Counting tournament score sequences

Anders Claesson Mark Dukes Atli Fannar Franklín
Sigurður Örn Stefánsson

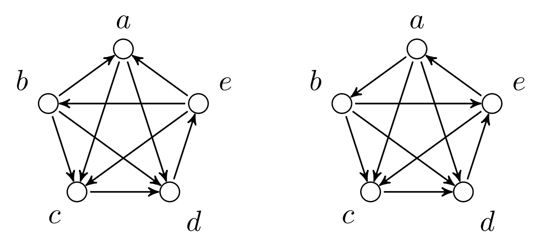
8 September 2022

Abstract

The score sequence of a tournament is the sequence of the out-degrees of its vertices arranged in nondecreasing order. The problem of counting score sequences of a tournament with n vertices is more than 100 years old (MacMahon 1920). In 2013 Hanna conjectured a surprising and elegant recursion for these numbers. We settle this conjecture in the affirmative by showing that it is a corollary to our main theorem, which is a factorization of the generating function for score sequences with a distinguished index. We also derive a closed formula and a quadratic time algorithm for counting score sequences.

1 Introduction

In 1953 Landau [7] used oriented complete graphs—also called *tournaments*—to model pecking orders. If the vertices of the complete graph represent players (rather than chickens), then the initial vertex of a directed edge signifies the winner of a game between the two end-point players. The number of wins of a player is equal to the number of outgoing edges from that vertex. A *score sequence* is a sequence of these number of wins given in a nondecreasing order. For instance, with 3 players there are two possible score sequences, namely $(0, 1, 2)$ and $(1, 1, 1)$. Note that non-isomorphic tournaments may give rise to the same score sequence. With 5 players there are, up to isomorphism, 12 tournaments but only 9 score sequences. To be even more specific, here are two non-isomorphic[1] tournaments:



[1]The longest directed path between any pair of vertices in the left-hand graph is 3, while in the right-hand graph there is a directed path of length 4 from c to b.

The score sequence associated with both is $(1,1,2,3,3)$. The following characterization of score sequences is known as Landau's theorem.

Theorem 1 (Landau [7]). *A sequence of integers $s = (s_0, \dots, s_{n-1})$ is a score sequence if and only if*

(1) $0 \le s_0 \le s_1 \le \cdots \le s_{n-1} \le n-1$,

(2) $s_0 + \cdots + s_{k-1} \ge \binom{k}{2}$ *for* $1 \le k < n$, *and*

(3) $s_0 + \cdots + s_{n-1} = \binom{n}{2}$.

Let S_n be the set of score sequences of length n. There is no known closed formula for the cardinalities $|S_n|$ or their generating function.

It should be noted that Landau was not the first person to study score sequences, or attempt to count them. MacMahon [8] used symmetric functions and hand calculations to determine $|S_n|$ for $n \le 9$ in 1920. Building on Landau's work, Narayana and Bent [9], in 1964, derived a multivariate recursive formula for determining $|S_n|$. They used it to give a table for $n \le 36$. In 1968 Riordan [10] gave a simpler and more efficient recursion, but unfortunately it turned out to be incorrect [11].

Let $[a,b]$ denote the interval of integers $\{a, a+1, \dots, b\}$. We may view a score sequence $s \in S_n$ as an endofunction $s : [0, n-1] \to [0, n-1]$. We now introduce the notion of a *pointed score sequence*. Define $S_n^\bullet$ as the Cartesian product $S_n^\bullet = S_n \times [0, n-1]$. We call the members of $S_n^\bullet$ *pointed score sequences*; e.g. there are 6 pointed score sequences in $S_3^\bullet$:

$$((0,1,2),0),\ ((0,1,2),1),\ ((0,1,2),2),$$
$$((1,1,1),0),\ ((1,1,1),1),\ ((1,1,1),2).$$

Let $(s,i) \in S_n^\bullet$. Depending on the context, the element i will be interpreted as a position (element in the domain) or a value (element in the codomain) of s. If i is a value, then the cardinality of the fiber $s^{-1}(i)$ is the number of times i occurs in s; this number may be zero. Let

$$S_n^\bullet(t) = \sum_{(s,i) \in S_n^\bullet} t^{|s^{-1}(i)|}$$

be the polynomial recording the distribution of the statistic $(s,i) \mapsto |s^{-1}(i)|$ on $S_n^\bullet$. As an example, $S_3^\bullet(t) = 2 + 3t + t^3$. Let

$$S^\bullet(x,t) = \sum_{n \ge 1} S_n^\bullet(t) x^n.$$

To present the bijection that is the main result of this paper, we will first introduce a particular type of multiset that is an essential ingredient in our deconstruction of a pointed score sequence. At first glance it is not obvious what the relevance of these multisets to score sequences is. The Erdős-Ginzburg-Ziv Theorem [3] is the following.

Theorem 2 (Erdős, Ginzburg, and Ziv [3]). *Each set of $2n-1$ integers contains some subset of n elements the sum of which is a multiple of n.*

We define EGZ_n as the set of multisets of size n with elements in the cyclic group $\mathbb{Z}_n$ whose sum is $\binom{n}{2}$ modulo n. This terminology is motivated by the simple one-to-one correspondence between EGZ_n and the sets considered by Erdős, Ginzburg, and Ziv that appear in the proof of the following lemma.

Lemma 3. *There is a one-to-one correspondence between* EGZ_n *and n-element subsets of* $[1, 2n-1]$ *whose sum is a multiple of n.*

Proof. Let $A = \{a_1, \ldots, a_n\}$ be a subset of $[1, 2n-1]$ such that $a_1 + \cdots + a_n$ is divisible by n. Without loss of generality we can further assume that $a_1 < a_2 < \cdots < a_n$. Let $b_i = a_i - i$. We claim that $A \mapsto \{b_1, \ldots, b_n\}$ is a bijection onto EGZ_n. Clearly, $i \leq a_i \leq n+i-1$ and hence $0 \leq b_i \leq n-1$. In this manner we can consider each b_i as an element of $\mathbb{Z}_n$. The sum $a_1+\cdots+a_n$ is divisible by n, by assumption, and hence

$$b_1 + \cdots + b_n \equiv a_1 + \cdots + a_n - 1 - 2 - \cdots - n \equiv -\binom{n+1}{2} \pmod{n}.$$

That this is congruent to $\binom{n}{2}$ modulo n follows from $\binom{n}{2} + \binom{n+1}{2} = n^2$. Conversely, given a multiset $B = \{b_1, \ldots, b_n\}$, with $b_1 \leq \cdots \leq b_n$, in EGZ_n we define the set $A = \{a_1, \ldots, a_n\}$ by $a_i = b_i + i$. One can verify that A is a subset of $[1, 2n-1]$ whose sum is divisible by n, and that $A \mapsto B$ is the inverse of the map above, but we omit the details. $\square$

The sequence of cardinalities $(|\mathrm{EGZ}_n|)_{n\geq 1}$ is entry A145855 in the OEIS [6]. As recorded in that OEIS entry, Jovović conjectured and Alekseyev [1] proved in 2008 that

$$|\mathrm{EGZ}_n| = \frac{1}{2n}\sum_{d|n}(-1)^{n-d}\varphi(n/d)\binom{2d}{d}, \tag{4}$$

where the sum runs over all positive divisors of n and φ is Euler's totient function. A generalization of this result was given by Chern [2] in 2019.

The zeros in a multiset $M \in \mathrm{EGZ}_n$ play a prominent role in our construction. We now introduce a generating function to record their number. For a multiset $M \in \mathrm{EGZ}_n$ let $|M|_i$ be the number of occurrences of i in M. Furthermore, let

$$\mathrm{EGZ}_n(t) = \sum_{M \in \mathrm{EGZ}_n} t^{|M|_0}$$

be the polynomial recording the distribution of zeros in multisets belonging to EGZ_n. For instance, EGZ_3 consists of the 4 multisets $\{0,0,0\}$, $\{0,1,2\}$, $\{1,1,1\}$ and $\{2,2,2\}$, and consequently $\mathrm{EGZ}_3(t) = 2 + t + t^3$ (looking at the distribution of 1s or 2s in EGZ_3 would result in the same polynomial). Define the generating functions

$$\mathrm{EGZ}(x,t) = \sum_{n\geq 1} \mathrm{EGZ}_n(t)x^n \quad \text{and} \quad S(x) = \sum_{n\geq 0} |S_n|x^n.$$

Our main result (Theorem 4) is a factorization of the generating function for pointed score sequences:

$$S^\bullet(x,t) = \mathrm{EGZ}(x,t)S(x).$$

We often view a pair $(s,i) \in S_n^\bullet$ as a score sequence with a distinguished position; i.e. $i \in \mathrm{dom}(s)$. Taking the alternative view that i is an element of the codomain of s we find that $S^\bullet(x,0)$ consists of terms stemming from pairs (s,i) such that $s^{-1}(i)$ is empty; i.e. i is outside the image of s. Thus, $S^\bullet(x,1) - S^\bullet(x,0)$ counts pairs (s,i) for which i is in the image of s. Let

$$\begin{aligned} S_n^\circ &= \{(s,i) \in S_n^\bullet : i \in \mathrm{Im}(s)\} \\ &= \{(s,i) \in S_n^\bullet : i = s_j \text{ for some } j \in [n]\} \end{aligned}$$

and let $S^\circ(x) = S^\bullet(x,1) - S^\bullet(x,0)$ be the corresponding generating function. For instance, S_3° consists of the 4 elements $((0,1,2),0)$, $((0,1,2),1)$, $((0,1,2),2)$, and $((1,1,1),1)$. We will show (in Corollary 11) that

$$S^\circ(x) = xC(x)S(x),$$

where $C(x) = (1-\sqrt{1-4x})/(2x)$ is the generating function for the Catalan numbers $C_n = \binom{2n}{n}/(1+n)$. This striking occurrence of the Catalan numbers was in fact the original inspiration for our work. It was in the summer of 2019 that we experimented with score sequences and conjectured the identity. Despite ample attempts we were for the longest time unable to prove it.

These results allow us to present the following recursion (Corollary 12):

$$|S_n| = \frac{1}{n}\sum_{k=1}^{n} |S_{n-k}||\mathrm{EGZ}_k|.$$

In other words, the logarithmic derivative of the generating function $S(x)$ is $\mathrm{EGZ}(x,1)/x$, a fact conjectured by Paul D. Hanna as recorded in the OEIS entry A000571 in 2013. This leads to the closed formula (Corollary 14)

$$|S_n| = \frac{1}{n!}\sum_{\pi\in\mathrm{Sym}(n)} \prod_{\ell\in C(\pi)} |\mathrm{EGZ}_\ell|,$$

where $C(\pi)$ is a partition encoding the cycle type of π. Alternatively,

$$|S_n| = \sum_{1\cdot m_1+2\cdot m_2+\cdots+n\cdot m_n = n} \frac{|\mathrm{EGZ}_1|^{m_1}|\mathrm{EGZ}_2|^{m_2}\cdots|\mathrm{EGZ}_n|^{m_n}}{1^{m_1}m_1!\,2^{m_2}m_2!\,\cdots\,n^{m_n}m_n!},$$

which is Corollary 15.

2 The main theorem and its bijection

Let the generating functions $S^\bullet(x,t)$, $\mathrm{EGZ}(x,t)$ and $S(x)$ be defined as in Section 1.

Theorem 4. *We have*

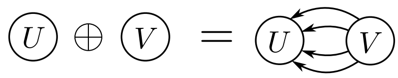
$$S^\bullet(x,t) = \mathrm{EGZ}(x,t)S(x).$$

We shall give a combinatorial proof of Theorem 4 using a bijection

$$\Phi : S_n^\bullet \to \bigcup_{k=1}^{n} \mathrm{EGZ}_k \times S_{n-k}$$

that maps a pointed score sequence to a pair consisting of a multiset and a score sequence. A property of this bijection is that, for $(M,v) = \Phi(s,i)$, the number of occurrences of i in s is equal to the multiplicity of zero in M. Before defining Φ we need to introduce several necessary concepts.

A nonempty directed graph is said to be *strongly connected* if there is a directed path between each pair of vertices of the graph. Note that we do not consider the empty graph to be strongly connected. A *strong score sequence* is one which stems from a strongly connected tournament. Equivalently (see Harary and Moser [5, Theorem 9]), $s = (s_0,\dots,s_{n-1})$, with $n \geq 1$, is a strong score sequence if the inequality (2) of Theorem 1 is always strict; that is, $s_0 + \cdots + s_{k-1} > \binom{k}{2}$ for $1 \leq k < n$. Let us define the *direct sum* of two score sequences $u \in S_k$ and $v \in S_\ell$ by $u \oplus v = uv'$, where v' is obtained from v by adding k to each of its letters and juxtaposition indicates concatenation. For instance, $(0) \oplus (0) \oplus (1,1,1) = (0,1,3,3,3)$. If U and V are tournaments having score sequences u and v, one may view the direct sum $u \oplus v$ as the score sequence of the tournament where arrows are placed between the vertices of U and V such that they all point towards U:



This may easily be seen to be independent of the choice of tournaments.

Lemma 5. *Let $s \in S_n$. If $s_0 + \cdots + s_{k-1} = \binom{k}{2}$ for some $k < n$, then $u = (s_0,\dots,s_{k-1})$ and $v = (s_k - k,\dots,s_{n-1} - k)$ are both score sequences, and $s = u \oplus v$.*

Proof. That u is a score sequence is clear from Landau's theorem. By definition we have $s = u \oplus v$, so it only remains to show that v is a score sequence and we will use Landau's theorem to do so, proving each of the three parts separately:

(1) Since $s_0 + \cdots + s_{k-1} = \binom{k}{2}$ and $s_0 + \cdots + s_k \geq \binom{k+1}{2}$ we have $s_k \geq \binom{k+1}{2} - \binom{k}{2} = k$ and hence $v_0 = s_k - k \geq 0$. Since s is weakly increasing it is clear that v is weakly increasing as well. Moreover, the length of v is $n-k$ and $v_{n-k-1} = s_{n-1} - k \leq n-1-k$.

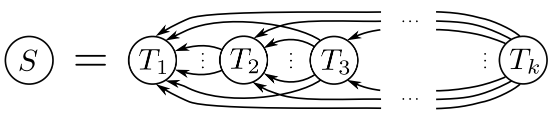
(2) For $1 \leq j < n-k$ we have

$$\begin{aligned} v_0 + v_1 + \cdots + v_{j-1} &= s_k + s_{k+1} + \cdots s_{k+j-1} - jk \\ &= s_0 + \cdots + s_{k+j-1} - (s_0 + \cdots + s_{k-1}) - jk \\ &\geq \binom{k+j}{2} - \binom{k}{2} - jk = \binom{j}{2}. \end{aligned}$$

(3) Similarly, $v_0 + v_1 + \cdots + v_{n-k-1} = \binom{n}{2} - \binom{k}{2} - (n-k)k = \binom{n-k}{2}$.

This concludes the proof. $\square$

A direct consequence of Lemma 5, above, is that every score sequence s can be uniquely written as a direct sum $s = t_1 \oplus t_2 \oplus \cdots \oplus t_k$ of nonempty strong score sequences; in this context, the t_i will be called the *strong summands* of s. In terms of underlying tournaments we have the picture:



We are now almost in a position to define the promised map Φ, but first a couple of definitions. Assume that we are given a score sequence $s = (s_0, s_1, \ldots, s_{n-1}) \in S_n$.

- For any integer j, let $s + j$ denote the sequence obtained by adding j to each element of s, reducing modulo n, and sorting the outcome in nondecreasing order. Note that $s+j$ need not be a score sequence even though s is. E.g. $s = (1,1,1)$ is a score sequence, but $s + 1 = (2,2,2)$ is not. On the other hand, if $s = (0,1,2)$ then $s + 1 = s$ is a score sequence. A characterization of when $s + j$ is a score sequence will be given in Lemma 7.

- Let $\mu(s + j)$ denote the multiset $\{s_0 + j, s_1 + j, \ldots, s_{n-1} + j\}$ with elements in the cyclic group $\mathbb{Z}_n$.

Given a pointed score sequence $(s, i) \in S_n^\bullet$, write $s = t_1 \oplus t_2 \oplus \cdots \oplus t_k$ and let j be the smallest index such that $i < |t_1 \oplus \cdots \oplus t_j|$. Another way to define j is as the smallest prefix $t_1 \oplus \cdots \oplus t_j$ of strong summands of s that begins $s_0, s_1, \ldots, s_i$. Define the two score sequences u and v by

$$u = t_1 \oplus \cdots \oplus t_j \quad \text{and} \quad v = t_{j+1} \oplus \cdots \oplus t_k.$$

Finally, we let

$$\Phi(s,i) = \big(\mu(u-i), v\big).$$

As an example, consider the score sequence $s = (0,2,2,3,3,5,7,7,7)$; its decomposition into strong summands is $s = (0) \oplus (1,1,2,2) \oplus (0) \oplus (1,1,1)$. With $i = 3$ we get $u = (0) \oplus (1,1,2,2) = (0,2,2,3,3)$, $v = (0) \oplus (1,1,1) = (0,2,2,2)$, $u - 3 = (0,0,2,4,4)$ and so $\Phi(s,3) = \big(\{0,0,2,4,4\}, (0,2,2,2)\big)$.

3 Proof of the main result

Our aim is to prove Theorem 4, but first we need to establish a number of lemmas. Let T_n be the set of strong score sequences of length n.

Lemma 6. *For any strong score sequence $s \in T_n$, the n multisets*

$$\mu(s+j) = \{s_0 + j, s_1 + j, \ldots, s_{n-1} + j\}, \quad \textit{for } j \in [0, n-1],$$

are all distinct.

Proof. Assume $j \in [0, n-1]$ is such that $\mu(s+j)$ and $\mu(s)$ are equal as multisets over $\mathbb{Z}_n$. Note that there is no loss of generality here: assuming that $\mu(s+j_1) = \mu(s+j_2)$ is equivalent to assuming that $\mu(s+j) = \mu(s)$ with $j = j_2 - j_1$. Let us write the values $s_0 + j, \ldots, s_{n-1} + j$ in nondecreasing order after reducing modulo n. Since $s_0 \leq \cdots \leq s_{n-1}$ the result must be a cyclic shift of the original order, say, $s_k + j, \ldots, s_{n-1} + j, s_0 + j, \ldots, s_{k-1} + j$ for some index k. As each s_i is less than n and $j < n$ we find that those elements must equal

$$s_k + j - n,\ \ldots,\ s_{n-1} + j - n,\ s_0 + j,\ \ldots,\ s_{k-1} + j. \tag{5}$$

Since we are assuming that $\mu(s+j)$ and $\mu(s)$ are equal as multisets over $\mathbb{Z}_n$, the sum of the elements listed in (5) must equal the sum of the elements $s_0 + s_1 + \cdots + s_{n-1} = \binom{n}{2}$. This implies $nj - (n-k)n = 0$, which gives $j = n - k$. Thus, the sequence (5) becomes

$$s_k - k,\ \ldots,\ s_{n-1} - k,\ s_0 + n - k,\ \ldots,\ s_{k-1} + n - k.$$

By assumption we have $s_0 = s_k - k$, $s_1 = s_{k+1} - k$, and so on; thus

$$\begin{aligned} s_0 + \cdots + s_{n-k-1} &= s_k - k + \cdots + s_{n-1} - k \\ &= \binom{n}{2} - (s_0 + \cdots + s_{k-1}) - (n-k)k \\ &= \binom{k}{2} + \binom{n-k}{2} - (s_0 + \cdots + s_{k-1}). \end{aligned}$$

Suppose $k \geq 1$. Since our score sequence s is strong we have $s_0 + \cdots + s_{k-1} > \binom{k}{2}$, but then $s_0 + \cdots + s_{n-k-1} < \binom{n-k}{2}$ contradicting that s is a score sequence. The only remaining possibility is that $k = 0$ and $j = 0$, which concludes the proof. $\square$

Consider the following example concerning the strong summands of a score sequence $s + j$. Let $s = (1, 1, 2, 2, 4, 6, 7, 7, 7, 8)$ and $j = 5$. We add 5 to each element of s to get the list of numbers 6, 6, 7, 7, 9, 11, 12, 12, 12, 13, which when reduced modulo $|s| = 10$ reads 6, 6, 7, 7, 9, 1, 2, 2, 2, 3; finally we sort this list to arrive at $s + 5 = (1, 2, 2, 2, 3, 6, 6, 7, 7, 9)$. Note that s and $s + 5$ share the same strong summands only arranged differently:

$$\begin{aligned} s &= (1, 1, 2, 2) \oplus (0) \oplus (1, 2, 2, 2, 3); \\ s + 5 &= (1, 2, 2, 2, 3) \oplus (1, 1, 2, 2) \oplus (0). \end{aligned}$$

As is detailed in the following lemma, this is not a coincidence. See also Fig. 1 where we show how this may be interpreted in terms of tournaments.

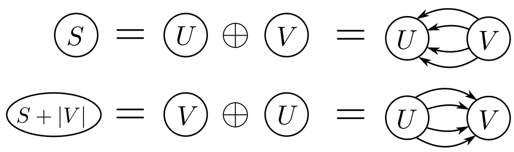


Figure 1: If $s = u \oplus v$, one may view $s + |v|$ as the score sequence of a tournament obtained by flipping the direction of all arrows pointing out of V towards U. Here, S, U, V are tournaments with score sequences s, u, v respectively.

Lemma 7. *Let s be any nonempty score sequence of length n and let $j \in [1, n-1]$. Then $s+j$ is a score sequence if and only if there are score sequences u and v with $|v| = j$, such that $s = u \oplus v$. In that case, $s + j = v \oplus u$.*

Proof. Let $s = (s_0, s_1, \ldots, s_{n-1})$. Define the sequences

$$\begin{aligned} u &= (s_0, s_1, \ldots, s_{n-j-1}); \\ v &= (s_{n-j} + j - n, s_{n-j+1} + j - n, \ldots, s_{n-1} + j - n) \end{aligned}$$

of length $n - j$ and j, respectively. Assume that u and v are score sequences. Clearly, $s = u \oplus v$. Since the length of v is j we find that

$$v \oplus u = (s_{n-j} + j - n, \ldots, s_{n-1} + j - n, s_0 + j, s_1 + j, \ldots, s_{n-j-1} + j).$$

If we consider the elements of this sequence modulo n we may add n to the first j of them. If we then sort the elements in nondecreasing order we obtain the sequence $(s_0 + j, \ldots, s_{n-1} + j)$. Thus, $s + j = v \oplus u$, which is a score sequence. For the other direction, assume that $s + j$ is a score sequence. Then by definition of $s + j$ there is an $\ell \in [0, n - 1]$ such that $s + j$ is

$$(s_\ell + j - n, s_{\ell+1} + j - n, \ldots, s_{n-1} + j - n, s_0 + j, s_1 + j, \ldots, s_{\ell-1} + j).$$

Since $s+j$ is a score sequence of length n, we find by item (3) in Landau's theorem that

$$\begin{aligned}\binom{n}{2} &= \sum_{i=\ell}^{n-1}(s_i+j-n)+\sum_{i=0}^{\ell-1}(s_i+j)\\ &= \sum_{i=0}^{n-1} s_i+(n-\ell)(j-n)+\ell j \;=\binom{n}{2}+n(\ell-n+j).\end{aligned}$$

Thus, $\ell = n-j$. Now consider the sum of the first j terms in $s+j$; let us call this sum K. By Landau's theorem we have

$$\begin{aligned}\binom{j}{2} \le K &= \sum_{i=n-j}^{n-1}(s_i+j-n)\\ &= \sum_{i=n-j}^{n-1} s_i + j(j-n)\\ &= \binom{n}{2} - \sum_{i=0}^{n-j-1} s_i + j(j-n)\\ &\le \binom{n}{2} - \binom{n-j}{2} + j(j-n) \;=\; \binom{j}{2}.\end{aligned}$$

In particular, the two inequalities above are in fact equalities and $K=\binom{j}{2}$. From Lemma 5 we deduce that u and v as previously defined are score sequences and that $s+j = v \oplus u$. As before, it is clear that $s = u \oplus v$. $\square$

The previous lemma may be equivalently stated as follows:

Lemma 8. *Let $s = t_1 \oplus t_2 \oplus \cdots \oplus t_k$ be any nonempty score sequence decomposed into its strong summands, and let $j \in [1, n-1]$. Then $s+j$ is a score sequence if and only if there is an ℓ such that*

$$j = |t_\ell| + |t_{\ell+1}| + \cdots + |t_k|$$

in which case $s+j = t_\ell \oplus t_{\ell+1} \oplus \cdots \oplus t_k \oplus t_1 \oplus \cdots \oplus t_{\ell-1}$.

Our next lemma will be essential for proving that the mapping f defined in Lemma 10, below, is surjective.

Lemma 9. *For all multisets $M \in \mathrm{EGZ}_n$ there is a score sequence $s \in S_n$ and a constant $j \in [0, n-1]$ such that $\mu(s+j) = M$.*

Proof. Let $M \in \mathrm{EGZ}_n$. By definition of EGZ_n, the sum of the members of M is $\binom{n}{2}$ modulo n. We start by showing that for some integer j the members of the multiset $M+j = \{x+j : x \in M\}$ will have sum exactly $\binom{n}{2}$ (without reducing the sum modulo n). Suppose that the sum of the members of M is greater than $\binom{n}{2}$. We will show that we can always choose a j such that $M+j$ has a smaller sum. For $k \in [0, n-1]$, let y_k be the number of element $x \in M$ such that $x \leq k$. Suppose that $y_k > k$ for all k. Then $y_0 \geq 1$ and there is at least one zero in M. Similarly, $y_1 \geq 2$ so in addition to that zero there is a value equal to at most one. Continuing like this we bound our values from above by $0, 1, \ldots, n-1$. The sum of these values is, however, at most $\binom{n}{2}$. Since we assumed that the sum is greater than $\binom{n}{2}$ we conclude that the assertion that $y_k > k$, for all k, is false. For the remainder of the argument, let k be the smallest index such that $y_k \leq k$.

If $k = 0$, then there are no zeroes in M, so subtracting one from all the elements in M simply causes the sum to drop by n since no modulo reductions occur. Thus, we can assume $k > 0$. Then $k-1 < y_{k-1} \leq y_k \leq k$ and hence $y_k = k$. We now subtract $k+1$ from all the elements in M. Exactly y_k of these values will be below zero. So we decrease the sum by $n(k+1)$ but reducing modulo n adds kn to the sum. The combined effect is to decrease the sum by n. In this way we can always decrease the sum down to $\binom{n}{2}$. The proof that the sum be can be increased in the same way is nearly identical.

Let $M = \{x_1, \ldots, x_n\} \in \mathrm{EGZ}_n$ with $x_1 \leq \cdots \leq x_n$. In light of the last two paragraphs, we can assume that $x_1 + \cdots + x_n = \binom{n}{2}$. The sequence $(x_1, \ldots, x_n)$ satisfies items (1) and (3) of Landau's theorem. To prove that $(x_1, \ldots, x_n)$ is a score sequence it remains to prove item (2), namely that $x_1 + \cdots + x_k \geq \binom{k}{2}$ for $k \in [1, n-1]$. Suppose that $x_1 + \cdots + x_k < \binom{k}{2}$ for some k in $[1, n-1]$. For now, let k be the largest such k. Since k is maximal we have $x_1 + \cdots + x_{k+1} \geq \binom{k+1}{2}$, which gives us $x_{k+1} \geq \binom{k+1}{2} - x_1 - \cdots - x_k > \binom{k+1}{2} - \binom{k}{2} = k$. If $x_k \geq k$ then we can decrement k until this no longer holds true; note that $x_1 + \cdots + x_{k-1} < \binom{k-1}{2}$ will continue to hold while $x_k \geq k$ and $x_1 + \cdots + x_k < \binom{k}{2}$. We eventually arrive at a k such that $x_1 + \cdots + x_k < \binom{k}{2}$, $x_k < k$ and $x_{k+1} \geq k$. The values $x_1, \ldots, x_k$ are thus precisely the values among $x_1, \ldots, x_n$ that are smaller than k. Define y_i as before and let $w_i = y_i - i - 1$. The number of values equal to r among $x_1, \ldots, x_k$ is $y_r - y_{r-1}$ as long as $r < k$, and so

$$x_1 + \cdots + x_k = 0 \cdot y_0 + 1 \cdot (y_1 - y_0) + \cdots + (k-1)(y_{k-1} - y_{k-2}).$$

This implies $(k-1)y_{k-1} - y_{k-2} - \cdots - y_0 = x_1 + \cdots + x_k < \binom{k}{2}$, but we know that $y_{k-1} = k$, so we must have $k^2 - \binom{k}{2} < y_{k-1} + \cdots + y_0$. Rewriting in terms of w_k we get $\binom{k+1}{2} < w_0 + \cdots + w_{k-1} + 1 + 2 + \cdots + k$, which is equivalent to $w_0 + \cdots + w_{k-1} > 0$. Thus, if we can choose a shift (i.e. choose a constant to add to the elements of M) such that the prefix sums of the w_i are non-positive, then the original set must be a score sequence.

Consider the values k such that $w_k = 0$ or, equivalently, $y_k = k+1$. If we subtract $k+1$ from every value in our multiset, then exactly $k+1$ of them will become negative. Thus, our sum decreases by $n(k+1)$ but the modulo reduction adds $n(k+1)$ back to the sum, so it remains unchanged. Let us consider what this does to our sequences (y_i) and (w_i). Denote the new sequences after the subtraction by (y'_i) and (w'_i). Consider the value of y'_i. It counts the number of values from 0 to i after the subtraction, which are values from $k+1$ to $k+1+i$ modulo n before the subtraction. We consider two cases.

(a) If $k+1+i < n$, then $y'_i = y_{k+1+i} - y_k$. Since $y_k = k+1$ this gives us in terms of w_i that $w'_i + i + 1 = w_{k+i+1} + k + i + 1 + 1 - k - 1$, which simplifies to $w'_i = w_{k+i+1}$.

(b) If $k+1+i \geq n$, then $y'_i = y_{n-1} - y_k + y_{k+1+i-n}$. We already know that $y_k = k+1$ and $y_{n-1} = n$. In terms of the w_i we have $w'_i + i + 1 = n - k - 1 + w_{k+1+i-n} + k + 1 + i - n + 1$, which reduces to $w'_i = w_{k+1+i-n}$.

Hence, in both cases we have $w'_i = w_{k+i+1}$ when indices are considered modulo n. We can thus cyclically permute the sequence (w_i) as long as the final value continues to be 0 and still have it correspond to a shift of our original multiset M such that the shift has sum $\binom{n}{2}$.

By the same argument as above (when $k = n$) we have

$$x_1 + \cdots + x_n = (n-1)y_{n-1} - y_{n-1} - y_{n-2} - \cdots - y_0.$$

Furthermore, $y_{n-1} = n$ and $x_1 + \cdots + x_n = \binom{n}{2}$. Thus, $y_0 + \cdots + y_{n-1} = \binom{n}{2}$, which gives us $w_0 + \cdots + w_{n-1} = 0$. Hence, we can consider the sequence of sums of values between the zeroes that appear in (w_i). This new sequence (g_i) also has sum zero and we can cyclically permute the sequence (w_i) in the way we want if and only if we can permute the sequence (g_i) to satisfy the same desired property. This is because either all values between two zeroes in (w_i) are of the same sign, or the positive values all come after the negative ones. The reason for this is that one cannot go from positive values to negative without intercepting zero since we decrease by at most one at a time. A sequence of integers with sum zero can be cyclically permuted to produce a sequence with non-positive prefix sums according to Raney's lemma [4]. This allows us to conclude that the prefix sums, as discussed above, are non-positive and so the original set must be a score sequence, which completes our proof. $\square$

Let $T(x)$ be the generating function for the number of strong score sequences according to length. Note that any nonempty score sequence s can be written $s = u \oplus t$, where u is a score sequence and t is the last strong summand of s, and thus $S(x) = 1 + S(x)T(x)$, or, equivalently, $S(x) = (1 - T(x))^{-1}$. Let

$$T_n^{\bullet} = T_n \times [0, n-1]$$

be the set of *pointed strong score sequences* of length n, and let

$$T_n^\bullet(t) = \sum_{(s,i)\in T_n^\bullet} t^{|s^{-1}(i)|} \quad \text{and} \quad T^\bullet(x,t) = \sum_{n\geq 1} T_n^\bullet(t)x^n.$$

Lemma 10. *Assume $n \geq 1$. Select those pointed score sequences (s,i) in $S_n^\bullet$ whose distinguished position is in the last strong summand, and denote the resulting set L_n. That is,*

$$L_n = \big\{(s,i) \in S_n^\bullet : s = t_1 \oplus \cdots \oplus t_k \text{ and } i \in [n - |t_k|, n-1]\big\}.$$

Then the mapping $f : L_n \to \mathrm{EGZ}_n$ defined by

$$f(s,i) = \mu(s-i)$$

is a bijection. Moreover, if $M = f(s,i)$, then $|M|_0 = |s^{-1}(i)|$. In terms of generating functions we have

$$\mathrm{EGZ}(x,t) = S(x)T^\bullet(x,t).$$

Proof. We start by proving that f is injective. To that end, assume that (u,i) and (v,j) in L_n are such that $f(u,i) = f(v,j)$; that is,

$$\mu(u-i) = \mu(v-j) = M, \quad \text{where } M \in \mathrm{EGZ}_n.$$

If $u = v$, then $i = j$ by Lemma 6, so we may assume that u and v are different score sequences. Without loss of generality we may further assume that $i \leq j$. Now, $\mu(u-i) = \mu(v-j)$ holds by assumption, and this multiset identity is true if and only if $\mu(u) = \mu(v + (j-i))$. Thus, $u = v + m$ with $m = j - i$. Assume that the decomposition of v into its strong summands is $v = t_1 \oplus \cdots \oplus t_k$. By Lemma 8, the score sequences u and v contain the same strong summands up to a cyclic permutation. To be more precise, there is an $\ell \geq 1$ such that

$$m = |t_\ell| + \cdots + |t_k| \quad \text{and} \quad u = t_{\ell+1} \oplus \cdots \oplus t_k \oplus t_1 \oplus \cdots \oplus t_\ell.$$

By definition of the set L_n we have $i \in [n-|t_\ell|, n-1]$ and $j \in [n-|t_k|, n-1]$. Therefore $m = j - i \leq n - 1 - (n - |t_\ell|) = |t_\ell| - 1$. This would however imply $|t_\ell| + \cdots + |t_k| < |t_\ell|$, which is impossible. Therefore no such i and j can exist and the mapping f is injective.

Next we prove surjectivity. We want to prove that every $M \in \mathrm{EGZ}_n$ is the image of some score sequence $s \in S_n$ with a distinguished element i in its last strong summand, i.e. $(s,i) \in L_n$. By Lemma 9 every $M \in \mathrm{EGZ}_n$ is the image of some such (s,j) where we place no restriction on j. If j is smaller than the size of the last strong summand of s, then f maps the pointed score sequence $(s, n-j)$ to our multiset M. If j is greater than the size of the last strong summand of s, then we can move the last summand to the front and decrease j by the corresponding size by Lemma 8. In this way we will eventually end up in the first case. Thus, we end up with a score sequence that maps to our multiset M.

We have shown that f is a bijection. Assume that $(s, i) \in L_n$ and $M = f(s, i)$. By definition of f it is clear that the number of zeros in M corresponds to the number of j for which $s_j = i$, and hence $|M|_0 = |s^{-1}(i)|$. Every $(s, i) \in L_n$ can be uniquely identified with a pair $(s', (t, j))$, where $s = s' \oplus t$, t is the last strong summand of s, and $j = i - |s'|$. In other words, s' is the score sequence s with its last strong summand t removed, and j is the distinguished position relative to t rather than s. Assuming $k = |s'|$, we have $s' \in S_k$ and $(t, j) \in T^{\bullet}_{n-k}$ with $|t^{-1}(j)| = |M|_0$. Consequently, $\mathrm{EGZ}(x, t) = S(x)T^{\bullet}(x, t)$ as claimed. □

We are now finally in a position to prove our main result.

Proof of Theorem 4. We will give a combinatorial proof of the power series identity $S^{\bullet}(x, t) = \mathrm{EGZ}(x, t)S(x)$ by showing that the mapping

$$\Phi : S^{\bullet}_n \to \bigcup_{k=1}^{n} \mathrm{EGZ}_k \times S_{n-k}$$

defined in Section 2 is a bijection. Recall that $\Phi(s, i) = (\mu(u - i), v)$ is defined by writing s in terms of its strong summands, say $s = t_1 \oplus t_2 \oplus \cdots \oplus t_k$, and then splitting the score sequence $s = u \oplus v$, where $u = t_1 \oplus \cdots \oplus t_j$ and j is the smallest index such that $i < |t_1 \oplus \cdots \oplus t_j|$, and v is the remainder. Let L_n be the set defined in Lemma 10. Note that $(u, i) \in L_{|u|}$. In fact, it is easy to see that $(s, i) \mapsto \big((u, i), v\big)$ is a bijection from $S^{\bullet}_n$ to $\cup_{k=1}^{n} L_k \times S_{n-k}$. Moreover, by Lemma 10, $(u, i) \mapsto \mu(u - i)$ is a bijection from L_k to EGZ_k, and hence Φ is a bijection as well. Finally, $(u, i) \mapsto \mu(u - i)$ has the property that $|\mu(u - i)|_0 = |u^{-1}(i)|$ and so $S^{\bullet}(x, t) = \mathrm{EGZ}(x, t)S(x)$, as claimed. □

Corollary 11. *We have*

$$S^{\circ}(x) = xC(x)S(x),$$

where $C(x)$ is the generating function for the Catalan numbers.

Proof. From Theorem 4 we get

$$S^{\circ}(x) = S^{\bullet}(x, 1) - S^{\bullet}(x, 0) = \big(\mathrm{EGZ}(x, 1) - \mathrm{EGZ}(x, 0)\big)S(x).$$

It thus suffices to show that $\mathrm{EGZ}(x, 1) - \mathrm{EGZ}(x, 0) = xC(x)$. Note that the coefficient of x^n in $\mathrm{EGZ}(x, 1) - \mathrm{EGZ}(x, 0)$ is the number of multisets of EGZ_n that contain at least one zero. On removing a single zero from such a multiset we are left with a multiset of size $n - 1$ whose sum is $\binom{n}{2}$ modulo n. There are C_{n-1} such multisets [12, Exercise 6.19jjj]. □

4 The number of score sequences

Since $S_n^\bullet = S_n \times [0, n-1]$ there is a simple relation $|S_n^\bullet| = n|S_n|$ between the number of pointed score sequences and the number of (plain) score sequences. The generating function for $|S_n^\bullet|$ is $S^\bullet(x,1)$ and letting $t = 1$ in the main theorem we have $S^\bullet(x,1) = \mathrm{EGZ}(x,1)S(x)$. From this and Equation 4 in Section 1 the corollary below immediately follows.

Corollary 12. *For $n \geq 1$,*

$$\begin{aligned} |S_n| &= \frac{1}{n}\sum_{k=1}^{n} |S_{n-k}||\mathrm{EGZ}_k| \\ &= \frac{1}{n}\sum_{k=1}^{n} |S_{n-k}|\frac{1}{2k}\sum_{d|k}(-1)^{k-d}\varphi(k/d)\binom{2d}{d}. \end{aligned}$$

Using dynamic programming and the sieve of Eratosthenes this allows us to calculate all values of $|S_k|$ for $k \leq n$ in $\mathcal{O}(n^2)$ time, assuming constant time integer operations. This is an improvement on earlier results by Narayana and Bent [9]. Their recursive formula can be implemented to find $|S_n|$ in $\mathcal{O}(n^3)$, but no faster since their recursive function must always visit $\mathcal{O}(n^3)$ states to do so; to get all $|S_k|$ for $k \leq n$ takes $\mathcal{O}(n^4)$ due to lack of overlap in the states recursively visited for different k.

Since $S(x) = (1 - T(x))^{-1}$ this can be extended to $|T_k|$. We can rewrite the equation as $T(x) = S(x) - (S(x) - 1)T(x)$, which gives us the recursion

$$|T_n| = |S_n| - \sum_{i=1}^{n-1} |T_i||S_{n-i}|.$$

We first calculate the values $|S_k|$ and use this recursion to calculate all the values $|T_k|$ for $k \leq n$ in $\mathcal{O}(n^2)$. This is the same method as used by Stockmeyer [13], just calculating the underlying $|S_k|$ faster which brings the total time complexity down from $\mathcal{O}(n^4)$ to $\mathcal{O}(n^2)$.

We shall derive a closed formula for $|S_n|$ from Corollary 12, but first we generalize the setting a little. Considering a power series $A(x) = \sum_{n\geq 0} a_n x^n$. The coefficient of x^n in $x\frac{d}{dx}A(x)$ is na_n. For this reason $x\frac{d}{dx}$ is sometimes called a pointing operator. The logarithmic derivative of $A(x)$ is $A'(x)/A(x)$. In the same vein, let the *logarithmic pointing* of $A(x)$ be

$$xA'(x)/A(x) = x\frac{d}{dx}\big(\log A(x)\big).$$

Lemma 13. *Let a power series $A(x) = \sum_{n\geq 0} a_n x^n$ with $a_0 = 1$ be given. Let $B(x) = \sum_{n\geq 1} b_n x^n$ be the logarithmic pointing of $A(x)$. Equivalently,* *the sequence of numbers* $\{b_n\}_{n\geq 0}$ *is implicitly given by* $b_0 = 0$ *and*

$$a_n = \frac{1}{n}\sum_{k=1}^{n} a_{n-k} b_k \qquad \text{for } n \geq 1.$$

Then

$$a_n = \frac{1}{n!}\sum_{\pi \in \mathrm{Sym}(n)} \prod_{\ell \in C(\pi)} b_\ell, \tag{6}$$

where $\mathrm{Sym}(n)$ *the symmetric group of degree* n *and* $C(\pi)$ *encodes the cycle type of* π*; i.e. there is an* $\ell \in C(\pi)$ *for each* ℓ*-cycle of* π.

Proof. For $n = 0$ there is a single member of $\mathrm{Sym}(n)$, namely the empty permutation. It has no cycles and (6) says $a_0 = 1$ as expected. We proceed by induction. Assume that $n \geq 1$ and consider the disjoint cycle structure of a permutation $\pi \in \mathrm{Sym}(n)$. Let α be the cycle of π that n belongs to, and let $k = |\alpha|$ be the length of that cycle. Then $\pi\alpha^{-1}$ fixes the elements of α and we can consider it a permutation of the remaining $n-k$ elements. Let π' be the "standardization" of $\pi\alpha^{-1}$; i.e. we replace the smallest element of $\pi\alpha^{-1}$ with 1, the next smallest with 2, etc. This defines a mapping $\pi \mapsto \pi'$ from $\mathrm{Sym}(n)$ to $\mathrm{Sym}(n-k)$; e.g. if $\pi = (1\,6\,2)(3\,5)(4\,7)$, then $\alpha = (4\,7)$, $\pi\alpha^{-1} = (1\,6\,2)(3\,5)$, and $\pi' = (1\,5\,2)(3\,4)$. How many permutations π map to the same π'? As many as there are choices for the cycle α, namely $\binom{n-1}{k-1}(k-1)! = (n-1)!/(n-k)!$; let us call this number d. Thus, $\pi \mapsto \pi'$ is a d-to-1 mapping, and in terms of equations:

$$\begin{aligned}
\frac{1}{n!}\sum_{\pi \in \mathrm{Sym}(n)} \prod_{\ell \in C(\pi)} b_\ell &= \frac{1}{n!}\sum_{k=1}^{n} \frac{(n-1)!}{(n-k)!} b_k \sum_{\pi' \in \mathrm{Sym}(n-k)} \prod_{\ell \in C(\pi')} b_\ell \\
&= \frac{1}{n}\sum_{k=1}^{n} \frac{1}{(n-k)!} b_k (n-k)! a_{n-k} \\
&= \frac{1}{n}\sum_{k=1}^{n} b_k a_{n-k} \\
&= a_n,
\end{aligned}$$

where in the second equality we used the induction hypothesis for a_{n-k}. □

Since the logarithmic pointing of $S(x)$ is $\mathrm{EGZ}(x,1)$ the lemma above applies and we have a closed form for the numbers $|S_n|$:

Corollary 14. *For* $n \geq 1$,

$$|S_n| = \frac{1}{n!}\sum_{\pi \in \mathrm{Sym}(n)} \prod_{\ell \in C(\pi)} |\mathrm{EGZ}_\ell|.$$

By collecting terms with equal cycle type we arrive at the following alternative formula.

Corollary 15. *For $n \geq 1$,*

$$|S_n| = \sum_{1 \cdot m_1 + 2 \cdot m_2 + \cdots + n \cdot m_n = n} \frac{|\mathrm{EGZ}_1|^{m_1} |\mathrm{EGZ}_2|^{m_2} \cdots |\mathrm{EGZ}_n|^{m_n}}{1^{m_1} m_1! \, 2^{m_2} m_2! \, \cdots \, n^{m_n} m_n!}.$$

References

[1] Max Alekseyev. https://oeis.org/A145855/a145855.txt, 2008.

[2] Shane Chern. An extension of a formula of Jovovic. *Integers*, 19:Paper No. A47, 2019.

[3] P. Erdős, A. Ginzburg, and A. Ziv. Theorem in the additive number theory. *Bull. Res. Council Israel Sect. F*, 10F(1):41–43, 1961.

[4] R.L. Graham, D.E. Knuth, and O. Patashnik. *Concrete Mathematics: A Foundation for Computer Science.* Addison-Wesley, 1994.

[5] Frank Harary and Leo Moser. The theory of round robin tournaments. *Amer. Math. Monthly*, 73:231–246, 1966.

[6] OEIS Foundation Inc. The Online Encyclopedia of Integer Sequences. `https://oeis.org`, 2021.

[7] H. G. Landau. On dominance relations and the structure of animal societies: III. The condition for a score structure. *Bull. Math. Biophys.*, 15:143–148, 1953.

[8] P. A. MacMahon. An American tournament treated by the calculus of symmetric functions. *Quart. J. Math.*, 49:1–36, 1920. Reprinted in *Percy Alexander MacMahon Collected Papers, Volume I*, George E. Andrews, ed., MIT Press, 1978.

[9] T. V. Narayana and D. H. Bent. Computation of the number of score sequences in round-robin tournaments. *Can. Math. Bull*, 7(1):133–136, 1964.

[10] John Riordan. The number of score sequences in tournaments. *J. Combinatorial Theory*, 5:87–89, 1968.

[11] John Riordan. Erratum: “The number of score sequences in tournaments”. *J. Combinatorial Theory*, 6:226, 1969.

[12] Richard P. Stanley. *Enumerative Combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and Appendix 1 by Sergey Fomin.

[13] Paul K. Stockmeyer. Counting various classes of tournament score sequences. *arXiv:2202.05238v1*, 2022.

Anders Claesson, University of Iceland, akc@hi.is

Mark Dukes, University College Dublin, mark.dukes@ucd.ie

Atli Fannar Franklín, University of Iceland aff6@hi.is

Sigurður Örn Stefánsson, University of Iceland, sigurdur@hi.is