Cent. Eur. J. Math. • 11(12) • 2013 • 2150-2175
DOI: 10.2478/s11533-013-0290-0

VERSITA

Central European Journal of **Mathematics**

# The free one-generated left distributive algebra: basics and a simplified proof of the division algorithm

**Review Article**

Richard Laver, Sheila K. Miller[1]*

[1] *Department of Mathematics, City University of New York, New York City College of Technology, 300 Jay Street, Brooklyn, New York 11201, USA*

**Received 12 June 2011; accepted 3 November 2012**

**Abstract:** The left distributive law is the law $a\cdot(b\cdot c)=(a\cdot b)\cdot(a\cdot c)$. Left distributive algebras have been classically used in the study of knots and braids, and more recently free left distributive algebras have been studied in connection with large cardinal axioms in set theory. We provide a survey of results on the free left distributive algebra on one generator, $\mathcal{A}$, and a new, simplified proof of the existence of a normal form for terms in $\mathcal{A}$. Topics included are: the confluence of $\mathcal{A}$, the linearity of the iterated left division ordering $<_{\mathrm{L}}$ of $\mathcal{A}$, the connections of $\mathcal{A}$ to the braid groups, and an extension $\mathcal{P}$ of $\mathcal{A}$ obtained by freely adding a composition operation. This is followed by a simplified proof of the division algorithm for $\mathcal{P}$, which produces a normal form for terms in $\mathcal{A}$ and is a powerful tool in the study of $\mathcal{A}$.

**MSC:** 06-XX, 08-XX, 17-XX, 20F36, 03E55

**Keywords:** Free left distributive algebra • Division algorithm • Normal form • Braid • Word

*Richard Laver passed away on September 19, 2012 after a long illness.*

## 1. Introduction

A left distributive algebra (LD) is a set, $L$, together with one binary operation $\cdot$ such that for all $a,b,c\in L$, $a\cdot(b\cdot c)=(a\cdot b)\cdot(a\cdot c)$, i.e., left translation $l_a(x)=a\cdot x$ is a homomorphism of $L$. Examples of LDs from classical mathematics which are prominent in the study of knot invariants are group conjugation (where $G$ is a group with operation $*$ and $g\cdot h=g*h*g^{-1}$) and the weighted mean (which is also right distributive): for fixed $r\in\mathbb{C}$, $p\cdot q=rp+(1-r)q$.

* *E-mail: smiller@citytech.cuny.edu*

In fact, for the above examples (assuming that $r \neq 1$ in the second example), and more generally for most classical LDs, left translation is an automorphism of the algebra. Such algebras are termed automorphic sets by Brieskorn [3] (where a comprehensive list is given) and racks by Fenn and Rourke [8]. See Joyce [10] for idempotent variants called quandles.

The braid groups act on direct powers of an automorphic set. Namely, for $2 \leq N \leq \infty$, let $B_N$ be the braid group on $N$ strands: $B_N$ is given by generators $\sigma_1, \sigma_2, \ldots, \sigma_i, \ldots$, $i < N$, subject to the conditions $\sigma_i\sigma_j = \sigma_j\sigma_i$ when $|i-j| > 1$ and $\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}$. Given an automorphic set $(L, \cdot)$, then for $2 \leq N \leq \infty$, $B_N$'s action on $L^N$ is given by

$$\sigma_j\langle l_0, \ldots, l_{j-1}, l_j, \ldots, l_i, \ldots\rangle_{i<N} = \langle l_0, \ldots, l_{j-1}l_j, l_{j-1}, \ldots, l_i, \ldots\rangle_{i<N}. \tag{1}$$

This paper is about the free left distributive algebra $\mathcal{A}$ on one generator, $x$. While it is not an automorphic set, $\mathcal{A}$ is closely connected with $B_N$ (the above group action of $B_N$ is a partial action on $\mathcal{A}^N$). There are a number of open questions about $\mathcal{A}$'s structure, see [17] for examples. $\mathcal{A}$ arises naturally in the study of large cardinal embeddings in set theory and is closely related to $B_N$. We present a simplified version of the proof of a division algorithm for (an extension of) $\mathcal{A}$ (first proved in [13] and simplified in [19]), preceded by a survey of definitions and results (with proofs for the reader's convenience) on the basics of $\mathcal{A}$ necessary to understand the proof. See the end of Section 3 for some applications of the division algorithm.

$\mathcal{A}$ is constructed by forming the collection $A$ of all terms in one generator $x$ and one binary operation $\cdot$ and letting $\mathcal{A} = A/\mathrm{LD}$, where, for $u, v \in A$, $u \equiv_{\mathrm{LD}} v$ if and only if $v$ can be obtained from $u$ by a series of substitutions of the form $a \cdot (b \cdot c) \leftrightarrow (a \cdot b) \cdot (a \cdot c)$. Henceforth we will write $ab$ for $a \cdot b$, and we will adopt the convention that

$$a_0 a_1 \cdots a_{n-1} a_n = ((((a_0 a_1) a_2) \cdots a_{n-1}) a_n).$$

A generator of a free LD is not a product of elements, therefore no automorphic set is free and no idempotent LD is free. Indeed, we will see that $\mathcal{A}$ has no idempotent elements (Theorem 3.3). The question arose of finding a representation of $\mathcal{A}$. Laver found such a representation in large cardinal embeddings from set theory (see Section 3). That such embeddings exist is a very strong axiom (it cannot be proved in ZFC, i.e., the usual axioms of mathematics: Zermelo–Fraenkel together with the Axiom of Choice). Subsequently an example in ZFC was found by Dehornoy via a binary operation on a subset of $B_\infty$.

This paper is organized as follows: In Section 2 we give a proof of Dehornoy's theorem that $A$ is confluent and another key fact, namely that $\mathcal{A}$ is linearly ordered by the relation $<_{\mathrm{L}}$, where $a <_{\mathrm{L}} b$ if and only if there exist $b_1, \ldots, b_n$ such that $b = ab_1 \cdots b_n$. In Section 3 the linearity of $<_{\mathrm{L}}$ is proved. The linear ordering has two parts: connectivity ($a \leq_{\mathrm{L}} b$ or $b \leq_{\mathrm{L}} a$) and irreflexivity ($a \nless_{\mathrm{L}} a$). Different proofs of these two parts were found by Laver and Dehornoy. We give Dehornoy's proof of connectivity, mention Laver's proof of irreflexivity, and give Dehornoy's proof of irreflexivity as simplified by Larue.

The division algorithm takes place not in $\mathcal{A}$ but in an extension $\mathcal{P}$ of $\mathcal{A}$ obtained by freely adding a composition operation $\circ$ such that left translation by $\cdot$ is still a homomorphism. In Section 4 we define $\mathcal{P} = P/\Sigma$, the free one-generated algebra satisfying a set of axioms $\Sigma$ which implies the LD law. We show that $\mathcal{P}$ is a conservative extension of $\mathcal{A}$: if $a_0, a_1 \in A$, $a_0 \equiv_{\mathrm{LD}} a_1$ if and only if $a_0 \equiv_\Sigma a_1$.

$\mathcal{A}$ and $\mathcal{P}$ satisfy left cancellation. Dehornoy showed that the braid group partially acts on direct powers of a left cancellative LD (via (1)). Equivalence of two members of $P$ can be naturally stated in terms of this partial group action. We close Section 3 with a definition of Dehornoy's linear ordering of the braid groups.

Sections 5–8 are given to proving the division algorithm, and Section 5 includes an outline of the proof. The idea of the division algorithm is as follows: For every $w \in \mathcal{P}$, the division algorithm produces a (seen to be unique) term representing $w$. The first step of the algorithm provides a $w_1 \in P$ such that $xw_1 \leq_{\mathrm{L}} w$. If $xw_1 = w$ or $x \circ w_1 = w$, the algorithm terminates. Otherwise it gives a $w_2$ such that $xw_1w_2 \leq_{\mathrm{L}} w$ (where $w_1$ is arbitrary but $w_2$ is bounded) and continues in this way until either $xw_1w_2 \cdots w_{n-1}w_n = w$ or $xw_1w_2 \cdots w_{n-1} \circ w_n$.

Simplifications in the proof alluded to above come from knowing in advance that $<_{\mathrm{L}}$ is a linear order, from uniformizing the normal forms originally used, and from several other simplified arguments; the simplified proof first appeared in Miller's thesis [19].

This paper is intended to be a proof of the basics and other lemmas leading up to the division algorithm, not a complete survey of LD theorems. For a survey of the Dehornoy order on $B_\infty$ and the connections of $B_\infty$ with free LDs, see [7]. For an account of connections of free LDs with large cardinals see [14, 15]. For a general account of results on free LD see [6].

## 2. Confluence

Let $w = x[(xx)(xx)]$ and $v = (xx)[x(xx)]$. To show that $w \equiv_{\mathrm{LD}} v$ we could give the short proof $w = x[(xx)(xx)] = x[x(xx)] = xx[x(xx)] = v$. In this proof we used the LD law in both directions, but one of the most basic properties of $\mathcal{A}$ is that $\mathcal{A}$ is confluent (Theorem 2.3), i.e., if two terms are equivalent, the equivalence can be proved by expanding both terms using only the forward direction of the LD law (allowing only substitutions of $(ab)(ac)$ for subterms of the form $a(bc)$).

An $a \in A$ will typically be represented by the unique parsing $a = xa_1a_2\cdots a_n$ (so if $a = c_0c_1\cdots c_k$ with $c_0 \neq x$ then for some $i \geq 1$, $c_0 = xa_1a_2\cdots a_i$).

### Definition 2.1.

For $u, v \in A$, write $u \to^* v$ if $v$ can be obtained from $u$ by replacing a subterm of $u$ of the form $a(bc)$ with $(ab)(ac)$. Write $u \to v$ if there exist $u_0, \ldots, u_n \in A$ such that $u = u_0 \to^* u_1 \to^* \ldots \to^* u_n = v$. So $u \to v$ implies $u \equiv_{\mathrm{LD}} v$.

### Proposition 2.2.

*If $xa_1a_2\cdots a_n \to xb_1b_2\cdots b_m$ then for all $i \leq n$ there is a $j \leq m$ such that $xa_1\cdots a_i \to xb_1\cdots b_j$.*

The above proposition can be proved by induction. It is true if the first → is replaced by $\to^*$. Proceed by induction.

The remainder of this section is given to proving that $\mathcal{A}$ is confluent. Numerous people independently worked out proofs of confluence; here we use the notation of Jech and Dougherty.

### Theorem 2.3 (Dehornoy (confluence) [5]).

*Given $u, v \in A$, $u \equiv_{\mathrm{LD}} v$ implies that there exists $w \in A$ such that $u \to w$ and $v \to w$.*

We will prove Theorem 2.3 by assigning to each $u \in A$ a $\delta u$ such that $u \to \delta u$ and if $u \equiv_{\mathrm{LD}} v$, then $\delta^m u \equiv_{\mathrm{LD}} \delta^n v$ for some $m$ and $n$.

### Definition 2.4.

For $u, v \in A$, let $u\#v$ be the result of distributing $u$ to every occurrence of $x$ in $v$, i.e., $u\#x = ux$, $u\#(bc) = (u\#b)(u\#c)$.

### Lemma 2.5.

(i) $uv \to u\#v$,

(ii) $u\#(v\#w) \to (u\#v)\#(u\#w)$.

(i) is immediate; for (ii) use induction on $w$. We now define $\delta u$, where $\delta u$ roughly replaces every subterm $a(bc)$ in $u$ by $(ab)(ac)$: For $u \in A$, put $\delta x = x$, $\delta(vw) = \delta v\#\delta w$.

### Lemma 2.6.

$u \to \delta u$.

***Proof.*** By induction on the length of $u$. Suppose $u = vw$, then $u = vw \to \delta v\delta w$ (by induction) $\to \delta v\#\delta w$ (by Lemma 2.5) $= \delta u$. □

### Lemma 2.7.

*$u \to^* v$ implies $v \to \delta u$.*

***Proof.*** By induction on the length of $u$. Suppose $u = rs \to^* v$. If $rs \to^* r's = v$ we have $r' \to \delta r$ by induction, and $s \to \delta s$ by Lemma 2.6. So $r's \to \delta r \delta s \to \delta r \# \delta s = \delta(rs) = \delta u$ by Lemma 2.6. The case $u = rs \to^* rs' = v$ is similar. Suppose $u = r(s_0 s_1) \to^* (rs_0)(rs_1) = v$. We have

$$(rs_0)(rs_1) \to (r\#s_0)(r\#s_1) = r\#(s_0 s_1),$$

and $r\#(s_0 s_1) \to \delta r \#(\delta s_0 \# \delta s_1) = \delta u$ (by Lemmas 2.5 and 2.6). □

### Lemma 2.8.

*If $u \to v$, then $\delta u \to \delta v$.*

***Proof.*** It suffices to show that $u \to^* v$ implies that $\delta u \to \delta v$. Let $u = rs$ and $v = r's$. If $r \to^* r'$ we have $\delta u = \delta(rs) = \delta r \delta s \to \delta r' \delta s$ (induction hypothesis) $= \delta v$. The case $u = rs$ with $s \to^* s'$ is similar. Suppose $u = r(s_0 s_1)$, and $v = rs_0(rs_1)$. Then $\delta u = \delta r \#(\delta s_0 \# \delta s_1) \to (\delta r \# \delta s_0)\#(\delta r \# \delta s_1)$ (by Lemma 2.6) $= \delta v$. □

***Proof of Theorem 2.3.*** Let $u \equiv_{\mathrm{LD}} v$ via the sequence $u = u_1, u_2, \dots, u_n = v$ where for each $i$, $u_i \to^* u_{i+1}$ or $u_{i+1} \to^* u_i$. Since $v \to \delta^n v$ for all $n$ (Lemma 2.6), we are done proving confluence by proving there is an $m$ such that $u \to \delta^m v$. To show $u \to \delta^m v$ for some $m$, prove by induction on $i$ that there exists a $j$ such that $u \to \delta^j u_i$.

Suppose $u \to \delta^k u_i$ for some $k$.

**Case 1:** $u_i \to^* u_{i+1}$. Then $\delta^k u_i \to \delta^k u_{i+1}$ by Lemma 2.8, so $u \to \delta^{k+1} u_{i+1}$ by Lemma 2.7.

**Case 2:** $u_{i+1} \to^* u_i$. Then $u_i \to \delta u_{i+1}$ by Lemma 2.7, so $u \to \delta^k u_i \to \delta^{k+1} u_{i+1}$ by Lemma 2.8. □

## 3. $<_{\mathrm{L}}$ linearly orders $\mathcal{A}$

The proof that $<_{\mathrm{L}}$ is a linear ordering of $\mathcal{A}$ has two parts: connectedness ($a \leq_{\mathrm{L}} b$ or $b \leq_{\mathrm{L}} a$) and irreflexivity ($a \not<_{\mathrm{L}} a$). Let us use the following notation: $x^{(0)} = x$; $x^{(n+1)} = xx^{(n)}$ ($= x^{(i)} x^{(n)}$ for all $i \leq n$ by induction on $i$).

### Lemma 3.1.

*For all $w \in \mathcal{A}$ there exists $n$ such that $w <_{\mathrm{L}} x^{(n)}$.*

To prove the lemma above, show by induction on $w$ that there is an $i$ such that for all $j \geq i$, $wx^{(j)} = x^{(j+1)}$ and take $n = i + 1$.

### Theorem 3.2 (connectivity).

*For $a, b \in \mathcal{A}$, $a \leq_{\mathrm{L}} b$ or $b \leq_{\mathrm{L}} a$.*

***Proof (Dehornoy).*** By Lemma 3.1, pick an $n$ such that $a, b <_{\mathrm{L}} x^{(n)}$. So $x^{(n)} = aa_1 a_2 \cdots a_k = bb_1 b_2 \cdots b_j$. By confluence (Theorem 2.3), pick $c = xc_1 c_2 \cdots c_m$ such that $aa_1 a_2 \cdots a_k \to c$ and $bb_1 b_2 \cdots b_j \to c$. By Proposition 2.2, $a \to xc_1 \cdots c_r$ and $b \to xc_1 \cdots c_s$ for some $r, s \leq m$, and thus $a \leq_{\mathrm{L}} b$ or $b \leq_{\mathrm{L}} a$. □

### Theorem 3.3 (irreflexivity).

*$<_{\mathrm{L}}$ is irreflexive on $\mathcal{A}$.*

It suffices to find an irreflexive LD because if $a = aa_1a_2\cdots a_n \in \mathcal{A}$, then a homomorphic image in any LD would give a similar relation. An irreflexive LD is a rarity, so we mention Laver's original proof of irreflexivity though it is not used in the sequel; the proof we will use in this paper is Dehornoy's version, which is a theorem of ZFC.

**Laver's version.** Let $V_\lambda$ be the collection of all sets of rank less than $\lambda$. A function $j\colon V_\lambda \to V_\lambda$ is a (nontrivial) elementary embedding if and only if $j$ is not the identity function and whenever a statement holds (in $V_\lambda$) of members $a_0, \dots, a_n \in V_\lambda$ then the same statement holds (in $V_\lambda$) of $j(a_0), \dots, j(a_n)$. More formally, for all first order formulas, $\Phi$,

$$(V_\lambda, \in) \models \Phi(a_0, \dots, a_n) \qquad \text{if and only if} \qquad (V_\lambda, \in) \models \Phi(j(a_0), \dots, j(a_n)).$$

Let $\mathcal{E}_{V_\lambda}$ be the set of all such nontrivial embeddings. Then the existence of $\mathcal{E}_{V_\lambda} \neq \emptyset$ is a very strong large cardinal axiom and thus is unprovable in ZFC.

If $\mathcal{E}_{V_\lambda} \neq \emptyset$ we may assume $\lambda$ is a limit ordinal (of cofinality $\omega$ [11]). Then $\mathcal{E}_{V_\lambda}$ is a left distributive algebra under the operation

$$j \cdot k = jk = \bigcup_{\alpha<\lambda} j(k \cap V_\alpha).$$

Laver showed [14] that $\mathcal{E}_{V_\lambda}$ is an irreflexive LD. Moreover, if $j \in \mathcal{E}_{V_\lambda}$, then $\mathcal{A}_j$, the subalgebra of $\mathcal{E}_{V_\lambda}$ generated by $\{j\}$, is isomorphic to $\mathcal{A}$.

**Dehornoy's proof.** Note that the action of the braid group (1) in the introduction acts on direct products of automorphic sets, but we do not need automorphic sets. It becomes a partial action on $L^\omega$ (e.g., $\sigma_i^{-1}(l_1, l_2, \dots, l_n)$ might not exist) for any left cancellative LD, $L$. For details on this see [5, 7, 16].

We begin by making an observation that motivates the definition of an operation $[\,\cdot\,]$ on $B_\infty$. Let $\vec{x} = \langle x, x, x, \dots\rangle$. Let $B_\infty$ be the collection of all braid words, i.e., all words of the form $\sigma_{i_1}^{\pm 1}\sigma_{i_2}^{\pm 1}\cdots\sigma_{i_n}^{\pm 1}$. Let sh be the extension of the shift operation on $B_\infty$ (i.e., the extension of the map $\mathrm{sh}(\sigma_i) = \sigma_{i+1}$ to an endomorphism of $B_\infty$).

*Claim.* For all $a \in A$, there is a braid word $\alpha_a$ such that $\alpha_a(\vec{x})$ exists (as defined in the introduction) and $\alpha_a(\vec{x}) = \langle a, x, x, \dots\rangle$.

*Proof of Claim.* We define $\alpha_a$ by induction on the length of $a$. Let $\alpha_x = \mathrm{id}$. Note that if $\alpha_b$ is a braid word and $\alpha_b(\overrightarrow{x}) = \langle b, x, x, x, \dots\rangle$ then

$$\mathrm{sh}\big(\alpha_b^{-1}\big)(\langle x, b, x, x, \dots\rangle) = \langle x, x, x, x, \dots\rangle.$$

If $a = bc$, take $\alpha_a = \alpha_b\,\mathrm{sh}(\alpha_c)\sigma_1\,\mathrm{sh}\big(\alpha_b^{-1}\big)$. Then we can see that

$$\langle x,x,x,x,\dots\rangle \xrightarrow{\alpha_b} \langle b,x,x,x,\dots\rangle \xrightarrow{\mathrm{sh}(\alpha_c)} \langle b,c,x,x,\dots\rangle \xrightarrow{\sigma_1} \langle bc,b,x,x,\dots\rangle \xrightarrow{\mathrm{sh}(\alpha_b^{-1})} \langle bc,x,x,x,\dots\rangle. \qquad \blacksquare$$

### Definition 3.4.

The Dehornoy bracket on $B_\infty$ is the function $\alpha[\beta] = \alpha\,\mathrm{sh}(\beta)\sigma_1\,\mathrm{sh}(\alpha^{-1})$.

This is easily checked to be a left distributive operation. We claim that Dehornoy's bracket is an irreflexive operation. Say the braid word $\alpha$ is $\sigma_1$-positive if and only if $\sigma_1$ occurs in $\alpha$ but there are no $\sigma_1^{-1}$ in $\alpha$.

Dehornoy's argument is that if $[\,\cdot\,]$ is reflexive, then there is a $\sigma_1$-positive braid word that is the identity. That is, if there is some $\alpha$ such that $\alpha[\alpha_1[\alpha_2[\cdots\alpha_n]]] = \alpha$, then $\alpha_1[\alpha_2[\cdots[\alpha_n]]] = \mathrm{id}$, where $\alpha_1[\alpha_2[\cdots[\alpha_n]]]$ is $\sigma_1$-positive. Dehornoy then showed no such $\alpha$ exists; we give Larue's proof [12].

### Lemma 3.5 ($\sigma_1$-proposition).

*If $\alpha \in B_\infty$ is written as a product of generators and their inverses, including at least one $\sigma_1$ and no $\sigma_1^{-1}$, then $\alpha \neq \mathrm{id}$.*

The $\sigma_1$-proposition can be proven using the Hurwitz–Artin action of the braid group $B_\infty$ as automorphisms of the free group $\langle \mathcal{F}_G, \mathrm{id}, \cdot\rangle$ on countably many generators $G = \{g_i : i \in \omega\}$.

### Definition 3.6.

The Hurwitz–Artin action of $B_\infty$ on $\mathcal{F}_G$ [1, 2, 9] is

$$(g_i)\sigma_i = g_i \cdot g_{i+1} \cdot g_i^{-1}, \qquad (g_{i+1})\sigma_i = g_i, \qquad \text{and if } j \neq i, i+1 \qquad (g_j)\sigma_i = g_j.$$

Thus the action of the inverses on the generators is

$$(g_i)\sigma_i^{-1} = g_{i+1}, \qquad (g_{i+1})\sigma_i^{-1} = {g_{i+1}}^{-1} \cdot g_i \cdot g_{i+1}, \qquad \text{and if } j \neq i, i+1 \qquad (g_j)\sigma_i^{-1} = g_j.$$

Establishing the following lemma will conclude the proof of the $\sigma_1$-proposition.

### Lemma 3.7.

*If a reduced word $f$ in $\mathcal{F}_G$ begins with $g_1$, and $\sigma \in B_\infty$ is a generator or the inverse of a generator, but not $\sigma_1^{-1}$, then the reduced form of $(f)\sigma$ also begins with $g_1$.*

***Proof.*** Suppose that $\sigma$ has been applied to every generator and the inverse of every generator in the reduced form of $f$ and that a fixed reduction has been applied to the result, yielding the reduced form of $(f)\sigma$.

There are two cases in which $(f)\sigma$ might not begin with $g_1$.

**Case 1: $\sigma = \sigma_i^{\pm 1}$ with $i > 1$.** In this case, as $i > 1$, all $g_1$ are unchanged by the action of $\sigma$ on (the reduced form of) $f$, and no $g_1$ result from the action. Thus the only way for $g_1$ to be cancelled from $f$ by the action of $\sigma$ is for $g_1^{-1}$ to already be present in the reduced form of $f$. Thus it must be that, for some $f_1, f_2 \in \mathcal{F}_G$, the reduced form of $f$ is $f = g_1 \cdot f_1 \cdot g_1^{-1} \cdot f_2$. But then $(f)\sigma = g_1 \cdot (f_1)\sigma \cdot g_1^{-1} \cdot (f_2)\sigma$, where $(f_1)\sigma = \mathrm{id}$, contradicting the claim that $f$ was in reduced form.

**Case 2: $\sigma = \sigma_1$.** $\sigma_1$ can produce $g_1^{-1}$ in two different ways: either by its action on $g_1^{\pm 1}$: $(g_1^{\pm 1})\sigma_1 = g_1 \cdot g_2^{\pm 1} \cdot g_1^{-1}$; or by its action on $g_2^{-1}$: $(g_2^{-1})\sigma_1 = g_1^{-1}$. We consider these cases separately.

*Case 2.1: The leading $g_1$ in the unreduced form of $(f)\sigma$ is cancelled by a $g_1^{-1}$ produced by the action of $\sigma_1$ on $g_1^{\pm 1}$.* Let the reduced form of $f$ displaying that instance of $g_1^{\pm 1}$ be $f = g_1 \cdot f_1 \cdot g_1^{\pm 1} \cdot f_2$. Thus

$$(f)\sigma_1 = g_1 \cdot g_2 \cdot g_1^{-1} \cdot (f_1)\sigma_1 \cdot g_1 \cdot g_2^{\pm 1} \cdot g_1^{-1} \cdot (f_2)\sigma_1.$$

Thus $g_2 \cdot g_1^{-1} \cdot (f_1)\sigma_1 \cdot g_1 \cdot g_1^{-1} \cdot (f_1)\sigma_1 \cdot g_1 \cdot g_2^{\pm 1} = \mathrm{id}$, so $(f_1)\sigma_1 = g_1 \cdot g_2^{-1} \cdot g_2^{\mp 1} \cdot g_1^{-1}$ and $f_1 = g_1^{-1} \cdot g_1^{\mp 1}$, which gives that $f$ was not in reduced from, a contradiction.

*Case 2.2: The leading $g_1$ in the unreduced form of $(f)\sigma$ is cancelled by $g_1^{-1}$ produced by the action of $\sigma_1$ on $g_2^{-1}$.* Now let the reduced form of $f$ displaying that instance of $g_2^{-1}$ be $f = g_1 \cdot f_1 \cdot g_2^{-1} \cdot f_2$. Thus

$$(f)\sigma_1 = g_1 \cdot g_2 \cdot g_1^{-1} \cdot (f_1)\sigma_1 \cdot g_1^{-1} \cdot (f_2)\sigma_1.$$

Thus $g_2 \cdot g_1^{-1} \cdot (f_1)\sigma_1 = \mathrm{id}$, so $(f_1)\sigma_1 = g_1 \cdot g_2^{-1}$ and $f_1 = g_1^{-1} \cdot g_2$, which again gives that $f$ was not in reduced from, a contradiction. ■

Definition 3.6 and Lemma 3.7 establish the $\sigma_1$-proposition (Lemma 3.5). In particular, suppose that $p$ is a $\sigma_1$-positive braid in $B_\infty$ (a product of generators and their inverses that begins with $\sigma_1$ and in which no $\sigma_1^{-1}$ appears). Then $p$ is of the form $p = p_1 \sigma_1 p_2$, where $p_1$ does not contain any instance of $\sigma_1^{\pm 1}$ and $p_2$ does not contain $\sigma_1^{-1}$. Then

$$(g_1^{-1})p = (g_1^{-1})p_1\sigma_1 p_2 = (g_1^{-1})\sigma_1 p_2 = (g_1 \cdot g_2^{-1} \cdot g_1)p_2.$$

As $g_1 \cdot g_2^{-1} \cdot g_1$ is in reduced form, begins with $g_1$, and there is no $\sigma_1^{-1}$ in $p_2$, we have that $(g_1 \cdot g_2^{-1} \cdot g_1)p_2 = (g_1^{-1})p \neq \mathrm{id}$. This concludes the proof of Lemma 3.5, and therefore the proof that $\mathcal{A}$ is irreflexive. □

### Theorem 3.8.

*$<_{\mathrm{L}}$ is a linear ordering of $\mathcal{A}$.*

Theorem 3.8 follows directly from Theorems 3.2 and 3.3.

### Corollary 3.9.

(i) *$ac <_{\mathrm{L}} ab$ if and only if $c <_{\mathrm{L}} b$.*

(ii) *$ac = ab$ if and only if $c = b$.*

(iii) *The word problem for $\mathcal{A}$ is solvable.*

Recall that the action of $B_\infty$ on automorphic sets described in (1) is a partial action on $L^\infty$, where $L$ is any left cancellative LD. Using the linearity of $<_{\mathrm{L}}$ on $\mathcal{A}$, define Dehornoy's linear ordering of the braid group, $<$, as follows. For details see [7].

For $\alpha, \beta \in B_\infty$, $\alpha < \beta$ if and only if for some $N < \infty$, there is $\vec{l} \in \mathcal{A}^N$ with $\alpha\vec{l}$ lexicographically less than $\beta\vec{l}$ with respect to $<_{\mathrm{L}}$.

Laver pointed out that $\alpha > \mathrm{id}$ if and only if $\alpha$ is $\sigma$-positive, where $\sigma$-positive means that the generator with least index appears only positively. See [7] for ten different proofs of the linearity of $<$.

### Remark.

The division algorithm and a many variable variation of it [16] have been used in the proofs of combinatorial facts about $\mathcal{A}$ and in other results. Examples include:

- $\alpha$ is a $B_N$-braid word greater than the identity if and only if $\alpha$ is equivalent to a $\sigma$-positive word in $B_N$. (It is somewhat surprising that this was a problem.)
- $B_N^+ = \{$positive braid words in $B_N\}$ is well-ordered under Dehornoy's ordering. Burckel [4] found a different proof (not using the division algorithm) and computed the associated ordinals.
- The one generator case of a conjecture of Moody [18] holds [17]: If $a, b, c, d \in \mathcal{A}$, $ab = cd$, $a$ and $b$ have no common left divisors, and $c$ and $d$ have no common left divisors, then $a = c$ and $b = d$.

For results and problems about a possible algorithm and applications of it in the many generator case, see [19, 20].

## 4. The extension $\mathcal{P}$ of $\mathcal{A}$

Let $a_0, w \in \mathcal{A}$ with $a_0 <_{\mathrm{L}} w$, i.e., $w = a_0 b_1 \cdots b_n$. One might hope that this can be witnessed canonically, i.e., that there were $a_1$ that was $<_{\mathrm{L}}$-greatest with $a_0 a_1 \leq_{\mathrm{L}} w$, and, in the case $a_0 a_1 <_{\mathrm{L}} w$, $a_2$ such that $a_0 a_1 a_2 \leq_{\mathrm{L}} w$, and so on until we get $w = a_0 a_1 a_2 \cdots a_n$. Then $a_1, \dots, a_n$ would be unique and this would be a division algorithm for $\mathcal{A}$. However, such $a_1$ need not exist in $\mathcal{A}$.

Namely, consider the case $w = xx(xxx)$ and $a_0 = x$. By left distributivity we have the following:

$$\begin{aligned} w = xx(xxx) &\equiv xx(xx)(xxx) \equiv x(xx)(xxx) \equiv x(xx)(xx)[x(xx)x] \\ &\equiv x(xxx)[x(xx)x] \equiv x(xxxx)(x(xxxx))[x(xxx)x] \equiv \dots \end{aligned}$$

More generally, we will see that, for $w$ given above, if $b \in \mathcal{A}$ and $xb <_{\mathrm{L}} w$, then $x(bx) <_{\mathrm{L}} w$ and there is no $b$ with $xb = w$.

We define an extension of the left distributive law by freely adding a composition symbol, $\circ$, to the language. Let $\Sigma$ be the following set of laws in the language $\{\cdot, \circ\}$:

$$(a\circ b)\circ c = a\circ(b\circ c), \qquad (a\circ b)c = a(bc), \qquad a(b\circ c) = ab\circ ac, \qquad a\circ b = ab\circ a.$$

The first two axioms are the normal properties of composition. The last two (with the aid of the second) assert that left multiplication is a homomorphism of the algebra. Namely, the second and fourth axioms together give left distributivity:

$$a(bc) = (a\circ b)c = (ab\circ a)c = (ab)(ac).$$

An example of a familiar algebra that satisfies $\Sigma$ is group conjugation, with operations $\cdot$ and $\circ$, as given in the introduction (with $* = \circ$).

Let $P$ be the collection of all terms in the language $\{\cdot, \circ\}$ and one generator, $x$. $\mathcal{P}$ is the free algebra satisfying the laws of $\Sigma$; $\mathcal{P} = P/\Sigma$. $(\mathcal{P}, \cdot)$ contains $(\mathcal{A}, \cdot)$ as a subalgebra. Then there is a natural linear ordering $<_{\mathrm{L}}$ of $\mathcal{P}$ which agrees with the $<_{\mathrm{L}}$ of $\mathcal{A}$. $\mathcal{P}$ acts as a type of completion of $\mathcal{A}$ (see Theorem 4.9) containing various least upper bounds, in particular those needed for the division algorithm.

We will occasionally blur the distinction between a term in elements of $P$ and a member of $\mathcal{P}$ (an equivalence class), and we will omit parentheses in $n$-fold compositions.

Preliminary to the normal form produced by the division algorithm (Theorem 8.1), there are two basic types of representations of members of $\mathcal{P}$. The second is given in Lemma 4.2; the first is: For $p, a_0, \dots, a_n \in \mathcal{P}$, write $p = a_0a_1\cdots a_{n-1}*a_n$ to mean that either $p = a_0a_1\cdots a_{n-1}a_n$ or $p = a_0a_1\cdots a_{n-1}\circ a_n$. When presented with such a parsing of $p$, we call $a_i$ the *subterms* of $p$. Note that each $a_i$ is a proper, literal subterm of $p$ (unless $p = a_0$).

We will now give a summary of useful definitions and lemmas pertaining to $\mathcal{P}$, beginning by extending the definition of $<_{\mathrm{L}}$ to $\mathcal{P}$: For $p, q \in \mathcal{P}$, $p <_{\mathrm{L}} q$ if and only if there exist $r_1, \dots, r_n \in \mathcal{P}$ such that $q = pr_1\cdots r_{n-1}*r_n$. We have that $<_{\mathrm{L}}$ is transitive, and $pq <_{\mathrm{L}} p\circ q = pq\circ p$, in fact

### *Lemma 4.1.*

*For $p, q, r \in \mathcal{P}$, if $q <_{\mathrm{L}} r$, then $pq <_{\mathrm{L}} p\circ q <_{\mathrm{L}} pr$.*

The proof of Lemma 4.1 is straightforward. Indeed, for $r = qa_1a_2\cdots a_{n-1}*a_n$,

$$pr = pq(pa_1)(pa_2)\cdots(pa_{n-1})*(pa_n) = (pq\circ p)a_1(pa_2)\cdots(pa_{n-1})*(pa_n) = (p\circ q)a_1(pa_1)\cdots(pa_{n-1})*(pa_n).$$

Thus $p\circ q <_{\mathrm{L}} pr$.

Here is the first version of the division algorithm: For $p <_{\mathrm{L}} q \in \mathcal{P}$ there is always $p_1 \in \mathcal{P}$ such that

$$p_1 \text{ is } <_{\mathrm{L}}\text{-greatest such that } pp_1 \leq_{\mathrm{L}} q.$$

If $pp_1 = q$ or $p\circ p_1 = q$ the algorithm terminates. Otherwise, let $p_2$ be $<_{\mathrm{L}}$-greatest such that $pp_1p_2 \leq_{\mathrm{L}} q$. Continuing in this manner, for some $n$, the sequence terminates with $pp_1p_2\cdots p_{n-1}*p_n = q$. For $w = p\circ q \in \mathcal{P}$, define the iterates $I_k(p,q)$ of $w$ by

$$I_0(p,q) = q, \qquad I_1(p,q) = p, \qquad I_2(p,q) = pq, \qquad \dots, \qquad I_{n+2}(p,q) = I_{n+1}(p,q)I_n(p,q).$$

In practice we use the iterates with $n > 0$, thus the sequence of positive iterates begins: $p, pq, pqp, pqp(pq), pqp(pq)(pqp), \dots$ Then inductively using the $a\circ b = ab\circ a$ law, $I_{n+1}(p,q)\circ I_n(p,q) = p\circ q$ and $I_1 <_{\mathrm{L}} I_2 <_{\mathrm{L}} \dots <_{\mathrm{L}} I_n <_{\mathrm{L}} \dots <_{\mathrm{L}} p\circ q$. In fact $p\circ q$ is the $<_{\mathrm{L}}$-least upper bound of the $I_n(p,q)$, $n \geq 1$. (See Corollary 5.5 and the discussion following Definition 5.12 in Section 5.)

Regarding the connection of $\mathcal{P}$ with $\mathcal{A}$, we define the second type of basic representation of a member of $\mathcal{P}$: $P^*$ is the set of formal compositions $a_0\circ a_1\circ\cdots\circ a_n$ of members of $A$.

### Lemma 4.2.

*Each $p \in P$ is represented an $a_0 \circ a_1 \circ \cdots \circ a_n \in P^*$, where $n = n_p$ is unique.*

**Proof.** The existence of such a representation is routine. If $p \in P$, a concrete example of $a_0 \circ a_1 \circ \cdots \circ a_n$ representing $p$ is $\overline{p}$, where $\overline{x} = x$, $\overline{(q \circ r)} = \overline{q} \circ \overline{r}$, and if $\overline{q} = b_0 \circ \cdots \circ b_m$ and $\overline{r} = c_0 \circ \cdots \circ c_k$, then

$$\overline{(qr)} = b_0(b_1(\cdots(b_m c_0))) \circ b_0(b_1(\cdots(b_m c_1))) \circ \cdots \circ b_0(b_1(\cdots(b_m c_k))).$$

To see that $n = n_p$ is unique, let, for $p \in P$, $N(p)$ be the number of essential compositions in $p$: $N(x) = 0$, $N(uv) = N(v)$, $N((u \circ v)) = N(u) + N(v) + 1$. Then $N$ is invariant under $\Sigma$. □

In the remainder of this section we show that $\Sigma$ is a conservative extension of $\{\mathrm{LD}\}$, i.e.: For $a, b \in A$, if $a \equiv_\Sigma b$ then $a \equiv_{\mathrm{LD}} b$. The next definition makes explicit the basic connection between braid groups and left distributive algebras.

### Definition 4.3.

For $\overset{*}{p} = a_0 \circ \cdots \circ a_n$, $\overset{*}{q} = b_0 \circ \cdots \circ b_n$, $a_0, \ldots, a_n, b_0, \ldots, b_n \in A$, write

$$\overset{*}{p} \sim \overset{*}{q}$$

if and only if there exists a sequence $\overset{*}{p} = \overset{*}{p}_0, \overset{*}{p}_1, \ldots, \overset{*}{p}_i, \ldots, \overset{*}{p}_m = \overset{*}{q}$ such that for each $i$, $\overset{*}{p}_{i+1}$ is obtained from $\overset{*}{p}_i$ by either: an internal application of the LD law or an application of the braid operation. Examples: $c_0 \circ \cdots \circ c_i \circ \cdots \circ c_n \leftrightarrow c_0 \circ \cdots \circ c_i' \circ \cdots \circ c_n$, where $c_i \equiv_{\mathrm{LD}} c_i'$; $c_0 \circ \cdots \circ c_i \circ c_{i+1} \circ \cdots \circ c_n \leftrightarrow c_0 \circ \cdots \circ c_i c_{i+1} \circ c_i \circ \cdots \circ c_n$.

### Theorem 4.4.

*For all $q, p \in P^*$, $q \equiv_\Sigma p \;\Leftrightarrow\; \overset{*}{q} \sim \overset{*}{p}$.*

Theorem 4.4 is proven by establishing Lemmas 4.6 and 4.7 below.

### Definition 4.5.

For $\overset{*}{a} = a_0 \circ a_1 \circ \cdots \circ a_n$, $\overset{*}{c} = c_0 \circ c_1 \circ \cdots \circ c_m \in P^*$, define

(i) $\overset{*}{a} \mathbin{\widehat{\circ}} \overset{*}{c} = a_0 \circ a_1 \circ \cdots \circ a_n \circ c_0 \circ c_1 \circ \cdots \circ c_m$,

(ii) $\overset{*}{a} \mathbin{\widehat{\cdot}} \overset{*}{c} = a_0(a_1(\cdots(a_n c_0))) \circ a_0(a_1(\cdots(a_n c_1))) \circ \cdots \circ a_0(a_1(\cdots(a_n c_m)))$.

### Lemma 4.6.

*For $\overset{*}{a}, \overset{*}{b}, \overset{*}{c}, \overset{*}{d} \in P^*$, if $\overset{*}{a} \sim \overset{*}{b}$ and $\overset{*}{c} \sim \overset{*}{d}$, then:*

(i) $\overset{*}{a} \mathbin{\widehat{\circ}} \overset{*}{c} \sim \overset{*}{b} \mathbin{\widehat{\circ}} \overset{*}{d}$,

(ii) $\overset{*}{a} \mathbin{\widehat{\cdot}} \overset{*}{c} \sim \overset{*}{b} \mathbin{\widehat{\cdot}} \overset{*}{d}$.

**Proof.** The proof of (i) is clear by first applying the transformations from $\overset{*}{a}$ to $\overset{*}{b}$ and then those from $\overset{*}{c}$ to $\overset{*}{d}$. Each transformation is as described in Definition 4.3, thus $\overset{*}{a} \mathbin{\widehat{\circ}} \overset{*}{c} \sim \overset{*}{b} \mathbin{\widehat{\circ}} \overset{*}{d}$. For (ii), proceed by induction. Observe that if $\overset{*}{c} = c_0 \circ \cdots \circ c_n$, $\overset{*}{d} = d_0 \circ \cdots \circ d_n$ with $c_0, \ldots, c_n, d_0, \ldots, d_n \in A$ and $\overset{*}{c} \sim \overset{*}{d}$, then $c_0(c_1(\cdots(c_n x))) \equiv_{\mathrm{LD}} d_0(d_1(\cdots(d_n x)))$. Furthermore, whenever $c_0(c_1(\cdots(c_n x))) \equiv_{\mathrm{LD}} d_0(d_1(\cdots(d_n x)))$, then for any $\overset{*}{w} \in P^*$,

$$(\overset{*}{w} \mathbin{\widehat{\cdot}} c_0)\big((\overset{*}{w} \mathbin{\widehat{\cdot}} c_1)(\cdots((\overset{*}{w} \mathbin{\widehat{\cdot}} c_n)x))\big) \equiv_{\mathrm{LD}} (\overset{*}{w} \mathbin{\widehat{\cdot}} d_0)\big((\overset{*}{w} \mathbin{\widehat{\cdot}} d_1)(\cdots((\overset{*}{w} \mathbin{\widehat{\cdot}} d_n)x))\big).$$ □

### Lemma 4.7.

*For $\overset{*}{a}, \overset{*}{b}, \overset{*}{c} \in P^*$, the following relations hold.*

(i) $\overset{*}{a} \mathbin{\widehat{\circ}} (\overset{*}{b} \mathbin{\widehat{\circ}} \overset{*}{c}) \sim (\overset{*}{a} \mathbin{\widehat{\circ}} \overset{*}{b}) \mathbin{\widehat{\circ}} \overset{*}{c}$,

(ii) $(\overset{*}{a} \mathbin{\widehat{\circ}} \overset{*}{b}) \mathbin{\widehat{\cdot}} \overset{*}{c} \sim \overset{*}{a} \mathbin{\widehat{\cdot}} (\overset{*}{b} \mathbin{\widehat{\cdot}} \overset{*}{c})$,

(iii) $\overset{*}{a} \mathbin{\widehat{\cdot}} (\overset{*}{b} \mathbin{\widehat{\circ}} \overset{*}{c}) \sim \overset{*}{a} \mathbin{\widehat{\cdot}} \overset{*}{b} \mathbin{\widehat{\circ}} \overset{*}{a} \mathbin{\widehat{\cdot}} \overset{*}{c}$,

(iv) $\overset{*}{a} \mathbin{\widehat{\circ}} \overset{*}{b} \sim \overset{*}{a} \mathbin{\widehat{\cdot}} \overset{*}{b} \mathbin{\widehat{\circ}} \overset{*}{a}$.

The proof is left to the reader.

**Proof of Theorem 4.4.** For $p', p \in P$, it is enough to consider the case when $p'$ is obtained from $p$ by a one-step application of a law of $\Sigma$. Then by Lemmas 4.6 and 4.7, $\overset{*}{p}{}' \sim \overset{*}{p}$. □

### Definition 4.8.

For $a_0 \circ a_1 \circ \cdots \circ a_n \in P^*$ let $\vec{a} = \langle a_0, a_1, \ldots, a_n \rangle$.

### Theorem 4.9.

(i) *If $a_0, \ldots, a_n, b_0, \ldots, b_n \in \mathcal{A}$, then $a_0 \circ \cdots \circ a_n = b_0 \circ \cdots \circ b_n \in \mathcal{P}$ if and only if there is $\alpha \in B_{N+1}$ such that $\alpha(\vec{a}) = \vec{b}$ (using the action defined in (1));*

(ii) *$\Sigma$ is a conservative extension of LD;*

(iii) *if $a, b \in A$ then $a <_{\mathrm{L}} b$ via $\Sigma$ implies $a <_{\mathrm{L}} b$ via LD; and*

(iv) *$<_{\mathrm{L}}$ is a linear ordering of $\mathcal{P}$.*

**Proof.** (i) is in essence Theorem 4.4.

(ii) If $a, b \in A$, $a \equiv_\Sigma b$, then (by Theorem 4.4), $\overset{*}{a} \sim \overset{*}{b}$. (Note that $a = \overset{*}{a}$ as there are no compositions in $a \in A$.) As $a, b \in A$, no braid operations are possible, which gives that $a \equiv_{\mathrm{LD}} b$. (iii) follows from (ii).

(iv) For any $a, b \in \mathcal{P}$, $a <_{\mathrm{L}} b$ via $\Sigma$ if and only if $ax <_{\mathrm{L}} bx$ via LD. (iv) follows directly from irreflexivity of $<_{\mathrm{L}}$ on $\mathcal{A}$ together with this observation. (Details left to the reader.) □

## 5. Normality and a restatement and strengthening of the division algorithm

Recall from Section 4 that the division algorithm (Theorem 8.1) says that for $p <_{\mathrm{L}} q$, $p, q \in \mathcal{P}$, $pp_1p_2 \cdots p_{n-1}p_n = q$ or $pp_1p_2 \cdots p_{n-1} \circ p_n = q$ for some $n$ (i.e., $pp_1p_2 \cdots p_{n-1} * p_n = q$) where $p_1$ is $<_{\mathrm{L}}$-greatest such that $pp_1 \leq_{\mathrm{L}} q$, $p_2$ is $<_{\mathrm{L}}$-greatest such that $pp_1p_2 \leq_{\mathrm{L}} q$, and so on.

For the example $p = x$, $q = w = xx(xxx)$ given in Section 4, $p_1 = x \circ x$, and the algorithm produces $q = x(x \circ x)x$.

The output $pp_1 \cdots p_{n-1} * p_n$ of an application of the division algorithm will turn out to be *normal*, where

### Definition 5.1.

$w = p_0p_1p_2 \cdots p_{n-1} * p_n \in \mathcal{P}$ is $p_0$-*normal* if and only if $p_2 \leq_{\mathrm{L}} p_0$, $p_i \leq_{\mathrm{L}} p_0p_1 \cdots p_{i-2}$ for $2 \leq i \leq n$, and, if $* = \circ$, $p_n <_{\mathrm{L}} p_0p_1 \cdots p_{n-2}$. (Note that $p_1$ is arbitrary.)

We will see that there can be at most one $p_0$-normal term representing $w$, namely the term given by the division algorithm. Whether a term is normal depends on how it is parsed, i.e., $xx(x \circ x)x$ is $xx$-normal but not $x$-normal. We make the convention that normality of $p_0p_1 \cdots p_{n-1} * p_n$ means $p_0$-normality. Note that this term is also normal when it is parsed as $cp_i \cdots p_{n-1} * p_n$ (where $c = p_0p_1 \cdots p_{i-1}$). In the case $* = \circ$, we require that $p_n <_{\mathrm{L}} p_0p_1 \cdots p_{n-2}$ rather than $p_n \leq_{\mathrm{L}} p_0p_1 \cdots p_{n-2}$, because if $p_n = p_0p_1 \cdots p_{n-2}$ we have

$$w = p_0p_1 \cdots p_{n-2}p_{n-1} \circ p_n = p_0p_1 \cdots p_{n-2}p_{n-1} \circ p_0p_1 \cdots p_{n-2} = p_0p_1 \cdots p_{n-2} \circ p_{n-1}$$

contrary to the goal that the $p_0$-normal representation of $w$ be unique.

Noticing that in the case of $p_n = p_0p_1 \cdots p_{n-2}$, the collapsing in the discussion of Definition 5.1 would continue if $p_{n-1} = p_0p_1 \cdots p_{n-3}$, etc., we make the following definition.

### Definition 5.2.

Suppose $p_0p_1 \cdots p_{n-1}p_n$ is $p_0$-normal and $w = p_0p_1p_2 \cdots p_{n-1} \circ p_n \in \mathcal{P}$. Define $\operatorname{collapse}(w)$ (with respect to $p_0$, which will be clear by context) to be $p_0p_1 \cdots p_{i-1} \circ p_i$ for the greatest $i \leq n$ such that $p_i <_{\mathrm{L}} p_0p_1 \cdots p_{i-2}$. If there is no such $i$, $\operatorname{collapse}(w) = p_0 \circ p_1$.

Then $\operatorname{collapse}(w)$ is a $p_0$-normal term representing $w$.

With the definition of normality, a restatement of the division algorithm theorem is: *If $p, q \in \mathcal{P}$, $p <_{\mathrm{L}} q$ then there is a (unique) $p$-normal term $pp_1 \cdots p_{i-1} * p_i$ equalling $q$ (namely the result of applying the division algorithm with base $p$ to $q$).* We now check how to compare $p$-normal terms. The key to the uniqueness and comparison of $p$-normal representations is the following lemma.

### Lemma 5.3.

*Suppose that $w, b_1, \ldots, b_n, a \in \mathcal{P}$ and that $b_1 <_{\mathrm{L}} a$. Then if $wb_1 \cdots b_{n-1} * b_n$ is normal, $wb_1 \cdots b_{n-1} * b_n <_{\mathrm{L}} wa$.*

***Proof.*** By induction on $i \leq n$, it suffices to show that $wb_1 \cdots b_i \circ b_{i+1} <_{\mathrm{L}} wa$. For $i = 1$, $b_1 <_{\mathrm{L}} a$ implies $w \circ b_1 < wa$ by Lemma 4.1. Suppose $wb_1 \cdots b_{i-1} \circ b_i <_{\mathrm{L}} wa$. By normality $b_{i+1} \leq_{\mathrm{L}} wb_1 \cdots b_{i-1}$, so

$$wa >_{\mathrm{L}} wb_1 \cdots b_{i-1} \circ b_i \ \text{(by the induction hypothesis)} \ = wb_1 \cdots b_{i-1}b_i \circ (wb_1 \cdots b_{i-1}) \geq_{\mathrm{L}} wb_1 \cdots b_{i-1}b_i \circ b_{i+1}.$$

This gives the induction step. □

### Corollary 5.4.

*For $p$-normal terms $r$ and $q = pp_1 \cdots p_{n-1} * p_n$, $q <_{\mathrm{L}} r$ if and only if*

- *$q = pp_1 \cdots p_{n-1}p_n$ and $r = pp_1 \cdots p_{n-1}p_n \cdots p_{k-1} * p_k$ for some $k > n$, or*
- *$r$ begins with $pp_1 \cdots p_{i-1} * r_i$ for some $i \leq n$ and $p_i <_{\mathrm{L}} r_i$, or*
- *$r = pp_1 \cdots p_{i-1} \circ p_i$ for some $i \leq n$ (where $i = n$ can occur only if $* = \cdot$ in $q$).*

The above may be rephrased as a lexicographic comparison. Namely, let the associated sequence of the $p$-normal term $pp_1 \cdots p_{n-1}p_n$ be $\langle p, p_1, p_2, \ldots, p_n \rangle$ and let the associated sequence of $pp_1 \cdots p_{n-1} \circ p_n$ be, letting $u = pp_1 \cdots p_{n-1}$,

$$\langle p, p_1, \ldots, p_{n-1}, p_n, u, up_n, up_nu, \ldots \rangle$$

so, the coordinates from $u$ onward are the sequence of iterates, $I_k(u, p_n)$, $k \geq 1$.

### Corollary 5.5.

*The lexicographic comparison of associated sequences of normal terms is equivalent to comparison under $<_{\mathrm{L}}$; namely $pp_1 \cdots p_{n-1} * p_n <_{\mathrm{L}} pq_1 \cdots q_{m-1} * q_m$ if and only if the associated sequence of the former is $<_{\mathrm{L}}$-lexicographically less than the associated sequence of the latter.*

## 5.1. Outline of the rest of the proof

In this (optional) section we give some motivational comments about hereditarily $p$-normal terms and particularly the ensuing definitions and results about horseshoe ($\sqsupset$) relations. We will discuss a strengthening of the division form theorem and indicate why $\sqsupset$-relations are needed.

To prove the division algorithm for $x$ terminates for all $w \in \mathcal{P}$, it suffices to show that the product $(r \cdot s)$ and composition $(r \circ s)$ of two $x$-normal terms $r$ and $s$ can each be expressed as an $x$-normal term. This suffices because the closure of $\{x\}$ under $\cdot$ and $\circ$ is $\mathcal{P}$.

Subsection 5.2 lists basic cases (given as Rules 1–6) where the $p$-normal representations

$$r = pr_1 \cdots r_{i-1} * r_i, \qquad s = ps_1 \cdots s_{j-1} * s_j$$

are such that the $p$-normal representations of $rs$ and $r \circ s$ can be specifically described. (The reader may wish to look over Rules 1–6 at this point.)

Up to this point $\mathcal{P}$ was sufficient for us to state the basic facts about normal forms that will be needed to complete the proof, but the induction required for the later parts of the proof is better described by letting the "$p$-normal representation" be a member of $P$ — not $\mathcal{P}$ — in order to be assured that we are working with well-founded trees.

So, making the change from $\mathcal{P}$ to $P$, define, for $w \in \mathcal{P}$ (or $P$) the division form (= $x$ division form) $|w|^x$ of $w$ to be the (unique if it exists) term $x_1x_2 \cdots x_{i-1} * x_i$ in $P$ which represents $w$ and is hereditarily $x$-normal, i.e., every subterm of $|w|^x$ is $x$-normal. (See Definition 5.7.)

An example of a term in $x$-DF (the set of all words in $x$-division form) is: $|w|^x = x(x(xx)x)x(xx) \circ (x \circ xxx)$. Note that every proper subterm of $|w|^x$ is $x$-normal, and $|w|^x$ itself is $x$-normal, the last step of that being that $(x \circ xxx) <_{\mathrm{L}} x(x(xx)x)x$, which holds as $xxx <_{\mathrm{L}} x(xx)$ and therefore $x \circ xxx <_{\mathrm{L}} x(x(xx))$ (using the fact that $b <_{\mathrm{L}} c$ implies $ab <_{\mathrm{L}} a \circ b <_{\mathrm{L}} ac$) so $(x \circ xxx) <_{\mathrm{L}} x(x(xx)) <_{\mathrm{L}} x(x(xx)x)x$.

Similarly, if $p >_{\mathrm{L}} x$ and $w \in \mathcal{P}$, then the $p$-division form $|w|^p$ of $w$ is the (unique if it exists) $p$-normal term in $P$, which represents $w$, all of whose subterms are $p$-normal, where each subterm $v$ of $|w|^p$ is $p$-normal if $v >_{\mathrm{L}} p$ and $x$-normal if $v \leq_{\mathrm{L}} p$. (See Definition 5.8.)

The main theorem (Theorem 8.1), *for all* $w, p \in \mathcal{P}\, |w|^{|p|}$ *exists*, prohibits certain infinite descents in $\mathcal{P}$; for example there is no infinite sequence $s_0, s_1, \ldots, s_k, \ldots$ of members of $\mathcal{P}$ such that each $s_k$ is $x$-normal and, letting $s_n = xa_1 \cdots a_{j-1} * a_j$, $s_{n+1}$ is one of the $a_k$'s. Namely, for $x$-DF, take $p = x$ and $w = s_0$.

There are variations of $p$-DF, each with its $\sqsupset$ relation. Let $\sqsupset' \subset P \times P$ be the horseshoe relation associated with a normal form $|\cdot|'$. (The definitions of the $\sqsupset'$ are in the succeeding section.) The $\sqsupset'$ relations describe when we can prove directly the existence of certain $|\cdot|'$ normal form terms from the existence of $|p|'$ and $|q|'$. The main results about $\sqsupset'$ are

### Theorem 5.6 (Main Horseshoe Theorem).

*If* $|p|' \sqsupset' |q|'$, *then*

(i) *for every* $|v|' \leq_{\mathrm{L}} |q|'$, $|p|' \sqsupset' |v|'$ *holds;*

(ii) $|pq|'$ *and* $|p \circ q|'$ *exist; and*

(iii) $|pq|' \sqsupset' |p|'$.

Proving the main theorem of the paper (Theorem 8.1) involves finding $|uv|^p$ and $|u \circ v|^p$, given $|u|^p$ and $|v|^p$, which in turn may involve changing bases, e.g., finding $|v|^p$ given $|v|^q$. The $\sqsupset$-definitions and theorems incorporate all that can be done in this matter using Rules 1–6 (described in subsection 5.2). For example: if $|u|^p = pp_1 \cdots p_{n-1} * p_n$ and $|t|^p = pp_1 \cdots p_{n-2}$, then $|u|^p \sqsupset^p |t|^p$ by definition of $\sqsupset^p$ (which we will see later) and by Theorem 5.6 (ii), $|ut|^p$ and $|u \circ t|^p$ exist. Furthermore, if $|v|^p \leq_{\mathrm{L}} |t|^p$ then $|u|^p \sqsupset^p |v|^p$ by Theorem 5.6 (i).

The changing of bases that will be needed is not a consequence of direct applications of Rules 1–6.

For $p, q \in \mathcal{P}$ let $p \twoheadrightarrow q$ mean that $|p|^x$ and $|q|^x$ exist and for every $w \in \mathcal{P}$, if $|w|^p$ exists, then $|w|^q$ exists. Let $p \twoheadleftarrow\!\!\!\!\twoheadrightarrow q$ if and only if $p \twoheadrightarrow q$ and $q \twoheadrightarrow p$. Let $S = \{p \in \mathcal{P} : p \twoheadleftarrow\!\!\!\!\twoheadrightarrow x\}$. We will show $S \supset \mathcal{A}$ and derive then that $S = \mathcal{P}$ (which is equivalent to the main theorem, Theorem 8.1). We have that $x \in S$. Assume $p \in S$ and $r \in S$. The remainder of this section is to indicate part of how to show that $pr \in S$ (in order to prove $S \supset \mathcal{A}$).

Firstly, $|pr|$ exists: since $p, r \in S$ we have $x \twoheadleftarrow\!\!\!\!\twoheadrightarrow p \twoheadleftarrow\!\!\!\!\twoheadrightarrow r$ and thus $|r|^p$ exists so $|p|^x |r|^p = |pr|^p$ and by $p \twoheadleftarrow\!\!\!\!\twoheadrightarrow x$, $|pr|^x$ exists.

The next step in showing $pr \in S$ is to show that the existence of $|w|^x$ together with $x \twoheadleftarrow\!\!\!\!\twoheadrightarrow p \twoheadleftarrow\!\!\!\!\twoheadrightarrow r$ imply $|w|^{pr}$ exists. We have $|w|^p$ exists (from the assumption $p \twoheadleftarrow\!\!\!\!\twoheadrightarrow x$), so suppose without loss of generality that $|w|^p = pp_1p_2 \cdots p_{n-1} * p_n$ is $p$-normal. By $p \twoheadleftarrow\!\!\!\!\twoheadrightarrow r$ we may change bases to write $p_1 = |p_1|^r = rr_1r_2 \cdots r_{k-1} * r_k$. So

$$w = p[rr_1r_2 \cdots r_{k-1} * r_k]p_2 \cdots p_{n-1} * p_n = \big[(pr)(pr_1) \cdots (pr_{k-1}) * (pr_k)\big]p_2 \cdots p_{n-1} * p_n.$$

By induction we obtain that each $pr_i$ and $p_j$, $j \geq 2$, have a $pr$-DF representation, and if $* = \cdot$ then the sequence $\big[(pr)(pr_1) \cdots (pr_{k-1}) \cdot (pr_k)\big]p_2 \cdots p_{n-1} * p_n$ is $pr$-normal and we are done with the $x \twoheadrightarrow pr$ part of the proof. Thus we need only further consider the case $* = \circ$.

A general iteration fact is: suppose $v = ul_1l_2 \cdots l_m l_{m+1} l_{m+2} \cdots l_{t-1} * l_t$ satisfies the condition that for each $l_i$, $|l_i|^u$ exists and $|l_i|^u \leq_{\mathrm{L}} ul_1 \cdots l_{i-2}$ (with strict inequality if $i = t$ and $* = \circ$), but with one exception: for some $m < t$, the normality condition is replaced by the weaker condition $ul_1l_2 \cdots l_m \sqsupset^u l_{m+1}$. Then $|v|^u$ exists by repeated application of Theorem 5.6; namely $|ul_1l_2 \cdots l_{m+1}|^u$ exists and $|ul_1l_2 \cdots l_{m+1}|^u \sqsupset^u ul_1l_2 \cdots l_m$, and since $ul_1l_2 \cdots l_m \geq_{\mathrm{L}} l_{m+2}$, it is also the case that $|ul_1l_2 \cdots l_{m+1}|^u \sqsupset^u l_{m+2}$. Continue this for $ul_1l_2 \cdots l_m l_{m+1} \geq_{\mathrm{L}} l_{m+3}$, etc., to get that $|v|^u$ exists. This general iteration fact about horseshoe relations enables us to deal with the case $* = \circ$ in which

$$\begin{aligned} w &= \big[(pr)(pr_1) \cdots (pr_{k-1}) * (pr_k)\big]p_2 \cdots p_{n-1} * p_n \\ &= \big[(pr)(pr_1) \cdots (pr_{k-1}) \circ (pr_k)\big]p_2 \cdots p_{n-1} * p_n = (pr)(pr_1) \cdots (pr_{k-1})(pr_kp_2)p_3 \cdots p_{n-1} * p_n. \end{aligned}$$

The last fact for this part of $x \twoheadrightarrow pr$ utilizes the lemma that allows us to consider a $pr$-DF word as having default $p$-DF rather than default $x$-DF so that given $x \twoheadleftarrow\!\!\!\!\twoheadrightarrow r \twoheadleftarrow\!\!\!\!\twoheadrightarrow p$ and $R \in r$-DF, then "$pR \sqsupset^{pr/p} p$", where $\sqsupset^{pr/p}$ is a $pr$-horseshoe with default value $|s|^p$ rather than $|s|^x$ for $|s|^{pr/p}$ when $s <_{\mathrm{L}} pr$.

## 5.2. Rules for representing by normal terms certain products and compositions of normal terms

In this section "normal" means $a_0$-normal.

### *Rule 1.*

Let $a = a_0a_1 \cdots a_n$ be normal. Suppose that $b \leq_{\mathrm{L}} a_0a_1 \cdots a_{n-1}$. Then $ab = a_0a_1 \cdots a_nb$ is normal. Moreover, $\mathrm{collapse}(a \circ b)$ is the normal sequence representing $a \circ b$.

### *Rule 2.*

Suppose $a = a_0a_1 \cdots a_{n-1}a_n$ and $b = a_0a_1 \cdots a_{n-1}b_n$ are normal, and that $a_n \circ b_n \leq_{\mathrm{L}} a_0 \cdots a_{n-2}$. Then the normal terms representing $ab$ and $a \circ b$ are

$$ab = a_0a_1 \cdots a_{n-1}(a_nb_n), \qquad a \circ b = a_0 \cdots a_{n-1}(a_n \circ b_n).$$

### *Rule 3.*

Suppose $a = a_0a_1 \cdots a_{n-1}a_n$ and $b = a_0a_1 \cdots a_{n-1} \circ b_n$ are normal and $a_n \circ b_n \leq_{\mathrm{L}} a_0a_1 \cdots a_{n-2}$. Let $u = a_0a_1 \cdots a_{n-1}$, $v = a_n$, and $w = b_n$. Then the normal terms representing $ab$ and $a \circ b$ are respectively $u(v \circ w)u \circ u(vw) = a_0 \cdots a_{n-1}(a_n \circ b_n)(a_0 \cdots a_{n-1}) \circ a_0 \cdots a_{n-1}(a_nb_n)$ and $\mathrm{collapse}(u \circ (v \circ w)) = \mathrm{collapse}(a_0 \cdots a_{n-1} \circ (a_n \circ b_n))$, as seen below.

$$ab = uv(u \circ w) = uv(uw \circ u) = uv(uwu \circ uw) = uv(uwu) \circ (uv(uw)) = u(v \circ w)u \circ u(vw).$$

Using the normal representation for $ab$ found above,

$$\begin{aligned} a \circ b = ab \circ a &= u(v \circ w)u \circ (u(vw) \circ uv) = u(v \circ w)u \circ (u(vw \circ v)) \\ &= u(v \circ w)u \circ (u(v \circ w)) = u(v \circ w) \circ u = \mathrm{collapse}(u \circ (v \circ w)). \end{aligned}$$

### *Rule 4.*

Let $a = a_0a_1\cdots a_{n-1}a_n$ and $b = a_0a_1\cdots a_{n-1}b_nb_{n+1}\cdots b_{k-1}b_k$ be normal with $k > n$ and $a_n \circ b_n \leq_{\mathrm{L}} a_0a_1\cdots a_{n-2}$. Then

$$\begin{aligned} ab &= a(a_0\cdots a_{n-1}b_nb_{n+1}\cdots b_{k-1}b_k) = a[a_0\cdots a_{n-1}b_nb_{n+1}](ab_{n+2})\cdots(ab_{k-1})(ab_k) \\ &= \big(a\circ(a_0a_1\cdots a_{n-1}b_n)\big)\,b_{n+1}(ab_{n+2})\cdots(ab_{k-1})(ab_k) \\ &= (a_0\cdots a_{n-1}a_n\circ a_0\cdots a_{n-1}b_n)b_{n+1}(ab_{n+2})\cdots(ab_{k-1})(ab_k) = a_0\cdots a_{n-1}(a_n\circ b_n)b_{n+1}(ab_{n+2})\cdots(ab_{k-1})(ab_k). \end{aligned}$$

The facts that $a_n\circ b_n \leq_{\mathrm{L}} a_0\cdots a_{n-2}$ and that $ab_i \leq_{\mathrm{L}} a(a_0\cdots a_n\cdots b_{i-2})$ for each $i \geq n+2$ give that the final term is normal. The normal representation of $a\circ b$ is computed as follows.

$$\begin{aligned} a\circ b = (a_0a_1\cdots a_n)\circ(a_0\cdots a_{n-1}b_nb_{n+1}\cdots b_{k-1}b_k) &= a(a_0\cdots a_{n-1}b_nb_{n+1}\cdots b_{k-1}b_k)\circ a \\ &= a_0\cdots a_{n-1}(a_n\circ b_n)b_{n+1}(ab_{n+2})\cdots(ab_{k-1})(ab_k)\circ a. \end{aligned}$$

### *Rule 5.*

Suppose $a = a_0a_1\cdots a_{n-1}a_n$ and $b = a_0a_1\cdots a_{n-1}b_nb_{n+1}\cdots b_{k-1}\circ b_k$ are normal with $a_n\circ b_n \leq_{\mathrm{L}} a_0a_1\cdots a_{n-2}$. The normal representation of $ab$ (when $k > n+1$) is (similarly to Rule 4),

$$ab = a_0\cdots a_{n-1}(a_n\circ b_n)b_{n+1}(ab_{n+2})\cdots(ab_{k-1})\circ(ab_k)$$

and that of $a\circ b$ is

$$\begin{aligned} a\circ b &= (a_0a_1\cdots a_{n-1}a_n)\circ(a_0\cdots a_{n-1}b_nb_{n+1}\cdots b_{k-1}\circ b_k) = a(a_0\cdots a_{n-1}b_nb_{n+1}\cdots b_{k-1}\circ b_k)\circ a \\ &= a_0\cdots a_{n-1}(a_n\circ b_n)b_{n+1}(ab_{n+2})\cdots(ab_{k-1})\circ ab_k\circ a = a_0\cdots a_{n-1}(a_n\circ b_n)b_{n+1}(ab_{n+2})\cdots(ab_{k-1})\circ(a\circ b_k). \end{aligned}$$

When $k = n+1$,

$$\begin{aligned} ab &= a_0\cdots a_{n-1}a_n(a_0\cdots a_{n-1}b_n\circ b_{n+1}) = a_0\cdots a_{n-1}a_n(a_0\cdots a_{n-1}b_nb_{n+1}\circ a_0\cdots a_{n-1}b_n) \\ &= a_0\cdots a_{n-1}a_n(a_0\cdots a_{n-1}b_nb_{n+1})\circ\big(a_0\cdots a_{n-1}a_n(a_0\cdots a_{n-1}b_n)\big) = a_0\cdots a_{n-1}(a_n\circ b_n)b_{n+1}\circ(a_0\cdots a_{n-1}(a_nb_n)). \end{aligned}$$

Recalling that $a_nb_n <_{\mathrm{L}} a_n\circ b_n \leq_{\mathrm{L}} a_0\cdots a_{n-2}$ by hypothesis, we see that the final expression is normal. For the normal representation of $a\circ b$ we have

$$\begin{aligned} a\circ b &= a_0\cdots a_{n-1}a_n\circ a_0\cdots a_{n-1}b_n\circ b_{n+1} = a_0\cdots a_{n-1}a_n(a_0\cdots a_{n-1}b_nb_{n+1})\circ a_0\cdots a_{n-1}a_n \\ &= a_0\cdots a_{n-1}(a_n\circ b_n)b_{n+1}\circ a_0\cdots a_{n-1}(a_nb_n)\circ a_0\cdots a_{n-1}a_n \\ &= a_0\cdots a_{n-1}(a_n\circ b_n)b_{n+1}\circ(a_0\cdots a_{n-1}(a_n\circ b_n)) = \mathrm{collapse}(a_0\cdots a_{n-1}(a_n\circ b_n)\circ b_{n+1}). \end{aligned}$$

### *Rule 6.*

Suppose $a = a_0a_1\cdots a_{n-1}\circ a_n$ is normal and $a_n\circ b \leq_{\mathrm{L}} a_0\cdots a_{n-2}$. Then the normal terms representing of $ab$ and $a\circ b$ are

$$ab = a_0a_1\cdots a_{n-1}(a_nb), \qquad a\circ b = \mathrm{collapse}(a_0a_1\cdots a_{n-1}\circ(a_n\circ b)).$$

## 5.3. Definition and basic facts about division form and a hereditary version of the division algorithm

### *Definition 5.7.*

Define DF, the set of division form terms, to be the set of hereditarily $x$-normal terms. Namely, $\mathrm{DF} \subseteq P$ is the smallest set such that $x \in \mathrm{DF}$ and

- If $a_1,\ldots,a_n \in \mathrm{DF}$, and $w = xa_1\cdots a_{n-1}*a_n$ is normal, then $w \in \mathrm{DF}$.

For $\mathrm{D} = \mathrm{DF}$ or one of the variants of division form defined below, proofs that the statement $\Phi(w)$ holds "by induction on $w \in \mathrm{D}$" involve defining a partial ordering $<_{\mathrm{D}}$ on D which is well-founded, i.e., there are no infinite descending sequences $w_0, w_1, \ldots$ with each $w_{i+1} <_{\mathrm{D}} w_i$; $<_{\mathrm{D}}$ is chosen so that for every $w$, $\Phi(w)$ holds if $\Phi(v)$ holds for all $v <_{\mathrm{D}} w$.

A simple case is when, like DF, D is defined as the union of $\mathrm{D}_0 \subset \mathrm{D}_1 \subset \ldots$ and the induction proceeds on rank $w =$ the least $\alpha$ with $w$ in $\mathrm{D}_\alpha$. Examples of (not necessarily compatible) bases are:

$$v <_{\mathrm{D}} w \text{ if and only if } \operatorname{length}(v) < \operatorname{length}(w), \qquad v <_{\mathrm{D}} w \text{ if } v \text{ is a proper subterm of } w.$$

Note that the length of the induction can be a (countable) $\alpha > \omega$. For suppose we are choosing our induction to have $p <_{\mathrm{D}} q$ whenever (for $p, q \in \mathrm{D}$) (a) $p$ is a proper (literal) subterm of $q$, or (b) $q = s \circ t$ and $p$ is an $I_k(s,t)$ for some $k$. For each D there is a well-founded partial ordering defined by the transitivization of conditions (a) and (b) above for $p, q \in \mathrm{D}$. Let $<_{\mathrm{D}}$ be the transitivization of all the $\langle p, q\rangle$ satisfying (a) or (b). To show this $<_{\mathrm{D}}$ is well-founded, suppose $r_0, r_1, r_2, \ldots$ is infinite with every $\langle r_{i+1}, r_i\rangle$ satisfying (a) or (b). By the well-foundedness of (a), some $r_i$ has no proper subterm $r_j$ for any $j > i$. Thus $r_i$ must be of the form $s \circ t$ with $r_{i+1}$ either of the form $I_l(s,t)$ for some $l < k$, or a subterm of $s$ or of $t$. These last cases must eventually be reached; some $r_j$ is a subterm of $s$ or of $t$ so is a subterm of $s \circ t = r_i$, a contradiction.

Using this well-founded partial ordering we define an ordinal-valued rank function satisfying rank $p <$ rank $q$ if and only if $p <_{\mathrm{D}} q$. This rank function is used for doing proofs by induction on the various division forms. Some examples of the D's other than DF are $p$-DF, $(p,r)$-DF and $pr/p$-DF, defined below.

### Definition 5.8.

For $p \in \mathcal{P}$, $p \in \mathrm{DF}$, define $p$-DF $\subseteq P$, the set of $p$-division form terms, to be the smallest set such that

- if $w \leq_{\mathrm{L}} p$ then $w \in p$-DF if and only $w \in \mathrm{DF}$, and
- if $w >_{\mathrm{L}} p$, $w = pa_1a_2\cdots a_{n-1} * a_n$, $w \in p$-DF if and only if $w$ is $p$-normal and each $a_i \in p$-DF.

For $p, r \in \mathrm{DF}$, we define $(p,r)$-division form using the concept of $(p,r)$-normal terms. A term $w \in P$ is $(p,r)$-normal if it is $p$-normal with one exception: terms of the form $p \circ q$ are $p$-normal but not $(p,r)$-normal; the $(p,r)$-normal form of these terms is $pq \circ p$. For all other terms (e.g., $w = pa_1 \circ a_2$), $p$-normal and $(p,r)$-normal are identical.

### Definition 5.9.

For $p, r \in \mathrm{DF}$, define the set $(p,r)$-DF $\subseteq P$ of $(p,r)$-division form terms, to be the smallest set such that

- if $w \leq_{\mathrm{L}} p$ and $w \in \mathrm{DF}$, then $w$ is in $(p,r)$-DF, and
- if $w = pa_1\cdots a_{n-1} * a_n >_{\mathrm{L}} p$ is $(p,r)$-normal, $a_1 \in r$-DF, and $a_2, \ldots, a_n \in (p,r)$-DF, then $w$ is in $(p,r)$-DF.

We now prove lemmas that illustrate that the $<_{\mathrm{L}}$-ordering between two normal terms is their iterated lexicographic ordering, thus there is at most one $p$-DF representation of any $w \in \mathcal{P}$. (Similarly there is at most one $(p,r)$-DF representation of $w \in \mathcal{P}$.)

### Lemma 5.10.

*If $u, v \in p$-DF and $u \neq v$, then $u <_{\mathrm{L}} v$ or $v <_{\mathrm{L}} u$.*

The proof is by induction on $\max\{\operatorname{length}(u), \operatorname{length}(v)\}$ using Corollary 5.4.

### Lemma 5.11.

*For $w, p \in P$, there is at most one $p$-DF term representing $w$. (Similarly there is at most one $(p,r)$-DF representation of $w \in \mathcal{P}$.)*

The proof follows immediately from Lemma 5.10 and the irreflexivity of $<_{\mathrm{L}}$.

### *Definition 5.12.*

For $w, p, r \in \mathcal{P}$, $|w|^p$ is the (unique by Lemma 5.11) member of $p$-DF representing $w$, when such a term exists. Let $|w| = |w|^x$. Similarly let $|w|^{p,r}$ denote the $(p, r)$-DF of $w \in \mathcal{P}$, when it exists.

A restatement of the division algorithm is: *if $p <_{\mathrm{L}} q$, there is a hereditarily $p$-normal sequence representing $q$.* A consequence of the division algorithm is that $p \circ q$ is the $<_{\mathrm{L}}$-least upper bound of $I_n(p, q)$, $n \geq 1$. (We write $I_n$ for $I_n(p, q)$ when the meaning is clear from context.) In particular, the associated sequence of $p \circ q$ is $I_\infty = p, pq, pqp, pqp(pq), \ldots = I_1, I_2, I_3, I_4, \ldots$, where for $k \geq 3$, $I_k = I_{k-1}I_{k-2}$. It is clear that $p \circ q$ is an upper bound of $I_\infty$. To see that it is least, note that the associated sequence of every upper bound $pu_1u_2 \ldots u_{n-1} * u_n$ of the iterates, if it lexicographically dominated every iterate $I_k(p, q)$, would lexicographically dominate $I_\infty(p, q)$ and therefore would lexicographically dominate (or equal) $p \circ q$. Therefore $p \circ q$ is the least upper bound of the iterates $I_n(p, q)$.

In Section 8 we will prove the full hereditary form of the division algorithm. Namely, we will prove: *for all $p, q \in \mathcal{P}$, $|p|^q$ exists.* For completeness we include the following definition and theorem. Theorem 5.13 gives that not only are $<_{\mathrm{L}}$ and $<_{\mathrm{L}}$-lexicographic comparison of associated sequences of division form terms equivalent orderings on $p$-DF, but also that every subterm $u$ of $v$ satisfies $u <_{\mathrm{L}} v$. Laver derived Theorem 5.13 from a normal form theorem [14]. McKenzie used the linearity of $<_{\mathrm{L}}$ on $\mathcal{A}$ and $\mathcal{P}$ to prove the equivalence of $<_s$ and $<_{\mathrm{L}}$.

For $u, v \in p$-DF denote: $u <_{\mathrm{Lex}} v$ if the associated sequence of $u$ is lexicographically less than the associated sequence of $v$ (see Corollary 5.5); $u <_s v$ if $u$ is a proper subterm of $v$.

### *Theorem 5.13.*

*For all $u, v \in p$-DF, $u <_s v$ iff $u <_{\mathrm{L}} v$ iff $u <_{\mathrm{Lex}} v$.*

## 6. The $\sqsupset$ relation

To show that for all $p, q \in \mathcal{P}$, $|p|^q$ exists, it suffices to show that if $|u|$ and $|v|$ exist (more generally $|u|^p$ and $|v|^p$ exist), then $|uv|$ and $|u \circ v|$ exist ($|uv|^p$ and $|u \circ v|^p$ exist). We will state and prove a result (Theorem 6.5) which gives conditions (namely "$u \sqsupset^p v$") under which the existence of $|uv|^p$ and $|u \circ v|^p$ follow by a direct inductive argument. Similar results hold for other types of $\sqsupset$ relations.

In what follows we are referring to terms in $P$ but will occasionally write equality (e.g., $|uv|^p \circ u = u \circ v$) when terms are not literally equal in $P$ but are equal in $\mathcal{P}$. This occurs only in the verification of normality conditions, which are unaffected by whether terms are considered to be in $P$ or $\mathcal{P}$. Such equivalences could be replaced by, in the case of the example, $|uv|^p \circ u \equiv uv \circ u \equiv u \circ v$. For the sake of notational and cognitive ease, we (mis)use equality instead.

### 6.1. Definitions

We now define the relation $\sqsupset^p$ for each $p \in$ DF. The atomic case $p = x$ is like the general case, with the exception that (i) does not occur in the atomic case.

### *Definition 6.1.*

For $p \in$ DF, $u, v \in p$-DF, define $u \sqsupset^p v$ by induction on $u$ in $p$-DF then on $v$ in $p$-DF.

(i) If $u <_{\mathrm{L}} p$, then $u \sqsupset^p v$ if and only if $u \sqsupset^x v$ and $u \circ v \leq_{\mathrm{L}} p$.

(ii) If $u = p$, then $u \sqsupset^p v$ for any $v$ in $p$-DF.

(iii) If $u = pa$, then $u \sqsupset^p v$ whenever $v \leq_{\mathrm{L}} p$ or $v = pb_1 \cdots b_{k-1} * b_k$ with $a \sqsupset^p b_1$.

(iv) If $u = p \circ a$, then $u \sqsupset^p v$ if and only if $a \sqsupset^p v$.

(v) If $u = pa_1 \cdots a_n$ with $n \geq 2$, then $u \sqsupset^p v$ if and only if either $v \leq_{\mathrm{L}} pa_1 \cdots a_{n-1}$ or $v = pa_1 \cdots a_{n-1} b_n b_{n+1} \cdots b_{k-1} * b_k$ with $a_n \sqsupset^p b_n$ and $a_n \circ b_n \leq_{\mathrm{L}} pa_1 \cdots a_{n-2}$.

(vi) If $u = pa_1 \cdots a_{n-1} \circ a_n$, $n \geq 2$ then $u \sqsupset^p v$ if and only if $a_n \sqsupset^p v$ and $a_n \circ v \leq_{\mathrm{L}} pa_1 \cdots a_{n-2}$.

Note that in the case $p = x$ it is impossible to satisfy the last condition of (vi) when $u = xa_1 \circ a_2$, though this poses no problems as no term of the form $u = xa_1 \circ a_2$ is in DF. In the sequel we will write $\sqsupset$ for $\sqsupset^x$. There exists an analogous definition for the relation $\sqsupset^{p,r}$ for $(p,r)$-DF terms.

### Definition 6.2.

For $u, v \in (p,r)$-DF define $u \sqsupset^{p,r} v$ by induction on $u$ then on $v$.

(a) If $u <_{\mathrm{L}} p$, then $u \sqsupset^{p,r} v$ if and only if $u \sqsupset v$ and $u \circ v \leq_{\mathrm{L}} p$.

(b) If $u = p$, $u \sqsupset^{p,r} v$ for any $v \leq_{\mathrm{L}} p, r$ with $v \in (p,r)$-DF.

(c) For $u = pq$, $u \sqsupset^{p,r} v$ if and only if $v \leq_{\mathrm{L}} p$ or if $v = pq' v_1 \cdots v_{m-1} * v_m$ where $q \sqsupset^r q'$.

(d) For $u = pa \circ p$, $u \sqsupset^{p,r} v$ if and only if $a \sqsupset^r v$ and $p \sqsupset^{p,r} v$.

(e) For $n \geq 1$, $u = pqu_1 \cdots u_{n-1} u_n$, $u \sqsupset^{p,r} v$ if and only if $v \leq_{\mathrm{L}} pqu_1 \cdots u_{n-1}$ or $v = pqu_1 \cdots u_{n-1} v_n v_{n+1} \cdots v_{m-1} * v_m$ with $u_n \sqsupset^{p,r} v_n$ and $u_n \circ v_n <_{\mathrm{L}} pqu_1 u_2 \cdots u_{n-2}$.

(f) For $n \geq 1$, $u = pqu_1 \cdots u_{n-1} \circ u_n$ (with $u_1 \leq_{\mathrm{L}} p$ if $n = 1$), $u \sqsupset^{p,r} v$ if and only if $u_n \sqsupset^{p,r} v$ and $u_n \circ v <_{\mathrm{L}} pqu_1 \cdots u_{n-2}$ (in the case of $n = 1$, this condition is $u_1 \circ v <_{\mathrm{L}} p$).

## 6.2. Basic $\sqsupset$ lemma

The $\sqsupset^p$ relation is not closed upward on the left, i.e., $u' >_{\mathrm{L}} u$ and $u \sqsupset^p v$ together do not imply $u' \sqsupset^p v$, but downward closure of $\sqsupset^p$ does occur on the right.

### Lemma 6.3.

*If $u, v, w \in p$-DF, $u \sqsupset^p v$, and $w <_{\mathrm{L}} v$, then $u \sqsupset^p w$.*

***Proof.*** By induction on the $p$-DF rank of $u$ then the $p$-DF rank of $v$. One first proves the lemma when $p = x$, which proceeds as below (except for case (i)) and is left to the reader.

**Case (i):** $u <_{\mathrm{L}} p$. As $w <_{\mathrm{L}} v$, $u \circ w <_{\mathrm{L}} u \circ v <_{\mathrm{L}} p$. Together with the $x$-DF version of this lemma which gives $u \sqsupset w$, we have $u \sqsupset^p w$.

**Case (ii):** $u = p$. $u \sqsupset^p w$ trivially.

**Case (iii):** $u = pa$. The result is clear when $w <_{\mathrm{L}} v \leq_{\mathrm{L}} p$, so suppose (without loss of generality) that $w = pc_1 \cdots c_{m-1} * c_m <_{\mathrm{L}} v = pb_1 \cdots b_{k-1} * b_k$ and $a \sqsupset^p b_1$. Then $c_1 \leq_{\mathrm{L}} b_1$, so $a \sqsupset^p c_1$ by the induction hypothesis, establishing this case.

**Case (iv):** $u = p \circ a$. By the induction hypothesis $a \sqsupset^p v$ implies $a \sqsupset^p w$, giving $u \sqsupset^p w$.

**Case (v):** $u = pa_1 \cdots a_{n-1} a_n$, $n > 1$. If $w <_{\mathrm{L}} v \leq_{\mathrm{L}} pa_1 \cdots a_{n-1}$, this case is clear, so suppose without loss of generality that

$$w = pa_1 \cdots a_{n-1} c_n c_{n+1} \cdots c_{m-1} * c_m <_{\mathrm{L}} v = pa_1 \cdots a_{n-1} b_n b_{n+1} \cdots b_{k-1} * b_k$$

with $a_n \sqsupset^p b_n$ and $a_n \circ b_n \leq_{\mathrm{L}} pa_1 \cdots a_{n-2}$. Then, as $c_n \leq_{\mathrm{L}} b_n$, we have $a_n \circ c_n \leq_{\mathrm{L}} pa_1 \cdots a_{n-2}$. Furthermore, $a_n \sqsupset^p c_n$ follows from the induction hypothesis, thus $u \sqsupset^p w$.

**Case (vi):** $u = pa_1 \cdots a_{n-1} \circ a_n$, $n > 1$. As $w <_{\mathrm{L}} v$, $a_n \circ w <_{\mathrm{L}} a_n \circ v \leq_{\mathrm{L}} pa_1 \cdots a_{n-2}$, so $a_n \sqsupset^p w$ follows from the induction hypothesis and $u \sqsupset^p w$ in this case as well. □

### Lemma 6.4.

*If $u, v, w \in (p,r)$-DF, $u \sqsupset^{p,r} v$, and $w <_{\mathrm{L}} v$, then $u \sqsupset^{p,r} w$.*

***Proof.*** By induction on the rank of $u$ then that of $v$ in $(p,r)$-DF, analogous to the proof of Lemma 6.3.

**Case (a):** $u <_L p$. Then $u \sqsupset^{p,r} v$ if and only if $u \sqsupset v$ and $u \circ v <_L p$. Note that $u \sqsupset v$ implies that $|v|^{p,r} = |v|^x$ and thus it must be the case that $v <_L p$ and that $u \circ v <_L p$. By Lemma 6.3, $u \sqsupset w$. As $u \circ w <_L u \circ v <_L p$, we have $u \sqsupset^{p,r} w$.

**Case (b):** $u = p$. Then $v \in (p,r)$-DF must also be in $r$-DF, as $|v|^{p,r} = |v|^r$ is only possible when $v \leq_L p, r$ and $|v|^{p,r} = |v|^r = |v|$. Furthermore, $w <_L v \leq_L p, r$ implies that $w \leq_L p, r$. Therefore (again by the $x$-DF version of Lemma 6.3), $u \sqsupset v$ implies $u \sqsupset w$, and $u \sqsupset^{p,r} w$ follows.

**Case (c):** $u = pq$. If $w <_L v \leq_L p$, then $u \sqsupset^{p,r} w$ is clear. Thus suppose that $v = pq'b_1 \cdots b_{n-1} * b_n$ where $q \sqsupset^r q'$ and $w = p\overline{q}c_1 \cdots c_{k-1} * c_k$ with $w <_L v$. In the case that $\overline{q} = q'$, then $q \sqsupset^r q'$ implies that $q \sqsupset^r \overline{q}$, and hence that $u \sqsupset^{p,r} w$. If $\overline{q} <_L q'$, then Lemma 6.3 for $r$-DF plus $q \sqsupset^r q'$ yields $q \sqsupset^r \overline{q}$, and, as $p(q \circ \overline{q}) \sqsupset^{p,r} c_1$ and $pq \sqsupset^{p,r} c_i$ for $1 \leq i \leq k$ by the induction hypothesis, we have that $u \sqsupset^{p,r} w$.

**Case (d):** $u = pq \circ p$. Then $u \sqsupset^{p,r} v$ implies $v \leq_L p, r$, hence $|v|^{p,r} = |v|^r = |v|$, so $v \in$ DF. As $w \leq_L v$, we also have that $w \in$ DF, so by the $r$-DF version of Lemma 6.3, we have that $q \sqsupset^r v$ implies $q \sqsupset^r w$. Furthermore, by the induction hypothesis, $p \sqsupset^{p,r} v$ implies $p \sqsupset^{p,r} w$, and hence $u \sqsupset^{p,r} v$ implies that $u \sqsupset^{p,r} w$.

**Case (e):** $u = pqa_1 \cdots a_{n-1}a_n$. When $w <_L v \leq_L pqa_1 \cdots a_{n-1}$, it is clear that $u \sqsupset^{p,r} w$. Therefore suppose $v = pqa_1 \cdots a_{n-1}b_n \cdots b_{m-1} * b_m$, with $a_n \sqsupset^{p,r} b_n$ and $a_n \circ b_n \leq_L pqa_1 \cdots a_{n-2}$.

**Case (e$_1$):** $w \leq_L pqa_1 \cdots a_{n-1}$. Then $u \sqsupset^{p,r} w$ is trivial.

**Case (e$_2$):** $w = pqa_1 \cdots a_{n-1}c_n \cdots c_{k-1} * c_k$ with $c_l <_L b_l$ for the first $l \geq n$ for which $c_l$ and $b_l$ differ. When $c_n = b_n$, it is clear that $a_n \sqsupset^{p,r} c_n$. When $c_n <_L b_n$ we have $a_n \sqsupset^{p,r} c_n$ by the induction hypothesis and $a_n \circ c_n <_L a_n \circ b_n \leq_L pqa_1 \cdots a_{n-1}$. Thus $u \sqsupset^{p,r} w$.

**Case (f):** $u = pqa_1 \cdots a_{n-1} \circ a_n$ and $a_n \sqsupset^{p,r} v$. Then $a_n \sqsupset^{p,r} w$ by induction, and $a_n \circ w <_L a_n \circ v <_L pqa_1 \ldots a_{n-2}$. □

## 6.3. The main $\sqsupset$ theorem

As above, we prove the $p$-DF version of the theorem supposing the (analogous) DF version. We include the $(p,r)$-DF version for completeness, though it is highly similar to the $p$-DF version.

***Theorem 6.5.***
*Suppose $u, v \in p$-DF and $u \sqsupset^p v$. Then $|uv|^p$ and $|u \circ v|^p$ exist, and $|uv|^p \sqsupset^p u$.*

***Proof.*** By induction on the $p$-DF rank of $u$ then that of $v$. The proof of the theorem for $x$-DF proceeds as below except the case (i), which is unnecessary in the case $p = x$.

**Case (i):** $u <_L p$. Then by the definition of $u \sqsupset^p v$ we have $u \sqsupset v$ and $u \circ v \leq_L p$. Therefore, by the base case of the induction, $u \sqsupset v$ implies $|uv| = |uv|^p$ and that $|u \circ v| = |u \circ v|^p$ exist and furthermore that $|uv| \sqsupset u$ (and thus $|uv|^p \sqsupset^p u$).

**Case (ii):** $u = p$. $|uv|^p = uv$ and $|u \circ v|^p = u \circ v$, and $uv = pv \sqsupset^p p$, which holds by (iii) of the definition of $\sqsupset^p$.

**Case (iii):** $u = pa$. If $v \leq_L p$, then $|uv|^p$ and $|u \circ v|^p$ exist by Rule 1, $|uv|^p = uv$ and $uv \sqsupset^p u$ is clear. Therefore suppose $v = pb_1 \cdots b_{k-1} * b_k$ and $a \sqsupset^p b_1$. If $k = 1$ and $* = \cdot$, then $|uv|^p$ and $|u \circ v|^p$ exist by Rule 2. If $k = 1$ and $* = \circ$, $|uv|^p$ and $|u \circ v|^p$ exist by Rule 3. If $k > 1$ the existence of $|uv|^p$ and $|u \circ v|^p$ follow from the induction hypothesis and Rules 4 (if $* = \cdot$) and 5 if $(* = \circ)$.

If $k = 1$ and $* = \cdot$, $|uv|^p = p|ab_1|^p \sqsupset^p pa$ as $|ab_1| \sqsupset^p a$ by the induction hypothesis. When $k = 1$ and $* = \circ$, Rule 3 gives that $|uv|^p = p|a \circ b_1|^p p \circ p|ab_1|^p$. By the case $k = 1$ and $* = \cdot$, we have that $p|ab_1|^p \sqsupset^p pa$, the condition necessary to assert that $|uv|^p \sqsupset^p u$.

When $k > 1$, $|uv|^p = p|a \circ b_1|^p|b_2|^p|pab_3|^p \cdots |pab_{k-1}|^p * |pab_k|^p$. For each $i > 2$, $|pab_i|^p$ exists by the induction hypothesis and $|uv|^p \sqsupset^p u$ is clear by inspection, utilizing, when $* = \circ$, that $|pab_k|^p \sqsupset^p pa$ by the induction hypothesis.

**Case (iv):** $u = p \circ a$. Then $a \sqsupset^p v$ by definition of $\sqsupset^p$, so $|uv|^p = p|av|^p$, $|u \circ v|^p = p \circ |a \circ v|^p$, and $|uv|^p = p|av|^p \sqsupset^p p \circ a$ as $|av|^p \sqsupset^p a$ by the induction hypothesis.

**Case (v):** $u = pa_1 \cdots a_n$, $n > 1$. If $v \leq_{\mathrm{L}} pa_1 \cdots a_{n-1}$, then $|uv|^p$, $|u \circ v|^p$ exist by Rule 1 and the induction hypothesis, and $|uv|^p = uv \sqsupset^p u$ follows the definition. If $v = pa_1 \cdots a_{n-1} b_n \cdots b_{k-1} * b_k$, then $|uv|^p$ and $|u \circ v|^p$ are as given by Rules 4 (if $* = \cdot$) and 5 (if $* = \circ$). $|uv|^p \sqsupset^p u$ follows directly if $* = \cdot$ and from the induction hypothesis applied to $|ub_k| \sqsupset^p u$ if $* = \circ$.

**Case (vi):** $u = pa_1 \cdots a_{n-1} \circ a_n$, $n > 1$. Then $a_n \sqsupset^p v$ implies $|a_n v|^p$ and $|a_n \circ v|^p$ exist and that $|a_n v|^p \sqsupset^p a_n$. Thus $|uv|^p$ and $|u \circ v|^p$ exist by Rule 6 and the induction hypothesis. Furthermore $|uv|^p = pa_1 \cdots a_{n-1} |a_n v|^p \sqsupset^p u$ as $|a_n v|^p \sqsupset^p a_n$. □

### *Theorem 6.6.*

*Suppose $u, v \in (p,r)$-DF and $u \sqsupset^{p,r} v$. Then $|uv|^{p,r}$ and $|u \circ v|^{p,r}$ exist, and $|uv|^{p,r} \sqsupset^{p,r} u$.*

***Proof.*** By induction on the $(p,r)$-DF rank $u$ and then on $v$.

**Case (a):** $u <_{\mathrm{L}} p$. Then $u \sqsupset^{p,r} v$ implies that $u \sqsupset v$ and $u \circ v <_{\mathrm{L}} p$. Thus $|uv|^{p,r} = |uv|$ and $|u \circ v|^{p,r} = |u \circ v|$ exist. Furthermore, $uvu <_{\mathrm{L}} uv \circ u = u \circ v <_{\mathrm{L}} p$. By Theorem 6.5, $|uv| \sqsupset u$, so $|uv|^{p,r} \sqsupset^{p,r} u$.

**Case (b):** $u = p$. Then $u \sqsupset^{p,r} v$ implies that $v$ is a term in DF such that $v <_{\mathrm{L}} p, r$. Thus $p|v|^r = p|v|$ is in $(p,r)$-DF and, as $u = p$, $u \circ v = p \circ v = pv \circ p$ (which is in $(p,r)$-DF). As $|uv|^{p,r} = pv$, and $pv \sqsupset^{p,r} w$ for all $w \leq_{\mathrm{L}} p$ by the definition of $\sqsupset^{p,r}$ we have $|uv|^{p,r} \sqsupset^{p,r} u$.

**Case (c):** $u = pq$ and either $v \leq_{\mathrm{L}} p$ or $v = pq' b_0 b_1 \cdots b_{n-1} * b_n$, where $q \sqsupset^r q'$.

**Subcase (c$_1$):** $v \leq_{\mathrm{L}} p$. Then $pqv = uv$ is in $(p,r)$-DF, as is $pq \circ v$, so $|uv|^{p,r}$ and $|u \circ v|^{p,r}$ exist. Furthermore, $|uv|^{p,r} \sqsupset^{p,r} u$ is clear.

**Subcase (c$_2$):** $v = pq'$. Then $uv = pq(pq') = p(qq')$. So $|uv|^{p,r} = p|qq'|^r$ and $|u \circ v|^{p,r} = p|q \circ q'|^r$. Both $|qq'|^r$ and $|q \circ q'|^r$ exist by $q \sqsupset^r q'$. The fact that $|qq'|^r \sqsupset^r q$ follows from the condition $q \sqsupset^r q'$ of the definition of $\sqsupset^{p,r}$ and Theorem 6.5. Finally $|uv|^{p,r} \sqsupset^{p,r} u$ follows from $|qq'|^r \sqsupset^r q$. Thus this case is established.

**Subcase (c$_3$):** $v = pq' b_1 \cdots b_{m-1} * b_m$ with $m \geq 1$. We have

$$|uv|^{p,r} = p|q \circ q'|^r |b_1|^{p,r} |pqb_2|^{p,r} \cdots |pqb_{m-1}|^{p,r} * |pqb_m|^{p,r},$$

where $|pqb_i|^{p,r}$ exists for $i > 2$ by the induction hypothesis. The displayed term is normal, and thus in $(p,r)$-DF. If $* = \cdot$ we have

$$|u \circ v|^{p,r} = |uv \circ u|^{p,r} = p|q \circ q'|^r |b_1|^{p,r} |pqb_2|^{p,r} \cdots |pqb_m|^{p,r} \circ pq.$$

If $* = \circ$, then

$$|u \circ v|^{p,r} = p|q \circ q'|^r |b_1|^{p,r} |pqb_2|^{p,r} \cdots |pqb_{m-1}|^{p,r} \circ |pq \circ b_m|^{p,r},$$

where the last term is obtained from $|pqb_m \circ pq|^{p,r} = |pq \circ b_m|^{p,r}$, which exists by the induction hypothesis. Furthermore, as $b_m <_{\mathrm{L}} pqb_0 \ldots b_{m-2}$, $|pq \circ b_m|^{p,r} <_{\mathrm{L}} p|q \circ q'|^r |b_1|^{p,r} |pqb_2|^{p,r} \cdots |pqb_{m-2}|^{p,r}$.

The representations of $u \circ v$ in each of the two cases above are normal, hence in $(p,r)$-DF, so it remains only to show that $|uv|^{p,r} \sqsupset u$. In the case that $*_v = \cdot$, this follows trivially from the definition. When $*_v = \circ$, it follows from the fact that $|pqb_m|^{p,r} \sqsupset^{p,r} pq$ together with $pqb_m \circ pq = pq \circ b_m <_{\mathrm{L}} p|q \circ q'|^r b_1 \cdots |pqb_{m-2}|^{p,r}$.

**Case (d):** $u = pq \circ p$. Then $u \sqsupset^{p,r} v$ if and only if $q \sqsupset^r v$ and $p \sqsupset^{p,r} v$. The existence of $|uv|^{p,r} = p|qv|^r$ and $|u \circ v|^{p,r} = p|q \circ v|^r \circ p$ follow from the $\sqsupset^r$ version of this theorem, so it remains only to show that $|uv|^{p,r} \sqsupset^{p,r} u$, i.e., that $p|qv|^r \sqsupset^{p,r} pq \circ p$. This follows from the definition of $\sqsupset^{p,r}$. (Use Rule 3 to get the $(p,r)$-DF representations of $p|qv|^r(p \circ q)$ and of $p|qv|^r \circ (p \circ q)$ noting that $pq \circ p \equiv p \circ q$.)

**Case (e):** $u = pqa_1 \cdots a_n$ with $n \geq 1$.

**Subcase (e$_1$):** $v <_{\mathrm{L}} pqa_1 \cdots a_{n-1}$. The existence of $|uv|^{p,r}$ and $|u \circ v|^{p,r}$ are clear. The fact that $|uv|^{p,r} \sqsupset^{p,r} u$ follows from the definition.

**Subcase (e$_2$):** $v = pqa_1 \cdots a_{n-1} b_n \cdots b_{m-1} * b_m$, $m \geq n$, with $a_n \sqsupset^{p,r} b_n$, $a_n \circ b_n <_{\mathrm{L}} pqa_1 \cdots a_{n-2}$. Then

$$|uv|^{p,r} = |pqa_1 \cdots a_{n-1}|^{p,r} |a_n \circ b_n|^{p,r} |b_{n+1}|^{p,r} |ub_{n+2}|^{p,r} \cdots |ub_{m-1}|^{p,r} * |ub_m|^{p,r},$$

which is normal by the induction hypothesis and the normality of $v$. Each subterm has a $(p,r)$-DF by the induction hypothesis. If $* = \cdot$, then we have

$$|u \circ v|^{p,r} = pqa_1 \cdots a_{n-1} |a_n \circ b_n|^{p,r} |b_{n+1}|^{p,r} |ub_{n+2}|^{p,r} \cdots |ub_m|^{p,r} \circ u.$$

If $* = \circ$,

$$|u \circ v|^{p,r} = pqa_1 \cdots a_{n-1} |a_n \circ b_n|^{p,r} |b_{n+1}|^{p,r} |ub_{n+2}|^{p,r} \cdots |ub_{m-1}|^{p,r} \circ |u \circ b_m|^{p,r}.$$

Either way $|u \circ v|^{p,r}$ exists by the induction hypothesis and the fact that $a_n \circ b_n <_{\mathrm{L}} pqa_1 \cdots a_{n-2}$, which yields normality. It is clear that $|uv|^{p,r} \sqsupset^{p,r} u$ when $* = \cdot$. When $* = \circ$, we use the fact that $|ub_m|^{p,r} \sqsupset^{p,r} u$ to see that $|uv|^{p,r} \sqsupset^{p,r} u$.

**Case (f):** $u = pqa_1 \cdots a_{n-1} \circ a_n$, $n > 1$. Then $v$ is such that $a_n \sqsupset^{p,r} v$ and $a_n \circ v <_{\mathrm{L}} pqa_1 \cdots a_{n-2}$. As $a_n \sqsupset^{p,r} v$ we have $|a_n v|^{p,r} \sqsupset^{p,r} a_n$ by the induction hypothesis. Therefore we have $|uv|^{p,r} = pqa_1 \cdots a_{n-1} |a_n v|^{p,r}$, $|u \circ v|^{p,r} = pqa_1 \cdots a_{n-1} \circ |a_n \circ v|^{p,r}$, and $|uv|^{p,r} \sqsupset^{p,r} u$. Thus the lemma is established. □

The following application of Theorems 6.5 and 6.6 is the key use of $\sqsupset$ results. In arguments below, it will be necessary to find $|w|^r$ given $|w|^s$ for various $w, r, s$ in $\mathcal{P}$. For some $w, r, s$ one encounters, in turning $|w|^s$ into $|w|^r$, a sequence which is almost $r$-normal except at one point where the coordinate is slightly larger than normality allows. The iterative use of the horseshoe results allows the sequence to be replaced by an $r$-normal sequence.

### Corollary 6.7.

*Suppose $w \in r$-DF and $wv_1 v_2 \cdots v_{n-1} * v_n$ is almost normal in the sense that*

(i) *$w, v_1, v_2, \ldots, v_n$ are in $r$-DF;*

(ii) *$w \sqsupset^r v_1$, $v_2 \leq_{\mathrm{L}} w$, $v_3 \leq_{\mathrm{L}} wv_1$, …, $v_n \leq_{\mathrm{L}} wv_1 \cdots v_{n-2}$ with strict inequality if $* = \circ$.*

*Then $|wv_1 \cdots v_{n-1} * v_n|^r$ exists.*

***Proof.*** By induction on $n$. $|wv_1|^r$ exists and $wv_1 \sqsupset^r v_2$ by Theorem 6.5. Suppose $|wv_1 \cdots v_{i-1}|^r \sqsupset^r v_i$. Then $|wv_1 \cdots v_{i-1} * v_i|^r$ exists and $|wv_1 \cdots v_{i-1} v_i|^r \sqsupset^r |wv_1 \cdots v_{i-1}|^r$ by Theorem 6.5. As $v_{i+1} \leq_{\mathrm{L}} |wv_1 \cdots v_{i-1}|^r$ by condition (ii), $|wv_1 \cdots v_i|^r \sqsupset^r v_{i+1}$ by Lemma 6.3. □

## 6.4. A variant of the main $\sqsupset$ theorem

Recall that for $w <_{\mathrm{L}} q$, $|w|^q = |w|^x$ when the latter exists. So the default form of $q$-DF for elements $w <_{\mathrm{L}} q$ is $x$-DF. We consider a variant of $q$-DF below when $p <_{\mathrm{L}} q$ where the default form for elements $w <_{\mathrm{L}} q$ is $p$-DF (as opposed to $x$-DF). Note that for $w <_{\mathrm{L}} p$, $|w|^{p,r} = |w|^p = |w|$.

### Definition 6.8.

Suppose $p, r, q \in$ DF and $p <_{\mathrm{L}} q$. Define $q/p$-DF to be the smallest set such that

- if $w \leq_{\mathrm{L}} q$ and $|w|^p$ exists then $|w|^p \in q/p$-DF, and
- if $w = qa_1 a_2 \cdots a_{n-1} * a_n$ is $q$-normal and each $a_i \in q/p$-DF, then $w \in q/p$-DF.

### Definition 6.9.

If $u, v \in q/p$-DF, define $u \sqsupset^{q/p} v$ by induction on the length of $u$ then the length of $v$.

(i) If $u <_{\mathrm{L}} q$, then $u \sqsupset^{q/p} v$ if and only if $u \sqsupset^p v$ and $u \circ v \leq_{\mathrm{L}} q$.

(ii) If $u = q$, then $u \sqsupset^{q/p} v$ for all $v$ in $q/p$-DF.

(iii) If $u = qa$, then $u \sqsupset^{q/p} v$ if and only if either $v \leq_{\mathrm{L}} q$ or if $v = qb_1 \cdots b_{m-1} * b_m$ where $a \sqsupset^{q/p} b_1$.

(iv) If $u = q \circ a$, then $u \sqsupset^{q/p} v$ if and only if $a \sqsupset^{q/p} v$.

(v) If $u = qa_1 \cdots a_n$, then $u \sqsupset^{q/p} v$ if and only if either $v \leq_{\mathrm{L}} qa_1 \cdots a_{n-1}$ or $v = qa_1 \cdots a_{n-1} b_n b_{n+1} \cdots b_{m-1} * b_m$, with $a_n \sqsupset^{q/p} b_n$ and $a_n \circ b_n \leq_{\mathrm{L}} qa_1 \cdots a_{n-2}$.

(vi) If $u = qa_1 \cdots a_{n-1} \circ a_n$, then $u \sqsupset^{q/p} v$ if and only if $a_n \sqsupset^{q/p} v$ and $a_n \circ v \leq_{\mathrm{L}} qa_1 \cdots a_{n-2}$.

### Theorem 6.10.

*If $u, v, w \in q/p$-DF, $u \sqsupset^{q/p} v$, and $w <_{\mathrm{L}} v$, then $u \sqsupset^{q/p} w$.*

**Proof.** By induction on the $q/p$-DF rank of $u$ then $v$.

**Case (i):** $u <_{\mathrm{L}} q$. The condition $u \circ v <_{\mathrm{L}} q$ implies that $v <_{\mathrm{L}} q$ (see Theorem 5.13), so $|v|^{q/p} = |v|^p$. Thus, if $w <_{\mathrm{L}} v$, then $w \in p$-DF with $u \circ w <_{\mathrm{L}} u \circ v$, so this case follows from Theorem 6.5.

**Case (ii):** $u = q$. Then $q \sqsupset^{q/p} w$ for all $w \in q/p$-DF (by Definition 6.9) so $u \sqsupset^{q/p} w$ certainly holds for $w <_{\mathrm{L}} v$.

**Case (iii):** $u = qa$. The case $w <_{\mathrm{L}} v \leq_{\mathrm{L}} q$ is trivial, so suppose $qc_1 \cdots c_{k-1} * c_k = w <_{\mathrm{L}} v = qb_1 \cdots b_{m_1} * b_m$. Then $c_1 \leq_{\mathrm{L}} b_1$ so $a \sqsupset^{q/p} c_1$ follows from the induction hypothesis applied to $a \sqsupset^{q/p} b_1$, and this case is established.

**Case (iv):** $u = q \circ a$. $u \sqsupset^{q/p} w$ if and only if $a \sqsupset^{q/p} w$, and $a \sqsupset^{q/p} w$ by the induction hypothesis.

**Case (v):** $u = qa_1 \cdots a_n$. The case $w <_{\mathrm{L}} v \leq_{\mathrm{L}} qa_1 \cdots a_{n-1}$ is trivial. If $w <_{\mathrm{L}} v = qb_1 \cdots b_{m_1} * b_m$, $u \sqsupset^{q/p} w$ follows as in case (iii), noting that $a_n \circ c_n <_{\mathrm{L}} a_n \circ b_n \leq_{\mathrm{L}} qa_1 \cdots a_{n-2}$.

**Case (vi):** $u = qa_1 \cdots a_{n-1} \circ a_n$. This case follows as in (iv) noting that $a_n \circ w <_{\mathrm{L}} a_n \circ v \leq_{\mathrm{L}} qa_1 \cdots a_{n-2}$. □

### Theorem 6.11.

*Suppose $u, v \in q/p$-DF and $u \sqsupset^{q/p} v$. Then $|uv|^{q/p}$ and $|u \circ v|^{q/p}$ exist, and $|uv|^{q/p} \sqsupset^{q/p} u$.*

**Proof.** The proof is as in Theorems 6.5 and 6.6; we provide it here for completeness. Proceed by induction on the $q/p$-DF rank of $u$ then $v$.

**Case (i):** $u <_{\mathrm{L}} q$. Then $u \sqsupset^{q/p} v$ means $u \sqsupset^p v$ with $u \circ v <_{\mathrm{L}} q$. Thus $|uv|^p$ and $|u \circ v|^p$ exist and $|uv|^p \sqsupset^p u$ by the $p$-DF version of this lemma. As $|uv|^p \circ u = u \circ v <_{\mathrm{L}} q$, $|uv|^{q/p} \sqsupset^{q/p} u$ and this case is established.

**Case (ii):** $u = q$. As $v \in q/p$-DF, $uv$ and $u \circ v \in q/p$-DF. Furthermore $uv \sqsupset^{q/p} u$ trivially (by case (iii) of the definition).

**Case (iii):** $u = qa$.

**Subcase (iii$_1$):** $v \leq_{\mathrm{L}} q$. Then $|uv|^{q/p} = qav$ and $|u \circ v|^{q/p} = \mathrm{collapse}(qa \circ v)$. By the first clause of (v) of the definition of $\sqsupset^{q/p}$, $qav \sqsupset^{q/p} qa$. Thus (iii$_1$) is established.

**Subcase (iii$_2$):** $v = qb_1 \cdots b_{m-1} * b_m$ **with** $a \sqsupset^{q/p} b_1$. Then

$$uv = qa(qb_1 \cdots b_{m-1} * b_m) = q(a \circ b_1)b_2(qb_3) \cdots (qb_{m-1}) * (qb_m) = q|a \circ b_1|^{q/p}|b_2|^{q/p}|qb_3|^{q/p} \cdots |qb_{m-1}|^{q/p} * |qb_m|^{q/p}.$$

$|a \circ b_1|^{q/p}$ exists by $a \sqsupset^{q/p} b_1$ together with the induction hypothesis, and each $|qb_i|^{q/p}$ exists for $3 \leq i \leq m$ by the induction hypothesis. This term is $q/p$-normal, so $q|a \circ b_1|^{q/p}|b_2|^{q/p}|qb_3|^{q/p} \cdots |qb_{m-1}|^{q/p} * |qb_m|^{q/p}$ is in $q/p$-DF. Furthermore, as $m \geq 1$, $q(a \circ b_1)b_2(qb_3) \cdots (qb_{m-1}) * (qb_m) \sqsupset^{q/p} q$ by parts (v) and (vi) of the definition of $\sqsupset^{q/p}$.

To complete the proof of case (iii) we establish the existence of $|u \circ v|^{q/p}$ by two further subcases.

**Subcase (iii$_{2,1}$):** $*_v = \cdot$. Then $u \circ v = q(a \circ b_1)b_2(qb_3) \cdots (qb_m) \circ q$ is clearly normal if $m \geq 2$ so $|u \circ v|^{q/p}$ exists by the induction hypothesis applied to the subterms. If $m = 1$,

$$u \circ v = qa(qb_1) \circ qa = q(ab_1) \circ qa = q(ab_1 \circ a) = q|a \circ b_1|^{q/p}$$

where $|a \circ b_1|^{q/p}$ exists by $a \sqsupset^{q/p} b_1$.

**Subcase (iii$_{2,2}$):** $*_v = \circ$. If $m \geq 3$ then

$$\begin{aligned} u \circ v &= qa \circ (qb_1 \cdots b_{m-1} \circ b_m) = (qa \circ qb_1 \cdots b_{m-1}) \circ b_m = qa(qb_1 \cdots b_{m-1}) \circ qa \circ b_m \\ &= q(a \circ b_1)b_2(qb_3) \cdots (qb_{m_1}) \circ (qa \circ b_m) = q|a \circ b_1|^{q/p} b_2 |qb_3|^{q/p} \cdots |qb_{m_1}|^{q/p} \circ |qa \circ b_m|^{q/p}. \end{aligned}$$

This term is $q/p$-normal, and the $q/p$-DF of each subterm exists by $a \sqsupset^{q/p} b_1$ and the induction hypothesis. If $m = 2$, $u \circ v = qa \circ qb_1 \circ b_2 = q(a \circ b_1) \circ b_2 = q|a \circ b_1|^{q/p} \circ b_2$. If $m = 1$, $u \circ v = q(a \circ b_1) = q|a \circ b_1|^{q/p}$. In both cases, $m = 1, 2$, each subterm has a $q/p$-DF by the induction hypothesis and the final representation is $q/p$-normal.

**Case (iv):** $u = q \circ a$. So $u \sqsupset^{q/p} v$ if and only if $a \sqsupset^{q/p} v$. Thus $uv = (q \circ a)v = q(av) = q|av|^{q/p}$, which exists by $a \sqsupset^{q/p} v$ and the induction hypothesis. $u \circ v = q \circ a \circ v = q \circ (a \circ v) = q \circ |a \circ v|^{q/p}$, which exists by $a \sqsupset^{q/p} v$ together with the induction hypothesis. For $|uv|^{q/p} \sqsupset^{q/p} u$: $q|av|^{q/p} \sqsupset^{q/p} q \circ a$ follows from part two of (iii) in the definition of $\sqsupset^{q/p}$ as $av \sqsupset^{q/p} a$ by the induction hypothesis.

**Case (v):** $u = qa_1 \cdots a_n$. In the case $v \leq_{\mathrm{L}} qa_1 \cdots a_{n-1}$, $uv = |uv|^{q/p}$, and $\mathrm{collapse}(u \circ v) = |u \circ v|^{q/p}$. Furthermore, $uv \sqsupset^{q/p} u$ by the first part of (v) in the definition of $\sqsupset^{q/p}$. Therefore suppose $v = qa_1 \cdots a_{n-1}b_n \cdots b_{m-1} * b_m$ is such that $a_n \sqsupset^{q/p} b_n$ and $a_n \circ b_n \leq_{\mathrm{L}} qa_1 \cdots a_{n_2}$ with $m \geq n+1$. (We will consider the case $m = n$ separately below.) Then

$$\begin{aligned} uv = qa_1 \cdots a_n(qa_1 \cdots a_{n-1}b_nb_{n+1} \cdots b_{m-1} * b_m) &= qa_1 \cdots a_{n-1}(a_n \circ b_n)b_{n+1}(ub_{n+2}) \cdots (ub_{m-1}) * (ub_m) \\ &= qa_1 \cdots a_{n-1}|a_n \circ b_n|^{q/p} b_{n+1}|ub_{n+2}|^{q/p} \cdots |ub_{m-1}|^{q/p} * |ub_m|^{q/p}. \end{aligned}$$

The $q/p$-DF of each subterm exists by $a_n \sqsupset^{q/p} b_n$ and the induction hypothesis, and the $q/p$-normality of the term follows from that of $u$ and $v$ and from $a_n \circ b_n \leq_{\mathrm{L}} qa_1 \cdots a_{n-2}$. Still in the case $m \geq n+1$ we have

$$u \circ v = uv \circ u = qa_1 \cdots a_{n-1}|a_n \circ b_n|^{q/p} b_{n+1}|ub_{n+2}|^{q/p} \cdots |ub_{m-1}|^{q/p} * |ub_m|^{q/p} \circ u.$$

The existence of $|u \circ v|^{q/p}$ follows from the existence of $|uv|^{q/p}$ and normality, taking the collapse if necessary. If $m = n$, then $u \circ v = qa_1 \cdots a_{n-1} * (a_n \circ b_n)$. If $m = n$ and $* = \cdot$, then by Rule 2 $|uv|^{q/p} = qa_1 \cdots a_{n-1}|a_nb_n|^{q/p}$ and $|u \circ v|^{q/p} = qa_1 \cdots a_{n-1}|a_n \circ b_n|^{q/p}$. When $m = n$ and $* = \circ$, by applying Rule 3 we obtain $|uv|^{q/p} = qa_1 \cdots a_{n-1}|a_n \circ a_n|^{q/p} \circ qa_1 \cdots a_{n-1}|a_nb_n|^{q/p}$ and $|u \circ v|^{q/p} = \mathrm{collapse}\big(qa_1 \cdots a_{n-1} \circ (a_n \circ b_n)\big)$, where each component can be written in $q/p$-DF by in the induction hypothesis.

**Case (vi):** $u = qa_1 \cdots a_{n-1} \circ a_n$ with $a_n \sqsupset^{q/p} v$ and $a_n \circ v \leq_{\mathrm{L}} qa_1 \cdots a_{n-2}$. Then $|uv|^{q/p} = qa_1 \cdots a_{n-1}|a_nv|^{q/p}$ and $uv \sqsupset^{q/p} u$ by $a_nv \sqsupset^{q/p} a_n$ (which follows from the induction hypothesis), and $|u \circ v|^{q/p} = qa_1 \cdots a_{n-1} \circ |a_n \circ v|^{q/p}$. □

## 7. Translating between division form bases

In this section we demonstrate lemmas giving circumstances under which for all $w$, whenever $|w|^p$ exists, $|w|^q$ exists.

### *Definition 7.1.*

For $p, q \in \mathcal{P}$, let $p \twoheadrightarrow q$ denote that $|p|$ and $|q|$ exist and for every $r \in \mathcal{P}$, if $|r|^p$ exists then $|r|^q$ exists. The definitions of $q \twoheadrightarrow (p, r)$ and $(p, r) \twoheadrightarrow q$ are analogous. Write $p \twoheadleftarrow\!\!\twoheadrightarrow q$ if and only if $p \twoheadrightarrow q$ and $q \twoheadrightarrow p$.

We will prove the main theorem, i.e., $|p|^q$ exists for all $p, q \in \mathcal{P}$, after establishing: *If $x \twoheadleftarrow\!\!\twoheadrightarrow p \twoheadleftarrow\!\!\twoheadrightarrow r$ then $x \twoheadleftarrow\!\!\twoheadrightarrow (p, r)$ and $x \twoheadleftarrow\!\!\twoheadrightarrow pr \twoheadleftarrow\!\!\twoheadrightarrow (p, r)$ and $x \twoheadrightarrow p \circ r$.*

### *Lemma 7.2.*

*If $x \twoheadleftarrow\!\!\twoheadrightarrow p \twoheadleftarrow\!\!\twoheadrightarrow r$, then $x \twoheadleftarrow\!\!\twoheadrightarrow (p, r)$.*

**Proof.** For $x \twoheadrightarrow (p,r)$ it suffices to show that $p \mathrel{\twoheadleftarrow\!\!\twoheadrightarrow} (p,r)$ as $x \mathrel{\twoheadleftarrow\!\!\twoheadrightarrow} p$. Proceeding by induction on the rank of $w \in p$-DF, we show that if $|w|^p$ exists, then $|w|^{p,r}$ exists.

For $p \twoheadrightarrow (p,r)$: If $w <_{\mathrm{L}} p$, $|w| = |w|^p = |w|^{p,r}$, so consider $|w|^p = pa_1a_2 \cdots a_{n-1} * a_n$. Then, as $p \mathrel{\twoheadleftarrow\!\!\twoheadrightarrow} r$, $|a_1|^r$ exists. Furthermore, the induction hypothesis yields $|a_i|^{p,r}$ for $i \geq 2$. Thus we have $|w|^{p,r} = p|a_1|^r|a_2|^{p,r} \cdots |a_{n-1}|^{p,r} * |a_n|^{p,r}$.

For $(p,r) \twoheadrightarrow p$: Given $w \in (p,r)$-DF, $|w|^{p,r} = |w|$ if $w \leq_{\mathrm{L}} p$, so suppose $|w|^{p,r} = pa_1a_2 \ldots a_{n-1} * a_n$. As $p \mathrel{\twoheadleftarrow\!\!\twoheadrightarrow} r$, each subterm $u$ of $w$ that is in $r$-DF can be replaced with $|u|^p$, which gives $|w|^p$ as normality is preserved. The special case $|w|^{p,r} = p|q|^r \circ p = p \circ |q|^r$ gives $p \circ |q|^p$ by the assumption $p \mathrel{\twoheadleftarrow\!\!\twoheadrightarrow} r$. □

Note that the assumption $x \mathrel{\twoheadleftarrow\!\!\twoheadrightarrow} p$ gives $pr \mathrel{\twoheadleftarrow\!\!\twoheadrightarrow} pr/p$, so Theorems 7.3 and 7.4 hold for $pr$-DF (and $(p,r)$-DF) as well as $pr/p$-DF (and $(p \circ r)/p$-DF).

### Theorem 7.3.

*Suppose $x \mathrel{\twoheadleftarrow\!\!\twoheadrightarrow} p \mathrel{\twoheadleftarrow\!\!\twoheadrightarrow} r$ and $w \in r$-DF. Then $|pw|^{pr/p}$ exists and $|pw|^{pr/p} \sqsupset^{pr/p} p$.*

**Proof.** We show $|pw|^{pr/p}$ exists and $|pw|^{pr/p} \sqsupset^{pr/p} p$ by induction on the $r$-DF rank of $|w|^r$ when $w >_{\mathrm{L}} r$.

First suppose $w <_{\mathrm{L}} r$. Then $|w|^r = |w|$. By the hypothesis that $x \twoheadrightarrow p$, $|w|^p$ exists so $p|w|^p$ is in $p$-DF and therefore $p|w|^p \sqsupset^p p$ by definition. As $|pw|^{pr/p} \sqsupset^{pr/p} p$ if and only if $p|w|^p \sqsupset^p p$ when $pw <_{\mathrm{L}} pr$, this case is established.

When $w = r$, $|pw|^{pr/p} \sqsupset^{pr/p} v$ for all $v \in pr/p$-DF, so $|pw|^{pr/p} \sqsupset^{pr/p} |p|$ as $|p| = |p|^{pr/p}$. If $w >_{\mathrm{L}} r$, we have

$$pw = p(rw_1 \cdots w_{m-1} * w_m) = pr(pw_1) \cdots (pw_{m-1}) * (pw_m).$$

Each $|pw_i|^{pr/p}$ exists by induction as $\operatorname{rank}|w_i|^r < \operatorname{rank}|w|^r$, and the displayed term is $pr/p$-normal by the $r$-normality of $w$. Thus $|pw|^{pr/p}$ exists. When $*_w = \cdot$ it is clear that $|pw|^{pr/p} \sqsupset^{pr/p} p$. Suppose then that $*_w = \circ$. The induction hypothesis gives that $|pw_m|^{pr/p} \sqsupset^{pr/p} p$, and $w_m <_{\mathrm{L}} rw_1 \cdots w_{m-2}$ yields $p \circ w_m <_{\mathrm{L}} pr(pw_1) \cdots (pw_{m-2})$. These are the conditions necessary to conclude $|pw|^{pr/p} \sqsupset^{pr/p} p$. □

### Theorem 7.4.

*Suppose $x \mathrel{\twoheadleftarrow\!\!\twoheadrightarrow} p \mathrel{\twoheadleftarrow\!\!\twoheadrightarrow} r$ and $w \in r$-DF. Then $|pw|^{(p \circ r)/p}$ exists, and $|pw|^{(p \circ r)/p} \sqsupset^{(p \circ r)/p} b$ for every $b \leq_{\mathrm{L}} p$.*

**Proof.** By induction on the $r$-DF rank of $|w|^r$.

**Case (i):** $w \leq_{\mathrm{L}} r$. Then $pw \leq_{\mathrm{L}} pr <_{\mathrm{L}} p \circ r$, so $|pw|^{(p \circ r)/p} = |pw|^p$. As $p \sqsupset^{(p \circ r)/p} w$ whenever $p \sqsupset^p w$ by definition of $\sqsupset^{(p \circ r)/p}$, $|pw|^p = |pw|^{(p \circ r)/p}$ exists by $p \mathrel{\twoheadleftarrow\!\!\twoheadrightarrow} x$. Furthermore, $|pw|^p \sqsupset^p b$ for all $b \leq_{\mathrm{L}} p$ by definition of $\sqsupset^p$, so $|pw|^{(p \circ r)/p} \sqsupset^{(p \circ r)/p} b$ for all $b \leq_{\mathrm{L}} p$ because $pw \circ b \leq_{\mathrm{L}} pw \circ p = p \circ w \leq_{\mathrm{L}} p \circ r$.

**Case (ii):** $w >_{\mathrm{L}} r$. Then $w = ra_1 \cdots a_{n-1} * a_n$.

**Subcase (ii$_1$):** $n > 1$ or $*_w = \cdot$. Then

$$pw = p(ra_1 \cdots a_{n-1} * a_n) = p(ra_1) \cdots (pa_{n-1}) * (pa_n) = (p \circ r)a_1(pa_2) \cdots (pa_{n-1}) * (pa_n).$$

Each of $|a_1|^{(p \circ r)/p}$, $|pa_i|^{(p \circ r)/p}$, $i > 1$, exist by the induction hypothesis. The final representation of the displayed term is normal (by the normality of $|w|^r$), so it is in $(p \circ r)/p$-DF once we put each of the components into $(p \circ r)/p$-DF. Thus $|pw|^{(p \circ r)/p}$ exists. If $*_w = \cdot$, then $|pw|^{(p \circ r)/p} \sqsupset^{(p \circ r)/p} b$ trivially. If $*_w = \circ$ and $n > 1$, we need only that $|pa_n|^{(p \circ r)/p} \sqsupset^{(p \circ r)/p} b$, which holds by the induction hypothesis. By Lemma 5.3 and the $r$-normality of $|w|^r$ we have that $pa_n \circ b \leq_{\mathrm{L}} pa_n \circ p = p \circ a_n <_{\mathrm{L}} (p \circ r)a_1(pa_2) \cdots (pa_{n-2})$.

**Case (ii$_2$):** $n = 1$ and $*_w = \circ$. Then $w = r \circ a_1$. Thus

$$pw = p(r \circ a_1) = p(ra_1 \circ r) = p(ra_1) \circ pr = (p \circ r)a_1 \circ pr$$

which is normal with respect to the $(p \circ r)/p$ parsing, so $|pw|^{(p \circ r)/p}$ exists and $|pw|^{(p \circ r)/p} = (p \circ r)|a_1|^{(p \circ r)/p} \circ pr \sqsupset^{(p \circ r)/p} p$. Furthermore, $b \leq_{\mathrm{L}} p$ implies that $pr \circ b \leq_{\mathrm{L}} p \circ r$, so $|pw|^{(p \circ r)/p} \sqsupset^{(p \circ r)/p} b$. □

### Lemma 7.5.

*If $x \twoheadleftarrow\!\!\twoheadrightarrow p \twoheadleftarrow\!\!\twoheadrightarrow r$ then $|pr| \twoheadrightarrow (p,r)$.*

Note that $x \twoheadrightarrow (p,r)$ and the existence of $|p|, |r|$ and $|pr|$ follows from $x \twoheadleftarrow\!\!\twoheadrightarrow p \twoheadleftarrow\!\!\twoheadrightarrow r$. In particular, if $x \twoheadleftarrow\!\!\twoheadrightarrow p \twoheadleftarrow\!\!\twoheadrightarrow r$, $r \in \mathrm{DF}$ by hypothesis, so $|r|^p$ exists, and thus $p|r|^p \in p$-DF. As $p \twoheadrightarrow x$, $|pr|$ exists. The fact that $x \twoheadrightarrow (p,r)$ is Lemma 7.2.

***Proof.*** By induction on the rank of $w \in pr/p$-DF we show that if $|w|^{pr/p}$ exists, then $|w|^{p,r}$ exists. If $w \leq_{\mathrm{L}} pr$, then $|w|^{pr/p} = |w|^p$, so by the hypotheses $p \twoheadrightarrow x$ and $x \twoheadrightarrow (p,r)$, $|w|^{p,r}$ exists. So suppose $w >_{\mathrm{L}} pr$ and

$$|w|^{pr/p} = pra_1a_2 \cdots a_{n-1} * a_n.$$

By the induction hypothesis, $a_1$ has a $(p,r)$-DF. If $a_1 \leq_{\mathrm{L}} p$, then $pr \sqsupset^{p,r} |a_1|^{p,r}$ and $pr|a_1|^{p,r} = pr|a_1|$ is in $(p,r)$-DF. Therefore suppose $|a_1|^{p,r} = pr'c_1 \cdots c_{m-1} *_{a_1} c_m$. As $r \sqsupset^r r'$ for all $r' \in r$-DF, $pr \sqsupset^{p,r} |a_1|^{p,r}$ so $|pra_1|^{p,r}$ exists and $|pra_1|^{p,r} \sqsupset^{p,r} b$ for all $b \leq_{\mathrm{L}} pr$, including $a_2$ (by the normality of $|w|^{pr/p}$), so $|pra_1|^{p,r} \sqsupset^{p,r} |a_2|^{p,r}$. We thus obtain $w \in (p,r)$-DF by repeated applications of Theorem 6.6 (i.e., the $(p,r)$-DF version of Corollary 6.7). □

### Lemma 7.6.

*If $x \twoheadleftarrow\!\!\twoheadrightarrow p \twoheadleftarrow\!\!\twoheadrightarrow r$ then $x \twoheadleftarrow\!\!\twoheadrightarrow |pr|$.*

***Proof.*** As $x \twoheadleftarrow\!\!\twoheadrightarrow p$, it suffices to show $p \twoheadrightarrow |pr|$ and $|pr| \twoheadrightarrow x$. In fact, as $pr/p \twoheadleftarrow\!\!\twoheadrightarrow pr$, we show $p \twoheadrightarrow pr/p$ by induction on the rank of $w$ in $p$-DF. If $w \in p$-DF with $w \leq_{\mathrm{L}} pr$, $|w|^p = |w|^{pr/p}$ by definition of $pr/p$-DF.

For $w >_{\mathrm{L}} pr$, $w = pa_1 \cdots a_{n-1} * a_n \in (p,r)$-DF exists by Lemma 7.2, where $|a_1|^r = rb_1b_2 \cdots b_{m-1} * b_m$. Thus

$$w = pa_1a_2 \cdots a_{n-1} * a_n = p(rb_1b_2 \cdots b_{m-1} *_{a_1} b_m)a_2 \cdots a_{n-1} * a_n = \big[(pr)(pb_1)(pb_2) \cdots (pb_{m-1}) *_{a_1} (pb_m)\big]a_2 \cdots a_{n-1} * a_n.$$

If $*_{a_1} = \cdot$, the induction hypothesis gives that the $pr/p$-DF of each subterm exists, so the $r$-normality of $a_1$ gives that

$$pr|pb_1|^{pr/p}|pb_2|^{pr/p} \cdots |pb_{m-1}|^{pr/p}|pb_m|^{pr/p}$$

is $pr/p$-normal. The $pr/p$-normality of the rest of $|w|^{pr/p}$ follows from the $(p,r)$-normality of $w$, and

$$|w|^{pr/p} = pr|pb_1|^{pr/p}|pb_2|^{pr/p} \cdots |pb_{m-1}|^{pr/p}|pb_m|^{pr/p}|a_2|^{pr/p} \cdots |a_{n-1}|^{pr/p} * |a_n|^{pr/p}.$$

In the case $*_{a_1} = \circ$,

$$w = (pr)(pb_1)(pb_2) \cdots (pb_{m-1})(pb_ma_2)a_3 \cdots a_{n-1} * a_n.$$

Each $|pb_i|^{pr/p}$ exists by the induction hypothesis, and Theorem 6.11 gives that $pb_m \sqsupset^{pr/p} a_2$ as $a_2 \leq_{\mathrm{L}} p$. This gives the existence of

$$|pr||pb_1|^{pr/p}|pb_2|^{pr/p} \cdots |pb_{m-1}|^{pr/p}|pb_ma_2|^{pr/p},$$

which we will denote by $u$ for brevity. We have that $u \sqsupset^{pr/p} a_3$ as $a_3 \leq_{\mathrm{L}} pa_1$, so $a_3$ satisfies condition (iii) of the definition of $\sqsupset^{pr/p}$. By Corollary 6.7, we get that $|w|^{pr/p}$ exists. For the $|pr| \twoheadrightarrow x$ direction, note that $|pr| \twoheadrightarrow (p,r)$ by Lemma 7.5 and $(p,r) \twoheadrightarrow x$ by Lemma 7.2. □

As $x \twoheadleftarrow\!\!\twoheadrightarrow p \twoheadleftarrow\!\!\twoheadrightarrow r$ implies $x \twoheadleftarrow\!\!\twoheadrightarrow (p,r) \twoheadleftarrow\!\!\twoheadrightarrow pr$ is now established, it remains only to show that $x \twoheadrightarrow p \circ r$.

### Lemma 7.7.

*Suppose $x \twoheadleftarrow\!\!\twoheadrightarrow p \twoheadleftarrow\!\!\twoheadrightarrow r$. Then $x \twoheadrightarrow p \circ r$.*

***Proof.*** It is enough to show that $x \twoheadrightarrow (p \circ r)/p$. Proceed by induction on the rank of $w \in \mathrm{DF}$. Given $w \in \mathrm{DF}$, Lemma 7.2 proves that $|w|^{p,r}$ exists. If $w \leq_{\mathrm{L}} pr$, $|w|^{(p \circ r)/p} = |w|^p$, and this case is established as $x \twoheadrightarrow p$. Thus suppose $w >_{\mathrm{L}} pr$ and $|w|^{p,r} = pa_1a_2 \cdots a_{n-1} * a_n = p(rb_1 \cdots b_{m-1} *_{a_1} b_m)a_2 \cdots a_{n-1} * a_n$, with $n \geq 1$ (it is possible that $a_1 = r$ unless $n = 1$).

In the case $*_{a_1} = \cdot$,

$$|w|^{(p \circ r)/p} = (p \circ r)|b_1|^{(p \circ r)/p}|pb_2|^{(p \circ r)/p} \cdots |pb_m|^{(p \circ r)/p}|a_2|^{(p \circ r)/p} \cdots |a_{n-1}|^{(p \circ r)/p} * |a_n|^{(p \circ r)/p},$$

where $|b_1|^{(p \circ r)/p}, |pb_2|^{(p \circ r)/p}, \ldots, |pb_m|^{(p \circ r)/p}, |a_2|^{(p \circ r)/p}, \ldots, |a_n|^{(p \circ r)/p}$ exist by the induction hypothesis and the $(p \circ r)/p$-normality of this representation of $w$ follows from the $(p, r)$-normality of $|w|^{p,r}$.

When $*_{a_1} = \circ$, $w = (p \circ r)b_1(pb_2) \cdots (pb_m a_2)a_3 \cdots a_{n-1} * a_n$. The existence of $|(p \circ r)b_1(pb_2) \cdots (pb_{m-1}) \circ (pb_m)|^{(p \circ r)/p}$ follows from the remarks made in the case where $*_{a_1} = \cdot$, and $|pb_m a_1|^{(p \circ r)/p}$ exists by Theorem 7.4, so to establish the existence of $|w|^{(p \circ r)/p}$ we need only show that there is a $(p \circ r)/p$-division form representation of

$$(p \circ r)|b_1|^{(p \circ r)/p}|pb_2|^{(p \circ r)/p} \cdots |pb_m a_2|^{(p \circ r)/p}|a_3|^{(p \circ r)/p} \ldots |a_{n-1}|^{(p \circ r)/p} * |a_n|^{(p \circ r)/p}.$$

The $(p \circ r)/p$-normality of this term up to and including $|pb_m a_2|^{(p \circ r)/p}$ follows directly from the facts that $pb_m <_{\mathrm{L}} (p \circ q)b_1(pb_2) \ldots (pb_{m-2})$ and that $pb_m a_1$ is $(p, r)$-normal, which yields $(p \circ r)/p$-normality.

To obtain $|p(rb_1 \cdots b_{m-1} \circ b_m)|^{(p \circ r)/p} \sqsupset^{(p \circ r)/p} |a_2|^{(p \circ r)/p}$, apply Theorems 6.5 and 7.4 and Lemma 6.3. Therefore $|p(rb_1 \cdots b_{m-1} \circ b_m)a_2|^{(p \circ r)/p}$ exists and $|p(rb_1 \cdots b_{m-1} \circ b_m)a_2|^{(p \circ r)/p} \sqsupset^{(p \circ r)/p} a_3$ as $a_3 \leq_{\mathrm{L}} pa_1 = p(rb_1 \cdots b_{m-1} \circ b_m)$. Corollary 6.7 gives that $|w|^{(p \circ r)/p}$ exists. □

## 8. The main result

***Theorem 8.1.***
*For every $p \in \mathcal{P}$, $|p|$ exists. More generally, for every $q \in \mathcal{P}$, $|p|^{|q|}$ exists.*

***Proof.*** To prove the theorem we first prove the following:

(I) for every $b \in A$, $|b|$ exists and $|b| \twoheadleftarrow\!\!\!\!\twoheadrightarrow x$;

(II) for every $b \in A$ and $p \in P$, $|p|^{|b|}$ exists.

(I) The set $\mathcal{M} = \{a \in \mathcal{A} : |a| \text{ exists and } |a| \twoheadleftarrow\!\!\!\!\twoheadrightarrow x\}$ contains $x$ and by Lemma 7.6 is closed under products. Thus $\mathcal{M} = \mathcal{A}$.

(II) By (I) it suffices to show that the claim holds for $b = x$. Given $p \in \mathcal{P}$, write $p = a_0 \circ a_1 \circ \cdots \circ a_n$ with each $a_i \in \mathcal{A}$. Then each $|a_i|$ exists by (I). Suppose inductively that $|a_{i+1} \circ a_{i+2} \circ \cdots \circ a_n|$ exists. Then by (I), $|a_{i+1} \circ a_{i+2} \circ \cdots \circ a_n|^{|a_i|}$ exists. Thus $|a_i \circ a_{i+1} \circ \cdots \circ a_n|^{|a_i|} = |a_i| \circ |a_{i+1} \circ a_{i+2} \circ \cdots \circ a_n|^{|a_i|}$ exists. So $|a_i \circ a_{i+1} \circ \cdots \circ a_n|$ exists by (I), giving, by repeated application, that $|p|$ exists.

To prove the theorem, given $p, q \in \mathcal{P}$ with $q = a_0 \circ \cdots \circ a_n$, then $|q|$ exists by (II). To show for all $p$, $|p|^{|q|}$ exists it suffices to show $q \twoheadleftarrow\!\!\!\!\twoheadrightarrow x$. We have each $a_i \twoheadleftarrow\!\!\!\!\twoheadrightarrow x$ by (I). Suppose inductively that $(a_{i+1} \circ a_{i+2} \circ \cdots \circ a_n) \twoheadleftarrow\!\!\!\!\twoheadrightarrow x$. Thus $a_i \twoheadleftarrow\!\!\!\!\twoheadrightarrow (a_{i+1} \circ a_{i+2} \circ \cdots \circ a_n) \twoheadleftarrow\!\!\!\!\twoheadrightarrow x$, so by Lemma 7.7, $x \twoheadleftarrow\!\!\!\!\twoheadrightarrow (a_i \circ a_{i+1} \circ \cdots \circ a_n)$. But since (by (II)) for every $p \in \mathcal{P}$, $|p|^x$ exists, we have for every $p \in \mathcal{P}$ that $|p|^{a_i \circ \cdots \circ a_n}$ exists. This gives $(a_0 \circ \cdots \circ a_n) \twoheadleftarrow\!\!\!\!\twoheadrightarrow x$. Therefore, for every $p, q \in \mathcal{P}$, $|p|^{|q|}$ exists. □

## References

[1] Artin E., Theorie der Zöpfe, Abh. Math. Sem. Univ. Hamburg, 1925, 4(1), 47–72

[2] Birman J.S., Braids, Links, and Mapping Class Groups, Ann. of Math. Stud., 82, Princeton University Press, Princeton, 1974

[3] Brieskorn E., Automorphic sets and braids and singularities, In: Braids, Santa Cruz, July 13–26, 1986, Contemp. Math., 78, American Mathematical Society, Providence, 1988, 45–115

[4] Burckel S., The wellordering on positive braids, J. Pure Appl. Algebra, 1997, 120(1), 1–17

[5] Dehornoy P., Braid groups and left distributive operations, Trans. Amer. Math. Soc., 1994, 345(1), 115–150

[6] Dehornoy P., Braids and Self-Distributivity, Progr. Math., 192, Birkhäuser, Basel, 2000

[7] Dehornoy P., Dynnikov I., Rolfsen D., Wiest B., Why are Braids Orderable?, Panor. Syntheses, 14, Société Mathématique de France, Paris, 2002

[8] Fenn R., Rourke C., Racks and links in codimension two, J. Knot Theory Ramifications, 1992, 1(4), 343–406

[9] Hurwitz A., Ueber Riemann'sche Flächen wit gegebenen Verzweigungspunkten, Math. Ann., 1891, 39(1), 1–60

[10] Joyce D., A classifying invariant of knots, the knot quandle, J. Pure Appl. Algebra, 1982, 23(1), 37–65

[11] Kunen K., Elementary embeddings and infinitary combinatorics, J. Symbolic Logic, 1971, 36(3), 407–413

[12] Larue D.M., Braid words and irreflexivity, Algebra Universalis, 1994, 31(1), 104–112

[13] Laver R., A division algorithm for the free left distributive algebra, In: Logic Colloquium '90, Helsinki, July 15–22, 1990, Lecture Notes Logic, 2, Springer, Berlin, 1993, 155-162

[14] Laver R., The left distributive law and the freeness of an algebra of elementary embeddings, Adv. Math., 1992, 91(2), 209–231

[15] Laver R., On the algebra of elementary embeddings of a rank into itself, Adv. Math., 1995, 110(2), 334–346

[16] Laver R., Braid group actions on left distributive structures, and well orderings in the braid groups, J. Pure Appl. Algebra, 1996, 108(1), 81–98

[17] Laver R., Miller S.K., Left division in the free left distributive algebra on one generator, J. Pure Appl. Algebra, 2010, 215(3), 276–282

[18] Laver R., Moody J.A., Well-foundedness conditions connected with left-distributivity, Algebra Univsersalis, 2002, 47(1), 65–68

[19] Miller S.K., Free Left Distributive Algebras, PhD thesis, University of Colorado, Boulder, 2007

[20] Miller S.K., Free left distributive algebras on $\kappa$ generators (in preparation)