

ON THE ENUMERATION AND ASYMPTOTIC GROWTH OF FREE QUASIGROUP WORDS

JONATHAN D.H. SMITH¹ AND STEFANIE G. WANG²

ABSTRACT. The paper counts the number of reduced quasigroup words of a particular length in a certain number of generators. Taking account of the relationship with the Catalan numbers, counting words in a free magma, we introduce the term peri-Catalan number for the free quasigroup word counts. The main result of the paper is an exact recursive formula for the peri-Catalan numbers, structured by the Euclidean Algorithm.

The Euclidean Algorithm structure does not readily lend itself to standard techniques of asymptotic analysis. However, conjectures for the asymptotic behavior of the peri-Catalan numbers, substantiated by numerical data, are presented. A remarkable aspect of the observed asymptotic behavior is the so-called asymptotic irrelevance of quasigroup identities, whereby cancelation resulting from quasigroup identities has a negligible effect on the asymptotic behavior of the peri-Catalan numbers for long words in a large number of generators.

1. INTRODUCTION

In the free magma on one generator x , the number of words of length n (i.e., with n occurrences of the letter x) is given by the Catalan number C_n (in the natural leaf-counting indexation of [8, Rem. 5.5]). More generally, the number of words of length n in the free magma on s generators is given as $s^n C_n$. Now consider the free quasigroup on s generators. The word problem was solved by Evans in a notable paper which gave one of the earliest applications of confluence (the “Diamond Lemma”) in rewriting systems [1] [2, §I.3]. We now quantify Evans’ normal form theorem by asking for the number of reduced words of length n , which we define as the *peri-Catalan number* P_n^s (§4.1). In the language of geometric group theory, we are interested in the growth rate of a free quasigroup (compare [4]).

In our terminology, the Greek qualifier $\pi\varepsilon\rho\iota$ has the sense of “around” (as in “perimeter”), in reference to the fact that the free quasigroup sits around the free magma. The close relationship between the Catalan numbers and the peri-Catalan numbers is reinforced by our conjectured formula

$$(1.1) \quad \lim_{n \rightarrow \infty} \frac{\log P_n^s}{\log C_n + n \log 3s - \log 3} \simeq 1 - \frac{0.019\dots}{s - \log((1 + \sqrt{5})/2)}$$

for the asymptotic behavior of the peri-Catalan numbers. We do not have an interpretation for the numerical constant $0.019\dots$ appearing in the numerator of the fraction on the right hand side of (1.1).

The main result of the paper is the recursive formula (4.15) for the peri-Catalan numbers P_n^s . The formula is structured around the Euclidean Algorithm. More

2010 *Mathematics Subject Classification.* 20N05, 08B20, 05A99.

Key words and phrases. quasigroup, triality, Catalan number, reduced word, golden section.

specifically, for each partition $n = (n - k) + k$ of length two, with $n - k \geq k$, and for each call to the Division Algorithm within the Euclidean Algorithm for computing $\gcd(n, k)$, there is a corresponding summand in the recursion formula. The appearance of the golden ratio in the conjectured formula (1.1) may be related to the fact that the Fibonacci numbers provide the worst cases for runs of the Euclidean Algorithm [7, §4.5.3, Th. F].

Section 2 gives the background definitions for equationally-defined quasigroups, as algebras $(Q, \cdot, /, \backslash)$ with basic operations of multiplication \cdot , right division $/$, and left division \backslash that satisfy identities such as $x \cdot (x \backslash y) = y$. This identity corresponds to one of the Latin square properties for the body of the multiplication table of a finite quasigroup, specifically that each element y appears in the row labeled by x , namely in the column labeled by $x \backslash y$ [14, Defn. 11.6(a)]. Thus $x \cdot (x \backslash y)$ is not reduced as a quasigroup word of length 3, and will not contribute to the peri-Catalan number P_3^2 .

Section 3 recalls the triality symmetry (S_3 -action) of the language of quasigroups. The recursion formula for P_n^s is based on an inductive process creating a quasigroup word of length n by applying one of the three basic operations to a pair of reduced words whose lengths sum to n (§4.2). The triality notation of Section 3 is used in §4.3 to give an efficient and non-repetitive treatment of the cancelations that take place in the inductive process. This application of triality is comparable to its use in the streamlining of the proof of Evans' normal form theorem [12]. The Euclidean Algorithm notation (generally following [14, §1.6]) is introduced in §4.4. The double inductive proof of the main Theorem 4.9 is then given in §4.5.

The recursive formula that is provided by Theorem 4.9 does not appear to be readily amenable to the standard methods of asymptotic analysis. Nevertheless, the concluding Section 5 provides some conjectured results based on an alternative approach. In particular, Conjecture 5.2 suggests the asymptotic behavior (1.1) for all generator counts s . The so-called *asymptotic irrelevance of quasigroup identities* (Conjecture 5.1) asserts that cancelation resulting from quasigroup identities has a negligible effect on the asymptotic behavior of the peri-Catalan numbers for long words in a large number of generators.

We use algebraic notation (with functions following, or as superfixes of, their arguments) as the default. Note that S_r denotes the symmetric group on r letters.

2. BACKGROUND

2.1. Magmas, rooted binary trees, and Catalan numbers. A *magma* (M, \cdot) is a set equipped with a single binary operation $M \times M \rightarrow M; (m_1, m_2) \mapsto m_1 \cdot m_2$. Elements of free magmas are described as (*magma*) *words*. Words in the free magma on a singleton alphabet $\{a\}$ are repeated concatenations of the generator a under the magma operation. Each word w is determined by a *rooted binary tree* B_w , defined recursively as follows:

- (a) The tree B_a consists of a single node (which is both root and leaf);
- (b) For words u, v , the tree $B_{u \cdot v}$ has a root with B_u as a left child, and B_v as a right child.

If B_w has n leaves, then the magma word w has *length* n . Although varying conventions are encountered, we define the n -th *Catalan number* C_n as the number of magma words of length n in the single generator a (compare [8, §5]). Then the free magma in s generators has $s^n C_n$ words of length n .

A magma (Q, \cdot) is a *combinatorial quasigroup* if knowing any two arguments in the equation $x \cdot y = z$ uniquely specifies the third argument for all $x, y, z \in Q$.

2.2. Quasigroups. An *equational quasigroup* $(Q, \cdot, /, \backslash)$ is a set equipped with three *basic* binary operations, multiplication \cdot , right division $/$, and left division \backslash such that for all $x, y \in Q$, the following identities are satisfied:

$$(2.1) \quad \begin{array}{ll} \text{(SL)} & x \cdot (x \backslash y) = y; \\ \text{(SR)} & y = (y/x) \cdot x; \\ \text{(IL)} & x \backslash (x \cdot y) = y; \\ \text{(IR)} & y = (y \cdot x)/x. \end{array}$$

Note that (IL), (IR) give the respective injectivity of the left multiplication

$$L(x): Q \rightarrow Q; y \mapsto xy$$

and right multiplication

$$R(x): Q \rightarrow Q; y \mapsto yx$$

for $x \in Q$, while (SL), (SR) give their surjectivity. Thus an equational quasigroup $(Q, \cdot, /, \backslash)$ yields a combinatorial quasigroup (Q, \cdot) . Conversely, a combinatorial quasigroup (Q, \cdot) yields an equational quasigroup $(Q, \cdot, /, \backslash)$ with $x/y = xR(y)^{-1}$ and $x \backslash y = yL(x)^{-1}$.

In an equational quasigroup $(Q, \cdot, /, \backslash)$, the three equations

$$(2.2) \quad x_1 \cdot x_2 = x_3, \quad x_3/x_2 = x_1, \quad x_1 \backslash x_3 = x_2$$

involving the basic operations are equivalent. Introducing the *opposite* operations

$$x \circ y = y \cdot x, \quad x // y = y/x, \quad x \backslash \backslash y = y \backslash x$$

on Q , the equations (2.2) are further equivalent to the equations

$$x_2 \circ x_1 = x_3, \quad x_2 // x_3 = x_1, \quad x_3 \backslash \backslash x_1 = x_2.$$

Thus each of the basic and opposite operations

$$(2.3) \quad (Q, \cdot), \quad (Q, /), \quad (Q, \backslash), \quad (Q, \circ), \quad (Q, //), \quad (Q, \backslash \backslash)$$

forms a (combinatorial) quasigroup. In particular, note that the identities (IR) in (Q, \backslash) and (IL) in $(Q, /)$ yield the respective identities

$$\begin{array}{ll} \text{(DL)} & x/(y \backslash x) = y, \\ \text{(DR)} & y = (x/y) \backslash x \end{array}$$

in the basic quasigroup divisions. The six quasigroups (2.3) are known as the *conjugates*, “parastrophes” [9] or “derived quasigroups” [6] of (Q, \cdot) .

2.3. Basic words and parsing trees. In the free quasigroup on an alphabet

$$\{a_1, a_2, \dots, a_s\}$$

of s letters, (*basic*) *quasigroup words* are repeated concatenations of the generators under the three basic quasigroup operations $\cdot, /, \backslash$. Each basic quasigroup word w is in one-one correspondence with a *basic parsing tree* T_w , defined recursively as follows:

- (a) For $1 \leq i \leq s$, the tree T_{a_i} is a single vertex annotated by a_i ;
- (b) For basic words u, v , the tree $T_{u \cdot v}$ has:
 - (i) a root annotated by the multiplication,
 - (ii) T_u as a left child, and T_v as a right child;
- (c) For basic words u, v , the tree $T_{u/v}$ has:
 - (i) a root annotated by the right division,

- (ii) T_u as a left child, and T_v as a right child;
- (d) For basic words u, v , the tree $T_{u \setminus v}$ has:
 - (i) a root annotated by the left division,
 - (ii) T_u as a left child, and T_v as a right child.

Remark 2.1. (a) A basic parsing tree stripped of its annotation yields a binary rooted tree.

(b) If a basic parsing tree T_u has n leaves, then the basic quasigroup word u has length n , involving $n - 1$ basic operations determined by the annotations of the nodes of T_u .

3. TRIALITY

3.1. Triality action of the symmetric group S_3 . The symmetric group S_3 on the 3-element set $\{1, 2, 3\}$ is presented as

$$\langle \sigma, \tau \mid \sigma^2 = \tau^2 = (\sigma\tau)^3 = 1 \rangle$$

writing σ and τ for the respective transpositions (12) and (23). The Cayley diagram of the presentation is

$$(3.1) \quad \begin{array}{ccccc} 1 & \iff & \tau & \longleftrightarrow & \tau\sigma \\ \updownarrow & & & & \updownarrow \\ \sigma & \iff & \sigma\tau & \longleftrightarrow & \sigma\tau\sigma \end{array}$$

with \leftrightarrow for right multiplication by σ and \iff for right multiplication by τ .

Now consider the full set

$$(3.2) \quad \{\cdot, \setminus, //, /, \backslash\backslash, \circ\}$$

of all quasigroup operations, both basic and opposite. For current purposes, it will be convenient to use postfix notation for binary operations, setting $x \cdot y = xy\mu$ and rewriting the first equation of (2.2) in the form

$$(3.3) \quad x_1x_2\mu = x_3.$$

The full set (3.2) is construed as the homogeneous space

$$(3.4) \quad \mu^{S_3} = \{\mu^g \mid g \in S_3\}$$

for a regular right permutation action of the symmetric group S_3 , such that (3.3) is equivalent to

$$(3.5) \quad x_{1g}x_{2g}\mu^g = x_{3g}$$

for each g in S_3 . This action is known as *triality*.^a The six binary operations, in their positions corresponding to the Cayley diagram (3.1), are displayed in Figure 1.

In the figure, the opposite of each operation μ^g is given by $\mu^{\sigma g}$, so passage to the opposite operation corresponds to left multiplication by the transposition σ in the symmetric group S_3 . The three pairs of opposite operations lie in the respective columns of Figure 1.

^aTriality in this sense is not to be confused with the distinct but related notion ultimately arising from the action of S_3 on the Dynkin diagram D_4 (compare, say, [5]).

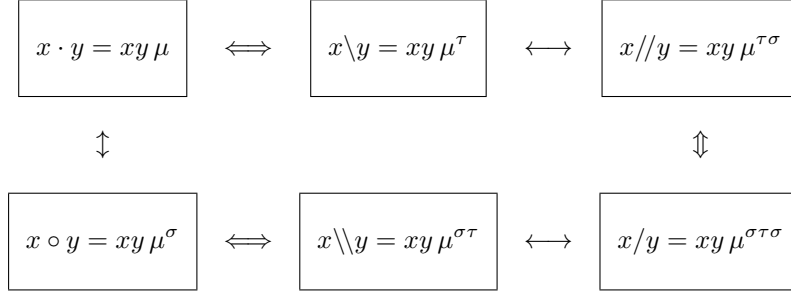


FIGURE 1. Symmetry of the quasigroup operations.

Remark 3.1. A deeper interpretation of Figure 1, and the relationship between the basic and the opposite quasigroup operations, is provided by the *orthant structure* discussed in [13, Ex. 12.1]. The basic operations constitute one set of representatives for the three doubleton orthants $\Omega^{+,+}, \Omega^{-,+}, \Omega^{+,-}$, while the opposite operations are the three remaining elements of the respective orthants.

3.2. Invertible elements of the monoid of binary words. Just as the left multiplication by σ in the symmetric group S_3 has a simple interpretation, so too does the left multiplication by τ . Let M be the complete set of all derived binary operations on a quasigroup, the binary operations that are built up by concatenation from the basic operations [10, p.293]. Together, these operations constitute the free algebra on two generators x, y in the variety \mathbf{Q} of quasigroups. A multiplication $*$ is defined on M by

$$(3.6) \quad xy(\alpha * \beta) = xxy\alpha\beta.$$

The right projection $xy\epsilon = y$ also furnishes a binary operation ϵ .

Lemma 3.2. *The set M of all derived binary quasigroup operations forms a monoid $(M, *, \epsilon)$ under the multiplication (3.6), with identity element ϵ .*

Proof. Observe that

$$xy(\alpha * \epsilon) = xxy\alpha\epsilon = xy\alpha \quad \text{and} \quad xy(\epsilon * \alpha) = xxy\epsilon\alpha = xy\alpha$$

for α in M , so ϵ is an identity element. Consider α, β, γ in M . Then

$$\begin{aligned} xy((\alpha * \beta) * \gamma) &= xxy(\alpha * \beta)\gamma \\ &= xxxy\alpha\beta\gamma \\ &= xxy\alpha(\beta * \gamma) = xy(\alpha * (\beta * \gamma)), \end{aligned}$$

confirming the associativity of the multiplication (3.6). \square

The significance of the left multiplication by τ then follows.

Proposition 3.3. *For each element g of S_3 , the binary operation μ^g is a unit of the monoid M , with inverse $\mu^{\tau g}$.*

Proof. The (IL) identity $x \setminus (x \cdot y) = y$ becomes $xy\mu\mu^\tau = y$ or $\mu * \mu^\tau = \epsilon$. Similarly the (SL) identity $x \cdot (x \setminus y) = y$ becomes $xy\mu^\tau\mu = x$ or $\mu^\tau * \mu = \epsilon$. This means that μ and μ^τ are mutual inverses.

The (IR) identity $(y \cdot x)/x = y$ is $x/(x \circ y) = y$. This becomes $xy\mu^\sigma\mu^{\tau\sigma} = y$ or $\mu^\sigma * \mu^{\tau\sigma} = \epsilon$. Similarly (SR), namely $(y/x) \cdot x = y$, may be written as $x \circ (x//y) = y$. This becomes $xy\mu^{\tau\sigma}\mu^\sigma = y$ or $\mu^{\tau\sigma} * \mu^\sigma = \epsilon$. Thus μ^σ and $\mu^{\tau\sigma}$ are mutual inverses.

The (DR) identity $(x/y) \setminus x = y$ is $x \setminus \setminus (x/y) = y$, which becomes $xy\mu^{\tau\sigma\tau}\mu^{\sigma\tau} = y$ or $\mu^{\tau\sigma\tau} * \mu^{\sigma\tau} = \epsilon$. Finally, the (DL) identity $x/(y \setminus x) = y$ is $x/(x \setminus \setminus y) = y$, which becomes $xy\mu^{\sigma\tau}\mu^{\tau\sigma\tau} = y$ or $\mu^{\sigma\tau} * \mu^{\tau\sigma\tau} = \epsilon$. Thus $\mu^{\sigma\tau}$ and $\mu^{\tau\sigma\tau}$ are mutual inverses. \square

Corollary 3.4. *The six quasigroup identities (SL), (IL), (SR), (IR), (DL), (DR) of §2.2 all take the form*

$$(3.7) \quad xxy\mu^{\tau g}\mu^g = y$$

for an element g of S_3 .

3.3. Full words and parsing trees. As described in §2.3, basic quasigroup words in the free quasigroup on an alphabet

$$(3.8) \quad \{a_1, a_2, \dots, a_s\}$$

of s generators are repeated concatenations of the letters under the three basic quasigroup operations $\cdot, /, \setminus$, and each word w is in one-one correspondence with a *basic parsing tree* T_w . It is sometimes convenient to consider an extended notion of quasigroup words and parsing trees. Thus *full quasigroup words* on the alphabet (3.8) are repeated concatenations of the generators under the full set (3.2) of all quasigroup operations, both basic and opposite. Full quasigroup words are in one-one correspondence with *full parsing trees*, which are defined recursively as follows:

- (a) For $1 \leq i \leq s$, the full parsing tree F_{a_i} is a single vertex annotated by a_i ;
- (b) For full parsing trees F_u, F_v and a basic or opposite operation μ^g from the set (3.4), the tree $F_{uv\mu^g}$ has:
 - (i) a base annotated by μ^g , along with
 - (ii) F_u as a left child, and F_v as a right child.

The basic parsing trees are precisely the full parsing trees having no node that is annotated by an opposite operation $\mu^\sigma, \mu^{\tau\sigma}$, or $\mu^{\sigma\tau}$ (compare Figure 1).

3.4. Nodal equivalence. A basic quasigroup word u of length n determines a unique basic parsing tree T_u with $n - 1$ nodes. The basic parsing tree represents a so-called *nodal equivalence class* \mathbf{F}_u of 2^{n-1} full parsing trees, sustaining a regular action of a permutation group $(S_2)^{n-1}$ known as the *nodal group* of the basic quasigroup word u . Each S_2 -factor of this group, indexed by a node of T_u , is known as the *nodal subgroup* for that node. At a given node of a full parsing tree with annotating operation μ^g , the non-trivial permutation of the nodal subgroup switches the two children of the node, and changes the node's annotation to $\mu^{\sigma g}$. It fixes the remainder of the tree.

The action of the nodal group of a basic quasigroup word on the nodal equivalence class \mathbf{F}_u of full parsing trees extends to an action on the set of corresponding full quasigroup words.

Example 3.5. Consider the basic quasigroup word $(a \cdot b)/c$ in the alphabet $\{a, b, c\}$. It determines the nodal equivalence class

$$\{F_{ab\mu c\mu^{\sigma\tau\sigma}}, F_{ba\mu^\sigma c\mu^{\sigma\tau\sigma}}, F_{cab\mu\mu^{\tau\sigma}}, F_{cba\mu^\sigma\mu^{\tau\sigma}}\}$$

of full parsing trees, represented by the basic parsing tree $T_{(a \cdot b)/c} = F_{ab\mu c\mu^{\sigma\tau\sigma}}$. Similarly, the basic quasigroup word $(a \cdot b)/c$ determines the nodal equivalence class

$$(3.9) \quad \{ab\mu c\mu^{\sigma\tau\sigma}, ba\mu^\sigma c\mu^{\sigma\tau\sigma}, cab\mu\mu^{\tau\sigma}, cba\mu^\sigma\mu^{\tau\sigma}\}$$

of full quasigroup words.

The regular action of the nodal group $(S_2)^2$ of the basic word $(a \cdot b)/c$ on the set (3.9) is displayed in the Cayley diagram

$$\begin{array}{ccc} ab\mu c\mu^{\sigma\tau\sigma} & \text{---} & ba\mu^\sigma c\mu^{\sigma\tau\sigma} \\ \parallel & & \parallel \\ cab\mu\mu^{\tau\sigma} & \text{---} & cba\mu^\sigma\mu^{\tau\sigma} \end{array}$$

where the (involutive) action of the nodal subgroup of the internal node \cdot is given by single lines, and the (involutive) action of the nodal subgroup of the root $/$ is given by double lines.

Each basic quasigroup word of length n is nodally equivalent to any member of a set of 2^{n-1} full quasigroup words. Conversely, each full quasigroup word is nodally equivalent to a unique basic quasigroup word, obtained by using nodal equivalence to replace any one of the opposite operations μ^σ , $\mu^{\tau\sigma}$, or $\mu^{\sigma\tau}$ by the respective basic operation μ , $\mu^{\sigma\tau\sigma}$, or μ^τ (compare Figure 1).

Convention 3.6. In later sections of the paper, the generic terms *quasigroup word* and *parsing tree* may refer to a basic quasigroup word or basic parsing tree, or to one of its nodally equivalent full quasigroup words or full parsing trees.

3.5. Reduced words and auxiliary bivariates. The following terminology is motivated by Evans' observation that the Diamond Lemma holds for quasigroup words [1] [2, §I.3].

Definition 3.7. (a) A (basic or full) quasigroup word is *reduced* if it will not reduce further via the quasigroup identities.

(b) A (basic or full) parsing tree representing a quasigroup word is *reduced* if its corresponding quasigroup word is reduced.

Remark 3.8. Note that a basic quasigroup word or parsing tree is reduced if and only if any of its nodally equivalent full quasigroup words or parsing trees is reduced.

Definition 3.9. Let s , a and b be positive integers. The *auxiliary bivariate* $m^s(a, b)$ denotes the number of $(a + b)$ -leaf parsing trees representing reduced quasigroup words in s arguments, with an a -leaf basic parsing tree on the left branch, a b -leaf basic parsing tree on the right branch, and a given (basic or opposite) quasigroup operation at the root vertex.

Remark 3.10. (a) The triality symmetry of the language of quasigroups (Figure 1) ensures that the auxiliary bivariate $m^s(a, b)$ is invariant under any change of the

choice of quasigroup operation at the root vertex of an $(a + b)$ -leaf parsing tree of the type considered in Definition 3.9.

(b) In particular, $m^s(a, b) = m^s(b, a)$, since the left hand side counting certain trees with μ^g at the root corresponds to the right hand side counting certain trees with the opposite operation $\mu^{\sigma g}$ at the root.

(c) By convention, whenever one of the arguments s, a, b of an auxiliary bivariate is nonpositive, the output of the auxiliary bivariate is zero.

4. COUNTING QUASIGROUP WORDS

4.1. Peri-Catalan Numbers.

Definition 4.1. Let n and s be natural numbers. The n -th s -peri-Catalan number, denoted P_n^s , gives the number of reduced basic quasigroup words of length n in the free quasigroup on an alphabet of s letters.

Since there are no constants in the language of quasigroups, $P_n^s = 0$ if $n = 0$ or $s = 0$. Now in our bookkeeping, there are C_n rooted binary trees with n leaves. Then over an s -element alphabet, each rooted binary tree may be annotated with any one of the s letters at each of its n leaves, along with any one of the three basic quasigroup operations at each of its $n - 1$ nodes (including the root), to yield a basic parsing tree. Thus the inequality

$$(4.1) \quad P_n^s \leq 3^{n-1} s^n C_n$$

provides an upper bound on the n -th s -peri-Catalan number. While the upper bound is exact only for $n < 3$, its asymptotic significance (the so-called ‘‘irrelevance of quasigroup identities’’) will be discussed in Section 5.

TABLE 1. The first ten peri-Catalan numbers for $s = 1, 2, 3$.

n	P_n^1	P_n^2	P_n^3
1	1	2	3
2	3	12	27
3	12	120	432
4	87	1,752	9,531
5	666	28,224	233,766
6	5,478	487,464	6,143,094
7	47,322	8,814,312	169,029,666
8	422,145	164,734,560	4,808,015,253
9	3,859,026	3,156,739,080	140,243,036,202
10	35,967,054	61,689,134,928	4,172,008,467,726

4.2. The inductive process. Consider $1 < n \in \mathbb{N}$ and $1 \leq k < n$. We may construct a quasigroup word of length n by adjoining a reduced word of length $n - k$ to a reduced word of length k , with one of the three basic quasigroup operations as the connective. By Definition 3.9, the number of reduced words obtained thus will be $m^s(n - k, k)$ for each connective.

Disregarding reductions, for each $1 \leq k < n$, we construct $3P_{n-k}^s P_k^s$ basic quasigroup words of length n in this fashion, and thus obtain

$$(4.2) \quad P_n^s = 3 \sum_{k=1}^{n-1} m^s(n-k, k) \leq 3 \sum_{k=1}^{n-1} P_{n-k}^s P_k^s$$

as an upper bound on the n -th s -peri-Catalan number.

The next section describes the reductions encountered in the inductive process. There, use of Convention 3.6 is made to invoke full quasigroup word equivalents of the basic quasigroup words as they are combined, in order to identify possible cancelations.

4.3. Root vertex cancelations. Let u and v be reduced quasigroup words, of respective lengths k and $n - k$, that are connected by a quasigroup operation μ^g to form $uv\mu^g$ at a step of the inductive process, as indicated by the outer level of boxes in Figure 2.

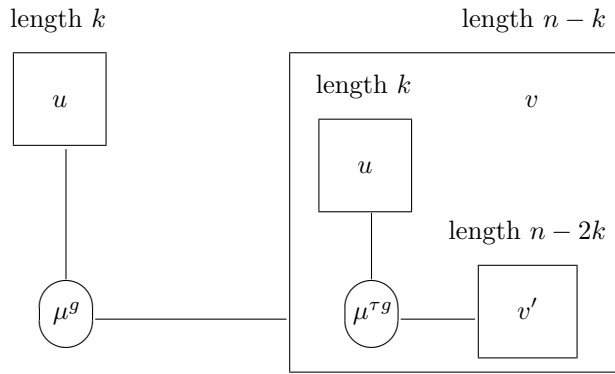


FIGURE 2. Root vertex cancelation

Proposition 4.2. *During the assembly of $uv\mu^g$ within the inductive process, with reduced word u of length k and reduced word v of length $n - k$, cancelation occurs if and only if there is a (necessarily reduced) word v' of length $n - 2k$ such that $v = uv'\mu^{\tau g}$.*

Proof. The unique cancelations available are of the form $uuv'\mu^{\tau g}\mu^g = v'$ described in Corollary 3.4. Since the word $v = uv'\mu^{\tau g}$ is reduced, it follows that the subword v' is also reduced. \square

The cancelation condition in Proposition 4.2 is illustrated by the inner level of boxes on the right hand side of Figure 2.

Proposition 4.3. *Consider $1 < n \in \mathbb{N}$ and $1 \leq k < n$.*

- (a) *The number of cancelations incurred during the inductive process when a word of length k is connected to a word of length $n - k$ by a given operation μ^g from the set (3.4) is $m^s(n - 2k, k)$. In particular, no cancelation occurs when $n = 2k$.*

(b) *The formula*

$$(4.3) \quad m^s(n-k, k) = P_{n-k}^s P_k^s - m^s(n-2k, k)$$

holds.

Proof. (a) A reduced word u of length k will appear at most once in a reduced word v of length $n-k$ with $\mu^{\tau g}$ in the root vertex. The number of such configurations (illustrated by the right-hand box in Figure 2) is given by $m^s(n-2k, k)$.

(b) The formula (4.3) results from (a) and Definition 3.9. \square

4.4. The Euclidean Algorithm. The next section presents the main recursive formula for peri-Catalan numbers. The formula is structured around the Euclidean Algorithm, using notation established in this paragraph following [14, §1.6].

For $1 < n \in \mathbb{N}$ and $1 \leq k \leq n-1$, let $r_{-1}^k = n$ and $r_0^k = k$. Consider the quotients q_l^k and remainders r_l^k for $1 \leq l \leq L_{k+1}$ as given below, resulting from calls to the Division Algorithm in the computation of $\gcd(n, k)$ by the Euclidean Algorithm:

$$(4.4) \quad r_{-1}^k = q_1^k r_0^k + r_1^k, \dots, r_{l-2}^k = q_l^k r_{l-1}^k + r_l^k, \dots, r_{L_k-1}^k = q_{L_k+1}^k r_{L_k}^k + r_{L_k+1}^k.$$

Here, $r_{L_k+1}^k = 0$ and $\gcd(n, k) = r_{L_k}^k$. Define quantities ϵ_l^k for $0 \leq l \leq L_k$ recursively by

$$(4.5) \quad \epsilon_0^k = 1 \quad \text{and} \quad \epsilon_{l+1}^k = \epsilon_l^k + q_{l+1}^k.$$

These quantities, along with the quotients q_l^k and remainders r_l^k , are used in the following section.

4.5. The recursive formula for peri-Catalan numbers.

Lemma 4.4. *Using the notation $r_{-1}^k = n$, $r_0^k = k$, $r_{-1}^k = q_1^k r_0^k + r_1^k$, and $\epsilon_0^k = 1$ from §4.4, the formula*

$$(4.6) \quad m^s(n-k, k) = (-1)^{q_1^k-1} m^s(r_0^k, r_1^k) + \sum_{j_0^k=1}^{q_1^k-1} (-1)^{\epsilon_0^k+j_0^k} P_{r_{-1}^k-j_0^k r_0^k}^s P_{r_0^k}^s$$

holds for $1 < n \in \mathbb{N}$ and $1 \leq k \leq n-1$.

Proof. It will be shown, by induction on i , that

$$(4.7) \quad m^s(n-k, k) = (-1)^i m^s(r_0^k, r_{-1}^k - (i+1)r_0^k) + \sum_{j_0^k=1}^i (-1)^{\epsilon_0^k+j_0^k} P_{r_{-1}^k-j_0^k r_0^k}^s P_{r_0^k}^s$$

for $0 \leq i < q_1^k$. Note that (4.7) for $i = q_1^k - 1$ yields (4.6). On the other hand, the base of the induction, namely (4.7) with $i = 0$, is given by Remark 3.10(b).

Now suppose that the induction hypothesis (4.7) holds for $0 \leq i < q_1^k - 1$. Then

$$\begin{aligned}
 m^s(n-k, k) &= (-1)^i m^s(r_0^k, r_{-1}^k - (i+1)r_0^k) + \sum_{j_0^k=1}^i (-1)^{\epsilon_0^k + j_0^k} P_{r_{-1}^k - j_0^k r_0^k}^s P_{r_0^k}^s \\
 &= (-1)^i \left[P_{r_{-1}^k - (i+1)r_0^k}^s P_{r_0^k}^s - m^s(r_0^k, r_{-1}^k - (i+2)r_0^k) \right] \\
 &\quad + \sum_{j_0^k=1}^i (-1)^{\epsilon_0^k + j_0^k} P_{r_{-1}^k - j_0^k r_0^k}^s P_{r_0^k}^s \\
 &= (-1)^{i+1} m^s(r_0^k, r_{-1}^k - (i+2)r_0^k) \\
 &\quad + (-1)^{\epsilon_0^k + (i+1)} P_{r_{-1}^k - (i+1)r_0^k}^s P_{r_0^k}^s + \sum_{j_0^k=1}^i (-1)^{\epsilon_0^k + j_0^k} P_{r_{-1}^k - j_0^k r_0^k}^s P_{r_0^k}^s \\
 &= (-1)^{i+1} m^s(r_0^k, r_{-1}^k - (i+2)r_0^k) + \sum_{j_0^k=1}^{i+1} (-1)^{\epsilon_0^k + j_0^k} P_{r_{-1}^k - j_0^k r_0^k}^s P_{r_0^k}^s
 \end{aligned}$$

by (4.3) and Remark 3.10(b), as required for the induction step. \square

Corollary 4.5. *The formula*

$$(4.8) \quad m^s(n-k, k) = (-1)^{\epsilon_1^k} m^s(r_0^k, r_1^k) + \sum_{j_0^k=1}^{q_1^k-1} (-1)^{\epsilon_0^k + j_0^k} P_{r_{-1}^k - j_0^k r_0^k}^s P_{r_0^k}^s$$

holds for $1 < n \in \mathbb{N}$ and $1 \leq k \leq n-1$.

Proof. Note that $(-1)^{\epsilon_1^k} = (-1)^{1+q_1^k} = (-1)^{q_1^k-1}$ by (4.5), so (4.8) is a restatement of (4.6). \square

Lemma 4.6. *Using the notation of §4.4, suppose that $m^s(n-k, k)$ is given by*

$$(4.9) \quad K(n, k) + (-1)^{\epsilon_l^k} m^s(r_{l-1}^k, r_l^k) + \sum_{i=0}^{l-1} \sum_{j_i^k=0}^{q_{i+1}^k-1} (-1)^{\epsilon_i^k + j_i^k} P_{r_{i-1}^k - j_i^k r_i^k}^s P_{r_i^k}^s$$

for $1 < n \in \mathbb{N}$, $1 \leq k \leq n-1$, and $0 < l \leq L_k$, with an additive term $K(n, k)$. Then $m^s(n-k, k) =$

$$(4.10) \quad K(n, k) + (-1)^{\epsilon_{i+1}^k} m^s(r_i^k, r_{i+1}^k) + \sum_{i=0}^l \sum_{j_i^k=0}^{q_{i+1}^k-1} (-1)^{\epsilon_i^k + j_i^k} P_{r_{i-1}^k - j_i^k r_i^k}^s P_{r_i^k}^s$$

follows.

Proof. It will be shown, by induction on h , that

$$\begin{aligned}
 (4.11) \quad m^s(n-k, k) &= K(n, k) + (-1)^{\epsilon_l^k + h} m^s(r_l^k, r_{l-1}^k - h r_l^k) \\
 &\quad + \sum_{j_i^k=0}^{h-1} (-1)^{\epsilon_l^k + j_i^k} P_{r_{l-1}^k - j_i^k r_l^k}^s P_{r_l^k}^s + \sum_{i=0}^{l-1} \sum_{j_i^k=0}^{q_{i+1}^k-1} (-1)^{\epsilon_i^k + j_i^k} P_{r_{i-1}^k - j_i^k r_i^k}^s P_{r_i^k}^s
 \end{aligned}$$

for $0 \leq h \leq q_{l+1}^k$. Recalling (4.5), note that (4.11) with $h = q_{l+1}^k$ yields the expression (4.10) for $m^s(n-k, k)$. On the other hand, the base of the induction, namely (4.11) with $h = 0$, is given by (4.9).

Starting from (4.11) with $h < q_{l+1}^k$, the induction step is given as $m^s(n-k, k) =$

$$\begin{aligned}
& K(n, k) + (-1)^{\epsilon_l^k + h} m^s(r_l^k, r_{l-1}^k - hr_l^k) \\
& + \sum_{j_l^k=0}^{h-1} (-1)^{\epsilon_l^k + j_l^k} P_{r_{l-1}^k - j_l^k r_l^k}^s P_{r_l^k}^s + \sum_{i=0}^{l-1} \sum_{j_i^k=0}^{q_{i+1}^k - 1} (-1)^{\epsilon_i^k + j_i^k} P_{r_{i-1}^k - j_i^k r_i^k}^s P_{r_i^k}^s \\
& = K(n, k) + (-1)^{\epsilon_l^k + h} \left[P_{r_{l-1}^k - hr_l^k}^s P_{r_l^k}^s - m^s(r_l^k, r_{l-1}^k - (h+1)r_l^k) \right] \\
& + \sum_{j_{l+1}^k=0}^{h-1} (-1)^{\epsilon_l^k + j_l^k} P_{r_{l-1}^k - j_l^k r_l^k}^s P_{r_l^k}^s + \sum_{i=0}^{l-1} \sum_{j_i^k=0}^{q_{i+1}^k - 1} (-1)^{\epsilon_i^k + j_i^k} P_{r_{i-1}^k - j_i^k r_i^k}^s P_{r_i^k}^s \\
& = K(n, k) + (-1)^{\epsilon_l^k + (h+1)} m^s(r_l^k, r_{l-1}^k - (h+1)r_l^k) \\
& + \sum_{j_l^k=0}^h (-1)^{\epsilon_l^k + j_l^k} P_{r_{l-1}^k - j_l^k r_l^k}^s P_{r_l^k}^s + \sum_{i=0}^{l-1} \sum_{j_i^k=0}^{q_{i+1}^k - 1} (-1)^{\epsilon_i^k + j_i^k} P_{r_{i-1}^k - j_i^k r_i^k}^s P_{r_i^k}^s
\end{aligned}$$

by (4.3) and Remark 3.10(b). \square

Proposition 4.7. *Let $1 < n \in \mathbb{N}$ and $1 < k < n$. Then $m^s(n-k, k)$ is specified by*

$$(4.12) \quad P_{r_{-1}^k}^s P_{r_0^k}^s + \sum_{i=0}^{L_k} \sum_{j_i^k=0}^{q_{i+1}^k - 1} (-1)^{\epsilon_i^k + j_i^k} P_{r_{i-1}^k - j_i^k r_i^k}^s P_{r_i^k}^s$$

in the notation of §4.4.

Proof. It will be proved, by induction on l , that $m^s(n-k, k) =$

$$(4.13) \quad P_{r_{-1}^k}^s P_{r_0^k}^s + (-1)^{\epsilon_{l+1}^k} m^s(r_l^k, r_{l+1}^k) + \sum_{i=0}^l \sum_{j_i^k=0}^{q_{i+1}^k - 1} (-1)^{\epsilon_i^k + j_i^k} P_{r_{i-1}^k - j_i^k r_i^k}^s P_{r_i^k}^s$$

holds for $0 \leq l \leq L_k$. Note that (4.13) is just (4.10) with

$$K(n, k) = P_{r_{-1}^k}^s P_{r_0^k}^s = P_n^s P_k^s$$

as the additive term.

The induction basis is given by (4.8) in Corollary 4.5. The induction step is given by Lemma 4.6. Then the proposition is given by (4.13) with $l = L_k$, since $r_{L_k+1}^k = 0$ and thus $m^s(r_{L_k}^k, r_{L_k+1}^k) = 0$. \square

Corollary 4.8. *Let $1 < n \in \mathbb{N}$ and $1 < k < n$. Then $m^s(n-k, k)$ is specified by*

$$(4.14) \quad \sum_{j_0^k=1}^{q_1^k - 1} (-1)^{\epsilon_0^k + j_0^k} P_{r_{-1}^k - j_0^k r_0^k}^s P_{r_0^k}^s + \sum_{i=1}^{L_k} \sum_{j_i^k=0}^{q_{i+1}^k - 1} (-1)^{\epsilon_i^k + j_i^k} P_{r_{i-1}^k - j_i^k r_i^k}^s P_{r_i^k}^s$$

in the notation of §4.4.

Proof. Note that $P_{r_{-1}^k}^s P_{r_0^k}^s$ cancels the $i = 0, j_i^k = 0$ summand in the term

$$\sum_{i=0}^{L_k} \sum_{j_i^k=0}^{q_{i+1}^k-1} (-1)^{\epsilon_i^k + j_i^k} P_{r_{i-1}^k - j_i^k r_i^k}^s P_{r_i^k}^s$$

of (4.12) in Proposition 4.7. Thus (4.14) specifies $m^s(n-k, k)$ in the same way that (4.12) does. \square

As there are no constants in the language of quasigroups, $P_0^s = 0$. Corollary 4.8 and (4.2) then yield the main result.

Theorem 4.9. *For $1 < n \in \mathbb{N}$, the n -th s -peri-Catalan number P_n^s is given by*

$$(4.15) \quad 3 \sum_{k=1}^{n-1} \left\{ \sum_{j_0^k=1}^{q_1^k-1} (-1)^{\epsilon_0^k + j_0^k} P_{r_{-1}^k - j_0^k r_0^k}^s P_{r_0^k}^s + \sum_{i=1}^{L_k} \sum_{j_i^k=0}^{q_{i+1}^k-1} (-1)^{\epsilon_i^k + j_i^k} P_{r_{i-1}^k - j_i^k r_i^k}^s P_{r_i^k}^s \right\}$$

in the notation of §4.4.

Example 4.10. We compute the fourth 2-peri-Catalan number as follows. For $k = 1, 2, 3$, the table

$k = 1$	$k = 2$	$k = 3$
$4 = 4 \cdot 1 + 0$	$4 = 2 \cdot 2 + 0$	$4 = 1 \cdot 3 + 1$
		$3 = 3 \cdot 1 + 0$

displays the relevant runs of the Euclidean Algorithm, yielding $L_1 = 0, L_2 = 0$, and $L_3 = 1$ in the notation of §4.4.

We then compute P_4^2 as

$$\begin{aligned} P_4^2 &= 3 \sum_{k=1}^3 \left\{ \sum_{j_0^k=1}^{q_1^k-1} (-1)^{\epsilon_0^k + j_0^k} P_{r_{-1}^k - j_0^k r_0^k}^s P_{r_0^k}^s + \sum_{i=1}^{L_k} \sum_{j_i^k=0}^{q_{i+1}^k-1} (-1)^{\epsilon_i^k + j_i^k} P_{r_{i-1}^k - j_i^k r_i^k}^s P_{r_i^k}^s \right\} \\ &= 3 \left\{ \sum_{j_0^1=1}^3 (-1)^{1+j_0^1} P_{4-j_0^1}^2 P_1^2 + \sum_{j_0^2=1}^1 (-1)^{1+j_0^2} P_{4-2j_0^2}^2 P_2^2 \right. \\ &\quad \left. + \sum_{j_0^3=1}^0 (-1)^{1+j_0^3} P_{4-3j_0^3}^2 P_3^2 + \sum_{j_1^3=0}^2 (-1)^{\epsilon_1^3 + j_1^3} P_{3-j_1^3}^2 P_1^2 \right\} \\ &= 3 \left\{ (P_3^2 P_1^2 - P_2^2 P_1^2 + P_1^2 P_1^2) + P_2^2 P_2^2 + (P_3^2 P_1^2 - P_2^2 P_1^2 + P_1^2 P_1^2) \right\} \\ &= 1,752 \end{aligned}$$

to obtain the value displayed in Table 1.

5. ASYMPTOTIC BEHAVIOR

Theorem 4.9, and in particular the form of the recursive expression (4.15) for the n -th s -peri-Catalan number P_n^s , do not readily lend themselves to an analytical asymptotic estimate comparable with the estimate

$$C_n \sim \frac{4^n}{\sqrt{\pi n^3}}$$

for the Catalan numbers that results from their generating function [3, I.2.4(33)]. Instead, we will sketch out a conjectured alternative approach, and present some numerical data in its support. The approach depends on the close relationship between the Catalan numbers and the peri-Catalan numbers.

5.1. Asymptotic irrelevance of quasigroup identities. The foundation of our approach is the following.

Conjecture 5.1 (Asymptotic irrelevance of quasigroup identities). *In the large, cancelation resulting from the quasigroup identities has a negligible effect on the asymptotic behavior of the peri-Catalan numbers P_n^s .*

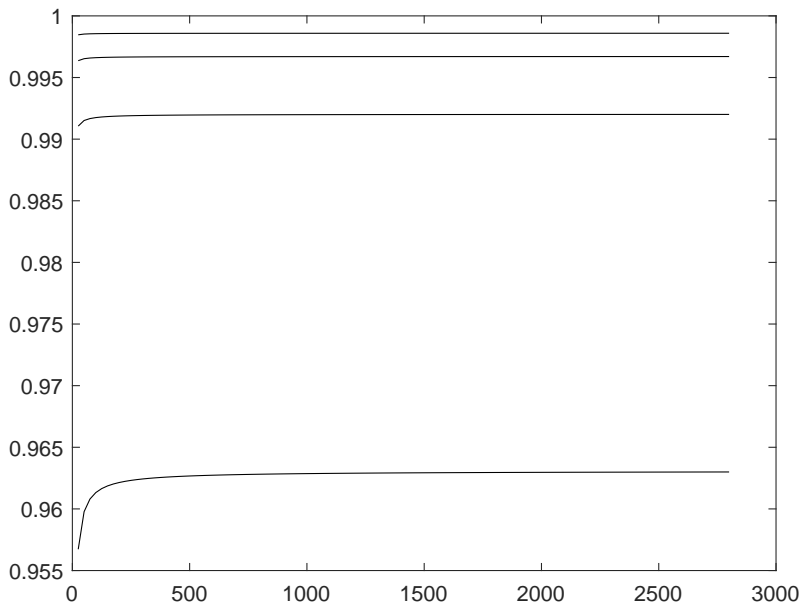
In Conjecture 5.1, the phrase “in the large” refers informally to large values of s for almost all values of n . Heuristically, referring to Figure 2, cancelation by a quasigroup identity would require the assignment of the k variables in the first copy of the word u to match exactly with the assignment of the k variables in the second copy of u . Such an event should be relatively unlikely when there is a large total of s variables available.

5.2. The effect of asymptotic irrelevance. We conjecture that

$$(5.1) \quad \lim_{s \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{\log P_n^s}{\log C_n + n \log 3s - \log 3} = 1$$

under the assumption of asymptotic irrelevance, and in particular, we conjecture that the indicated limits exist.

FIGURE 3. Plots of $\log P_n^s / (\log C_n + n \log 3s - \log 3)$ for $s = 1, 3, 6, 12$.



As initial evidence, we offer Figure 3, which displays plots of the quotients

$$(5.2) \quad \frac{\log P_n^s}{\log C_n + n \log 3s - \log 3}$$

for $s = 1, 3, 6, 12$, and $n \leq 2800$ on the horizontal axis. The curves for the respective s -values appear in order from the bottom to the top. Based on this evidence, we conjecture that, for a given value of s , the sequence of terms (5.2) increases monotonically with n . Now the denominator in (5.2) is the logarithm of the upper bound from (4.1), the total number of words of length n , on s generators, in the language of quasigroups. The numerator in (5.2) is the logarithm of P_n^s , the total number of reduced words of length n on s generators. Thus the terms (5.2) are bounded above by 1, and under our conjecture of monotonicity, the inner limit in (5.1) would exist.

Now again, based on evidence as initially exhibited in Figure 3, and then with greater detail as in §5.3 below, we conjecture that the sequence of these inner limits increases with s . The upper bound of 1 still applies, and thus the outer limit in (5.1) would exist. That this limit is 1 is essentially the content of the conjecture of asymptotic irrelevance.

While Figure 3 is useful for exhibiting qualitative features, in particular the monotonicities, it is less useful for gaining quantitative insight. Thus it is worth mentioning a second numerical experiment, where the values of $\log P_n^{12} - \log C_n$, over intervals of n up to $n = 2800$, were analyzed in Matlab. Linear regressions of these values were obtained as $3.576n - 1.102$ (with very tight error bars). The slope of the linear regression matches $\log 36 \simeq 3.583$ to three significant digits, while the (negated) intercept matches $\log 3 \simeq 1.099$ to the same accuracy.

5.3. Dependence on the number of generators. We present approximations to the values of

$$(5.3) \quad 1 - \lim_{n \rightarrow \infty} \frac{\log P_n^s}{\log C_n + n \log 3s - \log 3}$$

for $1 \leq s \leq 100$ in Table 2. The entries, which provide a numerical measure for the effect of cancelation by the quasigroup identities, are presented in the format “ $Me-N$ ” standing for $M \times 10^{-N}$. The quotient

$$\frac{\log P_{2000}^s}{\log C_{2000} + 2000 \log 3s - \log 3}$$

is used as a proxy for the limit. The function of s represented by Table 2 was fitted to the rational function

$$\frac{0.01929}{s - 0.4811}$$

by Matlab, with very tight 95% confidence bounds. On this basis, noting that the logarithm of the golden ratio is 0.4812, we formulate the following, which subsumes the unlabeled conjectures of §5.2.

Conjecture 5.2. *We have the approximate values*

$$\lim_{n \rightarrow \infty} \frac{\log P_n^s}{\log C_n + n \log 3s - \log 3} \simeq 1 - \frac{0.019\dots}{s - \log((1 + \sqrt{5})/2)}$$

for all generator counts s . In particular, the limit exists.

We do not currently have a conjectured interpretation for the numerator 0.019 . . .

TABLE 2. Decreasing impact of quasigroup identities as s increases

$1 \leq s \leq 25$	$26 \leq s \leq 50$	$51 \leq s \leq 75$	$76 \leq s \leq 100$
370e-4	561e-6	255e-6	161e-6
137e-4	536e-6	250e-6	159e-6
800e-5	514e-6	244e-6	157e-6
551e-5	493e-6	239e-6	154e-6
415e-5	474e-6	234e-6	152e-6
330e-5	456e-6	229e-6	150e-6
272e-5	439e-6	225e-6	148e-6
231e-5	424e-6	220e-6	146e-6
200e-5	409e-6	216e-6	144e-6
176e-5	396e-6	212e-6	142e-6
157e-5	383e-6	208e-6	140e-6
141e-5	371e-6	204e-6	138e-6
128e-5	359e-6	200e-6	136e-6
117e-5	349e-6	197e-6	135e-6
108e-5	340e-6	193e-6	133e-6
997e-6	329e-6	190e-6	131e-6
928e-6	320e-6	186e-6	130e-6
867e-6	311e-6	183e-6	128e-6
813e-6	303e-6	180e-6	126e-6
765e-6	295e-6	177e-6	124e-6
722e-6	288e-6	174e-6	123e-6
683e-6	281e-6	172e-6	122e-6
648e-6	274e-6	169e-6	121e-6
616e-6	268e-6	166e-6	119e-6
587e-6	261e-6	164e-6	118e-6

REFERENCES

- [1] T. Evans, “On multiplicative systems defined by generators and relations I. Normal form theorems”, *Proc. Cambridge Philos. Soc.* **47** (1951), 637–649.
- [2] T. Evans, “Varieties of loops and quasigroups”, pp. 1–26 in *Quasigroups and Loops: Theory and Applications* (O. Chein, H.O. Pflugfelder and J.D.H. Smith, eds.), Heldermann, Berlin, 1990.
- [3] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Cambridge University Press, Cambridge, 2009.
- [4] R.I. Grigorchuk, “Degrees of growth of finitely generated groups and the theory of invariant means” (Russian), *Izv. Akad. Nauk SSSR Ser. Mat.* **48** (1984), 939–985.
- [5] J.I. Hall and G.P. Nagy, “On Moufang 3-nets and groups with triality”, *Acta Sci. Math. (Szeged)* **67** (2001), 675–685.
- [6] I.M. James, “Quasigroups and topology,” *Math. Zeitschr.* **84** (1964), 329–342.
- [7] D.E. Knuth, *The Art of Computer Programming, Vol. 2. Seminumerical Algorithms*, 3rd ed., Addison-Wesley, Reading, MA, 1998.

- [8] L. Long and J.D.H. Smith, “Catalan loops”, *Math. Proc. Cambridge Philos. Soc.*, **149** (2010), no. 3, 445–453.
- [9] A. Sade, “Quasigroupes obéissant à certaines lois”, *Rev. Fac. Sci. Univ. Istanbul, Ser. A* **22** (1957), 151–184.
- [10] J.D.H. Smith and A.B. Romaowska, *Post-Modern Algebra*, Wiley, New York, NY, 1999.
- [11] J.D.H. Smith, *An Introduction to Quasigroups and Their Representations*, Chapman and Hall/CRC, Boca Raton, FL, 2007.
- [12] J.D.H. Smith, “Evans’ Normal Form Theorem revisited”, *Internat. J. Algebra Comput.* **17** (2007), 1577–1592.
- [13] J.D.H. Smith, “Groups, triality, and hyperquasigroups”, *J. Pure Appl. Algebra* **216** (2012), 811–825.
- [14] J.D.H. Smith, *Introduction to Abstract Algebra*, (2nd. ed.), Chapman and Hall/CRC, Boca Raton, FL, 2016.

¹ DEPT. OF MATHEMATICS, IOWA STATE UNIVERSITY, AMES, IOWA 50011, U.S.A.

¹ DEPT. OF MATHEMATICS & STATISTICS, SMITH COLLEGE, NORTHAMPTON, MASSACHUSETTS 01063, U.S.A.

Email address: ¹jdsmith@iastate.edu

Email address: ²stwang@smith.edu