Pascal’s Triangle and the Tower of Hanoi

Andreas M. Hinz

INTRODUCTION. The most genuine examples for the principle of complete induction are the arithmetic triangle (AT) and the Tower of Hanoi (TH). They also reveal an unexpected mathematical relation which will be developed here.

The AT has been studied by Blaise Pascal in a treatise published posthumously in 1665 and is therefore often called Pascal’s triangle. However, it was known before in Europe (Peter Apian, 1527), China (Jia Xian, 11th century), the Islamic world (al-Karaji, ca. 1000), and possibly India (Pingala, ca. –200). The TH is an invention of the French mathematician Édouard Lucas, who first published the puzzle in 1883. An account of its history and basic mathematical properties can be found in [4].

Recently, connections between the AT and the Sierpiński gasket (SG) have been observed. The SG is obtained from a closed equilateral triangle by deleting the open middle triangle and iterating this step for the remaining subtriangles ad infinitum. It turns out that its fractal geometry is the same as that of the AT modulo 2 (see [2, p. 10f]). On the other hand the SG can be viewed, in a certain sense, as the limit of the graph of the TH for an increasing number of discs (see Hinz and Schief [5]). So, by transitivity, there must be a link between the TH and the AT. Since the TH is closely related to binary structures, it is not surprising that this connection is again with AT mod 2.

The TH with \( n \in \mathbb{N}_0 \) discs will be identified with the graph \( TH_n \), whose vertices are the distributions of \( n \) discs among three pegs which are regular (i.e. no disc lies on a smaller one), and whose edges are legal moves of a single top disc, leading from one such distribution to another. This graph is simple, undirected, planar, and connected. Figure 1 shows the example \( n = 3 \). (The discs being numbered from 1 to \( n \), the pegs named 0, 1, and 2, the state with disc 1 on peg 0, disc 2 on peg 2, and disc 3 on peg 1 is abbreviated 021, for instance.)

![Figure 1](image1)

![Figure 2](image2)
ATₘ will denote the AT with m ∈ ℤ rows, counted as usual from the 0th row at the apex to the (m − 1)-th row at the base. Assume further that the geometry of ATₘ is as symmetric as possible, i.e. nearest neighbors are a unit apart. Then the basic observation is the following: The graph AT₂ⁿ mod 2, consisting of the odd numbers in AT₂ⁿ joined by an edge if one unit apart (see Figure 2), is isomorphic to THₙ.

1. THE PARITY OF BINOMIAL COEFFICIENTS. The parity of binomial coefficients \( \binom{\mu}{k} \) has recently played an important role in a paper of Jones and Matijasevič [6] in connection with Hilbert’s tenth problem, Gödel’s undecidability proposition, and computational complexity. They base their Lemma on the following theorem of Lucas [9, Section XXI].

**Theorem 0.** Let p be a prime. Then

\[
\binom{\mu}{k} = \prod_{i=0}^{n-1} \binom{\mu_i}{k_i} \mod p,
\]

where \( \mu_i \) and \( k_i \) are the p-ary digits (or pits) of \( \mu \) and \( k \), respectively.

Since Jones and Matijasevič only need the case \( p = 2 \), they could have relied on an older result of Kummer [7, p. 115f], namely, that the highest power of \( p \) dividing \( \binom{k+v}{k} \) is equal to the number of carries in the p-ary addition of \( k \) and \( v \), which for \( p = 2 \) means that \( \binom{\mu}{k} \) is odd if and only if \( k_i \leq \mu_i \) for all \( i \).

Lucas states in his famous book *Théorie des nombres* of 1891 (p. 420) that all binomial coefficients in a row of the AT are odd only if the row number is one less than a power of two:

\[
\forall 0 \leq k \leq \mu: \binom{\mu}{k} \text{ odd} \Leftrightarrow (\exists n \in \mathbb{N}_0: \mu = 2^n - 1), \tag{1}
\]

while the complementary statement, namely, they are all even (except the outer ones) if the row number is a power of two:

\[
\forall 0 < k < \mu: \binom{\mu}{k} \text{ even} \Leftrightarrow (\exists n \in \mathbb{N}_0: \mu = 2^n), \tag{2}
\]

is due to Fine. Fine also proved that odd binomial coefficients are sparse:

\[
\#\left\{\text{odd }\binom{\mu}{k} \in \text{AT}_m\right\} / \#(\text{AT}_m) \to 0, \quad \text{as } m \to \infty. \tag{3}
\]

All these results arise from Theorem 0 and can be extended to a general prime \( p \). As another consequence of Theorem 0, Glaisher represented the number of odd binomial coefficients in the \( \mu \)th row as:

\[
\forall \mu \in \mathbb{N}_0: \#\left\{\text{odd }\binom{\mu}{k}\right\} = 2^{\beta(\mu)}, \tag{4}
\]

where \( \beta(\mu) \) is the number of non-zero bits of \( \mu \).

The references to Fine and Glaisher, as well as to many other works on parity of binomial coefficients, can be found in Stolarsky [13], who was interested primarily
in the asymptotic behavior, as \( m \to \infty \), of
\[
B(m) := \# \left\{ \text{odd } \binom{\mu}{k} \in \text{AT}_m \right\},
\]
the number of odd binomial coefficients in the first \( m \) rows. It turns out that \( B(m) \) behaves essentially like \( m^s \), where \( s = \ln 3 / \ln 2 \) is the Hausdorff dimension of the SG (see [2, pp. 157–159]; roughly speaking, doubling the linear extension of the SG means tripling its measure, whence \( 2^s = 3 \)).

An explicit formula for \( B(m) \) can be obtained from the special case \( p = 2 \) of Corollary 4 in Roberts [10]:
\[
\forall m \in \mathbb{N}: B(m) = \sum_{i=0}^{n-1} m_i 2^{\frac{1}{n-1} m_i - 1} \cdot 3^i, \tag{5}
\]
for \( m = \sum_{i=0}^{n-1} m_i \cdot 2^i, m_i \in \{0, 1\} \).

Because of the isomorphy of \( \text{AT}_2^n \mod 2 \) and \( \text{TH}_n \) all these results can be reinterpreted in terms of the TH puzzle. For instance, \( B(m) \) is the number of states accessible from the perfect initial distribution (i.e., all discs are on the same peg) in less than \( m \) moves. However, most of these statements are easier derived from the properties of the TH. This and some additional results will be achieved in Section 3. Before doing that, it seems adequate to honor the person who stands for these things.

2. ÉDOUARD LUCAS (1842–1891). François Édouard Anatole Lucas was born on April 4th, 1842 in Amiens (France). Son of a worker, his talents earned him a scholarship for higher education. In 1861 he was accepted by the most prestigious French institutions of the time, the École polytechnique and the École normale. Lucas attended the latter and left it in 1864 as Agrégé des sciences mathématiques.

Édouard Lucas (1842–1891)

The employment at the Paris Observatory as an assistant of Leverrier was interrupted by his active participation in the Franco-Prussian war of 1870/71. His last twenty years Lucas held positions as a teacher of higher mathematics at the high schools of Moulins (72–76), Paris Charlemagne (76–79, 90–91), and Paris St. Louis (79–90). Being a mathematician out of line who was described as young, ardent, and energetic till the end of his life, this professional situation was not adequate since “his character of a noble independence, his spontaneous mind were not able to bend into the narrow mould of university or even high school
teaching, not more than his high intelligence, we may say his genius, could stay a prisoner of programmes. (A. Béligne)” His research interests being centered in number theory, Lucas himself felt that he was living “at a time and in a country where higher arithmetic is forsaken by mathematicians and public education.” So his main activities were focussed to learned societies of France and other countries and, of course, to his written works, which unfortunately are not accessible in collected form, but a catalogue of which has been compiled by Harkin [3].

Besides some papers on geometry, most of his articles and books concentrate on number theory, recurrent series, and recreational mathematics. He was the last “largest prime number record” holder in pre-computer age, has a series of numbers, namely 2, 1, 3, 4, 7, 11, …, called after him, and published, in addition to the famous TH of N. Claus de Siam (= Lucas d’Amiens), a collection of scientific puzzles, now apparently lost, which won a gold medal at the world’s fair of 1889. He left a couple of books unfinished, in particular the planned sequel of the Théorie des nombres. So large was the interest in his unpublished papers, that, as E. T. Bell once remarked, “the fantastic price of thirty thousand dollars was being asked for Lucas’s manuscripts. In all his life Lucas never had that much money.” One may doubt at least the last sentence, since it is known that Lucas donated a collection of calculating machines, among which those of Chebyshev and Roth, to a museum in Paris. This was in connection with his efforts to make mathematics popular, and it is said that he was an entertaining teacher in lectures for a general audience. Here and in his papers he took an interest in the history of mathematical problems, definitions, and theorems—not a very common attitude at his time. Lucas was actively involved in the publication of Fermat’s collected works and mentioned, an interesting detail in connection with the present note, that he lived for a while “No. 56 rue Monge in Paris, in the house built on the site of the one where Pascal died on August 19th, 1662.”

Édouard Lucas himself died in Paris, aged only 49, on October 3rd, 1891.

3. THE TOWER OF HANOI. The TH graphs as defined in the introduction can be obtained recursively in the following way: TH₀ has only one vertex (three empty pegs) and no edges (there are no discs to be moved); THₙ₊₁ is composed of three triangular THₙ graphs (movements in the n smaller discs are not affected by the largest disc which is on one of three possible pegs) joined at their base corners (disc n + 1 can move only if the other discs are on the peg not involved in that move).

Similarly, AT₂₀ mod 2 consists of just one 1, and AT₂ⁿ₊₁ mod 2 is constructed recursively in the same manner as THₙ₊₁ as can be seen from the following lemma.

**Lemma 1.** ∀ n ∈ ℕ₀ ∀ 0 ≤ ν, k < 2ⁿ: \( \binom{2^n + \nu}{k} \equiv \binom{\nu}{k} \mod 2. \)

This lemma is an immediate consequence of Theorem 0 (or of Kummer’s theorem) or can easily be proved by induction.

Thus the basic theorem of this paper is established, namely.

**Theorem 1.** For any \( n \in ℕ₀ \), \( AT₂ⁿ \mod 2 \) and \( THₙ \) are isomorphic.

With the help of this observation, properties of \( THₙ \), as developed e.g. in [4], will now be turned into statements about odd binomial coefficients. For instance,
from the trivial fact that there are precisely $3^n$ regular distributions of $n$ discs among three pegs it follows that for any $n \in \mathbb{N}$ ($\#(\text{graph})$ means the number of vertices):

$$\forall 2^{n-1} \leq m \leq 2^n: \frac{\#(AT_m \mod 2)}{\#(AT_m)} \leq \frac{\#(AT_{2^n} \mod 2)}{\#(AT_{2^n-1})}$$

$$= \frac{\#(TH_n)}{\#(AT_{2^n-1})} = \frac{3^n}{2^{n-2}(2^{n-1} + 1)},$$

which yields Fine’s result (3).

As mentioned in the introduction, $B(m)$, the number of odd binomial coefficients in the AT with $m \in \mathbb{N}$ rows, is, by Theorem 1, equal to the number of states of the TH which are accessible from a perfect starting configuration in at most $m - 1$ moves. Since it is obvious from the TH graph that $B(1) = 1$ and for any $n \in \mathbb{N}$

$$B\left(\sum_{i=0}^{n} m_i \cdot 2^i\right) = m_n \cdot 3^n + 2^{m_n}B\left(\sum_{i=0}^{n-1} m_i \cdot 2^i\right),$$

an induction on $n$, the length of the binary representation of $m$, yields Roberts’ formula (5).

Another look at the TH graph shows that

$$\forall n \in \mathbb{N}: 2^{n-1} \leq m < 2^n \Rightarrow 3^{n-1} \leq B(m) < 3^n,$$

from which follows the rough estimate of Stolarsky [13, Th. 1]:

$$\forall m \in \mathbb{N}: \frac{1}{3} < \frac{B(m)}{m^2} < 3.$$

Glaisher’s formula (4) for the number of odd binomial coefficients in row $\mu$ is a direct consequence of Proposition 5 in [4] which says that the number of states of the TH that are precisely $\mu$ steps away from a specific perfect state (here the top apex of $TH_n$) is $2^{\beta(\mu)}$. (1) and (2) are, of course, just special cases of this. Applying the same proposition to the lower left apex, however, one learns that the number of odd binomial coefficients at a (graph) distance $\bar{\nu}$ from this corner is $2^{\beta(\bar{\nu})}$. But these are just the odd numbers in the $(2^n - 1 - \bar{\nu})$-th diagonal of $AT_{2^n}$, consisting of binomial coefficients of the form $\binom{k + \nu}{k}$ with $\nu = 2^n - 1 - \bar{\nu}$, i.e. figurate numbers of order $\nu$ (so called as generalizations of triangular numbers ($\nu = 2$) and tetrahedral numbers ($\nu = 3$)). This can be summarized as follows:

**Proposition 1.** Among the first $2^n - \nu$ figurate numbers of order $\nu$ ($0 \leq \nu < 2^n$), $2^{\beta(\nu)}$ are odd, where $\beta(\nu)$ is the number of zero bits in the $n$-bit representation of $\nu$.

Although this result may not be too surprising, it is amazing how it came about from the TH. But there are some deeper insights which stem from considering yet another counting on the graph $TH_n$, namely the function $z_n$ that gives for an integer $\mu$ the number of states in $TH_n$ for which the difference of the distances to two distinct corners is exactly $\mu$. (By symmetry, this does not depend on the pair considered.) Before discussing the functions $z_n$ more detailed in the next section, their appearance in the AT should be pointed out.

Odd binomial coefficients in $AT_{2^n}$ for which the difference of the distances to the base corners is $\nu \in \mathbb{N}_0$ are those of the form $\binom{2k + \nu}{k}$ with $0 \leq k < (2^n - \nu)/2$. 

542 PASCAL’S TRIANGLE AND THE TOWER OF HANOI [June–July
Hence Theorem 1 gives for any \( n \in \mathbb{N}_0 \):

**Proposition 2.** Among the first \( [(2^n - \nu)/2] \) numbers in the \( \nu \)-th column of the \( AT \) \((0 \leq \nu < 2^n)\), \( z_n(\nu) \) are odd; here the columns are counted from the 0th at the center.

Note that the \( \nu \)-th column consists of the coefficients of the Chebyshev polynomial \( y_\nu \) in the development of \( \frac{1}{2}(2x)^{\nu+2k} \) (here \( y_0 = \frac{1}{2} \)).

Adding up the entries of the \( \nu \)-th subdiagonal of the \( AT \), one gets the Fibonacci number \( F_\nu \):

\[
F_\nu = \sum_{k=0}^{[\nu/2]} \binom{\nu - k}{k}.
\]

This representation can be found in Siebeck [12, p. 71] (the last term in his first formula should be \( 1 \cdot c^{(r-1)/2} \)). The geometrical interpretation of the \( AT \) has been given by, of course, Lucas [8, p. 138f], who also introduced in that article the name Fibonacci series. Since for odd binomial coefficients of the \( \nu \)-th subdiagonal \((0 \leq \nu < 2^n)\) the difference between the distances to the bottom right and top corner of \( AT_{2^n} \), respectively, is \( \nu = 2^n - 1 - \nu \), one has for any \( n \in \mathbb{N}_0 \):

**Proposition 3.** \( z_n(\nu) \) of the binomial coefficients in the \( \nu \)-th subdiagonal \((0 \leq \nu < 2^n)\) are odd.

4. THE FUNCTIONS \( z_n \). The following has been proved in [4, L.2].

**Lemma 2.** o) \( z_0(0) = 1 \), \( \forall \mu \in \mathbb{Z}\setminus\{0\} \): \( z_0(\mu) = 0 \);

\[ \forall \ n \in \mathbb{N}_0 \ \forall \mu \in \mathbb{Z}: z_n(\mu) = z_n(\mu - 2^n) + z_n(\mu) + z_n(\mu + 2^n); \]

i) \( \forall \ n \in \mathbb{N}_0 \ \forall \mu \in \mathbb{Z}: z_n(-\mu) = z_2(\mu), |\mu| \geq 2^n \Rightarrow z_n(\mu) = 0; \]

\[
z_n(0) = 1, \ z_n(1) = n, \ z_n(2^n - 1) = 1.
\]

Note that, since

\[
z_n(\mu) = \sum_{\xi \in \{-1,0,1\}^n} z_0 \left( \mu + \sum_{i=0}^{n-1} \xi_i \cdot 2^i \right)
\]

by induction from Lemma 2o, \( z_n(\mu) \) is just the number of ways \( \mu \) can be written as \( \sum_{i=0}^{n-1} \xi_i \cdot 2^i \) with \( \xi_i \in \{-1,0,1\} \). This shows that the \( z_n \) are not very easily accessible functions (cf. the discussion at the end of [5]). However, some special relations are feasible.

Let \( 2^a \leq \mu < 2^{a+1} \) for an \( a \in \mathbb{N}_0 \). Then (by Lemma 2i)

\[ \forall \ n \leq a: z_n(\mu) = 0 \]

and (by Lemma 2o)

\[ \forall \ n > a: z_{n+1}(2^n - \mu) = z_n(2^n - \mu), \]

whence (by induction)

\[ \forall \ k \in \mathbb{N}: z_{a+k}(\mu) = z_{a+1}(\mu) + (k - 1) z_{a+1}(2^{a+1} - \mu). \]

That is to say, for fixed \( \mu \), \( z_n(\mu) \) is eventually in arithmetic progression, while e.g. the lengths of the columns in Proposition 2 are essentially in geometric progression.

By Lemma 2, \( z_n(0) = 1 \), such that \( \forall \ k \neq 0: 2^{(2k)} \) by Proposition 2.

For the special cases \( \mu = 2^a, 2^a + 1, 2^a + 1 - 1 \ (a \in \mathbb{N}_0) \), (7), and Lemma 2 yield

\[
\forall k \in \mathbb{N}: z_{a+k}(\mu) = \begin{cases} 
  k \\
  a + (k - 1)(a + 1) \\
  1 + (k - 1)(a + 1),
\end{cases}
\]

respectively.

As an example, the sum for \( F_{22} \) in (6) is made up of \( z_5(9) = 7 \) odd and 5 even numbers by Proposition 3. (Though parity has been associated with gender, these values should not be taken as the numbers of male and female rabbits, since by definition of \( F_n \) they always appear in pairs!)

5. PASCAL'S TRIANGLE AND THE TOWER OF HANOI WITH MORE THAN THREE PEGS. It should be noted that the AT has been used in algorithms for a solution of the TH with more than three pegs; see e.g. Rohl and Gedeon [11]. Although in this paper, as in many others on the subject, minimality of the solution is claimed, there is no proof for that. As it stands, Monthly Problem 3918 [1939, p. 363] is still unsolved (cf. [1]), namely: What is the minimum number of moves required to transfer \( n \) discs from one of \( k \geq 3 \) pegs to another?

ACKNOWLEDGMENTS. This note was inspired by an article of I. Stewart (Warwick, England), a talk by J. M. Holte (St. Peter, Minnesota), and a hint of A. Douady (Paris, France), whom I all met at the ICM90 in Kyoto (Japan). I thank the Deutsche Forschungsgemeinschaft for travel support and Osanobu Yamada of the Ritsumeikan University in Kyoto for his kind hospitality.

REFERENCES

1. O. Dunkel, (Editorial Note), this Monthly, 48 (1941) 219.