**RESEARCH STATEMENT**

Mingjia Yang

## 1. Introduction

While I am broadly interested in all kinds of mathematics, my graduate study and research focus is in combinatorics. In particular, I study patterns in permutations and words, and more recently, integer partitions. Side note: my favorite area of math in high school is geometry (I studied in China before college) and my undergraduate favorite area is abstract algebra. But at this moment, it is experimental math and combinatorics! I find it a lot of fun to discover patterns using computers and then try to prove them by various means (or leave them as conjectures). Below I will summarize the main projects that I have worked on during my graduate school years and possibilities for future work. Many of them are accessible to undergraduate students. The first three projects (2.1.-2.3.) are especially of current interest.

## 2. Projects

2.1. **Systematic counting of restricted partitions.** [This is joint work with Doron Zeilberger]

Project link: `http://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/rpr.html`

One of the cornerstones of enumerative combinatorics (and number theory!) are integer partitions. Recall that a partition of a non-negative integer $n$ is a list of integers $(\lambda_1, \ldots, \lambda_k)$ such that $\lambda_1 \geq \cdots \geq \lambda_k \geq 1$ and $\lambda_1 + \cdots + \lambda_k = n$.

Already Euler considered the enumeration of sets of partitions obeying some *restrictions*. For example the number of partitions into **distinct** parts, let's call it $d(n)$, is given by the generating function ([An], p. 5)

$$\sum_{n=0}^{\infty} d(n) q^n = \prod_{i=1}^{\infty} (1+q^i) = \prod_{i=0}^{\infty} \frac{1}{1-q^{2i+1}} \quad .$$

More recently, Rogers and Ramanujan (with the help of MacMahon, see [1] and [25]) considered the problem of enumerating partitions with the property that the difference between consecutive parts is at least 2, i.e. for which

$$\lambda_i - \lambda_{i+1} \geq 2 \quad .$$

The First Rogers-Ramanujan identity states that these numbers, let's call them $d_2(n)$, also have a nice product generating function

$$\sum_{n=0}^{\infty} d_2(n) q^n = \prod_{i=0}^{\infty} \frac{1}{(1-q^{5i+1})(1-q^{5i+4})} \quad .$$

We can say that distinct partitions **avoid** the "pattern" $[a, a]$ and Rogers-Ramanujan partitions avoid both the pattern $[a, a]$ and the pattern $[a, a - 1]$.

This naturally leads to the question of enumerating partitions avoiding an *arbitrary* (finite) set of patterns, but first let's formally define the notion of a "pattern" in the context of partitions. Note that, for simplicity, definition for "pattern" is slightly different from above.

**Definition 2.1.** *A partition-pattern (pattern for short) is a list* $a = [a_1, \dots, a_r]$ *of length* $r \geq 1$ *of non-negative integers. A partition* $\lambda = (\lambda_1, \dots, \lambda_k)$ **contains** *the pattern* $a = [a_1, \dots, a_r]$ *if there exists* $1 \leq i \leq k - r$ *such that*

$$\lambda_i - \lambda_{i+1} = a_1 \quad , \quad \lambda_{i+1} - \lambda_{i+2} = a_2 \quad , \quad \dots \lambda_{i+r-1} - \lambda_{i+r} = a_r \quad .$$

Our goal is to devise an efficient algorithm, that inputs an **arbitrary** set of patterns, $P$, and an arbitrary positive integer $N$, and outputs the first $N$ terms of the sequence enumerating partitions of $n$ avoiding the set of patterns $P$.

A natural approach is to adapt the celebrated Goulden-Jackson [13] method to this new context. The Goulden-Jackson method is lucidly explained (and significantly extended) in the article [22]. I was able to implement this approach in Maple. However, as it turned out, while this approach is of considerable *theoretical* interest, it is less efficient than a more straightforward approach, and a sketch of the ideas for this approach is given below (Zeilberger, 2019; for details please refer to the paper):

Let's say we are partition the integer $n$, whose largest part is $m$. For any given set of patterns $A$, the computer automatically sets-up a scheme, introducing more general quantities, parameterized, in addition to the set of **global** conditions $A$, by a set of *local* conditions that should be avoided at the very beginning. Then, for each such set of beginning restrictions, $A'$, depending on $m'$ (the second largest part of the partition), either we are back to only the global conditions, $A$, i.e. the new $A'$ is the empty set, or if $m - m'$ happens to be one of the starting entries of $A$ or $A'$, the chopped partition, of $n - m$, in addition to obeying the global restrictions of $A$, must obey a brand-new kind of restrictions $A''$. So each 'state' $(m, m', A')$ gives rise to a state $(m', m'', A'')$ for some (possibly empty) set $A''$. Finding these "children" state is automatically done by the computer, setting up a quadratic-time scheme. At the end of the day, we are only interested in the case where $A' = \emptyset$, but we are forced to consider these auxiliary quantitities. Since there are only finitely many of them, and there are still only two arguments (namely $n$ and $m$, where $1 \leq m \leq n$), the algorithm remains quadratic time and quadratic memory, which is much faster than the brute force approach.

2.2. **Searching for partition identities.** [Joint work with Matthew Russell]

Even though the previous project will enable us to enumerate partitions avoiding an *arbitrary* (finite) set of patterns, many partition theoretic sum sides of partition identities requires more specific restrictions. Recall Schur's celebrated 1926 theorem: the number of partitions of $m$ into parts with minimal difference 3 and with no consecutive multiples of 3 is equal to the number of partitions of $m$ into distinct parts $\equiv 1, 2 \pmod 3$. We are interested in the "sum-side" of this identity, that is, the number of partitions of $m$ into parts with minimal difference 3 and with no consecutive multiples of 3.

Another example involves more complicated restrictions featured in the intruiging Kanade-Russell conjectures, for example, the sum-side in 3.1.1. of [18]:

(a) No consecutive parts allowed.
(b) Odd parts do not repeat.
(c) Even parts appear at most twice.
(d) If a part $2j$ appears twice then $2j \pm 3, 2j \pm 2$ are forbidden to appear at all.
(e) 2+2 is not allowed as a sub-partition.

In order to generalize the approach described at the end of 2.1., which is very efficient, we introduce some new notions and formulate these sum-sides in a different way, which turns out to be very flexible and also easy to feed to a computer. Here are the three major parameters that we use:

$A$: the set of patterns to avoid "globally" (same as before)

$Mod$: the list of patterns to avoid according to congruence conditions of the largest part of a sub-partition. If $Mod = [A_0, A_1, ..., A_{k-1}]$ (each $A_i$ is a set of patterns), then we are considering modulo $k$ and avoiding the sub-partitions $\{a_0, kj\}$, $\{a_1, kj+1\}$,..., $\{a_k, kj+(k-1)\}$ where $a_i$ can be any pattern in $A_i$ and $j$ can be any positive integer. (Note we are using the notation $\{v, k\}$ to denote the partition that starts with $k$ and has underlying pattern $v$).

$IC$: the list of sub-partitions to avoid (we call this "initial conditions")

In light of these notations we have–Schur: parts with minimal difference 3 translates to $A = \{[0], [1], [2]\}$. No consecutive multiples of 3 translates to $Mod = [\{[3]\}, \{\}, \{\}]$.

3.1.1. in [18]: No consecutive parts allowed translates to the global condition: $A = \{[1]\}$. Odd parts do not repeat, even parts appear at most twice, along with part (d) translate to $Mod = [\{[0,0],[0,3],[0,2],[2,0]\},\{[0],[3,0]\}]$. Part (e), that is, 2+2 is not allowed as a sub-partition translates to IC=[[2,2]].

Using these new notations I came up with an algorithm (for details, please see `http://sites.math.rutgers.edu/~my237/Pos_Ext`) that enables very efficient computation. Let $GP(m,n,A,Mod,B,IC)$ be the number of partitions of $n$, with largest part $m$, and the restrictions $A$, $Mod$, $IC$ described above, and $B$ and $B''$ are beginning restrictions. At the core of the algorithm is the recurrence relation:

$$GP(m,n,A,Mod,B,IC) = \sum_{\substack{1\le m'\le m \\ [m-m']\notin A\cup B'\cup Mod[i+1]}} GP(m',n-m,A,Mod,B'',IC) \quad .$$

With this algorithm and the help of Frank Garvan's q-series Maple package (in particular, the prodmake procedure that converts a $q$-series $f$ into an infinite product that agrees with $f$ to a some power of $q$), we are currently conducting the search over various parameter restrictions, with the help of Amarel cluster computing `https://oarc.rutgers.edu/amarel/`. We have already found many seemingly new Rogers-Ramanujan type identities, and has generalized one of them to an infinite family.

**Future work:**

1. Put more variations on the parameters to enlarge our search space.

2. Currently our approach only deal with conditions on contiguous sub-partitions. It will be nice to develop a general frame work/an efficient way to search for identities that avoid sub-partitions that are not necessarily contiguous.

3. During my talk at 2019 AMS Fall Southeastern Sectional Meeting, Drew Sills (see `http://home.dimacs.rutgers.edu/~asills/`) and Ali Uncu (see `https://risc.jku.at/m/ali-uncu/`) offered some great suggestions to make our algorithm more flexible, including allowing the initial conditions to contradict the global conditions and allowing Nandi-type conditions. We are currently working on implementing these.

2.3. **Relaxed Partitions.**

Project link: `http://sites.math.rutgers.edu/~my237/RP`

In this project, we took a road less traveled and studied an object which we call "relaxed partitions", or more specifically, $r$-partitions with $r$ to be specified. Unlike the traditional partitions where we require $\lambda_i - \lambda_{i+1} \geq 0$, for $r$-partitions we require $\lambda_i - \lambda_{i+1} \geq r$ where $r$ can be negative. For example, $(2, 3, 1, 1)$ is a $(-1)$-partition of 7.

Just as with traditional partitions, there are many questions one could ask about $r$-partitions. Perhaps one of the first questions to ask is: "for a fixed $r$, how many $r$-partitions of the integer $n$ do we have?" Using an easy recurrence relation, I programmed $\mathbf{NPr(n, r)}$, which answers this question for specific $n$ and $r$. But can we find a generating function for a given $r$? The answer turned out to be yes! And there is a nice closed form for it. By typing in a sequence of entries produced by $\mathbf{NPr(n, -1)}$ into the **OEIS**, we found that it is the reciprocal of

$$\frac{\sum_{k=0}^{\infty}(-1)^k q^{k^2}}{\prod_{i=1}^{k}(1-q^i)} \quad .$$

Recognizing this is the generating function for the "weighted" (with weight $(-1)^k$) number of partitions (traditional) of integer $n$ into parts with difference at least 2, Doron Zeilberger (2018) provided a short and elegant bijective proof for it. I generalized this to general $r$:

**Theorem 2.2.** *(Yang, 2018) The generating function for the number of $r$-partitions of $n$ is the reciprocal of*

$$\frac{\sum_{k=0}^{\infty}(-1)^k q^{k(2+(1-r)(k-1))/2}}{\prod_{i=1}^{k}(1-q^i)} \quad ,$$

*which is the "weighted" number of partitions of integer $n$ into parts with difference at least* $(-r+1)$.

Since this question was answered, we turned our focus to restricted $r$-partitions such that the first part and the number of the parts are fixed. Let $a_r(M, N, n)$ be the number of $r$ partitions with the first part equal to $M$ and exactly $N$ parts. To go with the notation in our Maple package **rPar**, let $Ftr(M, N, r, q)$ be the generating function for $a_r(M, N, n)$. Using a simple recurrence relation $Ftr(M, N, r, q)$ satisfies, I was able to program it in Maple and happily used Maple to conjecture (**and prove!**) the closed form for the case $q = 1$ (i.e., the total number of $r$ partitions with the first part equal to $M$ and exactly $N$ parts). It is:

**Theorem 2.3.** *(Yang, 2018)*

$$Ftr(M, N, r, 1) = \frac{(M-r)(M+(1-r)N-2)!}{(N-1)!(M-rN)!} \quad .$$

$$= \binom{M+(1-r)N-2}{N-1} + r\binom{M+(1-r)N-2}{N-2}$$

Although I was not yet able to find a closed form formula for the generating function $Ftr(M, N, r, q)$, I found out (using Maple) some initial terms (according to $N$) of it. Below are the first five terms produced by our Maple program, which can be proven easily using a recurrence relation. Note even though they look like rational functions, they are in fact polynomials.

$\mathbf{qGuessPol([seq(F(M1, 1, -1, q), M1 = 1..20)], M, q, 1)}$; yields

$$q^M$$

$\mathbf{qGuessPol([seq(F(M1, 2, -1, q), M1 = 1..20)], M, q, 1)}$; yields

$$\frac{q^{M+1}(q^{M+1}-1)}{(q-1)}$$

$\mathbf{qGuessPol([seq(F(M1, 3, -1, q), M1 = 1..20)], M, q, 1)}$; yields

$$\frac{q^{M+2}(q^{M+3}+q^2-q-1)(q^{M+1}-1)}{(q-1)^2(q+1)}$$

$\mathbf{qGuessPol([seq(F(M1, 4, -1, q), M1 = 1..20)], M, q, 1)}$; yields

$$\frac{q^{M+3}(q^{2M+8}+q^{M+7}-q^{M+5}-q^{M+4}-q^{M+3}+q^6-2q^4-q^3+2q+1)(q^{M+1}-1)}{(q-1)^3(q+1)(q^2+q+1)}$$

$\mathbf{qGuessPol([seq(F(M1, 5, -1, q), M1 = 1..20)], M, q, 1)}$; yields

$$\frac{q^{M+4}(q^{M+5}+q^4-q-1)(q^{2M+10}-q^{M+4}-q^{M+3}+q^8-q^5-2q^4+2q+1)(q^{M+1}-1)}{(q-1)^4(q+1)^2(q^2+1)(q^2+q+1))}$$

Drew Sills (see `http://home.dimacs.rutgers.edu/~asills/`), during his short visit to Rutgers and our brief meeting, observed that $F(M, N, -1, q)$ has denominator $(q; q)_N$ and a numerator of degree $N(M+N-1)$. He pointed out that it is plausible the numerator is a (possibly alternating) sum of polynomials that are a power of q times a Gaussian polynomial of the form $G(M, N) := GP(2N+M-1, N)$ where

$$GP(m, r) := \frac{(q^{m-r+1}, q)_r}{(q, q)_r}.$$

So far we have not made much progress in this, but this led me to discover an interesting pattern:

$$\begin{cases} G(2,1) = q^2+q+1 \\ F(2,2,-1,q) = q^5+q^4+q^3 \end{cases}$$

$$\begin{cases} G(2,2) = \underline{q^6+q^5+2q^4+2q^3+2q^2+q}+1 \\ F(2,3,-1,q) = \underline{q^9+q^8+2q^7+2q^6+2q^5+q^4} \end{cases}$$

$$\begin{cases} G(2,3) = \underline{q^{12}+q^{11}+2q^{10}+3q^9+4q^8+4q^7+5q^6+4q^5}+4q^4 \\ \qquad\qquad +3q^3+2q^2+q+1 \\ F(2,4,-1,q) = \underline{q^{14}+q^{13}+2q^{12}+3q^{11}+4q^{10}+4q^9+5q^8} \\ \qquad\qquad \underline{+4q^7}+3q^6+q^5 \end{cases}$$

Notice for each pair, the underlined parts have the same coefficients. This pattern continues indefinitely for bigger $M$ and $N$. This is an ongoing project to explore why this intriguing pattern exists and if it can point us to discover a closed form for $F(M,N,-1,q)$, or even $F(M,N,r,q)$.

**Connection to other combinatorial objects:**
With the help of **OEIS**, I found a direct connection between $F(M,N,-1,1)$ and Catalan's triangle. Since $F(M,N,-1,1) = \frac{(M-r)(M+(1-r)N-2)!}{(N-1)!(M-rN)!}$, it is clear that $F(M,N,-1,1) = C(M+N-1,N-1)$. And there turned out to be a nice a geometric interpretation.

There also seems to be a bit of connection between the standard Young Tableau and $F(M,N,-1,1)$. For example, typing the sequence $[seq(F(5,N,-1,1),N=1..20]$ into **OEIS**, we will find it can also represent the number of standard Young Tableau of shape $(N+3,N-2)$. (A003517) If we change the value of M, then we can find correspondence with other standard Young Tableau. I conjecture that $F(M,N,-1,1)$ is equal to the number of standard Young Tableau of shape $(N+\lceil M/2\rceil, N-\lfloor M/2\rfloor)$. For general $r$, we haven't found nice connections yet.

**Future work:** As mentioned earlier, it would be wonderful if we could discover why the intriguing pattern between the $F$ and $G$ exists. This may enable us to find a closed form for $F(M,N,-1,q)$, and perhaps give us insights for $F(M,N,r,q)$. During my talk at the 2019 AMS Spring Southeastern Sectional Meeting, Tim Huber (see `https://faculty.utrgv.edu/timothy.huber/`) mentioned that these generating functions' approximation of the Gaussian polynomials may indicate they count permutations by some statistic and may arise from a quotient of generating functions of a special form. This is still under exploration. It would also be nice if we can find connection between $F(M,N,r,1)$ and other combinatorial objects.

I am also interested in exploring plane partitions and "fractional counting", some initial Maple "playing-around" can be found on `http://sites.math.rutgers.edu/~my237/RP`.

### 2.4. Enumeration of words that contain pattern 123 exactly once.

Project link: `http://sites.math.rutgers.edu/~my237/One123`

Recall that a word $w = w_1...w_k$ is an ordered list of letters on some alphabet. To say a word contains a pattern (a certain permutation of $\{1,...,m\}$) $\sigma$ is to say there exist $1 \le i_1 < i_2 < ... < i_m \le k$ such that the subword $w_{i_1}...w_{i_m}$ is *order isomorphic* to $\sigma$ (for example, 246 is order isomorphic to 123). A word avoids the pattern $\sigma$ if it does not contain it.

Enumeration problems related to words avoiding patterns as well as permutations that contain the pattern 123 exactly once have been studied in great detail. However, the problem of enumerating words that contain the pattern 123 exactly once is new and was the focus of this project.

**Theorem 2.4.** *(Noonan, 1996* [20]*) The number of permutations with exaclty one 321 pattern is equal to* $\frac{3}{n}\binom{2n}{n+3}$.

In a paper published in 2011 [5], Burstein gave an elegant combinatorial proof of this theorem. Soon after, Zeilberger [Z1] was able to shorten Burstein's proof by using a bijection between a permutation with exactly one pattern 321, denoted as $\pi_1 c\pi_2 b\pi_3 a\pi_4$ $(a < b < c)$, with the pair $(\pi_1 b\pi_2 a, c\pi_3 b\pi_4)$ where $\pi_1 b\pi_2 a$ is a 321-avoiding permutation of $\{1, ..., b\}$ and $c\pi_3 b\pi_4$ is a 321-avoiding permutation of $\{b, ..., n\}$.

I was able to extend Zeilberger's bijective proof idea to words and came up with the following theorem:

**Theorem 2.5.** *(Yang, 2017) Let* $A(l_1, ..., l_n)$ *be the number of* 123*-avoiding words associated with the list* $[l_1, ..., l_n]$*. And let* $B(l_1, ..., l_n)$ *be the number of words associated with list* $[l_1, ..., l_n]$ *that contain the pattern 123 exactly once. Then we have*

$$B(l_1, ..., l_n) = \sum_{b=2}^{n-1}\sum_{j=0}^{l_b-1}(A(l_1, ..., l_{b-1}, j+1) - A(l_1, ..., l_{b-1}, j)) \cdot (A(l_b-j, l_{b+1}, ..., l_n) - A(l_b-j-1, l_{b+1}, ..., l_n)).$$

Building on the ideas from this result, I then extended Shar and Zeilberger's work [26] on generating functions enumerating 123-avoiding words (with $r$ occurrences of each letter) to words (with $r$ occurrences of each letter) having exactly one pattern 123 :

**Theorem 2.6.** *(Yang, 2017)*

$$h_r(x) = \frac{1}{x}\sum_{i=1}^{r}(g_r^{(0,\ i\ mod\ r)} - xg_r^{(0,\ i-1)} - \delta_{(i\ mod\ r,\ 0)})(g_r^{(0,\ (r+1-i)\ mod\ r)} - xg_r^{(0,\ r-i)} - \delta_{((r+1-i)\ mod\ r,\ 0)}).$$

(the definitions for $g_r$ and $h_r$ are somewhat complicated, details please see paper.)

Also in [26], for every positive integer $r$, Shar and Zeilberger found an algorithm for finding the defining algebraic equation for the ordinary generating function enumerating 123-avoiding words of length $rn$ where each of the $n$ letters of $\{1, 2, ..., n\}$ occurs exactly $r$ times.

Theorem 2.6 allowed us to produce an analogue of that, that is, a defining algebraic equation for the ordinary generating function enumerating words of length $rn$ where each of the $n$ letters of $\{1, 2, ..., n\}$ occurs exactly $r$ times, now with exactly one pattern 123. We used the same (as in Shar and Zeilberger's paper [26]) memory-intensive, and exponential time, Buchberger's algorithm for finding Gröbner bases, and our computer (running Maple) found the defining algebraic equation for $r = 2$:

$$x^4 (x+4)^2 F^4 + 2x^3 (x+4)(11x+23) F^3 - 4x \left(3x^4 - 10x^3 - 97x^2 - 146x + 1\right) F^2$$
$$+ \left(-168x^4 - 840x^3 - 744x^2 + 336x - 24\right) F + 144x^3 (x+2) = 0.$$

This took about a second. The minimal algebraic equation for $r = 3$ has 12 as the highest power for $F$ and the computation took about 20 seconds. Interested readers can find it on the website: `http://sites.math.rutgers.edu/~my237/One123`. Unfortunately, the case when $r = 4$ already took too long to compute (more than a month).

Finally, let $a_r(n)$ be the number of words of length $rn$ where each of the $n$ letters of $\{1, 2, ..., n\}$ occurs exactly $r$ times, with exactly one pattern 123. We used the Maple package **SCHUTZEN-BERGER** to derive recurrence relations for our sequences. Having obtained the defining algebraic equations of the generating functions for $a_r(n)$ in the cases $r = 2$ and $r = 3$, Manuel Kauers kindly helped us in finding the asymptotics for our sequences $a_2(n)$ and $a_3(n)$ (thanks to Kauers, the constants in front are fully rigorous and were computed via a step by step procedure; for details, please refer to [15]):

$$a_2(n) = \frac{3(13 - \sqrt{21})}{49} \cdot \frac{1}{\sqrt{\pi}} \cdot 12^n \cdot n^{-3/2} \cdot (1 + O(n^{-1})),$$

$$a_3(n) = \frac{-7 + 6\sqrt{7}}{56} \cdot \frac{1}{\sqrt{\pi}} \cdot 32^n \cdot n^{-3/2} \cdot (1 + O(n^{-1})).$$

2.5. **Increasing Consecutive Patterns in Words.** [This is joint work with Doron Zeilberger.] Project link: `http://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/icpw.html`

The enumeration of words/permutations avoiding a classical pattern (as in Section 2.1.), or a set of patterns, is in general quite difficult, and it is widely believed to be intractable for most patterns–hence it would be nice to have other notions for which the enumeration is more feasible. Such an analog was given, in 2003, by Elizalde and Noy, in a seminal paper [9], that introduced the study of the enumeration of permutations avoiding *consecutive* patterns. A permutation

$\pi = \pi_1 \cdots \pi_n$ avoids a consecutive pattern $\sigma = \sigma_1 \cdots \sigma_k$ if none of the $n-k+1$ length-$k$ consecutive subwords, $\pi_i \pi_{i+1} \cdots \pi_{i+k-1}$ of $\pi$, reduces to $\sigma$.

Algorithmic approaches to the enumeration of *permutations* avoiding sets of consecutive patterns were given by Nakamura, Baxter, and Zeilberger [21, 2]. Our project may be viewed as an extension, from permutations to words, of Nakamura's paper, who was also inspired by the Goulden-Jackson cluster method, but in a sense, is more straightforward, and closer in spirit to the original Goulden-Jackson cluster method ([13], that is beautifully exposited (and extended!) in [22]).

In this project we consider consecutive patterns of the form $1 \cdots r$, i.e. *increasing consecutive patterns*, and show how to count words in $1^{m_1} \cdots n^{m_n}$ avoiding the pattern $1 \cdots r$ (Theorem 2.7, that is originally due to Ira Gessel [11], we discovered this theorem independently using experimentation in Maple).

**Theorem 2.7.** *(Gessel [11]) For $n \geq 1, r \geq 2$, the generating function for words in $\{1, 2, \ldots, n\}$ avoiding the consecutive pattern $12 \cdots r$, let us call it $F_r(x_1, \ldots, x_n)$ is*

$$F_r(x_1, \ldots, x_n) = \frac{1}{1 - e_1 + e_r - e_{r+1} + e_{2r} - e_{2r+1} + e_{3r} - e_{3r+1} + \cdots} .$$

This immediately implies:

**Fundamental Recurrence** (Zeilberger, 2018): Let $f_r(\mathbf{m})$ be the number of words in the alphabet $\{1, \ldots, n\}$ with $m_1$ 1's, $m_2$ 2's, $\ldots$, $m_n$ $n$'s (where we abbreviate $\mathbf{m} = (m_1, \ldots, m_n)$) that avoid the consecutive pattern $1 \cdots r$. Also let $V_i$ be the set of $0-1$ vectors of length $n$ with $i$ ones, then

$$\begin{aligned} f_r(\mathbf{m}) = & \sum_{\mathbf{v} \in V_1} f_r(\mathbf{m} - \mathbf{v}) - \sum_{\mathbf{v} \in V_r} f_r(\mathbf{m} - \mathbf{v}) \\ & + \sum_{\mathbf{v} \in V_{r+1}} f_r(\mathbf{m} - \mathbf{v}) - \sum_{\mathbf{v} \in V_{2r}} f_r(\mathbf{m} - \mathbf{v}) \\ & + \sum_{\mathbf{v} \in V_{2r+1}} f_r(\mathbf{m} - \mathbf{v}) - \sum_{\mathbf{v} \in V_{3r}} f_r(\mathbf{m} - \mathbf{v}) \\ & + \sum_{\mathbf{v} \in V_{3r+1}} f_r(\mathbf{m} - \mathbf{v}) - \sum_{\mathbf{v} \in V_{4r}} f_r(\mathbf{m} - \mathbf{v}) + \cdots \quad . \end{aligned}$$

Using this recurrence, we devised an algorithm to **efficiently** count words in $1^s \cdots n^s$ avoiding the pattern $1 \cdots r$. Our implied algorithms are $O(n^{s+1})$ and hence yield many more terms, and, of course, new sequences for **OEIS**. I also provided a new proof of Theorem 2.7 by tweaking the Goulden-Jackson cluster method. Using this proof, along with a little more effort, I generalized Theorem 2.7 to counting words with a specified number of the pattern $12 \cdots r$ (Theorem 2.9), instead of just *avoiding*, that is, having *zero* occurrence of the pattern of interest.

**Definition 2.8.** *For any integer $k \geq 1$ and $r \geq 2$, $P_k^{(r)}(t)$ is defined as follows.*

If $k < r$, then it is 0. If $k = r$ then it is $t-1$, and if $k > r$ then

$$P_k^{(r)}(t) \,=\, (t-1)\sum_{i=1}^{r-1} P_{k-i}^{(r)}(t) \quad .$$

**Theorem 2.9** (Yang, 2018)**.** *For $k \geq 1, r \geq 2$, the generating function of $g_r(\mathbf{m}; t)$, let us call it $G_r(x_1, \ldots, x_n; t)$, is*

$$G_r(x_1, \ldots, x_n; t) = \frac{1}{1 - e_1 - \sum_{k=r}^{n} P_k^{(r)}(t) e_k} \quad .$$

**Maple Packages**: This project is accompanied by three Maple packages available from the webpage:

`http://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/icpw.html` .

These are

• `ICPW.txt`: For fast enumeration of sequences enumerating words avoiding increasing consecutive patterns.

• `ICPWt.txt`: For fast computation of sequences of weight-enumerators for words according to the number of increasing consecutive patterns ($t = 0$ reduces to the former case).

• `GJpats.txt`: For conjecturing generating functions (that still have to be proved by humans).

This page also has links to numerous input and output files. The input files can be modified to generate more data, if desired.

## References


[1] George Andrews, *The Theory of Partitions*, Addison-Wesley, Reading, MA, 1976. Reissued: Cambridge University Press, 1998.

[2] A. Baxter, B. Nakamura, and D. Zeilberger, Automatic generation of theorems and proofs on enumerating consecutive-Wilf classes, in *Advances in Combinatorics: Waterloo Workshop in Computer Algebra, W80, May 26-29, 2011* [Volume in honor of Herbert S. Wilf], edited by Ilias Kotsireas and Eugene Zima, Springer, 2013, pp. 121–138.

[3] Bóna, Permutations with one or two 132-subsequences, *Discrete Math.* **181** (1998), 267-274.
`http://www.sciencedirect.com/science/article/pii/S0012365X97000629?via%3Dihub`

[4] A. Burstein, *Enumeration of words with forbidden patterns*, Ph.D. thesis, University of Pennsylvania, 1998.

[5] A. Burstein, A short proof for the number of permutations containing pattern 321 exactly once, *The Electronic J. of Combinatorics* **18(2)** (2011), P21.
`http://www.combinatorics.org/ojs/index.php/eljc/article/view/v18i2p21/pdf`

[6] A. Burstein and T. Mansour, Words restricted by patterns with at most 2 distinct letters, *Electron. J. Combin.* **9** (2002), #R3.

[7] G. Chapuy, The asymptotic number of $12..d$-Avoiding Words with $r$ occurrences of each letter $1, 2, ..., n$ (2014)
`https://arxiv.org/abs/1412.6070`

[8] F. N. Davis and D. E. Barton, *Combinatorial Chance*, Hafner, New York, 1962.

[9] S. Elizalde and M. Noy, Consecutive patterns in permutations, *Adv. Appl. Math.* **30** (2003), 110–125 .

[10] Shalosh B. Ekhad and Doron Zeilberger, The "Monkey Typing Shakespeare" Problem for Compositions, *The Personal Journal of Shalosh B. Ekhad and Doron Zeilberger*, Jan. 12, 2019.
`http://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/kof.html`.

[11] I. M. Gessel, *Generating functions and enumeration of sequences*, PhD thesis, Massachusetts Institute of Technology, 1977.

[12] I. M. Gessel, Symmetric functions and P-recursiveness, *J. Combin. Theory Ser. A* **53** (1990), 257–285.

[13] Ian Goulden and David Jackson, An inversion theorem for cluster decomposition of sequences with distinguished subsequences, *J. London Math. Soc.(2)* **20** (1979), 567-576.

[14] M. Kauers, A Mathematica Package for Computing Asymptotic Expansions of Solutions of P-Finite Recurrence Equations, Technical Report RISC 11-04, Johannes Kepler University, April 2011.

[15] M. Kauers and P. Paule, *The Concrete Tetrahedron: Symbolic Sums, Recurrence Equations, Generating Functions, Asymptotic Estimates*, Springer (2011), Section 6.5.

[16] S. Kanade and M. C. Russell, IdentityFinder and some new identities of Rogers-Ramanujan type, *Exp. Math.* **24** (2015), 419-423. `https://arxiv.org/abs/1411.5346`.

[17] S. Kanade, D. Nandi and M. C. Russell, A variant of IdentityFinder and some new identities of Rogers-Ramanujan- MacMahontype. `https://arxiv.org/abs/1902.00790`.

[18] S. Kanade and M. C. Russell, Staircases to analytic sum-sides for many new integer partition identities of Rogers-Ramanujan type. *Electron. J. Combin.*, **26(1)**:Paper 1.6, 2019. `https://arxiv.org/pdf/1803.02515`.

[19] A. Mendes and J.B. Remmel, Permutations and words counted by consecutive patterns, *Adv. Appl. Math*, **37** (2006), 443–480.

[20] J. Noonan, The number of permutations containing exactly one increasing subsequence of length three, *Discrete Math.* **152** (1996), 307-313.
`http://www.sciencedirect.com/science/article/pii/0012365X9500247T`

[21] B. Nakamura, Computational approaches to consecutive pattern avoidance in permutations, *Pure Math. Appl.* (PU.M.A.) **22** (2011), 253–268.

[22] J. Noonan and D. Zeilberger, The Goulden-Jackson cluster method: extensions, applications, and implementations, *J. Difference Eq. Appl.* **5** (1999), 355–377.
`http://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/gj.html` .

[23] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, `http://oeis.org`.

[24] L. Pudwell, *Enumeration schemes for pattern-avoiding words and permutations*, Ph.D. thesis, Rutgers University, May 2008,
`http://faculty.valpo.edu/lpudwell/papers/pudwell_thesis.pdf`

[25] Andrew V. Sills, *An Invitation to the Rogers-Ramanujan Identities*, CRC Press, Boca Raton, 2018.

[26] N. Shar and D. Zeilberger, The (Ordinary) Generating Functions Enumerating 123-Avoiding Words with $r$ occurences of each of $1, 2, ..., n$ are Always Algebraic, *Annals of Combinatorics* **20** (2016), 387-396.
`http://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/words123.html`

[27] X. Wen, The symbolic Goulden-Jackson cluster method, *J. Difference Eq. Appl.* **11** (2005), 173–179.

[28] D. Zeilberger, Alexander Burstein's Lovely Combinatorial Proof of John Noonan's Beautiful Theorem that the number of $n$-permutations that contain the Pattern 321 Exactly Once Equals $(3/n)(2n!)/((n-3)!(n+3)!)$, *Personal Journal of Shalosh B. Ekhad and Doron Zeilberger*, Oct.18, 2011.

http://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/burstein.html

[29] D. Zeilberger, A Snappy Proof That 123-Avoiding Words are Equinumerous With 132-Avoiding Words, *Personal Journal of Shalosh B. Ekhad and Doron Zeilberger*, April 11, 2005. http://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/a123.html