¿

Some Deep and Original Questions about the "critical exponents" of Generalized Ballot Sequences

Shalosh B. EKHAD and Doron ZEILBERGER

Abstract: We numerically estimate the critical exponents of certain enumeration sequences that naturally generalize the famous Catalan and super-Catalan sequences, and raise deep and original questions about their exact values, and whether they are rational numbers.

Added April 7, 2021: Michael Wallner answered the questions raised in this article (in the 3D case). Once he writes it up, we will reference it.

Added May 27, 2021: Michael Wallner just posted his beautiful article

https://arxiv.org/abs/2105.12155

The promised donation to the OEIS in his honor has been made. Wallner also pointed out some minor typos. For the sake of historical accuracy, we did not correct them, and this second version of our paper is the same as the original one, after these remarks.

Preface

Some of the deepest, and most original, questions in enumeration are those of finding the so-called *critical exponents* of hard-to-compute sequences, most famously sequences enumerating self-avoiding walks. Physicists believe that the critical exponent is even more interesting than the so-called connective constant (since it is believed to be 'universal' and tells you 'how water boils', as opposed to 'what is the boiling temperature?', the latter depending on many contingent factors).

An important kind of sequences that occur a lot in enumerative combinatorics is the class of **P-recursive sequences** for which it is easy to derive exactly both connective constants (see below, the limit of the ratio of consecutive terms) and **critical exponents** (see below). In this deep and original article, we come up with many interesting and natural enumerative sequences, generalizing the Catalan numbers and their higher-dimensional counterparts (ballot numbers), that most probably are **not** P-recursive, yet seem to have nice asymptotics, similar to those of P-recursive sequences (and self-avoiding walks). We computed quite a few of them and estimated their critical exponents, but we can't even conjecture exact values, and whether or not they are rational. We are pledging donations to the OEIS in honor of the first identifier (and prover) of these critical exponents.

Maple Package

This article is accompanied by the Maple package `Capone.txt`, available from,

`https://sites.math.rutgers.edu/~zeilberg/tokhniot/Capone.txt` .

The front of this article

`https://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/capone.html` ,

contains numerous input and output files, briefly described in this article.

Ballot Sequences

One of the almost $\aleph_0$ combinatorial objects counted by the *Catalan numbers*, $\mathcal{C}_2(n) := \frac{(2n)!}{n!(n+1)!}$, are *ballot sequences*, that count the number of ways of counting votes of two candidates that each got n votes, in such a way that one of them was never behind in the partial count, as first proved by Bertrand [B], and given the *proof from the book* by André [A]. It is one of the *fundamental objects* listed in Stanley's fascinating book ([St], Theorem 1.5.1 (iv)).

But what if there are k candidates? What is the number of ways of counting the votes of k candidates, each of whom got at the end n votes, in such a way that the second candidate was never ahead of the first candidate, the third was never ahead of the second one, and in general, for $i = 1 \dots k-1$, Candidate $i+1$ was never ahead of Candidate i? Equivalently:

"How many ways are there of walking in the k-dimensional Manhattan lattice, with unit steps in each direction, from the origin $(0, \dots, 0)$ to the point $(n, \dots, n)$ always staying in the region

$$\{(x_1, \dots, x_k) \,|\, x_1 \geq x_2 \geq \dots \geq x_{k-1} \geq x_k\}?''$$

This number is famously

$$C_k(n) := \frac{(n\,k)!\,0!\,1!\,\cdots\,(k-1)!}{(n+k-1)!\,(n+k-2)!\,\dots\,n!} \quad .$$

Note that when $k = 2$, one gets the Catalan numbers.

The formula for $C_k(n)$ follows immediately from the so-called *Young-Frobenius* formula, equivalent to the *hook-length formula* (see [Kn], p. 58, Eq. (34)). The latter formula tells you that the number of ways of walking from the origin to $(a_1, \dots, a_k)$ (when $a_1 \geq a_2 \geq \dots \geq a_k$) **always** staying in $\{(x_1, \dots, x_k) \,|\, x_1 \geq x_2 \geq \dots \geq x_{k-1} \geq x_k\}$ is

$$\Delta(a_1+k-1, a_2+k-2, \dots, a_k) \cdot \frac{(a_1 + \dots + a_k)!}{(a_1+k-1)!\,(a_2+k-2)!\,\cdots\,a_k!} \quad ,$$

where $\Delta(x_1, \dots, x_k)$ is the *discriminant* function

$$\Delta(x_1, \dots, x_k) := \prod_{1 \leq i < j \leq k} (x_i - x_j) \quad .$$

Substituting $a_1 = n, a_2 = n, \dots, a_k = n$ immediately gives the above expression for $C_k(n)$.

What's nice about the Catalan numbers and their higher dimensional analogs is that they belong to a very important class of sequences, called P**-recursive sequences**.

A sequence $x(n)$ is P-recursive if there exists a **finite** positive integer L, called its **order**, and **polynomials** $p_i(n)$, for $i = 0, 1, \ldots, L$, such that

$$p_0(n)x(n) \,+\, p_1(n)x(n+1) \,+\, \ldots \,+\, p_L(n)x(n+L) \,=\, 0 \quad .$$

It so happens that for the Catalan sequence $C(n) = C_2(n)$, and even the super-Catalan sequences $C_k(n)$, the order L is 1. Such sequences are called **hypergeometric sequences**. But from a computational point of view, for *any* such sequence, once you have found a recurrence satisfied by it, you can compute the first 10000 (or whatever) terms very fast.

It is easy to see, using elementary linear algebra, that if an integer sequence is P-recursive, then the coefficients of the polynomials $p_i(n)$ above are all integers. This observation will be important later on.

It is easy to see ([KaP]) that a sequence $x(n)$ is P-recursive if and only if its **generating function**, alias **z-transform**,

$$X(z) \,:=\, \sum_{n=0}^{\infty} x(n)\, z^n \quad ,$$

satisfies a **linear differential equation with polynomial coefficients**. The sequence $\{x(n)\}$ is called **algebraic** if its generating function $X(z)$ satisfies an **algebraic equation**, i.e., for some positive integer M, called the **degree**, we have

$$q_0(z) \,+\, q_1(z)X(z) + q_2(z)X(z)^2 \,+ \ldots + q_M(z)X(z)^M \,=\, 0 \quad ,$$

for some **polynomials** of z, $q_0(z), \ldots, q_M(z)$. Once again it is easy to see that for integer sequences that are algebraic, the coefficients of the $q_i(z)$ can be taken to be integers.

If the degree of the defining equation, M, is 1, then the generating function is **rational** and the sequence belongs to the important subfamily of C-**finite sequences** (see [KaP], chapter 4).

The Catalan sequence $C(n)$ is algebraic since it famously satisfies the equation

$$z\, X(z)^2 \,-\, X(z) \,+\, 1 \,=\, 0 \quad .$$

Perhaps surprisingly, the higher-dimensional counterparts, $C_k(n)$, for $k \geq 3$, are **not** algebraic.

A classical theorem (that may be due to Comtet, see [KaP]) states that if a sequence has an algebraic generating function, then it is P-recursive, but of course, not vice-versa, witness the fact that $C_k(n)$ are **not** algebraic for $k \geq 3$.

Using standard techniques it is easy to prove the following fact.

Important Fact: Every P-recursive integer sequences, $x(n)$, has nice asymptotic expressions and there exist algorithms for finding it. For most 'natural' sequences the asymptotics is of the form

$$x(n) \asymp C \mu^n n^{\theta} \quad ,$$

where C is a constant, and, borrowing the terminology of mathematical physics, μ is called the **connective constant** and θ is called the **critical exponent**.

The Really Deep Sequences: Sequences enumerating Self-Avoiding Walks

Tony Guttmann, Iwan Jensen, and their collaborators used sophisticated (alas, still *exponential-time*) algorithms to crank out many terms of the sequences of self-avoiding walks in the 2-dimensional, triangular, square, and honeycomb lattices. Just search the OEIS for "Self-avoiding walks". The only rigorously proved connective constant is the one for the number of self-avoiding walks on the honeycomb lattice, proved to be $\sqrt{2+\sqrt{2}}$ in [DS] confirming a previous non-rigorous, but very convincing, 'physical derivation' by B. Nienhuis [N].

The **exact value** of the critical exponent for **all** (hence the *universality*) two-dimensional lattices (the triangular, cubic, and honeycomb) is **conjectured** to be $\frac{11}{32}$ (the physicists, for their own reasons, add 1, so for them it is $\frac{43}{32}$), but this is **wide open**, and worthy of a Fields medal.

Generalized Ballot Sequences in 2 dimensions

The general question is to enumerate the number of walks, with positive unit steps, in the two-dimensional square lattice from the origin to the point (an, bn) always staying in the region $b\,x_1 - a\,x_2 \geq 0$. Of course a and b are relatively prime. The special cases $a = 1$, general b (and $b = 1$, general a) give the **Fuss-Catalan numbers**, and were considered by R.C. Lyness (of "Lyness cycle" fame, one of the greatest high school teachers of all time) [L]. It is known (e.g. [BKK] [Ek] [TZ]) that the generating function is **always** algebraic, and there are efficient Maple implementations in [EK] and [TZ] for finding them. The connective constant is easily seen to be

$$\frac{(a+b)^{a+b}}{a^a\, b^b} \quad .$$

It turns out that the critical exponent is **always exactly** $-\frac{3}{2}$, and it is not hard to prove this using standard techniques.

To get the first 100 terms, as well as precise asymptotic *estimates* for **all** such sequences for $1 \leq a < b \leq 10$ and relatively prime a, and b, see the output file

`http://www.math.rutgers.edu/~zeilberg/tokhniot/oCapone1.txt` .

Some of these sequences (for small a and b) are in the OEIS.

The estimated critical exponents, using our **numerical method**, agree to (at least) ten decimal digits with the rigorously derived value of $-\frac{3}{2}$, giving us confidence at the estimates for the higher dimensional generalized ballot sequences, to be discussed next, for which, at present, there is **no** rigorous way (as far as we know).

Generalized Ballot Sequences in 3 dimensions

Things are starting to get interesting in three dimensions. To our surprise, **none** of these sequences were in the OEIS (viewed April 4, 2021).

The output file

`http://www.math.rutgers.edu/~zeilberg/tokhniot/oCapone2.txt` ,

contains enumerating sequences for the number of ways of walking from $(0,0,0)$ to $(a\,n\,,\,b\,n\,,\,c\,n)$ such that if $M = lcm(a,b,c)$, it stays in the region $(M/a)\,x \geq (M/b)\,y \geq (M/c)\,z \geq 0$ for all $1 \leq c \leq b \leq a \leq 4$ such that $gcd(a,b,c) = 1$.

We believe that except for the classical case $a = b = c = 1$, these sequences are **not** P-recursive, and hence we have no clue how to derive, rigorously, the critical exponents. We don't even know whether they are rational numbers, all we can do is crank out many terms and use **numerics**.

Here are numerical estimates for (a,b,c), as above, taken from the end of the above output file.

- (1,1,1) [the classical (proved!) case]: -4 .
- (2,1,1): -3.7312 .
- (2,2,1): -4.2884 .
- (3,1,1): -3.5976 .
- (3,2,1): -4.055 .
- (3,2,2): -3.8375 .
- (3,3,1): -4.4455 .
- (3,3,2): -4.1695 .
- (4,1,1): -3.515 .
- (4,2,1): -3.9091 .
- (4,3,1): -4.2454 .
- (4,3,2): -4.0237 .
- (4,3,3): -3.8834 .
- (4,4,1): -4.5453 .
- (4,4,3): -4.12019 .

Disclaimer: We have no rigorous error bars for the above values, so they may be only right to fewer decimal places. We used our own (admittedly crude) *home-made* asymptotics, implemented in procedure

`MyAsyM(L,mu,n,k)` ,

that inputs a sequence of positive numbers (integers in our case) believed to have an asymptotic behavior of the form (where μ, luckily, is known beforehand)

$$C\,\mu^n n^\theta (1 + \frac{c_1}{n} + \frac{c_2}{n^2} + \ldots)$$

and pretending that the asymptotic expansion only goes as far as the term $\frac{c_k}{n^k}$, takes the log of the last few terms of the sequence and solves for the unknowns $\log C$, θ and $c_1, c_2, \ldots, c_k$. Of course, our main interest is in θ.

Even though our algorithms are *polynomial time* (essentially dynamical programming), computing the first N terms is $O(N^3)$ (for the cubic lattice case). To get a more reliable estimate for the case $(2,1,1)$ we computed 400 terms in the output file

`http://www.math.rutgers.edu/~zeilberg/tokhniot/oCapone4.txt` ,

that yields the estimate -3.731220575 for the critical exponent for that sequence. We would love to know whether the exact value is a rational number, and are offering a donation of 100 dollars to the OEIS in honor of the first (proved) exact answer.

Generalized Ballot Sequences in 4 dimensions

Things start to get slow for higher dimensions, but for a few cases see this output file:

`http://www.math.rutgers.edu/~zeilberg/tokhniot/oCapone3.txt` .

Sequences enumerating the total number of n-step walks

The total number of n-step walks in the k-dimensional hyper-cubic lattice that stay in the region $x_1 \geq x_2 \geq \ldots \geq x_k \geq 0$ for, $2 \leq k \leq 12$, are all in the OEIS. These also count n-celled Standard Young tableaux with at most k rows, and thanks to the Robinson-Schenstead correspondence ([Kn]), the number of $12\ldots(k+1)$-avoiding *involutions* of size n.

• $k = 2$, A1405 (`https://oeis.org/A001405`). Easily seen to be given by $\binom{n}{[n/2]}$.

• $k = 3$, A1006 (`https://oeis.org/A001006`). First discovered, and proved, by Amitai Regev [R], to be given by the almost-as-famous Motzkin numbers.

Higer dimensional cases were treated in [BFK].

• $k = 4$: A5817 (`https://oeis.org/A005817`, see the references there).

• $k = 5$: A49401, (`https://oeis.org/A049401`).

• $k = 6$: A7579, (`https://oeis.org/A007579`).

- $k=7$: A7578, (https://oeis.org/A007578).
- $k=8$: A7580, (https://oeis.org/A007580).
- $k=9$: A212915, (https://oeis.org/A212915).
- $k=10$: A212916, (https://oeis.org/A212916).
- $k=11$: A229053, (https://oeis.org/A229053).
- $k=12$: A229068, (https://oeis.org/A229068).
- $k=13$ was not in the OEIS (viewed April 4, 2021). The first 16 terms are

$$1, 2, 4, 10, 26, 76, 232, 764, 2620, 9496, 35696, 140152, 568504, 2390479, 10349521, 46206511$$

It is well-known that these sequences are P-recursive for all k, so once a recurrence is found it is easy to get many terms, as well as precise asymptotics.

But what about other regions?

Even for the two-dimensional case, except for the sequence enumerating the number of n-step walks in the 2D square lattice that stay in the region $\{(x_1, x_2) \,|\, x_1 \geq 2x_2\}$, that is OEIS sequence A126042 (http://oeis.org/A126042), that is there for a different reason (and it would be an interesting exercise to prove that they are indeed the same), **none** of the other cases, even for the 2D case, seem to be there (yet).

Once again these (in the 2D case) are all algebraic, and henec P-recursive. We conjecture that for higher dimensions they are not P-recursive in general.

For the 2-dimensional case, see the output file:

https://sites.math.rutgers.edu/~zeilberg/tokhniot/oCapone5.txt .

For the 3-dimensional case, see the output file:

https://sites.math.rutgers.edu/~zeilberg/tokhniot/oCapone6.txt .

Here the critical exponents (separated according to n mod $a+b+c$) are (conjecturally!) *nice*: either $-\frac{1}{2}$ or -1.

Conclusion: Why is this Research both Deep and Original?

Determining the (proved!) **exact values** of critical exponents of enumeration sequences that do not seem to be P-recursive is very challenging. In the case of *self-avoiding walks*, it may get you a Fields medal. Since this is out of reach at present, it is fun and interesting to experiment with other kinds of walks, that are just minor tweaks of the classical ballot sequences, yet seem much harder, and we suspect that for dimensions three and higher are **not** P-recursive. But perhaps they belong to another class for which it would be possible to find the exact asymptotics? What's nice about the sequences considered in the present paper is that the algorithms for generating many terms of the sequences are **polynomial time**, hence with larger computers, one should be able to get many more terms. and derive more precise estimates.

Another kind of walks was considered in [KaZ], where the region was the classical one $x_1 \geq x_2 \geq \ldots \geq x_k$ but certain runs were forbidden. Surprisingly there the (conjectured!) critical exponents turned out, often, to be very simple.

If nothing else, we found lots of new sequences! Hopefully some of them will be entered in the OEIS by kind readers.

References

[A] D. André, *Solution directe du problème résolu par M. Bertrand*, Comptes Rendus Acad. Sci. Paris **105** (1887), 436-437.

[BFK] F. Bergeron, L. Favreau and D. Krob, *Conjectures on the enumeration of tableaux of bounded height*, Discrete Math **139** (1995), 463-468.

[B] J. Bertrand, *Solution d'une problème*, Comptes Rendus Acad. Sci. Paris **105** (1887), 369.

[BKK] Cyril Banderier, Christian Krattenthaler, Alan Krinik, Dmitry Kruchinin, Vladimir Kruchinin, David Tuan Nguyen, and Michael Wallner, *Explicit formulas for enumeration of lattice paths: basketball and the kernel method*, 2017.
`https://arxiv.org/abs/1609.06473v2`

[DS] H. Domini-Chupin and S. Smirnov, *The connective constant of the honeycomb lattice is* $\sqrt{2+\sqrt{2}}$, `https://arxiv.org/abs/1007.0575`.
Also in: Anal. of Math. (2) **175** (2012), 1653-1665.

[Ek] Bryan Ek, *Lattice Walk Enumeration*, `https://arxiv.org/abs/1803.10920`

[KaP] Manuel Kauers and Peter Paule, *"The Concrete Tetrahedron"*, Springer, 2011.

[KaZ] Manuel Kauers and Doron Zeilberger, *Counting Standard Young Tableaux with restricted runs*, Personal J. of Shalosh B. Ekhad and Doron Zeilberger.
`https://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/cyt.html` .

[Kn] Donald E. Knuth, *"The Art of Computer Programming"*, Second Edition, Addison-Wesley, 1998.

[L] R.C. Lyness, *Al Capone and the Death Ray*, The Mathematics Gazette **25** (1941), 283-287.

[available from JSTOR]

[N] B. Nienhuis, *Exact critical point and critical exponents of the $O(n)$ models in two dimensions*, Phys. Rev. Letters **49** (1982), 1062-1065.

[R] Amitai Regev, *Asymptotic values for degrees associated with strips of Young diagrams*, Adv. in Math. **41** (1981), 115-136.

[Sl] Neil Sloane, *The On-Line Encyclopedia of Integer Sequences*, `https://oeis.org` .

[St] Richard P. Stanley, *"Catatan numbers"*, Cambridge University Press, 2015.

[TZ] Thotsaporn "Aek" Thanatipanonda, Doron Zeilberger, *A Multi-Computational Exploration of Some Games of Pure Chance*, `https://arxiv.org/abs/1909.11546`.
Also in: J. of Symbolic Computation **104** (2021), 38-68.

[WimZ] Jet Wimp and Doron Zeilberger, *Resurrecting the asymptotics of linear recurrences*, J. Math. Anal. Appl. **111**, 162-177 (1985).

Shalosh B. Ekhad and Doron Zeilberger, Department of Mathematics, Rutgers University (New Brunswick), Hill Center-Busch Campus, 110 Frelinghuysen Rd., Piscataway, NJ 08854-8019, USA.
Email: `[ShaloshBEkhad, DoronZeil] at gmail dot com` .

Exclusively published in the Personal Journal of Shalosh B. Ekhad and Doron Zeilberger and arxiv.org

Original version, written: **April 4, 2021**.

This version, written: **May 27, 2021**.