mathfrak{U}$, i.e., weak orthogonality is indeed weaker than orthogonality. While this is folklore (Mennuni thanks E. Hrushovski for pointing this out), we could not find any example in print, so we record one.

**Example 2.2.** There is a theory with invariant $p, q$ such that $p \perp^{\mathrm{w}} q$ but $p \not\perp q$.

*Proof.* Let $L$ be a two-sorted language with sorts P, O (points, orders) and a relation symbol $x <_t y$ of arity $\mathrm{P}^2 \times \mathrm{O}$. The class $K$ of finite $L$-structures where, for every $d \in \mathrm{O}$, the relation $x <_d y$ is a linear order, is a (strong) amalgamation class. Let $T$ be the theory of the Fraïssé limit of $K$. Fix a small $M \models T$, and let $p, q$ be the 1-types of sort P defined as $p(x) = \{m <_d x <_d e \mid d \in \mathrm{O}(\mathfrak{U}), m \in M, e \in \mathrm{P}(\mathfrak{U}), e > M\}$ and $q(y) := \{e <_d y \mid d \in \mathrm{O}(\mathfrak{U}), e \in \mathrm{P}(\mathfrak{U})\}$. By quantifier elimination $p$, $q$ are complete, $p$ is $M$-invariant, and $q$ is $\varnothing$-definable, hence $\varnothing$-invariant.

Since $M$ is small, for every $d \in \mathrm{O}(\mathfrak{U})$ it is $<_d$-bounded, hence $p \perp^{\mathrm{w}} q$. Let $b$ be a point of sort O such that $M$ is $\leq_b$-cofinal in $\mathfrak{U}$, and set $B := \mathfrak{U}b$. Then $(q(y) \mid B) \vdash y \geq_b \mathrm{P}(\mathfrak{U})$ and $(p(x) \mid B) \vdash x \geq_b \mathrm{P}(\mathfrak{U})$, and both $x <_b y$ and $y <_b x$ are consistent with $(p(x) \mid B) \cup (q(y) \mid B)$, which is therefore not complete. $\square$

**Remark 2.3.** If $p \in S(A)$ is such that $p \perp^{\mathrm{w}} p$, then $p$ is realised in $\mathrm{dcl}(A)$. If $p, q \in S^{\mathrm{inv}}(\mathfrak{U})$ and $p \perp^{\mathrm{w}} q$, then $p(x) \otimes q(y) = q(y) \otimes p(x)$: they both coincide with (the unique completion of) $p(x) \cup q(y)$. Two definable sets $\varphi$, $\psi$ are orthogonal if and only if every definable subset of $\varphi^m(x) \wedge \psi^n(y)$ can be defined by a finite disjunction of formulas of the form $\theta(x) \wedge \eta(y)$. If two $M$-definable sets are orthogonal, then the definition of orthogonality still holds after replacing $\mathfrak{U}$ with $M$. Adding imaginaries preserves orthogonality, in the following sense. Let $\mathcal{A}$ be a family of sorts, and let $\tilde{\mathcal{A}}$ be a larger family, consisting of $\mathcal{A}$ together with imaginary sorts obtained as definable quotients of products of elements of $\mathcal{A}$. Let $\tilde{\mathcal{C}}$ be obtained similarly from another family of sorts $\mathcal{C}$. If $\mathcal{A}$ and $\mathcal{C}$ are orthogonal, then so are $\tilde{\mathcal{A}}$ and $\tilde{\mathcal{C}}$.

By [22, Proposition 3.13], if $p_0 \geq_{\mathrm{D}} p_1$ and $p_0 \perp^{\mathrm{w}} q$, then $p_1 \perp^{\mathrm{w}} q$. In particular, if $p_0 \geq_{\mathrm{D}} q$ and $p_0 \perp^{\mathrm{w}} q$, then $q$ is realised. As a consequence, $\perp^{\mathrm{w}}$ induces a well-defined relation on the domination poset, which we may expand to $(\widetilde{\mathrm{Inv}}(\mathfrak{U}), \geq_{\mathrm{D}}, \perp^{\mathrm{w}})$. By [22, Proposition 2.3.31] the map $\tilde{\iota}$ from Fact 1.1 is a homomorphism for both $\perp^{\mathrm{w}}$ and $\not\perp^{\mathrm{w}}$. We prove the analogous statements for orthogonality.

**Proposition 2.4.** *Let $p_0, p_1, q \in S^{\mathrm{inv}}(\mathfrak{U})$. If $p_0 \perp q$ and $p_0 \geq_{\mathrm{D}} p_1$, then $p_1 \perp q$. In particular, $\perp$ induces a well-defined relation on $\widetilde{\mathrm{Inv}}(\mathfrak{U})$.*

*Proof.* Fix $r$ witnessing $p_0 \geq_{\mathrm{D}} p_1$ and let $B \supseteq \mathfrak{U}$. Let $b \vDash p_1 \mid B$ and $c \vDash q \mid B$. By [22, Lemma 1.13], $(p_0 \mid B) \cup r \vdash (p_1 \mid B)$. Let $a$ be such that $ab \vDash (p_0 \mid B) \cup r$. Since $p_0 \perp q$, we have $(p_0 \mid B) \perp^{\mathrm{w}} (q \mid B)$, and hence $a \vDash p_0 \mid Bc$. Again by [22, Lemma 1.13] we have $(p_0 \mid Bc) \cup r \vdash (p_1 \mid Bc)$, therefore $b \vDash p_1 \mid Bc$. $\square$

**Proposition 2.5.** *In the setting of Fact 1.1, $\iota \restriction S^{\mathrm{inv}}_{\mathcal{A}^{<\omega}}(\mathcal{A}(\mathfrak{U}))$ is a $\perp$-homomorphism and a $\not\perp$-homomorphism, and so is the induced map $\tilde{\iota} : \widetilde{\mathrm{Inv}}(\mathcal{A}(\mathfrak{U})) \hookrightarrow \widetilde{\mathrm{Inv}}(\mathfrak{U})$.*

*Proof.* Let $p, q \in S^{\mathrm{inv}}_{\mathcal{A}^{<\omega}}(\mathcal{A}(\mathfrak{U}))$ be orthogonal and let $\mathfrak{U}_1 \succ \mathfrak{U}$ be $|\mathfrak{U}|^+$-saturated and $|\mathfrak{U}|^+$-strongly homogeneous. We show that, for $\varphi(x, y, z) \in L(\mathfrak{U})$ and $d \in \mathfrak{U}_1$, if $(\iota p(x) \otimes \iota q(y)) \mid \mathfrak{U}_1 \vdash \varphi(x, y, d)$ then $(\iota p \mid \mathfrak{U}d)(x) \cup (\iota q \mid \mathfrak{U}d)(y) \vdash \varphi(x, y, d)$. By full embeddedness, there are $\chi(x, y, w) \in L_{\mathcal{A}}(\mathcal{A}(\mathfrak{U}))$ and $e \in \mathcal{A}(\mathfrak{U}_1)$ such that $\mathfrak{U}_1 \vDash \forall x, y\ (\chi(x, y, e) \leftrightarrow \varphi(x, y, d))$. Because $(p \mid \mathcal{A}(\mathfrak{U})e) \perp^{\mathrm{w}} (q \mid \mathcal{A}(\mathfrak{U})e)$, there are $\theta_p(x, w), \theta_q(y, w) \in L_{\mathcal{A}}(\mathcal{A}(\mathfrak{U}))$ such that $(p \mid \mathcal{A}(\mathfrak{U})e) \vdash \theta_p(x, e)$, $(q \mid \mathcal{A}(\mathfrak{U})e) \vdash \theta_q(y, e)$, and $\mathcal{A}(\mathfrak{U}_1) \vDash \forall x, y\ ((\theta_p(x, e) \wedge \theta_q(y, e)) \to \chi(x, y, e))$. By invariance of $p, q$, we have

$$\pi_p(x) := \{\theta_p(x, e') \mid e' \in \mathfrak{U}_1, e \equiv_{\mathfrak{U}d} e'\} \subseteq \iota p \mid \mathfrak{U}_1,$$

$$\pi_q(y) := \{\theta_q(y, e') \mid e' \in \mathfrak{U}_1, e \equiv_{\mathfrak{U}d} e'\} \subseteq \iota q \mid \mathfrak{U}_1.$$

So $\pi_p, \pi_q$ are consistent. As $\operatorname{Aut}(\mathfrak{U}_1/\mathfrak{U}d)$ fixes them, they are equivalent to partial types $\sigma_p, \sigma_q$ over $\mathfrak{U}d$. But $\sigma_p \subseteq \iota p \mid \mathfrak{U}d$, $\sigma_q \subseteq \iota q \mid \mathfrak{U}d$, and $\sigma_p(x) \cup \sigma_q(y) \vdash \varphi(x, y, d)$, proving that $\perp$ is preserved.

Suppose there is $B$ with $\mathcal{A}(\mathfrak{U}) \subseteq B \subseteq \mathcal{A}(\mathfrak{U}_1)$ such that $(p \mid B) \not\perp^{\mathrm{w}} (q \mid B)$. By Remark 1.2, this yields $(\iota p \mid \mathfrak{U}B) \not\perp^{\mathrm{w}} (\iota q \mid \mathfrak{U}B)$, proving that $\not\perp$ is preserved as well.

The statement for $\tilde{\iota}$ follows from Proposition 2.4. □

**Lemma 2.6.** *Let $p, q_0, q_1 \in S(\mathfrak{U})$, with $p \perp^{\mathrm{w}} q_0$ and $(p(x) \cup q_0(y)) \geq_{\mathrm{D}} q_1(z)$, witnessed by $r \in S_{p\otimes q_0, q_1}(M)$. If $(r \upharpoonright x) \perp^{\mathrm{w}} (r \upharpoonright yz)$, then $q_0 \geq_{\mathrm{D}} q_1$, witnessed by $r \upharpoonright yz$. Hence, if $\mathcal{A}$, $\mathcal{C}$ are orthogonal families of sorts, $p \in S^{\mathrm{inv}}_{\mathcal{A}^{<\omega}}(\mathfrak{U})$, and $q_0, q_1 \in S^{\mathrm{inv}}_{\mathcal{C}^{<\omega}}(\mathfrak{U})$, if $(p \cup q_0) \geq_{\mathrm{D}} q_1$, then $q_0 \geq_{\mathrm{D}} q_1$.*

*Proof.* Routine, left to the reader. □

Recall that the product $\prod_{i\in I} P_i$ of a family of posets $(P_i, \leq_i)_{i\in I}$ is the cartesian product of the $P_i$ partially ordered by $(p_i)_{i\in I} \leq (q_i)_{i\in I}$ if $\forall i \in I \;\; p_i \leq_i q_i$.

**Corollary 2.7.** *Suppose that $\mathcal{A}$, $\mathcal{C}$ are orthogonal, fully embedded families of sorts. Assume that for every $p \in S^{\mathrm{inv}}(\mathfrak{U})$ there are some $p_{\mathcal{A}} \in S^{\mathrm{inv}}_{\mathcal{A}^{<\omega}}(\mathfrak{U})$ and $p_{\mathcal{C}} \in S^{\mathrm{inv}}_{\mathcal{C}^{<\omega}}(\mathfrak{U})$ such that $p \sim_{\mathrm{D}} p_{\mathcal{A}} \cup p_{\mathcal{C}}$. Then the map $[\![p]\!] \mapsto ([\![p_{\mathcal{A}}]\!], [\![p_{\mathcal{C}}]\!])$ is an isomorphism of posets $\widetilde{\operatorname{Inv}}(\mathfrak{U}) \to \widetilde{\operatorname{Inv}}(\mathcal{A}(\mathfrak{U})) \times \widetilde{\operatorname{Inv}}(\mathcal{C}(\mathfrak{U}))$. Moreover, if $\otimes$ respects $\geq_{\mathrm{D}}$ in $\mathfrak{U}$ (hence also in $\mathcal{A}(\mathfrak{U})$, $\mathcal{C}(\mathfrak{U})$), then this is also an isomorphism of monoids.*

*Proof.* Fact 1.1 yields embeddings of posets

$$\widetilde{\operatorname{Inv}}(\mathcal{A}(\mathfrak{U})) \hookrightarrow \widetilde{\operatorname{Inv}}(\mathfrak{U}) \quad \text{and} \quad \widetilde{\operatorname{Inv}}(\mathcal{C}(\mathfrak{U})) \hookrightarrow \widetilde{\operatorname{Inv}}(\mathfrak{U}).$$

We define a morphism of posets $\widetilde{\operatorname{Inv}}(\mathcal{A}(\mathfrak{U})) \times \widetilde{\operatorname{Inv}}(\mathcal{C}(\mathfrak{U})) \to \widetilde{\operatorname{Inv}}(\mathfrak{U})$ by setting $([\![p(x)]\!], [\![q(y)]\!]) \mapsto ([\![p(x) \cup q(y)]\!])$. It follows from orthogonality of $\mathcal{A}$ and $\mathcal{C}$ that this morphism is well defined: if $p' \sim_{\mathrm{D}} p$ and $q' \sim_{\mathrm{D}} q$, by just taking unions of domination witnesses we find that $p \cup q \sim_{\mathrm{D}} p' \cup q'$. As this map is injective by Lemma 2.6, it is enough to show that the natural candidate for its inverse, $[\![p]\!] \mapsto ([\![p_{\mathcal{A}}]\!], [\![p_{\mathcal{C}}]\!])$, is well defined and a morphism of posets. Both these statements follow from the observation that, if $(p_{\mathcal{A}} \cup p_{\mathcal{C}}) \sim_{\mathrm{D}} p \geq_{\mathrm{D}} q \sim_{\mathrm{D}} (q_{\mathcal{A}} \cup q_{\mathcal{C}})$, then by Lemma 2.6 we must have $p_{\mathcal{A}} \geq_{\mathrm{D}} q_{\mathcal{A}}$ and $p_{\mathcal{C}} \geq_{\mathrm{D}} q_{\mathcal{C}}$. The "moreover" part follows from Fact 1.1, and the fact that $\mathcal{AC}$ is fully embedded. □

**Example 2.8.** Let $A$, $C$ be structures in disjoint languages, $T$ the theory of their disjoint union, in families of sorts $\mathcal{A}, \mathcal{C}$. Then $\mathcal{A}$ and $\mathcal{C}$ are orthogonal, and invariant types in $\mathcal{A}$ are orthogonal to those in $\mathcal{C}$. Therefore, $\widetilde{\operatorname{Inv}}(\mathfrak{U})$ is isomorphic to $\widetilde{\operatorname{Inv}}(\mathcal{A}(\mathfrak{U})) \times \widetilde{\operatorname{Inv}}(\mathcal{C}(\mathfrak{U}))$, and is well defined as a monoid if and only if both factors are.

Orthogonality is preserved by the Morley product. The proof is folklore, and essentially the same as in the stable case, but we record it here for convenience.

**Proposition 2.9.** *If $p_0, p_1 \in S^{\mathrm{inv}}(\mathfrak{U})$ are orthogonal to $q$, then so is $p_0 \otimes p_1$.*

*Proof.* Let $ab \models p_0 \otimes p_1$ and $c \models q$. Because $p_1 \perp q$ we have $c \models q \mid \mathfrak{U}b$, and by definition of $\otimes$ we have $a \models p_0 \mid \mathfrak{U}b$. Since $p_0 \perp q$, this entails $c \models q \mid \mathfrak{U}ab$. □

## 3. Regular ordered abelian groups

In this section we study the domination monoid in certain theories of (linearly) ordered abelian groups, henceforth *oags*. Model-theoretically, the simplest oags are the (nontrivial) divisible ones. Their theory is o-minimal and their domination monoid was one of the first ones to be computed [17; 23]. It is isomorphic to the finite powerset semilattice $(\mathscr{P}_{<\omega}(\mathrm{CS}^{\mathrm{inv}}(\mathfrak{U})), \cup, \subseteq)$ of the set of invariant convex subgroups of $\mathfrak{U}$, and weakly orthogonal classes of types correspond to disjoint finite sets. Divisible oags eliminate quantifiers in the language $L_{\mathrm{oag}} := \{+, 0, -, <\}$. In this section we compute the domination monoid in the next simplest case.

**Definition 3.1.** A (nontrivial) oag is *discrete* if it has a minimum positive element, and *dense* otherwise. We view an oag $M$ as a structure in the *Presburger language* $L_{\mathrm{Pres}} := \{+, 0, -, <, 1, \equiv_n \mid n \in \omega\}$ by interpreting $+, 0, -, <$ in the natural way, 1 as the minimum positive element if $M$ is discrete and as 0 otherwise, and $\equiv_n$ as congruence modulo $nM$. An oag is *regular* if it eliminates quantifiers in $L_{\mathrm{Pres}}$.

**Fact 3.2** [7; 8; 27; 33; 34]. For an oag $M$, the following are equivalent.

(1) $M$ is regular.

(2) The only definable convex subgroups of $M$ are $\{0\}$ and $M$.

(3) The theory of $M$ has an archimedean model.

(4) For every $n > 1$, if the interval $[a, b]$ contains at least $n$ elements, then it contains an element divisible by $n$.

(5) Every quotient of $M$ by a nontrivial convex subgroup is divisible.

**Fact 3.3** [27; 34]. Every discrete regular $M$ is a model of *Presburger Arithmetic*, i.e., $M \equiv \mathbb{Z}$. If $M, N$ are dense regular, then $M \equiv N$ if and only if, for each $\mathfrak{p} \in \mathbb{P}$, either $M/\mathfrak{p}M$ and $N/\mathfrak{p}N$ are both infinite or they have the same finite size.

**Notation 3.4.** For the rest of the section we adopt the following (not entirely standard) conventions. Let $M$ be an oag and $A \subseteq M$. We denote by $A_{>0}$ the set $\{a \in A \mid a > 0\}$, by $\langle A\rangle$ the group generated by $A$, and by $\operatorname{div}(M)$ the divisible hull of $M$. We allow intervals to have endpoints in the divisible hull. In other words, an *interval* in $M$ is a set of the form $\{x \in M \mid a \sqsubset_0 x \sqsubset_1 b\}$, for suitable $a, b \in \operatorname{div}(M) \cup \{\pm\infty\}$ and $\{\sqsubset_0, \sqsubset_1\} \subseteq \{<, \leq\}$.

A *cut* $(L, R)$ is given by subsets $L, R \subseteq M$ such that $L \leq R$ and $L \cup R = M$. We call such a cut *realised* if $L \cap R \neq \varnothing$, and *nonrealised* otherwise. The *cut* $(L, R)$ of $c \in N > M$ is given by $L = \{m \in M \mid c \geq m\}$ and $R = \{m \in M \mid c \leq m\}$. The *cut* of a type $p \in S_1(M)$ is the cut of any $c \models p$. We say that $c \in N > M$ *fills* a cut $(L, R)$

if the latter equals the cut of $c$. For $a \in M$, we denote by $a^+$ the cut $(L, R)$ with $L = \{m \in M \mid m \le a\}$ and $R = \{m \in M \mid a < m\}$, and similarly for $a^-$. Analogous notions are defined for $a \in \operatorname{div}(M)$.

Every interval is definable: e.g., $(a/n, +\infty)$ is defined by $a < n \cdot x$. If $(L, R)$ is a cut then $|L \cap R| \le 1$. A type is realised if and only if its cut is. Let $L_{\mathrm{ab}} := \{0, +, -\}$.

**Remark 3.5.** By regularity, a 1-type over $M \vDash T$ is determined by a cut in $M$ and a choice of cosets modulo each $nM$ (if $M/nM$ is infinite a type may say that the coset $x + nM$ is not represented in $M$) consistent with the $L_{\mathrm{ab}}$-theory of $M$.

**Lemma 3.6.** *If $M$ is a dense regular oag then, for every $n > 0$, every coset of $nM$ is dense in $M$. In particular, given any nonrealised $p \in S_1(M)$, and any nonrealised $q_0 \in S_1(M \restriction L_{\mathrm{ab}})$, there is $q \in S_1(M)$ restricting to $q_0$ and in the same cut as $p$.*

*Proof.* By density and point (4) of Fact 3.2, every $nM$ is dense; as translations are homeomorphisms for the order topology, each coset of $nM$ is dense. □

***Imaginaries in regular ordered abelian groups.*** The first step to compute $\widetilde{\operatorname{Inv}}(\mathfrak{U}^{\mathrm{eq}})$ is to take care of the reduct to a certain fully embedded family of imaginary sorts, that suffice for weak elimination of imaginaries by a result of Vicaría [32]. Recall that $T$ has *weak elimination of imaginaries* if for every imaginary $e$ there is a real tuple $a$ such that $e \in \operatorname{dcl}^{\mathrm{eq}}(a)$ and $a \in \operatorname{acl}^{\mathrm{eq}}(e)$. For $\mathfrak{p} \in \mathbb{P}$ and $n \ge 1$, define $T_{\mathfrak{p}^n}$ as the $L_{\mathrm{ab}}$-theory of $\bigoplus_{i \in \omega} \mathbb{Z}/\mathfrak{p}^n\mathbb{Z}$. The following is well known.

**Fact 3.7.** (1) Let $A$ be an infinite abelian group. Then $A \vDash T_{\mathfrak{p}^n}$ if and only if $\mathfrak{p}A = \{a \in A \mid \mathfrak{p}^{n-1}a = 0\}$.

(2) $T_{\mathfrak{p}^n}$ has quantifier elimination and is totally categorical.

(3) If $A \vDash T_{\mathfrak{p}^n}$, then $\mathfrak{p}A$ is a model of $T_{\mathfrak{p}^{n-1}}$, and the induced structure on $\mathfrak{p}A$ is that of a pure abelian group.

(4) $T_{\mathfrak{p}^n}$ has weak elimination of imaginaries.

*Proof sketch.* For (4), as $T_{\mathfrak{p}^n}$ is stable, it suffices to show that canonical bases of types over models are interdefinable with real tuples [13, Proposition 3]. This is an application of the elementary divisor theorem, and is left to the reader. □

Let $T_{\mathfrak{p}^\infty}$ be the following multisorted theory:

- For every $n > 0$ there is a sort $\mathrm{Q}_{\mathfrak{p}^n}$, endowed with a copy of $L_{\mathrm{ab}}$.
- For every $n > 0$ there is a function symbol $\rho_{\mathfrak{p}^{n+1}} : \mathrm{Q}_{\mathfrak{p}^{n+1}} \to \mathrm{Q}_{\mathfrak{p}^n}$.
- $M \vDash T_{\mathfrak{p}^\infty}$ if and only if, for all $n > 0$, $\mathrm{Q}_{\mathfrak{p}^n}(M) \vDash T_{\mathfrak{p}^n}$ and $\rho_{\mathfrak{p}^{n+1}} : \mathrm{Q}_{\mathfrak{p}^{n+1}}(M) \to \mathrm{Q}_{\mathfrak{p}^n}(M)$ is a surjective group homomorphism with kernel $\mathfrak{p}^n\, \mathrm{Q}_{\mathfrak{p}^{n+1}}(M)$.

**Remark 3.8.** In an earlier version of this manuscript, we had claimed that $T_{\mathfrak{p}^\infty}$ has quantifier elimination. This does not hold. But one may show that it is enough to add function symbols $\lambda_n : \mathrm{Q}_{\mathfrak{p}^n} \to \mathrm{Q}_{\mathfrak{p}^{n+1}}$ for all $n$, interpreted as the definable group isomorphism $\mathrm{Q}_{\mathfrak{p}^n}(A) \to \mathfrak{p}\,\mathrm{Q}_{\mathfrak{p}^{n+1}}(A)$ mapping $a$ to $\mathfrak{p}\tilde{a}$ where $\tilde{a}$ is any element with $\rho_{\mathfrak{p}^{n+1}}(\tilde{a}) = a$. We thank A. Gehret for having pointed this out to us.

The quantifier elimination result above, which has been mentioned for the sake of completeness, will not be used below. Let $\hat{\kappa}$ be the monoid of cardinals not larger than $\kappa$, with the usual sum and order.

**Corollary 3.9.** (1) *The theory $T_{\mathfrak{p}^\infty}$ is complete, totally categorical,* 1*-based, and has weak elimination of imaginaries.*

(2) *In $T_{\mathfrak{p}^\infty}$, we have $\widetilde{\mathrm{Inv}}(\mathfrak{U}) \cong \mathbb{N}$ and, for each infinite cardinal $\kappa$, the monoid $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U})$ is* (*well defined and*) *isomorphic to $\hat{\kappa}$.*

(3) *More precisely, if* $\mathrm{tp}(a/\mathfrak{U})$ *is $M$-invariant, then there is a basis $b \in \mathrm{dcl}(Ma)$ of the $\mathbb{F}_{\mathfrak{p}}$-vector space $\mathrm{Q}_{\mathfrak{p}}(\mathrm{dcl}(\mathfrak{U}a))$ over $\mathfrak{U}$, and* $\mathrm{tp}(a/\mathfrak{U})$ *is domination-equivalent to* $\mathrm{tp}(b/\mathfrak{U})$, *witnessed by* $\mathrm{tp}(ab/M)$ *in both directions, and the isomorphism above sends its domination-equivalence class to the cardinality of $b$.*

*Proof.* Statement (1) is immediate from Fact 3.7 and the fact that abelian groups are 1-based. As for (2), each of the sorts $\mathrm{Q}_{\mathfrak{p}^n}$ is stable unidimensional, that is, if $p \perp q$ then one of $p, q$ is algebraic, and it follows easily that so is $T_{\mathfrak{p}^\infty}$. The conclusion for finitary types then follows from [22, Corollary 5.19], and the version for $*$-types is similar.

To prove (3), if $b \in \mathrm{dcl}(Ma)$ is a basis of $\mathrm{Q}_{\mathfrak{p}}(\mathrm{dcl}(Ma))$ over $M$, by $M$-invariance it is also a basis of $\mathrm{Q}_{\mathfrak{p}}(\mathrm{dcl}(\mathfrak{U}a))$ over $\mathfrak{U}$. Because in unidimensional theories the domination-equivalence class of a tuple is determined by its weight [22, Remark 5.12], it suffices to show that the cardinality $\kappa$ of $b$ equals the weight $w(\mathrm{tp}(a/\mathfrak{U}))$. For $T_{\mathfrak{p}^n}$ this is well known, and as $\mathrm{Q}_{\mathfrak{p}^n}(\mathfrak{U})$ is a fully embedded model of $T_{\mathfrak{p}^n}$, the result is easily seen to transfer to $T_{\mathfrak{p}^\infty}$. □

We now consider a regular oag $M$. Since it is well known that Presburger arithmetic eliminates imaginaries (by definable choice), we may assume that $M$ is dense.

We view $M$ as a structure in the language with one sort for the oag itself, endowed with $L_{\mathrm{oag}}$, one sort $\mathrm{Q}_{\mathfrak{p}^n}$ for each prime $\mathfrak{p}$ and each $n > 0$, endowed with $L_{\mathrm{ab}}$ and interpreted as the group $M/\mathfrak{p}^n M$, functions $\pi_{\mathfrak{p}^n}$ for the quotient map from $M$ to $M/\mathfrak{p}^n M$ and functions $\rho_{\mathfrak{p}^{n+1}}$ for the canonical surjections $M/\mathfrak{p}^{n+1}M \to M/\mathfrak{p}^n M$. Moreover, for every prime $\mathfrak{p}$ we definably expand the language on $(\mathrm{Q}_{\mathfrak{p}^n})_{n>0}$ so that the multisorted structure $(\mathrm{Q}_{\mathfrak{p}^n}(M))_{n>0}$ has quantifier elimination.

For every $\mathfrak{p} \in \mathbb{P}$, let $d_{\mathfrak{p}} \in \mathbb{N} \cup \{\infty\}$ be such that $(M : \mathfrak{p}M) = \mathfrak{p}^{d_{\mathfrak{p}}}$. Set $T := \mathrm{Th}(M)$. The proof of the following lemma is straightforward from Lemma 3.6 and quantifier elimination for the one-sorted theory of $M$ in $L_{\mathrm{Pres}}$, and we leave it to the reader.

**Lemma 3.10.** *The theory $T$ eliminates quantifiers. For $\mathfrak{U} \vDash T$, the following holds. For every $\mathfrak{p}$ prime and $n > 0$, the sort $\mathrm{Q}_{\mathfrak{p}^n}(\mathfrak{U})$ equipped with the natural $L_{\mathrm{ab}}$-structure* *is fully embedded. If* $d_{\mathfrak{p}} = \infty$, *the structure given by* $(\mathrm{Q}_{\mathfrak{p}^n}(\mathfrak{U}))_{n>0}$, *together with the maps* $\rho_{\mathfrak{p}^{n+1}}$ *and the natural* $L_{\mathrm{ab}}$*-structure on each sort, is fully embedded and a model of* $T_{\mathfrak{p}^\infty}$*. If* $d_{\mathfrak{p}}$ *is finite, every sort* $\mathrm{Q}_{\mathfrak{p}^n}(\mathfrak{U})$ *is finite. If* $\mathfrak{p}$, $\mathfrak{q}$ *are distinct primes, then* $\mathrm{Q}_{\mathfrak{p}^n}(\mathfrak{U})_{n>0}$ *and* $\mathrm{Q}_{\mathfrak{q}^n}(\mathfrak{U})_{n>0}$ *are orthogonal.*

**Definition 3.11.** Denote by $\mathcal{Q}$ the family of sorts $\{\mathrm{Q}_{\mathfrak{p}^n} \mid \mathfrak{p} \in \mathbb{P}, n > 0\}$. If $q = \mathrm{tp}(c/\mathfrak{U})$ is a $*$-type, possibly with coordinates in the sorts in $\mathcal{Q}$, for each $\mathfrak{p} \in \mathbb{P}$, let $\kappa_{\mathfrak{p}}(q)$ be the dimension of the $\mathbb{F}_{\mathfrak{p}}$-vector space $\mathrm{dcl}(\mathfrak{U}c)/\mathfrak{p}(\mathrm{dcl}(\mathfrak{U}c))$ over $\mathfrak{U}/\mathfrak{p}\mathfrak{U}$. Let $\mathbb{P}_T$ be the set of primes $\mathfrak{p}$ such that if $M \vDash T$ then $\mathfrak{p}M$ has infinite index, and denote by $\prod_{\mathbb{P}_T} \hat{\kappa}$ the monoid of $\mathbb{P}_T$-indexed sequences of cardinals smaller or equal than $\kappa$ with pointwise cardinal sum, equipped with the product (partial) order.

**Corollary 3.12.** *The family of sorts* $\mathcal{Q}$, *equipped with the* $L_{\mathrm{ab}}$*-structure on each sort and the maps* $\rho_{\mathfrak{p}^{n+1}}$, *is fully embedded. When viewed as a standalone structure,* $\otimes$ *respects* $\geq_{\mathrm{D}}$ *and* $\widetilde{\mathrm{Inv}}_\kappa(\mathcal{Q}(\mathfrak{U})) \cong \prod_{\mathbb{P}_T} \hat{\kappa}$.

*Proof.* This follows from Lemma 3.10, Corollary 3.9, and Fact 1.1. Compatibility of $\otimes$ with $\geq_{\mathrm{D}}$ is a consequence of stability, see [22, Propositions 1.21 and 1.25]. □

**Fact 3.13** [32, Theorem 5.1]. The theory $T$ has weak elimination of imaginaries.

**Remark 3.14.** Vicaría [32] proves a more general result, of which Fact 3.13 is a special case. Note that she adds sorts for quotients of the form $M/nM$ for all $n > 0$. As $M/nM$ is definably isomorphic to $\prod_{i=1}^m M/\mathfrak{p}_i^{n_i}M$, where $n = \prod_{i=1}^m \mathfrak{p}_i^{n_i}$ is the decomposition of $n$ into prime powers, it suffices to add the sorts $\mathrm{Q}_{\mathfrak{p}^n}$.

Observe that, for the above to go through, we need to have in our language the sorts $\mathrm{Q}_{\mathfrak{p}^n}$ even when they are finite. Alternatively, one may dispense with the finite $\mathrm{Q}_{\mathfrak{p}^n}$ by naming enough constants, e.g., by naming a model.

***Moving to the right of a convex subgroup.***

**Assumption 3.15.** Until the end of the section, $T$ is the complete $L_{\mathrm{Pres}}$-theory of a regular oag. Imaginary sorts are not in our language until further notice.

**Definition 3.16.** Let $B \subseteq M$. A cut $(L, R)$ is *right of* $B$ if $R \cap B = \varnothing$ and $B$ is cofinal in $L$. An element $c \in N > M$ is *right of* $B$ if its cut is, and a type $q \in S_1(M)$ is *right of* $B$ if any of its realisations is. A convex subgroup $H$ of $\mathfrak{U}$ is called $(A$-$)$*invariant* if there is an $(A$-$)$invariant type to its right.

**Remark 3.17.** Let $p \in S_1(\mathfrak{U})$ be an $M$-invariant type. If its cut $(L, R)$ is definable, then it is $M$-definable. If not, then exactly one between the cofinality of $L$ and the coinitiality of $R$ is small, and $M$ contains a set cofinal in $L$ or coinitial in $R$.

*Proof.* The case of a definable cut is clear, so let us assume $(L, R)$ is a nondefinable cut of $\mathfrak{U}$. In particular, $L \neq \varnothing \neq R$. If $L \cap M$ is not cofinal in $L$, there is $\ell \in L$ with $L \cap M < \ell$, so by regularity of $\mathfrak{U}$ and saturation there is $\ell_0 \in L$ divisible by all

$n \geq 1$ such that $L \cap M < \ell_0 < \ell$. Similarly, if $R \cap M$ is not coinitial in $M$ there is $r_0 \in R$, which is divisible by all $n \geq 1$, such that $r_0 < R \cap M$. By Remark 3.5 it follows that $\operatorname{tp}(\ell_0/M) = \operatorname{tp}(r_0/M)$, showing that $(L, R)$ is not $M$-invariant. □

In particular, in a regular oag a nontrivial convex subgroup $H$ of $\mathfrak{U}$ is invariant if and only if the cofinality of $H$ or the coinitiality of $(\mathfrak{U} \setminus H)_{>0}$ is small, while the trivial subgroup $\{0\}$ is invariant if and only if $\mathfrak{U}$ is dense.

**Lemma 3.18.** *In the theory of a regular oag, suppose that $p \in S_1^{\mathrm{inv}}(\mathfrak{U})$ and $f$ is a definable function such that $f_* p$ is not realised. Then $p \sim_{\mathrm{D}} f_* p$.*

*Proof.* Clearly $p \geq_{\mathrm{D}} f_* p$. By [7, Corollary 1.10], $f$ is piecewise affine. As $f_* p$ is not realised, $f$ cannot be constant at $p$, so it is invertible at $p$ and

$$f_* p \geq_{\mathrm{D}} f_*^{-1}(f_* p) = p.$$ □

**Proposition 3.19.** *In Presburger arithmetic, every invariant 1-type is domination-equivalent to a type right of an invariant convex subgroup.*

*Proof.* By Lemma 3.18 it suffices to show that, for every nonrealised $p \in S_1^{\mathrm{inv}}(\mathfrak{U})$ there is a definable $f$ such that $f_* p$ is right of an invariant convex subgroup. By Fact 3.2, $\mathfrak{U}/\mathbb{Z}$ is divisible, and it is easy to see that $\mathfrak{U}/\mathbb{Z}$ inherits saturation and strong homogeneity from $\mathfrak{U}$. The conclusion follows by lifting the analogous result [17, Corollary 13.11] (see also [23, Proposition 4.8]) from $\mathfrak{U}/\mathbb{Z}$. □

In the rest of the subsection we generalise the above to the regular case.

**Assumption 3.20.** Until the end of the subsection, $M$ denotes a dense regular oag, and $\mathfrak{U}$ a monster model of $T := \operatorname{Th}(M)$.

**Proposition 3.21.** *Let $b \in \mathfrak{U} \setminus M$ be divisible by every $n > 1$ and let $B := \langle Mb \rangle = M + \mathbb{Q}b$. If $M_{>0}$ is coinitial in $B_{>0}$, then $M \prec B \prec \mathfrak{U}$.*

*Proof.* The inclusion $M \subseteq B$ is *pure*, i.e., for every $n > 1$ we have $nB \cap M = nM$. Moreover, if $c = a + \gamma b$, with $a \in M$ and $\gamma \in \mathbb{Q}$, then for every $n$ we clearly have $c - a \in nB$, and hence $B/nB$ may be naturally identified with $M/nM$.

Because $M$ is dense and $M_{>0}$ is coinitial in $B_{>0}$, it follows that $B$ is as well dense. Let $c < d \in B$ and $n > 1$. By assumption, $(0, d - c)$ intersects $M$, so it contains an interval $I$ of $M$, and hence represents all elements of $M/nM$ by Lemma 3.6. These can be identified with the elements of $B/nB$, as observed above, so there is $e \in I$ such that $c + e \in nB$. Clearly, $c + e \in (c, d)$, and hence $B$ is regular by Fact 3.2.

By Fact 3.3 and the identification of $M/nM$ with $B/nB$, we obtain $B \equiv M$. Since $M$ is pure in $B$, it is an $L_{\mathrm{Pres}}$-substructure of $B$, and the conclusion follows by quantifier elimination in $L_{\mathrm{Pres}}$. □

Recall that an extension $A < B$ of oags is an *i-extension* if there is no $b \in B_{>0}$ such that the set $\{a \in A \mid a < b\}$ is closed under sum.

**Lemma 3.22.** *Let $H < M < N$, with $M$ dense regular and $H$ convex. The set of elements of $N$ right of $H$ is closed under sum. In particular, $N$ is an i-extension if and only if $H \mapsto H \cap M$ is a bijection between the convex subgroups of $N$ and $M$.*

*Proof.* If $H = M$, the statement is trivial. If $H = \{0\}$, let $0 < c, d < M_{>0}$ and pick $a \in M_{>0}$. By density, there is $b \in M$ with $0 < b < a$, and since $b$ and $a - b$ are both in $M_{>0}$ we conclude $c + d < b + a - b = a$. If $H$ is proper nontrivial, by Fact 3.2 the quotient $M/H$ is divisible, and the conclusion follows from the previous case applied to $M/H$ as a subgroup of the quotient of $N$ by the convex hull of $H$. □

**Proposition 3.23.** *Every $M \vDash T$ has a maximal elementary i-extension.*

*Proof.* This is easy, see, e.g., [21, Proposition 4.2.17]. □

**Proposition 3.24.** *Suppose $M \vDash T$ has no proper elementary i-extension and let $p \in S_1(M)$ be nonrealised. Then there are $a \in M$ and $\beta \in \mathbb{Z} \setminus \{0\}$ such that, if $f(t) = a + \beta t$, then the pushforward $f_* p$ is right of a convex subgroup.*

*Proof.* Let $b \vDash p$, and suppose first that $b$ is divisible by every $n$. Consider $B := \langle Mb \rangle = M + \mathbb{Q}b$. If there are $a' \in M$ and $\beta' \in \mathbb{Q}$ such that $0 < a' + \beta' b < M_{>0}$, by Lemma 3.22 multiplying by the denominator of $\beta'$ yields a positive element smaller than $M_{>0}$, so we obtain the conclusion with the convex subgroup $\{0\}$. If instead there is no such $a' + \beta' b$, then $M_{>0}$ is coinitial in $B_{>0}$, and by Proposition 3.21 $B \succ M$. By maximality of $M$, there must be convex subgroups $H_0 \subsetneq H_1$ of $B$ such that $H_0 \cap M = H_1 \cap M$. Hence any positive $a + \beta b \in H_1 \setminus H_0$ is right of $H_0 \cap M$. We conclude again by clearing the denominator of $\beta$ and using Lemma 3.22.

This shows the conclusion when $b$ is divisible by all $n$. In the general case, by Lemma 3.6, there is $c \in \mathfrak{U}$ with the same cut in $M$ as $b$ which is divisible by every $n$. As we just proved, there is $f(t) := a + \beta t$, with $\beta \in \mathbb{Z}$ and $a \in M$, such that the cut of $f(c)$ in $M$ is that of a convex subgroup. Because $f(t)$ sends intervals to intervals, it sends cuts to cuts, and hence the cut of $f(b)$ equals that of $f(c)$. □

**Corollary 3.25.** *For every nonrealised $p(x) \in S_1^{\mathrm{inv}}(\mathfrak{U})$ there is a definable function $f$ such that $(f_* p)(y)$ is right of an invariant convex subgroup, and domination-equivalent to $p$, witnessed by any small type containing $y = f(x)$.*

*Proof.* If $p$ is $M$-invariant, up to enlarging $M$ we may assume that it has no proper elementary i-extension. Let $f(t)$ be an $M$-definable function given by Proposition 3.24 applied to $p \upharpoonright M$. Then $f_* p$ is $M$-invariant, and its cut is either the one to the left of $(f_* p \upharpoonright M)(\mathfrak{U})$ or the one to its right, which are both cuts right of convex subgroups of $\mathfrak{U}$ by Lemma 3.22. Now apply Lemma 3.18. □

***Computing the domination monoid.*** By Fact 3.13, regular oags weakly eliminate imaginaries after adding the sorts $\mathrm{Q}_{\mathrm{p}^n}$. As already remarked, this implies that passing to $T^{\mathrm{eq}}$ does not affect the poset $\widetilde{\mathrm{Inv}}(\mathfrak{U})$, nor its well-definedness as a monoid. Hence, we will conflate the two settings, and refer to our theory in this language as $T^{\mathrm{eq}}$, reserving $T$ for the 1-sorted $L_{\mathrm{Pres}}$-theory of a regular oag.

**Assumption 3.26.** Until the end of the section, we work in $T^{\mathrm{eq}}$.

**Lemma 3.27.** *Let $H_0 \subsetneq H_1$ be convex subgroups of $M \vDash T$ and, for $i < 2$, let $q_i(x^i) \in S_1(M)$ be right of $H_i$. Suppose that there is no prime $\mathfrak{p} \in \mathbb{P}$ such that both $q_i(x^i)$ prove that $x^i$ is in a new coset modulo some $\mathfrak{p}^{\ell_i}$ Then $q_0 \perp^{\mathrm{w}} q_1$.*

*Proof.* By Lemma 3.22 the cut of every $k_0x^0+k_1x^1$ is determined by $q_0(x^0) \cup q_1(x^1)$, and we conclude by assumption and quantifier elimination. □

**Proposition 3.28.** *Suppose that $q_H(x) \in S_1^{\mathrm{inv}}(\mathfrak{U})$ is right of the convex subgroup $H$ and prescribes realised cosets modulo every $n$ for $x$. For an invariant $*$-type $q$ with all coordinates in the home sort, the following are equivalent.*

(1) *For every (equivalently, some) $b \vDash q$, no type right of $H$ is realised in $\langle \mathfrak{U}b \rangle$.*

(2) *$q_H \perp^{\mathrm{w}} q$.*

(3) *$q_H$ commutes with $q$.*

(4) *$q_H \perp q$.*

*Moreover, if $q'$ is a $*$-type with no coordinates in the home sort, then $q_H \perp q'$.*

*Proof.* To show (1) ⇒ (2), consider $q_H(x) \cup q(y)$. By assumption on $q_H$ we only need to deal with inequalities of the form $kx + \sum_{i<|y|} k_i y_i + d \geq 0$, but (1) gives immediately that the cut of $kx$ in $\langle \mathfrak{U}b \rangle$ is determined. If (1) fails, as witnessed by $f(b)$, say, then $q_H(x) \otimes q(y)$ and $q(y) \otimes q_H(x)$ disagree on the formula $f(y) < x$, proving (3) ⇒ (1), and (2) ⇒ (3) holds for every type in every theory.

We prove (2) ⇒ (4), the converse being trivial. Suppose that $B \supseteq \mathfrak{U}$ is such that $(q_H \mid B) \not\perp^{\mathrm{w}} (q \mid B)$. The cosets modulo every $n$ of a realisation of $q_H$ are all realised in $\mathfrak{U}$, so there must be some inequality of the form $kx + \sum_{i<|y|} k_i y_i + d \geq 0$, with $k_i \in \mathbb{Z}$ and $d \in \langle B \rangle$, that is not decided. Hence, if (4) fails, it fails for a 1-type $\tilde{q}$, namely the pushforward of $q$ under the map $y \mapsto \sum_{i<|y|} k_i y_i$. By Corollary 3.25 and Proposition 2.4, we may assume $\tilde{q}$ is right of a convex subgroup. Therefore $q_H(x)$ and $\tilde{q}(z)$ are weakly orthogonal by (2) and [22, Proposition 3.13], to the right of distinct (by weak orthogonality) convex subgroups, but the cut in $\langle B \rangle$ of $kx + z$ is not determined by $(q_H \mid B)(x) \cup (\tilde{q} \mid B)(z)$. This contradicts Lemma 3.22.

Now we consider the "moreover" part. By Proposition 2.4 we may replace $q'$ with any domination-equivalent type, so we may assume, using Corollary 3.12 and Proposition 2.5, that $q'(z)$ is the type of an independent tuple, with $z_i \in \mathrm{Q}_{\mathfrak{p}_i}$. Let $H'$ be any invariant convex subgroup different from $H$, let $p_i$ be the 1-type right of $H'$ in a new coset modulo $\mathfrak{p}_i$ and congruent to 0 modulo every other prime, and let $q$ be the tensor product, in any order, of the $p_i$. Clearly $q \geq_{\mathrm{D}} q'$ and, by construction, if $b \vDash q$ then no type right of $H$ is realised in $\langle \mathfrak{U} b\rangle$, so we conclude by Proposition 2.4. □

**Definition 3.29.** Let $q$ be an invariant global $*$-type, and $c \vDash q$. Let $\mathcal{H}(q)$ be the set of cuts of convex subgroups of $\mathfrak{U}$ filled in $\langle \mathfrak{U} c\rangle$.

**Theorem 3.30.** *If $p, q$ are invariant $*$-types, then $p \geq_{\mathrm{D}} q$ if and only if $\mathcal{H}(p) \supseteq \mathcal{H}(q)$ and $\forall \mathfrak{p} \in \mathbb{P}\ \kappa_{\mathfrak{p}}(p) \geq \kappa_{\mathfrak{p}}(q)$. Hence, $[\![q]\!]$ is determined by $\mathcal{H}(q)$ and $\mathfrak{p} \mapsto \kappa_{\mathfrak{p}}(q)$.*

*Proof.* Let $c \vDash q$, and write $c = c^0 c^1$, with $c^0$ a tuple in the home sort and $c^1$ a tuple from the sorts $\mathrm{Q}_{\mathfrak{p}^n}$. By enlarging $c_1$ with at most $|c|$ points of $\operatorname{dcl}(\mathfrak{U} c)$ if necessary, we may assume that it contains bases of all $\mathbb{F}_{\mathfrak{p}}$-vector spaces $\mathrm{Q}_{\mathfrak{p}}(\operatorname{dcl}(\mathfrak{U} c))$ over $\mathfrak{U}$. Observe that this is harmless domination-wise, and that it does not impact compatibility of $\otimes$ with $\geq_{\mathrm{D}}$ by [22, Proposition 1.23].

Index on a suitable cardinal $\kappa$, bounded by the cardinality of $c^0$, the (necessarily invariant) convex subgroups $H_j$ whose cuts are filled in $\langle \mathfrak{U} c^0\rangle$. Note that, by Corollary 3.25, we have $\kappa \neq 0$ unless $c^0$ is realised.

For $j < \kappa$, let $q_j(y_j)$ be the type right of $H_j$ divisible by every nonzero integer. By Lemma 3.27 and Proposition 3.28, the $q_j$ are orthogonal, and it follows from Proposition 2.9 and compactness that their union is a complete type; call it $q_{\mathrm{H}}(y)$. Let $q_{\mathrm{Q}}(z) := \operatorname{tp}(c_1/\mathfrak{U})$. By Propositions 3.28 and 2.9, $q_{\mathrm{H}} \perp q_{\mathrm{Q}}$. We prove that $q(x)$ is domination-equivalent to $q'(yz) := q_{\mathrm{H}}(y) \otimes q_{\mathrm{Q}}(z)$. If $c^0 \in \mathfrak{U}$, equivalently if $q_{\mathrm{H}}$ is realised, this is trivial, so we assume this is not the case.

To show $q'(yz) \geq_{\mathrm{D}} q(x)$, let $b \in \operatorname{dcl}(\mathfrak{U} c)$ be maximal amongst the tuples with each $b_k$ in the cut of an invariant convex subgroup, and such that if $k < k'$ then $\langle b_k\rangle_{>0} < \langle b_{k'}\rangle_{>0}$. A maximal such $b$ exists because the size of $b$ is at most that of $c^0$, by looking at $\mathbb{Q}$-linear dimension over $\mathfrak{U}$ in the divisible hull. Since $c^0 \notin \mathfrak{U}$, by Corollary 3.25 there is a point of $\operatorname{dcl}(\mathfrak{U} c)$ in the cut of an invariant convex subgroup, and hence $b$ is nonempty. By [7, Corollary 1.10] definable functions are piecewise affine and, by clearing denominators using Lemma 3.22, we may assume that $b \in \langle \mathfrak{U} c^0\rangle$.

Write $b_k = f_k(c^0)$, for suitable affine functions $f_k$. Let $M \prec^+ \mathfrak{U}$ be large enough to contain the parameters of the $f_k$, such that $q$ and $q'$ are $M$-invariant, and such that $M$ has no proper elementary i-extension. Let $r \in S_{qq'}(M)$ contain the following.

(1) For each $k$, by choice of $q'$ there is $j < \kappa$ such that $y_j$ is in the same cut as $b_k$ according to $q'$. If the cut of $b_k$ has small cofinality on the right, put in $r$ the formula $f_k(x) > y_j$; if it has small cofinality on the left, put in $r$ the formula $f_k(x) < y_j$.

(2) For each $j < |c^1|$, the formula $x_{|c^0|+j} = z_j$.

By Lemma 3.10, point (2) above, the fact that $c^1$ contains bases of all $\mathbb{F}_{\mathfrak{p}}$-vector spaces $\mathrm{Q}_{\mathfrak{p}}(\operatorname{dcl}(\mathfrak{U} c))$ over $\mathfrak{U}$, and Corollary 3.9, to prove $q' \geq_{\mathrm{D}} q$ it suffices to show that $q' \cup r$ decides the cut in $\mathfrak{U}$ of every $\sum_i \delta_i x_i$. We first prove a special case.

**Claim.** *$q' \cup r$ entails the quantifier-free $\{+, 0, -, <\}$-type of the $f_k(x)$ over $\mathfrak{U}$.*

*Proof of Claim.* It is enough to show that the cut of every $\sum_k \beta_k f_k(x)$ in $\mathfrak{U}$ is decided, where only finitely many $\beta_k \in \mathbb{Z}$ are nonzero. By choice of $r$ and Remark 3.17, $q' \cup r$ determines the cut of each $f_k(x)$ over $\mathfrak{U}$. Moreover, $r$ contains the information that $\langle f_k(x)\rangle_{>0} < \langle f_{k'}(x)\rangle_{>0}$ for $k < k'$. By this, the fact that the $f_k(x)$ are right of convex subgroups, and Lemma 3.22, the cut of $\sum_k \beta_k f_k(x)$ must be that of $\operatorname{sign}(\beta_k) f_k(x)$, with $k$ the largest such that $\beta_k \neq 0$. □

As $M$ has no proper elementary i-extension, given a term $\sum_i \delta_i x_i$, by Proposition 3.24 we can compose with an $M$-definable injective affine function and reduce to a term $\sum_i \gamma_i x_i + d$, with $d \in M$ and $\gamma_i \in \mathbb{Z}$, with cut in $M$ right of a convex subgroup. As $\operatorname{tp}\bigl(\sum_i \gamma_i x_i + d/\mathfrak{U}\bigr)$ is $M$-invariant, $\sum_i \gamma_i x_i + d$ is in the cut of an $M$-invariant convex subgroup of $\mathfrak{U}$. By maximality of $b$, there must be $k$ and positive integers $n, m$ such that $nb_k \leq m\bigl(\sum_i \gamma_i x_i + d\bigr) \leq (n+1)b_k$. Thus $r \vdash nf_k(x) \leq m\bigl(\sum_i \gamma_i x_i + d\bigr) \leq (n+1) f_k(x)$, and by the Claim $q' \geq_{\mathrm{D}} q$.

Similar arguments show $q \geq_{\mathrm{D}} q'$ and that, if $\mathcal{H}(p) \supseteq \mathcal{H}(q)$ and $\kappa_{\mathfrak{p}}(p) \geq \kappa_{\mathfrak{p}}(q)$ for all $\mathfrak{p} \in \mathbb{P}$, and $p'$ is defined analogously to $q'$, then $p' \geq_{\mathrm{D}} q'$. That $\mathcal{H}(p) \supseteq \mathcal{H}(q)$ is necessary to have $p \geq_{\mathrm{D}} q$ follows from Proposition 3.28 and [22, Proposition 3.13]. As $\forall \mathfrak{p} \in \mathbb{P}\ \ \kappa_{\mathfrak{p}}(p) \geq \kappa_{\mathfrak{p}}(q)$, if for some $\mathfrak{p} \in \mathbb{P}$ we have $\kappa_{\mathfrak{p}}(q) > \kappa_{\mathfrak{p}}(p)$ then we easily find a type in the quotient sorts dominated by $q$ but not by $p$, a contradiction. □

**Proposition 3.31.** *For all invariant $*$-types $p, q$ and $\mathfrak{p} \in \mathbb{P}$, we have*

$$\mathcal{H}(p \otimes q) = \mathcal{H}(p) \cup \mathcal{H}(q) \quad \textit{and} \quad \kappa_{\mathfrak{p}}(p \otimes q) = \kappa_{\mathfrak{p}}(p) + \kappa_{\mathfrak{p}}(q).$$

*Proof.* By Proposition 3.28, $\mathcal{H}(q)$ is precisely the set of convex invariant subgroups $H$ such that $q \not\perp q_H$. By Proposition 2.9, we therefore have the first statement. The second one is an easy consequence of the definition of $\otimes$. □

Note that if $q \in S^{\mathrm{inv}}_{<\kappa^+}(\mathfrak{U})$ then $|\mathcal{H}(q)|$ and each $\kappa_{\mathfrak{p}}(q)$ are at most $\kappa$.

**Definition 3.32.** We denote by $\mathrm{CS}^{\mathrm{inv}}(\mathfrak{U})$ the set of invariant convex subgroups of $\mathfrak{U}$, and by $\mathscr{P}_{\leq\kappa}(\mathrm{CS}^{\mathrm{inv}}(\mathfrak{U}))$ the monoid of its subsets of size at most $\kappa$ with union, partially ordered by inclusion.

**Corollary 3.33** (Theorem D). *For $T$ the theory of a regular oag and $\kappa$ a small infinite cardinal, $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}^{\mathrm{eq}})$ is well defined, and $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}^{\mathrm{eq}}) \cong \mathscr{P}_{\leq\kappa}(\mathrm{CS}^{\mathrm{inv}}(\mathfrak{U})) \times \prod_{\mathbb{P}_T} \hat{\kappa}$.*

*Proof.* Compatibility of $\otimes$ and $\geq_{\mathrm{D}}$ follows from Theorem 3.30 and Proposition 3.31. The same results show that the map $[\![p]\!]$ to $(\mathcal{H}(p), \mathfrak{p} \mapsto \kappa_{\mathfrak{p}}(p))$ is well defined, an embedding of posets, and a morphism of monoids. Surjectivity is easily checked. □

In general, the embedding $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}) \hookrightarrow \widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}^{\mathrm{eq}})$ is not surjective, although its image may be easily computed. We state the result of this computation, which we leave to the reader, and of the analogous ones for finitary types. Denote by $\prod^{\mathrm{bdd}}_{\mathbb{P}_T} \omega$ the submonoid of $\prod_{\mathbb{P}_T} \hat{\omega}$ consisting of bounded sequences of natural numbers.

**Corollary 3.34.** *The monoids* $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U})$, $\widetilde{\mathrm{Inv}}(\mathfrak{U})$, $\widetilde{\mathrm{Inv}}(\mathfrak{U}^{\mathrm{eq}})$ *are all well defined*, *and*

$$\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}) \cong \left(\mathscr{P}_{\leq\kappa}(\mathrm{CS}^{\mathrm{inv}}(\mathfrak{U})) \times \prod_{\mathbb{P}_T} \hat{\kappa}\right) \setminus \{(a,b) \mid a = \varnothing, b \neq 0\},$$

$$\widetilde{\mathrm{Inv}}(\mathfrak{U}^{\mathrm{eq}}) \cong \left(\mathscr{P}_{<\omega}(\mathrm{CS}^{\mathrm{inv}}(\mathfrak{U})) \times \prod_{\mathbb{P}_T}^{\mathrm{bdd}} \omega\right) \setminus \{(a,b) \mid a = \varnothing, \operatorname{supp}(b) \text{ infinite}\},$$

$$\widetilde{\mathrm{Inv}}(\mathfrak{U}) \cong \left(\mathscr{P}_{<\omega}(\mathrm{CS}^{\mathrm{inv}}(\mathfrak{U})) \times \prod_{\mathbb{P}_T}^{\mathrm{bdd}} \omega\right) \setminus \{(a,b) \mid a = \varnothing, b \neq 0\}.$$

## 4. Pure short exact sequences

We study pure short exact sequences of abelian structures $0 \to \mathcal{A} \xrightarrow{\iota} \mathcal{B} \xrightarrow{\nu} \mathcal{C} \to 0$, where $\mathcal{A}$ and $\mathcal{C}$ may be equipped with extra structure. We view them as multisorted structures, and use the relative quantifier elimination results from [2] to describe the domination poset in terms of $\mathcal{A}$ and $\mathcal{C}$. A decomposition of the form $\widetilde{\mathrm{Inv}}(\mathcal{A}(\mathfrak{U})) \times \widetilde{\mathrm{Inv}}(\mathcal{C}(\mathfrak{U}))$ only holds in special cases; in general we will need to look at $*$-types and introduce a family of imaginaries of $\mathcal{A}$ which depends on $\mathcal{B}$.

We refer the reader to [2, Section 4.5] for definitions. We adopt almost identical notations, with the following differences. We write $\mathcal{A}$ for an abelian structure and $L$ for its language. We denote by $\mathcal{F}$ a fundamental family of pp formulas for $\mathcal{B}$. The corresponding family of quotient sorts of $\mathcal{A}$ is denoted by $\mathcal{A}_{\mathcal{F}}$. An $\mathcal{A}$-*sort* is simply a sort in $\mathcal{A}$. We write, e.g., $t(x)$ for a tuple of terms, 0 for a tuple of zeroes of the appropriate length, etc. Tuples of the same length may be added, and tuples of appropriate lengths used as arguments, as in $f(t(x,0) - d) = 0$.

**Example 4.1.** In the simplest abelian structures, namely abelian groups, we have that $\mathcal{F} := \{\exists y\, x = n \cdot y \mid n \in \omega\}$ is always fundamental. In an arbitrary abelian structure, one may always resort to taking as $\mathcal{F}$ the trivially fundamental set of all pp formulas.

**Remark 4.2.** In an $L$-abelian structure, each $L$-term $t(x)$ is built from homomorphisms of abelian groups by taking $\mathbb{Z}$-linear combinations and compositions. Hence, $t(x)$ is itself a homomorphism of abelian groups.

A short exact sequence of abelian groups $0 \to \mathrm{A} \to \mathrm{B} \to \mathrm{C} \to 0$ is pure if and only if, for each $n$, we have $n\mathrm{B} \cap \mathrm{A} = n\mathrm{A}$. This holds, e.g., if C is torsion-free, and in particular in the two examples below. We may take as $\mathcal{F}$ that of Example 4.1.

**Example 4.3.** Suppose that the expansion $L^*_{\mathrm{ac}}$ endows A, C with the structure of oags. Note that one then recovers, definably, an oag structure on B, induced by declaring that $\iota(\mathrm{A})$ is convex. Because of this, and of the fact that the kernel of a morphism of oags is convex, this setting is equivalent to that of a short exact sequence of oags. This will be used in Section 5, with B an oag and A a suitably chosen convex subgroup. The sorts $\mathrm{A}_\varphi$ coincide with the quotients $\mathrm{A}/n\mathrm{A}$.

**Example 4.4.** In the valued field context (Section 6) we will deal with the sequence $1 \to \mathrm{k}^\times \to \mathrm{RV} \setminus \{0\} \to \Gamma \to 0$, which is pure since $\Gamma$ is torsion-free. The extra structure in $L^*_{\mathrm{ac}}$ is induced by the field structure on k and the order on $\Gamma$. The sorts $\mathrm{A}_\varphi$ are in this case $\mathrm{k}^\times/(\mathrm{k}^\times)^n$.

We may and will assume that, for each variable $x$ from an $\mathcal{A}$-sort $\mathrm{A}_s$, the formula $\varphi := x = 0$ is in $\mathcal{F}$, and identify $\mathrm{A}_s$ with $\mathrm{A}_\varphi = \mathrm{A}_s/0\mathrm{A}_s$. In other words, $\mathcal{A} \subseteq \mathcal{A}_\mathcal{F}$.

**Remark 4.5.** As pp formulas commute with cartesian products, every split short exact sequence is pure. Since purity is first-order, a short exact sequence is pure in case some elementarily equivalent structure splits. Note that, even if a short exact sequence splits, it need not do so definably, and that the definition of expanded pure short exact sequence does *not* allow to add splitting maps. If we add one, then matters simplify considerably. For example, if in $L^*_{\mathrm{ac}}$ there is no symbol involving $\mathcal{A}$ and $\mathcal{C}$ jointly, a splitting map makes the short exact sequence interdefinable with the disjoint union of $\mathcal{A}$ and $\mathcal{C}$, where $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ decomposes as a product (Example 2.8).

**Fact 4.6** [2, Remark 4.21]. Let $\varphi(x^{\mathrm{a}}, x^{\mathrm{b}}, x^{\mathrm{c}})$ be an $L^*_{\mathrm{abcq}}$-formula with $x^{\mathrm{a}}, x^{\mathrm{b}}, x^{\mathrm{c}}$ tuples of variables from the $\mathcal{A}_\mathcal{F}$-sorts, $\mathcal{B}$-sorts and $\mathcal{C}$-sorts respectively. There are an $L^*_{\mathrm{acq}}$-formula $\psi$ and special terms $\sigma_i$ such that, in the $L^*_{\mathrm{abcq}}$-theory of all expanded pure short exact sequences, we have $\varphi(x^{\mathrm{a}}, x^{\mathrm{b}}, x^{\mathrm{c}}) \leftrightarrow \psi(x^{\mathrm{a}}, \sigma_1(x^{\mathrm{b}}), \dots, \sigma_m(x^{\mathrm{b}}), x^{\mathrm{c}})$.

**Corollary 4.7.** *The $L^*_{\mathrm{acq}}$-reduct is fully embedded. In particular, $\mathcal{A}$ and $\mathcal{C}$ are orthogonal if and only if they are such in the $L^*_{\mathrm{acq}}$-reduct.*

We show that expanded pure short exact sequences are controlled, domination-wise, by their $L^*_{\mathrm{acq}}$-part, provided we pass to $*$-types. This is a necessity since, in general, there are finite tuples from $\mathcal{B}$ that cannot be domination-equivalent to any finitary tuple from the $L^*_{\mathrm{acq}}$-reduct; see Remark 4.17.

**Proposition 4.8.** *In an expanded pure short exact sequence of $L$-abelian structures, let $\mathcal{F}$ be a fundamental family for $\mathcal{B}$, and let $\kappa \geq |L|$ be a small cardinal. There is a family of $\kappa$-tuples of definable functions $\{\tau^p \mid p \in S_\kappa(\mathfrak{U})\}$ such that*:

(1) *Each function in $\tau^p$ is defined at realisations of $p$.*

(2) *Each $\tau^p$ is partitioned as $(\rho^p, \nu^p)$, where each function in $\rho^p$ is either the identity on some $\mathrm{A}_\varphi$, or has domain a cartesian product of $\mathcal{B}$-sorts and codomain one of the $\mathrm{A}_\varphi$, and each function in $\nu^p$ is either the identity on a $\mathcal{C}$-sort, or one of the $\nu_s$.*

(3) *For each $p \in S_\kappa(\mathfrak{U})$ we have $p \sim_{\mathrm{D}} \tau^p_* p$.*

(4) *For each $p_0, p_1 \in S^{\mathrm{inv}}_\kappa(\mathfrak{U})$ we have $p_0 \otimes p_1 \sim_{\mathrm{D}} \tau^{p_0}_* p_0 \otimes \tau^{p_1}_* p_1$.*

*Proof.* Let $abc \models p(x^{\mathrm{a}}, x^{\mathrm{b}}, x^{\mathrm{c}})$, in the notation of Fact 4.6. Define the tuples $\nu^p$ and $\rho^p$ as follows. For each coordinate in $x^{\mathrm{c}}$ of sort $\mathrm{C}_s$, put in $\nu^p$ the corresponding identity map on $\mathrm{C}_s$. For each coordinate in $x^{\mathrm{b}}$ of sort $\mathrm{B}_s$, put in $\nu^p$ the corresponding map $\nu_s : \mathrm{B}_s \to \mathrm{C}_s$. For each coordinate in $x^{\mathrm{a}}$ of sort $\mathrm{A}_\varphi$, put in $\rho^p$ the corresponding identity map on $\mathrm{A}_\varphi$. For each finite tuple of $L_{\mathrm{b}}$-terms $t(x^{\mathrm{b}}, w)$ and $\varphi \in \mathcal{F}$, if there is $d \in \mathfrak{U}$ such that $p \vdash t(x^{\mathrm{b}}, 0) - d \in \nu^{-1}(\varphi(\mathcal{C}))$, choose such a $d$, call it $d_{p,\varphi,t,x^{\mathrm{b}}}$, and put in $\rho^p$ the map $\rho_\varphi(t(x^{\mathrm{b}}, 0) - d_{p,\varphi,t,x^{\mathrm{b}}})$.

Let $\tau^p$ be the concatenation of $\rho^p$ and $\nu^p$, let $q(y) := \tau^p_* p(x)$, let $D_p$ be the set of all $d_{p,\varphi,t,x^{\mathrm{b}}}$ as above, and let $r(x, y) \in S_{pq}(D_p)$ contain $y = \tau(x)$. Clearly $p \cup r \vdash q$. By Fact 4.6, to show $q \cup r \vdash p$ it suffices to prove that $q \cup r$ recovers the formulas $\varphi(x^{\mathrm{a}}, d^{\mathrm{a}}, \sigma_1(x^{\mathrm{b}}, d^{\mathrm{b}}), \dots, \sigma_m(x^{\mathrm{b}}, d^{\mathrm{b}}), x^{\mathrm{c}}, d^{\mathrm{c}})$ implied by $p$, where the $\sigma_i$ are special terms, $\varphi$ is an $\mathcal{L}^*_{\mathrm{acq}}$-formula, and the $d^\bullet$ are tuples of parameters from the appropriate sorts of $\mathfrak{U}$. Let us say that $q \cup r$ *has access to* the term (with parameters) $\sigma(x^{\mathrm{b}}, d)$ if for some $\mathfrak{U}$-definable function $f$ we have $q(y) \cup r(x, y) \vdash f(y) = \sigma(x^{\mathrm{b}}, d)$. We show that $q \cup r$ has access to all special terms with parameters, and hence $q \cup r \vdash p$.

By construction, $q \cup r$ has access to each $\nu_s(x^{\mathrm{b}}_i)$. Because $\nu$ is a homomorphism of $L$-structures, $q \cup r$ also has access to each $\nu(t_0(x^{\mathrm{b}}, d))$, for $t_0$ an $L_{\mathrm{b}}$-term. In particular, $q \cup r$ decides whether a given tuple $t(x^{\mathrm{b}}, d)$ of $L_{\mathrm{b}}$-terms with parameters is in $\nu^{-1}(\varphi(\mathcal{C}))$ or not. If not, then $q \cup r$ entails $\rho_\varphi(t(x^{\mathrm{b}}, d)) = 0$.

If instead $q \cup r \vdash t(x^{\mathrm{b}}, d) \in \nu^{-1}(\varphi(\mathcal{C}))$, by Remark 4.2 we have

$$t(x^{\mathrm{b}}, d) = t(x^{\mathrm{b}}, 0) + t(0, d),$$

and by construction and the fact that $p$ is consistent with $q \cup r$ we have that $p$ entails $t(x^{\mathrm{b}}, 0) - d_{p,\varphi,t,x^{\mathrm{b}}} \in \nu^{-1}(\varphi(\mathcal{C}))$. As this formula is over $D_p$, it is in $r$. Hence

$$q \cup r \vdash t(0, d) + d_{p,\varphi,t,x^{\mathrm{b}}} = t(x^{\mathrm{b}}, 0) + t(0, d) - (t(x^{\mathrm{b}}, 0) - d_{p,\varphi,t,x^{\mathrm{b}}}) \in \nu^{-1}(\varphi(\mathcal{C})).$$

But $t(0, d) + d_{p,\varphi,t,x^{\mathrm{b}}} \in \mathfrak{U}$, and $\rho_\varphi \restriction \nu^{-1}(\varphi(\mathcal{C}))$ is a homomorphism of $L$-structures. Because of this, and because $q \cup r$ has access to $\rho_\varphi(t(x^{\mathrm{b}}, 0) - d_{p,\varphi,t,x^{\mathrm{b}}})$ by construction, it also has access to

$$\rho_\varphi(t(x^{\mathrm{b}}, 0) - d_{p,\varphi,t,x^{\mathrm{b}}}) + \rho_\varphi(t(0, d) + d_{p,\varphi,t,x^{\mathrm{b}}}) = \rho_\varphi(t(x^{\mathrm{b}}, d)).$$

We are left to prove (4). By definition of $\otimes$, if

$$p_0(x) \otimes p_1(y) \vdash t(x^{\mathrm{b}}, y^{\mathrm{b}}, d) \in \nu^{-1}(\varphi(\mathcal{C})),$$

then there is $\tilde{b} \in \mathfrak{U}$ with $p_0(x) \vdash t(x^{\mathrm{b}}, \tilde{b}, d) \in \nu^{-1}(\varphi(\mathcal{C}))$. Hence, by arguing as above, $p_0 \vdash t(x^{\mathrm{b}}, 0, 0) - d_{p_0,\varphi,t,x^{\mathrm{b}}} \in \nu^{-1}(\varphi(\mathcal{C}))$. So $p_0(x) \otimes p_1(y)$ entails

$$\nu^{-1}(\varphi(\mathcal{C})) \ni t(x^{\mathrm{b}}, y^{\mathrm{b}}, d) - t(x^{\mathrm{b}}, 0, 0) + d_{p_0,\varphi,t,x^{\mathrm{b}}} = t(0, y^{\mathrm{b}}, 0) + t(0, 0, d) + d_{p_0,\varphi,t,x^{\mathrm{b}}}$$

and because $t(0,0,d)+d_{p_0,\varphi,t}\in\mathfrak{U}$, by construction we have

$$p_1(y)\vdash t(0,y^{\mathrm{b}},0)-d_{p_1,\varphi,t,y^{\mathrm{b}}}\in\nu^{-1}(\varphi(\mathcal{C})).$$

Similar arguments show that, in order to have access to $\rho_\varphi(t(x^{\mathrm{b}},y^{\mathrm{b}},d))$, it is enough to have access to $\rho_\varphi(t(x^{\mathrm{b}},0,0)-d_{p_0,\varphi,t,x^{\mathrm{b}}})$ together with $\rho_\varphi(t(0,y^{\mathrm{b}},0)-d_{p_1,\varphi,t,y^{\mathrm{b}}})$, and the conclusion follows. □

**Corollary 4.9** (Theorem C). *Suppose that $\mathfrak{U}$ is an expanded pure short exact sequence of $L$-abelian structures and $\kappa\ge|L|$ is a small cardinal.*

(1) *There is an isomorphism of posets $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U})\cong\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}\restriction L^*_{\mathrm{acq}})$.*

(2) *If $\otimes$ respects $\ge_{\mathrm{D}}$ in $\mathfrak{U}\restriction L^*_{\mathrm{acq}}$, then the same is true in $\mathfrak{U}$, and the above is also an isomorphism of monoids.*

(3) *If $\mathcal{A}$ and $\mathcal{C}$ are orthogonal, then there is an isomorphism of posets $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U})\cong\widetilde{\mathrm{Inv}}_\kappa(\mathcal{A}_{\mathcal{F}}(\mathfrak{U}))\times\widetilde{\mathrm{Inv}}_\kappa(\mathcal{C}(\mathfrak{U}))$. Moreover, if $\otimes$ respects $\ge_{\mathrm{D}}$ in both $\mathcal{A}_{\mathcal{F}}(\mathfrak{U})$ and $\mathcal{C}(\mathfrak{U})$, then the same is true in $\mathfrak{U}$, and the above is also an isomorphism of monoids.*

*Proof.* By Fact 1.1 we have an embedding of posets $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}\restriction L^*_{\mathrm{acq}})\hookrightarrow\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U})$. This embedding is surjective by Proposition 4.8, its inverse being induced by the maps $\tau$, hence an isomorphism. For (2), by Proposition 4.8 we may apply Proposition 1.3 to the family of sorts $\mathcal{A}_{\mathcal{F}}\,\mathcal{C}$. We conclude by combining (2) with Corollary 2.7. □

**Remark 4.10.** Variants of Fact 4.6 for settings such as abelian groups augmented by an absorbing element are presented in [2, Section 4]. These yield variants of Proposition 4.8 and its consequences, with no significant difference in the proofs.

Specialised to abelian groups, the results above enjoy a form of local finiteness.

**Notation 4.11.** For the rest of the section, $L$ is just the language of abelian groups, and $\mathcal{F}$ the family of formulas $\{\exists y\, x=n\cdot y\mid n\in\omega\}$. We will write $\rho_n:\mathrm{B}\to\mathrm{A}/n\mathrm{A}$ in place of $\rho_\varphi:\mathrm{B}\to\mathrm{A}_\varphi$, and identify A with A/0A for notational convenience.

**Definition 4.12.** A $*$-type $p(x)$ is *locally finitary* if $x$ has finitely many coordinates of each sort.

**Proposition 4.13.** *Consider a pure short exact sequence of abelian groups equipped with an $L^*_{\mathrm{abcq}}$-structure. Let $p(x)$ be a locally finitary global type. Then, in Proposition 4.8, we may choose $\tau^p$ in such a way that $\tau^p_* p$ is locally finitary.*

*Proof.* Write $p(x)=p(x^{\mathrm{a}},x^{\mathrm{b}},x^{\mathrm{c}})$ as in the proof of Proposition 4.8, and recall that an $L$-term is just a $\mathbb{Z}$-linear combination. For each $n\in\omega$, consider the subgroup

$$K^p_n:=\{k\in\mathbb{Z}^{|x^{\mathrm{b}}|}\mid\exists d\in\mathrm{B}(\mathfrak{U})\ p\vdash k\cdot x^{\mathrm{b}}-d\in\nu^{-1}(n\mathrm{C})\}\quad\text{of }\mathbb{Z}^{|x^{\mathrm{b}}|},$$

say generated by $k^n_0,\dots,k^n_{m(n)}$. Choose $d_{p,n,i,x^{\mathrm{b}}}$ witnessing $k^n_i\in K^p_n$. Proceed as in Proposition 4.8 but, instead of putting in $\rho^p$ each $\rho_\varphi(t(x^{\mathrm{b}},0)-d_{p,\varphi,t,x^{\mathrm{b}}})$, use a locally finite $\rho^p$ extending $(\rho_n(k_i^n \cdot x^{\mathrm{b}} - d_{p,n,i,x^{\mathrm{b}}}))_{n\in\omega, i\le m(n)}$. Besides this, $\tau^p$ contains a finite tuple of identity maps and finitely many $\nu$, therefore $\tau_*^p p$ is locally finitary.

The proof of Proposition 4.8 now goes through, with a pair of modifications which we now sketch. The first one concerns proving access to each $\rho_n(t(x^{\mathrm{b}}, d))$. Fix $n$ and $t(x^{\mathrm{b}}, d)$. Without loss of generality we have that $d$ is a singleton and $t(x^{\mathrm{b}}, d) = \ell \cdot x^{\mathrm{b}} - d$. If $p \vdash t(x^{\mathrm{b}}, d) \in \nu^{-1}(n\mathrm{C})$, by definition we have $\ell \in K_n^p$, so we may write $\ell = \sum_{i\le m(n)} e_i k_i^n$ for suitable $e_i \in \mathbb{Z}$. This allows us to rewrite

$$\begin{aligned} t(x^{\mathrm{b}}, d) = \ell \cdot x^{\mathrm{b}} - d &= \Bigg(\sum_{i\le m(n)} e_i k_i^n\Bigg) \cdot x^{\mathrm{b}} - d \\ &= \sum_{i\le m(n)} e_i (k_i^n \cdot x^{\mathrm{b}} - d_{p,n,i,x^{\mathrm{b}}}) + \sum_{i\le m(n)} e_i\, d_{p,n,i,x^{\mathrm{b}}} - d. \end{aligned}$$

Since $\ell \cdot x^{\mathrm{b}} - d$ and all $k_i^n \cdot x^{\mathrm{b}} - d_{p,n,i,x^{\mathrm{b}}}$ are in $\nu^{-1}(n\mathrm{C})$, so is $\sum_{i\le m(n)} e_i\, d_{p,n,i,x^{\mathrm{b}}} - d$. Since $\rho_n \restriction \nu^{-1}(n\mathrm{C})$ is a homomorphism and $\sum_{i\le m(n)} e_i\, d_{p,n,i,x^{\mathrm{b}}} - d \in \mathfrak{U}$, we have that $q \cup r$ has access to $\rho_n(t(x^{\mathrm{b}}, d))$.

Finally, proving (4) of Proposition 4.8 boils down to showing $K_n^{p\otimes q} = K_n^p \times K_n^q$, where we identify, e.g., $K_n^p$ with $K_n^p \times \{0\}$. Since by construction $K_n^p \cap K_n^q = \{0\}$, one only needs to show generation. We leave the easy proof to the reader. $\square$

**Remark 4.14.** In the case of abelian groups, we therefore have an analogue of Corollary 4.9 where $\kappa$-types are replaced by locally finitary $\omega$-types.

**Corollary 4.15.** *Let $\mathfrak{U}$ be an expanded pure short exact sequences of abelian groups where, for all $n > 0$, the sort $\mathrm{A}/n\mathrm{A}$ is finite. If $\mathrm{A}$ and $\mathrm{C}$ are orthogonal, there is an isomorphism of posets $\widetilde{\mathrm{Inv}}(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}(\mathrm{A}(\mathfrak{U})) \times \widetilde{\mathrm{Inv}}(\mathrm{C}(\mathfrak{U}))$. If $\otimes$ respects $\ge_{\mathrm{D}}$ in $\mathrm{A}$ and $\mathrm{C}$, then $\otimes$ respects $\ge_{\mathrm{D}}$, and the above is an isomorphism of monoids.*

*Proof.* Use Proposition 4.13 and observe that for each $p$ we may replace $\tau^p$ by its composition with the projection on the nonrealised coordinates of $\tau_*^p p$ and still have the same results. If $\mathrm{A}/n\mathrm{A}$ is finite for all $n > 0$ and $p$ is finitary, this yields another finitary type. The conclusion now follows as in the proof of Corollary 4.9. $\square$

**Remark 4.16.** The $A/nA$ are in general necessary to obtain a product decomposition. For example, let A be a regular oag divisible by all $\mathfrak{p} \in \mathbb{P} \setminus \{2\}$, and with $[\mathrm{A} : 2\mathrm{A}]$ infinite, and let C be a nontrivial divisible oag. The expanded short exact sequence $0 \to \mathrm{A} \to \mathrm{B} \to \mathrm{C} \to 0$ induces a group ordering on B (Example 4.3). Let $p(y)$ concentrate on B, at $+\infty$, in a new coset modulo 2B. For every nonrealised 1-type $q$ of an element of sort A divisible by all $n$, we have $p \perp^{\mathrm{w}} q$. It follows that $p$ cannot dominate any nonrealised $p'$ in a cartesian power of A: such a $p'$ must have a coordinate in a nonrealised cut, and hence dominate a type $q$ as above. Hence, if we had a product decomposition as in Corollary 4.15, then $p$ would be domination-equivalent to a type in a cartesian power of C. This is a contradiction, because C is orthogonal to $(\mathrm{A}/n\mathrm{A})_{n<\omega}$, while $p$ dominates a nonrealised type in $\mathrm{A}/2\mathrm{A}$.

**Remark 4.17.** Analogously, $\omega$-types are a necessity: let A be a regular oag with each $[\mathrm{A} : n\mathrm{A}]$ infinite, C a nontrivial divisible oag, and take as $p \in S_{\mathrm{B}}(\mathfrak{U})$ the type at $+\infty$ in a new coset of each $n\mathrm{A}$. For each $n > 1$, there is a nonrealised 1-type $q_n$ of sort $\mathrm{A}/n\mathrm{A}$ such that $p \geq_{\mathrm{D}} q_n$. One shows that the only way for a finitary type in $((\mathrm{A}/n\mathrm{A})_{n\in\omega}, \mathrm{C})$ to dominate all of the $q_n$ is to have a nonrealised coordinate in the sort A, hence to dominate a type orthogonal to $p$.

## 5. Finitely many definable convex subgroups

Using the previous two sections we may describe $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ in oags with finitely many $L_{\mathrm{oag}}$-definable convex subgroups. The arguments still work if the subgroups are defined "by fiat" using additional predicates, so we work in this setting.

**Definition 5.1.** Let $G$ be an oag with unary predicates $H_0, \dots, H_s$, each defining a convex subgroup, with $0 = H_0 \subsetneq H_1 \subsetneq \dots \subsetneq H_{s-1} \subsetneq H_s = G$, and such that $G$ has no other definable convex subgroup. Denote by $\mathbb{T}$ the union of the set of prime powers with $\{0\}$ and work with the following sorts. For $0 \leq i < s$, a sort $\mathrm{S}_i$ for $G/H_i$, carrying $L_{\mathrm{oag}}$ together with predicates for $H_j/H_i$ for $i < j < s$. For $1 \leq i \leq s$ and $n \in \mathbb{T}$, sorts $\mathrm{Q}_{i,n}$ for $H_i/(nH_i + H_{i-1})$, carrying $L_{\mathrm{ab}}$ if $n \neq 0$ and $L_{\mathrm{oag}}$ if $n = 0$. We denote by $\mathcal{Q}_i$ the family of sorts $(\mathrm{Q}_{i,n})_{n\in\mathbb{T}}$. We include the canonical projection and inclusion maps together with, for each $n \in \mathbb{T}$ and $1 \leq i \leq s-1$, the maps $\rho_{n,i} : \mathrm{S}_{i-1} \to \mathrm{Q}_{i,n}$ as in Notation 4.11, relative to the short exact sequence $0 \to \mathrm{Q}_{i,0} \to \mathrm{S}_{i-1} \to \mathrm{S}_i \to 0$.

For $1 \leq i < s$ the short exact sequence $0 \to H_i/H_{i-1} \to G/H_{i-1} \to G/H_i \to 0$ is pure and, as pointed out in Example 4.4, interdefinable with an expanded pure short exact sequence of abelian groups.

**Lemma 5.2.** *Every $H_{i+1}/H_i$ is regular. For each $i \neq j$, the sort $\mathrm{S}_i$ is fully embedded as an oag, the family $\mathcal{Q}_i$ (with $L_{\mathrm{oag}}$-structure on $\mathrm{Q}_{i,0}$, $L_{\mathrm{ab}}$-structure on other sorts, and projection maps) is fully embedded, orthogonal to $\mathrm{S}_i$, and orthogonal to $\mathcal{Q}_j$.*

*Proof.* Apply Fact 3.2 to $H_{i+1}/H_i$, whose only definable convex subgroups are itself and $\{0\}$. The rest is by Corollary 4.7, Remark 2.3, and induction on $i$. □

**Theorem 5.3.** *Let $G$ be as in Definition 5.1, and $\kappa$ a small infinite cardinal. Then $\otimes$ respects $\geq_{\mathrm{D}}$, and $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}^{\mathrm{eq}}) \cong \prod_{i=1}^{s} \widetilde{\mathrm{Inv}}_\kappa(\mathcal{Q}_i(\mathfrak{U}))$.*

*Proof.* By the previous lemma, Corollaries 4.9, 3.33 and induction we get that $\otimes$ respects $\geq_{\mathrm{D}}$, and $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}) \cong \prod_{i=1}^{s} \widetilde{\mathrm{Inv}}_\kappa(\mathcal{Q}_i(\mathfrak{U}))$.

If the $H_i$ are $L_{\mathrm{oag}}$-definable,[2] a result of Vicaría [32, Theorem 5.1] yields weak elimination of imaginaries in the language with sorts $\mathrm{S}_i/n\mathrm{S}_i$ for $0 \le i < s$ and $n \in \mathbb{T}$,[3] and one may check that her proof goes through also in the case where the $H_i$ are explicitly named by predicates, i.e., not necessarily $L_{\mathrm{oag}}$-definable.

After adding the sorts from Vicaría's result, for $1 \le i \le s$ the short exact sequences $0 \to \mathrm{Q}_{i,n} \to \mathrm{S}_{i-1}/n\mathrm{S}_{i-1} \to \mathrm{S}_i/n\mathrm{S}_i \to 0$ are fully embedded, and Corollary 4.9 may thus be applied to these. From this, we obtain an embedding $\prod_{i=1}^{s} \widetilde{\mathrm{Inv}}_\kappa(\mathcal{Q}_i(\mathfrak{U})) \hookrightarrow \widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}^{\mathrm{eq}})$. As $\mathrm{Q}_{s,n} = \mathrm{S}_{s-1}/n\mathrm{S}_{s-1}$, by induction on $i$ one obtains surjectivity of this embedding. We leave to the reader to check this, along with the proof of transfer of compatibility of $\otimes$ and $\ge_{\mathrm{D}}$, by showing that every $*$-type is dominated by its image among a suitable tuple of definable maps. $\square$

## 6. Benign valued fields

In this section $T$ is a complete $\mathcal{RV}$-expansion of a theory of henselian valued fields with elimination of K-quantifiers and "enough saturated maximal models" (see below for the precise definitions). We show the existence of an isomorphism $\widetilde{\mathrm{Inv}}(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}(\mathcal{RV}(\mathfrak{U}))$. In particular, our results hold in any *benign* valued field in the sense of [30],[4] i.e., in any henselian valued field which is of equicharacteristic 0, or algebraically closed, or algebraically maximal Kaplansky of characteristic $\mathfrak{p} > 0$.

Associate to a valued field K the pure (Example 4.4) short exact sequence $1 \to \mathrm{k}^\times \to \mathrm{K}^\times/(1+\mathfrak{m}) \to \Gamma \to 0$. Add absorbing elements $0, 0, \infty$, and view it as a short exact sequence of abelian monoids $1 \to \mathrm{k} \to \mathrm{K}/(1+\mathfrak{m}) \to \Gamma \cup \{\infty\} \to 0$. We may harmlessly conflate the two settings (Remark 4.10) and write $\Gamma$ for $\Gamma \cup \{\infty\}$.

The middle term $\mathrm{K}/(1+\mathfrak{m})$ is called the *leading term structure* RV, and comes with a natural map $\mathrm{rv}: \mathrm{K} \to \mathrm{K}/(1+\mathfrak{m}) = \mathrm{RV}$ through which the valuation $v: \mathrm{K} \to \Gamma$ factors. Besides the structure of a (multiplicatively written) monoid, RV is equipped with a "partially defined sum": a ternary relation defined by

$$\oplus(x_0, x_1, x_2) \overset{\mathrm{def}}{\iff} \exists y_0, y_1, y_2 \in \mathrm{K}\left(y_2 = y_0 + y_1 \wedge \bigwedge_{i<3} \mathrm{rv}(y_i) = x_i\right).$$

When there is a unique $x_2$ such that $\oplus(x_0, x_1, x_2)$, we write $x_0 \oplus x_1 = x_2$, and say that $x_0 \oplus x_1$ *is well defined*. It turns out that $\mathrm{rv}(x) \oplus \mathrm{rv}(y)$ is well defined if and only if $v(x+y) = \min\{v(x), v(y)\}$. If we say that $\bigoplus_{i<\ell} x_i$ *is well defined*, we mean that, regardless of the choice of parentheses and order of the summands, the "sum" is well defined and always yields the same result.

[2]Oags with finitely many definable convex subgroups are known as the oags *of finite regular rank*. Note that every $H_i$ must be fixed by every automorphism, and is therefore $\varnothing$-definable.

[3]Vicaría uses sorts indexed by $n \in \omega$; as in Remark 3.14, it suffices to work with $n \in \mathbb{T}$.

[4]Definition 1.57 of [30] allows $\{k\}$-$\{\Gamma\}$-expansions in the definition of benign. Since we are shortly going to allow more general expansions, the difference is immaterial for our purposes.

Let $\mathcal{RV}$ be the expansion of $1 \to \mathrm{k} \xrightarrow{\iota} \mathrm{RV} \xrightarrow{v} \Gamma \to 0$ by the field structure on k and the order on $\Gamma$. This induces an expansion of RV, which is precisely that given by multiplication and $\oplus$ [2, Lemma 5.17], is biinterpretable with $\mathcal{RV}$, and can be axiomatised independently [30, Appendix B]. Hence, we may view RV as a standalone structure $(\mathrm{RV}, \cdot, \oplus)$, fully embedded in (K, RV, rv), and in $\mathcal{RV}$.

By the (short) five lemma, an extension of valued fields is immediate, i.e., does not change k nor $\Gamma$, if and only if it does not change $\mathcal{RV}$.

In this section, $L$ has sorts K, k, RV, $\Gamma$, function symbols $\mathrm{rv}:\mathrm{K}\to\mathrm{RV}$, $\iota:\mathrm{k}\to\mathrm{RV}$, $v:\mathrm{RV}\to\Gamma$. We abuse the notation and also write $v$ for the composition $v\circ\mathrm{rv}$. The sorts K and k carry disjoint copies of the language of rings, $\Gamma=\Gamma\cup\{\infty\}$ carries the (additive) language of ordered groups, together with an absorbing element $\infty$ and an extra constant symbol $v(\mathrm{Char}(\mathrm{K}))$, and RV carries the (multiplicative) language of groups, together with an absorbing element 0 and a ternary relation symbol $\oplus$. We denote by $\mathcal{RV}$ the reduct to the sorts k, RV, $\Gamma$. There may be other arbitrary symbols on $\mathcal{RV}$, i.e., as long as they do not involve K. An *$\mathcal{RV}$-expansion* of a theory $T'$ of valued fields is a complete $L$-theory $T\supseteq T'$. Until the end of the section, $T$ denotes such a theory. We identify k with the image of its embedding $\iota$ in RV.

**Remark 6.1.** Angular components factor through the map rv, yielding a splitting of $\mathcal{RV}$. Therefore, the Denef–Pas language (and each of its {k, $\Gamma$}-expansions)[5] may be seen as an $\mathcal{RV}$-expansion. In that case $\mathcal{RV}$ is definably isomorphic to $\mathrm{k}\times\Gamma$.

**Fact 6.2.** Fix a language $L$ as above. The theory of all $\mathcal{RV}$-expansions of benign valued fields eliminates K-sorted quantifiers.

*Proof.* In equicharacteristic this follow from [9, Théorème 2.1]. The residue characteristic 0 case is explicitly done in [4, Theorem B] (see also [20, Corollary 2.2]), the algebraically maximal Kaplansky case in [20, Theorem 2.6] (see [15, Corollary A.3] for a modern treatment). The algebraically closed case is folklore (see, e.g., [18, Fact 2.4]). □

**Remark 6.3.** If $T$ eliminates K-quantifiers, then every formula is equivalent to one of the form $\varphi\big(x, \mathrm{rv}(f_0(y)), \dots, \mathrm{rv}(f_m(y))\big)$, where $\varphi(x, z_0, \dots, z_m)$ is a formula in $\mathcal{RV}$, $x$ and $z$ tuples of $\mathcal{RV}$-variables, $y$ a tuple of K-variables, and the $f_i$ polynomials over $\mathbb{Z}$. In particular, $\mathcal{RV}$ (with the restriction of $L$ to its sorts) is fully embedded.

*Proof.* By inspecting the formulas without K-sort quantifiers and observing that, for example, if $y$ is of sort K then $T\vdash y=0\leftrightarrow \mathrm{rv}(y)=0$. □

**Definition 6.4.** Let $K_0\subseteq K_1$ be an extension of valued fields. A basis $(a_i)_i$ of a $K_0$-vector subspace of $K_1$ is *separating* if for all finite tuples $d$ from $K_0^\ell$ and pairwise distinct $i_j$, we have $v\big(\sum_{j<\ell} d_j a_{i_j}\big)=\min_{j<\ell}(v(d_j)+v(a_{i_j}))$.

[5] A {k, $\Gamma$}-expansion is one where the new symbols only involve the sorts k and $\Gamma$, possibly simultaneously. If we want to exclude the latter possibility, we speak of {k}-{$\Gamma$}-expansions.

**Fact 6.5.** A basis $(a_i)_i$ is separating if and only if each sum $\bigoplus_{j<\ell} \operatorname{rv}(d_j)\operatorname{rv}(a_{i_j})$ is well defined. If this is the case, it equals $\operatorname{rv}\bigl(\sum_{j<\ell} d_j\, a_{i_j}\bigr)$.

**Lemma 6.6.** *Let $p \in S^{\mathrm{inv}}_{\mathrm{K}^{\le\omega}}(\mathfrak{U}, M_0)$, $M_0 \preceq M \prec^+ \mathfrak{U} \subseteq B$, $a \vDash p \mid B$, and $(f_i)_{i\in I}$ a family of $M$-definable functions $\mathrm{K}^\omega \to \mathrm{K}$ such that $(f_i(a))_{i\in I}$ is a separating basis of the $\mathrm{K}(M)$-vector space they generate. If $M$ is $|M_0|^+$-saturated, or $p$ is definable, then $(f_i(a))_{i\in I}$ is a separating basis of the $\mathrm{K}(B)$-vector space they generate.*

*Proof.* Towards a contradiction, suppose there are an $L(M)$-formula

$$\varphi(x, w) := v\Biggl(\sum_{i<\ell} w_i\, f_i(x)\Biggr) > \min_{i<\ell}\{v(w_i) + v(f_i(x))\}$$

and $d \in B^{|w|}$ such that $a \vDash \varphi(x, d)$. Let $H$ be the set of parameters appearing in $\varphi(x, w)$. Choose $\tilde{d} \in M$ with $\tilde{d} \equiv_{M_0 H} d$ if $M$ is $|M_0|^+$-saturated, or in $d_p\,\varphi$ if $p$ is definable. Then $a \vDash \varphi(x, \tilde{d})$ contradicts that $(f_i(a))_{i\in I}$ is separating over $M$. $\square$

Hence, saturation of $M$ allows to lift separating bases. As maximality of $M$ guarantees their existence (see Lemma 6.13 below), we give the following definition.

**Definition 6.7.** We say that $T$ has *enough saturated maximal models* if for every $\kappa > |L|$, for every $M_0 \vDash T$ of size at most $\kappa$ there is $M \succ M_0$ of size at most $2^{2^\kappa}$ which is maximally complete and $|M_0|^+$-saturated.

**Remark 6.8.** If we restrict to definable types, saturation is not necessary to lift separating bases (see Lemma 6.6), and it is enough to assume only "enough maximal models" for weak versions of the results of this section to go through.

**Proposition 6.9.** *Let $T$ be an $\mathcal{RV}$-expansion of a theory of henselian valued fields eliminating* K*-quantifiers, where every $M \vDash T$ has a unique maximal immediate extension up to isomorphism over $M$. If $M' \vDash T$ is maximal, $\kappa > |L|$, and $\mathcal{RV}(M')$ is $\kappa$-saturated, then $M'$ is $\kappa$-saturated.*

The proposition above is folklore, but we include a proof for convenience. As pointed out to us by the referee, uniqueness of the maximal immediate extension is not needed, and maximality of $M'$ may be relaxed to requiring that chains of balls of length smaller than $\kappa$ have nonempty intersection; the result then follows by using Swiss cheese decomposition. Nevertheless, the proof below has the advantage that it can be adapted to more general contexts, which we will need in Proposition 8.1.

*Proof.* If $\kappa$ is limit $\kappa$-saturation equals $\lambda$-saturation for all $\lambda < \kappa$, so we may assume $\kappa$ is successor, and hence regular. It suffices to prove that if $M \equiv M'$ is $\kappa$-saturated, then the set $\mathcal{S}$ of partial elementary maps between $M$ and $M'$ with domain of size less than $\kappa$ has the back-and-forth property. In fact, we only need the "forth" part (and the "back" part is true by $\kappa$-saturation of $M$). So assume $f \in \mathcal{S}$, with

$$f : A = (\mathrm{K}(A), \mathcal{RV}(A)) \to A' = (\mathrm{K}(A'), \mathcal{RV}(A'))$$

and suppose that $A \subseteq B \subseteq M$, with $|B| < \kappa$. In order to extend $f$ to some $g \in \mathcal{S}$ with domain containing $B$, consider the following two constructions.

**Construction 1.** Enlarge $A$ to an elementary substructure. That is, there are $A_1 \supseteq A$ and $f_1 : A_1 \to A_1'$ extending $f$ such that $f_1 \in \mathcal{S}$ and $A_1 \preceq M$. To do this, we find $A_1'$ with $A' \subseteq A_1' \preceq M'$ and $|A_1'| < \kappa$ using the Löwenheim–Skolem theorem, and invoke $\kappa$-saturation of $M$ to obtain the desired $A_1$, $f_1$.

**Construction 2.** For a given $\hat{B}$ such that $A \subseteq \hat{B} \subseteq M$ and $|\hat{B}| < \kappa$, enlarge $\mathcal{RV}(A)$ so that it contains $\mathcal{RV}(\hat{B})$. That is, there are $A_1 \supseteq A$ and $f_1 : A_1 \to A_1'$ extending $f$ such that $f_1 \in \mathcal{S}$ and $\mathcal{RV}(A_1) \supseteq \mathcal{RV}(\hat{B})$. To do this, it suffices to set $A_1 = (\mathrm{K}(A), \mathcal{RV}(\hat{B}))$ and extend $f$ on $\mathcal{RV}$ using $\kappa$-saturation of $\mathcal{RV}(M')$; by elimination of K-quantifiers, the extension is still an elementary map.

By repeated applications of the constructions above, we find an elementary chain $(M_n)_{n\in\omega}$ of elementary submodels of $M$, with $A \subseteq M_0$, and $f_n \in \mathcal{S}$ with domain $M_n$ such that $f_0 \supseteq f$, $f_{n+1} \supseteq f_n$, and that if $B_n$ is the structure generated by $M_n B$ then $\mathcal{RV}(B_n) \subseteq \mathcal{RV}(M_{n+1})$. Let $M_\omega := \bigcup_{n\in\omega} M_n$ and let $f_\omega := \bigcup_{n\in\omega} f_n$. Since $\kappa$ is regular and uncountable we have $f \in \mathcal{S}$, and by construction the structure $B_\omega$ generated by $M_\omega B$ is K-generated and an immediate extension of $M_\omega$. Since $M'$ is maximal and the maximal immediate extension of $M_\omega$ is uniquely determined up to $M_\omega$-isomorphism, we may extend $f_\omega$ to a map $g \in \mathcal{S}$ with domain $B_\omega \supseteq B$. □

**Remark 6.10.** Above (and in {k}-{Γ}-expansions of the Denef–Pas language), if k and Γ are orthogonal it suffices to assume that $\mathrm{k}(M')$ and $\Gamma(M')$ are $\kappa$-saturated.

**Corollary 6.11.** *Suppose that $T$ satisfies the assumptions of Proposition 6.9, and furthermore that every maximal immediate extension of every $M \models T$ is an elementary extension. Then $T$ has enough saturated maximal models.*

*Proof.* Given $\kappa > |L|$ and $M_0 \models T$ of size $|M_0| \leq \kappa$, find $M_1 \succ M_0$ which is $|M_0|^+$-saturated of size $|M_1| \leq 2^{|M_0|}$. Let $M$ be a maximal immediate extension of $M_1$. Then $\mathcal{RV}(M) = \mathcal{RV}(M_1)$, and the latter is $|M_0|^+$-saturated because $M_1$ is. By assumption, $M \succ M_1$, and by Proposition 6.9 $M$ is $|M_0|^+$-saturated. To conclude, observe that, since by Krull's inequality [10, Proposition 3.6] we have $|\mathrm{K}| \leq \mathrm{k}^\Gamma$, we obtain

$$|M| \leq |\mathrm{k}(M)|^{|\Gamma(M)|} = |\mathrm{k}(M_1)|^{|\Gamma(M_1)|} \leq (2^{|M_0|})^{2^{|M_0|}} = 2^{2^{|M_0|}}. \qquad \square$$

**Corollary 6.12.** *Every $\mathcal{RV}$-expansion of a benign $T$ has enough saturated maximal models.*

*Proof.* Since the assumptions of Fact 6.2 are preserved by taking maximal immediate extensions (which are unique by [19, Theorem 5]) elementarity follows from elimination of K-quantifiers. We conclude by Corollary 6.11. □

**Lemma 6.13.** *Let $p, q \in S^{\mathrm{inv}}_{\mathrm{K}^{<\omega}}(\mathfrak{U}, M_0)$, let $(a, b) \vDash p \otimes q$ and $M_0 \prec M \prec^+ \mathfrak{U}$.*

(1) *If $M$ is maximally complete, then there are polynomials $(f_i)_{i<\omega}$ in $\mathrm{K}(M)[x]$ such that $(f_i(a))_{i<\omega}$ is a separating basis of $\mathrm{K}(M)[a]$ as a $\mathrm{K}(M)$-vector space.*

(2) *If $M$ is $|M_0|^+$-saturated then, for each $(f_i)_{i<\omega}$ as above, $\{f_i(a) \mid i < \omega\}$ is a separating basis of $\mathrm{K}(\mathfrak{U})[a]$.*

(3) *If $(f_i^p(a))_{i<\omega}$, $(f_j^q(b))_{j<\omega}$ are separating bases of $\mathrm{K}(\mathfrak{U})[a]$ and $\mathrm{K}(\mathfrak{U})[b]$, then $(f_i^p(a) \cdot f_j^q(b))_{i,j<\omega}$ is a separating basis of $\mathrm{K}(\mathfrak{U})[ab]$.*

*Proof.* Part (1) is by [5, Lemma 3] (see also [17, Lemma 12.2]) and does not require saturation, and part (2) is by Lemma 6.6 applied to $(f_i)_{i<\omega}$. So we only need to prove (3). By the definition of $\otimes$, the tuple $(f_i^p(a) \cdot f_j^q(b))_{i,j<\omega}$ is linearly independent, and clearly it generates $\mathrm{K}(\mathfrak{U})[ab]$ as a $\mathrm{K}(\mathfrak{U})$-vector space. Let us check that this basis is separating. Let $B$ be the structure generated by $\mathfrak{U}b$. By Lemma 6.6, $(f_i^p(a))_{i<\omega}$ is a separating basis of the $\mathrm{K}(B)$-vector space $\mathrm{K}(B)[a]$, so we have

$$\begin{aligned} v\Bigl(\sum_{i,j} d_{ij} f_i^p(a) f_j^q(b)\Bigr) &= v\Bigl(\sum_i \Bigl(\sum_j d_{ij} f_j^q(b)\Bigr) f_i^p(a)\Bigr) \\ &= \min_i \Bigl(v\Bigl(\sum_j d_{ij} f_j^q(b)\Bigr) + v(f_i^p(a))\Bigr) \\ &= \min_i \Bigl(\min_j \bigl(v(d_{ij}) + v(f_j^q(b))\bigr) + v(f_i^p(a))\Bigr) \\ &= \min_{i,j} \bigl(v(d_{ij}) + v(f_j^q(b)) + v(f_i^p(a))\bigr) \\ &= \min_{i,j} \bigl(v(d_{ij}) + v(f_j^q(b) \cdot f_i^p(a))\bigr). \end{aligned}$$ □

**Proposition 6.14.** *Suppose that $T$ eliminates $\mathrm{K}$-quantifiers and has enough saturated maximal models. For every $p \in S^{\mathrm{inv}}(\mathfrak{U})$ there is $q \in S^{\mathrm{inv}}_{\mathcal{RV}^\omega}(\mathfrak{U})$ such that $p \sim_{\mathrm{D}} q$. More precisely, let $p(x, z) \in S^{\mathrm{inv}}(\mathfrak{U}, M_0)$, where $x$ is a tuple of $\mathrm{K}$-variables and $z$ a tuple of $\mathcal{RV}$-variables. Let $(a, c) \vDash p(x, z)$, let $M \succ M_0$ be $|M_0|^+$-saturated and maximally complete, and let $(f_i)_{i<\omega}$ be given by Lemma 6.13 applied to $a$ and $M$. Then $p$ is domination-equivalent to the $*$-type $q(y, t) := \mathrm{tp}(\mathrm{rv}(f_i(a))_{i<\omega}, c/\mathfrak{U})$, witnessed by $r(x, z, y, t) := \mathrm{tp}(a, c, \mathrm{rv}(f_i(a))_{i<\omega}, c/M)$.*

*Proof.* That $p \cup r \vdash q$ is trivial. By elimination of $\mathrm{K}$-quantifiers (Fact 6.2), to prove $q \cup r \vdash p$ it is enough to show that $q \cup r$ has access to every $\mathrm{rv}(f(x))$, that is, that for every $f \in \mathrm{K}(\mathfrak{U})[x]$, there is a $\mathfrak{U}$-definable function $g$ such that $q \cup r \vdash \mathrm{rv}(f(x)) = g(y)$. Write $f(x) = \sum_{i<\ell} d_i f_i(x)$. By Fact 6.5, we have $\mathrm{rv}(f(a)) = \bigoplus_{i<\ell} \mathrm{rv}(d_i)\,\mathrm{rv}(f_i(a))$, and we only need to ensure that this information is in $q \cup r$. But by Fact 6.5 whether the $(f_i(a))_{i<\omega}$ form a separating basis or not only depends on the type of their images in RV, which is part of $q$ by definition. □

The work done so far is enough to obtain an infinitary version of Theorem B. After stating such a version, we will proceed to finitise it.

**Remark 6.15.** Separating bases of vector spaces of uncountable dimension need not exist. Nevertheless, a $*$-type version of Lemma 6.13 still holds, with the $f_i(a)$ now enumerating separating bases of all finite dimensional subspaces of $\mathrm{K}(M)[a]$.

**Corollary 6.16.** *If $\kappa$ is a small infinite cardinal, there is an isomorphism of posets $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}_\kappa(\mathcal{RV}(\mathfrak{U}))$. If $\otimes$ respects $\geq_{\mathrm{D}}$ on $*$-types in $\mathcal{RV}(\mathfrak{U})$, then the same holds in $\mathfrak{U}$, and the above is also an isomorphism of monoids.*

*Proof.* By the $*$-type versions of Lemma 6.13 and Propositions 6.14 and 1.3. $\square$

**Lemma 6.17.** *Let $M_0 \prec^+ M \prec^+ \mathfrak{U}$, let $e \vDash q \in S^{\mathrm{inv}}_{\mathrm{RV}^\omega}(\mathfrak{U}, M_0)$. Let $I \subseteq \omega$ be such that $(v(e_i))_{i\in I}$ generates $\mathbb{Q}\langle\Gamma(\mathfrak{U})v(e)\rangle$ over $\mathbb{Q}\,\Gamma(\mathfrak{U})$ as $\mathbb{Q}$-vector spaces. Let $G \subseteq \mathrm{RV}$ be the multiplicative group generated by $\mathrm{RV}(\mathfrak{U})e$. Let $(g_j)_{j\in J} \subseteq \mathrm{k} \cap G$ be such that $\mathrm{k} \cap G \subseteq \mathrm{acl}(\mathfrak{U}(g_j)_{j\in J})$ and $J$ is countable. Let $b := (e_i, g_j \mid i \in I, j \in J)$. Then there is $M \prec N \prec^+ \mathfrak{U}$ such that $e$ and $b$ are interalgebraic over $N$.*

*Proof.* By assumption, for $\ell \in \omega \setminus I$ there are $n_\ell > 0$, $d_\ell \in \mathfrak{U}$, a finite $I_0 \subseteq I$ and, for $i \in I_0$, integers $n_{\ell,i} \in \mathbb{Z}$, with $n_\ell v(e_\ell) = v(d_\ell) + \sum_{i\in I_0} n_{\ell,i} v(e_i)$. By $M_0$-invariance, we may assume $d_\ell \in M$. Let $h_\ell(x)$ be the $M$-definable function $h_\ell(y) := (y_\ell^{n_\ell})/\big(d_\ell \prod_{i\in I_0} y_i^{n_{\ell,i}}\big)$. By construction, we have $v(h_\ell(e)) = 0$, and hence $h_\ell(e) \in G \cap \mathrm{k}^\times$, so by assumption $h_\ell(e) \in \mathrm{acl}(\mathfrak{U}(g_j)_{j\in J})$. Let $N \succ M$ be small such that $\{h_\ell(e) \mid \ell \in \omega \setminus I\} \subseteq \mathrm{acl}(N(g_j)_{j\in J})$ and $\{g_j \mid j \in J\}$ is contained in the group generated by $\mathrm{RV}(N)e$. By definition of $h_\ell$, for each $\ell \in \omega \setminus I$, we therefore have $e_\ell^{n_\ell} \in \mathrm{acl}(Nb)$. As $\Gamma$ is ordered and the kernel of $v : \mathrm{RV} \to \Gamma$ is the multiplicative group of a field, RV has finite $n$-torsion for each $n$, so $e_\ell$ is algebraic over $e_\ell^{n_\ell}$, and hence $e \in \mathrm{acl}(Nb)$. $\square$

**Theorem 6.18** (Theorem B). *For $T$ an $\mathcal{RV}$-expansion of a theory of valued fields with enough saturated maximal models eliminating* K*-quantifiers (e.g., a benign one), there is an isomorphism of posets $\widetilde{\mathrm{Inv}}(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}(\mathcal{RV}(\mathfrak{U}))$. If $\otimes$ respects $\geq_{\mathrm{D}}$ in $\mathcal{RV}(\mathfrak{U})$, then $\otimes$ respects $\geq_{\mathrm{D}}$ in $\mathfrak{U}$, and the above is an isomorphism of monoids.*

*Proof.* Fix $p(x, z) \in S^{\mathrm{inv}}(\mathfrak{U})$ and $ac \vDash p$, where $x$ is a tuple of K-variables and $z$ a tuple of $\mathcal{RV}$-variables. Let $(f_i)_{i<\omega}$ be given by Lemma 6.13. As usual, denote by $\mathfrak{U}(a)$ the field generated by $a$ over $\mathfrak{U}$. As $\mathrm{trdeg}(\mathfrak{U}(a)/\mathfrak{U})$ is finite, by the Abhyankar inequality so is $\dim_{\mathbb{Q}}(\mathbb{Q}\,\Gamma(\mathfrak{U}(a))/\mathbb{Q}\,\Gamma(\mathfrak{U}))$. Let $m$ be such that $v(f_i(a))_{i<m}$ generates $\mathbb{Q}\,\Gamma(\mathfrak{U}(a))$ over $\mathbb{Q}\,\Gamma(\mathfrak{U})$. Again by the Abhyankar inequality, $\mathrm{trdeg}(\mathrm{k}(\mathfrak{U}(a))/\mathrm{k}(\mathfrak{U}))$ is finite. By the choice of the $f_j$ and Fact 6.5, we may choose a transcendence basis $(g_j \mid j < n)$ of $\mathrm{k}(\mathfrak{U}(a))$ over $\mathrm{k}(\mathfrak{U})$, which is contained in the group generated by $\mathrm{RV}(\mathfrak{U})\big(\mathrm{rv}(f_i(a))\big)_{i<\omega}$. Write each $g_j$ as $h_j(a)$, for suitable definable functions $h_j$. We may now apply Lemma 6.17 to $e = \big(\mathrm{rv}(f_i(a))\big)_{i<\omega}$, the $g_j$ defined above, and $I = \{i \in \omega \mid i < m\}$. Together with Proposition 6.14, we obtain

$$p \sim_{\mathrm{D}} p' := \operatorname{tp}\bigl(\operatorname{rv}(f_i(a))_{i<m}, (h_j(a))_{j<n}, c/\mathfrak{U}\bigr) \tag{1}$$

Therefore, every (finitary) type is equivalent to one in $\mathcal{RV}$. By full embeddedness of $\mathcal{RV}$, and Fact 1.1, we obtain the required isomorphism of posets.

By Proposition 1.3 it is enough to show that if $p', q'$ are obtained from $p, q$ as in (1) above, then $p \otimes q \sim_{\mathrm{D}} p' \otimes q'$. Denote by

$$\rho^p(x, z) := \bigl(\operatorname{rv}(f_i^p(x))_{i<m_p}, (h_j^p(x))_{j<n_p}, \operatorname{id}^p(z)\bigr)$$

the tuple of definable functions from (1), and similarly for $q$ and $p \otimes q$. By point (3) of Lemma 6.13 we may take as $(f_i^{p\otimes q})_{i<\omega}$ (a reindexing on $\omega$ of) the concatenation of $(f_i^p)_{i<\omega}$ with $(f_i^q)_{i<\omega}$. By the properties of $\otimes$, the concatenation of $(f_i^p(a))_{i<m_p}$ and $(f_i^q(b))_{i<m_q}$ is a basis of the vector space

$$\mathbb{Q}\bigl\langle\Gamma(\mathfrak{U})\bigl(v(f_i^p(a))\bigr)_{i<\omega}\bigl(v(f_i^q(b))\bigr)_{i<\omega}\bigr\rangle$$

over $\mathbb{Q}\,\Gamma(\mathfrak{U})$, and so as $(f_i^{p\otimes q})_{i<m_{p\otimes q}}$ we may take the concatenation of $(f_i^p)_{i<m_p}$ with $(f_i^q)_{i<m_q}$. Similarly, as $(h_j^{p\otimes q})_{j<n_{p\otimes q}}$ we may take the concatenation of the respective tuples for $p$ and $q$, and ultimately we obtain that as $\rho^{p\otimes q}$ we may take the concatenation of $\rho^p$ with $\rho^q$. By (1), we have $p \otimes q \sim_{\mathrm{D}} p' \otimes q'$ and we are done. □

For $\{\mathrm{k}, \Gamma\}$-expansions, we are in the setting of Section 4, so we may combine the above with, e.g., Theorem C or Corollary 4.15. We spell out two nice cases; the special subcases of ACVF and RCVF were previously known (see the introduction).

**Corollary 6.19** (Theorem A). *Let $T$ be a complete $\{\mathrm{k}\}$-$\{\Gamma\}$-expansion of a benign theory of valued fields where, for all $n > 1$, the group $\mathrm{k}^\times/(\mathrm{k}^\times)^n$ is finite. There is an isomorphism of posets $\widetilde{\operatorname{Inv}}(\mathfrak{U}) \cong \widetilde{\operatorname{Inv}}(\mathrm{k}(\mathfrak{U})) \times \widetilde{\operatorname{Inv}}(\Gamma(\mathfrak{U}))$. If $\otimes$ respects $\geq_{\mathrm{D}}$ in $\mathrm{k}$ and $\Gamma$, then $\otimes$ respects $\geq_{\mathrm{D}}$, and the above is an isomorphism of monoids.*

*Proof.* Apply Theorem 6.18. By Fact 6.2, if the extra structure on $\mathcal{RV}$ involves only $\mathrm{k}$ and $\Gamma$, and never both at the same time, then the sorts $\mathrm{k}$ and $\Gamma$ are orthogonal. As $\mathcal{RV}$ is an expanded pure short exact sequence, we conclude by Corollary 4.15. □

**Corollary 6.20.** *Let $T$ be a complete $\{\mathrm{k}\}$-$\{\Gamma\}$-expansion of a benign theory of valued fields, and let $\mathcal{A}_{\mathrm{k}}$ denote the family of sorts $(\mathrm{k}^\times/(\mathrm{k}^\times)^n)_{n\in\omega}$. For $\kappa \geq |L|$, there is an isomorphism of posets $\widetilde{\operatorname{Inv}}_\kappa(\mathfrak{U}) \cong \widetilde{\operatorname{Inv}}_\kappa(\mathcal{A}_{\mathrm{k}}(\mathfrak{U})) \times \widetilde{\operatorname{Inv}}_\kappa(\Gamma(\mathfrak{U}))$. If $\otimes$ respects $\geq_{\mathrm{D}}$ in $\mathcal{A}_{\mathrm{k}}$ and $\Gamma$, then $\otimes$ respects $\geq_{\mathrm{D}}$, and the above is an isomorphism of monoids.*

*Proof.* As in Corollary 6.19, but using Corollary 4.9 instead of Corollary 4.15. □

In special cases, results such as the previous corollaries may also be obtained by using domination by a family of sorts in the sense of [12, Definition 1.7] (see [23, Section 6]). This kind of domination was proven in the algebraically closed case in [17], in the real closed case in [12], and in the equicharacteristic zero case in [31].

In algebraically or real closed valued fields, the decomposition

$$\widetilde{\mathrm{Inv}}(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}(\mathrm{k}(\mathfrak{U})) \times \widetilde{\mathrm{Inv}}(\Gamma(\mathfrak{U}))$$

remains valid after passing to $T^{\mathrm{eq}}$, as can be shown using resolutions [12; 17; 23]. A natural question is whether Theorem 6.18 generalises to $T^{\mathrm{eq}}$, or at least to $T^{\mathcal{G}}$, the expansion of $T$ by the geometric sorts of [16].

**Question 6.21.** Let $T$ be an $\mathcal{RV}$-expansion of a theory of valued fields with enough saturated maximal models eliminating K-quantifiers. Are there conditions guaranteeing that the isomorphism $\widetilde{\mathrm{Inv}}(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}(\mathcal{RV}(\mathfrak{U}))$ holds in $T^{\mathcal{G}}$, or even in $T^{\mathrm{eq}}$? Does compatibility of $\geq_{\mathrm{D}}$ with $\otimes$ transfer?

## 7. Mixed characteristic henselian valued fields

Let K be henselian of characteristic $(0, \mathfrak{p})$ for $\mathfrak{p} \in \mathbb{P}$. For $n \in \omega$, we define $\mathfrak{m}_n := \{x \in \mathrm{K} \mid v(x) > v(\mathfrak{p}^n)\}$. Let $\mathrm{RV}_n$ be the multiplicative monoid $\mathrm{RV}_n := \mathrm{K}/(1+\mathfrak{m}_n)$ and $\mathrm{RV}_n^\times := \mathrm{RV}_n \setminus \{0\}$. For each $n$, denote by $\mathrm{rv}_n : \mathrm{K} \to \mathrm{RV}_n$ the quotient map. For $m > n$, we have natural maps $\mathrm{rv}_{m,n} : \mathrm{RV}_m \to \mathrm{RV}_n$, and the valuation $v : \mathrm{K} \to \Gamma$ induces maps $\mathrm{RV}_n \to \Gamma$, still denoted by $v$. The kernel $\mathrm{k}_n$ of $v$ fits in a short exact sequence $1 \to \mathrm{k}_n \to \mathrm{RV}_n \xrightarrow{v} \Gamma \to 0$. We have relations $\oplus_n$, defined analogously to $\oplus$, and again well defined precisely when $v(x+y) = \min\{v(x), v(y)\}$. For $n = 0$ we recover the notions from the previous section. The following generalises Fact 6.5.

**Fact 7.1.** A basis $(a_i)_i$ is separating if and only if, for each $n \in \omega$, each sum $\mathrm{rv}_n(d_0)\,\mathrm{rv}_n(a_{i_0}) \oplus_n \ldots \oplus_n \mathrm{rv}_n(d_\ell)\,\mathrm{rv}_n(a_{i_\ell})$ is well defined, if and only if this happens for $n = 0$. If this is the case, then the sum equals $\mathrm{rv}_n\bigl(\sum_{j \leq \ell} d_j\, a_{i_j}\bigr)$.

In this section, $L$ is a language as follows. We have sorts K, $\Gamma$ and, for each $n \in \omega$, sorts $\mathrm{k}_n, \mathrm{RV}_n$. There are function symbols $\mathrm{rv}_n : \mathrm{K} \to \mathrm{RV}_n$, $\iota : \mathrm{k}_n \to \mathrm{RV}_n$, $v : \mathrm{RV}_n \to \Gamma$. The sort K carries a copy of the language of rings, while the sort $\Gamma = \Gamma \cup \{\infty\}$ carries the (additive) language of ordered groups, together with an absorbing element $\infty$ and an extra constant symbol $v(\mathfrak{p})$. Each $\mathrm{RV}_n$ and $\mathrm{k}_n$ carries the (multiplicative) language of groups, together with an absorbing element 0 and a ternary relation symbol $\oplus_n$. We denote by $\mathcal{RV}_*$ the reduct to the sorts $\mathrm{k}_n, \mathrm{RV}_n, \Gamma$. There may be other arbitrary symbols on $\mathcal{RV}_*$, i.e., as long as they do not involve K.

An $\mathcal{RV}_*$*-expansion* of a theory $T'$ of henselian valued fields of characteristic $(0, \mathfrak{p})$ is a complete $L$-theory $T \supseteq T'$, with the sorts and symbols above interpreted in the natural way. Until the end of the section, $T$ denotes such a theory. We will freely confuse the sort $\mathrm{k}_n$ with the image of its embedding in $\mathrm{RV}_n$. By [4, Theorem B] (see also [14, Proposition 4.3]) $T$ eliminates K-quantifiers, so $\mathcal{RV}_*$ is fully embedded.

**Proposition 7.2.** *Suppose $T$ eliminates* K*-quantifiers and has enough saturated maximal models. For every $p \in S^{\mathrm{inv}}(\mathfrak{U})$ there is $q \in S^{\mathrm{inv}}_{\mathcal{RV}_*^\omega}(\mathfrak{U})$ such that $p \sim_{\mathrm{D}} q$. More precisely, let $p(x, z) \in S^{\mathrm{inv}}(\mathfrak{U}, M_0)$, where $x$ is a tuple of* K*-variables and $z$ a tuple of $\mathcal{RV}_*$-variables. Let $(a, c) \vDash p(x, z)$, let $M \succ M_0$ be $|M_0|^+$-saturated and maximally complete, and let $(f_i)_{i<\omega}$ be given by the $*$-type version of Lemma 6.13 applied to $a$ and $M$ (see Remark 6.15). Then $p \sim_{\mathrm{D}} q(y, t) := \mathrm{tp}(\mathrm{rv}_n(f_i(a))_{i,n<\omega}, c/\mathfrak{U})$, witnessed by*

$$r(x, z, y, t) := \mathrm{tp}(a, c, \mathrm{rv}_n(f_i(a))_{i,n<\omega}, c/M).$$

*If $\kappa \geq |L|$ is small, there is an isomorphism of posets $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}_\kappa(\mathcal{RV}_*(\mathfrak{U}))$. If $\otimes$ respects $\geq_{\mathrm{D}}$ in $\mathcal{RV}_*(\mathfrak{U})$, then the same holds in $\mathfrak{U}$, and the above is an isomorphism of monoids.*

*Proof.* Adapt the proofs of Lemma 6.13, Proposition 6.14 and Corollary 6.16, replacing Facts 6.2 and 6.5 by [4, Theorem B] and Fact 7.1 respectively. □

The assumptions of Proposition 7.2 are satisfied in a number of cases of interest. Besides the algebraically closed case, we note the following.

**Remark 7.3.** Every $\mathcal{RV}_*$-expansion of a finitely ramified henselian valued field has enough saturated maximal models.

*Proof.* Finite ramification ensures immediate extensions are precisely those where $\mathcal{RV}_*$ does not change. By this and [14, Proposition 4.3], maximal immediate extensions are elementary, and by [10, Corollary 4.29] they are also unique. We may therefore adapt the proof of Proposition 6.9, replacing $\mathcal{RV}$ with $\mathcal{RV}_*$. □

**Remark 7.4.** $\mathcal{RV}_*$ may be viewed as a short exact sequence of abelian structures, each consisting of an inverse system of abelian groups. Since $\Gamma$ is torsion-free, this sequence is pure.[6] Hence, the results from Section 4 apply to this setting, e.g., by taking as $\mathcal{F}$ the family of all pp formulas.

If k eliminates imaginaries, we can get rid of those arising from $\mathcal{F}$ and obtain a product decomposition. We state a special case as an example application of the results above. We thank the referee for pointing out the "moreover" part.

**Corollary 7.5.** *In the theory of the Witt vectors over $\mathbb{F}_{\mathrm{p}}^{\mathrm{alg}}$, the domination monoid is well defined. If $\kappa$ is a small infinite cardinal, then*

$$\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}_\kappa(\mathrm{k}(\mathfrak{U})) \times \widetilde{\mathrm{Inv}}_\kappa(\Gamma(\mathfrak{U})) \cong \hat{\kappa} \times \mathscr{P}_{\leq\kappa}\big(\mathrm{CS}^{\mathrm{inv}}(\Gamma(\mathfrak{U}))\big).$$

*Moreover*, $\widetilde{\mathrm{Inv}}(\mathfrak{U}) \cong \hat{\omega} \times \mathscr{P}_{<\omega}\big(\mathrm{CS}^{\mathrm{inv}}(\Gamma(\mathfrak{U}))\big)$.

[6] Another way of seeing this is that, in a saturated enough model of $T$, the valuation map has a section, inducing a compatible system of angular components, i.e., a splitting of $\mathcal{RV}_*$.

*Proof.* The residue field k is fully embedded. Moreover, $\mathrm{k}_n = W_n(\mathrm{k})^\times$ for each $n$, where $W_n(\mathrm{k})$ is the truncated ring of Witt vectors over k, and $\mathrm{k}_n$ is in definable bijection with $\mathrm{k}^{n-1} \times \mathrm{k}^\times$ (see [30, Corollary 1.62 and Proposition 1.67]). The computation of $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U})$ follows. As for $\widetilde{\mathrm{Inv}}(\mathfrak{U})$, using discreteness of the value group it is possible to build a prodefinable surjection $\mathrm{K} \to \mathrm{k}^\omega$ [30, proof of Remark 3.23]; together with the argument above, this gives the "moreover" part. □

**Remark 7.6.** The product decomposition fails for finitary types: the surjection $\mathrm{K} \to \mathrm{k}^\omega$ yields a 1-type in K dominating the type of an infinite independent k-tuple.

However, finitisation is possible in the case of the $\mathfrak{p}$-adics.

**Corollary 7.7** (Theorem E). *Let $T$ be a complete* $\{\Gamma\}$*-expansion of* $\mathrm{Th}(\mathbb{Q}_\mathfrak{p})$. *There is an isomorphism of posets* $\widetilde{\mathrm{Inv}}(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}(\Gamma(\mathfrak{U}))$. *If* $\otimes$ *respects* $\geq_{\mathrm{D}}$ *in* $\Gamma(\mathfrak{U})$, *then the same holds in* $\mathfrak{U}$, *and the above is also an isomorphism of monoids. In particular, in* $\mathrm{Th}(\mathbb{Q}_\mathfrak{p})$, $\otimes$ *respects* $\geq_{\mathrm{D}}$, *and* $(\widetilde{\mathrm{Inv}}(\mathfrak{U}), \otimes, \geq_{\mathrm{D}}) \cong \big(\mathscr{P}_{<\omega}\big(\mathrm{CS}^{\mathrm{inv}}(\Gamma(\mathfrak{U}))\big), \cup, \supseteq\big)$.

*Proof.* By Remark 7.3 we may apply Proposition 7.2. Since each $\mathrm{k}_n$ is finite, each $\mathrm{RV}_n$ is a finite cover of $\Gamma$, so each element of $\mathrm{RV}_n$ is interalgebraic with an element of $\Gamma$. Thus if $p(x, z) \in S^{\mathrm{inv}}(\mathfrak{U}, M_0)$, where $x$ is a tuple of K-variables and $z$ a tuple of $\mathcal{RV}_*$-variables, and if $ac \models p$, then $\dim_\mathbb{Q}\big(\mathbb{Q}\,\Gamma\big(\mathrm{dcl}(\mathfrak{U}(ac))\big)/\mathbb{Q}\,\Gamma(\mathfrak{U})\big) \leq |xz|$ by the Abhyankar inequality, so there is a finitary invariant type in $\Gamma$ which is interalgebraic with the type $q(y, t) \sim_{\mathrm{D}} p$ found in Proposition 7.2. We conclude by Proposition 1.3. The "in particular" part then follows from Corollary 3.34. □

The infinite ramification case remains open.

**Problem 7.8.** Compute $\widetilde{\mathrm{Inv}}(\mathfrak{U})$ in an infinitely ramified mixed characteristic henselian valued field that is not algebraically closed.

## 8. D-henselian valued fields with many constants

Here we deal with certain differential valued fields. As the proofs are adaptations of those in Section 6, we give sketches and leave it to the reader to fill in the details.

We let $T$ be a complete theory with sorts K, k, $\Gamma$, RV, as in Section 6, naturally interpreted, and use the notation $\mathcal{RV}$. The fields k and K have characteristic 0 and both carry a derivation $\partial$ (denoted by the same symbol), commuting with the residue map. The valued differential field K is *monotone*, i.e., $v(\partial x) \geq v(x)$, has *many constants*,[7] i.e., for every $\gamma \in \Gamma$ there is $x \in \mathrm{K}$ with $\partial x = 0$ and $v(x) = \gamma$, and is *D-henselian*, i.e., the following holds. If $P(X) \in \mathcal{O}\{X\} = \mathcal{O}[\partial^i X]_{i\in\omega}$ is a differential polynomial over the valuation ring $\mathcal{O}$, and $a \in \mathcal{O}$ is such that $v(P(a)) > 0$ and for some $i$ we have $v(\mathrm{d}P/\mathrm{d}(\partial^i X))(a) = 0$, then there is $b \in \mathcal{O}$ such that $P(b) = 0$ and $v(a - b) > 0$. The family of sorts $\mathcal{RV}$ may carry additional structure.

[7] Here we follow the terminology of [1]. In [28], this condition is called having *enough constants*.

The derivation $\partial$ on K induces a map $\partial_{\mathrm{RV}}$ on RV which, for all $\gamma \in \Gamma$, fixes $v^{-1}(\gamma) \cup \{0\}$ setwise, defined by $\partial_{\mathrm{RV}}(\mathrm{rv}(x)) = \mathrm{rv}(\partial(x))$ if $v(\partial(x)) = v(x)$, and $\partial_{\mathrm{RV}}(\mathrm{rv}(x)) = 0$ otherwise, which extends the derivation $\partial$ on k.

By [28, Theorem 6.4 and Corollary 5.8] (see also [1, Corollary 8.3.3]) the theory $T$ given by the list of properties above (in a fixed language) eliminates K-quantifiers.

**Proposition 8.1.** *The theory $T$ has enough saturated maximal models.*

*Proof sketch.* By [28, Remark 6.2], $k$ is linearly surjective in the terminology of [1], so by [1, Theorem 7.4.3] $T$ has uniqueness of maximal immediate extensions. The maximal immediate extension $N$ of $M$ is monotone and D-henselian by [1, Lemma 6.3.5 and Theorem 7.4.3] with many constants. As $T$ eliminates K-quantifiers, $M \prec N$, so the proofs of Proposition 6.9 and Corollary 6.11 may be adapted. □

**Theorem 8.2.** *Let $\kappa$ be a small infinite cardinal. There is an isomorphism of posets $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}_\kappa(\mathcal{RV}(\mathfrak{U}))$. If $\otimes$ respects $\geq_{\mathrm{D}}$ in $\mathcal{RV}(\mathfrak{U})$, then the same holds in $\mathfrak{U}$, and the above is also an isomorphism of monoids.*

*Proof sketch.* By elimination of K-quantifiers, $\mathcal{RV}(M)$ is fully embedded in $M$. If we replace "polynomial" by "differential polynomial", $\mathrm{K}(M)[a]$ by $\mathrm{K}(M)\{a\}$, and so on, in the statements of Lemma 6.13 and Proposition 6.14, essentially the same proofs go through. We can then conclude as in the proof of Corollary 6.16. □

**Lemma 8.3.** *$\partial_{\mathrm{RV}}$ is definable from the short exact sequence structure, the differential field structure on* k, *and a predicate for $C := \{c \in \mathrm{RV} \mid \partial_{\mathrm{RV}}(c) = 0\}$.*

*Proof.* Suppose $a \in \mathrm{RV}$ and $v(a) \notin \{0, \infty\}$. Since K has many constants, there is $c \in \mathrm{RV}(M)$ with $\partial_{\mathrm{RV}}(c) = 0$ and $v(c) = v(a)$. Then we have $a/c \in \mathrm{k}(\mathfrak{U})$ and $\partial_{\mathrm{RV}}(a) = c\partial(a/c)$. Because this does not depend on the choice of $c$, the function $y = \partial_{\mathrm{RV}}(x)$ is $\varnothing$-definable by the formula

$$\varphi(x, y) := \exists z \in C\big((v(z) = v(x)) \wedge (y = z\partial(x/z))\big).$$ □

If $L$ had a section of the valuation, or an angular component compatible with $\partial$, we could recover $C$ from the constant field of k, and conclude by (the $*$-type version of) Remark 4.5. Yet, the absence of definable splitting is not a serious obstacle. For simplicity, we only give a result in the model companion $\mathsf{VDF}_{\mathcal{EC}}$.

**Theorem 8.4** (Theorem F). *In $\mathsf{VDF}_{\mathcal{EC}}$, for every small infinite cardinal $\kappa$, the monoid $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U})$ is well defined, and we have isomorphisms*

$$\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}_\kappa(\mathrm{k}(\mathfrak{U})) \times \widetilde{\mathrm{Inv}}_\kappa(\Gamma(\mathfrak{U})) \cong \prod_{\delta(\mathfrak{U})}^{\leq\kappa} \hat{\kappa} \times \mathscr{P}_{\leq\kappa}\big(\mathrm{CS}^{\mathrm{inv}}(\Gamma(\mathfrak{U}))\big),$$

*where $\delta(\mathfrak{U})$ is a cardinal, and $\prod_{\delta(\mathfrak{U})}^{\le\kappa}\hat{\kappa}$ denotes the submonoid of $\prod_{\delta(\mathfrak{U})}\hat{\kappa}$ consisting of $\delta(\mathfrak{U})$-sequences with support of size at most $\kappa$.*

*Proof.* By Theorem 8.2 we reduce to $\mathcal{RV}$. Let $L_C := L_{\mathrm{ab}} \cup \{C\}$, with $C$ a unary predicate. Expand the language of $\mathcal{RV}$ by a predicate $C$ on each sort, interpreted as the constants in both k and RV and as the full $\Gamma$ in $\Gamma$, obtaining a short exact sequence of $L_C$-abelian structures (to be precise, of abelian structures augmented by an absorbing element, see Remark 4.10) expanded by the differential field structure on k and the order on $\Gamma$. By Lemma 8.3, we may apply the material from Section 4, say, by taking as a fundamental family that of all pp $L_C$-formulas, provided we show that $\mathcal{RV}$ is pure. If $M \vDash \mathsf{VDF}_{\mathcal{EC}}$ is $\aleph_1$-saturated then, since $M$ has many constants, we may find a section $s : \Gamma(M) \to \mathrm{RV}(M)$ of the valuation with image included in $C(\mathrm{RV}(M))$. Hence the short exact sequence $\mathcal{RV}(M)$ of $L_C$-abelian structures splits, so is pure by Remark 4.5. Since k is a model of $\mathsf{DCF}_0$, which eliminates imaginaries, we may get rid of the auxiliary sorts $\mathrm{A}_\varphi$. We conclude by Corollary 3.33 and the fact that $\mathsf{DCF}_0$ is $\omega$-stable multidimensional (see [22, Section 5] for the relation between our setting and that of domination via forking in stable theories). □

**Remark 8.5.** In $\mathsf{VDF}_{\mathcal{EC}}$, finitisation is not to be expected (e.g., by [26, Proposition 4.2]), and in fact not possible: one may construct a 1-type $p \in S^{\mathrm{inv}}_{\mathrm{K}}(\mathfrak{U})$ with $((v \circ \partial^n)_* p)_{n\in\omega}$ nonalgebraic and pairwise weakly orthogonal, and hence not domination-equivalent.

Computing the image of the home sort in finitely many variables seems difficult.

**Remark 8.6.** Most arguments in this section may be adapted to $\sigma$-henselian valued difference fields of residue characteristic 0. An analogue of Theorem 8.2 goes through, using quantifier reduction to $\mathcal{RV}$ and a $\sigma$-Kaplansky theory yielding uniqueness and elementarity of maximal immediate extensions [11, Theorems 5.8 and 7.3]. In every completion of the model companion of the *isometric* case (see [6]), in sufficiently saturated models there is a section of the valuation with values in the fixed field, and hence one may obtain the decomposition $\widetilde{\mathrm{Inv}}_\kappa(\mathfrak{U}) \cong \widetilde{\mathrm{Inv}}_\kappa(\mathrm{k}(\mathfrak{U})) \times \widetilde{\mathrm{Inv}}_\kappa(\Gamma(\mathfrak{U}))$, by regarding $\mathcal{RV}$ as a pure short exact sequence of $\mathbb{Z}[\sigma]$-modules, and using elimination of imaginaries in $\mathsf{ACFA}_0$. The same goes through in the *multiplicative* setting, provided that, in the notation of [25], $\rho$ is transcendental. This applies, e.g., to the model companion of the *contractive* case (see [3]).

## Acknowledgements

Mennuni thanks E. Hrushovski for the useful discussions around orthogonality of invariant types. We thank A. Gehret for Remark 3.8, and the referee for providing extensive and thorough feedback that helped improve our paper.

## References


[1] M. Aschenbrenner, L. van den Dries, and J. van der Hoeven, *Asymptotic differential algebra and model theory of transseries*, Annals of Mathematics Studies **195**, Princeton University Press, Princeton, NJ, 2017. MR Zbl

[2] M. Aschenbrenner, A. Chernikov, A. Gehret, and M. Ziegler, "Distality in valued fields and related structures", *Trans. Amer. Math. Soc.* **375**:7 (2022), 4641–4710. MR Zbl

[3] S. Azgin, "Valued fields with contractive automorphism and Kaplansky fields", *J. Algebra* **324**:10 (2010), 2757–2785. MR Zbl

[4] S. A. Basarab, "Relative elimination of quantifiers for Henselian valued fields", *Ann. Pure Appl. Logic* **53**:1 (1991), 51–74. MR Zbl

[5] W. Baur, "Die Theorie der Paare reell abgeschlossener Körper", pp. 25–34 in *Logic and algorithmic* (Zurich, 1980), Monogr. Enseign. Math. **30**, Univ. Genève, Geneva, 1982. MR Zbl

[6] L. Bélair, A. Macintyre, and T. Scanlon, "Model theory of the Frobenius on the Witt vectors", *Amer. J. Math.* **129**:3 (2007), 665–721. MR Zbl

[7] R. Cluckers and I. Halupczok, "Quantifier elimination in ordered abelian groups", *Confluentes Math.* **3**:4 (2011), 587–615. MR Zbl

[8] P. F. Conrad, "Regularly ordered groups", *Proc. Amer. Math. Soc.* **13** (1962), 726–731. MR Zbl

[9] F. Delon, *Quelques propriétés des corps valués en théorie des modèles*, Ph.D. thesis, Université Paris, 1982.

[10] L. van den Dries, "Lectures on the model theory of valued fields", pp. 55–157 in *Model theory in algebra, analysis and arithmetic*, Lecture Notes in Math. **2111**, Springer, Heidelberg, 2014. MR Zbl

[11] S. Durhan and G. Onay, "Quantifier elimination for valued fields equipped with an automorphism", *Selecta Math.* (*N.S.*) **21**:4 (2015), 1177–1201. MR Zbl

[12] C. Ealy, D. Haskell, and J. Maříková, "Residue field domination in real closed valued fields", *Notre Dame J. Form. Log.* **60**:3 (2019), 333–351. MR Zbl

[13] D. Evans, A. Pillay, and B. Poizat, "Le groupe dans le groupe", *Algebra i Logika* **29**:3 (1990), 368–378, 382. MR

[14] J. Flenner, "Relative decidability and definability in Henselian valued fields", *J. Symbolic Logic* **76**:4 (2011), 1240–1260. MR Zbl

[15] Y. Halevi and A. Hasson, "Eliminating field quantifiers in strongly dependent Henselian fields", *Proc. Amer. Math. Soc.* **147**:5 (2019), 2213–2230. MR Zbl

[16] D. Haskell, E. Hrushovski, and D. Macpherson, "Definable sets in algebraically closed valued fields: elimination of imaginaries", *J. Reine Angew. Math.* **597** (2006), 175–236. MR Zbl

[17] D. Haskell, E. Hrushovski, and D. Macpherson, *Stable domination and independence in algebraically closed valued fields*, Lecture Notes in Logic **30**, Cambridge University Press, Cambridge, 2008. MR Zbl

[18] M. Hils, M. Kamensky, and S. Rideau, "Imaginaries in separably closed valued fields", *Proc. Lond. Math. Soc.* (3) **116**:6 (2018), 1457–1488. MR Zbl

[19] I. Kaplansky, "Maximal fields with valuations", *Duke Math. J.* **9** (1942), 303–321. MR Zbl

[20] F.-V. Kuhlmann, "Quantifier elimination for Henselian fields relative to additive and multiplicative congruences", *Israel J. Math.* **85**:1-3 (1994), 277–306. MR Zbl

[21] R. Mennuni, "Invariant types in model theory", *Bull. Symb. Log.* **26**:3-4 (2020), 296–297. MR Zbl

[22] R. Mennuni, "Product of invariant types modulo domination-equivalence", *Arch. Math. Logic* **59**:1-2 (2020), 1–29. MR Zbl

[23] R. Mennuni, "The domination monoid in o-minimal theories", *J. Math. Log.* **22**:1 (2022), art. id. 2150030. MR Zbl

[24] R. Mennuni, "Weakly binary expansions of dense meet-trees", *MLQ Math. Log. Q.* **68**:1 (2022), 32–47. MR Zbl

[25] K. Pal, "Multiplicative valued difference fields", *J. Symbolic Logic* **77**:2 (2012), 545–579. MR Zbl

[26] S. Rideau, "Imaginaries and invariant types in existentially closed valued differential fields", *J. Reine Angew. Math.* **750** (2019), 157–196. MR Zbl

[27] A. Robinson and E. Zakon, "Elementary properties of ordered abelian groups", *Trans. Amer. Math. Soc.* **96** (1960), 222–236. MR Zbl

[28] T. Scanlon, "Quantifier elimination for the relative Frobenius", pp. 323–352 in *Valuation theory and its applications*, *II* (Saskatoon, SK, 1999), Fields Inst. Commun. **33**, Amer. Math. Soc., Providence, RI, 2003. MR Zbl

[29] S. Shelah, *Classification theory and the number of nonisomorphic models*, 2nd ed., Studies in Logic and the Foundations of Mathematics **92**, North-Holland Publishing Co., Amsterdam, 1990. MR Zbl

[30] P. Touchard, "Burden in Henselian valued fields", *Ann. Pure Appl. Logic* **174**:10 (2023), art. id. 103318. MR Zbl

[31] M. Vicaria, "Residue field domination in henselian valued fields of equicharacteristic zero", preprint, 2021. arXiv 2109.08243

[32] M. Vicaría, "Elimination of imaginaries in ordered abelian groups with bounded regular rank", *J. Symb. Log.* **88**:4 (2023), 1639–1654. MR Zbl

[33] V. Weispfenning, "Quantifier eliminable ordered abelian groups", pp. 113–126 in *Algebra and order* (Luminy-Marseille, 1984), Res. Exp. Math. **14**, Heldermann, Berlin, 1986. MR Zbl

[34] E. Zakon, "Generalized archimedean groups", *Trans. Amer. Math. Soc.* **99** (1961), 21–40. MR Zbl

Received November 10, 2021. Revised February 13, 2024.

MARTIN HILS
INSTITUT FÜR MATHEMATISCHE LOGIK UND GRUNDLAGENFORSCHUNG
UNIVERSITÄT MÜNSTER
MÜNSTER
GERMANY
hils@uni-muenster.de

ROSARIO MENNUNI
DIPARTIMENTO DI MATEMATICA
UNIVERSITÀ DI PISA
PISA
ITALY
r.mennuni@posteo.net

PACIFIC JOURNAL OF MATHEMATICS
Vol. 328, No. 2, 2024
https://doi.org/10.2140/pjm.2024.328.325

# INVERSE SEMIGROUP FROM METRICS ON DOUBLES III: COMMUTATIVITY AND (IN)FINITENESS OF IDEMPOTENTS

VLADIMIR MANUILOV

**We have shown recently that, given a metric space $X$, the coarse equivalence classes of metrics on the two copies of $X$ form an inverse semigroup $M(X)$. Here we study the property of idempotents in $M(X)$ of being finite or infinite, which is similar to this property for projections in $C^*$-algebras. We show that if $X$ is a free group then the unit of $M(X)$ is infinite, while if $X$ is a free abelian group then it is finite. As a by-product, we show that the inverse semigroup $M(X)$ is not a quasiisometry invariant. We also show that $M(X)$ is commutative if it is Clifford, and give a geometric description of spaces $X$ for which $M(X)$ is commutative.**

## 1. Introduction

Given metric spaces $(X, d_X)$ and $(Y, d_Y)$, a metric $d$ on $X \sqcup Y$ that extends the metrics $d_X$ on $X$ and $d_Y$ on $Y$, depends only on the values $d(x, y)$, $x \in X$, $y \in Y$, and it may be not easy to check which functions $d : X \times Y \to (0, \infty)$ determine a metric on $X \sqcup Y$. The problem of description of all such metrics is difficult due to the lack of a nice algebraic structure on the set of metrics, but, passing to coarse equivalence of metrics, we get an algebraic structure, namely, that of an inverse semigroup [Manuilov 2021a]. Recall that two metrics, $b$, $d$, on a space $Z$ are coarsely equivalent, $b \sim d$, if there exist monotone functions $\varphi, \psi : [0, \infty) \to [0, \infty)$ such that

$$\lim_{t\to\infty} \varphi(t) = \lim_{t\to\infty} \psi(t) = \infty$$

and

$$\varphi(d(z_1, z_2)) \leq b(z_1, z_2) \leq \psi(d(z_1, z_2))$$

for any $z_1, z_2 \in Z$. We denote by $[d]$ the coarse equivalence class of a metric $d$. Our standard reference on metric spaces is [Burago et al. 2001].

Let $\mathcal{M}(X, Y)$ denote the set of all metrics $d$ on $X \sqcup Y$ such that:

- The restriction of $d$ onto $X$ and $Y$ are $d_X$ and $d_Y$ respectively.
- $\inf_{x \in X, y \in Y} d(x, y) > 0$.

The research was supported by RSF, project No. 23-21-00068.
*MSC2020:* 20M18, 51F30.
*Keywords:* inverse semigroup, metric, finite projection.

Coarse equivalence classes of metrics in $\mathcal{M}(X, Y)$ can be considered as morphisms from $X$ to $Y$ [Manuilov 2019], where the composition $b \circ d$ of a metric $d$ on $X \sqcup Y$ and a metric $b$ on $Y \sqcup Z$ is given by the metric determined by

$$(b \circ d)(x, z) = \inf_{y \in Y}[d(x, y) + b(y, z)], \quad x \in X, z \in Z.$$

When $Y = X$, we call $X \sqcup X$ the double of $X$. In what follows we identify the double of $X$ with $X \times \{0, 1\}$, and write $X$ for $X \times \{0\}$ (resp., $x$ for $(x, 0)$) and $X'$ for $X \times \{1\}$ (resp., $x'$ for $(x, 1)$). We also write $\mathcal{M}(X)$ for $\mathcal{M}(X, X)$.

The main result of [Manuilov 2021a] is that the semigroup $M(X) = \mathcal{M}(X)/\sim$ (with respect to this composition) of coarse equivalence classes of metrics on the double of $X$ is an inverse semigroup with the unit element $\mathbf{1}$ and the zero element $\mathbf{0}$, and the unique pseudoinverse for $[d] \in M(X)$ is the coarse equivalence class of the metric $d^*$ given by $d^*(x, y') = d(x', y)$, $x, y \in X$.

Recall that a semigroup $S$ is an inverse semigroup if for any $s \in S$ there exists a unique $t \in S$ (denoted by $s^*$ and called a pseudoinverse) such that $s = sts$ and $t = tst$ [Lawson 1998]. Philosophically, inverse semigroups describe local symmetries in a similar way as groups describe global symmetries, and technically, the construction of the (reduced) group $C^*$-algebra of a group generalizes to that of the (reduced) inverse semigroup $C^*$-algebra [Paterson 1999]. It is known that any two idempotents in an inverse semigroup $S$ commute, and that there is a partial order on $S$ defined by $s \leq t$ if $s = ss^*t$. Our standard references for inverse semigroups are [Lawson 1998] and [Howie 1995].

Close relation between inverse semigroups and $C^*$-algebras allows to use classification of projections in $C^*$-algebras for idempotents in inverse semigroups. Namely, as in $C^*$-algebra theory, we call two idempotents, $e, f \in E(S)$ von Neumann equivalent (and write $e \sim f$) if there exists $s \in S$ such that $s^*s = e$, $ss^* = f$. An idempotent $e \in E(S)$ is called *infinite* if there exists $f \in E(S)$ such that $f \preceq e$, $f \neq e$, and $f \sim e$. Otherwise $e$ is *finite*. An inverse semigroup is *finite* if every idempotent is finite, and is *weakly finite* if it is unital and the unit is finite. A commutative unital inverse semigroup is patently finite.

In [Manuilov 2021b] we gave a geometric description of idempotents in the inverse semigroup $M(X)$ (there are two types of idempotents, named type I and type II) and showed in Lemma 3.3 of [loc. cit.] that the type is invariant under the von Neumann equivalence. In Part I, we study the property of weak finiteness for $M(X)$ (i.e., finiteness of the unit element) and discuss its relation to geometric properties of $X$.

We start with several examples of finite or infinite idempotents, and then show that if $X$ is a free group then $M(X)$ is not weakly finite, while if $X$ is a free abelian group then it is weakly finite. We also show that the inverse semigroup $M(X)$ is not a quasiisometry invariant. The property of being weakly finite is also not a coarse invariant. We don't know if it is a quasiisometry invariant.

In Part II, we give a geometric description of spaces, for which the inverse semigroup $M(X)$ is commutative, and show that the condition of being a Clifford inverse semigroup (i.e., that $ss^* = s^*s$ for any $s \in S$) guarantees that $M(X)$ is commutative.

# Part I. Weak finiteness of $M(X)$

## 2. Geometric description of weak finiteness

Two maps $f, g : X \to X$ are called *equivalent* if there exists $C > 0$ such that $d_X(f(x), g(x)) < C$ for any $x \in X$. A map $f : X \to X$ is an *almost isometry* if there exists $C > 0$ such that:

- $|d_X(f(x), f(y)) - d_X(x, y)| < C$ for any $x, y \in X$.
- For any $y \in X$ there exists $x \in X$ such that $d_X(f(x), y) < C$.

(The latter condition provides existence of an "inverse" map $g : X \to X$ such that $f \circ g$ and $g \circ f$ are equivalent to the identity map; this map is also an almost isometry, but with possibly greater constant $C$; if $f$ is surjective then this property is superfluous.) We call $f$ a $C$-almost isometry when we need an explicit value of the constant $C$.

In a metric space, it makes sense to define equivalence of subsets as follows: for $A, B \subset X$ we say that $A \sim B$ if there exists $C > 0$ such that $A \subset N_C(B)$ and $B \subset N_C(A)$, where $N_C(Y) = \{x \in X : d_X(x, Y) < C\}$ denotes the $C$-neighborhood of $Y \subset X$. In particular, a subset $A \subset X$ is equivalent to $X$ if it is a $C$-net, i.e., if there exists $C > 0$ such that for any $x \in X$ there exists $y \in A$ with $d_X(x, y) < C$.

**Theorem 2.1.** *The following are equivalent*:

(1) *$M(X)$ is weakly finite.*

(2) *If there exists an almost isometry $X \to A \subset X$ then the subset $A$ is equivalent to $X$.*

*Proof.* For $B \subset X$, set

$$d^B(x, y') = \inf_{u \in B} [d_X(x, u) + 1 + d_X(u, y)].$$

Then $d_X$ is a metric on the double of $X$, and $[d^B]$ is an idempotent in $M(X)$ [Manuilov 2021a]. It was shown in Lemma 3.3 of [Manuilov 2021b] that if $d$ is a metric on the double of $X$ and $[d^*][d] = [d^B]$ then there exists $A \subset X$ such that $[d][d^*] = [d^A]$.

Suppose that there exists a $C$-almost isometry $f : X \to A$ for some $A \subset X$ and for some $C > 0$. Then set

$$d(x, y') = \inf_{u \in X} [d_X(x, u) + C + d_X(f(u), y)].$$

It was shown in Lemma 3.2 of [Manuilov 2019] that this defines a metric on the double of $X$. Then

$$\begin{aligned} d^* \circ d(x, x') &= \inf_{y \in X}[d(x, y') + d^*(y, x')] = 2 \inf_{y \in X} d(x, y') \\ &= 2 \inf_{u, y \in X}[d_X(x, u) + C + d_X(f(u), y)] \le 2C \end{aligned}$$

(we might take $y = f(u)$ and $u = x$), hence $[d^*][d] = \mathbf{1}$.

$$\begin{aligned} d \circ d^*(x, x') &= \inf_{y \in X}[d^*(x, y') + d(y, x')] = 2 \inf_{y \in X} d(y, x') \\ &= 2 \inf_{u, y \in X}[d_X(y, u) + C + d_X(f(u), x)] \\ &= 2C + 2 \inf_{u \in X} d_X(x, f(u)) = 2C + 2d_X(x, f(X)) \end{aligned}$$

(taking $u = y$), so, using that $f(X)$ is $C$-dense in $A$, we see that

$$|d \circ d^*(x, x') - d_X(x, A)| \le 4C,$$

hence $[d][d^*] = [d^A]$ by Proposition 3.2 of [Manuilov 2021a]. If $M(X)$ is weakly finite then $[d^A] = \mathbf{1}$, hence, by Proposition 4.2 of [Manuilov 2021a], $X$ lies in a $C$-neighborhood of $A$ for some $C > 0$.

In the opposite direction, let $M(X)$ be not weakly finite. Then there exists a metric $d$ on the double of $X$ such that $[d^*][d] = \mathbf{1}$, but $[d][d^*] \ne \mathbf{1}$. By Lemma 3.3 of [Manuilov 2021b], $[d \circ d^*] = [d^A]$, where $A \subset X$ is constructed as follows. As $[d^*][d] = \mathbf{1}$, there exists $C > 0$ such that

$$d^* \circ d(x, x') = 2d(x, X') < 2C$$

for any $x \in X$, i.e., for any $x \in X$ there exists $y \in X$ such that $d(x, y') < C$. Then $A = \{y \in X : d(X, y') < C\}$.

Given $x \in X$, there may be several $y$'s such that $d(x, y') < C$. Choose one of them and set $f(x) = y$. It follows from

$$d(X, f(x)') \le d(x, f(x)') < C$$

that $f(x) \in A$. If $x_1, x_2 \in X$ then the triangle inequality for the quadrangle $x_1, x_2, f(x_1)', f(x_2)'$ gives $|d_X(f(x_1), f(x_2)) - d_X(x_1, x_2)| < 2C$. If $z \in A$ then $d(X, z') < C$, hence there exists $x \in X$ such that $d(x, z') < C$. Then

$$d_X(z, f(x)) = d_X(z', f(x)') \le d(z', x) + d(x, f(x)') < 2C,$$

hence $f$ is a $2C$-almost isometry. Finally, the condition $[d][d^*] \ne \mathbf{1}$ implies that $A$ is not equivalent to $X$. □

## 3. Some examples

The following example shows that in $M(X)$, for an appropriate $X$, we can imitate examples of partial isometries and projections in a Hilbert space.

**Example 3.1.** Let $l^1(\mathbb{N})$ be the space of infinite $l^1$ sequences, with the metric given by the $l^1$-norm, and let

$$X_n = \{(0, \dots, 0, t, 0, \dots) : t \in [0, \infty)\}$$

with $t$ at the $n$-th place, $n \in \mathbb{N}$. Set

$$X = \bigcup_{n=1}^{\infty} X_n \subset l^1(\mathbb{N}), \quad A = \bigcup_{n=2}^{\infty} X_n \subset l^1(\mathbb{N}).$$

The set $A$ is not equivalent to $X$, and there is an obvious isometry $f : X \to A$ that isometrically maps $X_n$ to $X_{n+1}$, $n \in \mathbb{N}$. Thus, $\mathbf{1}$ is infinite. Let $d$ be a metric on the double of $X$ induced by $f$. Although $d$ seems similar to a one-sided shift in a Hilbert space, it behaves differently: $h = [d \circ d^*]$ is orthogonally complemented, i.e., there exists $e \in E(M(X))$ such that $e \vee h = \mathbf{1}$, $e \wedge h = \mathbf{0}$ (recall that $E(M(X))$ is a lattice [Manuilov 2021b]), but the complement $e$ is not a minimal idempotent, i.e., there exists a lot of idempotents $j \in E(M(X))$ such that $j \leq e$, $j \neq e$.

On the other hand, if $X \subset [0, \infty)$ with the standard metric then the inverse semigroup $M(X)$ is commutative [Manuilov 2021a, Proposition 7.1], hence any idempotent can be equivalent only to itself, hence is finite. In Part II, we shall give a geometric description of all metric spaces with commutative $M(X)$, which is then patently finite.

The next example shows that the picture may be more complicated.

**Proposition 3.2.** *There exists an amenable space $X$ of bounded geometry and $s \in M(X)$ such that $s^*s = \mathbf{1}$, but $ss^* \neq \mathbf{1}$.*

*Proof.* Consider $l_\infty(\mathbb{N})$ with sup metric, and let

$$x_n = (\log 2, \log 3, \dots, \log(n-1), \log n, 0, 0, \dots) \in l_\infty(\mathbb{N}),$$
$$X = \{x_n : n \in \mathbb{N}\} \subset l_\infty(\mathbb{N}); \quad A = \{x_{2n} : n \in \mathbb{N}\}.$$

Set

$$f : X \to A; \quad f(x_n) = x_{2n}, n \in \mathbb{N}.$$

Given $n < m$, we have

$$d_X(x_n, x_m) = \log m, \quad d_X(f(x_n), f(x_m)) = d_X(x_{2n}, x_{2m}) = \log(2m),$$

hence

$$d_X(f(x_n), f(x_m)) - d_X(x_n, x_m) = \log(2m) - \log m = \log 2.$$

As $f$ is surjective, it is an almost isometry.

Note that

$$d_X(x_{2n-1}, x_{2m}) = \begin{cases} \log(2m) & \text{if } 2n-1 < 2m, \\ \log(2n-1) & \text{if } 2n-1 > 2m, \end{cases}$$

hence

$$d_X(x_{2n-1}, A) = \inf_{m\in\mathbb{N}} d_X(x_{2n-1}, x_{2m}) = \log(2m),$$

thus $A \subset X$ is not equivalent to $X$, hence $M(X)$ is not weakly finite.

Note that $X$ is amenable. Set $F_n = \{x_1, \dots, x_n\} \subset X$. Let $N_r(A)$ denote the $r$-neighborhood of the set $A$. Then $N_r(F_n) \setminus F_n$ is empty when $\log(n+1) > r$, hence $\{F_n\}_{n\in\mathbb{N}}$ is a Følner sequence. For $r = \log m$, the ball $B_r(x_n)$ of radius $r$ centered at $x_n$ contains either no other points besides $x_n$ (if $n \geq m+1$), or it consists of the points $x_1, \dots, x_m$ (if $n \leq m$), hence the metric on $X$ is of bounded geometry. In fact, this space is of asymptotic dimension zero. □

## 4. Case of free groups

In this section we show that $M(X)$ is not weakly finite for two classes of groups, both of which include free groups.

Let $X = \Gamma$ be a finitely generated group with the word length metric $d_X$. Consider the following Property (I):

(i1) $X = Y \sqcup Z$, and for any $D > 0$ there exists $z \in Z$ such that $d_X(z, Y) > D$.

(i2) There exist $g, h \in \Gamma$ such that $gY \subset Y$, $hZ \subset Y$ and $gY \cap hZ = \varnothing$.

(i3) There exists $C > 0$ such that $|d_X(gy, hz) - d_X(y, z)| < C$ for any $y \in Y$, $z \in Z$.

Property (I) looks similar to nonamenability, but, at least formally, is neither stronger nor weaker than nonamenability.

**Lemma 4.1.** *The free group $\mathbb{F}_2$ on two generators satisfies Property (I).*

*Proof.* Let $a$ and $b$ be the generating elements of $\mathbb{F}_2$, and let $Y \subset X$ be the set of all reduced words in $a$, $a^{-1}$, $b$ and $b^{-1}$ that begin with $a$ or $a^{-1}$, $Z = X \setminus Y$. Let $g = ab$, $h = a^2$. Clearly, $gY \subset Y$ and $hZ \subset Y$.

If $z$ begins with $a^n$, $n > D$, then $d_X(z, Y) \geq n$.

If $y \in Y$, $z \in Z$ then

$$d_X(aby, a^2z) = |y^{-1}b^{-1}a^{-1}a^2z| = |y^{-1}b^{-1}az| = |y^{-1}z| + 2 = d_X(y, z) + 2,$$

as the word $y^{-1}b^{-1}az$ cannot be reduced any further ($y^{-1}$ ends with $a^{\pm}$, and $z$ either begins with $b^{\pm}$, or is an empty word). □

**Theorem 4.2.** *Let $X = \Gamma$ be a group with Property (I). Then $X$ is not weakly finite.*

*Proof.* We shall prove that there exists an almost isometry $f : X \to A \subset X$, where $A$ is not equivalent to $X$.

Let $X = Y \sqcup Z$, $g, h \in \Gamma$ satisfy the conditions of Property (I). Define a map $f : X \to X$ by setting

$$f(x) = \begin{cases} gx & \text{if } x \in Y; \\ hx & \text{if } x \in Z. \end{cases}$$

The maps $f|_Y$ and $f|_Z$ are left multiplications by $g$ and $h$, respectively, hence are isometries. If $y \in Y$, $z \in Z$ then (i3) holds for some $C > 0$, hence

$$|d_X(f(x), f(y)) - d_X(x, y)| < C$$

holds for any $x, y \in X$. Set $A = f(X)$, then $f$ is an almost isometry from $X$ to $A$. By (i1), $A$ is not equivalent to $X$. □

Our next argument also works for free groups, but refers to non-co-Hopfian groups, i.e., groups isomorphic to a proper subgroup.

**Theorem 4.3.** *Let $X = G$ be a finitely generated group with the word length metric, and let $A = H \subset G$ be an infinite index subgroup. Suppose that there exists a map $f : G \to H$ that is both an isomorphism and an almost isometry. Then $X$ is not weakly finite.*

*Proof.* We need only to check that $A$ is not equivalent to $X$. Suppose it is, i.e., there exists $C > 0$ such that for any $x \in X$ there exists $y \in H$ with $d_X(x, y) < C$. As $H$ is of infinite index, there are infinitely many different cosets $Hg_i$, $g_i \in G$, $i \in \mathbb{N}$. Let $h_i \in H$ satisfy $d_X(g_i, h_i) < C$, $i \in \mathbb{N}$, which means that $|g_i^{-1}h_i| < C$. As $G$ is finitely generated, the set of group elements $g$ with $|g| < C$ is finite, so there exist $i \neq j$ such that $g_i^{-1}h_i = g_j^{-1}h_j$, or, equivalently, $h_i^{-1}g_i = h_j^{-1}g_j$, hence $Hg_i = Hg_j$ — a contradiction. □

**Remark 4.4.** It is easy to find examples of isomorphisms that are also almost isometries. Indeed, if $\gamma \in G$ then the map $f(g) = \gamma^{-1}g\gamma$ is an example: it follows from $d_X(f(g_1), f(g_2)) = |\gamma^{-1}g_1^{-1}g_2\gamma|$ and $d_X(g_1, g_2) = |g_1^{-1}g_2|$ that

$$|d_X(f(g_1), f(g_2)) - d_X(g_1, g_2)| \leq 2|\gamma|$$

for any $g_1, g_2 \in G$. There are many examples when the subgroup $H = \gamma^{-1}G\gamma$ is of infinite index in $G$, e.g., if $G$ is a free group, and $\gamma$ is not a generator.

## 5. Case of abelian groups

A positive result is given by the following theorem.

**Theorem 5.1.** *Let $X = \mathbb{R}^n$, with a norm $\|\cdot\|$, and let the metric $d_X$ be determined by the norm $\|\cdot\|$. Then $M(X)$ is weakly stable.*

*Proof.* We have to show that if $f : X \to X$ is a $C$-almost isometry for some $C > 0$ then $f(X)$ is equivalent to $X$. Suppose the contrary: for any $n \in \mathbb{N}$ there exists $x_n \in X$ such that $d_X(x_n, f(X)) > n$.

First, note that we can replace $f$ by another almost isometry $g$, which is continuous and close to $f$. Namely, choose a triangulation of $X$ by simplices with length of edges greater than $C$ and with a uniform lower bound for their volumes. Then set $g(v) = f(v)$ for all vertices and extend this map to the inner points of the simplices by linearity. Then $g : X \to X$ is continuous and there exists $C' > 0$ depending on the dimension of $X$ and on the norm $\|\cdot\|$, such that $d_X(f(x), g(x)) < C'$ for any $x \in X$. As $f$ was a $C$-almost isometry, $g$ is a $D$-almost isometry, where $D = 2C' + C$.

Let $x_0$ denote the origin of $X$. Without loss of generality, we may assume that $f(x_0) = x_0$ (we may compose $f$ with a translation).

Denote by $S_R$ the sphere of radius $R$ centered at $x_0$. Then $g(x)$ lies between $S_{R-D}$ and $S_{R+D}$ for any $x \in S_R$. Let $d_X(x_0, x_n) = R_n$. Clearly, $\lim_{n\to\infty} R_n = \infty$. Passing to a subsequence, we may assume that $\lim_{n\to\infty} R_{n+1} - R_n = \infty$. Then, once again, we can replace $g$ by a continuous $D'$-almost isometry $h : X \to X$ with $\sup_{x\in X} d_X(f(x), h(x)) < D'$ for some $D' > 0$ such that $h(S_{R_n}) \subset S_{R_n}$.

As $d_X(x_n, f(X)) > n$, $d_X(x_n, h(X)) > n - D'$, hence $x_n \notin h(S_{R_n})$ when $n > D'$. Thus, the map $h|_{S_{R_n}} : S_{R_n} \to S_{R_n}$ is not surjective. Then, by the Borsuk–Ulam theorem, there exists a pair of antipodal points $y_1, y_2 \in S_{R_n}$ such that $h(y_1) = h(y_2) = z$. But this contradicts the almost isometricity of $h$:

$$|d_X(h(y_1), h(y_2)) - d_X(y_1, y_2)| = |d_X(z, z) - d_X(y_1, y_2)| = |0 - 2R_n| = 2R_n$$

is not bounded. □

**Corollary 5.2.** *Let $X = \mathbb{Z}^n$ with an $l_p$-metric, $1 \le p \le \infty$. Then $M(X)$ is weakly finite.*

*Proof.* By Proposition 9.2 of [Manuilov 2021a], $M(\mathbb{Z}^n) = M(\mathbb{R}^n)$. □

**Corollary 5.3.** *$M(X)$ is weakly finite for any finitely generated free abelian group $X$ with a word length metric with respect to any finite set of generators.*

## 6. $M(X)$ doesn't respect equivalences

**Proposition 6.1.** *The inverse semigroup $M(X)$ is not a coarse invariant.*

*Proof.* The space $X$ from Proposition 3.2 is coarsely equivalent to the space $Y = \{n^2 : n \in \mathbb{N}\}$ with the standard metric, which we denote by $b_X$. Indeed, for $n < m$, we have $b_X(x_n, x_m) = m^2 - n^2$ and $d_X(x_n, x_m) = \log(m + 1)$. As $m^2 - (m - 1)^2 = 2m - 1 > \log(m + 1)$ for $m > 1$, we have $d_X(x, y) \le b_X(x, y)$ for any $x, y \in X$, and taking $f(t) = 2e^t$, we have $b_X(x, y) \le f(d_X(x, y))$ for any $x, y \in X$.

For the metric $d_X$ from Proposition 3.2, the inverse semigroup $M(X, d_X)$ is not commutative ($[d^*d] \neq [dd^*]$), while the inverse semigroup $M(X, b_X)$ is commutative by Proposition 7.1 of [Manuilov 2021b]. □

**Theorem 6.2.** *The inverse semigroup $M(X)$ is not a quasiisometry invariant.*

*Proof.* Let $X = \mathbb{N}$ be endowed with the metric $b_X$ given by $b_X(n, m) = |2^n - 2^m|$, $n, m \in \mathbb{N}$, and let $y_n = s(n)4^{[n/2]}$, where $s(n) = (-1)^{[(n-1)/2]}$ and $[t]$ is the greatest integer not exceeding $t$. Let $d_X$ be the metric on $X$ given by $d_X(n, m) = |y_n - y_m|$, $n, m \in \mathbb{N}$. The two metrics are quasiisometric. Indeed, suppose that $n > m$. If $s(n) = -s(m)$ then

$$d_X(n, m) = 4^{[n/2]} + 4^{[m/2]} \leq 4^{n/2+1} + 4^{m/2+1} = 4(2^n + 2^m) \leq 12b_X(n, m);$$
$$d_X(n, m) = 4^{[n/2]} + 4^{[m/2]} \geq 4^{n/2} + 4^{m/2} \geq 2^n - 2^m = b_X(n, m).$$

We use here that $(2^r + 1)/(2^r - 1) \leq 3$ for any $r = n - m \in \mathbb{N}$. If $s(n) = s(m)$ then

$$d_X(n, m) = 4^{[n/2]} - 4^{[m/2]} \leq 4^{n/2+1} - 4^{m/2} = 4 \cdot 2^n - 2^m \leq 7b_X(n, m).$$

We use here that $(4 \cdot 2^r - 1)/(2^r - 1) \leq 7$ for any $r = n - m \in \mathbb{N}$. To obtain an estimate in other direction, note that $s(n) = s(m)$ implies that $[n/2] \geq [m/2] + 1$, and that $n - m \neq 2$. If $n = m + 1$ then

$$d_X(m + 1, m) = 3 \cdot 4^{[m/2]} \geq \tfrac{3}{2} \cdot 2^m = \tfrac{3}{2} b_X(m + 1, m),$$

If $n \geq m + 3$ then

$$d_X(n, m) = 4^{[n/2]} - 4^{[m/2]} \geq 4^{n/2} - 4^{m/2+1} = 2^n - 4 \cdot 2^m \geq \tfrac{4}{7} b_X(n, m).$$

We use here that $(2^r - 4)/(2^r - 1) \geq \frac{4}{7}$ for any $r = n - m \geq 3$. Thus,

$$\tfrac{3}{7} b_X(n, m) \leq d_X(n, m) \leq 12 \cdot b_X(n, m)$$

for any $n, m \in \mathbb{N}$, so the two metrics are quasiisometric.

We already know that $M(X, b_X)$ is commutative, so it remains to expose two noncommuting elements in $M(X, d_X)$.

Let

$$X = \{(y_n, 0) : n \in \mathbb{N}\}, \quad X' = \{(-y_n, 1) : n \in \mathbb{N}\},$$

and let $d$ be the metric on $X \sqcup X'$ induced from the standard metric on the plane $\mathbb{R}^2$, $s = [d]$. Note that $-y_n = y_{n-1}$ if $y_n > 0$ and $n > 1$, and $-y_n = y_{n+1}$ if $y_n < 0$. Hence, $d^* = d$ and $s^2 = \mathbf{1}$.

Let

$$A_+ = \{y_n : n \in \mathbb{N}; y_n > 0\}, \quad A_- = \{y_n : n \in \mathbb{N}; y_n < 0\},$$

$X = A_+ \sqcup A_-$, and let the metrics $d_+$ and $d_-$ on $X \sqcup X'$ be given by

$$d_\pm(n, m') = \inf_{k \in A_\pm} [d_X(n, k) + 1 + d_X(k, m)],$$

$e = [d_+]$, $f = [d_-]$. Then $es = \mathbf{0}$, while $se = f$, i.e., $e$ and $s$ do not commute. $\square$

Note that, unlike $M(X)$, the set $E(M(X))$ of idempotents of $M(X)$ is a coarse invariant. This follows from the geometric description of idempotents in [Manuilov 2021b].

# Part II. When $M(X)$ is commutative

## 7. $R$-spaces

**Definition 7.1.** A metric space $X$ is an *$R$-space* ($R$ for rigid) if, for any $C > 0$ and any two sequences $\{x_n\}_{n\in\mathbb{N}}$, $\{y_n\}_{n\in\mathbb{N}}$ of points in $X$ satisfying

(7-1) $$|d_X(x_n, x_m) - d_X(y_n, y_m)| < C \quad \text{for any } n, m \in \mathbb{N}$$

there exists $D > 0$ such that $d_X(x_n, y_n) < D$ for any $n \in \mathbb{N}$.

**Example 7.2.** As $M(X)$ is commutative for any $X \subset [0, \infty)$, it would follow from Theorem 8.2 below that such $X$ is an $R$-space. A less trivial example is a planar spiral $X$ given by $r = e^\varphi$ in polar coordinates with the metric induced from the standard metric on the plane. Indeed, take any two sequences $\{x_n\}_{n\in\mathbb{N}}$, $\{y_n\}_{n\in\mathbb{N}}$, in $X$. Without loss of generality we may assume that $x_1 = y_1 = 0$ is the origin. If these sequences satisfy (7-1) then

$$|d_X(0, x_n) - d_X(0, y_n)| < C$$

for some fixed $C > 0$ (we take $m = 1$). If $x_n = (r_n, \varphi_n)$, $y_n = (s_n, \psi_n)$ then $d_X(0, x_n) = r_n$, $d_X(0, y_n) = s_n$, and we have $|r_n - s_n| < C$. Then $x_n$ and $y_n$ lie in a ring of width $C$, say $R \le r \le R + C$. If $R$ is sufficiently great then

$$d_X(x_n, y_n) \le (\log(R + C) - \log R)(R + C),$$

which is bounded as a function of $R$.

Consider the set $AI(X)$ of all equivalence classes of almost isometries of $X$. It is easy to see that it is a group with respect to the composition. A metric space $X$ is called AI-rigid [Kar et al. 2016] if the group $AI(X)$ is trivial.

**Proposition 7.3.** *A countable $R$-space $X$ is AI-rigid.*

*Proof.* Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence of all points of $X$, and let $f : X \to X$ be an almost isometry. Set $y_n = f(x_n)$. Then there exists $C > 0$ such that

$$|d_X(f(x_n), f(x_m)) - d_X(x_n, x_m)| < C$$

for any $n, m \in \mathbb{N}$, hence there exists $D > 0$ such that

$$d_X(x_n, f(x_n)) = d_X(x_n, y_n) < D$$

for any $n \in \mathbb{N}$, i.e., $f$ is equivalent to the identity map, hence $X$ is AI-rigid. $\square$

**Example 7.4.** Euclidean spaces $\mathbb{R}^n$, $n \geq 1$, are not $R$-spaces, as they have a nontrivial symmetry. The Archimedean spiral $r = \varphi$ is not an $R$-space, as it is $\pi$-dense in $\mathbb{R}^2$.

## 8. Criterion of commutativity

Let $a, b : T \to [0, \infty)$ be two functions on a set $T$. We say that $a \preceq b$ if there exists a monotone increasing function $\varphi : [0, \infty) \to [0, \infty)$ with $\lim_{s\to\infty} \varphi(s) = \infty$ (we call such functions *reparametrizations*) such that $a(t) \leq \varphi(b(t))$ for any $t \in T$.

The following lemma should be known, but we could not find a reference.

**Lemma 8.1.** *Let $a, b : T \to [0, \infty)$ be two functions. If $a \preceq b$ is not true then there exists $C > 0$ and a sequence $(t_n)_{n\in\mathbb{N}}$ of points in $T$ such that $b(t_n) < C$ for any $n \in \mathbb{N}$ and $\lim_{n\to\infty} a(t_n) = \infty$.*

*Proof.* If $a \preceq b$ is not true then for any reparametrization $\varphi$ there exists $t \in T$ such that $a(t) > \varphi(b(t))$. Suppose that for any $C > 0$, the value $\max\{a(t) : b(t) \leq C\}$ is finite. Then set

$$f(C) = \max(\max\{a(t) : b(t) \leq C\}, C).$$

This gives a reparametrization $f$. If $b(t) = C$ then $a(t) \leq f(C) = f(b(t))$ — a contradiction. Thus, there exists $C > 0$ such that $\max\{a(t) : b(t) \leq C\} = \infty$. It remains to choose a sequence $(t_n)_{n\in\mathbb{N}}$ in the set $\{t \in T : b(t) \leq C\}$ such that $a(t_n) > n$. $\square$

**Theorem 8.2.** *$X$ is an $R$-space if and only if $M(X)$ is commutative.*

*Proof.* Let $X$ be an $R$-space. We shall show that any $s \in M(X)$ is a projection. It would follow that $M(X)$ is commutative. First, we shall show that any $s \in M(X)$ is selfadjoint. Let $d \in \mathcal{M}(X)$, $[d] = s$. Suppose that $[d^*] \neq [d]$. This means that either $d^* \preceq d$ or $d \preceq d^*$ is not true, where $d$ and $d^*$ are considered as functions on $T = X \times X'$. Without loss of generality we may assume that $d^* \preceq d$ is not true. Then there exist sequences $(x_n)_{n\in\mathbb{N}}$ in $X$ and $(y')_{n\in\mathbb{N}}$ in $X'$ and $L > 0$ such that $d(x_n, y'_n) < L$ for any $n \in \mathbb{N}$ and $\lim_{n\to\infty} d(y_n, x'_n) = \infty$ (recall that $d^*(x, y') := d(y, x')$).

Take $n, m \in \mathbb{N}$. Since $d(x_n, y'_n) < L$, $d(x_m, y'_m) < L$, we have

$$|d_X(x_n, x_m) - d_X(y_n, y_m)| = |d_X(x_n, x_m) - d_X(y'_n, y'_m)| < 2L,$$

and, since $X$ is an $R$-space, there exists $D > 0$ such that $d_X(x_n, y_n) < D$ for any $n \in \mathbb{N}$.

Then, using the triangle inequality for the quadrangle $x_n y_n x'_n y'_n$, we get

$$\begin{aligned} d(y_n, x'_n) &\le d_X(y_n, x_n) + d(x_n, y'_n) + d_X(y'_n, x'_n) \\ &= d_X(y_n, x_n) + d(x_n, y'_n) + d_X(y_n, x_n) < D + L + D, \end{aligned}$$

which contradicts the condition $\lim_{n\to\infty} d(y_n, x'_n) = \infty$.

Now, let us show that $[d] \in M(X)$ is idempotent if $X$ is an $R$-space. Let $a(x) = d(x, X')$, $b(x) = d(x, x')$. It was shown in [Manuilov 2021a, Theorem 3.1 and remark at the end of Section 11] that if $[d]$ is selfadjoint then it is idempotent if and only if $b \preceq a$. Suppose that the latter is not true. Then there exists $L > 0$ and a sequence $\{x_n\}_{n\in\mathbb{N}}$ of points in $X$ such that $d(x_n, X') < L$ for any $n \in \mathbb{N}$ and $\lim_{n\to\infty} d(x_n, x'_n) = \infty$. In particular, this means that there exists a sequence $\{y_n\}_{n\in\mathbb{N}}$ of points in $X$ such that $d(x_n, y'_n) < L$ for any $n \in \mathbb{N}$. Since $[d]$ is selfadjoint, for any $L > 0$ there exists $R > 0$ such that if $d(x, y') < L$ then $d(x', y) < R$.

It follows from the triangle inequality for the quadrangle $x_n x_m y'_n y'_m$ that

$$|d_X(x_n, x_m) - d_X(y_n, y_m)| = |d_X(x_n, x_m) - d_X(y'_n, y'_m)| \le d(x_n, y'_n) + d(x_m, y'_m) < 2L$$

for any $n, m \in \mathbb{N}$, hence, the property of being an $R$-space implies that there exists $D > 0$ such that $d_X(x_n, y_n) < D$ for any $n \in \mathbb{N}$. Therefore,

$$d(x_n, x'_n) \le d_X(x_n, y_n) + d(y_n, x'_n) < D + R$$

for any $n \in \mathbb{N}$ — a contradiction with $\lim_{n\to\infty} d(x_n, x'_n) = \infty$.

In the opposite direction, suppose that $X$ is not an $R$-space. i.e., that there exists $C > 0$ and sequences $\{x_n\}_{n\in\mathbb{N}}$, $\{y_n\}_{n\in\mathbb{N}}$ of points in $X$ such that (7-1) holds and $\lim_{n\to\infty} d_X(x_n, y_n) = \infty$.

Note that these sequences cannot be bounded. Indeed, if there exists $R > 0$ such that $d_X(x_1, x_n) < R$ for any $n \in \mathbb{N}$ then

$$d_X(y_1, y_n) \le d_X(x_1, x_n) + C = R + C$$

for any $n \in \mathbb{N}$, but then

$$d_X(x_n, y_n) \le d_X(x_n, x_1) + d_X(x_1, y_1) + d_X(y_1, y_n) < R + d_X(x_1, y_1) + R + C,$$

which contradicts $\lim_{n\to\infty} d_X(x_n, y_n) = \infty$. Passing to a subsequence, we may assume that

$$d_X(x_k, x_n) > k, \quad d_X(x_k, y_n) > k, \quad d_X(y_k, x_n) > k, \quad d_X(y_k, y_n) > k$$

for any $n < k$, and $d_X(x_k, y_k) > k$ for any $k \in \mathbb{N}$. In particular, this means that

$$d_X(x_k, y_n) > k \quad \text{for any } k, n \in \mathbb{N}. \tag{8-1}$$

Let us define two metrics on the double of $X$ and show that they don't commute. For $x, y \in X$ set

$$d_1(x, y') = \min_{n\in\mathbb{N}} [d_X(x, x_n) + C + d_X(y_n, y)];$$

$$d_2(x, y') = \min_{n\in\mathbb{N}} [d_X(x, y_n) + C + d_X(x_n, y)]$$

(it is clear that the minimum is attained on some $n \in \mathbb{N}$ as $x_n, y_n \to \infty$). Let us show that $d_1$ is a metric on $X \sqcup X'$ (the case of $d_2$ is similar).

Due to symmetry, it suffices to check the two triangle inequalities for the triangle $xzy'$, $z \in X$:

$$\begin{aligned}
&d_1(x, y') + d_1(z, y') \\
&\qquad = \min_{n\in\mathbb{N}}[d_X(x, x_n) + C + d_X(y_n, y)] + \min_{m\in\mathbb{N}}[d_X(z, x_m) + C + d_X(y_m, y)] \\
&\qquad = d_X(x, x_{n_x}) + d_X(y_{n_x}, y) + d_X(y, y_{n_z}) + d_X(z, x_{n_z}) + 2C \\
&\qquad \geq d_X(x, x_{n_x}) + d_X(y_{n_x}, y_{n_z}) + d_X(z, x_{n_z}) + 2C \\
&\qquad \geq d_X(x, x_{n_x}) + (d_X(x_{n_x}, x_{n_z}) - C) + d_X(z, x_{n_z}) + 2C \\
&\qquad = d_X(x, x_{n_x}) + d_X(x_{n_x}, x_{n_z} + d_X(z, x_{n_z}) + C \\
&\qquad \geq d_X(x, z) + C \geq d_X(x, z).
\end{aligned}$$

and

$$\begin{aligned}
d_1(x, y') &= \min_{n\in\mathbb{N}}[d_X(x, x_n) + C + d_X(y_n, y)] \\
&\leq d_X(x, x_{n_z}) + d_X(y_{n_z}, y) + C \\
&\leq d_X(x, z) + d_X(z, x_{n_z}) + d_X(y_{n_z}, y) + C = d_X(x, z) + d_1(z, y').
\end{aligned}$$

Let us evaluate $(d_2 \circ d_1)(x_k, x'_k)$ and $(d_1 \circ d_2)(x_k, x'_k)$.

Taking fixed values $n = m = k$, $u = y_k$, we get

$$\begin{aligned}
&(d_2 \circ d_1)(x_k, x'_k) \\
&\quad = \inf_{u\in X}\{\min_{n\in\mathbb{N}}[d_X(x_k, x_n) + C + d_X(y_n, u)] + \min_{m\in\mathbb{N}}[d_X(u, y_m) + C + d_X(x_m, x_k)]\} \\
&\quad \leq \inf_{u\in X}\{[d_X(x_k, x_k) + C + d_X(y_k, u)] + [d_X(u, y_k) + C + d_X(x_k, x_k)]\} \\
&\quad = [d_X(x_k, x_k) + C] + [C + d_X(x_k, x_k)] \\
&\quad = 2C.
\end{aligned}$$

Using the triangle inequality for the triangle $x_n x_m u$ and (8-1), we get

$$\begin{aligned}
&(d_1 \circ d_2)(x_k, x'_k) \\
&\quad = \inf_{u\in X}\{\min_{n\in\mathbb{N}}[d_X(x_k, y_n) + C + d_X(x_n, u)] + \min_{m\in\mathbb{N}}[d_X(u, x_m) + C + d_X(y_m, x_k)]\} \\
&\quad \geq \inf_{u\in X}\{\min_{n\in\mathbb{N}}[d_X(x_k, y_n) + d_X(x_n, u)] + \min_{m\in\mathbb{N}}[d_X(u, x_m) + d_X(y_m, x_k)]\} \\
&\quad \geq \min_{n,m\in\mathbb{N}}[d_X(x_k, y_n) + d_X(x_n, x_m) + d_X(y_m, x_k)] > k + d_X(x_n, x_m) + k > 2k.
\end{aligned}$$

Thus, for the sequence $\{x_k\}_{k\in\mathbb{N}}$ of points in $X$, the distances $(d_2 \circ d_1)(x_k, x'_k)$ are uniformly bounded, while $\lim_{k\to\infty}(d_1 \circ d_2)(x_k, x'_k) = \infty$, hence the metrics $d_2 \circ d_1$ and $d_1 \circ d_2$ are not equivalent, i.e., $[d_2][d_1] \neq [d_1][d_2]$. $\square$

Recall that an inverse semigroup $S$ is Clifford (see [Howie 1995], Theorem 4.2.1) if $s^*s = ss^*$ for any $s \in S$. If $S$ is commutative then it is patently Clifford, but not the other way. Nevertheless, for inverse semigroups of the form $M(X)$ these two properties are the same.

**Corollary 8.3.** *If $M(X)$ is Clifford then $X$ is an $R$-space* (*and $M(X)$ is commutative*).

*Proof.* Let $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$ are sequences in $X$ satisfying (7-1), and let $d_1$, $d_2$ are metrics on the double of $X$ defined above. Note that $d_1^* = d_2$, and let $s = [d_1]$. We have $s^*s \neq ss^*$, which contradicts that $M(X)$ is Clifford. □

## References

[Burago et al. 2001] D. Burago, Y. Burago, and S. Ivanov, *A course in metric geometry*, Grad. Stud. Math. **33**, Amer. Math. Soc., Providence, RI, 2001. MR Zbl

[Howie 1995] J. M. Howie, *Fundamentals of semigroup theory*, London Mathematical Society Monographs. New Series **12**, The Clarendon Press, Oxford University Press, New York, 1995. MR Zbl

[Kar et al. 2016] A. Kar, J.-F. Lafont, and B. Schmidt, "Rigidity of almost-isometric universal covers", *Indiana Univ. Math. J.* **65**:2 (2016), 585–613. MR Zbl

[Lawson 1998] M. V. Lawson, *Inverse semigroups: The theory of partial symmetries*, World Scientific Publishing Co., River Edge, NJ, 1998. MR Zbl

[Manuilov 2019] V. M. Manuilov, "Roe bimodules as morphisms of discrete metric spaces", *Russ. J. Math. Phys.* **26**:4 (2019), 470–478. MR Zbl

[Manuilov 2021a] V. Manuilov, "Metrics on doubles as an inverse semigroup", *J. Geom. Anal.* **31**:6 (2021), 5721–5739. MR Zbl

[Manuilov 2021b] V. Manuilov, "Metrics on doubles as an inverse semigroup II", *J. Math. Anal. Appl.* **496**:2 (2021), Paper No. 124821, 16. MR Zbl

[Paterson 1999] A. L. T. Paterson, *Groupoids, inverse semigroups, and their operator algebras*, Progress in Mathematics **170**, Birkhäuser, Boston, 1999. MR Zbl

Received January 15, 2024.

VLADIMIR MANUILOV
MOSCOW CENTER FOR FUNDAMENTAL AND APPLIED MATHEMATICS
MOSCOW STATE UNIVERSITY
MOSCOW
RUSSIA
manuilov@mech.math.msu.su

PACIFIC JOURNAL OF MATHEMATICS
Vol. 328, No. 2, 2024
https://doi.org/10.2140/pjm.2024.328.339

# THE NUMBER OF $\mathbb{F}_q$-POINTS ON DIAGONAL HYPERSURFACES WITH MONOMIAL DEFORMATION

DERMOT MCCARTHY

**We consider the family of diagonal hypersurfaces with monomial deformation**

$$D_{d,\lambda,h} : x_1^d + x_2^d + \cdots + x_n^d - d\lambda\, x_1^{h_1} x_2^{h_2} \ldots x_n^{h_n} = 0,$$

**where $d = h_1 + h_2 + \cdots + h_n$ with $\gcd(h_1, h_2, \ldots, h_n) = 1$. We first provide a formula for the number of $\mathbb{F}_q$-points on $D_{d,\lambda,h}$ in terms of Gauss and Jacobi sums. This generalizes a result of Koblitz, which holds in the special case $d \mid q-1$. We then express the number of $\mathbb{F}_q$-points on $D_{d,\lambda,h}$ in terms of a $p$-adic hypergeometric function previously defined by the author. The parameters in this hypergeometric function mirror exactly those described by Koblitz when drawing an analogy between his result and classical hypergeometric functions. This generalizes a result by Sulakashna and Barman, which holds in the case $\gcd(d, q-1) = 1$. In the special case $h_1 = h_2 = \cdots = h_n = 1$ and $d = n$, i.e., the Dwork hypersurface, we also generalize a previous result of the author which holds when $q$ is prime.**

## 1. Introduction

Counting the number of solutions to equations over finite fields using character sums dates back to the works of Gauss and Jacobi. A renewed interest in such problems followed subsequent important contributions from Hardy and Littlewood [1922] and Davenport and Hasse [1935]. In a seminal paper, Weil [1949] gives an exposition on the topic up to that point (as well as going on to make his famous conjectures on the zeta functions of algebraic varieties). Specifically, he develops a formula for the number of solutions over $\mathbb{F}_q$, the finite field with $q$ elements, and its extensions, of $a_0x_0^{n_0} + a_1x_1^{n_1} + \cdots + a_kx_k^{n_k} = 0$, in terms of what we now call Gauss sums and Jacobi sums. The techniques involved have since become standard practice and can be found in many well-known text books, e.g., [Berndt et al. 1998; Ireland and Rosen 1990]. Since then, many authors have used and adapted the techniques outlined in Weil's paper to study other equations, e.g.,

*MSC2020:* primary 11G25, 33E50; secondary 11S80, 11T24, 33C99.
*Keywords:* diagonal hypersurface, Gauss sum, Jacobi sum, finite field hypergeometric function, $p$-adic hypergeometric function, counting points.

[Delsarte 1951; Furtado Gomide 1949; Koblitz 1983]. Of particular interest is the work of Koblitz [1983] where he examines the family of diagonal hypersurfaces with monomial deformation

$$D_{d,\lambda,h} : x_1^d + x_2^d + \cdots + x_n^d - d\lambda\, x_1^{h_1} x_2^{h_2} \ldots x_n^{h_n} = 0, \tag{1-1}$$

where $h_i \in \mathbb{Z}^+$, with $\gcd(h_1, h_2, \ldots, h_n) = 1$, and $d = h_1 + h_2 + \cdots + h_n$. Koblitz's main result [1983, Theorem 2] gives a formula for the number of $\mathbb{F}_q$-points on $D_{d,\lambda,h}$ in the terms of Gauss and Jacobi sums, in the case $d \mid q-1$. Using the analogy between Gauss sums and the gamma function, he notes that the main term in his formula can be considered a finite field analogue of a classical hypergeometric function. The purpose of this paper is to study $D_{d,\lambda,h}$ more generally, i.e., when the condition $d \mid q-1$ is removed. Firstly, we generalize Koblitz's result and provide a formula for the number of $\mathbb{F}_q$-points on $D_{d,\lambda,h}$ in terms of Gauss and Jacobi sums without the condition $d \mid q-1$. We then express the number of $\mathbb{F}_q$-points on $D_{d,\lambda,h}$ in terms of a $p$-adic hypergeometric function previously defined by the author. The parameters in this hypergeometric function mirror exactly those described by Koblitz when drawing an analogy between his result and classical hypergeometric functions. This generalizes a result of [Sulakashna and Barman 2022], which holds in the case $\gcd(d, q-1) = 1$. We also examine the special case when $h_1 = h_2 = \cdots = h_n = 1$ and $d = n$, i.e., the Dwork hypersurface, and generalize a previous result of the author, which holds when $q$ is prime.

## 2. Statement of results

Let $q = p^r$ be a prime power and let $\mathbb{F}_q$ denote the finite field with $q$ elements. Let $\widehat{\mathbb{F}_q^*}$ denote the group of multiplicative characters of $\mathbb{F}_q^*$. We extend the domain of $\chi \in \widehat{\mathbb{F}_q^*}$ to $\mathbb{F}_q$ by defining $\chi(0) := 0$ (including for the trivial character $\varepsilon$) and denote $\bar{\chi}$ as the inverse of $\chi$. Let $T$ be a fixed generator of $\widehat{\mathbb{F}_q^*}$. Let $\theta$ be a fixed nontrivial additive character of $\mathbb{F}_q$ and for $\chi \in \widehat{\mathbb{F}_q^*}$ we define the Gauss sum $g(\chi) := \sum_{x \in \mathbb{F}_q} \chi(x)\theta(x)$. For $\chi_1, \chi_2, \ldots, \chi_k \in \widehat{\mathbb{F}_q^*}$, we define the Jacobi sum $J(\chi_1, \chi_2, \ldots, \chi_k) := \sum_{t_i \in \mathbb{F}_q, t_1+t_2+\cdots+t_k=1} \chi_1(t_1)\chi_2(t_2)\cdots\chi_k(t_k)$.

We consider the family of diagonal hypersurfaces with monomial deformation described in (1-1). Let $t := \gcd(d, q-1)$ throughout and define

$$W := \left\{ w = (w_1, w_2, \ldots, w_n) \in \mathbb{Z}^n : 0 \le w_i < t,\ \sum_{i=1}^{n} w_i \equiv 0 \pmod t \right\}. \tag{2-1}$$

Define an equivalence relation $\sim_h$ on $W$ by

$$w \sim_h w' \text{ if } w - w' \text{ is a multiple modulo } t \text{ of } h = (h_1, h_2, \ldots, h_n). \tag{2-2}$$

We denote the class containing $w$ by $[w]$. If $h = (1, 1, \ldots, 1)$ we write $\sim_1$. We note, in this case, that each class contains a representative $w$ where some $w_i = 0$, for $1 \le i \le n$. We will write $[w_0]$ to indicate that we have chosen such a representative for a particular class.

Our first result provides a formula for the number of $\mathbb{F}_q$-points on $D_{d,\lambda,h}$ in terms of Gauss and Jacobi sums, without the condition $d \mid q-1$. We will use $\mathbb{A}^n(\mathbb{F}_q)$ and $\mathbb{P}^n(\mathbb{F}_q)$ to denote the affine and projective $n$-spaces, respectively, over $\mathbb{F}_q$. We denote the subset of elements in these spaces where all coordinates are nonzero by $\mathbb{A}^n(\mathbb{F}_q^*)$ and $\mathbb{P}^n(\mathbb{F}_q^*)$.

**Theorem 2.1.** *Let $N_q(D_{d,\lambda,h})$ be the number of points in $\mathbb{P}^{n-1}(\mathbb{F}_q)$ on $D_{d,\lambda,h}$. Then*

$$N_q(D_{d,\lambda,h}) = \frac{q^{n-1}-1}{q-1} - \sum_{w^*} J(T^{w_1\frac{q-1}{t}}, T^{w_2\frac{q-1}{t}}, \dots, T^{w_n\frac{q-1}{t}}) + \frac{1}{q-1}\sum_{s,w} \frac{g(T^{w_1\frac{q-1}{t}+h_1 s})g(T^{w_2\frac{q-1}{t}+h_2 s})\dots g(T^{w_n\frac{q-1}{t}+h_n s})}{g(T^{ds})} T^{ds}(d\lambda),$$

*where the first sum is over all $w^* = (w_1, w_2, \dots, w_n) \in W$ such that $0 < w_i < t$ for all $i$, and the second sum is over all $s \in \left\{0, 1, \dots, \frac{q-1}{t} - 1\right\}$ and all $w = (w_1, w_2, \dots, w_n) \in W$.*

Theorem 2.1 generalizes [Koblitz 1983, Theorem 2], which holds in the case $d \mid q-1$. Using an analogy between Gauss sums and the gamma function, Koblitz noted that the second summand in his formula, which corresponds to the second summand in Theorem 2.1 above with $t = d$, can be considered a finite field analogy of the classical hypergeometric function

$$\prod_{i=1}^{n} \Gamma\left(\tfrac{w_i}{d}\right) \cdot {}_dF_{d-1}\left[\begin{matrix} \cdots\cdots \frac{w_i}{dh_i} + \frac{b_i}{h_i} \cdots\cdots \\ \frac{1}{d} \quad \frac{2}{d} \quad \cdots \qquad \frac{d-1}{d} \end{matrix} \;\middle|\; \lambda^d h_1^{h_1} \dots h_n^{h_n}\right], \tag{2-3}$$

where the top line parameters range over all $i = 1, \dots, n$ and, for each $i$, all $b_i = 0, \dots, h_i - 1$. The main purpose of this paper is to express $N_q(D_{d,\lambda,h})$ in terms of a $p$-adic hypergeometric function previously defined by the author, whereby the parameters in this $p$-adic hypergeometric function mirror exactly those described by Koblitz in (2-3) above.

Next, we rewrite Theorem 2.1 in a way more amenable to manipulation when we pass to the $p$-adic setting.

**Corollary 2.2.**

$$N_q(D_{d,\lambda,h}) = \frac{q^{n-1}-1}{q-1} - \frac{1}{q} \sum_{\substack{w \in W \\ \text{some } w_i = 0}} \prod_{i=1}^{n} g(T^{w_i\frac{q-1}{t}}) + \frac{1}{q(q-1)} \sum_{s,w} \prod_{i=1}^{n} g(T^{w_i\frac{q-1}{t}+h_i s}) g(T^{-ds})\, T^{ds}(-d\lambda)$$

*where the first sum is over all $w = (w_1, w_2, \ldots, w_n) \in W$ such that at least one $w_i = 0$, and the second sum is over* ***either*** *all $s \in \left\{0, 1, \ldots, \frac{q-1}{t} - 1\right\}$ and all $w = (w_1, w_2, \ldots, w_n) \in W$* ***or*** *all $s \in \{0, 1, \ldots, q-2\}$ and all $w = (w_1, w_2, \ldots, w_n) \in W/\sim_h$. In the latter case, the sum is independent of the choice of equivalence class representatives.*

We now define our $p$-adic hypergeometric function. Let $\mathbb{Z}_p$ denote the ring of $p$-adic integers, $\mathbb{Q}_p$ the field of $p$-adic numbers, $\overline{\mathbb{Q}_p}$ the algebraic closure of $\mathbb{Q}_p$, and $\mathbb{C}_p$ the completion of $\overline{\mathbb{Q}_p}$. Let $\mathbb{Z}_q$ be the ring of integers in the unique unramified extension of $\mathbb{Q}_p$ with residue field $\mathbb{F}_q$. Recall that for each $x \in \mathbb{F}_q^*$, there is a unique Teichmüller representative $\omega(x) \in \mathbb{Z}_q^\times$ such that $\omega(x)$ is a $(q-1)$-st root of unity and $\omega(x) \equiv x \pmod p$. Therefore, we define the Teichmüller character to be the primitive character $\omega : \mathbb{F}_q^* \to \mathbb{Z}_q^\times$ given by $x \mapsto \omega(x)$, which we extend with $\omega(0) := 0$.

**Definition 2.3** [McCarthy 2013, Definition 5.1]. Let $q = p^r$ for $p$ an odd prime. Let $\lambda \in \mathbb{F}_q$, $m \in \mathbb{Z}^+$ and $a_i, b_i \in \mathbb{Q} \cap \mathbb{Z}_p$ for $1 \le i \le m$. Then define

$$
\begin{aligned}
&{}_mG_m\left[\begin{matrix} a_1, & a_2, & \ldots, & a_m \\ b_1, & b_2, & \ldots, & b_m \end{matrix} \;\middle|\; \lambda \right]_q \\
&:= \frac{-1}{q-1} \sum_{s=0}^{q-2} (-1)^{sm}\, \overline{\omega}^s(\lambda) \\
&\quad \times \prod_{i=1}^{m} \prod_{k=0}^{r-1} \frac{\Gamma_p\big(\big\langle \big(a_i - \frac{s}{q-1}\big) p^k \big\rangle\big)}{\Gamma_p(\langle a_i p^k \rangle)} \frac{\Gamma_p\big(\big\langle \big(-b_i + \frac{s}{q-1}\big) p^k \big\rangle\big)}{\Gamma_p(\langle -b_i p^k \rangle)} (-p)^{-\left\lfloor \langle a_i p^k \rangle - \frac{sp^k}{q-1} \right\rfloor - \left\lfloor \langle -b_i p^k \rangle + \frac{sp^k}{q-1} \right\rfloor}.
\end{aligned}
$$

We note that the value of ${}_mG_m[\cdots]$ depends only on the fractional part of the $a_i$ and $b_i$ parameters, and is invariant if we change the order of the parameters. Our main result expresses $N_q(D_{d,\lambda,h})$ in terms of this function.

**Theorem 2.4.** *Let $q = p^r$ for $p$ an odd prime. Then, for $p \nmid d h_1 \cdots h_n$,*

$$
\begin{aligned}
N_q(D_{d,\lambda,h}) = \frac{q^{n-1}-1}{q-1} - \frac{(-1)^n}{q} \sum_{\substack{w \in W \\ \text{some } w_i = 0}} C(w) \\
+ \frac{(-1)^n}{q} \sum_{[w] \in W/\sim_h} C(w)\, {}_dG_d\left[\begin{matrix} \cdots\cdots \frac{w_i}{th_i} + \frac{b_i}{h_i} \cdots\cdots \\ 1 \;\; \frac{1}{d} \;\; \frac{2}{d} \;\; \cdots \;\; \frac{d-1}{d} \end{matrix} \;\middle|\; (\lambda^d h_1^{h_1} \cdots h_n^{h_n})^{-1} \right]_q
\end{aligned}
$$

*where the top line parameters in ${}_dG_d$ are the list*

$$\left[ \frac{w_i}{th_i} + \frac{b_i}{h_i} \;\middle|\; i = 1, \ldots, n;\, b_i = 0, 1, \ldots, h_i - 1 \right],$$

*and*

$$C(w) := \prod_{i=1}^{n}\prod_{a=0}^{r-1} \Gamma_p\big(\big\langle\big(\tfrac{w_i}{t}\big)p^a\big\rangle\big)(-p)^{\left\langle\left(\frac{w_i}{t}\right)p^a\right\rangle}. \tag{2-4}$$

As we can see, the parameters of ${}_dG_d$ in Theorem 2.4 mirror exactly those in (2-3) (when $d \mid q-1$ and so $t = d$) up to inversion of the argument $\lambda^d h_1^{h_1}\cdots h_n^{h_n}$. This inversion is a feature of the definition of the function ${}_mG_m$. Because we are summing over $W/\sim_h$, we can remove this inversion while also swapping the top and bottom line parameters, which gives a more natural representation, in the opinion of the author. This can be seen more clearly later, in Corollary 2.9, where we get an all integral bottom line parameters.

**Corollary 2.5.** *Let $q = p^r$ for $p$ an odd prime. Then, for $p \nmid dh_1\cdots h_n$,*

$$N_q(D_{d,\lambda,h}) = \frac{q^{n-1}-1}{q-1} - \frac{(-1)^n}{q}\sum_{\substack{w\in W\\ \text{some } w_i=0}} C(w)$$
$$+\frac{(-1)^n}{q}\sum_{[w]\in W/\sim_h} C(-w)\,{}_dG_d\left[\begin{matrix} 1 & \frac{1}{d} & \frac{2}{d} & \cdots & \frac{d-1}{d} \\ & \cdots\cdots & \frac{w_i}{th_i}+\frac{b_i}{h_i} & \cdots\cdots & \end{matrix}\;\middle|\;\lambda^d h_1^{h_1}\cdots h_n^{h_n}\right]_q.$$

Ideally, in Theorem 2.4 and Corollary 2.5, we would like to combine both sums into a single hypergeometric term. In general, it seems that this is not possible. However, it can be achieved in two special cases as we see in the next two results. The first is when $\gcd(d, q-1) = 1$ and the second is when all $h_i = 1$, i.e., the Dwork hypersurface.

**Corollary 2.6.** *Let $q = p^r$ for $p$ an odd prime. If $\gcd(d, q-1) = 1$ then, for $p \nmid dh_1\cdots h_n$,*

$$N_q(D_{d,\lambda,h}) = \frac{q^{n-1}-1}{q-1} + (-1)^n{}_{d-1}G_{d-1}\left[\begin{matrix} \frac{1}{d} & \frac{2}{d} & \cdots & \frac{d-1}{d} \\ & \cdots & \frac{b_i}{h_i} & \cdots \end{matrix}\;\middle|\;\lambda^d h_1^{h_1}\cdots h_n^{h_n}\right]_q$$

*where the bottom line parameters in ${}_{d-1}G_{d-1}$ are the list*

$$\left[\frac{b_i}{h_i}\;\middle|\; i = 1,\ldots,n;\, b_i = 0,1,\ldots,h_i-1\right]$$

*with exactly one zero removed.*

Corollary 2.6 is Theorem 1.2 of [Sulakashna and Barman 2022].

When $h_1 = h_2 = \cdots = h_n = 1$ and $d = n$ in (1-1), we recover the Dwork hypersurface, which we will denote $D_\lambda$, i.e.,

$$D_\lambda : x_1^n + x_2^n + \cdots + x_n^n - n\lambda\, x_1x_2\ldots x_n = 0.$$

We now provide formulas for the number of $\mathbb{F}_q$-points on $D_\lambda$, first in terms of Gauss and Jacobi sums, and then in terms of the $p$-adic hypergeometric function. For a given $w = (w_1, w_2, \ldots, w_n) \in W$, define $n_k$ to be the number of $k$'s appearing in $w$, i.e., $n_k = |\{w_i \mid 1 \le i \le n,\, w_i = k\}|$. We then let $S_w := \{k \mid 0 \le k \le t-1,\, n_k = 0\}$ and $S_w^c$ denote its complement in $\{0, 1, \ldots, t-1\}$. So the elements of $S_w$ are the numbers from 0 to $t-1$, inclusive, which do not appear in $w$. We define the following lists:

$$A_w : \left[\frac{t-k}{t} \;\middle|\; k \in S_w\right] \cup \left[\frac{b}{n} \;\middle|\; 0 \le b \le n-1,\, b \not\equiv 0 \left(\bmod \frac{n}{t}\right)\right]; \tag{2-5}$$

$$B_w : \left[\frac{t-k}{t} \text{ repeated } n_k - 1 \text{ times} \;\middle|\; k \in S_w^c\right]. \tag{2-6}$$

We note both lists contain $n - |S_w^c|$ numbers.

**Corollary 2.7** (corollary to Theorem 2.1). *Let $N_q(D_\lambda)$ be the number of points in $\mathbb{P}^{n-1}(\mathbb{F}_q)$ on $D_\lambda$. Let $t = \gcd(n, q-1)$. Then, for $\lambda \neq 0$,*

$$N_q(D_\lambda) = \frac{q^{n-1}-1}{q-1} + \frac{1}{q(q-1)} \sum_{s,w} \left[\prod_{k \in S_w^c} \frac{g(T^{k\frac{q-1}{t}+s})^{n_k-1}}{g(T^{-k\frac{q-1}{t}-s})} T^{k\frac{q-1}{t}+s}(-1)\, q\right] g(T^{-ns})\, T^{ns}(-n\lambda)$$

*where the sum is over **either** all $s \in \left\{0, 1, \ldots, \frac{q-1}{t} - 1\right\}$ and all $w = (w_1, w_2, \ldots, w_n) \in W$ **or** all $s \in \{0, 1, \ldots, q-2\}$ and all $w = (w_1, w_2, \ldots, w_n) \in W/\sim_1$. In the latter case, the sum is independent of the choice of equivalence class representatives.*

**Theorem 2.8.** *Let $q = p^r$ for $p$ an odd prime. Let $N_q(D_\lambda)$ be the number of points in $\mathbb{P}^{n-1}(\mathbb{F}_q)$ on $D_\lambda$ for some $\lambda \in \mathbb{F}_q^*$. Let $t = \gcd(n, q-1)$ and let $C(w)$ be defined by* (2-4). *Then, for $p \nmid n$,*

$$N_q(D_\lambda) = \frac{q^{n-1}-1}{q-1} + (-1)^n \sum_{[w_0] \in W/\sim_1} C(w_0)\, {}_l G_l \left[\begin{matrix} A_{w_0} \\ B_{w_0} \end{matrix} \;\middle|\; \lambda^n \right]_q .$$

Theorem 2.8 generalizes Theorem 2.2 in [McCarthy 2017] which holds for $q = p$. Finally, if we let $\gcd(n, q-1) = 1$ in Theorem 2.8, or we let $h_1 = h_2 = \cdots = h_n = 1$ in Corollary 2.6, it easy to see that we arrive at the following result.

**Corollary 2.9.** *If $\gcd(n, q-1) = 1$ then, for $p \nmid n$,*

$$N_q(D_\lambda) = \frac{q^{n-1}-1}{q-1} + (-1)^n {}_{n-1}G_{n-1} \left[\begin{matrix} \frac{1}{n} & \frac{2}{n} & \cdots & \frac{n-1}{n} \\ 1 & 1 & \cdots & 1 \end{matrix} \;\middle|\; \lambda^n \right]_q .$$

Corollary 2.9 generalizes Corollary 2.3 in [McCarthy 2017] which holds for $q = p$.

## 3. Preliminaries

We start by recalling some properties of Gauss and Jacobi sums. See [Berndt et al. 1998; Ireland and Rosen 1990] for further details, noting that we have adjusted results to take into account $\varepsilon(0) = 0$, where $\varepsilon$ is the trivial character. We first note that $G(\varepsilon) = -1$. For $\chi \in \widehat{\mathbb{F}_q^*}$,

$$G(\chi)G(\bar{\chi}) = \begin{cases} \chi(-1)q & \text{if } \chi \neq \varepsilon, \\ 1 & \text{if } \chi = \varepsilon. \end{cases} \tag{3-1}$$

For $\chi_1, \chi_2, \ldots, \chi_k \in \widehat{\mathbb{F}_q^*}$ and $\alpha \in \mathbb{F}_q$, we define the generalized Jacobi sum

$$J_\alpha(\chi_1, \chi_2, \ldots, \chi_k) := \sum_{t_i \in \mathbb{F}_q, t_1+t_2+\cdots+t_k=\alpha} \chi_1(t_1)\chi_2(t_2)\cdots\chi_k(t_k).$$

When $\alpha = 1$ we recover the usual Jacobi sum as defined in Section 2.

**Proposition 3.1.** *For* $\chi_1, \chi_2, \ldots, \chi_k \in \widehat{\mathbb{F}_q^*}$,

$$J_0(\chi_1, \chi_2, \ldots, \chi_k) = \begin{cases} (q-1)^k - (q-1)\, J(\chi_1, \chi_2, \ldots, \chi_k) & \text{if } \chi_1, \chi_2, \ldots, \chi_k \text{ all trivial}, \\ -(q-1)J(\chi_1, \chi_2, \ldots, \chi_k) & \text{if } \chi_1\chi_2\cdots\chi_k \text{ trivial but at least} \\ & \text{one of } \chi_1, \chi_2, \ldots, \chi_k \text{ nontrivial}, \\ 0 & \text{if } \chi_1\chi_2\cdots\chi_k \text{ nontrivial}. \end{cases}$$

**Proposition 3.2.** *For* $\chi_1\chi_2\cdots\chi_k$ *trivial but at least one of* $\chi_1, \chi_2, \ldots, \chi_k$ *nontrivial then*

$$J(\chi_1, \chi_2, \ldots, \chi_k) = -\chi_k(-1)J(\chi_1, \chi_2, \ldots, \chi_{k-1}) .$$

**Proposition 3.3.** *For* $\chi_1, \chi_2, \ldots, \chi_k$ *all trivial*,

$$J(\chi_1, \chi_2, \ldots, \chi_k) = [(q-1)^k + (-1)^{k+1}]/q.$$

**Proposition 3.4.** *For* $\chi_1, \chi_2, \ldots, \chi_k$ *not all trivial*,

$$J(\chi_1, \chi_2, \ldots, \chi_k) = \begin{cases} \dfrac{G(\chi_1)G(\chi_2)\cdots G(\chi_k)}{G(\chi_1\chi_2\cdots\chi_k)} & \text{if } \chi_1\chi_2\cdots\chi_k \neq \varepsilon, \\ -\dfrac{G(\chi_1)G(\chi_2)\cdots G(\chi_k)}{q} & \text{if } \chi_1\chi_2\cdots\chi_k = \varepsilon. \end{cases}$$

We now recall the $p$-adic gamma function. For further details, see [Koblitz 1980]. Let $p$ be an odd prime. For $n \in \mathbb{Z}^+$ we define the $p$-adic gamma function as

$$\Gamma_p(n) := (-1)^n \prod_{\substack{0<j<n \\ p \nmid j}} j$$

and extend it to all $x \in \mathbb{Z}_p$ by setting $\Gamma_p(0) := 1$ and $\Gamma_p(x) := \lim_{n \to x} \Gamma_p(n)$ for $x \neq 0$, where $n$ runs through any sequence of positive integers $p$-adically approaching $x$. This limit exists, is independent of how $n$ approaches $x$, and determines a continuous function on $\mathbb{Z}_p$ with values in $\mathbb{Z}_p^*$. The function satisfies the following product formula.

**Theorem 3.5** [Gross and Koblitz 1979, Theorem 3.1]. *If $h \in \mathbb{Z}^+$, $p \nmid h$ and $0 \le x < 1$ with $(q-1)x \in \mathbb{Z}$, then*

$$\prod_{a=0}^{r-1}\prod_{b=0}^{h-1}\Gamma_p\big(\big\langle \tfrac{x+b}{h}p^a\big\rangle\big) = \omega(h^{(q-1)x})\prod_{a=0}^{r-1}\Gamma_p(\langle xp^a\rangle)\prod_{b=1}^{h-1}\Gamma_p\big(\big\langle \tfrac{b}{h}p^a\big\rangle\big). \tag{3-2}$$

We note that in the original statement of Theorem 3.5 in [Gross and Koblitz 1979], $\omega$ is the Teichmüller character of $\mathbb{F}_p^*$. However, the result above still holds as $\omega|_{\mathbb{F}_p^*}$ is the Teichmüller character of $\mathbb{F}_p^*$.

The Gross–Koblitz formula allows us to relate Gauss sums and the $p$-adic gamma function. Let $\pi \in \mathbb{C}_p$ be the fixed root of $x^{p-1}+p=0$ that satisfies $\pi \equiv \zeta_p - 1 \pmod{(\zeta_p-1)^2}$.

**Theorem 3.6** [Gross and Koblitz 1979, Theorem 1.7]. *For $j \in \mathbb{Z}$,*

$$g(\overline{\omega}^j) = -\pi^{(p-1)\sum_{a=0}^{r-1}\left\langle\frac{jp^a}{q-1}\right\rangle}\prod_{a=0}^{r-1}\Gamma_p\big(\big\langle\tfrac{jp^a}{q-1}\big\rangle\big).$$

We now recall some results of [Weil 1949; Koblitz 1983; Furtado Gomide 1949]. Note that the definitions and notation used for characters and for Gauss and Jacobi sums vary among those papers and differ from what's defined in this paper. So, we have adjusted the statement of their results accordingly. For $d \in \mathbb{Z}^+$, let $D_d$ denote the diagonal hypersurface

$$D_d : x_1^d + x_2^d + \cdots + x_n^d = 0.$$

**Theorem 3.7** [Weil 1949]. *Let $N_q^A(D_d)$ be the number of points in $\mathbb{A}^n(\mathbb{F}_q)$ on $D_d$. Let $t := \gcd(d, q-1)$. Then*

$$N_q^A(D_d) = q^{n-1} - (q-1)\sum_{w^*} J(T^{w_1\frac{q-1}{t}}, T^{w_2\frac{q-1}{t}}, \ldots, T^{w_n\frac{q-1}{t}}),$$

*where the sum is over all $w^* = (w_1, w_2, \ldots, w_n) \in W$ such that $0 < w_i < t$.*

Using similar methods to those in [Weil 1949; Koblitz 1983, Theorem 2] it is easy to see that

**Theorem 3.8.** *Let $N_q^{A,*}(D_d)$ be the number of points in $\mathbb{A}^n(\mathbb{F}_q^*)$ on $D_d$ Let $t := \gcd(d, q-1)$. Then*

$$N_q^{A,*}(D_d) = \sum_{w\in W} J_0(T^{w_1\frac{q-1}{t}}, T^{w_2\frac{q-1}{t}}, \ldots, T^{w_n\frac{q-1}{t}}).$$

The next result appears in [Koblitz 1983] in the homogenous case, and in general in [Furtado Gomide 1949]. We note that [Furtado Gomide 1949] contains a minor error. A term is omitted in the determination, but is easily fixed.

**Theorem 3.9** [Furtado Gomide 1949; Koblitz 1983, Theorem 1]. *Let $N_q^{A,*}$ be the number of points in $\mathbb{A}^n(\mathbb{F}_q^*)$ on*

$$\sum_{i=1}^{r} a_i x_1^{m_{1i}} x_2^{m_{2i}} \dots x_n^{m_{ni}} = 0$$

*for some $a_i \in \mathbb{F}_q^*$, $m_{ji} \in \mathbb{Z}_{\geq 0}$, such that for a given $i$, $m_{ji}$ are not all zero. Then*

$$N_q^{A,*} = \frac{1}{q}\big[(q-1)^n + (-1)^r (q-1)^{n-r+1}\big] \\ - (q-1)^{n-r+1} \sum_{\alpha} T^{-\alpha_1}(a_1) T^{-\alpha_2}(a_2) \dots T^{-\alpha_r}(a_r) J(T^{\alpha_1}, T^{\alpha_2}, \dots, T^{\alpha_r}),$$

*where the sum is over all $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_r) \neq 0$ satisfying $0 \leq \alpha_i < q-1$, $\sum_{i=1}^{r} \alpha_i \equiv 0 \pmod{q-1}$, and $\sum_{i=1}^{r} m_{ji}\alpha_i \equiv 0 \pmod{q-1}$ for all $j \in \{1, 2, \dots, n\}$.*

A key step in proving the main results of this paper is to adapt Theorem 3.9 to $D_{d,\lambda,h}$.

**Corollary 3.10.** *Let $t := \gcd(d, q-1)$. For $\lambda \neq 0$,*

$$N_q^{A,*}(D_{d,\lambda,h}) = \sum_{s,w} J(T^{w_1 \frac{q-1}{t} + h_1 s}, T^{w_2 \frac{q-1}{t} + h_2 s}, \dots, T^{w_n \frac{q-1}{t} + h_n s})\, T^{ds}(d\lambda)$$

*where the sum is over all $s \in \big\{0, 1, \dots, \frac{q-1}{t} - 1\big\}$ and all $w = (w_1, w_2, \dots, w_n) \in W$.*

Corollary 3.10 generalizes Corollary 1 in [Koblitz 1983], which holds in the case $d \mid q-1$.

## 4. Proofs

*Proof of Corollary 3.10.* We take $r = n+1$; $a_i = 1$, for $i = 1, \dots, n$, and $a_r = -d\lambda$; $m_{ji} = d$ if $i = j$ and zero otherwise, and, $m_{jr} = h_j$, for all $j = 1, \dots, n$, in Theorem 3.9. This yields

$$\text{(4-1)} \quad N_q^{A,*}(D_{d,\lambda,h}) \\ = \tfrac{1}{q}\big[(q-1)^n + (-1)^{n+1}\big] - \sum_{\alpha} T^{-\alpha_{n+1}}(-d\lambda) J(T^{\alpha_1}, T^{\alpha_2}, \dots, T^{\alpha_{n+1}})$$

where the sum is over all $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_{n+1}) \neq 0$ satisfying $0 \leq \alpha_i < q-1$, $\sum_{i=1}^{n+1} \alpha_i \equiv 0 \pmod{q-1}$, and $d\,\alpha_j + h_j\,\alpha_{n+1} \equiv 0 \pmod{q-1}$ for all $j = 1, 2, \dots, n$.

The condition $d\,\alpha_j + h_j\,\alpha_{n+1} \equiv 0 \pmod{q-1}$, for all $j \in \{1, 2, \dots, n\}$, implies $t = \gcd(d, q-1)$ divides $h_j\,\alpha_{n+1}$ for all $j \in \{1, 2, \dots, n\}$. If $l^e$ is a prime power dividing $t$ but not $\alpha_{n+1}$, then $l$ divides $h_j$ for all $j \in \{1, 2, \dots, n\}$. This is a contradiction, as $\gcd(h_1, \dots, h_n) = 1$. Therefore, $l^e$ divides $\alpha_{n+1}$, which implies $t$ divides $\alpha_{n+1}$. So $\frac{\alpha_{n+1}}{t} \in \left\{0, 1, \dots, \frac{q-1}{t} - 1\right\}$. Let

$$s \equiv -\left(\frac{d}{t}\right)^{-1} \frac{\alpha_{n+1}}{t} \left(\operatorname{mod} \frac{q-1}{t}\right)$$

such that $s \in \left\{0, 1, \dots, \frac{q-1}{t} - 1\right\}$. Then $s$ runs around $\left\{0, 1, \dots, \frac{q-1}{t} - 1\right\}$ as $\frac{\alpha_{n+1}}{t}$ does.

We now express the conditions on $\alpha$ in terms of $s$. Firstly,

$$\begin{aligned} d\,\alpha_j \equiv -h_j\,\alpha_{n+1} \pmod{q-1} \quad &\Longrightarrow \quad \frac{d}{t}\alpha_j \equiv -h_j \frac{\alpha_{n+1}}{t} \left(\operatorname{mod} \frac{q-1}{t}\right) \\ &\Longrightarrow \quad \alpha_j \equiv h_j s \left(\operatorname{mod} \frac{q-1}{t}\right). \end{aligned}$$

So $\alpha_j = h_j s + w_j \frac{q-1}{t}$ for $w_j \in \{0, 1, \dots, t-1\}$, for $j \in \{1, 2, \dots, n\}$. Also,

$$\text{(4-2)} \qquad \frac{\alpha_{n+1}}{t} \equiv -\left(\frac{d}{t}\right)s \left(\operatorname{mod} \frac{q-1}{t}\right) \quad \Longrightarrow \quad \alpha_{n+1} \equiv -ds \pmod{q-1}.$$

Using the fact that $\sum_{i=j}^{n} h_j = d$, it is easy to see that

$$\text{(4-3)} \qquad \sum_{j=1}^{n} w_j = \sum_{j=1}^{n} \frac{t}{q-1}(\alpha_j - h_j s) = \frac{t}{q-1}\left(\sum_{j=1}^{n} \alpha_j - ds\right).$$

Combining (4-2) and (4-3) we get that

$$\sum_{j=1}^{n} w_j \equiv 0 \pmod{t} \quad \Longleftrightarrow \quad \sum_{i=1}^{n+1} \alpha_i \equiv 0 \pmod{q-1}.$$

Substituting for $\alpha$, (4-1) becomes

$$\text{(4-4)} \quad \begin{aligned} N_q^{A,*}(D_{d,\lambda,h}) = \frac{1}{q}[(q-1)^n + (-1)^{n+1}] \\ - \sum_{s,w} T^{ds}(-d\lambda) J\big(T^{w_1\frac{q-1}{t}+h_1 s}, \dots, T^{w_n\frac{q-1}{t}+h_n s}, T^{-ds}\big), \end{aligned}$$

where the sum is over all $s \in \left\{0, 1, \dots, \frac{q-1}{t} - 1\right\}$ and all $w = (w_1, w_2, \dots, w_n)$, such that $0 \le w_i < t$ and $\sum_{i=1}^{n} w_i \equiv 0 \pmod{t}$, and such that not all of $s, w_1, w_2, \dots, w_n$ are zero.

As $\sum_{i=1}^{n} w_i \frac{q-1}{t} + h_i s - ds \equiv 0 \pmod{q-1}$, by Proposition 3.2 we have

$$J\big(T^{w_1\frac{q-1}{t}+h_1 s}, \dots, T^{w_n\frac{q-1}{t}+h_n s}, T^{-ds}\big) = -J\big(T^{w_1\frac{q-1}{t}+h_1 s}, \dots, T^{w_n\frac{q-1}{t}+h_n s}\big) T^{-ds}(-1),$$

and by Proposition 3.3 we have

$$J(\underbrace{T^0, T^0, \dots, T^0}_{n \text{ times}}) = \frac{1}{q}[(q-1)^n + (-1)^{n+1}],$$

completing the proof. $\square$

*Proof of Theorem 2.1.* We follow Koblitz [1983, Theorem 2] and note that

$$N_q(D_{d,\lambda,h}) - N_q^*(D_{d,\lambda,h}) = N_q(D_{d,0,h}) - N_q^*(D_{d,0,h}). \tag{4-5}$$

We know

$$N_q(D_{d,0,h}) = \frac{N_q^A(D_d) - 1}{q-1} = \frac{q^{n-1}-1}{q-1} - \sum_{w^*} J(T^{w_1\frac{q-1}{t}}, T^{w_2\frac{q-1}{t}}, \dots, T^{w_n\frac{q-1}{t}}) \tag{4-6}$$

by Weil's result, Theorem 3.7 above;

$$N_q^*(D_{d,\lambda,h}) = \frac{1}{q-1}\sum_{s,w} J(T^{w_1\frac{q-1}{t}+h_1 s}, T^{w_2\frac{q-1}{t}+h_2 s}, \dots, T^{w_n\frac{q-1}{t}+h_n s})\, T^{ds}(d\lambda)$$

when $\lambda \neq 0$, by Corollary 3.10; and

$$N_q^*(D_{d,0,h}) = N_q^*(D_d) = \frac{1}{q-1}\sum_{w} J_0(T^{w_1\frac{q-1}{t}}, T^{w_2\frac{q-1}{t}}, \dots, T^{w_n\frac{q-1}{t}})$$

by Theorem 3.8.

Using Propositions 3.1, 3.3 and 3.4, we get that for $\lambda \neq 0$,

$$\begin{aligned}
&(q-1)\big(N_q^*(D_{d,\lambda,h}) - N_q^*(D_{d,0,h})\big)\\
&= \sum_{\substack{s,w\\ s\neq 0}} J(T^{w_1\frac{q-1}{t}+h_1 s}, T^{w_2\frac{q-1}{t}+h_2 s}, \dots, T^{w_n\frac{q-1}{t}+h_n s})\,T^{ds}(d\lambda)\\
&\quad + \sum_{w} J(T^{w_1\frac{q-1}{t}}, T^{w_2\frac{q-1}{t}}, \dots, T^{w_n\frac{q-1}{t}}) - \sum_{w} J_0(T^{w_1\frac{q-1}{t}}, T^{w_2\frac{q-1}{t}}, \dots, T^{w_n\frac{q-1}{t}})\\
&= \sum_{\substack{s,w\\ s\neq 0}} J(T^{w_1\frac{q-1}{t}+h_1 s}, T^{w_2\frac{q-1}{t}+h_2 s}, \dots, T^{w_n\frac{q-1}{t}+h_n s})\,T^{ds}(d\lambda)\\
&\quad + q\sum_{\substack{w\\ w\neq 0}} J(T^{w_1\frac{q-1}{t}}, T^{w_2\frac{q-1}{t}}, \dots, T^{w_n\frac{q-1}{t}}) + q\,J(\varepsilon,\varepsilon,\dots,\varepsilon) - (q-1)^n\\
&= \sum_{\substack{s,w\\ s\neq 0}} J(T^{w_1\frac{q-1}{t}+h_1 s}, T^{w_2\frac{q-1}{t}+h_2 s}, \dots, T^{w_n\frac{q-1}{t}+h_n s})\,T^{ds}(d\lambda)\\
&\quad + q\sum_{\substack{w\\ w\neq 0}} J(T^{w_1\frac{q-1}{t}}, T^{w_2\frac{q-1}{t}}, \dots, T^{w_n\frac{q-1}{t}}) + (-1)^{n+1}\\
&= \sum_{s,w} \frac{g(T^{w_1\frac{q-1}{t}+h_1 s})\, g(T^{w_2\frac{q-1}{t}+h_2 s}) \dots g(T^{w_n\frac{q-1}{t}+h_n s})}{g(T^{ds})}\,T^{ds}(d\lambda).
\end{aligned} \tag{4-7}$$

Combining (4-5), (4-6) and (4-7), which trivially hold for $\lambda = 0$ also, yields the result. □

*Proof of Corollary 2.2.* Applying (3-1) and Proposition 3.4 to Theorem 2.1 we get that

$$
\begin{aligned}
&N_q(D_{d,\lambda,h})\\
&\quad= \frac{q^{n-1}-1}{q-1} + \frac{1}{q}\sum_{w^*}\prod_{i=1}^{n} g(T^{w_i\frac{q-1}{t}}) - \frac{1}{q-1}\sum_{w}\prod_{i=1}^{n} g(T^{w_i\frac{q-1}{t}})\\
&\qquad + \frac{1}{q(q-1)}\sum_{\substack{s,w\\ s\neq 0}}\prod_{i=1}^{n} g(T^{w_i\frac{q-1}{t}+h_i s})g(T^{-ds})\,T^{ds}(-d\lambda)\\
&\quad= \frac{q^{n-1}-1}{q-1} + \frac{1}{q}\sum_{w^*}\prod_{i=1}^{n} g(T^{w_i\frac{q-1}{t}}) - \frac{1}{q-1}\sum_{w}\prod_{i=1}^{n} g(T^{w_i\frac{q-1}{t}})\Big(1-\frac{1}{q}\Big)\\
&\qquad + \frac{1}{q(q-1)}\sum_{s,w}\prod_{i=1}^{n} g(T^{w_i\frac{q-1}{t}+h_i s})g(T^{-ds})\,T^{ds}(-d\lambda)\\
&\quad= \frac{q^{n-1}-1}{q-1} - \frac{1}{q}\sum_{\substack{w\\ \text{some } w_i=0}}\prod_{i=1}^{n} g(T^{w_i\frac{q-1}{t}})\\
&\qquad + \frac{1}{q(q-1)}\sum_{s,w}\prod_{i=1}^{n} g(T^{w_i\frac{q-1}{t}+h_i s})g(T^{-ds})\,T^{ds}(-d\lambda),
\end{aligned}
$$

where the last sum is over all $s \in \left\{0,1,\dots,\frac{q-1}{t}-1\right\}$ and all $w=(w_1,w_2,\dots,w_n)\in W$, as required. To get the alternative summation limits, we note that

$$
\begin{aligned}
&\sum_{s=0}^{\frac{q-1}{t}-1}\sum_{w\in W}\prod_{i=1}^{n} g(T^{w_i\frac{q-1}{t}+h_i s})g(T^{-ds})\,T^{ds}(-d\lambda) \qquad (4\text{-}8)\\
&\qquad= \sum_{s=0}^{\frac{q-1}{t}-1}\sum_{j=0}^{t-1}\sum_{[w]\in W/\sim} g(T^{(w_i+jh_i)\frac{q-1}{t}+h_i s})g(T^{-ds})\,T^{ds}(-d\lambda)\\
&\qquad= \sum_{s=0}^{\frac{q-1}{t}-1}\sum_{j=0}^{t-1}\sum_{[w]\in W/\sim} g(T^{w_i\frac{q-1}{t}+h_i(s+j\frac{q-1}{t})})g(T^{-ds})\,T^{ds}(-d\lambda)\\
&\qquad= \sum_{s=0}^{q-2}\sum_{[w]\in W/\sim}\prod_{i=1}^{n} g(T^{w_i\frac{q-1}{t}+h_i s})g(T^{-ds})\,T^{ds}(-d\lambda).
\end{aligned}
$$

This sum is independent of the choice of equivalence class representatives $[w]$, as changing representative can be countered by a simple change of variable in $s$. $\square$

*Proof of Corollary 2.7.* We start from Corollary 2.2 with $h = (1, 1, \ldots, 1)$ and $d = n$, and rewrite using the notation described in Section 2, i.e.,

$$N_q(D_\lambda) = \frac{q^{n-1}-1}{q-1} - \frac{1}{q} \sum_{\substack{w \in W \\ 0 \in S_w^c}} \prod_{k \in S_w^c} g(T^{k\frac{q-1}{t}})^{n_k} + \frac{1}{q(q-1)} \sum_{s,w} \prod_{k \in S_w^c} g(T^{k\frac{q-1}{t}+s})^{n_k} g(T^{-ns})\, T^{ns}(-n\lambda), \tag{4-9}$$

where $t = \gcd(n, q-1)$, and the second sum is over all $s \in \left\{0, 1, \ldots, \frac{q-1}{t} - 1\right\}$ and all $w \in W$. We proceed in the same fashion as the proof of Theorem 2.2 in [McCarthy 2017]. By (3-1) it is easy to see that

$$\sum_{\substack{w \\ 0 \in S_w^c}} \prod_{k \in S_w^c} g(T^{k\frac{q-1}{t}})^{n_k} = \sum_{\substack{w \\ 0 \in S_w^c}} \Bigg[\prod_{k \in S_w^c} \frac{g(T^{k\frac{q-1}{t}})^{n_k-1}}{g(T^{-k\frac{q-1}{t}})}\Bigg]\Bigg[\prod_{k \in S_w^c \setminus \{0\}} T^{k\frac{q-1}{t}}(-1)\, q\Bigg]. \tag{4-10}$$

We now focus on the second sum in (4-9). If $T^{k\frac{q-1}{t}+s} = \varepsilon$ then $k\frac{q-1}{t} + s \equiv 0 \pmod{q-1}$, which can only happen if $s \equiv 0 \left(\mathrm{mod}\ \frac{q-1}{t}\right)$, in which case $s = 0$. So, if $s \neq 0$ then $T^{k\frac{q-1}{t}+s} \neq \varepsilon$. Again using (3-1), we see that, for $\lambda \neq 0$,

$$\begin{aligned}
&\sum_{w \in W} \sum_{s=0}^{\frac{q-1}{t}} \prod_{k \in S_w^c} g(T^{k\frac{q-1}{t}+s})^{n_k} g(T^{-ns})\, T^{ns}(-n\lambda) \\
&= \sum_{w \in W} \sum_{s=1}^{\frac{q-1}{t}} \Bigg[\prod_{k \in S_w^c} \frac{g(T^{k\frac{q-1}{t}+s})^{n_k-1}}{g(T^{-k\frac{q-1}{t}-s})} T^{k\frac{q-1}{t}+s}(-1)\, q\Bigg] g(T^{-ns})\, T^{ns}(-n\lambda) \\
&\quad - \sum_{w \in W} \Bigg[\prod_{k \in S_w^c} \frac{g(T^{k\frac{q-1}{t}})^{n_k-1}}{g(T^{-k\frac{q-1}{t}})}\Bigg]\Bigg[\prod_{k \in S_w^c \setminus \{0\}} T^{k\frac{q-1}{t}}(-1)\, q\Bigg] \\
&= \sum_{w \in W} \sum_{s=0}^{\frac{q-1}{t}} \Bigg[\prod_{k \in S_w^c} \frac{g(T^{k\frac{q-1}{t}+s})^{n_k-1}}{g(T^{-k\frac{q-1}{t}-s})} T^{k\frac{q-1}{t}+s}(-1)\, q\Bigg] g(T^{-ns})\, T^{ns}(-n\lambda) \\
&\quad + \sum_{w \in W} \Bigg[\prod_{k \in S_w^c} \frac{g(T^{k\frac{q-1}{t}})^{n_k-1}}{g(T^{-k\frac{q-1}{t}})}\Bigg]\Bigg[\prod_{k \in S_w^c} T^{k\frac{q-1}{t}}(-1)\, q - \prod_{k \in S_w^c \setminus \{0\}} T^{k\frac{q-1}{t}}(-1)\, q\Bigg] \\
&= \sum_{w \in W} \sum_{s=0}^{\frac{q-1}{t}} \Bigg[\prod_{k \in S_w^c} \frac{g(T^{k\frac{q-1}{t}+s})^{n_k-1}}{g(T^{-k\frac{q-1}{t}-s})} T^{k\frac{q-1}{t}+s}(-1)\, q\Bigg] g(T^{-ns})\, T^{ns}(-n\lambda) \\
&\quad + (q-1) \sum_{\substack{w \\ 0 \in S_w^c}} \Bigg[\prod_{k \in S_w^c} \frac{g(T^{k\frac{q-1}{t}})^{n_k-1}}{g(T^{-k\frac{q-1}{t}})}\Bigg]\Bigg[\prod_{k \in S_w^c \setminus \{0\}} T^{k\frac{q-1}{t}}(-1)\, q\Bigg].
\end{aligned} \tag{4-11}$$

Accounting for (4-10) and (4-11) in (4-9) yields

$$N_q(D_\lambda)=\frac{q^{n-1}-1}{q-1}$$
$$+\frac{1}{q(q-1)}\sum_{w\in W}\sum_{s=0}^{\frac{q-1}{t}}\Bigg[\prod_{k\in S_w^c}\frac{g(T^{k\frac{q-1}{t}+s})^{n_k-1}}{g(T^{-k\frac{q-1}{t}-s})}T^{k\frac{q-1}{t}+s}(-1)\,q\Bigg]g(T^{-ns})\,T^{ns}(-n\lambda).$$

To get the alternative summation limit, proceed in the same manner as in (4-8). $\square$

*Proof of Theorem 2.4.* We start from Corollary 2.2, which we rewrite as

$$N_q(D_{d,\lambda,h})=\frac{q^{n-1}-1}{q-1}-\frac{1}{q}\sum_{\substack{w\in W\\ \text{some } w_i=0}}\prod_{i=1}^{n}g(T^{w_i\frac{q-1}{t}})+\frac{1}{q(q-1)}\sum_{[w]\in W/\sim}R_{[w]}, \tag{4-12}$$

where

$$R_{[w]}:=\sum_{s=0}^{q-2}\prod_{i=1}^{n}g(T^{w_i\frac{q-1}{t}+h_is})g(T^{-ds})\,T^{ds}(-d\lambda).$$

We note $R_{[w]}$ is independent of the choice of equivalence class representative.

We now let $T=\overline{\omega}$ and apply the Gross–Koblitz formula, Theorem 3.6, to both summands in (4-12). From the first summand we get that

$$\begin{aligned}\prod_{i=1}^{n}g(T^{w_i\frac{q-1}{t}})&=(-1)^n(-p)^{\sum_{i=1}^{n}\sum_{a=0}^{r-1}\langle(\frac{w_i}{t})p^a\rangle}\prod_{i=1}^{n}\prod_{a=0}^{r-1}\Gamma_p\big(\langle(\tfrac{w_i}{t})p^a\rangle\big)\\&=(-1)^n\,C(w).\end{aligned} \tag{4-13}$$

The second, $R_{[w]}$, yields

$$\begin{aligned}R_{[w]}=(-1)^{n+1}\sum_{s=0}^{q-2}\Bigg[\prod_{a=0}^{r-1}\prod_{i=1}^{n}\Gamma_p\big(\langle(\tfrac{w_i}{t}+\tfrac{h_is}{q-1})p^a\rangle\big)\Bigg]\Bigg[\prod_{a=0}^{r-1}\Gamma_p\big(\langle(\tfrac{-ds}{q-1})p^a\rangle\big)\Bigg]\\ \times(-p)^v\,\overline{\omega}^{ds}(-d\lambda),\end{aligned} \tag{4-14}$$

where

$$\begin{aligned}v&=\sum_{a=0}^{r-1}\sum_{i=1}^{n}\langle(\tfrac{w_i}{t}+\tfrac{h_is}{q-1})p^a\rangle+\sum_{a=0}^{r-1}\langle(\tfrac{-ds}{q-1})p^a\rangle\\&=\sum_{a=0}^{r-1}\sum_{i=1}^{n}(\tfrac{w_i}{t}+\tfrac{h_is}{q-1})p^a+\sum_{a=0}^{r-1}(\tfrac{-ds}{q-1})p^a-\sum_{a=0}^{r-1}\sum_{i=1}^{n}\lfloor(\tfrac{w_i}{t}+\tfrac{h_is}{q-1})p^a\rfloor-\sum_{a=0}^{r-1}\lfloor(\tfrac{-ds}{q-1})p^a\rfloor\\&=\sum_{a=0}^{r-1}\sum_{i=1}^{n}(\tfrac{w_i}{t})p^a-\sum_{a=0}^{r-1}\sum_{i=1}^{n}\lfloor(\tfrac{w_i}{t}+\tfrac{h_is}{q-1})p^a\rfloor-\sum_{a=0}^{r-1}\lfloor(\tfrac{-ds}{q-1})p^a\rfloor\in\mathbb{Z}\end{aligned}$$

as $\sum_{i=1}^{n}h_i=d$ and $\sum_{i=1}^{n}w_i\equiv 0 \pmod t$.

We will now use Theorem 3.5 to expand the terms involving the $p$-adic gamma function in (4-14). Let $k \in \mathbb{Z}$ such that

$$k \le \frac{w_i}{t} + \frac{h_i s}{q-1} < k+1.$$

Then $0 \le x := \frac{w_i}{t} + \frac{h_i s}{q-1} - k < 1$ and $(q-1)x \in \mathbb{Z}$. So, by Theorem 3.5, with $h = h_i$ and $p \nmid h_i$,

$$\prod_{a=0}^{r-1} \prod_{b=0}^{h_i-1} \Gamma_p\Big(\Big\langle\Big(\frac{w_i}{th_i} + \frac{s}{q-1} + \frac{b-k}{h_i}\Big)p^a\Big\rangle\Big)$$
$$= \omega\big(h_i^{w_i \frac{q-1}{t} + h_i s}\big) \prod_{a=0}^{r-1} \Gamma_p\Big(\Big\langle\Big(\frac{w_i}{t} + \frac{h_i s}{q-1}\Big)p^a\Big\rangle\Big) \prod_{b=1}^{h_i-1} \Gamma_p\Big(\Big\langle\Big(\frac{b}{h_i}\Big)p^a\Big\rangle\Big).$$

As $\{b \mid b = 0, 1, \ldots, h_i - 1\} \equiv \{b - k \mid b = 0, 1, \ldots, h_i - 1\} \pmod{h_i}$ we have

$$\prod_{a=0}^{r-1} \prod_{b=0}^{h_i-1} \Gamma_p\Big(\Big\langle\Big(\frac{w_i}{th_i} + \frac{s}{q-1} + \frac{b}{h_i}\Big)p^a\Big\rangle\Big) \tag{4-15}$$
$$= \omega\big(h_i^{w_i \frac{q-1}{t} + h_i s}\big) \prod_{a=0}^{r-1} \Gamma_p\Big(\Big\langle\Big(\frac{w_i}{t} + \frac{h_i s}{q-1}\Big)p^a\Big\rangle\Big) \prod_{b=1}^{h_i-1} \Gamma_p\Big(\Big\langle\Big(\frac{b}{h_i}\Big)p^a\Big\rangle\Big).$$

Similarly, with $k \in \mathbb{Z}$ chosen such that $0 \le x := \frac{w_i}{t} - k < 1$, we apply Theorem 3.5 to get that

$$\prod_{a=0}^{r-1} \prod_{b=0}^{h_i-1} \Gamma_p\Big(\Big\langle\Big(\frac{w_i}{th_i} + \frac{b}{h_i}\Big)p^a\Big\rangle\Big) \tag{4-16}$$
$$= \omega\big(h_i^{w_i \frac{q-1}{t}}\big) \prod_{a=0}^{r-1} \Gamma_p\Big(\Big\langle\Big(\frac{w_i}{t}\Big)p^a\Big\rangle\Big) \prod_{b=1}^{h_i-1} \Gamma_p\Big(\Big\langle\Big(\frac{b}{h_i}\Big)p^a\Big\rangle\Big).$$

Combining (4-15) and (4-16) we have, for $p \nmid h_i$,

$$\prod_{a=0}^{r-1} \Gamma_p\Big(\Big\langle\Big(\frac{w_i}{t} + \frac{h_i s}{q-1}\Big)p^a\Big\rangle\Big) \tag{4-17}$$
$$= \prod_{a=0}^{r-1} \prod_{b=0}^{h_i-1} \frac{\Gamma_p\Big(\Big\langle\Big(\frac{w_i}{th_i} + \frac{s}{q-1} + \frac{b}{h_i}\Big)p^a\Big\rangle\Big)}{\Gamma_p\Big(\Big\langle\Big(\frac{w_i}{th_i} + \frac{b}{h_i}\Big)p^a\Big\rangle\Big)} \prod_{a=0}^{r-1} \Gamma_p\Big(\Big\langle\Big(\frac{w_i}{t}\Big)p^a\Big\rangle\Big)\, \overline{\omega}^s(h_i^{h_i}).$$

A final application of Theorem 3.5, this time with $k \in \mathbb{Z}$ such that $0 \le x := k - \frac{ds}{q-1} < 1$ and $p \nmid d$, we get, after some simplification, that

$$\prod_{a=0}^{r-1} \Gamma_p\Big(\Big\langle\Big(\frac{-ds}{q-1}\Big)p^a\Big\rangle\Big) = \prod_{a=0}^{r-1} \prod_{b=0}^{d-1} \frac{\Gamma_p\Big(\Big\langle\Big(\frac{-b}{d} - \frac{s}{q-1}\Big)p^a\Big\rangle\Big)}{\Gamma_p\Big(\Big\langle\Big(\frac{-b}{d}\Big)p^a\Big\rangle\Big)}\, \overline{\omega}^s(d^{-d}). \tag{4-18}$$

Accounting for (4-17) and (4-18) in (4-14) and making the change of variable $s \to (q-1)-s$ we get that

$$\text{(4-19)}\quad R_{[w]} = (-1)^{n+1} \prod_{i=1}^{n} \prod_{a=0}^{r-1} \Gamma_p\big(\langle\langle(\tfrac{w_i}{t})p^a\rangle\rangle\big) \sum_{s=0}^{q-2} (-1)^{sd} \times \Bigg[\prod_{i=1}^{n}\prod_{b=0}^{h_i-1}\prod_{a=0}^{r-1} \frac{\Gamma_p\big(\langle\langle(\frac{w_i}{th_i}+\frac{b}{h_i}-\frac{s}{q-1})p^a\rangle\rangle\big)}{\Gamma_p\big(\langle\langle(\frac{w_i}{th_i}+\frac{b}{h_i})p^a\rangle\rangle\big)}\Bigg] \times \Bigg[\prod_{b=0}^{d-1}\prod_{a=0}^{r-1} \frac{\Gamma_p\big(\langle\langle(\frac{-b}{d}+\frac{s}{q-1})p^a\rangle\rangle\big)}{\Gamma_p\big(\langle\langle(\frac{-b}{d})p^a\rangle\rangle\big)}\Bigg](-p)^y\, \overline{\omega}^s\Bigg(\Bigg[\lambda^d \prod_{i=1}^{n} h_i^{h_i}\Bigg]^{-1}\Bigg),$$

where

$$y = \sum_{a=0}^{r-1}\sum_{i=1}^{n} (\tfrac{w_i}{t})p^a - \sum_{a=0}^{r-1}\sum_{i=1}^{n} \big\lfloor(\tfrac{w_i}{t}-\tfrac{h_i s}{q-1})p^a\big\rfloor - \sum_{a=0}^{r-1}\big\lfloor(\tfrac{ds}{q-1})p^a\big\rfloor,$$

and we have used the fact that $\overline{\omega}(-1) = -1$. Let

$$z := -\Bigg[\sum_{i=1}^{n}\sum_{a=0}^{r-1}\sum_{b=0}^{h_i-1} \big\lfloor\langle(\tfrac{w_i}{th_i}+\tfrac{b}{h_i})p^a\rangle - \tfrac{sp^a}{q-1}\big\rfloor + \sum_{a=0}^{r-1}\sum_{b=0}^{d-1}\big\lfloor\langle(\tfrac{-b}{d})p^a\rangle + \tfrac{sp^a}{q-1}\big\rfloor\Bigg].$$

Using the fact that $\lfloor mx\rfloor = \sum_{b=0}^{m-1}\lfloor x+\frac{b}{m}\rfloor$ we get that

$$y - z = \sum_{a=0}^{r-1}\sum_{i=1}^{n}(\tfrac{w_i}{t})p^a - \sum_{a=0}^{r-1}\sum_{i=1}^{n}\sum_{b=0}^{h_i-1}\big\lfloor(\tfrac{w_i}{th_i}-\tfrac{s}{q-1})p^a+\tfrac{b}{h_i}\big\rfloor - \sum_{a=0}^{r-1}\sum_{b=0}^{d-1}\big\lfloor\tfrac{sp^a}{q-1}+\tfrac{b}{d}\big\rfloor + \sum_{i=1}^{n}\sum_{a=0}^{r-1}\sum_{b=0}^{h_i-1}\big\lfloor\langle(\tfrac{w_i}{th_i}+\tfrac{b}{h_i})p^a\rangle - \tfrac{sp^a}{q-1}\big\rfloor + \sum_{a=0}^{r-1}\sum_{b=0}^{d-1}\big\lfloor\langle(\tfrac{-b}{d})p^a\rangle+\tfrac{sp^a}{q-1}\big\rfloor$$

As $\gcd(p,d)=1$, $\{b \mid b=0,1,\dots,d-1\} \equiv \{bp^a \mid b = 0,1,\dots,d-1\} \pmod d$ and so

$$\sum_{b=0}^{d-1}\big\lfloor\langle(\tfrac{-b}{d})p^a\rangle+\tfrac{sp^a}{q-1}\big\rfloor = \sum_{b=0}^{d-1}\big\lfloor\langle(\tfrac{b}{d})p^a\rangle+\tfrac{sp^a}{q-1}\big\rfloor = \sum_{b=0}^{d-1}\big\lfloor\langle\tfrac{b}{d}\rangle+\tfrac{sp^a}{q-1}\big\rfloor = \sum_{b=0}^{d-1}\big\lfloor\tfrac{b}{d}+\tfrac{sp^a}{q-1}\big\rfloor.$$

Similarly, as $\gcd(p,h_i)=1$,

$$\begin{aligned}\sum_{b=0}^{h_i-1}\big\lfloor\langle(\tfrac{w_i}{th_i}+\tfrac{b}{h_i})p^a\rangle - \tfrac{sp^a}{q-1}\big\rfloor &= \sum_{b=0}^{h_i-1}\big\lfloor\langle(\tfrac{w_i}{th_i})p^a+\tfrac{b}{h_i}\rangle - \tfrac{sp^a}{q-1}\big\rfloor \\ &= \sum_{b=0}^{h_i-1}\big\lfloor(\tfrac{w_i}{th_i})p^a+\tfrac{b}{h_i} - \big\lfloor(\tfrac{w_i}{th_i})p^a+\tfrac{b}{h_i}\big\rfloor - \tfrac{sp^a}{q-1}\big\rfloor\end{aligned}$$

$$= \sum_{b=0}^{h_i-1} \left\lfloor \left(\tfrac{w_i}{th_i} - \tfrac{s}{q-1}\right) p^a + \tfrac{b}{h_i} \right\rfloor - \sum_{b=0}^{h_i-1} \left\lfloor \left(\tfrac{w_i}{th_i}\right) p^a + \tfrac{b}{h_i} \right\rfloor$$

$$= \sum_{b=0}^{h_i-1} \left\lfloor \left(\tfrac{w_i}{th_i} - \tfrac{s}{q-1}\right) p^a + \tfrac{b}{h_i} \right\rfloor - \left\lfloor \left(\tfrac{w_i}{t}\right) p^a \right\rfloor.$$

So

$$y - z = \sum_{a=0}^{r-1} \sum_{i=1}^{n} \left(\tfrac{w_i}{t}\right) p^a - \sum_{a=0}^{r-1} \sum_{i=1}^{n} \left\lfloor \left(\tfrac{w_i}{t}\right) p^a \right\rfloor = \sum_{a=0}^{r-1} \sum_{i=1}^{n} \left\langle \left(\tfrac{w_i}{t}\right) p^a \right\rangle.$$

Thus

$$\frac{1}{q-1} R_{[w]} = (-1)^n C(w) \tag{4-20}$$

$$\times \frac{-1}{q-1} \sum_{s=0}^{q-2} (-1)^{sd} \left[ \prod_{i=1}^{n} \prod_{b=0}^{h_i-1} \prod_{a=0}^{r-1} \frac{\Gamma_p\left(\left\langle \left(\frac{w_i}{th_i} + \frac{b}{h_i} - \frac{s}{q-1}\right) p^a \right\rangle\right)}{\Gamma_p\left(\left\langle \left(\frac{w_i}{th_i} + \frac{b}{h_i}\right) p^a \right\rangle\right)} \right]$$

$$\times \left[ \prod_{b=0}^{d-1} \prod_{a=0}^{r-1} \frac{\Gamma_p\left(\left\langle \left(\frac{-b}{d} + \frac{s}{q-1}\right) p^a \right\rangle\right)}{\Gamma_p\left(\left\langle \left(\frac{-b}{d}\right) p^a \right\rangle\right)} \right] (-p)^z \, \overline{\omega}^s \left( \left[ \lambda^d \prod_{i=1}^{n} h_i^{h_i} \right]^{-1} \right)$$

$$= (-1)^n C(w) \, {}_dG_d \left[ \begin{matrix} \cdots\cdots \frac{w_i}{th_i} + \frac{b}{h_i} \cdots\cdots \\ 1 \;\; \frac{1}{d} \;\; \frac{2}{d} \;\; \cdots \;\; \frac{d-1}{d} \end{matrix} \;\middle|\; (\lambda^d h_1^{h_1} \cdots h_n^{h_n})^{-1} \right]_q .$$

Substituting for (4-13) and (4-20) in (4-12), we get the required result. □

*Proof of Corollary 2.5.* In Theorem 2.4, we make the change of variables $w \to -w \pmod{t}$, which is a bijection on $W/\sim$, and $s \to (q-1) - s$ in the expansion of ${}_dG_d$ by definition. □

*Proof of Corollary 2.6.* If $t = \gcd(d, q-1) = 1$ then $w = (0, 0, \ldots, 0)$ is the only element in $W$ and $C(0) = 1$. So, by Corollary 2.5

$$N_q(D_{d,\lambda,h}) = \frac{q^{n-1}-1}{q-1} + \frac{(-1)^n}{q} \left( -1 + {}_dG_d \left[ \begin{matrix} 0 \;\; \frac{1}{d} \;\; \frac{2}{d} \;\; \cdots \;\; \frac{d-1}{d} \\ \cdots\cdots \frac{b_i}{h_i} \cdots\cdots \end{matrix} \;\middle|\; \lambda^d h_1^{h_1} \cdots h_n^{h_n} \right]_q \right).$$

The first bottom line parameter in ${}_dG_d$ is $\frac{0}{h_1} = 0$. We will "cancel" the zero from both top and bottom to get the required ${}_{d-1}G_{d-1}$. From Definition 2.3 we see that the contribution to the summand of the top and bottom line zero is

$$\prod_{k=0}^{r-1} \frac{\Gamma_p\left(\left\langle \left(0 - \frac{s}{q-1}\right) p^k \right\rangle\right)}{\Gamma_p(\langle 0 p^k \rangle)} \frac{\Gamma_p\left(\left\langle \left(0 + \frac{s}{q-1}\right) p^k \right\rangle\right)}{\Gamma_p(\langle 0 p^k \rangle)} (-p)^{-\left\lfloor \langle 0p^k \rangle - \frac{sp^k}{q-1} \right\rfloor - \left\lfloor \langle 0p^k \rangle + \frac{sp^k}{q-1} \right\rfloor}$$

which, by Theorem 3.6 and (3-1), equals

$$g(\overline{\omega}^{-s}) g(\overline{\omega}^s) = \begin{cases} \overline{\omega}^s(-1)\, q & \text{if } s \neq 0, \\ 1 & \text{if } s = 0. \end{cases}$$

We also note that when $s = 0$ the summand in Definition 2.3 equals 1. Therefore,

$$ {}_{d}G_{d}\left[\begin{matrix} 0, & a_2, & \dots, & a_n \\ 0, & b_2, & \dots, & b_n \end{matrix} \;\middle|\; \lambda \right]_q = 1 + q \cdot {}_{d-1}G_{d-1}\left[\begin{matrix} a_2, & \dots, & a_n \\ b_2, & \dots, & b_n \end{matrix} \;\middle|\; \lambda \right]_q $$

as required. □

*Proof of Theorem 2.8.* We start from Corollary 2.7 and proceed in the same fashion as the second half of the proof of Theorem 2.2 in [McCarthy 2017]. We let $T = \overline{\omega}$ and apply the Gross–Koblitz formula, Theorem 3.6, to get

$$ N_q(D_\lambda) = \frac{q^{n-1}-1}{q-1} + \frac{1}{q(q-1)} \sum_{[w]\in W/\sim_1} R_{[w]}, \tag{4-21} $$

where

$$ R_{[w]} = \sum_{s=0}^{q-2} (-1)^{n+1} (-p)^{v}\, \overline{\omega}^{ns}(-n\lambda) \prod_{a=0}^{r-1} \Gamma_p\big(\big\langle\big(\tfrac{-ns}{q-1}\big)p^a\big\rangle\big) \\ \times \prod_{k\in S_w^c} \overline{\omega}(-1)^{k\frac{q-1}{t}+s}\, q \prod_{a=0}^{r-1} \frac{\Gamma_p\big(\big\langle\big(\frac{k}{t}+\frac{s}{q-1}\big)p^a\big\rangle\big)^{n_k-1}}{\Gamma_p\big(\big\langle\big(\frac{-k}{t}-\frac{s}{q-1}\big)p^a\big\rangle\big)} $$

with

$$ v = \sum_{k\in S_w^c} \frac{n_k k}{t} \sum_{a=0}^{r-1} p^a - \sum_{k\in S_w^c} (n_k-1) \sum_{a=0}^{r-1} \big\lfloor \big(\tfrac{k}{t}+\tfrac{s}{q-1}\big)p^a \big\rfloor \\ + \sum_{k\in S_w^c} \sum_{a=0}^{r-1} \big\lfloor \big(-\tfrac{k}{t}-\tfrac{s}{q-1}\big)p^a \big\rfloor - \sum_{a=0}^{r-1} \big\lfloor \big(\tfrac{-ns}{q-1}\big)p^a \big\rfloor. $$

As $p \nmid n$ we derive from (4-18) that

$$ \prod_{a=0}^{r-1} \Gamma_p\big(\big\langle\big(\tfrac{-ns}{q-1}\big)p^a\big\rangle\big) \\ = \prod_{a=0}^{r-1} \frac{\prod_{k=0}^{t-1} \Gamma_p\big(\big\langle\big(\frac{k}{t}-\frac{s}{q-1}\big)p^a\big\rangle\big) \prod\limits_{\substack{b=0\\ b\not\equiv 0 \ (\mathrm{mod}\ \frac{n}{t})}}^{n-1} \Gamma_p\big(\big\langle\big(\frac{b}{n}-\frac{s}{q-1}\big)p^a\big\rangle\big)}{\prod_{b=0}^{n-1} \Gamma_p\big(\big\langle\big(\frac{b}{n}\big)p^a\big\rangle\big)}\, \overline{\omega}^{s}(n^{-n}). $$

So, after some manipulation,

$$R_{[w]} = (-1)^{n+1} \sum_{s=0}^{q-2} (-p)^{v}\, \overline{\omega}^{ns}(-\lambda) \left[ \prod_{k\in S_w} \prod_{a=0}^{r-1} \frac{\Gamma_p\big(\big\langle\big(\frac{t-k}{t} - \frac{s}{q-1}\big)p^a\big\rangle\big)}{\Gamma_p\big(\big\langle\frac{t-k}{t}p^a\big\rangle\big)} \right] \times \left[ \prod_{\substack{b=0\\ b\not\equiv 0 \ (\mathrm{mod}\ \frac{n}{t})}}^{n-1} \prod_{a=0}^{r-1} \frac{\Gamma_p\big(\big\langle\big(\frac{b}{n} - \frac{s}{q-1}\big)p^a\big\rangle\big)}{\Gamma_p\big(\big\langle\frac{b}{n}p^a\big\rangle\big)} \right] \times \left[ \prod_{k\in S_w^c} \prod_{a=0}^{r-1} \frac{\Gamma_p\big(\big\langle\big(-\frac{t-k}{t} + \frac{s}{q-1}\big)p^a\big\rangle\big)^{n_k-1}}{\Gamma_p\big(\big\langle -\frac{t-k}{t}p^a\big\rangle\big)^{n_k-1}} \right] F[w], \tag{4-22}$$

where

$$F[w] := \left[ \prod_{k\in S_w^c} \prod_{a=0}^{r-1} \Gamma_p\big(\big\langle\big(\tfrac{-k}{t}\big)p^a\big\rangle\big) \right]^{-1} \left[ \prod_{k\in S_w^c} \prod_{a=0}^{r-1} \Gamma_p\big(\big\langle\tfrac{k}{t}p^a\big\rangle\big)^{n_k-1} \right] \left[ \prod_{k\in S_w^c} \overline{\omega}(-1)^{k\frac{q-1}{t}+s}\, q \right].$$

Applying the Gross–Koblitz formula, Theorem 3.6, in reverse and (3-1) we get that

$$\begin{aligned} \prod_{k\in S_w^c} \prod_{a=0}^{r-1} \Gamma_p\big(\big\langle\big(\tfrac{-k}{t}\big)p^a\big\rangle\big)\Gamma_p\big(\big\langle\tfrac{k}{t}p^a\big\rangle\big) &= \prod_{k\in S_w^c} g\big(\overline{\omega}^{-k\frac{q-1}{t}}\big) g\big(\overline{\omega}^{k\frac{q-1}{t}}\big) (-p)^{-\sum_{a=0}^{r-1} \langle(\frac{-k}{t})p^a\rangle + \langle(\frac{k}{t})p^a\rangle} \\ &= (-1)^{r|S_w^c\setminus\{0\}|} \prod_{k\in S_w^c} \overline{\omega}(-1)^{k\frac{q-1}{t}}. \end{aligned}$$

Thus

$$F[w] = (-1)^{r|S_w^c\setminus\{0\}|}\, q^{|S_w^c|}\, \overline{\omega}(-1)^{s|S_w^c|} \prod_{k\in S_w^c} \prod_{a=0}^{r-1} \Gamma_p\big(\big\langle\tfrac{k}{t}p^a\big\rangle\big)^{n_k}. \tag{4-23}$$

If we let

$$-z = \sum_{k\in S_w^c} (n_k - 1) \sum_{a=0}^{r-1} \Big\lfloor \big\langle -\tfrac{t-k}{t}p^a \big\rangle + \tfrac{sp^a}{q-1} \Big\rfloor + \sum_{k\in S_w} \sum_{a=0}^{r-1} \Big\lfloor \big\langle \tfrac{t-k}{t}p^a \big\rangle - \tfrac{sp^a}{q-1} \Big\rfloor + \sum_{\substack{b=0\\ b\not\equiv 0 \ (\mathrm{mod}\ \frac{n}{t})}}^{n-1} \Big\lfloor \big\langle \tfrac{b}{n}p^a \big\rangle - \tfrac{sp^a}{q-1} \Big\rfloor,$$

then, after a lengthy but straightforward calculation, we find that

$$v - z = -r\,|S_w^c \setminus \{0\}| + \sum_{i=1}^{n} \sum_{a=0}^{r-1} \big\langle \tfrac{w_i}{t} p^a \big\rangle. \tag{4-24}$$

Accounting for (4-23) and (4-24) in (4-22), and then (4-21), yields the result. □

## 5. Concluding remarks

When $d \mid q-1$ it is possible express the results of Koblitz, and those in this paper, in terms of hypergeometric functions over finite fields, as defined in [Greene 1987], or using a normalized version defined in [McCarthy 2012]. For example, see [Goodson 2017a; McCarthy 2017; Nakagawa 2021] for related results. To extend these results beyond $q \equiv 1 \pmod{d}$ it is necessary to move to the $p$-adic setting as we have done in this paper. Other results where the $p$-adic hypergeometric function, ${}_mG_m$, is used to count points on certain hypersurfaces, which are special cases of the results in this paper, can be found in [Barman et al. 2016; Goodson 2017b].

## References

[Barman et al. 2016] R. Barman, H. Rahman, and N. Saikia, "Counting points on Dwork hypersurfaces and $p$-adic hypergeometric functions", *Bull. Aust. Math. Soc.* **94**:2 (2016), 208–216. MR Zbl

[Berndt et al. 1998] B. C. Berndt, R. J. Evans, and K. S. Williams, *Gauss and Jacobi sums*, Wiley, New York, 1998. MR Zbl

[Davenport and Hasse 1935] H. Davenport and H. Hasse, "Die Nullstellen der Kongruenzzetafunktionen in gewissen zyklischen Fällen", *J. Reine Angew. Math.* **172** (1935), 151–182. MR Zbl

[Delsarte 1951] J. Delsarte, "Nombre de solutions des équations polynomiales sur un corps fini", exposé 39 in *Séminaire Bourbaki*, 1948/1951, W. A. Benjamin, 1951. Reprinted as pp. 321–329 in *Séminaire Bourbaki* **1**, Soc. Math. France, Paris, 1995. MR Zbl

[Furtado Gomide 1949] E. Furtado Gomide, "On the theorem of Artin–Weil", *Bol. Soc. Mat. São Paulo* **4** (1949), 1–18. MR Zbl

[Goodson 2017a] H. Goodson, "A complete hypergeometric point count formula for Dwork hypersurfaces", *J. Number Theory* **179** (2017), 142–171. MR Zbl

[Goodson 2017b] H. Goodson, "Hypergeometric functions and relations to Dwork hypersurfaces", *Int. J. Number Theory* **13**:2 (2017), 439–485. MR Zbl

[Greene 1987] J. Greene, "Hypergeometric functions over finite fields", *Trans. Amer. Math. Soc.* **301**:1 (1987), 77–101. MR Zbl

[Gross and Koblitz 1979] B. H. Gross and N. Koblitz, "Gauss sums and the $p$-adic $\Gamma$-function", *Ann. of Math.* (2) **109**:3 (1979), 569–581. MR Zbl

[Hardy and Littlewood 1922] G. H. Hardy and J. E. Littlewood, "Some problems of "Partitio Numerorum", IV: the singular series in Waring's Problem and the value of the number $G(k)$", *Math. Z.* **12**:1 (1922), 161–188. MR Zbl

[Ireland and Rosen 1990] K. Ireland and M. Rosen, *A classical introduction to modern number theory*, 2nd ed., Graduate Texts in Mathematics **84**, Springer, 1990. MR Zbl

[Koblitz 1980] N. Koblitz, *p-adic analysis: a short course on recent work*, London Math. Soc. Lecture Note Series **46**, Cambridge University Press, 1980. MR Zbl

[Koblitz 1983] N. Koblitz, "The number of points on certain families of hypersurfaces over finite fields", *Compositio Math.* **48**:1 (1983), 3–23. MR Zbl

[McCarthy 2012] D. McCarthy, "Transformations of well-poised hypergeometric functions over finite fields", *Finite Fields Appl.* **18**:6 (2012), 1133–1147. MR Zbl

[McCarthy 2013] D. McCarthy, "The trace of Frobenius of elliptic curves and the $p$-adic gamma function", *Pacific J. Math.* **261**:1 (2013), 219–236. MR Zbl

[McCarthy 2017] D. McCarthy, "The number of $\mathbb{F}_p$-points on Dwork hypersurfaces and hypergeometric functions", *Res. Math. Sci.* **4** (2017), art. id. 4. MR Zbl

[Nakagawa 2021] A. Nakagawa, "Artin $L$-functions of diagonal hypersurfaces and generalized hypergeometric functions over finite fields", 2021. arXiv 2111.15054

[Sulakashna and Barman 2022] Sulakashna and R. Barman, "Number of $\mathbb{F}_q$-points on diagonal hypersurfaces and hypergeometric function", 2022. arXiv 2210.11732

[Weil 1949] A. Weil, "Numbers of solutions of equations in finite fields", *Bull. Amer. Math. Soc.* **55** (1949), 497–508. MR Zbl

Received December 9, 2023.

DERMOT MCCARTHY
DEPARTMENT OF MATHEMATICS & STATISTICS
TEXAS TECH UNIVERSITY
LUBBOCK, TX
UNITED STATES

dermot.mccarthy@ttu.edu

PACIFIC JOURNAL OF MATHEMATICS
Vol. 328, No. 2, 2024
https://doi.org/10.2140/pjm.2024.328.361

# DEFORMATION OF PAIRS AND SEMIREGULARITY

TAKEO NISHINOU

**We study relative deformation of a map into a Kähler manifold whose image is a divisor. We show that if the map satisfies a condition called semiregularity, then it allows relative deformations if and only if the cycle class of the image remains Hodge in the family. This gives a refinement of the so-called variational Hodge conjecture. We also show that the semiregularity of maps is related to classical notions such as Cayley–Bacharach conditions and d-semistability.**

## 1. Introduction

Let $\pi : \mathfrak{X} \to D$ be a deformation of a compact Kähler manifold $X_0$ of dimension $n \geq 2$ over a disk $D$ in the complex plane. Let $C_0$ be a compact reduced curve (when $n = 2$) or a compact smooth complex manifold of dimension $n-1$ (when $n > 2$). Let $\varphi_0 : C_0 \to X_0$ be a map which is an immersion, that is, for any $p \in C_0$, there is an open neighborhood $p \in V_p \subset C_0$ such that $\varphi_0|_{V_p}$ is an embedding. Then, the image of $\varphi_0$ determines an integral cohomology class $[\varphi_0(C_0)]$ of type $(1, 1)$, that is, a Hodge class which is the Poincaré dual of the cycle $\varphi_0(C_0)$. Note that the class $[\varphi_0(C_0)]$ naturally determines an integral cohomology class of each fiber of $\pi$. Therefore, it makes sense to ask whether this class remains Hodge in these fibers or not. Clearly, the condition that the class $[\varphi_0(C_0)]$ remains Hodge is necessary for the existence of deformations of the map $\varphi_0$ to other fibers.

The notion of *semiregularity* plays a role in this context, though it was introduced [21] and developed [16] originally for submanifolds of codimension one in a *fixed* complex manifold. The main result of these studies is that if a submanifold of codimension one is semiregular, then the obstruction to deforming it in the ambient manifold vanishes. Bloch [5] generalized the notion of semiregularity to subvarieties of any codimension in a projective manifold which are local complete intersection. He generalized the results of [16; 21] to this case, and also related the notion of semiregularity to deformation of pairs. Namely, he proved that if $C_0$ is a subvariety of a projective manifold $X_0$ which is local complete intersection and semiregular, and if the class $[C_0]$ remains Hodge in an algebraic family $\mathfrak{X} \to \mathbb{C}$

*MSC2020:* 14C30, 32G10.
*Keywords:* deformation theory, Hodge theory, semiregularity.

whose central fiber is $X_0$, then there is a deformation of $C_0$ relative to the base. In other words, a local complete intersection subvariety which is semiregular satisfies the *variational Hodge conjecture*. More precisely, the variational Hodge conjecture asks the existence of a family of cycles of the class $[C_0]$ which need not restrict to $C_0$ on the central fiber. Therefore, Bloch's theorem in fact shows that the semiregularity gives a result stronger than the variational Hodge conjecture. We also note that Ran [20] generalized Bloch's result to cases where a weaker version of semiregularity holds. However, although Bloch's and Ran's theorems guarantee the existence of a relative deformation of a cycle on the central fiber $X_0$, it gives little control of the geometry of the deformed cycle.

More recently, the notion of semiregularity has been generalized to maps between varieties [6; 11]. In [11], maps between compact Kähler manifolds were investigated, and it was shown that if the map is semiregular, then it deforms in a fixed target manifold. In [6], the notion of semiregularity was generalized to a very broad context using cotangent complexes, and many known results were generalized. The case of maps was also considered (see [6, Theorem 7.23]), but not in the context of the variational Hodge conjecture as we will do. See also [2; 3; 12; 17; 19] for recent developments related to semiregularity.

Our purpose is to show that the semiregularity in fact suffices to control the geometry of the deformed cycles when the cycle is of codimension one, and also that we can extend the result to maps to a family of Kähler manifolds. Recall that $\varphi_0 : C_0 \to X_0$ is an immersion where $\dim C_0 = \dim X_0 - 1$.

**Theorem 1.** *Assume that the map $\varphi_0$ is semiregular in the sense of Definition 4. If the class $[\varphi_0(C_0)]$ remains Hodge, then the map $\varphi_0$ deforms to other fibers.*

For example, if the image $\varphi_0(C_0)$ has normal crossing singularity, then there is a natural map $\tilde{\varphi}_0 : \tilde{C}_0 \to X_0$, where $\tilde{C}_0$ is the normalization of $C_0$ (when $n > 2$, $C_0 = \tilde{C}_0$). Then, if $\tilde{\varphi}_0$ is semiregular, Theorem 1 implies that it deforms to a general fiber and the singularity of the image remains the same (e.g., it gives a relative equigeneric deformation when $n = 2$).

On the other hand, if the image $\varphi_0(C_0)$ has normal crossing singularity, the semiregularity turns out to be related to some classical notions appeared in different contexts. Namely, we will prove the following (see Corollary 17).

**Theorem 2.** *Assume that the subvariety $\varphi_0(C_0)$ is semiregular in the classical sense. That is, the inclusion of $\varphi_0(C_0)$ into $X_0$ is semiregular in the sense of Definition 4. Then, if the map $H^0(\varphi_0(C_0), \mathcal{N}_\iota) \to H^0(\varphi_0(C_0), \mathcal{S})$ is surjective, the map $\varphi_0$ is semiregular. In particular, if the class $[\varphi_0(C_0)]$ remains Hodge on the fibers of $\mathfrak{X}$, the map $\varphi_0$ can be deformed to general fibers of $\mathfrak{X}$.*

Here, $\mathcal{N}_\iota$ is the normal sheaf of $\varphi_0(C_0)$ in $X_0$ and $\mathcal{S}$ is the *infinitesimal normal sheaf* of the variety $\varphi_0(C_0)$, see Section 6 for the definition. A variety with normal crossing singularity is called *d-semistable* if the infinitesimal normal sheaf is trivial, see [8]. The notion of d-semistability is known to be related to the existence of log-smooth deformations (see [13; 14]). By the above theorem, it turns out that it is also related to deformations of pairs, see Corollary 19.

In the case where $n = 2$, if $X_0$ is a K3 surface, any immersion $\varphi_0$ is semiregular, and Theorem 1 can be applied to any such $\varphi_0$. This result is well known and was proved, for example, using the twistor family associated with the hyperkähler structure of K3 surfaces. Theorem 1 gives a generalization of it to general surfaces. In general, we need to check whether a given map $\varphi_0$ is semiregular or not. For that purpose, Theorem 31 in [18] combined with Theorem 1 above implies the following. Let $\varphi_0 : C_0 \to X_0$ be an immersion such that the image $\varphi_0(C_0)$ is a reduced nodal curve. Let $p : C_0 \to \varphi_0(C_0)$ be the natural map (which is a partial normalization of $\varphi_0(C_0)$) and $P = \{p_i\}$ be the set of nodes of $\varphi_0(C_0)$ whose inverse image by $p$ consists of two points.

**Theorem 3.** *Assume that $\varphi_0(C_0)$ is semiregular in the classical sense and the class $[\varphi_0(C_0)]$ remains Hodge on the fibers of $\mathfrak{X}$. Then, the map $\varphi_0$ deforms to general fibers of $\mathfrak{X}$ if for each $p_i \in P$, there is a first-order deformation of $\varphi_0(C_0)$ which smooths $p_i$, but does not smooth the other nodes of $P$.* □

The condition in Theorem 20 is related (in a sense opposite) to the classical *Cayley–Bacharach condition*, see [4], which requires that if a first-order deformation does not smooth the nodes $P \setminus \{p_i\}$, then it does not smooth $p_i$, either. Using this, we can also deduce a geometric criterion for the existence of deformations of pairs, see Corollary 21.

***Notation.*** We will work in the complex analytic category. Later in the paper, we will study nonconstant maps $\varphi_0 : C_0 \to X_0$ from a variety $C_0$ to a Kähler manifold $X_0$ and their deformations. We denote by $\mathfrak{X}$ a family of compact Kähler manifolds over a disk $D \subset \mathbb{C}$ whose central fiber is $X_0$. A deformation of $\varphi_0$ over $\operatorname{Spec} \mathbb{C}[t]/t^{k+1}$ will be written as $\varphi_k : C_k \to X_k = \mathfrak{X} \times_D \operatorname{Spec} \mathbb{C}[t]/t^{k+1}$. By the image of a map $\varphi_0$ or $\varphi_k$, we mean the analytic locally ringed space with the annihilator structure, see [9, Chapter I, Definition 1.45]. That is, if $U$ is an open subset of $C_k$ with the induced structure of an analytic locally ringed space, and $V$ is an open subset of $X_k$ such that $\varphi_k(U)$ is closed in $V$, we associate the structure sheaf

$$\mathcal{O}_V / \mathcal{A}nn_{\mathcal{O}_V}((\varphi_k)_* \mathcal{O}_U)$$

to the image $\varphi_k(U)$.

## 2. Semiregularity for local embeddings

Let $n$ and $p$ be positive integers with $p < n$. Let $M$ be a complex variety (not necessarily smooth or reduced) of dimension $n - p$ and $X$ a compact Kähler manifold of dimension $n$. Let $\varphi : M \to X$ be a map which is an immersion, that is, for any $p \in M$, there is an open neighborhood $p \in U_p \subset M$ such that $\varphi|_{U_p}$ is an embedding. We assume that the image is a local complete intersection. Then, the normal sheaf $\mathcal{N}_\varphi$ is locally free of rank $p$. Define the locally free sheaves $\mathcal{K}_\varphi$ and $\omega_M$ on $M$ by

$$\mathcal{K}_\varphi = \wedge^p \mathcal{N}_\varphi^\vee \quad \text{and} \quad \omega_M = \mathcal{K}_\varphi^\vee \otimes \varphi^* \mathcal{K}_X,$$

where $\mathcal{K}_X$ is the canonical sheaf of $X$.

When $\varphi$ is an inclusion, the natural inclusion

$$\varepsilon : \mathcal{N}_\varphi^\vee \to \varphi^* \Omega_X^1$$

gives rise to an element

$$\begin{aligned}\wedge^{p-1}\varepsilon \in \operatorname{Hom}_{\mathcal{O}_M}(\wedge^{p-1}\mathcal{N}_\varphi^\vee, \varphi^*\Omega_X^{p-1}) &= \Gamma\big(M, (\varphi^*\Omega_X^{n-p+1})^\vee \otimes \varphi^*\mathcal{K}_X \otimes \mathcal{K}_\varphi^\vee \otimes \mathcal{N}_\varphi^\vee\big) \\ &= \operatorname{Hom}_{\mathcal{O}_X}(\Omega_X^{n-p+1}, \omega_M \otimes \mathcal{N}_\varphi^\vee).\end{aligned}$$

This induces a map on cohomology:

$$\wedge^{p-1}\varepsilon : H^{n-p-1}(X, \Omega_X^{n-p+1}) \to H^{n-p-1}(M, \omega_M \otimes \mathcal{N}_\varphi^\vee).$$

When $\varphi$ is not an inclusion, then $\Gamma\big(M, (\varphi^*\Omega_X^{n-p+1})^\vee \otimes \varphi^*\mathcal{K}_X \otimes \mathcal{K}_\varphi^\vee \otimes \mathcal{N}_\varphi^\vee\big)$ is not necessarily isomorphic to $\operatorname{Hom}_{\mathcal{O}_X}(\Omega_X^{n-p+1}, \omega_M \otimes \mathcal{N}_\varphi^\vee)$, but the map

$$\wedge^{p-1}\varepsilon : H^{n-p-1}(X, \Omega_X^{n-p+1}) \to H^{n-p-1}(M, \omega_M \otimes \mathcal{N}_\varphi^\vee)$$

is still defined.

**Definition 4.** We call $\varphi$ *semiregular* if the natural map $\wedge^{p-1}\varepsilon$ is surjective.

In this paper, we are interested in the case where $p = 1$ and $M$ is reduced when $n = 2$, and $M$ is smooth when $n > 2$. In this case, we have $\omega_M \otimes \mathcal{N}_\varphi^\vee \cong \varphi^*\mathcal{K}_X$ and the map $\wedge^{p-1}\varepsilon$ will be

$$H^{n-2}(X, \mathcal{K}_X) \to H^{n-2}(M, \varphi^*\mathcal{K}_X).$$

**Remark 5.** As we mentioned in the introduction, in [6; 11], Buchweitz–Flenner and Iacono also considered semiregularity of maps between varieties in broader contexts. In the case of maps we consider in this paper, their definitions coincide with ours.

## 3. Local calculation

Let $\pi : \mathfrak{X} \to D$ be a deformation of a compact Kähler manifold $X_0$ of dimension $n \geq 2$. Here, $D$ is a disk on the complex plane centered at the origin. Let

$$\{(U_i, (x_{i,1}, \dots, x_{i,n})\}$$

be a coordinate system of $X_0$. Taking $D$ small enough, the sets

$$\{\mathfrak{U}_i = U_i \times D, (x_{i,1}, \ldots, x_{i,n}, t)\}$$

gives a coordinate system of $\mathfrak{X}$. Precisely, we fix an isomorphism between $\mathfrak{U}_i$ and a suitable open subset of $\mathfrak{X}$ which is compatible with $\pi$ and the inclusion $U_i \to X_0$. Here, $t$ is a coordinate on $D$ pulled back to $\mathfrak{U}_i$. The functions $x_{i,l}$ are also pulled back to $\mathfrak{U}_i$ from $U_i$ by the natural projection.

Take coordinate neighborhoods $\mathfrak{U}_i, \mathfrak{U}_j$ and $\mathfrak{U}_k$. On the intersections of these open subsets, the coordinate functions on one of them can be written in terms of those on another. Namely, on $\mathfrak{U}_i \cap \mathfrak{U}_j$, $x_{i,l}$ can be written as $x_{i,l}(\boldsymbol{x}_j, t)$, here we write

$$\boldsymbol{x}_j = (x_{j,1}, \ldots, x_{j,n}).$$

Similarly, on $\mathfrak{U}_j \cap \mathfrak{U}_k$, we have $x_{j,l} = x_{j,l}(\boldsymbol{x}_k, t)$. Then, on $\mathfrak{U}_i \cap \mathfrak{U}_j \cap \mathfrak{U}_k$, we have

$$x_{i,l} = x_{i,l}(\boldsymbol{x}_k, t) = x_{i,l}(\boldsymbol{x}_j(\boldsymbol{x}_k, t), t).$$

For simplicity we often write $x_{i,l}(\boldsymbol{x}_k, t)$ as $x_{i,l}(\boldsymbol{x}_k)$ and $x_{i,l}(\boldsymbol{x}_j(\boldsymbol{x}_k, t), t)$ as $x_{i,l}(\boldsymbol{x}_j(\boldsymbol{x}_k))$.

Let $X_t = \pi^{-1}(t)$ be the fiber of the family $\pi$ over $t \in D$. Assume that the map

$$\varphi_0 : C_0 \to X_0$$

exists from a variety $C_0$ of dimension $n-1$ to $X_0$, which is an immersion.

We can take an open covering $\{V_i\}$ of $C_0$ such that the restriction of $\varphi_0$ to $V_i$ is an embedding, the image $\varphi_0(V_i)$ is contained in $U_i$ and is defined by an equation $f_{i,0} = 0$ for some holomorphic function $f_{i,0}$. Moreover, we assume that if $V_i \cap V_j$ is nonempty, we have

$$\varphi_0(V_i \cup V_j) \cap (U_i \cap U_j) = \varphi_0(V_i \cap V_j).$$

Let $\operatorname{Spec} \mathbb{C}[t]/t^{m+1}$ be the $m$-th order infinitesimal neighborhood of the origin of $D$. Note that

$$\{U_{i,m} = \mathfrak{U}_i \times_D \operatorname{Spec} \mathbb{C}[t]/t^{m+1}\}$$

gives a covering by coordinate neighborhoods of $X_m = \mathfrak{X} \times_D \operatorname{Spec} \mathbb{C}[t]/t^{m+1}$. We write by $x_{i,l,m}$ the restriction of $x_{i,l}$ to $U_{i,m}$. Let us write

$$\boldsymbol{x}_{i,m} = \{x_{i,1,m}, \ldots, x_{i,n,m}\}.$$

Assume that we have constructed an $m$-th order deformation $\varphi_m : C_m \to X_m$ of $\varphi_0$. Here, $m$ is a nonnegative integer and $C_m$ is an $m$-th order deformation of $C_0$. Let $V_{i,m}$ be the locally ringed space obtained by restricting the structure of a locally ringed space on $C_m$ to $V_i$.

Let $\{f_{i,m}(\boldsymbol{x}_{i,m}, t)\}$ be the set of local defining functions of $\varphi_m(V_{i,m})$ in $U_{i,m}$. We will often write $f_{i,m}(\boldsymbol{x}_{i,m}, t)$ as $f_{i,m}(\boldsymbol{x}_{i,m})$ for notational simplicity. In particular, on the intersection $U_{i,m} \cap U_{j,m}$, there is an invertible function $g_{ij,m}$ which satisfies

$$f_{i,m}(\boldsymbol{x}_{i,m}(\boldsymbol{x}_{j,m}, t), t) = g_{ij,m}(\boldsymbol{x}_{j,m}, t)\, f_{j,m}(\boldsymbol{x}_{j,m}, t) \mod t^{m+1}.$$

Define a holomorphic function $\nu_{ij,m}$ on $U_{i,m} \cap U_{j,m}$ by

$$\begin{aligned} t^{m+1}\nu_{ij,m}(\boldsymbol{x}_{j,m+1}) &= t^{m+1}\nu_{ij,m}(\boldsymbol{x}_{j,0}) \\ &= f_{i,m}(\boldsymbol{x}_{i,m+1}(\boldsymbol{x}_{j,m+1})) - g_{ij,m}(\boldsymbol{x}_{j,m+1})\, f_{j,m}(\boldsymbol{x}_{j,m+1}), \end{aligned}$$

which is an equality over $\mathbb{C}[t]/t^{m+2}$. Note that $\nu_{ij,m}$ can be regarded as a function on $U_i \cap U_j$.

**Proposition 6.** *Assume that the intersection $U_i \cap U_j \cap U_k$ is nonempty. Then, on $U_i \cap U_j \cap U_k \cap \varphi_0(V_i)$, the following identities hold*:

$$\nu_{ik,m}(\boldsymbol{x}_{k,m+1}) = \nu_{ij,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1})) + g_{ij,0}(\boldsymbol{x}_{j,0}(\boldsymbol{x}_{k,0}))\, \nu_{jk,m}(\boldsymbol{x}_{k,m+1})$$

*and*

$$\nu_{ij,m} = -g_{ij,0}\, \nu_{ji,m}.$$

**Remark 7.** The equality

$$U_i \cap U_j \cap U_k \cap \varphi_0(V_i) = U_i \cap U_j \cap U_k \cap \varphi_0(V_j) = U_i \cap U_j \cap U_k \cap \varphi_0(V_k)$$

holds by the way we took $\{V_i\}$.

*Proof.* We have

$$\boldsymbol{x}_{i,m+1}(\boldsymbol{x}_{k,m+1}) \equiv \boldsymbol{x}_{i,m+1}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1})) \mod t^{m+2},$$

on $U_{i,m+1} \cap U_{j,m+1} \cap U_{k,m+1}$. Then,

$$\begin{aligned} &t^{m+1}\nu_{ik,m}(\boldsymbol{x}_{k,m+1}) \\ &= f_{i,m}(\boldsymbol{x}_{i,m+1}(\boldsymbol{x}_{k,m+1})) - g_{ik,m}(\boldsymbol{x}_{k,m+1})\, f_{k,m}(\boldsymbol{x}_{k,m+1}) \\ &= f_{i,m}\big(\boldsymbol{x}_{i,m+1}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1}))\big) - g_{ij,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1}))\, f_{j,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1})) \\ &\qquad + g_{ij,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1}))\, f_{j,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1})) - g_{ik,m}(\boldsymbol{x}_{k,m+1})\, f_{k,m}(\boldsymbol{x}_{k,m+1}) \\ &= t^{m+1}\nu_{ij,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1})) + g_{ij,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1}))(f_{j,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1})) \\ &\qquad - g_{jk,m}(\boldsymbol{x}_{k,m+1})\, f_{k,m}(\boldsymbol{x}_{k,m+1})) \\ &\qquad + g_{ij,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1}))\, g_{jk,m}(\boldsymbol{x}_{k,m+1})\, f_{k,m}(\boldsymbol{x}_{k,m+1}) \\ &\qquad - g_{ik,m}(\boldsymbol{x}_{k,m+1})\, f_{k,m}(\boldsymbol{x}_{k,m+1}) \\ &= t^{m+1}\nu_{ij,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1})) + t^{m+1} g_{ij,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1}))\, \nu_{jk,m}(\boldsymbol{x}_{k,m+1}) \\ &\qquad + (g_{ij,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1}))\, g_{jk,m}(\boldsymbol{x}_{k,m+1}) - g_{ik,m}(\boldsymbol{x}_{k,m+1}))\, f_{k,m}(\boldsymbol{x}_{k,m+1}), \end{aligned}$$

mod $t^{m+2}$. Since

$$(g_{ij,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1}))\, g_{jk,m}(\boldsymbol{x}_{k,m+1}) - g_{ik,m}(\boldsymbol{x}_{k,m+1}))\, f_{k,m}(\boldsymbol{x}_{k,m+1}) \equiv 0 \mod t^{m+1},$$

we have

$$g_{ij,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1}))\, g_{jk,m}(\boldsymbol{x}_{k,m+1}) \equiv g_{ik,m}(\boldsymbol{x}_{k,m+1}) \mod t^{m+1}.$$

Therefore, we have

$$\begin{aligned}
&(g_{ij,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1}))\, g_{jk,m}(\boldsymbol{x}_{k,m+1}) - g_{ik,m}(\boldsymbol{x}_{k,m+1}))\, f_{k,m}(\boldsymbol{x}_{k,m+1})\\
&\equiv (g_{ij,m}(\boldsymbol{x}_{j,m+1}(\boldsymbol{x}_{k,m+1}))\, g_{jk,m}(\boldsymbol{x}_{k,m+1}) - g_{ik,m}(\boldsymbol{x}_{k,m+1}))\, f_{k,0}(\boldsymbol{x}_{k,m+1}) \mod t^{m+2}.
\end{aligned}$$

Since $f_{k,0}(\boldsymbol{x}_k) = 0$ on the image of $\varphi_0|_{V_k}$, we have the first identity. The second identity follows from this by taking $k = i$. □

Note that the pull back $\varphi_0^*|_{V_i \cap V_j} g_{ij,0}$ of the set of functions $\{g_{ij,0}\}$ is the set of transition functions for the normal sheaf of $\varphi_0$. Thus, the proposition shows that the pull back $\{\varphi_0^*|_{V_i \cap V_j} \nu_{ij,m+1}\}$ of the set of functions $\{\nu_{ij,m+1}\}$ behaves as a Čech 1-cocycle with values in the normal sheaf $\mathcal{N}_{\varphi_0}$ of $\varphi_0$. Then, the following is a straightforward generalization of an argument in [16, Section 3], whose proof we omit.

**Lemma 8.** *The cohomology class of the cocycle $\{\varphi_0^*|_{V_i \cap V_j} \nu_{ij,m+1}\}$ represents the obstruction to deforming $\varphi_m$ one step further.* □

The assumption that $\varphi_0 : C_0 \to X_0$ is an immersion and $\dim C_0 = \dim X_0 - 1$ is crucial for this lemma. For simplicity, we will write $\{\varphi_0^*|_{V_i \cap V_j} \nu_{ij,m+1}\}$ as $\{\nu_{ij,m+1}\}$ if no confusion would occur.

## 4. Explicit description of the Kodaira–Spencer class

Let $\pi : \mathfrak{X} \to D$ be a deformation of a compact Kähler manifold $X_0$ as before. We have the exact sequence

$$0 \to \pi^*\Omega_D^1 \to \Omega_{\mathfrak{X}}^1 \to \Omega_{\mathfrak{X}/D}^1 \to 0$$

The Kodaira–Spencer class is, by definition, the corresponding class in

$$\mu \in \operatorname{Ext}^1(\Omega_{\mathfrak{X}/D}^1, \pi^*\Omega_D^1).$$

**Lemma 9.** *The class $\mu$ is represented by the Čech 1-cocycle*

$$\mu_{ij} = \sum_{l=1}^{n} \frac{\partial x_{i,l}(\boldsymbol{x}_j, t)}{\partial t} \partial_{x_{i,l}}\, dt.$$

*Proof.* See [10, Section II.1]. □

From now on, we drop $dt$ from these expressions since it plays no role below. Restricting this to a presentation over $\mathbb{C}[t]/t^{m+1}$, we obtain the Kodaira–Spencer class for the deformation $X_{m+1} := \mathfrak{X} \times_D \operatorname{Spec} \mathbb{C}[t]/t^{m+2}$. We denote this class by $\mu_m$.

Assume that we have constructed an $m$-th order deformation $\varphi_m : C_m \to X_m$ of $\varphi_0$. Let $\mathcal{N}_{m/D}$ be the relative normal sheaf of $\varphi_m$ and

$$p_m : \varphi_m^* \mathcal{T}_{X_m/D} \to \mathcal{N}_{m/D}$$

be the natural map, where $\mathcal{T}_{X_m/D}$ is the relative tangent sheaf of $X_m$. Pulling $\mu_m$ back to $C_m$ and taking the image by $p_m$, we obtain a class $\bar{\mu}_m \in H^1(C_m, \mathcal{N}_{m/D})$.

As before, let $\{f_{i,m}(\boldsymbol{x}_{i,m}, t)\}$ be the set of local defining functions of $\varphi_m(V_{i,m})$ on $U_{i,m}$.

**Lemma 10.** *The class $\bar{\mu}_m$ is represented by the pull back of*

$$\eta_{ij,m} = \sum_{l=1}^{n} \frac{\partial x_{i,l}(\boldsymbol{x}_j, t)}{\partial t} \partial_{x_{i,l}} f_{i,m}(\boldsymbol{x}_i, t)$$

*to $C_m$.*

*Proof.* We check the cocycle condition. Namely, we have

$$\begin{aligned}
&\eta_{ik,m} - \eta_{ij,m} - g_{ij,m}\,\eta_{jk,m} \\
&\quad = \sum_{l=1}^{n} \frac{\partial x_{i,l}(\boldsymbol{x}_k, t)}{\partial t} \partial_{x_{i,l}} f_{i,m}(\boldsymbol{x}_i, t) - \sum_{l=1}^{n} \frac{\partial x_{i,l}(\boldsymbol{x}_j, t)}{\partial t} \partial_{x_{i,l}} f_{i,m}(\boldsymbol{x}_i, t) \\
&\qquad\qquad - g_{ij,m} \sum_{l=1}^{n} \frac{\partial x_{j,l}(\boldsymbol{x}_k, t)}{\partial t} \partial_{x_{j,l}} f_{j,m}(\boldsymbol{x}_i, t) \\
&\quad = \sum_{l=1}^{n} \frac{\partial x_{i,l}(\boldsymbol{x}_k, t)}{\partial t} \partial_{x_{i,l}} f_{i,m}(\boldsymbol{x}_i, t) - \sum_{l=1}^{n} \frac{\partial x_{i,l}(\boldsymbol{x}_j, t)}{\partial t} \partial_{x_{i,l}} f_{i,m}(\boldsymbol{x}_j, t) \\
&\qquad\qquad - g_{ij,m} \sum_{l=1}^{n} \frac{\partial x_{j,l}(\boldsymbol{x}_k, t)}{\partial t} \partial_{x_{j,l}} (g_{ij,m}^{-1} f_{i,m}(\boldsymbol{x}_i(\boldsymbol{x}_j, t), t)) \\
&\quad = (\mu_{ik} - \mu_{ij} - \mu_{jk}) f_{i,m} - g_{ij,m} f_{i,m}(\boldsymbol{x}_i(\boldsymbol{x}_j, t), t) \sum_{l=1}^{n} \frac{\partial x_{j,l}(\boldsymbol{x}_k, t)}{\partial t} \partial_{x_{j,l}} (g_{ij,m}^{-1}).
\end{aligned}$$

Since $\mu_{ik} - \mu_{ij} - \mu_{jk} = 0$ by the cocycle condition, and $f_{i,m}(\boldsymbol{x}_i(\boldsymbol{x}_j, t), t)$ is zero on the image of $\varphi_m$, we see that $\eta_{ik,m} = \eta_{ij,m} + g_{ij,m}\,\eta_{jk,m}$ on $C_m$. Also, note that $g_{ij,m}$ is the transition function of the normal sheaf $\mathcal{N}_{m/D}$. Then, it is clear that $\eta_{ij,m}$ represents the class $\bar{\mu}_m$. □

Recall that a complex analytic cycle of codimension $r$ in a Kähler manifold determines a cohomology class of type $(r, r)$, which is the Poincaré dual of the homology class of the cycle. Let $\zeta_{C_0} \in H^1(X_0, \Omega^1_{X_0/D})$ be the class corresponding to the image of $\varphi_0$. Note that since the family $\mathfrak{X}$ is differential geometrically trivial, the class $\zeta_{C_0}$ determines a cohomology class in $H^2(\mathfrak{X}, \mathbb{C})$. We denote it by $\tilde{\zeta}_{C_0}$. Then, we have:

**Lemma 11.** *When $\varphi_0$ is semiregular, the class $\tilde{\zeta}_{C_0}$ remains Hodge in $X_{m+1}$ if and only if the class $\bar{\mu}_m$ is zero.*

*Proof.* Since we are assuming we have constructed $\varphi_m : C_m \to X_m$, the class $\tilde{\zeta}_{C_0}$ is Hodge on $X_m$. That is,

$$\tilde{\zeta}_{C_0}|_{X_m} \in H^1(X_m, \Omega^1_{X_m/D}).$$

Bloch [5, Proposition 4.2] showed that $\tilde{\zeta}_{C_0}$ remains Hodge on $X_{m+1}$ if and only if the cup product

$$\tilde{\zeta}_{C_0}|_{X_m} \cup \mu_m \in H^2(X_m, \mathcal{O}_{X_m})$$

is zero. This is the same as the claim that the cup product $\tilde{\zeta}_{C_0}|_{X_m} \cup \mu_m \cup \alpha$ is zero for any $\alpha \in H^{2n-2}(X_m, \mathbb{C})$. On the other hand, we have:

**Claim 12.** *The cup product $\tilde{\zeta}_{C_0}|_{X_m} \cup \mu_m \cup \alpha$ is zero for any $\alpha \in H^{2n-2}(X_m, \mathbb{C})$ if and only if the cup product $\bar{\mu}_m \cup \varphi_m^* \alpha$ is zero on $C_m$.*

*Proof of Claim 12.* By definition of $\tilde{\zeta}_{C_0}|_{X_m}$, the class $\tilde{\zeta}_{C_0}|_{X_m} \cup \mu_m \cup \alpha$ is zero if and only if the class $\varphi_m^* \mu_m \cup \varphi_m^* \alpha$ is zero. Note that the cohomology group $H^{2n-2}(X_m, \mathbb{C})$ decomposes as

$$H^{2n-2}(X_m, \mathbb{C}) \cong H^n(X_m, \Omega^{n-2}_{X_m/D}) \oplus H^{n-1}(X_m, \Omega^{n-1}_{X_m/D}) \oplus H^{n-2}(X_m, \mathcal{K}_{X_m/D}),$$

here, $\mathcal{K}_{X_m/D}$ is the relative canonical sheaf. By dimensional reason, the cup product between $\varphi_m^* \mu_m$ and the pull back of the classes in

$$H^n(X_m, \Omega^{n-2}_{X_m/D}) \oplus H^{n-1}(X_m, \Omega^{n-1}_{X_m/D})$$

is zero. Therefore, we can assume that $\alpha$ belongs to $H^{n-2}(X_m, \mathcal{K}_{X_m/D})$, and so the class $\varphi_m^* \alpha$ belongs to $H^{n-2}(C_m, \varphi_m^* \mathcal{K}_{X_m/D})$. On the other hand, $\varphi_m^* \mu_m$ belongs to $H^1(C_m, \varphi_m^* \mathcal{T}_{X_m/D})$ and we have the natural map

$$H^1(C_m, \varphi^* \mathcal{T}_{X_m/D}) \to H^1(C_m, \mathcal{N}_{m/D}).$$

Here, $\bar{\mu}_m$ is the image of $\varphi_m^* \mu_m$ by this map. Recall that the dual of $H^1(C_m, \mathcal{N}_{m/D})$ is given by $H^{n-2}(C_m, \varphi_m^* \mathcal{K}_{X_m/D})$. So, it follows that the cup product $\varphi_m^* \mu_m \cup \varphi_m^* \alpha$ reduces to $\bar{\mu}_m \cup \varphi_m^* \alpha$. This proves the claim. $\square$

It immediately follows that if $\bar{\mu}_m$ is zero, then $\tilde{\zeta}_{C_0}$ remains Hodge in $X_{m+1}$. For the converse, assume that $\tilde{\zeta}_{C_0}$ remains Hodge in $X_{m+1}$. There is a natural map

$$\iota : H^{2n-2}(X_m, \mathbb{C}) \to H^1(C_m, \mathcal{N}_{m/D})^\vee$$

as in the proof of the claim. Namely, for a class $\alpha$ of

$$H^{2n-2}(X_m, \mathbb{C}) = H^n(X_m, \Omega^{n-2}_{X_m/D}) \oplus H^{n-1}(\Omega^{n-1}_{X_m/D}) \oplus H^{n-2}(\Omega^n_{X_m/D})$$

and $\beta \in H^1(C_m, \mathcal{N}_{m/D})$, let $\iota(\alpha)(\beta)$ be the cup product $\beta \cup \varphi_m^* \alpha$ composed with the trace map $H^{n-1}(C_m, \omega_{C_m}) \to \mathbb{C}$. Here, $\omega_{C_m}$ is the dualizing sheaf of $C_m$, see [1]. The restriction of this map to $X_0$ is a surjection by the semiregularity of $\varphi_0$. Since the surjectivity is an open condition, $\iota$ is also a surjection. This shows that $\bar{\mu}_m \cup \varphi_m^* \alpha$ is zero for any $\alpha \in H^{2n-2}(X_m, \mathbb{C})$ is equivalent to the claim that $\bar{\mu}_m$ is zero. □

Thus, when the class $\tilde{\zeta}_{C_0}$ remains Hodge in $X_{m+1}$, we can write $\bar{\mu}_m$ as the coboundary of a Čech 0-cochain with values in $\mathcal{N}_{m/D}$ on $C_m$. We choose one such representative $\{\delta_i\}$ where $\delta_i \in \Gamma(V_{i,m}, \mathcal{N}_{m/D})$ such that

$$\delta_i - g_{ij,m}\,\delta_j = \eta_{ij,m}.$$

Here, $\{\eta_{ij,m}\}$ is a representative of $\bar{\mu}_m$ (see Lemma 10). Also, note that by the exact sequence

$$0 \to \mathcal{O}_{U_{i,m}} \to \mathcal{O}_{U_{i,m}}(\varphi_m(V_{i,m})) \to \mathcal{N}_{m/D}|_{V_{i,m}} \to 0,$$

there is a section $\tilde{\delta}_i$ of $\mathcal{O}_{U_{i,m}}(\varphi_m(V_{i,m}))$ which maps to $\delta_i$. Then, we have a lift of $\eta_{ij,m}$ to an open subset of $X_0$ as follows.

**Lemma 13.** *When the class $\tilde{\zeta}_{C_0}$ remains Hodge in $X_{m+1}$, the section*

$$\tilde{\eta}_{ij,m} = \tilde{\delta}_i(\boldsymbol{x}_i(\boldsymbol{x}_j, t), t) - g_{ij,m}(\boldsymbol{x}_j, t)\,\tilde{\delta}_j(\boldsymbol{x}_j, t)$$

*of $\mathcal{O}_{U_{i,m}\cap U_{j,m}}(\varphi_m(V_{i,m} \cap V_{j,m}))$ coincides with $\eta_{ij,m}$ when restricted to $V_{i,m}$.* □

## 5. Proof of Theorem 1

As we mentioned in the introduction, in [16], it was shown that if $C_0 \subset X_0$ is a submanifold of codimension one that is semiregular, then the obstruction to deforming $C_0$ in $X_0$ vanishes. The point of their proof is to construct a Čech 1-cocycle on $X_0$ with values in the sheaf $\mathcal{O}_X(C_0)$, whose restriction to $C_0$ is the relevant obstruction class. Then, the vanishing of such a class in cohomology is a straightforward consequence of the definition of semiregularity. Thus, it is important to represent the obstruction as a restriction of a Čech cocycle on the *ambient* space. In the case which was studied in [16], the construction of such a cocycle on the ambient space can be done by a direct calculation. In our case of maps where $\varphi_0(C_0)$ may be singular, we need an additional argument which is a variant of that in [18]. Also, we need to take account of the effect of the deformation of the ambient space, but it is covered by Lemma 13. In this section, we unify these arguments and complete the proof of the main theorem.

Recall that the obstruction to deforming $\varphi_m$ is given by a cocycle

$$\{\varphi_0^*|_{V_i \cap V_j} \nu_{ij,m+1}\} \quad \text{on } C_0,$$

where $\nu_{ij,m}$ is defined by

$$t^{m+1}\,\nu_{ij,m}(\boldsymbol{x}_j) = f_{i,m}(\boldsymbol{x}_i(\boldsymbol{x}_j,t),t) - g_{ij,m}(\boldsymbol{x}_j,t)\,f_{j,m}(\boldsymbol{x}_j,t).$$

For the explicit calculation of the obstruction, we eliminate $g_{ij,m}(\boldsymbol{x}_j,t)$ from this expression as follows.

**Lemma 14.** *On $U_{i,m}\cap U_{j,m}$, we have*

$$(*)\quad (m+1)\,t^m\,\frac{\nu_{ij,m}(\boldsymbol{x}_j)}{f_{i,m}(\boldsymbol{x}_i,t)} = \frac{1}{f_{i,m}(\boldsymbol{x}_i,t)}\left(\frac{\partial f_{i,m}(\boldsymbol{x}_i,t)}{\partial t}+\tilde{\delta}_i\right) - \frac{1}{f_{j,m}(\boldsymbol{x}_i,t)}\left(\frac{\partial f_{j,m}(\boldsymbol{x}_j,t)}{\partial t}+\tilde{\delta}_j\right),$$

*modulo functions holomorphic on the image of $\varphi_m|_{V_{i,m}\cap V_{j,m}}$.*

*Proof.* First, by differentiating the equation

$$t^{m+1}\,\nu_{ij,m}(\boldsymbol{x}_j) = f_{i,m}(\boldsymbol{x}_i(\boldsymbol{x}_j,t),t) - g_{ij,m}(\boldsymbol{x}_j,t)\,f_{j,m}(\boldsymbol{x}_j,t)$$

with respect to $t$, we have

$$\begin{aligned}
(m+1)\,t^m\,\nu_{ij,m}(\boldsymbol{x}_j) &= \frac{\partial f_{i,m}(\boldsymbol{x}_i,t)}{\partial t} + \sum_{l=1}^{n}\frac{\partial x_{i,l}(\boldsymbol{x}_j,t)}{\partial t}\frac{\partial f_{i,m}(\boldsymbol{x}_i,t)}{\partial x_{i,l}} \\
&\qquad - g_{ij,m}(\boldsymbol{x}_j,t)\frac{\partial f_{j,m}(\boldsymbol{x}_j,t)}{\partial t} - \frac{\partial g_{ij,m}(\boldsymbol{x}_j,t)}{\partial t}f_{j,m}(\boldsymbol{x}_j,t) \\
&= \frac{\partial f_{i,m}(\boldsymbol{x}_i,t)}{\partial t} - g_{ij,m}(\boldsymbol{x}_j,t)\frac{\partial f_{j,m}(\boldsymbol{x}_j,t)}{\partial t} + \eta_{ij,m} - \frac{\partial g_{ij,m}(\boldsymbol{x}_j,t)}{\partial t}f_{j,m}(\boldsymbol{x}_j,t)
\end{aligned}$$

on $V_{i,m}\cap V_{j,m}$. Since $f_{j,m}$ is zero on the image of $\varphi_m|_{V_{j,m}}$, we can ignore the last term. By the same reason, we can replace $\eta_{ij,m}$ by $\tilde{\eta}_{ij,m}$ introduced in Lemma 13, and we can regard the above equation as an equation on $U_{i,m}\cap U_{j,m}$.

Dividing this by $f_{i,m}(\boldsymbol{x}_i,t)$, we have

$$\begin{aligned}
(*)\quad (m+1)\,t^m\,\frac{\nu_{ij,m}(\boldsymbol{x}_j)}{f_{i,m}(\boldsymbol{x}_i,t)} &= \frac{1}{f_{i,m}(\boldsymbol{x}_i,t)}\frac{\partial f_{i,m}(\boldsymbol{x}_i,t)}{\partial t} - \frac{g_{ij,m}(\boldsymbol{x}_j,t)}{f_{i,m}(\boldsymbol{x}_i,t)}\frac{\partial f_{j,m}(\boldsymbol{x}_j,t)}{\partial t} + \frac{\eta_{ij,m}}{f_{i,m}(\boldsymbol{x}_i,t)} \\
&= \frac{1}{f_{i,m}(\boldsymbol{x}_i,t)}\frac{\partial f_{i,m}(\boldsymbol{x}_i,t)}{\partial t} - \frac{g_{ij,m}(\boldsymbol{x}_j,t)}{f_{i,m}(\boldsymbol{x}_i,t)}\frac{\partial f_{j,m}(\boldsymbol{x}_j,t)}{\partial t} + \frac{\tilde{\delta}_i}{f_{i,m}(\boldsymbol{x}_i,t)} - \frac{g_{ij,m}\tilde{\delta}_j}{f_{i,m}(\boldsymbol{x}_i,t)} \\
&= \frac{1}{f_{i,m}(\boldsymbol{x}_i,t)}\left(\frac{\partial f_{i,m}(\boldsymbol{x}_i,t)}{\partial t}+\tilde{\delta}_i\right) - \frac{1}{f_{j,m}(\boldsymbol{x}_i,t)}\left(\frac{\partial f_{j,m}(\boldsymbol{x}_j,t)}{\partial t}+\tilde{\delta}_j\right)
\end{aligned}$$

modulo functions holomorphic on $C_m$. Note that this is an equation over $\mathbb{C}[t]/t^{m+1}$, and so we have

$$\frac{g_{ij,m}(\boldsymbol{x}_j, t)\, f_{j,m}(\boldsymbol{x}_j, t)}{f_{i,m}(\boldsymbol{x}_i, t)} = 1. \qquad \square$$

Let

$$\left[\frac{1}{f_{i,m}(\boldsymbol{x}_i, t)}\left(\frac{\partial f_{i,m}(\boldsymbol{x}_i, t)}{\partial t} + \tilde{\delta}_i\right)\right]_m$$

be the coefficient of $t^m$ in

$$\frac{1}{f_{i,m}(\boldsymbol{x}_i, t)}\left(\frac{\partial f_{i,m}(\boldsymbol{x}_i, t)}{\partial t} + \tilde{\delta}_i\right).$$

Note that the above equation still holds when we replace

$$\frac{1}{f_{i,m}(\boldsymbol{x}_i, t)}\left(\frac{\partial f_{i,m}(\boldsymbol{x}_i, t)}{\partial t} + \tilde{\delta}_i\right) \quad \text{and} \quad \frac{1}{f_{j,m}(\boldsymbol{x}_i, t)}\left(\frac{\partial f_{j,m}(\boldsymbol{x}_j, t)}{\partial t} + \tilde{\delta}_j\right)$$

by

$$\left[\frac{1}{f_{i,m}(\boldsymbol{x}_i, t)}\left(\frac{\partial f_{i,m}(\boldsymbol{x}_i, t)}{\partial t} + \tilde{\delta}_i\right)\right]_m \quad \text{and} \quad \left[\frac{1}{f_{j,m}(\boldsymbol{x}_i, t)}\left(\frac{\partial f_{j,m}(\boldsymbol{x}_j, t)}{\partial t} + \tilde{\delta}_j\right)\right]_m,$$

respectively. Also, we can think of

$$\left[\frac{1}{f_{i,m}(\boldsymbol{x}_i, t)}\left(\frac{\partial f_{i,m}(\boldsymbol{x}_i, t)}{\partial t} + \tilde{\delta}_i\right)\right]_m$$

as a function on $U_i$ by forgetting $t^m$.

Now, introduce any Riemannian metric on $X_0$. Recall that we fixed an open covering $\{V_i\}$ of $C_0$. If $V_i$ does not contain a singular point of $C_0$, we write $\mathring{V}_i = V_i$. If $V_i$ contains a singular point of $C_0$, we write by $\mathring{V}_i$ the complement of a small closed disk around the singular point in $V_i$. For each $\mathring{V}_i$, let $N_{\varphi_0}|_{\mathring{V}_i}$ be the normal bundle of $\varphi_0$ restricted to $\mathring{V}_i$. Let $S_\delta|_{\mathring{V}_i}$ be the circle bundle of radius $\delta$ in $N_{\varphi_0}|_{\mathring{V}_i}$. Here, $\delta$ is a small positive real number. If $\delta$ is small enough, the exponential map gives an embedding of $S_\delta|_{\mathring{V}_i}$ into a small neighborhood of the image $\varphi_0(\mathring{V}_i)$. We can assume that the image of $S_\delta|_{\mathring{V}_i}$ is disjoint from $\varphi_0(V_i)$ even if $V_i$ contains a singular point of $C_0$. Note that the bundles $S_\delta|_{\mathring{V}_i}$ on each $\mathring{V}_i$ glue and give a circle bundle $S_\delta$ on the open subset $\bigcup_i \mathring{V}_i$ of $C_0$. When $n \geq 3$, this is actually a bundle over $C_0$.

Note that the obstruction class

$$[\varphi_0^*|_{V_i \cap V_j} \nu_{ij,m+1}] \in H^1(C_0, \mathcal{N}_{\varphi_0})$$

is zero if and only if the pairing of it with any class in $H^{n-2}(C_0, \varphi^*\mathcal{K}_{X_0})$ is zero. By semiregularity, any class in $H^{n-2}(C_0, \varphi^*\mathcal{K}_{X_0})$ is a restriction of an element of $H^{n-2}(X_0, \mathcal{K}_{X_0})$. Let $\Theta$ be any closed $C^\infty$ $(2n-2)$-form on $X_0$. In particular, $\Theta$ represents a class in

$$H^{2n-2}(X_0, \mathbb{C}) = H^{n-2}(X_0, \mathcal{K}_{X_0}) \oplus H^{n-1}(X_0, \Omega^{n-1}_{X_0}) \oplus H^n(X_0, \Omega^{n-2}_{X_0}).$$

Here, $\Omega^i_{X_0}$ is the sheaf of holomorphic $i$-forms on $X_0$. Integrating the restriction of the singular $(2n-2)$-form

$$\left[\frac{1}{f_{i,m}(\boldsymbol{x}_i, t)}\left(\frac{\partial f_{i,m}(\boldsymbol{x}_i, t)}{\partial t} + \tilde{\delta}_i\right)\right]_m \Theta$$

to the circle bundle along the fibers, we obtain a closed $(2n-3)$-forms $\gamma_i$ on $\mathring{V}_i$. Then, we have:

**Lemma 15.** *On $\mathring{V}_i \cap \mathring{V}_j$, the limit $\lim_{\delta \to 0} \gamma_i - \gamma_j$ exists, and is $m+1$ times the fiberwise pairing between $\varphi_0^*|_{V_i \cap V_j} \nu_{ij,m}(\boldsymbol{x}_j)$ and $\varphi_0^*\Theta$.*

*Proof.* Note that $\nu_{ij,m}(\boldsymbol{x}_j)$ is a local section of the normal sheaf $\mathcal{N}_{\varphi_0}$ of $\varphi_0 : C_0 \to X_0$. Thus, it naturally pairs with the pull back of $\Theta$ and gives a $(2n-3)$-form on $\mathring{V}_i \cap \mathring{V}_j$. Then, the claim is a consequence of the equation $(*)$ and standard estimates of integrations.

Now, if $C_0$ is nonsingular (in particular if $n \geq 3$), let $\mathcal{C}^{2n-3}(C_0)$ be the sheaf of smooth closed $(2n-3)$-forms on $C_0$. It has a resolution

$$0 \to \mathcal{C}^{2n-3}(C_0) \to \mathcal{A}^{2n-3}(C_0) \to \mathcal{A}^{2n-2}(C_0) \to 0$$

by flabby sheaves. Here, $\mathcal{A}^i(C_0)$ is the sheaf of complex valued smooth $i$-forms on $C_0$. Thus, the cohomology group $H^1(C_0, \mathcal{C}^{2n-3})$ is naturally isomorphic to $H^{2n-2}(C_0, \mathbb{C}) \cong H^{n-1}(C_0, \mathcal{K}_{C_0})$.

By Lemma 15, as the radius $\delta$ goes to zero, the Čech 1-cocycle $\{\gamma_{ij}\}$ with values in closed $(2n-3)$-forms obtained as the differences of $\{\gamma_i\}$ converges to the obstruction class $[\nu_{ij,m}]$ paired with the pull back of $\Theta$ by $\varphi_0$, considered as a class in $H^1(C_0, \mathcal{C}^{2n-3})$. However, by the above isomorphism between $H^1(C_0, \mathcal{C}^{2n-3})$ and $H^{n-1}(C_0, \mathcal{K}_{C_0})$, this class is the same as the obstruction class paired with $\varphi_0^*\Theta$. Thus, the obstruction to deforming $\varphi_m$ vanishes if and only if the limit class in Lemma 15 vanishes for any $\Theta \in H^{n-2}(X_0, \mathcal{K}_{X_0})$.

If $C_0$ is nonsingular, $\{\gamma_i\}$ is defined on a genuine open covering of $C_0$. Thus, the Čech cocycle $\{\gamma_{ij}\}$ vanishes for all $\delta$. So, the limit also vanishes. This finishes the proof of Theorem 1 for $C_0$ nonsingular.

When $n = 2$ and $C_0$ is singular, $\bigcup_i \mathring{V}_i$ covers only an open subset of $C_0$. However, one can show that the Čech 1-cocycle defined by $\gamma_{ij} = \gamma_i - \gamma_j$ still does not depend on the radius $\delta$, and $\lim_{\delta \to 0} \gamma_i - \gamma_j$ gives the obstruction class paired with $\varphi_0^*\Theta$. Thus, it suffices to prove the vanishing of the class $[\gamma_{ij}]$ for a small $\delta$. This can be reduced to an application of the Stokes theorem. See [18] for full details. $\square$

## 6. Criterion for semiregularity

In this section, we give necessary conditions for a map $\varphi_0 : C_0 \to X_0$ to be semiregular. It turns out that some classical notions which appeared in different contexts such as Cayley–Bacharach condition and d-semistability are related to relative deformations of maps.

***The case n > 2.*** First, we consider the case $n > 2$. Let $\pi : \mathfrak{X} \to D$ be a family of $n$-dimensional Kähler manifolds. Let $\varphi_0 : C_0 \to X_0$ be a map from a compact smooth complex manifold of dimension $n-1$ which is an immersion. We also assume that the image $\varphi_0(C_0)$ has normal crossing singularity.

Consider the exact sequence on $\varphi_0(C_0)$ given by

$$0 \to \iota^* \mathcal{K}_{X_0} \to p_* \varphi_0^* \mathcal{K}_{X_0} \to \mathcal{Q} \to 0,$$

where $\iota : \varphi_0(C_0) \to X_0$ is the inclusion, and $p : C_0 \to \varphi_0(C_0)$ is the normalization. The sheaf $\mathcal{Q}$ is defined by this sequence. It is supported on the singular locus $\operatorname{sing}(\varphi_0(C_0))$ of $\varphi_0(C_0)$. We have an associated exact sequence of cohomology groups

$$\begin{aligned} \cdots \to H^{n-2}(\varphi_0(C_0), \iota^* \mathcal{K}_{X_0}) &\to H^{n-2}(\varphi_0(C_0), p_* \varphi_0^* \mathcal{K}_{X_0}) \\ &\to H^{n-2}(\varphi_0(C_0), \mathcal{Q}) \\ &\to H^{n-1}(\varphi_0(C_0), \iota^* \mathcal{K}_{X_0}) \\ &\to H^{n-1}(\varphi_0(C_0), p_* \varphi_0^* \mathcal{K}_{X_0}) \\ &\to H^{n-1}(\varphi_0(C_0), \mathcal{Q}). \end{aligned} \tag{1}$$

By dimensional reason, we have $H^{n-1}(\varphi_0(C_0), \mathcal{Q}) = 0$. Also, note that

$$H^i(\varphi_0(C_0), p_* \varphi_0^* \mathcal{K}_{X_0}) \cong H^i(C_0, \varphi_0^* \mathcal{K}_{X_0})$$

for $i = n-2, n-1$, by the Leray spectral sequence. Therefore, if $\varphi_0(C_0)$ is semiregular in the classical sense, that is, the natural map

$$H^{n-2}(X_0, \mathcal{K}_{X_0}) \to H^{n-2}(\varphi_0(C_0), \iota^* \mathcal{K}_{X_0})$$

is surjective, then the map $\varphi_0$ is semiregular if and only if the map

$$H^{n-2}(\varphi_0(C_0), \iota^* \mathcal{K}_{X_0}) \to H^{n-2}(\varphi_0(C_0), p_* \varphi_0^* \mathcal{K}_{X_0})$$

is surjective.

**Corollary 16.** *Assume that $\varphi_0(C_0)$ is semiregular in the classical sense and the class $[\varphi_0(C_0)]$ remains Hodge on the fibers of $\mathfrak{X}$. Then, if the map*

$$H^{n-2}(\varphi_0(C_0), \iota^* \mathcal{K}_{X_0}) \to H^{n-2}(\varphi_0(C_0), p_* \varphi_0^* \mathcal{K}_{X_0})$$

*is surjective, $\varphi_0$ can be deformed to general fibers of $\mathfrak{X}$.* □

On the other hand, consider the exact sequence

$$0 \to p_* \mathcal{N}_{\varphi_0} \to \mathcal{N}_\iota \to \mathcal{S} \to 0,$$

of sheaves on $\varphi_0(C_0)$, where $\mathcal{S}$ is defined by this sequence. The associated exact sequence of cohomology groups is

$$\begin{aligned} 0 \to H^0(\varphi_0(C_0), p_* \mathcal{N}_{\varphi_0}) &\to H^0(\varphi_0(C_0), \mathcal{N}_\iota) \\ &\to H^0(\varphi_0(C_0), \mathcal{S}) \\ &\to H^1(\varphi_0(C_0), p_* \mathcal{N}_{\varphi_0}) \\ &\to H^1(\varphi_0(C_0), \mathcal{N}_\iota) \to \cdots \end{aligned} \tag{2}$$

We have

$$H^i(\varphi_0(C_0), p_* \mathcal{N}_{\varphi_0}) \cong H^i(C_0, \mathcal{N}_{\varphi_0})$$

again by the Leray spectral sequence. Note that the group $H^i(C_0, \mathcal{N}_{\varphi_0})$ is isomorphic to the dual of $H^{n-1-i}(C_0, \varphi_0^* \mathcal{K}_{X_0}),\ i = 0, 1$. Similarly, the group $H^i(\varphi_0(C_0), \mathcal{N}_\iota)$ is isomorphic to the dual of $H^{n-1-i}(\varphi_0(C_0), \iota^* \mathcal{K}_{X_0}),\ i = 0, 1$.

Comparing the dual of the cohomology exact sequence (1) with (2), we obtain $H^{n-2}(\varphi_0(C_0), \mathcal{Q})^\vee \cong H^0(\varphi_0(C_0), \mathcal{S})$. In particular, we can restate Corollary 16 as follows.

**Corollary 17.** *Assume that $\varphi_0(C_0)$ is semiregular in the classical sense and the class $[\varphi_0(C_0)]$ remains Hodge on the fibers of $\mathfrak{X}$. Then, if the map*

$$H^0(\varphi_0(C_0), \mathcal{N}_\iota) \to H^0(\varphi_0(C_0), \mathcal{S})$$

*is surjective, $\varphi_0$ can be deformed to general fibers of $\mathfrak{X}$.* □

The sheaf $\mathcal{S}$ is the *infinitesimal normal sheaf* of the singular locus of $\varphi_0(C_0)$, as we will see below. Recall that we assume that the image $\varphi_0(C_0)$ has normal crossing singularity. Then, for any point $p \in \varphi_0(C_0)$, we can take a coordinate system $(z_1, \ldots, z_n)$ on a neighborhood $U$ of $p$ in $X_0$ so that $U \cap \varphi_0(C_0)$ is given by $x_1 \cdots x_k = 0,\ 1 \le k \le n$. Let $\mathcal{I}_j$ be the ideal of $\mathcal{O}_U$ generated by $x_j$ and let $\mathcal{I}$ be the ideal defining $\varphi_0(C_0) \cap U$ in $U$. Then,

$$\mathcal{I}_1/\mathcal{I}_1\mathcal{I} \otimes \cdots \otimes \mathcal{I}_k/\mathcal{I}_k\mathcal{I}$$

gives an invertible sheaf on the singular locus of $\varphi_0(C_0) \cap U$. Globalizing this construction, we obtain an invertible sheaf on the singular locus of $\varphi_0(C_0)$. Then, the dual invertible sheaf of this is called the infinitesimal normal sheaf of the singular locus of $\varphi_0(C_0)$, see [8]. We note that the infinitesimal normal sheaf is canonically isomorphic to the sheaf (see [8, Proposition 2.3])

$$\mathrm{Ext}^1_{\mathcal{O}_{\varphi_0(C_0)}}(\Omega^1_{\varphi_0(C_0)}, \mathcal{O}_{\varphi_0(C_0)}).$$

**Lemma 18.** *The sheaf $\mathcal{S}$ is isomorphic to the infinitesimal normal sheaf.*

*Proof.* Note that the sheaf $\mathcal{I}_1/\mathcal{I}_1\mathcal{I} \otimes \cdots \otimes \mathcal{I}_k/\mathcal{I}_k\mathcal{I}$ is generated by the element $x_1 \otimes \cdots \otimes x_k$. The sheaf $p_*\mathcal{N}_{\varphi_0}$ is given by

$$\bigoplus_{i=1}^{k} \mathrm{Hom}(\mathcal{I}_i/\mathcal{I}_i^2, \mathcal{O}_U) \quad \text{on } U.$$

The sheaf $\mathcal{N}_\iota$ is given by $\mathrm{Hom}(\mathcal{I}/\mathcal{I}^2, \mathcal{O}_U)$. The sheaf $\mathcal{N}_\iota$ is an invertible sheaf and generated by the morphism which maps $x_1 \cdots x_k$ to $1 \in \mathcal{O}_U$. In particular, by multiplying any $x_1 \cdots \check{x}_i \cdots x_k$, the generator is mapped into the image of $p_*\mathcal{N}_{\varphi_0} \to \mathcal{N}_\iota$, namely, the image of the generator of $\mathrm{Hom}(\mathcal{I}_i/\mathcal{I}_i^2, \mathcal{O}_U)$. Also, note that the ideal of the singular locus of $\varphi_0(C_0)$ is generated by $x_1 \cdots \check{x}_i \cdots x_k$, $i = 1, \dots, k$. From these, it is easy to see that the cokernel of the map $p_*\mathcal{N}_{\varphi_0} \to \mathcal{N}_\iota$ is isomorphic to the dual of $\mathcal{I}_1/\mathcal{I}_1\mathcal{I} \otimes \cdots \otimes \mathcal{I}_k/\mathcal{I}_k\mathcal{I}$. □

Recall that the infinitesimal normal sheaf is related to deformations of $\varphi_0(C_0)$ which smooth the singular locus, see [8]. In particular, $\varphi_0(C_0)$ is called *d-semistable* if the infinitesimal normal sheaf is trivial, and d-semistable variety carries a log structure log smooth over a standard log point, so that one can study its deformations via log smooth deformation theory [13; 14; 15].

By Corollary 17, the infinitesimal normal sheaf plays a crucial in the deformation theory even if it is not d-semistable.

On the other hand, the notion of d-semistability gives a sufficient condition for the existence of deformations in this situation, too, as follows.

**Corollary 19.** *Let the image $\varphi_0(C_0)$ be very ample and $H^1\big(X_0, \mathcal{O}_{X_0}(\varphi_0(C_0))\big) = 0$. Let $\varphi_0(C_0)$ be d-semistable and the singular locus of $\varphi_0(C_0)$ is connected. Then, the map $\varphi_0$ is semiregular.*

*Proof.* First, we note that the subvariety $\varphi_0(C_0)$ is semiregular in the classical sense. Namely, consider the cohomology exact sequence

$$\cdots \to H^1\big(X_0, \mathcal{O}_{X_0}(\varphi_0(C_0))\big) \to H^1(\varphi_0(C_0), \mathcal{N}_\iota) \to H^2(X_0, \mathcal{O}_{X_0}) \to \cdots,$$

here $\iota : \varphi_0(C_0) \to X_0$ is the inclusion and $\mathcal{N}_\iota$ is the normal sheaf of it. When $H^1\big(X_0, \mathcal{O}_{X_0}(\varphi_0(C_0))\big) = 0$, the map $H^1(\varphi_0(C_0), \mathcal{N}_\iota) \to H^2(X_0, \mathcal{O}_{X_0})$ is injective. Since this map is the dual of the semiregularity map

$$H^{n-2}(X_0, \mathcal{K}_{X_0}) \to H^{n-2}(\varphi_0(C_0), \iota^*\mathcal{K}_{X_0}),$$

it follows that $\varphi_0(C_0)$ is semiregular.

To prove that $\varphi_0$ is semiregular, it suffices to show the map

$$H^0(\varphi_0(C_0), \mathcal{N}_\iota) \to H^0(\varphi_0(C_0), \mathcal{S})$$

is surjective. When $\varphi_0(C_0)$ is d-semistable, the sheaf $\mathcal{S}$ is the trivial line bundle on the singular locus of $\varphi_0(C_0)$. Since we assume that the singular locus is connected, it suffices to show that the map $H^0(\varphi_0(C_0), \mathcal{N}_\iota) \to H^0(\varphi_0(C_0), \mathcal{S})$ is not the zero map. This in turn is equivalent to the claim that the injection

$$H^0(\varphi_0(C_0), p_*\mathcal{N}_{\varphi_0}) \to H^0(\varphi_0(C_0), \mathcal{N}_\iota)$$

is not an isomorphism. Since $\varphi_0(C_0)$ is very ample, there is a section $s$ of $\mathcal{O}_X(\varphi_0(C_0))$ which does not entirely vanish on the singular locus of $\varphi_0(C_0)$. Then, if $\sigma$ is a section of $\mathcal{O}_X(\varphi_0(C_0))$ defining $\varphi_0(C_0)$, the sections $\sigma + \tau s$, where $\tau \in \mathbb{C}$ is a parameter, deforms $\varphi_0(C_0)$, and the nonvanishing of $s$ on the singular locus of $\varphi_0(C_0)$ implies that this smooths a part of the singular locus of $\varphi_0(C_0)$. Since the sections of $H^0(\varphi_0(C_0), p_*\mathcal{N}_{\varphi_0})$ give first-order deformations which does not smooth the singular locus, it follows that the map

$$H^0(\varphi_0(C_0), p_*\mathcal{N}_{\varphi_0}) \to H^0(\varphi_0(C_0), \mathcal{N}_\iota)$$

is not an isomorphism. This proves the claim. □

***The case n = 2.*** Now, let us consider the case $n = 2$. Although we can work in a more general situation, we assume $\varphi_0(C_0)$ is a reduced nodal curve for simplicity. However $C_0$ need not be smooth. Let $p : C_0 \to \varphi_0(C_0)$ be the natural map, which is a partial normalization. In this case, we can deduce very explicit criterion for the semiregularity. Again, we have the exact sequence

$$\begin{aligned} 0 \to H^0(\varphi_0(C_0), p_*\mathcal{N}_{\varphi_0}) &\to H^0(\varphi_0(C_0), \mathcal{N}_\iota) \\ &\to H^0(\varphi_0(C_0), \mathcal{S}) \\ &\to H^1(\varphi_0(C_0), p_*\mathcal{N}_{\varphi_0}) \to H^1(\varphi_0(C_0), \mathcal{N}_\iota) \to \cdots, \end{aligned}$$

and if $\varphi_0(C_0)$ is semiregular in the classical sense, then $\varphi_0$ is semiregular if and only if the map $H^0(\varphi_0(C_0), \mathcal{N}_\iota) \to H^0(\varphi_0(C_0), \mathcal{S})$ is surjective. Let $P = \{p_i\}$ be the set of nodes of $\varphi_0(C_0)$ whose inverse image by $p$ consists of two points. Then, the sheaf $\mathcal{S}$ is isomorphic to $\oplus_i \mathbb{C}_{p_i}$, where $\mathbb{C}_{p_i}$ is the skyscraper sheaf at $p_i$. By an argument similar to the one in the previous subsection, we proved Theorem 20 below in [18].

**Theorem 20.** *Assume that $\varphi_0(C_0)$ is semiregular in the classical sense. Then, the map $\varphi_0$ is semiregular if and only if for each $p_i \in P$, there is a first-order deformation of $\varphi_0(C_0)$ which smooths $p_i$, but does not smooth the other nodes of $P$.* □

For applications, it will be convenient to write this in a geometric form. Consider the exact sequence

$$0 \to \mathcal{O}_{X_0} \to \mathcal{O}_{X_0}(\varphi_0(C_0)) \to \mathcal{N}_\iota \to 0$$

of sheaves on $X_0$ and the associated cohomology sequence

$$0 \to H^0(X_0, \mathcal{O}_{X_0}) \to H^0\big(X_0, \mathcal{O}_{X_0}(\varphi_0(C_0))\big) \\ \to H^0(\varphi_0(C_0), \mathcal{N}_\iota) \to H^1(X_0, \mathcal{O}_{X_0}) \to \cdots.$$

Let $V$ be the image of the map $H^0(\varphi(C_0), \mathcal{N}_\iota) \to H^1(X_0, \mathcal{O}_{X_0})$. Since we are working in the analytic category, we have the exact sequence

$$0 \to \mathbb{Z} \to \mathcal{O}_{X_0} \to \mathcal{O}^*_{X_0} \to 0$$

of sheaves on $X$. Let $\overline{V}$ be the image of $V$ in $Pic^0(X_0) = H^1(X_0, \mathcal{O}^*_{X_0})$. In [18], we proved the following.

**Corollary 21.** *In the situation of Theorem 20, the map $\varphi_0$ is unobstructed if for each $p_i \in P$, there is an effective divisor $D$ such that $\mathcal{O}_X(\varphi_0(C_0) - D) \in \overline{V}$ which avoids $p_i$ but passes through all points in $P \setminus \{p_i\}$.*

A particularly nice case is when the map $H^0(\varphi_0(C_0), \mathcal{N}_\iota) \to H^1(X_0, \mathcal{O}_{X_0})$ is surjective. This is the case when $\varphi_0(C_0)$ is sufficiently ample. Then, if for each $p_i \in P$ there is an effective divisor $D$ which is algebraically equivalent to $\varphi_0(C_0)$ which avoids $p_i$ but passes through all points in $P \setminus \{p_i\}$, the map $\varphi_0$ is semiregular. This is, in a sense, the opposite to the classical *Cayley–Bacharach property*, see, for example, [4].

Combined with Theorem 1, we have:

**Corollary 22.** *Assume that $\varphi_0(C_0)$ is reduced, nodal and semiregular in the classical sense and the class $[\varphi_0(C_0)]$ remains Hodge on the fibers of $\mathfrak{X}$. Then, the map $\varphi_0$ deforms to general fibers of $\mathfrak{X}$ if the condition in Theorem 20 or Corollary 21 is satisfied.* $\square$

In the case of $n = 2$, the original exact sequence

$$\begin{aligned} \cdots \to H^0(\varphi_0(C_0), \iota^*\mathcal{K}_{X_0}) &\to H^0(\varphi_0(C_0), p_*\varphi_0^*\mathcal{K}_{X_0}) \\ &\to H^0(\varphi_0(C_0), \mathcal{Q}) \\ &\to H^1(\varphi_0(C_0), \iota^*\mathcal{K}_{X_0}) \\ &\to H^1(\varphi_0(C_0), p_*\varphi_0^*\mathcal{K}_{X_0}) \to H^1(\varphi_0(C_0), \mathcal{Q}) \end{aligned}$$

before taking the dual is sometimes also useful. In this case, if $\varphi_0(C_0)$ is semiregular in the classical sense, then $\varphi_0$ is semiregular if and only if the map

$$H^0(\varphi_0(C_0), \iota^*\mathcal{K}_{X_0}) \to H^0(\varphi_0(C_0), p_*\varphi_0^*\mathcal{K}_{X_0}) \cong H^0(C_0, \varphi_0^*\mathcal{K}_X)$$

is surjective. For example, when the canonical sheaf $\mathcal{K}_{X_0}$ is trivial, then it is clear that this map is surjective and also $\varphi_0(C_0)$ is semiregular in the classical sense. In fact, in this case it is not necessary to assume that the image $\varphi_0(C_0)$ is nodal or reduced, and any immersion $\varphi_0$ from a reduced curve $C_0$ is semiregular. It is known that when $X_0$ is a K3 surface and the image $\varphi_0(C_0)$ is reduced, then the map $\varphi_0$ deforms to general fibers if the class $[\varphi_0(C_0)]$ remains Hodge. This claim is proved using the twistor family associated with the hyperkähler structure of K3 surfaces, see, for example, [7]. Corollary 22 gives a generalization of this fact to general surfaces.

## Acknowledgements

The author would like to thank the referees for informative comments. The author was supported by JSPS KAKENHI grant number 18K03313.

## References

[1] A. Altman and S. Kleiman, *Introduction to Grothendieck duality theory*, Lect. Notes Math. **146**, Springer, Berlin, 1970. Zbl

[2] R. Bandiera, E. Lepri, and M. Manetti, "Semiregularity maps and deformations of modules over Lie algebroids", preprint, 2022. To appear in *Pure Appl. Math. Q.* arXiv 2208.00694

[3] R. Bandiera, E. Lepri, and M. Manetti, "$L_\infty$ liftings of semiregularity maps via Chern–Simons classes", *Adv. Math.* **435** (2023), art. id. 109358. MR Zbl

[4] W. P. Barth, K. Hulek, C. A. M. Peters, and A. V. de Ven, *Compact complex surfaces*, A Series of Modern Surveys in Mathematics **3**, Springer, Berlin, 2004. Zbl

[5] S. Bloch, "Semi-regularity and deRham cohomology", *Invent. Math.* **17** (1972), 51–66. MR Zbl

[6] R.-O. Buchweitz and H. Flenner, "A semiregularity map for modules and applications to deformations", *Compositio Math.* **137**:2 (2003), 135–210. MR Zbl

[7] X. Chen, F. Gounelas, and C. Liedtke, "Curves on K3 surfaces", *Duke Math. J.* **171**:16 (2022), 3283–3362. MR Zbl

[8] R. Friedman, "Global smoothings of varieties with normal crossings", *Ann. of Math.* (2) **118**:1 (1983), 75–114. MR Zbl

[9] G.-M. Greuel, C. Lossen, and E. Shustin, *Introduction to singularities and deformations*, Springer, Berlin, 2007. MR Zbl

[10] P. A. Griffiths, "Periods of integrals on algebraic manifolds, II", *Amer. J. Math.* **90**:3 (1968), 805–865. Zbl

[11] D. Iacono, "A semiregularity map annihilating obstructions to deforming holomorphic maps", *Canad. Math. Bull.* **54**:3 (2011), 472–486. MR Zbl

[12] D. Iacono and M. Manetti, "Semiregularity and obstructions of complete intersections", *Adv. Math.* **235** (2013), 92–125. MR Zbl

[13] K. Kato, "Logarithmic structures of Fontaine–Illusie", pp. 191–224 in *Algebraic analysis, geometry, and number theory* (Baltimore, MD, 1988), Johns Hopkins Univ. Press, Baltimore, MD, 1989. MR Zbl

[14] F. Kato, "Log smooth deformation theory", *Tohoku Math. J.* (2) **48**:3 (1996), 317–354. MR Zbl

[15] Y. Kawamata and Y. Namikawa, "Logarithmic deformations of normal crossing varieties and smoothing of degenerate Calabi–Yau varieties", *Invent. Math.* **118**:3 (1994), 395–409. MR

[16] K. Kodaira and D. C. Spencer, “A theorem of completeness of characteristic systems of complete continuous systems”, *Amer. J. Math.* **81** (1959), 477–500. MR Zbl

[17] E. Lepri and M. Manetti, “Connections and $L_\infty$ liftings of semiregularity maps”, *J. Geom. Phys.* **168** (2021), art. id. 104303. MR Zbl

[18] T. Nishinou, “Obstructions to deforming maps from curves to surfaces”, *J. Math. Soc. Japan* **76**:1 (2024), 51–71. MR Zbl

[19] J. P. Pridham, “Semiregularity as a consequence of Goodwillie’s theorem”, preprint, 2012. arXiv 1208.3111

[20] Z. Ran, “Hodge theory and the Hilbert scheme”, *J. Differential Geom.* **37**:1 (1993), 191–198. MR Zbl

[21] F. Severi, “Sul teorema fondamentale dei sistemi continui di curve sopra una superficie algebrica”, *Ann. Mat. Pura Appl.* (4) **23** (1944), 149–181. MR Zbl

Received October 4, 2022. Revised December 10, 2023.

Takeo Nishinou
Department of Mathematics
Rikkyo University
Tokyo
Japan

nishinou@rikkyo.ac.jp

# Guidelines for Authors

Authors may submit articles at msp.org/pjm/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 BibTeX is preferred but not required. Any bibliographical citation style may be used but tags will be converted to the house format (see a current issue for examples).

Figures, whether prepared electronically or hand-drawn, must be of publication quality. Figures prepared electronically should be submitted in Encapsulated PostScript (EPS) or in a form that can be converted to EPS, such as GnuPlot, Maple or Mathematica. Many drawing tools such as Adobe Illustrator and Aldus FreeHand can produce EPS output. Figures containing bitmaps should be generated at the highest possible resolution. If there is doubt whether a particular figure is in an acceptable format, the authors should check with production by sending an email to pacific@math.berkeley.edu.

Each figure should be captioned and numbered, so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text (“the curve looks like this:”). It is acceptable to submit a manuscript will all figures at the end, if their placement is specified in the text by means of comments such as “Place Figure 1 here”. The same considerations apply to tables, which should be used sparingly.

Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal’s preferred fonts and layout.

Page proofs will be made available to authors (or to the designated corresponding author) at a website in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

PACIFIC JOURNAL OF MATHEMATICS

Volume 328 No. 2 February 2024