Séminaire Lotharingien de Combinatoire **92** (2025), Article B92c

A PROBABILISTIC PROOF OF EULER'S PENTAGONAL NUMBER THEOREM

SHANE CHERN

ABSTRACT. We present a probabilistic proof of Euler's pentagonal number theorem based on a shuffling model.

To George Andrews, a great advisor, on his $\lfloor (\tanh 1; \tanh 1)_\infty^{-1} \rfloor$*-th birthday.*

1. Prologue

Euler's *pentagonal number theorem* is recognized as one of the pioneering discoveries in the analysis of the infinite. Formally, it is the following series expansion.

Theorem 1.1. *We have*

$$\prod_{j\ge 1}(1-q^j) = \sum_{n=-\infty}^{\infty} (-1)^n q^{\frac{n(3n+1)}{2}}. \tag{1.1}$$

This result was first communicated from Euler to Daniel Bernoulli in a letter dated November 30, 1740[1]; however, no proof was included therein. The seemingly earliest demonstration offered by Euler [3] was published in 1760, which is a retelling of the idea in his letter to Goldbach [4, Brief 144] on June 9, 1750. Euler's proof, as briefly presented at the end of this paper, is inductive, relying on an ingenious arithmetic manipulation of the series.

Ever since its discovery, the pentagonal number theorem has shown broad and profound connections to other fields. As observed by Legendre [8, pp. 132–133, §458], the series expansion (1.1) may be combinatorially interpreted in terms of integer partitions; this interpretation motivated Franklin's bijective proof [6]

2020 *Mathematics Subject Classification.* 60C05, 05A19, 11P84.

Key words and phrases. Euler's pentagonal number theorem, probabilistic proof, shuffling model.

[1]Euler's original letter, unfortunately, has not been preserved but the lines regarding the pentagonal number theorem were alluded to briefly in Bernoulli's reply dated January 28, 1741. See [5, Brief 52].

half a century later. Also, Goncharova [7] noticed that the pentagonal number theorem arises as a natural implication from the calculation of the cohomology of the negative part of the Witt algebra, and it is just a tip of a big picture that bonds Lie algebras, partition theory, and q-series. We refer the interested reader to a terse exposition by Andrews [1], and a more detailed summary on Euler's discovery by Bell [2].

The objective of the present paper is to interpret the pentagonal number theorem in a different way. To be specific, we propose a shuffling model in the next section and utilize it to provide a new probabilistic proof.

2. A shuffling model

Let $\mathbb{N}_+$ be the set of positive integers, and $\mathbb{N} := \mathbb{N}_+ \cup \{0\}$ the set of natural numbers. Given $a, b \in \mathbb{N}$, we denote by $[a, b]$ the set of natural numbers n such that $a \le n \le b$, and by $[a, \infty)$ the set of natural numbers n such that $n \ge a$.

Throughout, we always assume that q is a real number such that $0 < q < 1$.

In a classical *probability model* $(\Omega, \mathbf{P})$, the *sample space* Ω describes the set of all possible outcomes, and the *probability function* $\mathbf{P} : \Omega \to [0, 1]$ assigns a probability with value taken from $[0, 1]$ to each element in Ω such that

$$\sum_{\omega \in \Omega} \mathbf{P}_\omega = 1.$$

We may also attach an *event space* E, which is a subset of Ω, to the model in order to describe the outcomes of particular interest. By abuse of notation, we denote by $\mathbf{P}_E$ the total probability of all elements in the event space E. This *probability space* is written as $(\Omega, E, \mathbf{P})$.

Now fix a sequence $\mathfrak{s}_k := (s_0, \ldots, s_{k-1}, s_k)$ of $k + 1$ different natural numbers for a certain $k \ge 0$. We define a probability model $\text{Ш}_{\mathfrak{s}_k} = (\Omega_{\mathfrak{s}_k}, \mathbf{P}^{(\mathfrak{s}_k)})$ where the sample space is

$$\Omega_{\mathfrak{s}_k} = \big\{ \boldsymbol{s}_0, \boldsymbol{s}_1, \ldots, \boldsymbol{s}_{k-1}, \boldsymbol{s}_k \big\},$$

with[2]

$$\boldsymbol{s}_i = \boldsymbol{s}_i(\mathfrak{s}_k) := \begin{cases} (s_0, \ldots, s_{k-1}, s_k), & \text{if } i = 0, \\ (s_0, \ldots s_{k-i-1}, s_k, s_{k-i}, \ldots, s_{k-1}), & \text{if } i \in [1, k-1], \\ (s_k, s_0, \ldots, s_{k-1}), & \text{if } i = k. \end{cases} \tag{2.1}$$

[2]It is notable that throughout this paper, *three* different fonts for "s" are used, including s, $\boldsymbol{s}$ and $\mathfrak{s}$, and they should be carefully distinguished, especially for the latter two (in *bold* and *Fraktur* styles).

The probability function is then assigned by

$$\mathbf{P}_{\boldsymbol{s}_i}^{(\mathfrak{s}_k)} = \begin{cases} 1-q, & \text{if } i=0, \\ q^i(1-q), & \text{if } i \in [1,k-1], \\ q^k, & \text{if } i=k. \end{cases} \tag{2.2}$$

Note that the outcomes in the sample space are sequences permuted by the original one. Intuitively, we *swap* s_k with its preceding entry with a probability q through a step-by-step procedure, and the program is terminated when a certain step of swapping is not executed or s_k already reaches the leftmost position.

This procedure more or less aligns with *card shuffling*. So although less accurate[3], in the present paper, for S a subset of $\mathbb{N}$, we say a sequence $\mathfrak{s} := (s_0, s_1, s_2, \ldots)$ is a *shuffle* (instead of using the usual term of *permutation*) of S if every number in $\mathfrak{s}$ belongs to S and every number in S appears exactly once in $\mathfrak{s}$.

We then describe our *shuffling model* Ш according to a *Markov process*:

Definition 2.1 (Markov process for the shuffling model Ш)**.**

- **Time 0:** We begin with the sequence (0);
- **Time $\boldsymbol{k}$ ($\boldsymbol{k \geq 1}$):** For the sequence $\mathfrak{s}_{k-1} = (s_0, \ldots, s_{k-1})$ derived at Time $k-1$, we append the number k to the right end to get $\hat{\mathfrak{s}}_k = (s_0, \ldots, s_{k-1}, k)$ (so that the new sequence $\hat{\mathfrak{s}}_k$ is of length $k+1$) and then apply the probability model $Ш_{\hat{\mathfrak{s}}_k}$.

If we only look at the outcomes of this stochastic process at Time k and denote the associated probability model as $Ш_k := (\Omega_k, \mathbf{P}^{(k)})$, it is clear that the sample space Ω_k covers all shuffles of $[0,k]$. Let $\mathfrak{s}_k = (s_0, \ldots s_{k-i-1}, k, s_{k-i}, \ldots, s_{k-1})$ be such a shuffle wherein $i \in [0,k]$. Then by removing k, the sequence $\mathfrak{s}_{k-1} = (s_0, \ldots, s_{k-1})$ is a shuffle of $[0, k-1]$ according to the construction of the Markov process, and more importantly, it is *exactly* the one derived at Time $k-1$ in this Markov process, thereby lying in the sample space of $Ш_{k-1}$. We also iteratively have the probability evaluation:

$$\mathbf{P}_{\mathfrak{s}_k}^{(k)} = \mathbf{P}_{\mathfrak{s}_{k-1}}^{(k-1)} \cdot P_i,$$

[3]The term *shuffle* is usually reserved for permutations where one divides a sequence into a front and a tail part and then forms a permutation in which the order of the elements from the front part is preserved, and also the order of the elements from the tail part (thus imitating a riffle shuffle of cards). For example, in the *shuffle algebra*, the multiplication is exactly defined using this shuffle.

where $\mathbf{P}^{(k)}_{\mathfrak{s}_k}$ denotes the probability of the event $\mathfrak{s}_k$ with respect to the probability model Ш_k and

$$P_i = \begin{cases} 1-q, & \text{if } i=0,\\ q^i(1-q), & \text{if } i\in[1,k-1],\\ q^k, & \text{if } i=k, \end{cases}$$

according to the probability model $\text{ш}_{\hat{\mathfrak{s}}_k}$.

Now we read our *shuffling model* $\text{Ш}=(\Omega,\mathbf{P})$ as Ш_∞. Then the sample space Ω of Ш covers all shuffles of $\mathbb{N}$.

Example 2.2. Consider the shuffle $(2,0,3,1,4,5,6,7\ldots)$ of $\mathbb{N}$. Then in the Markov process, at Time 0 the resulting event is (0); at Time 1 the resulting event is $(0,1)$; at Time 2 the resulting event is $(2,0,1)$; at Time 3 the resulting event is $(2,0,3,1)$; at Time 4 the resulting event is $(2,0,3,1,4)$; and so on.

3. Event spaces

In this section, we describe the event spaces of our interest attached to the shuffling model $\text{Ш}=(\Omega,\mathbf{P})$.

For $\mathfrak{s}_k=(s_0,\ldots,s_k)$ a generic sequence of different natural numbers of length $k+1$, recall that we have defined a probability model $\text{ш}_{\mathfrak{s}_k}=(\Omega_{\mathfrak{s}_k},\mathbf{P}^{(\mathfrak{s}_k)})$ in the previous section, where we have in particular labeled elements in the sample space $\Omega_{\mathfrak{s}_k}$ by $\boldsymbol{s}_0,\ldots,\boldsymbol{s}_k$ as recorded in (2.1). Now in this model, we further attach an event space $E_{\mathfrak{s}_k}\subset\Omega_{\mathfrak{s}_k}$ and call the probability space $\text{щ}_k=\text{щ}_k(\mathfrak{s}_k):=(\Omega_{\mathfrak{s}_k},E_{\mathfrak{s}_k},\mathbf{P}^{(\mathfrak{s}_k)})$ a *shuffling (restricted to $E_{\mathfrak{s}_k}$)*. For any element in $\Omega_{\mathfrak{s}_k}$, we say it is *produced* by this shuffling щ_k if it belongs to $E_{\mathfrak{s}_k}$; otherwise, it is *not* produced by щ_k.

We make the following categorization:

- Type $\triangleright$: The shuffling щ_k is *leftmost* if the event space $E_{\mathfrak{s}_k}$ is $\{\boldsymbol{s}_k\}$;
- Type $\triangleleft$: The shuffling щ_k is *anti-leftmost* if the event space $E_{\mathfrak{s}_k}$ is $\{\boldsymbol{s}_0,\ldots,\boldsymbol{s}_{k-1}\}$;
- Type $\diamond$: The shuffling щ_k is *free* if the event space $E_{\mathfrak{s}_k}$ is $\{\boldsymbol{s}_0,\ldots,\boldsymbol{s}_{k-1},\boldsymbol{s}_k\}$.

Next, we move on to the shuffling model Ш. Recall that Ш is defined based on a Markov process introduced in Definition 2.1 of the previous section. Given any outcome $\mathfrak{s}\in\Omega$ of the model Ш, it is derived by the infinite chain

$$\mathfrak{s}_0\to\mathfrak{s}_1\to\cdots\to\mathfrak{s}_k\to\cdots,$$

where, for the moment, $\mathfrak{s}_k$ is the shuffle obtained at Time k of our Markov process. Now it is notable that $\mathfrak{s}_k$, called the *projection* of $\mathfrak{s}$ onto the model Ш_k, belongs to the sample space Ω_k. In light of what has been done at Time k of the Markov process, $\mathfrak{s}_k$ may also be viewed as an outcome of the model $ш_{\hat{\mathfrak{s}}_k}$ where $\hat{\mathfrak{s}}_k = (s_0, \ldots, s_{k-1}, k)$ with $(s_0, \ldots, s_{k-1}) =: \mathfrak{s}_{k-1}$.

The takeaway in our analysis is that once $\mathfrak{s}_{k-1} \in \Omega_{k-1}$ is determined at Time $k-1$ of the Markov process, we will not only focus on all possible outcomes of the above $ш_{\hat{\mathfrak{s}}_k}$. Instead, we look at a shuffling $ш_k = ш_k(\hat{\mathfrak{s}}_k) = (\Omega_{\hat{\mathfrak{s}}_k}, E_{\hat{\mathfrak{s}}_k}, \mathbf{P}^{(\hat{\mathfrak{s}}_k)})$ restricted to a certain event space $E_{\hat{\mathfrak{s}}_k} \subset \Omega_{\hat{\mathfrak{s}}_k}$. More precisely, we require that $ш_k$ belongs to one of the three types introduced earlier: "leftmost", "anti-leftmost", or "free". By doing so, our attention is indeed restricted to an event space $E_k \subset \Omega_k$ attached to the model $Ш_k$.

Example 3.1. Consider the Markov process where the outcomes at Time 0, 1, and 2 are given by

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

Then at Time 3,

- The outcome produced by a leftmost shuffling $ш_3$ includes $(3, 2, 0, 1)$;
- The outcomes produced by an anti-leftmost shuffling $ш_3$ include $(2, 3, 0, 1)$, $(2, 0, 3, 1)$, and $(2, 0, 1, 3)$;
- The outcomes produced by a free shuffling $ш_3$ include $(3, 2, 0, 1)$, $(2, 3, 0, 1)$, $(2, 0, 3, 1)$, and $(2, 0, 1, 3)$.

The same idea may be repeated at each time of the Markov process.

For the moment, let us mark each index in the *time set* $\mathbb{N}_+$ with a chosen subscript:

$$\mathbf{T} := \big\{ k_{\Box_k} : k \in \mathbb{N}_+ \big\},$$

where $\Box_k \in \{\triangleright, \triangleleft, \Diamond\}$; such $\mathbf{T}$ is called a *marked time set*.

Now starting with $E_0(\mathbf{T}) = \big\{(0)\big\} = \Omega_0$, we may iteratively construct an event space $E_k(\mathbf{T}) \subset \Omega_k$ attached to the model $Ш_k$ by

$$E_k(\mathbf{T}) := \bigcup_{\mathfrak{s}_{k-1} \in E_{k-1}(\mathbf{T})} \left\{ \begin{matrix} \text{outcomes of the shuffling} \\ ш_k = (\Omega_{\hat{\mathfrak{s}}_k}, E_{\hat{\mathfrak{s}}_k}, \mathbf{P}^{(\hat{\mathfrak{s}}_k)}) \text{ of Type } \Box_k \end{matrix} \right\},$$

where as before we write $\hat{\mathfrak{s}}_k = (s_0, \ldots, s_{k-1}, k)$ with $(s_0, \ldots, s_{k-1}) =: \mathfrak{s}_{k-1}$. Eventually, we arrive at an event space attached to our shuffling model $Ш = (\Omega, \mathbf{P})$:

$$E(\mathbf{T}) := E_\infty(\mathbf{T}) \subset \Omega.$$

Example 3.2. We fix the marked time set as $\{1_\Diamond, 2_\triangleleft, 3_\triangleright, \ldots\}$. Since we witness a free ($\Diamond$) shuffling at Time 1, then

$$E_1(\{1_\Diamond, 2_\triangleleft, 3_\triangleright, \ldots\}) = \big\{(0, 1), (1, 0)\big\};$$

next, since we witness an anti-leftmost ($\triangleleft$) shuffling at Time 2, then

$$E_2(\{1_\Diamond, 2_\triangleleft, 3_\triangleright, \ldots\}) = \big\{(0,1,2),(0,2,1),(1,0,2),(1,2,0)\big\};$$

next, since we witness a leftmost ($\triangleright$) shuffling at Time 3, then

$$E_3(\{1_\Diamond, 2_\triangleleft, 3_\triangleright, \ldots\}) = \big\{(3,0,1,2),(3,0,2,1),(3,1,0,2),(3,1,2,0)\big\}.$$

If we define

$$I := \big\{k \in \mathbb{N}_+ : \square_k = \triangleright\big\},$$
$$J := \big\{k \in \mathbb{N}_+ : \square_k = \triangleleft\big\},$$

it is clear that I and J, as subsets of $\mathbb{N}_+$, are *disjoint*. Conversely, suppose I and J are two disjoint subsets of $\mathbb{N}_+$. Then we may recover a marked time set $\{1_{\square_1}, 2_{\square_2}, 3_{\square_3}, \ldots\}$ according to the rule:

$$\square_k = \begin{cases} \triangleright, & \text{if } k \in I, \\ \triangleleft, & \text{if } k \in J, \\ \Diamond, & \text{if } k \notin I \text{ and } k \notin J. \end{cases}$$

Accordingly, we have an event space $E(\{1_{\square_1}, 2_{\square_2}, 3_{\square_3}, \ldots\})$ attached to the model Ш, and we abbreviate it by the notation:

$$\big\langle I_\triangleright, J_\triangleleft \big\rangle := E(\{1_{\square_1}, 2_{\square_2}, 3_{\square_3}, \ldots\}).$$

Example 3.3. Consider the shuffle

$$(6,1,0,4,2,3,5,7,8,9,10,\ldots) \in \Omega.$$

It belongs to the event spaces $\big\langle \varnothing_\triangleright, \varnothing_\triangleleft \big\rangle$, $\big\langle \{1\}_\triangleright, \varnothing_\triangleleft \big\rangle$, $\big\langle \{1,6\}_\triangleright, \varnothing_\triangleleft \big\rangle$, $\big\langle \varnothing_\triangleright, \{2\}_\triangleleft \big\rangle$, $\big\langle \varnothing_\triangleright, \{4,5\}_\triangleleft \big\rangle$, $\big\langle \{1\}_\triangleright, \{2,4\}_\triangleleft \big\rangle$ (and many others) but not the space $\big\langle \{2\}_\triangleright, \varnothing_\triangleleft \big\rangle$ (because the projection of shuffles in this space onto $Ш_2$ should be either $(2,0,1)$ or $(2,1,0)$) nor $\big\langle \varnothing_\triangleright, \{1\}_\triangleleft \big\rangle$ (because the projection of shuffles in this space onto $Ш_1$ should be $(0,1)$).

We close this section with two observations.

Proposition 3.4. *Let I and J be two disjoint subsets of $\mathbb{N}_+$, and I' and J' another two disjoint subsets of $\mathbb{N}_+$. If there exists a positive integer k such that $k \in I$ and $k \in J'$, then*

$$\big\langle I_\triangleright, J_\triangleleft \big\rangle \cap \big\langle I'_\triangleright, J'_\triangleleft \big\rangle = \varnothing. \tag{3.1}$$

Proof. Assume that there exists a certain shuffle $\mathfrak{s}$ in both $\big\langle I_\triangleright, J_\triangleleft \big\rangle$ and $\big\langle I'_\triangleright, J'_\triangleleft \big\rangle$. We project $\mathfrak{s}$ onto the model $Ш_k$ to get a shuffle $\mathfrak{s}_k \in \Omega_k$. Since $\mathfrak{s} \in \big\langle I_\triangleright, J_\triangleleft \big\rangle$ and $k \in I$, the shuffle $\mathfrak{s}_k$ is produced by a leftmost shuffling at Time k. Meanwhile, we see from the assumptions $\mathfrak{s} \in \langle I'_{\triangleright}, J'_{\triangleleft}\rangle$ and $k \in J'$ that $\mathfrak{s}_k$ is also produced by an anti-leftmost shuffling. However, the two properties cannot occur simultaneously, and hence we are led to a contradiction. $\square$

Next, we recall that the union of two sets S_1 and S_2 is written as $S_1 \sqcup S_2$ once S_1 and S_2 are disjoint; this union is called a *disjoint union*.

Proposition 3.5. *Let I and J be disjoint subsets of $\mathbb{N}_+$. Then for any subset K of $\mathbb{N}_+$ that is disjoint from both I and J,*

$$\langle I_{\triangleright}, J_{\triangleleft}\rangle = \bigsqcup_{K' \subset K} \langle (I \sqcup K')_{\triangleright}, (J \sqcup (K \backslash K'))_{\triangleleft}\rangle. \tag{3.2}$$

In particular, if K is a singleton, namely, $K = \{k\}$ for a certain $k \in \mathbb{N}_+$, then

$$\langle I_{\triangleright}, J_{\triangleleft}\rangle = \langle (I \sqcup \{k\})_{\triangleright}, J_{\triangleleft}\rangle \sqcup \langle I_{\triangleright}, (J \sqcup \{k\})_{\triangleleft}\rangle. \tag{3.3}$$

Proof. We only need the fact that at an arbitrary Time $k \in K$, the event space of a free shuffling Щ_k is the disjoint union of the event spaces of a leftmost Щ_k and an anti-leftmost Щ_k. $\square$

4. Probabilities

In the previous section, we have determined the desired event spaces $\langle I_{\triangleright}, J_{\triangleleft}\rangle$ attached to the shuffling model $\text{Ш} = (\Omega, \mathbf{P})$ where $I, J \subset \mathbb{N}_+$ are disjoint. Now our task is to analyze the probability $\mathbf{P}_{\langle I_{\triangleright}, J_{\triangleleft}\rangle}$.

Recall that the model Ш is defined according to a Markov process. In particular, at Time k, we focus on a shuffling $\text{Щ}_k = \text{Щ}_k(\hat{\mathfrak{s}}_k)$, which is a probability space $(\Omega_{\hat{\mathfrak{s}}_k}, E_{\hat{\mathfrak{s}}_k}, \mathbf{P}^{(\hat{\mathfrak{s}}_k)})$ revolving around a certain event space $E_{\hat{\mathfrak{s}}_k}$, where $\mathfrak{s}_{k-1} = (s_0, \ldots, s_{k-1})$ is the outcome at Time $k-1$ and $\hat{\mathfrak{s}}_k = (s_0, \ldots, s_{k-1}, k)$. Note that if Щ_k is of one of the three types categorized in the previous section, the event space $E_{\hat{\mathfrak{s}}_k}$ is also decided accordingly. In light of the probability function (2.2), we compute that

$$\mathbf{P}^{(\hat{\mathfrak{s}}_k)}_{E_{\hat{\mathfrak{s}}_k}} = \begin{cases} q^k, & \text{if } \text{Щ}_k \text{ is leftmost,} \\ 1 - q^k, & \text{if } \text{Щ}_k \text{ is anti-leftmost,} \\ 1, & \text{if } \text{Щ}_k \text{ is free.} \end{cases}$$

In particular, this probability evaluation is *independent* of $\hat{\mathfrak{s}}_k$, or equivalently, the outcome $\mathfrak{s}_{k-1}$ at Time $k-1$.

This independence, together with the nature of a Markov process, immediately gives the following probability evaluation for the event space $\langle I_{\triangleright}, J_{\triangleleft}\rangle$ attached to the shuffling model Ш.

Proposition 4.1. *Let I and J be disjoint subsets of $\mathbb{N}_+$. Then*

$$\mathbf{P}_{\langle I_\triangleright, J_\triangleleft\rangle} = \prod_{i\in I} q^i \prod_{j\in J}(1-q^j). \tag{4.1}$$

Proposition 4.1 has two implications.

Corollary 4.2. *Let I and J be two disjoint subsets of $\mathbb{N}_+$ with I being infinite. Then*

$$\mathbf{P}_{\langle I_\triangleright, J_\triangleleft\rangle} = 0. \tag{4.2}$$

Corollary 4.3. *Let I and I' be two finite subsets of $\mathbb{N}_+$, not necessarily disjoint, and let J be a subset of $\mathbb{N}_+$ that is disjoint from both I and I'. Suppose that*

$$\sum_{i\in I} i = \sum_{i'\in I'} i'.$$

Then

$$\mathbf{P}_{\langle I_\triangleright, J_\triangleleft\rangle} = \mathbf{P}_{\langle I'_\triangleright, J_\triangleleft\rangle}. \tag{4.3}$$

Finally, we recall that, in a probability model, if two event spaces E_1 and E_2 are disjoint, then

$$\mathbf{P}_{E_1\sqcup E_2} = \mathbf{P}_{E_1} + \mathbf{P}_{E_2}.$$

Therefore, Proposition 3.5 implies the following relation.

Proposition 4.4. *Let I and J be disjoint subsets of $\mathbb{N}_+$. Then for any subset K of $\mathbb{N}_+$ that is disjoint from both I and J,*

$$\mathbf{P}_{\langle I_\triangleright, J_\triangleleft\rangle} = \sum_{K'\subset K} \mathbf{P}_{\langle (I\sqcup K')_\triangleright, (J\sqcup(K\backslash K'))_\triangleleft\rangle}. \tag{4.4}$$

In particular, if K is a singleton, namely, $K=\{k\}$ for a certain $k\in\mathbb{N}_+$, then

$$\mathbf{P}_{\langle I_\triangleright, J_\triangleleft\rangle} = \mathbf{P}_{\langle (I\sqcup\{k\})_\triangleright, J_\triangleleft\rangle} + \mathbf{P}_{\langle I_\triangleright, (J\sqcup\{k\})_\triangleleft\rangle}. \tag{4.5}$$

5. The pentagonal number theorem

The key to our probabilistic proof of the pentagonal number theorem is the following relation.

Theorem 5.1. *For every nonnegative integer N,*

$$\begin{aligned}\mathbf{P}_{\langle \varnothing_\triangleright, [1,\infty)_\triangleleft\rangle} = \sum_{n=0}^{N}(-1)^n\big(\mathbf{P}_{\langle [n+1,2n]_\triangleright, \varnothing_\triangleleft\rangle} - \mathbf{P}_{\langle [n+1,2n+1]_\triangleright, \varnothing_\triangleleft\rangle}\big)&\\ - (-1)^N \sum_{k\geq 1} \mathbf{P}_{\langle [k+1+N,k+1+2N]_\triangleright, [1+N,k+N]_\triangleleft\rangle}.&\end{aligned} \tag{5.1}$$

We show this result by induction on N.

Base case. We need to prove the case where $N = 0$:

$$\begin{aligned}\mathbf{P}_{\langle\varnothing_\triangleright,[1,\infty)_\triangleleft\rangle} &= \mathbf{P}_{\langle\varnothing_\triangleright,\varnothing_\triangleleft\rangle} - \mathbf{P}_{\langle\{1\}_\triangleright,\varnothing_\triangleleft\rangle} - \sum_{k\geq 1}\mathbf{P}_{\langle\{k+1\}_\triangleright,[1,k]_\triangleleft\rangle}\\ &= \mathbf{P}_{\langle\varnothing_\triangleright,\varnothing_\triangleleft\rangle} - \sum_{k\geq 1}\mathbf{P}_{\langle\{k\}_\triangleright,[1,k-1]_\triangleleft\rangle}.\end{aligned}$$

Note that $\left\langle\varnothing_\triangleright,\varnothing_\triangleleft\right\rangle$ contains all shuffles of $\mathbb{N}$ so $\left\langle\varnothing_\triangleright,\varnothing_\triangleleft\right\rangle = \Omega$. Meanwhile, given an arbitrary shuffle $\mathfrak{s} \in \Omega$, we project it onto each model Ш_k to get an outcome shuffle $\mathfrak{s}_k \in \Omega_k$ at Time k. Then either every $\mathfrak{s}_k$ is produced by an anti-leftmost shuffling (this case corresponds to $\left\langle\varnothing_\triangleright,[1,\infty)_\triangleleft\right\rangle$) or there exists a unique time k such that $\mathfrak{s}_k$ is the first shuffle produced by a leftmost shuffling (this case corresponds to $\left\langle\{k\}_\triangleright,[1,k-1]_\triangleleft\right\rangle$). Furthermore, given two different positive integers k and k', we assume without loss of generality that $k < k'$. Then $k \in \{k\}$ and $k \in [1, k'-1]$. In light of Proposition 3.4,

$$\left\langle\{k\}_\triangleright,[1,k-1]_\triangleleft\right\rangle \cap \left\langle\{k'\}_\triangleright,[1,k'-1]_\triangleleft\right\rangle = \varnothing.$$

Hence,

$$\left\langle\varnothing_\triangleright,\varnothing_\triangleleft\right\rangle = \left\langle\varnothing_\triangleright,[1,\infty)_\triangleleft\right\rangle \sqcup \left(\bigsqcup_{k\geq 1}\left\langle\{k\}_\triangleright,[1,k-1]_\triangleleft\right\rangle\right).$$

Application of Proposition 4.4 leads us to the required relation. □

Inductive step. It suffices to show that, for each nonnegative integer N,

$$\begin{aligned}\sum_{k\geq 1}\mathbf{P}_{\langle[k+1+N,k+1+2N]_\triangleright,[1+N,k+N]_\triangleleft\rangle} &= \mathbf{P}_{\langle[2+N,2+2N]_\triangleright,\varnothing_\triangleleft\rangle} - \mathbf{P}_{\langle[2+N,3+2N]_\triangleright,\varnothing_\triangleleft\rangle}\\ &\quad - \sum_{k\geq 1}\mathbf{P}_{\langle[k+2+N,k+3+2N]_\triangleright,[2+N,k+1+N]_\triangleleft\rangle}.\end{aligned}$$

We start by noting that, for each $k \geq 1$, the sets $[k+1+N, k+1+2N]$, $[2+N, k+N]$ and $\{1 + N\}$ are pairwise disjoint. It follows from Proposition 4.4 that

$$\begin{aligned}\sum_{k\geq 1}\mathbf{P}_{\langle[k+1+N,k+1+2N]_\triangleright,[1+N,k+N]_\triangleleft\rangle} &= \sum_{k\geq 1}\mathbf{P}_{\langle[k+1+N,k+1+2N]_\triangleright,[2+N,k+N]_\triangleleft\rangle}\\ &\quad - \sum_{k\geq 1}\mathbf{P}_{\langle(\{1+N\}\sqcup[k+1+N,k+1+2N])_\triangleright,[2+N,k+N]_\triangleleft\rangle}.\end{aligned}$$

We split the first sum on the right-hand side as

$$\sum_{k\geq 1}\mathbf{P}_{\langle [k+1+N,k+1+2N]_\triangleright,[2+N,k+N]_\triangleleft\rangle} = \mathbf{P}_{\langle [2+N,2+2N]_\triangleright,\varnothing_\triangleleft\rangle} + \sum_{k\geq 1}\mathbf{P}_{\langle [k+2+N,k+2+2N]_\triangleright,[2+N,k+1+N]_\triangleleft\rangle}.$$

We further note that the sets $[k+2+N, k+2+2N]$, $[2+N, k+N]$ and $\{k+1+N\}$ are pairwise disjoint. Another application of Proposition 4.4 yields

$$\sum_{k\geq 1}\mathbf{P}_{\langle [k+2+N,k+2+2N]_\triangleright,[2+N,k+1+N]_\triangleleft\rangle} = \sum_{k\geq 1}\mathbf{P}_{\langle [k+2+N,k+2+2N]_\triangleright,[2+N,k+N]_\triangleleft\rangle} - \sum_{k\geq 1}\mathbf{P}_{\langle [k+1+N,k+2+2N]_\triangleright,[2+N,k+N]_\triangleleft\rangle}.$$

Finally, we see from Corollary 4.3 that

$$\sum_{k\geq 1}\mathbf{P}_{\langle (\{1+N\}\sqcup[k+1+N,k+1+2N])_\triangleright,[2+N,k+N]_\triangleleft\rangle} = \sum_{k\geq 1}\mathbf{P}_{\langle [k+2+N,k+2+2N]_\triangleright,[2+N,k+N]_\triangleleft\rangle}.$$

This implies that

$$\sum_{k\geq 1}\mathbf{P}_{\langle [k+1+N,k+1+2N]_\triangleright,[1+N,k+N]_\triangleleft\rangle} = \mathbf{P}_{\langle [2+N,2+2N]_\triangleright,\varnothing_\triangleleft\rangle} - \sum_{k\geq 1}\mathbf{P}_{\langle [k+1+N,k+2+2N]_\triangleright,[2+N,k+N]_\triangleleft\rangle}.$$

By singling out the $k = 1$ term of the sum in the above expression, we obtain the desired part $\mathbf{P}_{\langle [2+N,3+2N]_\triangleright,\varnothing_\triangleleft\rangle}$. □

Now we are ready to conclude the pentagonal number theorem.

Proof of Theorem 1.1. Letting $N \to \infty$ in (5.1) and recalling Corollary 4.2, we have

$$\mathbf{P}_{\langle \varnothing_\triangleright,[1,\infty)_\triangleleft\rangle} = \sum_{n\geq 0}(-1)^n\big(\mathbf{P}_{\langle [n+1,2n]_\triangleright,\varnothing_\triangleleft\rangle} - \mathbf{P}_{\langle [n+1,2n+1]_\triangleright,\varnothing_\triangleleft\rangle}\big).$$

Application of Proposition 4.1 yields

$$\prod_{j\geq 1}(1-q^j) = \sum_{n\geq 0}(-1)^n\big(q^{\frac{n(3n+1)}{2}} - q^{\frac{(n+1)(3n+2)}{2}}\big).$$

This is exactly (1.1). □

6. Epilogue

Before putting down the curtain, let us flash back to Euler's original proof of the pentagonal number theorem in [3]. Adopting the *q-Pochhammer symbol* for $n \in \mathbb{N} \cup \{\infty\}$,

$$(A;q)_n := \prod_{j=0}^{n-1}(1 - Aq^j),$$

the key observation of Euler, although not explicitly stated, is the following.

Theorem 6.1. *For every nonnegative integer N,*

$$\begin{aligned}(q;q)_\infty = &\sum_{n=0}^{N}(-1)^n\big(q^{\frac{n(3n+1)}{2}} - q^{\frac{(n+1)(3n+2)}{2}}\big) \\ &- (-1)^N \sum_{k\geq 1} q^{\frac{(1+N)(2k+2+3N)}{2}}(q^{1+N};q)_k.\end{aligned} \tag{6.1}$$

It is worth noting that our proof can be understood as a translation of Euler's proof into the language of classical probability theory. To tie the knot, we simply apply Proposition 4.1 and derive that

$$\begin{aligned}\mathbf{P}_{\langle [k+1+N,k+1+2N]_\triangleright,[1+N,k+N]_\triangleleft\rangle} &= \prod_{i=k+1+N}^{k+1+2N} q^i \prod_{j=1+N}^{k+N}(1-q^j) \\ &= q^{\frac{(1+N)(2k+2+3N)}{2}}(q^{1+N};q)_k.\end{aligned}$$

In view of (5.1), this is clearly the last piece of expression required for (6.1).

Acknowledgements. This work was supported by the Austrian Science Fund (No. 10.55776/F1002). I would like to thank the anonymous referee for the inspiring suggestions that have improved the quality of this paper to a great extent. I am also grateful to Hugh Thomas for the comments on Proposition 4.4 in an earlier draft.

References

1. G. E. Andrews, Euler's pentagonal number theorem, *Math. Mag.* **56** (1983), no. 5, 279–284.
2. J. Bell, A summary of Euler's work on the pentagonal number theorem, *Arch. Hist. Exact Sci.* **64** (2010), no. 3, 301–373.
3. L. Euler, Demonstratio theorematis circa ordinem in summis divisorum observatum, in: *Novi Commentarii Academiae Scientiarum Imperialis Petropolitanae, Tom. V*, 75–83, Academiae Scientiarvm, Petropoli, 1760.

4. L. Euler, *Opera omnia. Series Quarta A: Commercium epistolicum. Vol. IV: Leonhardi Euleri commercium epistolicum cum Christiano Goldbach*, Birkhäuser, Basel, 2015.
5. L. Euler, *Opera omnia. Series Quarta A: Commercium epistolicum. Vol. III: Briefwechsel mit Daniel Bernoulli. Teil I: Einleitungen, Briefwechsel 1726–1743*, Birkhäuser, Cham, 2016.
6. F. Franklin, Sur le développement du produit infini $(1-x)(1-x^2)(1-x^3)(1-x^4)\ldots$, *C. R. Acad. Sci. Paris Sér. A* **92** (1881), 448–450.
7. L. V. Goncharova, The cohomologies of Lie algebras of formal vector fields on the line, *Funct. Anal. Appl.* **7** (1973), no. 2, 91–97.
8. A. M. Legendre, *Théorie des nombres. Vol. II. Third edition*, Chez Firmin Didot Frère, Paris, 1830. (Reprinted: Blanchard, Paris, 1955).

Fakultät für Mathematik, Universität Wien, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria

Email address: `chenxiaohang92@gmail.com, xiaohangc92@univie.ac.at`