INVOLUTIONS AND THEIR PROGENIES

TEWODROS AMDEBERHAN AND VICTOR H. MOLL

Abstract. Any permutation has a disjoint cycle decomposition and concept generates an equivalence class on the symmetry group called the cycle-type. The main focus of this work is on permutations of restricted cycle-types, with particular emphasis on the special class of involutions and their partial sums. The paper provides generating functions, determinantal expressions, asymptotic estimates as well as arithmetic and combinatorial properties.

1. Introduction

For $n \in \mathbb{N}$, the group of permutations in n symbols $\{a_1, a_2, \cdots, a_n\}$ is called the symmetric group, denoted by $\mathfrak{S}_n$. A *cycle* $\rho \in \mathfrak{S}_n$ is a permutation of the form, in a one-line notation, $\rho = (a_{i_1}\, a_{i_2}\, \cdots a_{i_r})$. The notation indicates that all the entries of the cycle are distinct and $\rho(a_{i_j}) = a_{i_{j+1}}$ for $1 \leq j \leq r-1$ and $\rho(a_{i_r}) = a_{i_1}$. The cycle ρ is said to have *length* r, written as $r = L(\rho)$. Every permutation $\pi \in \mathfrak{S}_n$ can be written as a product of cycles $\pi = \rho_1\rho_2\cdots\rho_m$. This decomposition is not unique, but if the cycles are assumed to be disjoint and the lengths are taken in weakly decreasing order, then $\{L(\rho_1), L(\rho_2), \cdots, L(\rho_m)\}$ is uniquely determined by π, called the *cycle type* of π.

The following notation is used: for $1 \leq \ell,\, t \leq n$,

$$C_{n,\ell} = \{\pi \in \mathfrak{S}_n \Big| \text{ with every cycle in } \pi \text{ of length at most } \ell\}, \tag{1.1}$$

$$\alpha_t(\pi) = \text{ number of cycles in } \pi \in \mathfrak{S}_n \text{ of length } t. \tag{1.2}$$

and the cardinality of $C_{n,\ell}$ is denoted by $d_{n,\ell} = \#C_{n,\ell}$.

Definition 1.1. A permutation π in $\mathfrak{S}_n$ is called an *involution* if $\pi^2(j) = j$, for $1 \leq j \leq n$. The set of involutions in $\mathfrak{S}_n$ is denoted by $\mathrm{Inv}(n)$. The cardinality of this set, denoted by $I_1(n)$, is called the *involution number*.

The factorization of π as a product of disjoint cycles shows that any cycle in the factorization of an involution has length 1 or 2. This implies $\mathrm{Inv}(n) = C_{n,2}$ and thus $I_1(n) = d_{n,2}$. It follows that if $\pi \in \mathrm{Inv}(n)$ is an involution, then $\alpha_1(\pi) + 2\alpha_2(\pi) = n$.

Example 1.2. Every permutation of 2 symbols (a transposition) is an involution and for $n = 3$ there are 4 involutions

$$\pi_1 = (1)(2)(3),\ \pi_2 = (12),\ \pi_3 = (13), \pi_4 = (23). \tag{1.3}$$

The cycles (123) and (132) are the only elements of S_3 that are not involutions. Therefore $I_1(2) = 2$ and $I_1(3) = 4$.

Date: February 8, 2015.

1991 *Mathematics Subject Classification.* Primary 05A15, 11B75.

Key words and phrases. involutions, valuations, asymptotics.

Elementary properties of the numbers $I_1(n)$ are described in Section 2. These include a second order recurrence, an exponential generating function as well as an explicit finite sum. These are generalized to the *involution polynomials* $I_1(n;t)$ in Section 3 which are intimately linked to the (probabilistic) Hermite polynomials defined by

$$H_n(t) = n! \sum_{j=0}^{\lfloor n/2 \rfloor} \frac{(-1)^j}{j!(n-2j)!} \frac{t^{n-2j}}{2^j} \tag{1.4}$$

with generating function

$$\sum_{n=0}^{\infty} H_n(t) \frac{x^n}{n!} = \exp\left(xt - \tfrac{1}{2}x^2\right). \tag{1.5}$$

The involution polynomials have a combinatorial interpretation as the generating function for fixed points of permutation in $\mathfrak{S}_n$. Arithmetic properties of $I_1(n)$ are presented in Section 4. Particular emphasis is given to the 2-adic valuation of $I_1(n)$. Recall that, for $x \in \mathbb{N}$ and p prime, the *p-adic valuation* of x, denoted by $\nu_p(x)$, is the highest power of p that divides x. An odd prime p is called *efficient* if p does not divide $I_1(j)$ for $0 \leq j \leq p-1$. Otherwise it is called *inefficient*. The prime $p = 3$ is efficient and $p = 5$ is inefficient since $I_1(4) = 10$. A periodicity argument is used to show that $\nu_p(I_1(n) = 0$; i.e., p is efficient. Morever, for a prime p, it is shown that either p divides $I_1(n)$ infinitely often or never. In the case of an inefficient prime, it is conjecture that the p-adic valuation of the sequence $I_1(n)$ can be given in terms of a tree $\mathbb{T}_p$. This phenomena is illustrated for the prime $p = 5$. It is an open question to characterize efficient (or inefficient) primes. The partial sums of $I_1(n)$, denoted by a_n, are discussed in Section 5. Their arithmetic properties are presented in Section 6. For instance, an explicit expression for their 2-adic valuation is given there. The valuations for odd primes are also conjectured to have a tree structure. This is illustrated in the case $p = 5$. Section 7 considers the statistics of the sequence $d_{n,\ell}$ in (1.1). This is a generalization of $I_1(n) = C_{n,2}$. Finally, the asymptotic behavior of $d_{n,\ell} = |C_{n,\ell}|$ is given in Section 8.

2. Basic properties of the involution numbers

This section discusses fundamental properties of $I_1(n)$. Some of them are well-known but proofs are included here for the convenience of the reader.

Theorem 2.1. *The sequence $I_1(n)$ satisfies the recurrence*

$$I_1(n) = I_1(n-1) + (n-1)I_1(n-2), \text{ for } n \geq 2, \tag{2.1}$$

with initial conditions $I_1(0) = I_1(1) = 1$.

Proof. There are $I_1(n-1)$ involutions that fix n. The number of involutions that contain a cycle $(j\, n)$, with $1 \leq j \leq n-1$ is $n-1$ times the number of involutions containing the cycle $(n-1,\, n)$. This is $(n-1)I_1(n-2)$. □

The recurrence above generates the values

n	0	1	2	3	4	5	6	7	8	9	10
$I_1(n)$	1	1	2	4	10	26	76	232	764	2620	9496

This is sequence $A000085$ in OEIS.

The recurrence (2.1) now enables to write a generating function for $\{I_1(n)\}$.

Theorem 2.2. *The exponential generating function for $I_1(n)$ is*

$$\sum_{n=0}^{\infty} \frac{I_1(n)}{n!} x^n = \exp(x + \tfrac{1}{2}x^2). \tag{2.2}$$

Proof. On the basis of (2.1) verify that both sides of (2.2) satisfy $f'(x) = (1+x)f(x)$ and the value $f(0) = 1$. □

Cauchy's product formula on e^x and $e^{x^2/2}$ allows to express $I_1(n)$ as a finite sum.

Corollary 2.3. *The involution numbers $I_1(n)$ are given by*

$$I_1(n) = \sum_{j=0}^{\lfloor n/2 \rfloor} \binom{n}{2j} \binom{2j}{j} \frac{j!}{2^j}. \tag{2.3}$$

The numbers $\binom{2j}{j}\frac{j!}{2^j}$ appearing in (2.3) are now shown to be of the same parity.

Corollary 2.4. *For $j \in \mathbb{N}$, the numbers $(2j)!/(j!2^j)$ are odd integers.*

Proof. The identity

$$\frac{(2j)!}{j!2^j} = \frac{(2j)(2j-1)\cdots(j+1)}{2^j} \tag{2.4}$$

shows that the denominator is a power of 2. To compute this power, use Legendre's formula

$$\nu_2(n!) = n - s_2(n), \tag{2.5}$$

where $s_2(n)$ is the sum of the digits of n in its binary expansion. Therefore,

$$\nu_2\left(\frac{(2j)!}{j!2^j}\right) = (2j - s_2(2j)) - (j - s_2(j)) - j = 0, \tag{2.6}$$

in view of $s_2(2j) = s_2(j)$. □

A second recurrence for the involution numbers is presented next. This result appears to be new to the authors.

Theorem 2.5. *For $n, m \in \mathbb{N}$, the involution numbers satisfy*

$$I_1(n+m) = \sum_{k \geq 0} k! \binom{n}{k} \binom{m}{k} I_1(n-k) I_1(m-k). \tag{2.7}$$

Proof. Split up the set $[n+m]$ into two disjoint subsets A and B, of n and m letters, respectively. Count the involutions in $\mathrm{Inv}(n+m)$ acoording to the number k of cross-permutations that make up a cycle (ab), with $a \in A$ and $b \in B$. The letters a and b can be chosen in $\binom{n}{k}\binom{m}{k}$ ways and $k!$ ways to place the k cycles (ab). The remaining elements in A (respectively B) allow $I_1(n-k)$ (respectively $I_1(m-k)$) involutions. To complete the argument, summing $k!\binom{n}{k}\binom{m}{k}I_1(n-k)I_1(m-k)$ over k. □

As a direct consequence of (in fact, equivalent to) Theorem 2.5 the following analytic statement is recorded. This result bypasses the need for an otherwise messy chain rule for derivatives.

Corollary 2.6. *Higher order derivatives of the function* $f(x) = \exp\left(x + x^2/2\right)$ *are computed by the umbral*

$$\frac{d^m}{dx^m} f(x) = f(x) \sum_{k=0}^{m} \binom{m}{k} I_1(m-k) x^k := f(x)(x + I_1)^m. \tag{2.8}$$

Proof. From Theorem 2.2, $\dfrac{d^m}{dx^m} f(x) = \displaystyle\sum_n I_1(n+m) \frac{x^n}{n!}$. The right-hand side of Theorem 2.5 implies

$$\sum_n \sum_k k! \binom{n}{k} \binom{m}{k} I_1(n-k) I_1(m-k) \frac{x^n}{n!} =$$

$$\sum_k \binom{m}{k} I_1(m-k) x^k \sum_n I_1(n-k) \frac{x^{n-k}}{(n-k)!}$$

$$= f(x) \sum_k \binom{m}{k} I_1(m-k) x^k.$$

The claim follows. □

The recurrence (2.7) is now used to prove periodicity of $I_1(n) \bmod p^r$.

Theorem 2.7. *Let* p *be a prime and* $r \in \mathbb{N}$. *Then the function* $n \mapsto I_1(n) \bmod p^r$ *is periodic, and its minimum period divides* p^r.

Proof. Write $n = cp^r + t$ with $0 \leq t < p^r$. Theorem 2.5 gives

$$I_1(cp^r + t) = \sum_{k=0}^{t} k! \binom{cp^r}{k} \binom{t}{k} I_1(cp^r - k) I_1(t-k). \tag{2.9}$$

For $k > 0$, $\binom{cp^r}{k} k! = (cp^r)(cp^r - 1) \cdots (cp^r - k + 1) \equiv 0 \bmod p^r$ yields

$$I_1(cp^r + t) \equiv I_1(cp^r) I_1(t) \bmod p^r. \tag{2.10}$$

Using Theorem 2.5 again

$$I_1(2p^r) = \sum_{k=0}^{p^r} k! \binom{p^r}{k}^2 I_1)(p^r - k)^2 \equiv I_1(p^r)^2 \bmod p^r \tag{2.11}$$

and then induction on c gives

$$I_1(cp^r) \equiv I_1(p^r)^c \bmod p^r. \tag{2.12}$$

The next step is to show that $I_1(p^r) \equiv 1 \bmod p^r$. Then (2.10) and (2.12) imply the required periodicity. Observe first that for $m \not\equiv 0 \bmod p$,

$$\binom{p^r}{m} = \frac{p^r}{m} \binom{p^r - 1}{m-1} \equiv 0 \bmod p^r, \tag{2.13}$$

so that

$$I_1(p^r) \equiv \sum_{m=0}^{r-1} \binom{p^r}{2mp} \binom{2mp}{mp} \frac{(mp)!}{2^{mp}} \bmod p^r, \tag{2.14}$$

where the upper bound arises from $\nu_p((mp)!) \geq m + \lfloor m/r \rfloor \geq r$ if $m \geq r$.

The final step is to show that

$$\nu_p\binom{p^r}{2mp} = \begin{cases} r-1 & \text{if } m \not\equiv 0 \bmod p \\ r-2 & \text{if } m \equiv 0 \bmod p. \end{cases} \tag{2.15}$$

This would imply $I_1(p^r) \equiv 1 \bmod p^r$ since $\nu_p((mp)!) \geq 2$ for $m \geq 2$. The periodicity of $I_1(n) \bmod p^r$ follows from here.

To prove (2.15), recall Legendre's formula

$$\nu_p(x!) = \frac{x - s_p(x)}{p-1} \tag{2.16}$$

where $s_p(x)$ is the digit sum of x in base p. This gives

$$\begin{aligned} \nu_p\left(\binom{p^r}{2mp}\right) &= \frac{-s_p(p^r) + s_p(2mp) + s_p(p^r - 2mp)}{p-1} \\ &= \frac{-1 + s_p(2m) + s_p(p^{r-1} - 2m)}{p-1}. \end{aligned} \tag{2.17}$$

Write $2m = \sum_{i=0}^{r-2} u_i p^i$ with $0 \leq u_i \leq p-1$. Then

$$\begin{aligned} p^{r-1} - 2m &= p^{r-1} - \sum_{i=0}^{r-2} u_i p^i = 1 + \sum_{i=0}^{r-2} (p-1-u_i)p^i \\ &= (p-u_0) + \sum_{i=1}^{r-2} (p-1-u_i)p^i. \end{aligned}$$

If $m \not\equiv 0 \bmod p$, then

$$s_p(p^{r-1} - 2m) = (p-u_0) + \sum_{i=1}^{r-2} (p-1-u_i). \tag{2.18}$$

On the other hand, if $\nu_p(m) = a$, a direct calculation leads to

$$s_p(p^{r-1} - 2m) = (p-u_a) + \sum_{i=a+1}^{r-2} (p-1-u_i). \tag{2.19}$$

The claim (2.15) now follows from (2.17). □

Corollary 2.8. *Assume $I_1(n) \not\equiv 0 \bmod p$ for $0 \leq n \leq p-1$ and p an odd prime. Then $\nu_p(I_1(n)) \equiv 0$.*

Corollary 2.9. *A prime p divides the sequence $I_1(n)$ infinitely often or never at all.*

In the process of discovering the previous congruences, the following result was obtained by the authors. Even though it is not related yet to the material that follows, it is of intrinsic interest and thus placed here for future use. In the sequel, $\lambda \vdash n$ means λ is a partition of n.

Proposition 2.10. *Let $\lambda = (\lambda_1, \cdots, \lambda_k) \vdash n$ with $\lambda_1 \geq \cdots \geq \lambda_k \geq 1$. Denote $\binom{pn}{p\lambda} = \binom{pn}{p\lambda_1, \cdots, p\lambda_k}$.*
a) If $p \geq 3$ is a prime, then $\binom{pn}{p\lambda} \equiv \binom{n}{\lambda} \bmod p^2$.
b) If $p \geq 5$ is a prime, then $\binom{pn}{p\lambda} \equiv \binom{n}{\lambda} \bmod p^3$.

Proof. The case $k = 2$ is considered first and show $\binom{pn}{pb} \equiv \binom{n}{b} \bmod p^2$. Take an $n \times p$ rectangular grid. Choose pb of these squares in $\binom{pn}{pb}$ ways and paint them red. One option is to paint b entire rows red, call this type-1. This can be done in $\binom{n}{b}$ different ways. In all other cases, there exist *at least two rows* containing t_1, t_2 red squares, respectively, where $0 < t_1, t_2 < p$. Two such coloring are considered equivalent if one is produced from the other by a *cyclic shift* of the squares in each row *independently*. This generates equivalence classes and the number of elements in each of the latter class is then divisible by p^2. Thus, modulo p^2, only type-1 coverings remain.

To prove the general case, choose $p\lambda_i$ of these squares and paint them with color c_i, $1 \leq i \leq k$, in $\binom{pn}{p\lambda}$ ways. Then proceed as in the case $k = 2$. The general case also follows from the special case $k = 2$ and the identity

$$\begin{aligned}\binom{pn}{p\lambda} &= \binom{pn}{p\lambda_1}\binom{p(n-\lambda_1)}{p\lambda_2}\cdots\binom{p(n-\lambda_1-\cdots-\lambda_{k-1}}{p\lambda_k} \\ &\equiv \binom{n}{\lambda_1}\binom{n-\lambda_1}{\lambda_2}\cdots\binom{n-\lambda_1-\cdots-\lambda_{k-1}}{\lambda_k} \bmod p^2 \\ &= \binom{n}{\lambda}.\end{aligned}$$

A similar argument reveals the second congruence. □

3. The involution polynomials

This section introduces a sequence of polynomials generalizing the involution numbers $I_1(n)$. To this end, modify (2.1) so that $I_1(n;1) = I_1(n)$.

Definition 3.1. The *involution polynomials* $I_1(n;t)$ are defined by the recurrence

$$I_1(n;t) = tI_1(n-1;t) + (n-1)I_1(n-2;t), \tag{3.1}$$

with initial conditions $I_1(0;t) = 1$ and $I_1(1;t) = t$.

Proposition 3.2. *The involution polynomials are expressible as*

$$I_1(n;t) = \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2j} \frac{(2j)!}{2^j j!} t^{n-2j}. \tag{3.2}$$

Proof. A direct calculation shows that the right-hand side of (3.2) satisfies the recurrence (3.1) with the same initial conditions as $I_1(n,t)$. □

Theorem 3.1. There is an exponential generating function for the involution polynomials

$$\sum_{n=0}^{\infty} I_1(n;t)\frac{x^n}{n!} = \exp\left(tx + \tfrac{1}{2}x^2\right). \tag{3.3}$$

Proof. Multiply the recurrence (3.1) by $x^n/n!$ and sum over $n \geq 2$ to produce

$$\sum_{n=2}^{\infty} I_1(n;t)\frac{x^n}{n!} = \sum_{n=1}^{\infty} tI_1(n;t)\frac{x^{n+1}}{(n+1)!} + \sum_{n=0}^{\infty} I_1(n;t)\frac{x^{n+2}}{(n+1)n!}. \tag{3.4}$$

Denote the generating function by $h(x,t)$. The recurrence implies $\dfrac{\partial h}{\partial x} = (x+t)h$ and the proof follows from a standard argument. □

Note 3.3. The generating function (1.5) shows the relation

$$I_1(n;t) = \imath^n H_n(-\imath t) \tag{3.5}$$

between the involution polynomials $I_1(n;t)$ and the Hermite polynomials $H_n(t)$.

The next result offers a combinatorial interpretations of the involution polynomials.

Proposition 3.4. *The involution polynomials can be expressed as*

$$I_1(n;t) = \sum_{\pi \in Inv(n)} t^{\alpha_1(\pi)}, \tag{3.6}$$

where $\alpha_1(\pi)$ is the number of fixed points of π.

Proof. Let $g_n(t)$ be the right-hand side in (3.6). Rearrange the set of involutions $\pi \in \mathrm{Inv}(n)$ into two groups according to whether $\pi(n) = n$ or not. In the first case $\pi = \pi_1$ with $\pi_1 \in \mathrm{Inv}(n-1)$. The involution π has the same number of 2-cycles as π_1 and the extra fixed point n. Therefore the term $t^{c_1(\pi)}$ in $g_n(t)$ cancels a unique term in $tg_{n-1}(t)$. In the second case, let $\pi(n) = k$ with $1 \le k \le n-1$. Then π is π_2 times the cycle (nk); that is, $\pi = \pi_2(nk)$, with $\pi_2 \in \mathrm{Inv}(n-2)$. The permutation π_2 has the same number of fixed points as π. Thus, $t^{c_1(\pi)}$ in $g_n(t)$ cancels a unique term in $g_{n-2}(t)$. Summing over n gives the relation

$$g_n(t) = tg_{n-1}(t) + g_{n-2}(t), \tag{3.7}$$

since every term on both sides has been canceled in the previous description. The polynomials $g_n(t)$ and $I_1(n;t)$ satisfy the same recurrence with matching initial conditions. This establishes the assertion. □

4. Arithmetic properties of the numbers $I_1(n)$.

This section discusses the p-adic valuation of the sequence $\{I_1(n)\}$. The analysis begins with the prime $p = 2$.

Theorem 4.1. *The 2-adic valuation of $I_1(n)$ is given by*

$$\nu_2(I_1(n)) = \begin{cases} k & \text{if } n = 4k \\ k & \text{if } n = 4k+1 \\ k+1 & \text{if } n = 4k+2 \\ k+2 & \text{if } n = 4k+3 \end{cases} \tag{4.1}$$

This is equivalent to $\nu_2(I_1(n)) = \left\lfloor \frac{n}{2} \right\rfloor - 2\left\lfloor \frac{n}{4} \right\rfloor + \left\lfloor \frac{n+1}{4} \right\rfloor$.

Proof. Let $n \in \mathbb{N}$ and assume the result is valid up to $n-1$. The proof is divided into four cases according to the residue of n modulo 4. The symbol O_i stands for an odd number.

Case 1: $n = 4k$. The induction hypothesis states that

$$\nu_2(I_1(n-1)) = k+1,\ \nu_2(I_1(n-2)) = k \text{ and } \nu_2(n-1) = 0. \tag{4.2}$$

The recurrence (2.1) implies that $I_1(n) = 2^{k+1}O_1 + 2^k O_2 = 2^k(2O_1 + O_2)$, for some O_1, O_2 odd integers. This proves $\nu_2(I_1(n)) = k$.

Case 2: $n = 4k+1$. The argument is similar to Case 1.

Case 3: $n = 4k + 2$. By induction hypothesis, $I_1(n-1) = 2^k O_1$ and $I_1(n-2) = 2^k O_2$. The recurrence (2.1) now yields $I_1(n) = 2^k (O_1 + O_2 O_3)$, with $O_1 + O_2 O_3$ even, so that $\nu_2(I_1(n)$ is not determined from here. It is necessary to iterate (2.1) to obtain $I_1(n) = nI_1(n-2)+(n-2)I_1(n-3)$. The result now follows immediately.

Case 4: $n = 4k + 3$. The recurrence (2.1) now needs to be iterated twice to produce $I_1(n) = 2(n-1)I_1(n-3) + n(n-3)I_1(n-4)$. Induction gives $I_1(n) = 2^{k+2}\left[O_1 + 2^{1+\nu_2(k)} O_2\right]$, showing that $\nu_2(I_1(n)) = k + 2$.

An alternative proof follows from the recurrence (2.7). Write $n = 4k + r$ for $0 \leq r \leq 3$ and proceed by induction on k. The result follows directly from the identities

$$\begin{aligned}
I_1(4k+1) &= I_1(4k) + 4kI_1(4k-1) \\
I_1(4k+2) &= 2I_1(4k) + 8kI_1(4k-1) + 4k(4k-1)I_1(4k-2) \\
I_1(4k+3) &= 4I_1(4k) + 24kI_1(4k-1) + 12k(4k-1)I_1(4k-2) \\
&\qquad +6\binom{4k}{3} I_1(4k-3) \\
I_1(4k+4) &= 10I_1(4k) + 64kI_1(4k-1) + 48k(4k-1)I_1(4k-2) \\
&\qquad +24\binom{4k}{3} I_1(4k-3) + 24\binom{4k}{4} I_1(4k-4).
\end{aligned}$$

□

The case of $\nu_p(I_1(n))$ for p an odd prime is considered next. Lemma 2.8 shows that if $I_1(n) \not\equiv 0 \bmod p$ for $0 \leq n \leq p-1$, then $\nu_p(I_1(n)) \equiv 0$.

Definition 4.2. The prime p is called *efficient* if $I_1(n) \not\equiv 0 \bmod p$, for every n in the range $0 \leq n \leq p-1$. Otherwise, it is called *inefficient*.

Lemma 2.8 shows that $\nu_p(I_1(n)) \equiv 0$ if p is an efficient prime.

Example 4.3. The values $I_1(0) = 1$, $I_1(1) = 1$, $I_1(2) = 2$ show that $p = 3$ is efficient. Therefore $\nu_3(I_1(n)) \equiv 0$. The prime $p = 5$ is inefficient since $I_1(4) = 10$ is divisible by 5. Similarly $p = 7$ is efficient, in view of the table

n	0	1	2	3	4	5	6
$I_1(n)$	1	1	2	4	10	26	76
$\text{Mod}(I_1(n), 7)$	1	1	2	4	3	5	6

Among the first 100 primes, there are 62 inefficient ones. These are listed in the table below.

(4.3)

5	13	19	23	29	31	43	53
59	61	67	73	79	83	89	97
103	131	137	151	157	163	173	179
181	191	197	199	211	229	233	239
241	281	293	307	317	347	359	367
373	379	389	397	409	419	421	431
433	443	449	457	461	463	479	487
491	499	509	521	523	541		

The p-adic valuation $\nu_p(I_1(n))$ for inefficient primes is (conjecturally) described by a tree structure $\mathbb{T}_p$ and certain modular classes. The case $p = 5$ is prototypical.

Each vertex V of the tree $\mathbb{T}_5$ corresponds to a subset of $\mathbb{N}$. The vertex V is called *terminal* if $\{\nu_5(I_1(n)) : n \in V\}$ reduces to a single value; that is, $\nu_5(I_1(n))$ is independent of $n \in V$; otherwise it is called *non-terminal*. The description of the tree $\mathbb{T}_5$ uses the notation $\Omega_5 := \{0, 1, 2, 3, 4\}$.

The construction begins with a *root vertex* V_0 that represents all $\mathbb{N}$. Since $\nu_5(I_1(n))$ is not a constant function, the vertex V_0 is non-terminal. The root is now split into five different vertices, denoted by $V_{1,k} : k \in \Omega_5$, with

$$V_{1,k} = \{n \in \mathbb{N} : n \equiv k \bmod 5\}. \tag{4.4}$$

These five vertices form the *first level*. Theorem 2.7 shows that

$$I_1(k + 5n) \equiv I_1(k) \bmod 5 \tag{4.5}$$

for $k \in \Omega_5$. The values $I_1(0) = 1$, $I_1(1) = 1$, $I_1(2) = 2$, $I_1(3) = 4$, $I_1(4) = 10$ give

$$\nu_5(V_{1,k}) = 0 \text{ for } 0 \le k \le 3 \text{ and } \nu_5(V_{1,4}) \ge 1. \tag{4.6}$$

Thus, $V_{1,k}$ is a terminal vertex for $0 \le k \le 3$ and $V_{1,4}$ is non-terminal.

In order to determine the valuation of numbers associated to the vertex $V_{1,4}$, that is, numbers of the form $5n_1 + 4$, split the index n_1 according to its residue modulo 5 and write $5n_1 + 4 = 5^2n_2 + 5k + 4$, with $k \in \Omega_5$. Then

$$\nu_5(I_1(5^2n_2 + 5k + 4)) \ge 1, \text{ for } k \in \Omega_5. \tag{4.7}$$

The *second level* is formed by vertices $V_{2,k}$ corresponding to the sets $\{n \in \mathbb{N} : n \equiv 5k + 4 \bmod 5^2\}$. Theorem 2.7 gives

$$I_1(5^2n_2 + 5k + 4) \equiv I_1(5k + 4) \bmod 5^2, \text{ for every } k \in \Omega_5. \tag{4.8}$$

Therefore if $I_1(5k+4) \not\equiv 0 \bmod 5^2$, it follows that $\nu_5(V_{2,k}) = 1$ and $V_{2,k}$ is a terminal vertex. The values

$$I_1(4) \equiv 10,\ I_1(9) \equiv 20,\ I_1(14) \equiv 5,\ I_1(19) \equiv 15,\ I_1(24) \equiv 0 \bmod 5^2, \tag{4.9}$$

show that $\nu_5(V_{2,k}) = 1$, for $k \in \Omega_5$, $k \ne 4$ and, in the single remaining case, $\nu_5(V_{2,4}) \ge 2$.

Conjecture. Assume p is an inefficient prime. Then, for every $n \in \mathbb{N}$, the n-th level of the tree $\mathbb{T}_p$ contains a single non-terminal vertex. This level contains $p - 1$ vertices with valuation $n - 1$ and the single non-terminal vertex has valuation at least n. This determines the tree $\mathbb{T}_p$ and the valuations $\nu_p(I_1(n))$.

5. Partial sums of involution numbers

What happens if the term $\binom{n}{2k}$ is replaced by $\binom{n}{2k+1}$ in the formula

$$I_1(n) = \sum_{k\ge 0} \frac{(2k)!}{k!\,2^k}\binom{n}{2k}? \tag{5.1}$$

It is perhaps convenient to also shift n and define

$$a_n = \sum_{k\ge 0} \frac{(2k)!}{k!\,2^k}\binom{n+1}{2k+1}. \tag{5.2}$$

The next result shows that a_n is actually closely tied to I_1.

Theorem 5.1. *If $n \in \mathbb{N}$, then*

$$a_n = \sum_{j=0}^{n} I_1(j). \tag{5.3}$$

Proof. This is immediate from a simple binomial identity so that

$$\begin{aligned}
\sum_{j=0}^{n} I_1(j) &= \sum_{j=0}^{n} \sum_{k=0}^{\lfloor j/2 \rfloor} \binom{j}{2k} \frac{(2k)!}{k!2^k} \\
&= \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(2k)!}{k!2^k} \sum_{j=\lfloor k/2 \rfloor}^{n} \binom{j}{2k} \\
&= \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(2k)!}{k!2^k} \binom{j+1}{2k+1} \\
&= a_n.
\end{aligned}$$

□

A recurrence for a_n is routinely generated by the WZ-method (see [4, 5]).

Proposition 5.1. The sequence a_n satisfies the recurrence

$$a_n = 2a_{n-1} + (n-2)a_{n-2} - (n-1)a_{n-3}, \text{ for } n \geq 3, \tag{5.4}$$

with initial conditions $a_0 = 1$, $a_1 = 2$ and $a_2 = 4$.

The first few values are tabulated below.

n	0	1	2	3	4	5	6	7	8	9	10
a_n	1	2	4	8	18	44	120	352	1116	3736	13232

This sequence does not appear in OEIS.

Given any sequence $\{q_n\}$ with ordinary generating function $f(x)$, then the partial sums $q_1 + \cdots + q_n$ have the ordinary generating function $f(x)/(1-x)$. The corresponding statement for exponential generating functions is given below.

Lemma 5.2. If $w(x) = \displaystyle\sum_{n=0}^{\infty} c_n \frac{x^n}{n!}$ and $u_n = c_0 + \cdots + c_n$, then

$$\sum_{n=0}^{\infty} u_n \frac{x^n}{n!} = w(x) + e^x \int_0^x e^{-t} w(t)\, dt. \tag{5.5}$$

Proof. Start with $u_n = c_n + u_{n-1}$, multiply through by $x^{n-1}/(n-1)!$ and sum over n. The outcome is the differential equation $g'(x) - g(x) = w'(x)$. Now solve this linear differential equation to obtain the result. □

Corollary 5.3. The exponential generating function for the sequence $\{a_n\}$ is

$$\sum_{n=0}^{\infty} a_n \frac{x^n}{n!} = e^{x+x^2/2} + e^x \int_0^x e^{t^2/2}\, dt. \tag{5.6}$$

Proof. Using $\displaystyle\sum_{n=0}^{\infty} I_1(n) \frac{x^n}{n!} = \exp\left(x + x^2/2\right)$, the claim follows from Lemma 5.2. □

Corollary 5.4. The sequence $\{a_n\}$ satisfies

$$\sum_{k=1}^{n}(-1)^{n-k}\binom{n}{k}a_{k-1} = \begin{cases}(2m)!/2^m\,m! & \text{if } n=2m+1,\\ 0, & \text{if } n=2m.\end{cases} \tag{5.7}$$

Proof. Using the notation of Lemma 5.2,

$$\int_0^x e^{t^2/2}\,dt = e^{-x}[g(x)-w(x)] = e^{-x}\sum_{n=0}^{\infty}[a_n - I_1(n)]\frac{x^n}{n!} = e^{-x}\sum_{n=1}^{\infty}a_{n-1}\frac{x^n}{n!}.$$

Now write e^{-x} as a series and multiply out to arrive at the assertion. □

Corollary 5.5. The following identity holds:

$$\sum_{j=1}^{m}\binom{2m}{2j}a_{2j-1} = \sum_{j=1}^{m}\binom{2m}{2j-1}a_{2j-2}. \tag{5.8}$$

Proof. This is Corollary 5.4 for $n=2m$. □

6. Arithmetic properties of the sequence a_n

The next statement is the corresponding counterpart to Theorem 4.1.

Theorem 6.1. *The 2-adic valuation of the sequence a_n is given by*

$$\nu_2(a_n) = \begin{cases}k, & \text{if } n=4k-3,\\ k+1, & \text{if } n=4k-2,\\ k, & \text{if } n=4k,\\ \nu_2(k)+k+2, & \text{if } n=4k-1.\end{cases} \tag{6.1}$$

Proof. The inductive proof distinguishes the four values of n modulo 4.

Case 1: $n=4k-3$. Then, the induction hypothesis shows that

$$a_{n-1} = 2^{k-1}O_1,\ a_{n-2} = 2^{k+1+\nu_2(k-1)}O_2,\ \text{and } a_{n-3} = 2^kO_3 \tag{6.2}$$

with O_j odd integers. Then the recurrence (5.4) implies

$$a_{4k-3} = 2^k\left[O_1 + (4k-5)2^{1+\nu_2(k-1)}O_2 - (k-1)2^2O_3\right]. \tag{6.3}$$

Thus $\nu_2(a_{4k-3}) = k$.

Case 2: $n=4k$. Then

$$a_{n-1} = 2^{\nu_2(k)+k+2}O_1,\ a_{n-2} = 2^{k+1}O_2,\ \text{and } a_{n-3} = 2^kO_3 \tag{6.4}$$

and (5.4) implies

$$a_{4k} = 2^k\left[2^{\nu_2(k)+3}O_1 + 2^2(2k-1)O_2 - (4k-1)O_3\right] \tag{6.5}$$

and $\nu_2(a_{4k}) = k$ follows form here.

Case 3: $n=4k-2$. Then, as in the proof of Theorem 4.1, the recurrence needs to be iterated to produce

$$a_{4k-2} = 4ka_{4k-4} + (4k-7)a_{4k-5} - 8(k-1)a_{4k-6}. \tag{6.6}$$

The induction hypothesis gives

$$a_{4k-2} = 2^{k+1}\left[kO_1 + (4k-7)2^{\nu_2(k-1)}O_2 - 4(k-1)O_3\right]. \tag{6.7}$$

For k odd, the first term in the square bracket is odd and the other two are even. For k even, the second term is odd and the other two are even. In either case, $\nu_2(a_{4k-2}) = k+1$.

Case 4: $n = 4k-1$. The statement to be proved is

$$\nu_2(a_n) = \nu_2(4k) + k. \tag{6.8}$$

Observe that

$$\begin{aligned} a_{4k-1} &= \sum_{j=0}^{2k-1} \frac{(2j)!}{j!2^j}\binom{4k}{2j+1} \\ &= 4k\sum_{j=0}^{2k-1} \frac{(2j)!}{j!2^j}\frac{1}{2j+1}\binom{4k-1}{2j}. \end{aligned} \tag{6.9}$$

Therefore, it suffices to show that $\nu_2(b_k) = k$ where

$$b_k = \sum_{j=0}^{2k-1} \frac{(2j)!}{j!2^j}\frac{1}{2j+1}\binom{4k-1}{2j}. \tag{6.10}$$

It should be noted that not all summands in b_k are integers.

The proof of this last step is based on the valuations of the sum

$$F(\alpha,\beta,k) = \sum_{j=0}^{2k-1} (2j+\alpha)^\beta \frac{(2j)!}{j!2^j}\binom{4k-1}{2j}. \tag{6.11}$$

Observe that

$$F(\alpha,0,k) = \sum_{j=0}^{2k-1} \frac{(2j)!}{j!2^j}\binom{4k-1}{2j} = I_1(4k-1). \tag{6.12}$$

The next lemma relates $F(\alpha,1,k)$ with the involution numbers.

Lemma 6.2. *Let α, $k \in \mathbb{N}$. Then*

$$F(\alpha,1,k) = \alpha I_{4k-1} + 2(4k-1)(2k-1)I_{4k-3}. \tag{6.13}$$

Proof. Simply observe that

$$F(\alpha,1,k) = \alpha F(\alpha,0,k) + 2\sum_{j=1}^{2k-1} j\cdot\frac{(2j)!}{j!2^j}\binom{4k-1}{2j}$$

and then check that the last sum is $2(4k-1)(2k-1)I_{4k-3}$. □

The 2-adic valuation of $F(\alpha,\beta,k)$ is computed next when α, $\beta \in \mathbb{N}$ and α is odd.

Theorem 6.3. *Let α, $\beta \in \mathbb{N}$ with α odd. Then*

$$\nu_2(F(\alpha,\beta,k)) = \begin{cases} k+1 & \text{if } \beta \text{ is even,} \\ k & \text{if } \beta \text{ is odd.} \end{cases} \tag{6.14}$$

Proof. The case $\beta = 0$ is Theorem 4.1. The case $\beta = 1$ is obtained from the identity (6.13) and Theorem 4.1. The rest of the proof is divided according to the parity of β.

Case 1: $\beta > 1$ odd. Expand $(2j+\alpha)^\beta$ by the binomial theorem to obtain

$$F(\alpha,\beta,k) = \sum_{\ell=0}^{\beta} \binom{\beta}{\ell} \alpha^{\beta-\ell} \sum_{j=0}^{2k-1} \binom{4k-1}{2j} \frac{(2j)!}{j!2^j} 2^\ell j^\ell. \tag{6.15}$$

The term corresponding to $\ell = 0$ is

$$t_{\ell=0} := \alpha^\beta \sum_{j=0}^{2k-1} \binom{4k-1}{2j} \frac{(2j)!}{j!2^j} = \alpha^\beta I_1(4k-1). \tag{6.16}$$

Theorem 4.1 gives its 2-adic valuation as

$$\nu_2(t_{\ell=0}) = \nu_2(I_{4k-1}) = k+1. \tag{6.17}$$

The term for $\ell = 1$ is

$$t_{\ell=1} := \beta\alpha^{\beta-1} \sum_{j=0}^{2k-1} \binom{4k-1}{2j} \frac{(2j)!}{j!2^j} \cdot 2j = 2(4k-1)(2k-1) I_1(4k-3) \tag{6.18}$$

and its 2-adic valuation is

$$\nu_2(t_{\ell=1}) = \nu_2(I_{4k-3}) = k. \tag{6.19}$$

For the remaining terms in the sum $F(\alpha,\beta,k)$ use the identity

$$j^\ell = \sum_{r=1}^{\ell} c_r \frac{j!}{(j-r)!} \tag{6.20}$$

where $c_r \in \mathbb{Z}$ (these are the Stirling numbers, but only their integrality matters here). This leads to the expression

$$\sum_{\ell=2}^{\beta} \binom{\beta}{\ell} \alpha^{\beta-\ell} \sum_{r=1}^{\ell} c_r \sum_{j=0}^{2k-1} \binom{4k-1}{2j} \frac{(2j)!}{(j-r)!2^{j-\ell}}. \tag{6.21}$$

The theorem now follows from the fact that the internal sum has 2-adic valuation at least $k+1$. This implies that $\ell = 1$ controls the valuation. In order to verify this statement, observe that

$$\begin{aligned}
\sum_{j=0}^{2k-1} \binom{4k-1}{2j} \frac{(2j)!}{(j-r)!2^{j-\ell}} &= 2^{\ell-r} \sum_{j=0}^{2k-1} \binom{4k-1}{2j} \frac{(2j)!}{(j-r)!2^{r-j}} \\
&= 2^{\ell-r} \frac{(4k-1)!}{(4k-2r-1)!} \sum_{m=0}^{2k-r} \binom{4k-2r-1}{2m} \frac{(2m)!}{m!2^m} \\
&= 2^{\ell-r} \frac{(4k-1)!}{(4k-2r-1)!} I_{4k-2r-1}.
\end{aligned}$$

A direct application of Theorem 4.1 shows that

$$\nu_2\left(\frac{2^{\ell-r}(4k-1)!}{(4k-2r-1)!} I_{4k-2r-1}\right) \geq \ell - r + \left(r + \left\lfloor \frac{r}{2} \right\rfloor\right) + \left(k + \left\lfloor \frac{r}{2} \right\rfloor - 1\right) \geq \ell + k - 1.$$

The statement about the valuation of the internal sums is now immediate since $\ell \geq 2$.

Case 2: β even. As in the case β odd, the valuations of the internal sums are bounded from below by $\ell+k-1$. In particular, the lower bound is at least $k+2$ if $\ell \geq 3$. This leads to the decomposition

$$F(\alpha,\beta,k) = X_1(\alpha,\beta,k) + X_2(\alpha,\beta,k) + X_3(\alpha,\beta,k) \tag{6.22}$$

where

$$\begin{aligned} X_1(\alpha,\beta,k) &= \alpha^\beta \sum_{j=0}^{2k-1} \binom{4k-1}{2j} \frac{(2j)!}{j!2^j} \\ X_2(\alpha,\beta,k) &= \sum_{\ell=1}^{2} \binom{\beta}{\ell} \alpha^{\beta-\ell} \sum_{j=0}^{2k-1} \binom{4k-1}{2j} \frac{(2j)!}{j!2^j} (2j)^\ell \\ X_3(\alpha,\beta,k) &= \sum_{\ell=3}^{\beta} \binom{\beta}{\ell} \alpha^{\beta-\ell} \sum_{j=0}^{2k-1} \binom{4k-1}{2j} \frac{(2j)!}{j!2^j} (2j)^\ell. \end{aligned} \tag{6.23}$$

Then $\nu_2(X_3(\alpha,\beta,k)) \geq k+2$. It is now shown that $\nu_2(X_1(\alpha,\beta,k)) = k+1$ and $\nu_2(X_2(\alpha,\beta,k)) \geq k+3$. This proves the formula for the valuation of $F(\alpha,\beta,k)$ when β is even.

To prove the statement about the valuation of X_1 use the identity $X_1(\alpha,\beta,k) = \alpha^\beta I_1(4k-1)$ and Theorem 4.1. The proof of the corresponding formula for X_2 starts with the expression

$$\begin{aligned} X_2(\alpha,\beta,k) &= \beta\alpha^{\beta-1} \sum_{j=0}^{2k-1} \binom{4k-1}{2j} \frac{(2j)!}{j!2^j} (2j) \\ &+ \binom{\beta}{2} \alpha^{\beta-2} \sum_{j=0}^{2k-1} \binom{4k-1}{2j} \frac{(2j)!}{j!2^j} (4j^2) \end{aligned} \tag{6.24}$$

and then use $4j^2 = 4j(j-1)+4j$ and the identities

$$\begin{aligned} \sum_{j=0}^{2k-1} \binom{4k-1}{2j} \frac{(2j)!}{j!2^j} (2j) &= (4k-1)(4k-2) I_{4k-3} \\ \sum_{j=0}^{2k-1} \binom{4k-1}{2j} \frac{(2j)!}{j!2^j} (4j^2) &= (4k-1)(4k-2)(4k-3)(4k-4) I_{4k-5} \end{aligned}$$

to arrive at

$$\begin{aligned} X_2(\alpha,\beta,k) &= 2\beta\alpha^{\beta-2}(\alpha+\beta-1)(2k-1)(4k-1) I_{4k-3} \\ &+ 8\binom{\beta}{2} \alpha^{\beta-2}(4k-1)(2k-1)(4k-3)(k-1) I_{4k-5}. \end{aligned}$$

Then Theorem 4.1 implies

$$\nu_2(\beta) + \nu_2(\alpha+\beta-1) + 1 + \nu_2(I_{4k-3}) = \nu_2(\beta) + \nu_2(\alpha+\beta-1) + k \geq k+2, \tag{6.25}$$

and the valuation of the second term is

$$\nu_2(\beta) - 1 + 3 + \nu_2(I_{4k-5}) = \nu_2(\beta) + 2 + k \geq k+3. \tag{6.26}$$

The statement about $\nu_2(X_2)$ is established . The formula for $\nu_2(F(\alpha,\beta,k))$, when β is even, follows from these results. □

The remainder of the proof of Theorem 6.1 has been reduced to verifying that $\nu_2(b_k) = k$, where $b_k = F(1, -1, k)$ is defined in (6.10).

For $m \in \mathbb{N}$ and a odd, Euler's theorem yields

$$a^{-1} \equiv a^{\varphi(2^m)-1} = a^{2^{m-1}-1} \bmod 2^m. \tag{6.27}$$

Therefore

$$\begin{aligned} F(1,-1,k) &\equiv \sum_{j=0}^{2k-1} (2j+1)^{2^{m-1}-1} \frac{(2j)!}{j!2^j} \binom{4k-1}{2j} \\ &= F(1, 2^{m-1}-1, k) \bmod 2^m. \end{aligned}$$

Since $2^{m-1} - 1$ is odd, Proposition 6.3 gives

$$\nu_2(F(1, 2^{m-1}-1, k)) = k. \tag{6.28}$$

Now choose $m = k$ to compute

$$F(1, -1, k) \equiv F(1, 2^{k-1} - 1, k) \equiv 0 \bmod 2^k \tag{6.29}$$

and then choose $m = k + 1$ to obtain

$$F(1, -1, k) \equiv F(1, 2^k - 1, k) \not\equiv 0 \bmod 2^{k+1}. \tag{6.30}$$

It follows that $\nu_2(F(1, -1, k)) = k$, as desired. The proof of Theorem 6.1 is now complete. □

Note 6.4. For p odd, the p-adic valuation of a_n also exhibits some interesting patterns which will be investigated in the future. For instance, when $p = 3$, it is noted that

$$\nu_3(a_n) = 0 \text{ if } n \not\equiv 8 \bmod 9 \tag{6.31}$$

and

$$\nu_3(a_{9n+8}) = \begin{cases} 0 & \text{if} \quad n \equiv 0 \bmod 3, \\ 0 & \text{if} \quad n \equiv 1 \bmod 3, \\ \nu_3(n+1) & \text{if} \quad n \equiv 2 \bmod 3. \end{cases} \tag{6.32}$$

Similar formulas may be tested out experimentally for other primes.

7. Permutations of restricted length

Let $n \in \mathbb{N}$ and $0 \le \ell \le n$. This section considers the set

$$C_{n,\ell} = \{\pi \in \mathfrak{S}_n \Big| \text{ every cycle in } \pi \text{ is of length at most } \ell\}$$

and its cardinality $d_{n,\ell} = \#C_{n,\ell}$.

Proposition 7.1. *The numbers $d_{n,\ell}$ satisfy the recurrence*

$$d_{n+1,\ell} = d_{n,\ell} + \frac{n!}{(n-1)!} d_{n-1,\ell} + \frac{n!}{(n-2)!} d_{n-2,\ell} + \cdots + \frac{n!}{(n-\ell+1)!} d_{n+1-\ell,\ell}.$$

Equivalently

$$\sum_{n=0}^{\infty} d_{n,\ell} \frac{x^n}{n!} = \exp\left(x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots + \frac{x^\ell}{\ell}\right). \tag{7.1}$$

Proof. For $1 \leq j \leq \ell$, the number 1 is in exactly $n!/(n-j+1)!$ cycles of length j. Choose labels from $[n+1]$. Now count the elements in $C_{n+1,\ell}$ according to the length of the cycle containing the number 1. □

Definition 7.2. For n, $\ell \in \mathbb{N}$ and indeterminates Y_1, $Y_2, \cdots, Y_\ell$, the square matrix $M_{n,\ell}(\mathbf{Y})$, of size n, has entries

$$M_{n,\ell}(k,j) = \begin{cases} \imath^{j-k} Y_{j-k+1} & \text{if } 0 \leq j-k \leq \ell-1, \\ \imath j & \text{if } k = j+1, \\ 0, & \text{otherwise.} \end{cases} \tag{7.2}$$

Example 7.3. Let $n = 5$ and $\ell = 4$, then

$$M_{5,4}(\mathbf{Y}) = \begin{pmatrix} Y_1 & \imath Y_2 & \imath^2 Y_3 & \imath^3 Y_4 & 0 \\ \imath & Y_1 & \imath Y_2 & \imath^2 Y_3 & \imath^3 Y_4 \\ 0 & 2\imath & Y_1 & \imath Y_2 & \imath^2 Y_3 \\ 0 & 0 & 3\imath & Y_1 & \imath Y_2 \\ 0 & 0 & 0 & 4\imath & Y_1 \end{pmatrix}. \tag{7.3}$$

A multivariate generalization of Proposition 7.1 is given in the form of the above Toeplitz matrices.

Theorem 7.1. The exponential generating function for the determinants of $M_{n,\ell}(\mathbf{Y})$ is

$$\sum_{n=0}^{\infty} \det(M_{n,\ell}(\mathbf{Y})) \frac{x^n}{n!} = \exp\left(Y_1 x + Y_2 \frac{x^2}{2} + \cdots + Y_\ell \frac{x^\ell}{\ell} \right). \tag{7.4}$$

Proof. Fix ℓ and let $g_{n,\ell}(\mathbf{Y}) = \det(M_{n,\ell}(\mathbf{Y}))$. Use successive Laplace expansion of $g_{n,\ell}$ along the last column to obtain

$$\begin{aligned} g_{n,\ell}(\mathbf{Y}) = g_{n-1,\ell}(\mathbf{Y}) Y_1 + (n-1) g_{n-2,\ell}(\mathbf{Y}) Y_2 + (n-1)(n-2) g_{n-3,\ell}(\mathbf{Y}) Y_3 + \cdots + \\ + (n-1)(n-2)\cdots(n-\ell+1) g_{n-\ell,\ell}(\mathbf{Y}) Y_\ell. \end{aligned}$$

Now set $F_\ell(x;\mathbf{Y}) = \displaystyle\sum_{n=0}^{\infty} g_{n,\ell}(\mathbf{Y}) \frac{x^n}{n!}$. The recurrence shows that $F_\ell(x;\mathbf{Y})$ matches the right-hand side of (7.4). □

Comparing coefficients in the expansion (7.4) gives a statistic on the set $C_{n,\ell}$.

Theorem 7.2. Let $n \in \mathbb{N}$ and $0 \leq \ell \leq n$. Recall $\alpha_t(\pi)$ = number of t-cycles in $\pi \in \mathfrak{S}_n$. Then

$$\det M_{n,\ell}(Y_1, \cdots, Y_\ell) = \sum_{\pi \in C_{n,\ell}} Y_1^{\alpha_1(\pi)} \cdots Y_2^{\alpha_\ell(\pi)}. \tag{7.5}$$

Example 7.4. Let $n = 5$ and $\ell = 4$. Then the cycle-index polynomial is computed as

$$\begin{aligned} \det M_{5,4}(\mathbf{Y}) &= Y_1^5 + 10 Y_1^3 Y_2 + 20 Y_1^2 Y_3 + 15 Y_1 Y_2^2 + 30 Y_1 Y_4 + 20 Y_2 Y_3 \\ &= \sum_{\pi \in C_{5,4}} Y_1^{\alpha_1(\pi)} Y_2^{\alpha_2(\pi)} Y_3^{\alpha_3(\pi)} Y_4^{\alpha_4(\pi)} \end{aligned}$$

and it encodes the statistic on the set $C_{5,4}$. For instance, 20 permutations in $\mathfrak{S}_5$ are a product of a 2-cycle and a 3-cycle. Also, there are $\#C_{5,4} = \det M_{5,4}(\mathbf{1}) = 96$

permutations formed by cycles of length 4 or less. This is $5! = 120$ minus the 24 cycles of length 5.

Note 7.5. The special case $Y_j = 1$ (for all j) produces

$$d_{n,\ell} = \det(M_{n,\ell})(\mathbf{1}). \tag{7.6}$$

8. Asymptotics

This section considers the asymptotic behavior, as $n \to \infty$ with ℓ fixed, of the numbers $d_{n,\ell} = \#C_{n,\ell}$, counting the number of permutations in $\mathfrak{S}_n$ with every cycle of length at most ℓ. Their exponential generating function is

$$f_\ell(z) = \sum_{n=0}^{\infty} d_{n,\ell} \frac{z^n}{n!} = \exp\left(z + \frac{z^2}{2} + \cdots + \frac{z^\ell}{\ell}\right). \tag{8.1}$$

Several authors provide asymptotic expansions for $f_\ell(z)$ (see Moser-Wyman [3] and Knuth [2] when $\ell = 2$; Wimp-Zeilberger [7] for any ℓ using methods from Birkhoff-Trjitzinsky). In this section, the general case is revisited using the *saddle-point* technique in order to generate a first-order estimate.

Cauchy's integral formula gives

$$d_{n,\ell} = \frac{n!}{2\pi\imath} \oint_C \frac{f_\ell(z)}{z^{n+1}} dz \tag{8.2}$$

where C is a simple closed curve around the origin. In the analysis presented here C is a circle of radius r. Therefore

$$\begin{aligned} d_{n,\ell} &= \frac{n!}{2\pi\imath} \int_{|z|=r} f_\ell(z) \exp(-n \log z) \frac{dz}{z} \\ &= \frac{n!}{2\pi\imath} \int_{|z|=r} \exp\left(z + \frac{z^2}{2} + \cdots + \frac{z^\ell}{\ell} - n \log z\right) \frac{dz}{z} \end{aligned} \tag{8.3}$$

The saddle-point method ([1, Theorem X]; also [3])gives

$$d_{n,\ell} \sim \frac{n!}{\sqrt{2\pi\ell n}} \exp\left(r + r^2/2 + \cdots + r^\ell/\ell - n \log r\right) \tag{8.4}$$

where the saddle point $r_+ \in \mathbb{R}^+$ is defined by the equation

$$\frac{d}{dr}\left(r + r^2/2 + \cdots + r^\ell/\ell - n \log r\right) = 1 + r + r^2 + \cdots + r^{\ell-1} - \frac{n}{r} = 0. \tag{8.5}$$

This is equivalent to $r + r^2 + \cdots + r^\ell = n$. To obtain information about the saddle point r_+, it is convenient to rewrite (8.5) in the form

$$r\left(\frac{1 - r^{-\ell}}{1 - r^{-1}}\right)^{1/\ell} - \eta = 0, \qquad \text{for } \eta \in \mathbb{C}. \tag{8.6}$$

This defines $r = r(\eta)$. For $\eta = n^{1/\ell}$, this becomes $r(n^{1/\ell}) = r_+$.

The limiting value $r/\eta \to 1$, as $r \to \infty$, suggests the asymptotic expansion

$$r = r(\eta) = \eta + \alpha_0 + \frac{\alpha_{-1}}{\eta} + \frac{\alpha_{-2}}{\eta^2} + \cdots, \text{ as } \eta \to \infty. \tag{8.7}$$

The exponent in (8.4) is now written as $\Phi(\eta) - \eta^\ell \log \eta$, with

$$\Phi(\eta) = r(\eta) + \frac{1}{2} r(\eta)^2 + \cdots + \frac{1}{\ell} r(\eta)^\ell - \eta^\ell \log \frac{r(\eta)}{\eta}. \tag{8.8}$$

The expansion (8.7) leads to

$$\Phi(\eta) = \beta_\ell \eta^\ell + \cdots + \beta_1 \eta + \beta_0 + \frac{\beta_{-1}}{\eta} + \cdots \tag{8.9}$$

and the relevant contributions to the behavior of (8.4) come from the positive powers in this expansion. To compute these contributions observe that, for $k > 0$,

$$\beta_k = \frac{1}{2\pi\imath}\int_C \frac{\Phi(\eta)}{\eta^{k+1}} d\eta = \frac{1}{2\pi\imath}\int_C \frac{\Phi'(\eta)}{k\eta^k} d\eta, \tag{8.10}$$

where the second expression is obtained by using Cauchy's integral formula on the expansion of $\Phi'(\eta)$. The contour C is made to pass through the point r_+. Therefore, using $\Phi'(\eta) = \eta^{\ell-1} - \ell\eta^{\ell-1}\log\frac{r}{\eta}$ it follows that, for $0 < k < \ell$,

$$\begin{aligned} \beta_k &= \frac{1}{2\pi\imath}\int_C \frac{1}{k\eta^{k-\ell+1}} d\eta - \frac{1}{2\pi i}\int_C \frac{\ell\log\frac{r}{\eta}}{k\eta^{k-\ell+1}} d\eta \\ &= -\frac{1}{2\pi\imath}\int_C \frac{\ell\log\frac{r}{\eta}}{k\eta^{k-\ell+1}} d\eta \end{aligned} \tag{8.11}$$

since the first integral vanishes. This is now written as

$$\begin{aligned} \beta_k &= -\frac{1}{2\pi\imath}\frac{\ell}{k(\ell-k)}\int_C \log\frac{r}{\eta}\frac{d}{d\eta}\eta^{\ell-k}\, d\eta \\ &= \frac{1}{2\pi\imath}\frac{\ell}{k(\ell-k)}\int_C \left(\frac{r'}{r} - \frac{1}{\eta}\right)\eta^{\ell-k} d\eta \\ &= \frac{1}{2\pi\imath}\frac{\ell}{k(\ell-k)}\int_C \frac{r'}{r}\eta^{\ell-k} d\eta \end{aligned} \tag{8.12}$$

since the integral of $\eta^{\ell-k-1}$ vanishes. Now make the change of variables $r \mapsto r(\eta)$ to obtain

$$\beta_k = \frac{1}{2\pi\imath}\int_{C_1} \frac{\ell}{k(\ell-k)}\left(\frac{1-r^{-\ell}}{1-r^{-1}}\right)^{\frac{\ell-k}{k}} r^{\ell-k-1} dr. \tag{8.13}$$

A residue calculation then gives

$$\beta_k = \frac{\ell}{k(\ell-k)}\frac{1}{(\ell-k)!}\left(\frac{\ell-k}{\ell}\right)\left(\frac{\ell-k}{\ell}+1\right)\cdots\left(\frac{\ell-k}{\ell}+k-1\right). \tag{8.14}$$

An easier calculation, left to the reader, gives $\beta_0 = -\dfrac{1}{\ell}\displaystyle\sum_{j=2}^{\ell}\frac{1}{j}$ and $\beta_\ell = \dfrac{1}{\ell}$.

Combining the above, using $\eta^\ell = n$ in the form

$$e^{-\eta^\ell \log\eta} = n^{-n/\ell} \tag{8.15}$$

and invoking Stirling's formula $n! \sim \sqrt{2\pi n}\, n^n e^{-n}$ yields the following result.

Theorem 8.1. *Let $\ell \in \mathbb{N}$ be fixed. Then*

$$d_{n,\ell} \sim \frac{1}{\sqrt{\ell}} n^{n(1-1/\ell)} \exp\left(-\frac{1}{\ell}\sum_{j=2}^{\ell}\frac{1}{j} + n\left(\frac{1}{\ell}-1\right) + \sum_{k=1}^{\ell-1} n^{\frac{\ell-k}{\ell}}\frac{\ell}{k!(\ell-k)k}\prod_{j=0}^{k-1}\left(\frac{k}{\ell}+j\right)\right).$$

Notice that in the case $\ell = 2$ the theorem above gives the well-known asymptotic formula [6, p. 187]

$$d_{n,2} \sim \frac{n^{n/2}}{\sqrt{2}} e^{-\frac{1}{4}-\frac{n}{2}+\sqrt{n}}. \tag{8.16}$$

Acknowledgments. The second author acknowledges the partial support of NSF-DMS 1112656.

References

[1] W. K. Hayman. A generalization of Stirling's formula. *J. Reine Angew Math.*, 196:67–95, 1956.
[2] D. E. Knuth. *Art of Computer Programming. Sorting and Searching*, volume 3. Addison-Wesley, Reading, Mass., 1st. edition, 1997.
[3] I. Moser and M. Wyman. Asymptotic expansions. *Canad. J. Math.*, 8:225–233, 1956.
[4] I. Nemes, M. Petkovsek, H. Wilf, and D. Zeilberger. How to do MONTHLY problems with your computer. *Amer. Math. Monthly*, 104:505–519, 1997.
[5] M. Petkovsek, H. Wilf, and D. Zeilberger. *A=B*. A. K. Peters, 1st. edition, 1996.
[6] H. S. Wilf. *generatingfunctionology*. Academic Press, 1st edition, 1990.
[7] J. Wimp and D. Zeilberger. Resurrecting the asymptotics of linear recurrences. *J. Math. Anal. Appl.*, 111:162–176, 1985.

Department of Mathematics, Tulane University, New Orleans, LA 70118
E-mail address: tamdeber@tulane.edu

Department of Mathematics, Tulane University, New Orleans, LA 70118
E-mail address: vhm@tulane.edu