Generalizing Rules of Three

Ewin Tang*

August 20, 2017

1 Introduction

This paper explores The Rule of Three, the phenomenon where for some sequences in a ring, all elements in the sequence satisfy a commutation relation precisely when the relation is satisfied by subsets of size three and fewer.

This behavior has been seen in some notable cases. Kirillov [3] shows that elementary symmetric polynomials in noncommuting variables commute (and, in some cases, all Schur functions) when elementary symmetric polynomials of degree at most three commute when restricted to at most three of the variables. Generalizing this, Blasiak and Fomin [1] give a wider theory for rules of three of generating functions over rings via rules of three in sums and products.

In this report we detail our attempts to apply the general theory to interesting specific cases. Namely, we consider Schur Q-functions and loop symmetric functions. In the former case, we give a conjecture for a rule of three, and give progress towards a proof for the said conjecture, emulating the structure of the proof given for super elementary symmetric functions in [1]. In the latter case, we find negative results.

2 Background

First, we give some notation. Let $e_k$ and $h_k$ denote the $k$th elementary symmetric polynomial and $k$th homogeneous symmetric polynomial, respectively. We use the standard definition, in

---

*This research was carried out as part of the 2017 REU program at the School of Mathematics at the University of Minnesota, Twin Cities, and was supported by NSF RTG grant DMS-1148634. The author would like to thank Elizabeth Kelley, Pavlo Pylyavskyy, and Vic Reiner for their invaluable mentorship and support.
both the commutative and noncommutative sense:

\[
e_k(x_1, \ldots, x_n) = \sum_{1 \leq j_1 < \cdots < j_k \leq n} x_{j_k} \cdots x_{j_1}
\]

\[
h_k(x_1, \ldots, x_n) = \sum_{1 \leq j_1 \leq \cdots \leq j_k \leq n} x_{j_1} \cdots x_{j_k}
\]

For \(x_1, \ldots, x_n\) and \(S \subset [n]\) with \(S = \{s_1 < s_2 < \cdots < s_k\}\), let \(e_i(x_S)\) denote the polynomial with \(x_{s_1}\) through \(x_{s_k}\) as variables. However, when speaking of \(x_S\) in other contexts, it will denote the product \(x_{s_k}x_{s_{k-1}} \cdots x_{s_1}\). Note that we multiply in descending order; when we want to explicitly give the order for multiplying we will add an arrow in the superscript. For example, \(x^\uparrow_S\) will be multiplying in ascending order. Finally, let \([x, y] = xy - yx\) be the commutator in the standard sense.

[6] gives some definitions that we reproduce here. A strict partition of \(n\) is a sequence \(\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_\ell) \in \mathbb{Z}^\ell\) such that \(\lambda_1 > \lambda_2 > \cdots > \lambda_\ell\). The corresponding shifted diagram of \(\lambda\) is an array of square cells in which the \(i\)th row has \(\lambda_i\) cells, and is shifted \(i - 1\) units to the right with respect to the top row.

A (semistandard) shifted Young tableau \(T\) of shape \(\lambda\) is a filling of a shifted diagram \(\lambda\) with letters from the alphabet \(A = \{1', 1, 2', 2, \cdots\}\) such that:

- Rows and columns are weakly increasing;
- Each column has at most one \(k\) for \(k \in \{1, 2, \cdots\}\);
- Each row has at most one \(k\) for \(k \in \{1', 2', \cdots\}\).

These allow us to define the Schur Q-functions \(Q_\lambda\) for a shifted partition \(\lambda\).

**Definition.** \(Q_\lambda(x_1, \ldots, x_n) \in Q[x_1, \ldots, x_n]\) can be defined in any of the following equivalent ways: [7]

(a) For a tableau \(T\) with content \((a_1, a_2, \ldots)\) (we ignore the distinction between primed and unprimed entries here), define \(x^T = x_1^{a_1}x_2^{a_2} \cdots\). Then \(Q_\lambda = \sum_{\text{shape}(T) = \lambda} x^T\), where we sum over shifted tableaus \(T\).

(b) Let \(q_k\) be the \(k\)th coefficient of the product of generating functions \(\prod_{i=1}^n f_i\), where

\[
f_i = \frac{1 + x_i t}{1 - x_i t} = 1 + 2(x_i t + x_i^2 t^2 + \cdots)
\]

And let

\[
Q_{(k_1, k_2)} = q_{k_1}q_{k_2} + \sum_{i=1}^{k_2} (-1)^i q_{k_1+i} q_{k_2-i}
\]

Then for \(\lambda\) with \(\ell\) parts,

\[
Q_\lambda = \begin{cases} 
\text{pf}[Q_{(\lambda_i, \lambda_j)}]_{1 \leq i < j \leq \ell} & \ell \text{ even} \\
\text{pf}[Q_{(\lambda_i, \lambda_j)}]_{1 \leq i < j \leq \ell+1} & \ell \text{ odd}
\end{cases}
\]

where pf is the Pfaffian, and \(\lambda_{\ell+1} = 0\) when \(\ell\) is odd.
These are symmetric (although that is not obvious), and in fact, the Schur Q-functions generate a subalgebra of the space of symmetric functions, namely (5 2.5)

\[ Q[k] = Q[p_{2k+1}] \quad \text{where} \quad p_k = \sum_i x_i^k \]

Further, (5 2.6)

\[ q_k = \sum_{i_1, i_3, \ldots, i_p} \frac{2^{i_1+i_3+\cdots+i_p}}{i_1! \cdots i_p!} \left( \frac{p_1}{1} \right)^{i_1} \left( \frac{p_2}{p} \right)^{i_2} \]

So this immediately gives us that \( q_{2k+1} \) is an algebraically independent generating set of \( \Lambda \) (at least in the commutative case).

3 Results

3.1 Schur-Q functions, noncommutative case hook reading

To move Schur Q-functions to the noncommutative case, we need to decide on a method for reading the tableau. Following [6], we will use the hook reading, which preserves elementary properties of commutative Schur Q-functions. In this context, \( Q(k) \), which we will refer to as \( q_k \), is defined as follows:

**Definition.** \( q_k(x_1, \ldots, x_n) \in \mathbb{Q}\langle x_1, \ldots, x_n \rangle \) can be defined in any of the following equivalent ways:

(a) The standard tableaux definition of \( q_k \), except \( x^T \) is defined by reading the unprimed entries right to left, then the primed entries left to right. For example, if

\[ T = \begin{array}{ccccccc}
1 & 1' & 2' & 4 & 5' & 5 & 6
\end{array}\]

then \( x^T = (x_6 x_4 x_1)(x_1 x_2 x_5 x_5) \).

(b) Let \( g_i = 1 + x_i t \) and let \( h_i = (1 - x_i t)^{-1} \). \( q_k \) is the \( k \)th coefficient of the product of generating functions

\[ \prod_{i=n}^1 g_i \prod_{i=n}^1 h_i \]

(c) \( q_k \) is the sum of monomials whose subscripts strictly decrease, then nonstrictly increase, with a factor of two added.

Evidence from Sage suggests that a rule of 3 is possible. Namely, with three commuting variables, \([q_3, q_5]\) is contained in the ideal generated by all of the relations only using \( q_k \) for \( k \) at most three. Further, with two commuting variables, all relations of degree at most \( k + 1 \) (that is, \([q_i, q_j]\) for \( i + j \leq k + 1 \)) are contained in the ideal generated by the relations involving \( q_i \) and \( q_1 \) for \( i \) at most \( k \). Notice that we only need to worry about \( q_k \) for \( k \) odd, since commutation relations hold for even \( k \) when they hold for odd \( k \) by [2].
Consider some noncommuting variables $x_1, \ldots, x_N$, and notate the elements of $S \subset [N]$ by $S = \{s_1 < s_2 < \cdots < s_n\}$. Then we define $q_k(x_S)$ as the Schur Q-function in $n$ variables where $x_i$ is replaced with $x_{s_i}$. We conjecture the following:

**Conjecture 3.1.** Let $x_1, \ldots, x_N, y_1, \ldots, y_N$ be elements of a ring $A$. The following are equivalent:

- $q_k(x_S)$ and $q_\ell(y_S)$ commute for all $S, k, \ell$.
- the above holds when $k = 1$ or $\ell = 1$, and for all $S$

We have found no evidence to support anything stronger; the statement does not hold when the second case is restricted to $|S| \leq 3$, and it does not hold when the second case specifies $kl \leq 5$.

We can rewrite this in terms of generating functions, as done in [1], most similarly to Lemma 8.2 in the cited paper.

**Conjecture 3.2.** Let $A$ be a ring, and let $x_1, \ldots, x_N, y_1, \ldots, y_N \in A$ satisfy

$[x_a, y_b] = [y_a, x_b]$ for all $1 \leq a \leq b \leq N$

Further, let

$$a_i = 1 + x_i t, \quad b_i = 1 - x_i t,$$

$$\alpha_i = 1 + y_i s, \quad \beta_i = 1 - y_i s$$

Notice that this implies that all expresssions of the same index commute: $x_i, y_i, a_i, b_i, \alpha_i, \beta_i$ commute with each other (along with their inverses). Further, suppose that the following relations are satisfied:

$$\left[ \sum_{i \in S} x_i, \alpha_S(\beta_S)^{-1} \right] = 0$$

$$\left[ \sum_{i \in S} y_i, a_S(b_S)^{-1} \right] = 0$$

Then $a_{[N]}(b_{[N]})^{-1} \alpha_{[N]}(\beta_{[N]})^{-1} = \alpha_{[N]}(\beta_{[N]})^{-1} a_{[N]}(b_{[N]})^{-1}$.

We would like to give this statement without explicitly defining $a_i, b_i, \alpha_i$, and $\beta_i$, instead giving only the relations that are necessary for the proof. However, we do not know the full list of relations necessary. This is covered more in [3,5]

We can prove that rephrasing the conjecture in terms of generating functions indeed gives us an equivalent statement.

**Theorem 3.3.** 3.1 and 3.2 are equivalent.
Proof. Using $a_i, b_i, \alpha_i, \beta_i$ as specified, the two relations in [3.1] are equivalent as follows:

\[
\sum_{i \in S} x_i, a_S(\beta_S)^{-1} = \begin{bmatrix} q_1(x_S), \sum_{i \geq 0} q_i(y_S)s^i \end{bmatrix} = 0 \iff [q_1(x_S), q_i(y_S)] = 0 \ \forall i
\]

\[
\sum_{i \in S} y_i, a_S(b_S)^{-1} = \begin{bmatrix} q_1(y_S), \sum_{i \geq 0} q_i(x_S)t^i \end{bmatrix} = 0 \iff [q_1(y_S), q_i(x_S)] = 0 \ \forall i
\]

The $[x_a, y_b] = [y_a, x_b]$ condition is implied by $[q_1, q_1] = 0$ for $|S| \leq 2$. Further,

\[
[a_n(b_n)^{-1}, a_n(\beta_n)^{-1}] = \begin{bmatrix} \sum_{i \geq 0} q_i(x_S)t^i, \sum_{j \geq 0} q_j(y_S)s^j \end{bmatrix} = 0
\]

if and only if $[q_i(x_n), q_j(y_n)] = 0 \ \forall i, j$

While this only gives us the statement for the set of variables indexed by $[n]$, the set of conditions hold for all $S \subset [n]$, so we can restrict to some $T$ and get the statement for every subset as well.

We can reproduce some lemmas that act as an analogue for those in the proof of Lemma 8.2 in [1].

**Lemma 3.4** (6.1 analogue). Let $R$ be a ring, and let $x_i, x_j, \alpha_i, \beta_i, \alpha_j, \beta_j$ satisfy the following conditions:

- all elements of the same index commute
- $[x_i + x_j, \alpha_j \alpha_i \beta_i \beta_j] = 0$

Then $(\alpha_j [x_i, x_j] - [x_i, \alpha_j] \alpha_i \beta_i \beta_j = \alpha_j \alpha_i ([x_j, \beta_i] \beta_j - \beta_i [\beta_j, x_i])$.

**Remark 3.5.** In the proof of Blasiak-Fomin 8.2, this lemma is enough to prove the statement for $N = 2$. However, considering the two equations we get from applying 3.4 to the relations from [3.2]

\[
(\alpha_2 \alpha_1, x_2) - [x_2, \alpha_2 \alpha_1] \beta_2^{-1} \beta_2^{-1} = \alpha_2 \alpha_1 ([x_2, \beta_1^{-1}] \beta_2^{-1} - \beta_1^{-1} \beta_2^{-1}, x_1)]
\]

\[
(a_2 a_1, y_2) - [y_2, a_2 a_1] b_1^{-1} b_2^{-1} = a_2 a_1 ([y_2, b_1^{-1}] b_2^{-1} - b_1^{-1} [b_2^{-1}, y_1])
\]

Sage computation confirms that these relations, along with the commutation relations given in [3.2] are not enough to prove that $a_2 a_1 b_1 b_2$ and $\alpha_2 a_1 \beta_1 \beta_2$ commute. Thus, an attempt at a proof must use more than what is presented in this paper; we suspect this may involve a relation in $b_s$ and $\beta_s$, since we have a relation among $a_s$ and $\alpha_s$; arithmetic suggests that having more analogous relations would give progress towards the desired result.

This also shows that we cannot generalize [3.2] to arbitrary $a_i$, $b_i$, $\alpha_i$, and $\beta_i$, as is done in Lemma 8.2 in Blasiak-Fomin. We will need to find the complete list of relations necessary in the specific case to do so.
We have also considered the weakened version of allowing \( x_i \) and \( x_j \) to commute when \( |i - j| > 1 \), as is done in [2].

**Conjecture 3.6.** [3.1 and 3.2 hold when \([x_i, x_j] = [x_i, y_j] = [y_i, y_j] = 0 \) for \(|i - j| > 1\).

This allows us to give an analogue to 6.4.

**Lemma 3.7 (6.4 analogue).** Suppose:

\[
\left[ \sum_{i \in S} x_i, \alpha_S^\uparrow \beta_S^\downarrow \right] = 0 \text{ for all } S \subset \{1, \ldots, N\}
\]

And suppose \([x_i, \alpha_k] = [x_i, \beta_k] = 0 \) for nonadjacent \( i, k \). Then, for any \( S = \{s_1 < \cdots < s_m\}, \)

\[
([x_{s_{m-1}}, \alpha_{s_{m-1}} - \alpha_{s_{m}}[\alpha_{s_{m-1}}, x_{s_{m}}], \alpha_U \beta_S = \alpha_S \beta_U \beta_{s_{m-1}}[\beta_{s_{m-1}}, x_{s_{m-1}}] - [x_{s_{m}}, \beta_{s_{m-1}}] \beta_{s_{m}})
\]

**Proof.** Let \( T = S \setminus s_m \) and let \( U = S \setminus \{s_m, s_{m-1}\} \). Then

\[
\left[ \sum_{i \in S} x_i, \alpha_S^\uparrow \beta_S^\downarrow \right] = \left[ x_{s_m}, \alpha_{s_m} \alpha_T \beta_T \beta_{s_m} \right] + \left[ \sum_{i \in T} x_i, \alpha_{s_m} \alpha_T \beta_T \beta_{s_m} \right]
\]

\[
= \alpha_{s_m} [x_{s_m}, \alpha_T \beta_T] \beta_{s_m} + \left[ \sum_{i \in T} x_i, \alpha_{s_m} \right] \alpha_T \beta_S + \alpha_S \beta_T \left[ \sum_{i \in T} x_i, \beta_{s_m} \right]
\]

And the statement follows.

3.2 Negative Results

When generalizing Schur \( Q \)-functions to the noncommutative case, we can also consider the less natural choice of reading shifted tableaus in descending order from right to left. This is equivalent to using the generating polynomial \( f \) in the non-commutative setting.

That is, define \( q_k \) in the non-commutative case as the \( k \)th coefficient of the product of generating functions \( \prod f_i \) from 1 to \( n \), where

\[
f_i = (1 + x_i t)(1 - x_i t)^{-1}
\]

In this case, we immediately get the following result from Theorem 2.5 of [1]:

**Corollary 3.8.** The following are equivalent:

- \([q_k(u_S), q_\ell(u_S)] = 0 \) for any \( k, \ell, S \subset N \)
- the above holds for \( |S| = 2, 3 \).
However, this theorem, unlike typical rules of three, does not allow for two sets of non-commuting variables. Notice that Conjecture 2.3 of [1], if proved, would allow for two sets of non-commuting variables.

Also notice that this corollary has no restriction on $k$ and $\ell$. Computational evidence suggests that we cannot restrict $k, \ell \leq 5$, and we also cannot restrict to either $k$ or $\ell$ being 1.

We define loop elementary symmetric functions on $n$ flavors through the generating function called the whirl matrix, [4]

\[
\begin{bmatrix}
1 & a_i^{(1)} & 0 & \cdots & 0 \\
0 & 1 & a_i^{(2)} & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
a_i^{(n)} & 0 & 0 & \cdots & 1
\end{bmatrix}
\]

The product of these are

\[P_{ij} = \sum_{k=0}^{\frac{i-j+m}{n}} e^{(i)}_{kn-(i-j)} t^k\]

where the superscript on $e$ is the flavor of the first element in all of the monomials, $n$ is the number of flavors, and $m$ is the number of generating functions in the product.

Computing the three-flavor three-variable case suggests that a rule of three for these functions does not exist: when degree one loop symmetric polynomials commute with degree two and degree three polynomials, we do not get commutation between degree two and degree three polynomials. Because we are working with only one set of non-commuting variables, this has been true for all of the rules of three mentioned in this paper.

4 Further Directions

In following the form of Blasiak and Fomin’s proof of Lemma 8.2, we have achieved the following in the standard case and in the weakened case (where nonadjacent variables commute).

\[
\begin{align*}
6.1 & \quad 6.2 & \quad 6.3 \\
8.1 & \quad \triangle & \quad 6.4 \\
& \quad \square & \\
\end{align*}
\]

In the diagram, $\triangle$ represents the $N = 2$ case, and $\square$ represents a full proof. Thus, any progress on 8.1 would result in a significant result.

Beyond proving the conjecture, we can also ask when commutativity of $q_k$, the elementary Schur Q-functions, extends to all Schur Q-functions, as is explored in [2]. In the weakened
case, this phenomenon may occur, since the question reduces to deciding whether the Pfaffian formula holds true in the noncommutative case.

A Code

All code mentioned is located at: https://github.com/ewin-t/ruleof3qschur

The README in the repository explains the precise content of the code.

References