Counting Words in Free Quasigroups

Stefanie G. Wang

Smith College
Department of Mathematics

Loops 2019
July 9, 2019

Overview

- Background on quasigroup words and rooted trees
- Quasigroup conjugates and nodal equivalence
- s-peri-Catalan numbers
- The closed formula for P_n^s
- Asymptotics for P_n^s

Quasigroup Conjugates

In an equational quasigroup $(Q, \cdot, /, \backslash)$, we have the *opposite* operations:

$$x \circ y = y \cdot x\,, \qquad x//y = y/x\,, \qquad x\backslash\backslash y = y\backslash x. \qquad (1)$$

Basic and opposite operations yield the following combinatorial quasigroups known as *conjugates* or *parastrophes*:

$$(Q, \cdot), \quad (Q, /), \quad (Q, \backslash), \quad (Q, \circ), \quad (Q, //), \quad (Q, \backslash\backslash) \qquad (2)$$

The identities (IR) in $(Q, \backslash)$ and (IL) in (Q,/) yield the respective identities

$$\text{(DL)} \quad x/(y\backslash x) = y\,,$$
$$\text{(DR)} \quad y = (x/y)\backslash x$$

Basic parsing trees

In the free quasigroup on an alphabet $\{a_1, a_2, \ldots, a_s\}$, (*basic*) *quasigroup words* are repeated concatenations of the generators under the three basic quasigroup operations $\cdot, /, \backslash$. 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;

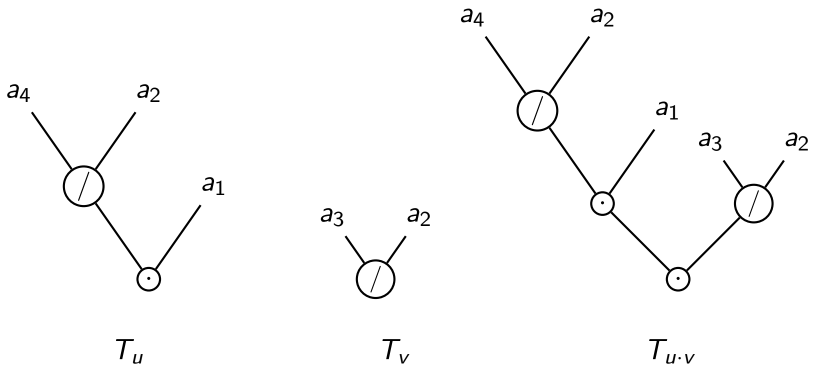
(b) For basic words u, v, the tree $T_{u \cdot v}$ has:
  (i) a root annotated by the multiplication,
  (ii) T_u as a left child, and T_v as a right child;

(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 $T_{u \backslash v}$ has:
  (i) a root annotated by the left division,
  (ii) T_u as a left child, and T_v as a right child.

Basic parsing trees

Let $u = (a_4/a_2) \cdot a_1$ and $v = a_3/a_2$.



Action of S_3 on quasigroup operations

Writing σ and τ for the respective transpositions (12) and (23), the full set $\{\cdot, \backslash, //, /, \backslash\backslash, \circ\}$ of basic and opposite quasigroup operations is construed as the homogeneous space

$$\mu^{S_3} = \{\mu^g \mid g \in S_3\} \tag{3}$$

for a regular right permutation action of the symmetric group S_3.

$$\begin{array}{ccccc}
\boxed{x \cdot y = xy\,\mu} & \Longleftrightarrow & \boxed{x \backslash y = xy\,\mu^{\tau}} & \longleftrightarrow & \boxed{x // y = xy\,\mu^{\tau\sigma}} \\
\updownarrow & & & & \Updownarrow \\
\boxed{x \circ y = xy\,\mu^{\sigma}} & \Longleftrightarrow & \boxed{x \backslash\backslash y = xy\,\mu^{\sigma\tau}} & \longleftrightarrow & \boxed{x / y = xy\,\mu^{\sigma\tau\sigma}}
\end{array}$$

Monoid of binary words

Let M be the complete set of all derived binary operations on a quasigroup. A multiplication $*$ is defined on M by

$$xy(\alpha * \beta) = x\,xy\alpha\,\beta\,. \tag{4}$$

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

Lemma

The set M of all derived binary quasigroup operations forms a monoid $(M, *, \epsilon)$ under the multiplication (4), with identity element ϵ.

Proposition

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

Corollary

The six quasigroup identities $(\mathrm{SL}), (\mathrm{IL}), (\mathrm{SR}), (\mathrm{IR}), (\mathrm{DL}), (\mathrm{DR})$ *all take the form*

$$x\,xy\mu^{\tau g}\,\mu^g = y \tag{5}$$

for an element g of S_3.

Full parsing trees

Full quasigroup words on the alphabet $\{a_1, a_2, \ldots, a_s\}$ are repeated concatenations of the generators under the full set $\{\cdot, \backslash, //, /, \backslash\backslash, \circ\}$.

(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), 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.

Nodal equivalence

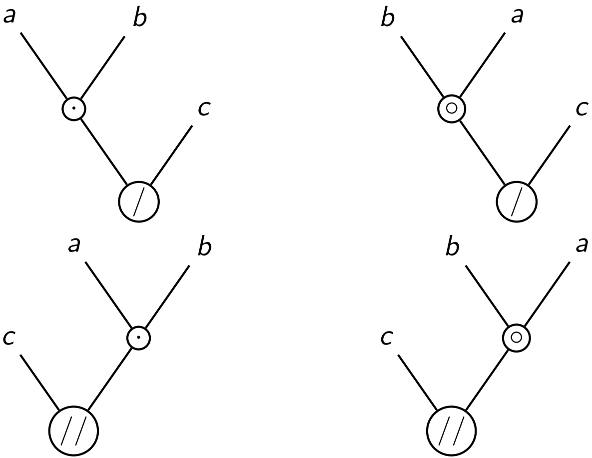
- The basic parsing tree represents a so-called *nodal equivalence class* $\mathbf{F}_u$ of 2^{n-1} full parsing trees, sustaining a regular action of a permutation group $(S_2)^{n-1}$ known as the *nodal group* of the basic quasigroup word u.
- At a given node of a full parsing tree with annotating 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.

Nodal equivalence example

Consider the basic quasigroup word $(a \cdot b)/c$. It determines the nodal equivalence class

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

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


a
b
c
b
a
c
a
b
c
b
a
c

s-peri-Catalan numbers

Definition

(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.

Definition

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.

Auxiliary bivariate function

Definition

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 child, a b-leaf basic parsing tree on the right child, and a given (basic or opposite) quasigroup operation at the root vertex.

Auxiliary bivariates and nodal equivalence

- 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 6.
- 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.
- By convention, whenever one of the arguments s, a, b of an auxiliary bivariate is nonpositive, the output of the auxiliary bivariate is zero.

To construct a length n quasigroup word:

1. Adjoin a reduced word u of length $n - k$ to a reduced word v of length k, and
2. take one of the three basic quasigroup operations as the connective.

Then

$$P_n^s = 3\sum_{k=1}^{n-1} m^s(n-k,k) \leq 3\sum_{k=1}^{n-1} P_{n-k}^s P_k^s \tag{6}$$

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

Number of reductions

Proposition

During the assembly of $uv\,\mu^{g}$ *within the inductive process, cancellation 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 cancellations available are of the form $u\,uv'\mu^{\tau g}\,\mu^{g} = v'$ described in Corollary 3. Since the word $v = uv'\,\mu^{\tau g}$ is reduced, it follows that the subword v' is also reduced. □

Root vertex cancellation

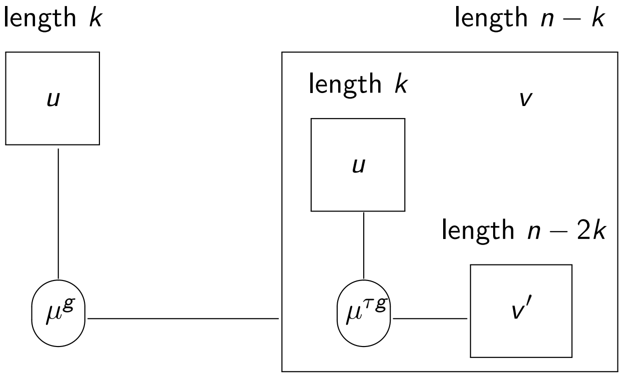


Figure: Root vertex cancellation

Unpacking the auxiliary bivariate

Proposition

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

(a) *The number of cancellations 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 *is* $m^s(n-2k,k)$. *In particular, no cancellation occurs when* $n = 2k$.

(b) *The formula*

$$m^s(n-k,k) = P^s_{n-k}P^s_k - m^s(n-2k,k) \tag{7}$$

holds.

Euclidean Algorithm notation

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:

$$r_{-1}^k = q_1^k r_0^k + r_1^k \,, \ldots, r_{l-2}^k = q_l^k r_{l-1}^k + r_l^k \,, \ldots, \tag{8}$$

$$r_{L_k-1}^k = q_{L_k+1}^k r_{L_k}^k + r_{L_k+1}^k \,. \tag{9}$$

Here, $r_{L_k+1}^k = 0$ and $\gcd(n,k) = r_{L_k}^k$.

$$\epsilon_0^k = 1 \quad \text{and} \quad \epsilon_{l+1}^k = \epsilon_l^k + q_{l+1}^k \,. \tag{10}$$

Lemma

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

$$m^s(n-k,k) = (-1)^{q_1^k - 1} m^s(r_0^k, r_1^k) + \sum_{j_0^k=1}^{q_1^k - 1} (-1)^{\epsilon_0^k + j_0^k} P^s_{r_{-1}^k - j_0^k r_0^k} P^s_{r_0^k} \tag{11}$$

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

Proof of Lemma

Proof:

Through induction on i, we will show:

$$m^s(n-k,k) = (-1)^i m^s\left(r_0^k, r_{-1}^k - (i+1) r_0^k\right) + \sum_{j_0^k=1}^{i} (-1)^{\epsilon_0^k + j_0^k} P^s_{r_{-1}^k - j_0^k r_0^k} P^s_{r_0^k} \tag{12}$$

for $0 \leq i < q_1^k$. Note that (12) for $i = q_1^k - 1$ yields (11). On the other hand, the base of the induction, namely (12) with $i = 0$, is given by $m^s(a,b) = m^s(b,a)$.

Proof cont.

Now suppose that the induction hypothesis (12) holds for $0 \leq i < q_1^k - 1$. Then

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

by (7) and $m^s(a,b) = m^s(b,a)$, as required for the induction step. $\square$

Full unpacking of $m^s(n-k,k)$

Proposition

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

$$\sum_{j_0^k=1}^{q_1^k-1}(-1)^{\epsilon_0^k+j_0^k}P^s_{r_{-1}^k-j_0^k r_0^k}P^s_{r_0^k}+\sum_{i=1}^{L_k}\sum_{j_i^k=0}^{q_{i+1}^k-1}(-1)^{\epsilon_i^k+j_i^k}P^s_{r_{i-1}^k-j_i^k r_i^k}P^s_{r_i^k}.$$

Proof sketch:

Induction basis given by previous lemma. For the induction step, suppose $m^s(n-k,k) =$

$$\sum_{j_0^k=1}^{q_1^k-1} (-1)^{\epsilon_0^k + j_0^k} P^s_{r_{-1}^k - j_0^k r_0^k} P^s_{r_0^k} + (-1)^{\epsilon_{l+1}^k} m^s(r_l^k, r_{l+1}^k) \tag{13}$$

$$+ \sum_{i=1}^{l} \sum_{j_i^k=0}^{q_{i+1}^k-1} (-1)^{\epsilon_i^k + j_i^k} P^s_{r_{i-1}^k - j_i^k r_i^k} P^s_{r_i^k} \tag{14}$$

for $0 < l < L_k$.

The formula for P_n^s

Previously:

$$P_n^s = 3\sum_{k=1}^{n-1} m^s(n-k,k)$$

The formula for P_n^s

Previously:

$$P_n^s = 3\sum_{k=1}^{n-1} m^s(n-k,k)$$

Theorem

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

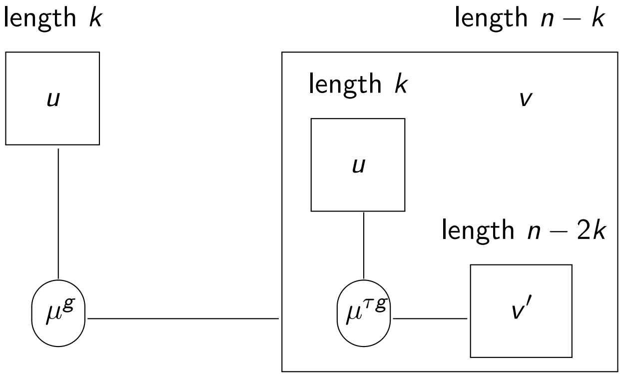
$$3\sum_{k=1}^{n-1}\left\{\sum_{j_0^k=1}^{q_1^k-1}(-1)^{\epsilon_0^k+j_0^k}P^s_{r^k_{-1}-j_0^k r_0^k}P^s_{r_0^k}+\sum_{i=1}^{L_k}\sum_{j_i^k=0}^{q_{i+1}^k-1}(-1)^{\epsilon_i^k+j_i^k}P^s_{r^k_{i-1}-j_i^k r_i^k}P^s_{r_i^k}\right\}.$$

First ten s-peri-Catalan numbers

Table: 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

Recall, cancellations must have the following format:



Numerical observations

- For any n, s, the most cancellations occur when $k = 1$.
- When adjoining an arbitrary length k reduced word to an arbitrary $n - k$ reduced word, the probability of a cancellation occurring is $< 1/P_k^s$.
- There are $3^{n-1}s^n C_n$ magma words of length n in s generators.

Growth of magma words and quasigroup words

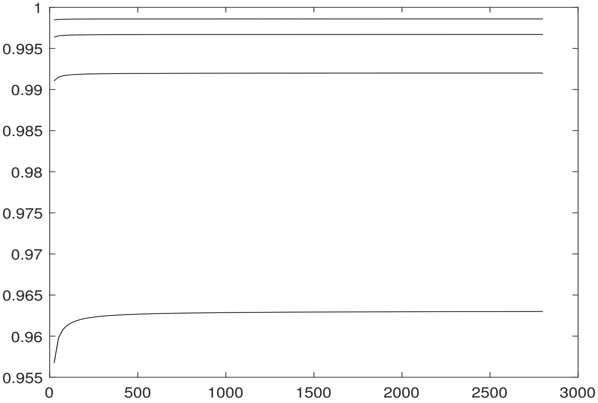
Conjecture (Asymptotic irrelevance of quasigroup identities)

In the large, cancellation resulting from the quasigroup identities has a negligible effect on the asymptotic behavior of the peri-Catalan numbers P_n^s.

We conjecture that

$$\lim_{s\to\infty}\lim_{n\to\infty}\frac{\log P_n^s}{\log C_n + n\log 3s - \log 3} = 1 \tag{15}$$

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



Thank you for your attention!