Tracing the Constant e: From Derangements to Transcendence

Cici Chen

March 2026

1 Introduction

The constant e is one of the most important numbers in mathematics, yet it is often only briefly introduced. During a recent lesson on exponential and logarithmic functions, my teacher's introduction to e inspired me to explore its origin and significance more deeply.

The presence of e may seem coincidental, yet appears with remarkable consistency. Why does this number arise so naturally? Is it merely a coincidence, or does it reflect a deeper mathematical structure?

In this paper, we follow the appearance of e from modern problems to its historical origins and deeper properties. Through this journey, we glimpse the collective brilliance of mathematical thought developed across generations. This paper aims to show that e is not accidental, but an inevitable consequence of fundamental mathematical structures.

2 Why does 1/e appear in the hat-check problem?

Imagine a scenario where there is a lively party with n guests. Each puts their hat in a central basket and then randomly picks one hat from the basket. The game seems simple, yet it hides a surprising mathematical fact: as the number of guests increases, the probability that no one gets their own hat does not approach zero; instead, it stabilizes around a constant of approximately 36.79%. Why is this the case?

Let there be n guests at the party, and let A_n denote the total number of derangements (no guest ends up with their own hat).

First, the nth guest cannot take their own hat, so there are $n-1$ choices for him to pick a hat belonging to someone else. Without loss of generality, assume that the nth guest takes the kth hat ($k \neq n$). We now divide this into two cases:

1. If the kth guest's hat is exactly the one taken by the nth guest, the remaining $n-2$ guests only need to form a derangement among themselves. The number of derangements in this case is A_{n-2}.
2. If the kth guest's hat is *not* taken by the nth guest, we need to form a derangement among the $n-1$ guests excluding the nth guest. The number of derangements in this case is A_{n-1}.

We have the recurrence relation for derangements:

$$A_n = (n-1)(A_{n-1} + A_{n-2}), \quad n \geq 3$$

with initial conditions $A_1 = 0,\ A_2 = 1,\ A_3 = 2$.

Dividing both sides by $n!$:

$$\frac{A_n}{n!} = \frac{(n-1)A_{n-1}}{n!} + \frac{(n-1)A_{n-2}}{n!}$$

Let $b_n = \frac{A_n}{n!}$, then $b_1 = 0,\ b_2 = \frac{1}{2},\ b_3 = \frac{1}{3}$.

$$nb_n = (n-1)b_{n-1} + b_{n-2}$$

$$b_n = \frac{n-1}{n}b_{n-1} + \frac{1}{n}b_{n-2}$$

$$b_n = b_{n-1} + \frac{1}{n}\left(b_{n-2} - b_{n-1}\right)$$

$$\frac{b_n - b_{n-1}}{b_{n-1} - b_{n-2}} = -\frac{1}{n}$$

By repeated substitution, we obtain:

$$\frac{b_{n-1} - b_{n-2}}{b_{n-2} - b_{n-3}} = -\frac{1}{n-1}, \quad \frac{b_{n-2} - b_{n-3}}{b_{n-3} - b_{n-4}} = -\frac{1}{n-2}, \quad \ldots, \quad \frac{b_3 - b_2}{b_2 - b_1} = -\frac{1}{3}$$

Multiplying all the equations:

$$\frac{b_n - b_{n-1}}{b_{n-1} - b_{n-2}} \cdot \frac{b_{n-1} - b_{n-2}}{b_{n-2} - b_{n-3}} \cdots \frac{b_3 - b_2}{b_2 - b_1} = (-1)^{n-2} \cdot \frac{1}{n} \cdot \frac{1}{n-1} \cdots \frac{1}{3}$$

$$\frac{b_n - b_{n-1}}{b_2 - b_1} = (-1)^{n-2} \cdot \frac{1}{n(n-1)\cdots 3}$$

Substituting the initial values $b_1 = 0,\ b_2 = \frac{1}{2}$, (so $b_2 - b_1 = \frac{1}{2}$):

$$b_n - b_{n-1} = (-1)^n \cdot \frac{1}{n!}$$

$$b_n - b_{n-1} = (-1)^n \cdot \frac{1}{n!}, \quad b_{n-1} - b_{n-2} = (-1)^{n-1} \cdot \frac{1}{(n-1)!}, \quad \ldots, \quad b_2 - b_1 = (-1)^2 \cdot \frac{1}{2!}$$

Summing up all the equations:

$$b_n - b_1 = \sum_{k=2}^{n} \frac{(-1)^k}{k!}$$

Since $b_1 = 0$, we finally obtain:

$$b_n = \sum_{k=2}^{n} \frac{(-1)^k}{k!}$$

$$\sum_{k=0}^{n} \frac{(-1)^k}{k!} = \sum_{k=2}^{n} \frac{(-1)^k}{k!}$$

$$b_n = \sum_{k=0}^{n} \frac{(-1)^k}{k!}$$

$$n = 1 :\ b_1 = 0, \quad n = 2 :\ b_2 = \frac{1}{2}$$

Therefore,

$$\forall n \geq 1, \quad b_n = \sum_{k=0}^{n} \frac{(-1)^k}{k!}$$

$$a_n = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}$$

The number of derangements of n hats is

$$n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

As the total number of permutations is $n!$. Therefore, the probability that no person gets their own hat is

$$\frac{n!\sum_{k=0}^{n}\frac{(-1)^k}{k!}}{n!}=\sum_{k=0}^{n}\frac{(-1)^k}{k!}.$$

When $n\to\infty$, for the series $\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}$, let $a_k=\frac{1}{k!}$.

For $\forall k\geq 0$, we have $k!<(k+1)!$, so $a_k>a_{k+1}$, which means $\{a_k\}$ is strictly decreasing.

Furthermore, $\lim_{k\to\infty} a_k=\lim_{k\to\infty}\frac{1}{k!}=0$.

By the Leibniz test, the series $\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}$ converges, so the limit of the series exists. The Maclaurin series expansion of the exponential function is

$$e^x=\sum_{k=0}^{\infty}\frac{x^k}{k!}.$$

Setting $x=-1$, we obtain:

$$e^{-1}=\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}=\frac{1}{e}\approx 0.367879441.$$

This is the probability that no one receives their own hat in the derangement problem, which is approximately 36.79%.

3 Why does $1/e$ appear in the optimal blind date strategy?

Suppose there are 100 blind dates. You have to date them one by one and confirm or reject them on the spot.

What is the probability of choosing the best candidate?
What strategy should you use?

For blind date selection, everyone has an internal standard based on appearance, education, family background, etc. A candidate above this standard is considered satisfactory. This standard is also determined by the overall quality of the candidate pool.

Therefore, the optimal strategy is: first date the first r candidates and reject all of them to establish this internal standard. After that, as soon as the strongest candidate appears, select them immediately.

Let there be n candidates in total, and let the globally best candidate be at position k, so $k \geq r+1$.

Let P be the probability of selecting the globally optimal candidate.

To successfully choose the kth person, two conditions must be met:

1. The kth person is globally optimal, so the probability is $\frac{1}{n}$.
2. The second-best candidate globally appears in the first r candidates. If the second-best candidate is between $r+1$ and $k-1$, you will select that person and will not wait for the kth person. The probability of this condition is $\frac{r}{k-1}$.

Therefore,

$$P = \sum_{k=r+1}^{n} \frac{1}{n} \cdot \frac{r}{k-1} = \frac{r}{n} \sum_{k=r+1}^{n} \frac{1}{k-1}$$

where r and n are constants. Let $t = k-1$, so $k = t+1$.

The original expression becomes:

$$\frac{r}{n} \sum_{k=r+1}^{n} \frac{1}{k-1} = \frac{r}{n} \sum_{t=r}^{n} \frac{1}{t}$$

$$\sum_{t=r}^{n} \frac{1}{t} = \sum_{t=1}^{n} \frac{1}{t} - \sum_{t=1}^{r} \frac{1}{t}$$

Using the approximation for the harmonic series:

$$\sum_{t=1}^{n} \frac{1}{t} = \ln n + \gamma$$

where γ is the Euler-Mascheroni constant, with $\gamma \approx 0.5772$.

$$\sum_{t=r}^{n} \frac{1}{t} = (\ln n + \gamma) - (\ln r + \gamma) = \ln n - \ln r$$

Thus,

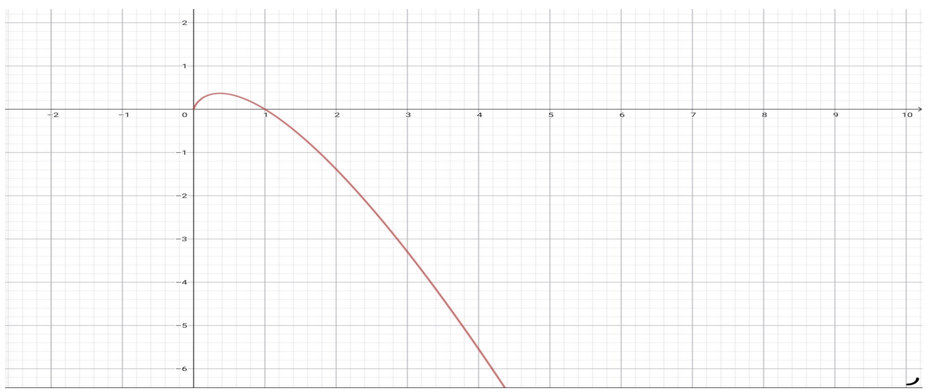
$$P = \frac{r}{n}(\ln n - \ln r) = \frac{r}{n} \ln \frac{n}{r}$$

Let X be the ratio of the exploration sample r (the number of initial candidates) to the total number of candidates:

$$X = \frac{r}{n}$$

Thus,

$$P(X) = X \ln \frac{1}{X} = -X \ln X$$



This is the graph of $P(X) = -X \ln X$. When $P(X)$ reaches its maximum value, $P'(X) = 0$.

$$P'(X) = -\ln X - 1.$$

Setting $P'(X) = 0$, we get $X = \frac{1}{e}$. The maximum value

$$P_{\max} = -\frac{1}{e} \ln\left(\frac{1}{e}\right) = \frac{1}{e}.$$

Therefore, when the total number of candidates $n = 100$, we need to date $100 \times \frac{1}{e} \approx 37$ candidates first. After that, we should accept the first candidate who is better than all the previous ones, and the probability of successfully selecting the best candidate is $\frac{1}{e} \approx 36.79\%$.

The theoretical prototype of this problem is the **Secretary Problem**, and the **37% Rule** is also known as **Optimal Stopping**

Theory. It is applied to maximize benefits in scenarios featuring uncertainty, irreversibility and limited information.

This theory can be used in many aspects of life, such as online shopping, job hunting, house hunting and finding a partner. This strategy consists of two steps:

1. Use the first 37% of the total sample as a reference pool, and take the best option among them as the benchmark.
2. If an option better than the benchmark appears in the remaining 63% of the sample, choose it decisively.

What Optimal Stopping Theory reveals is that life is made up of a series of choices. It teaches us to be brave in exploring, yet also dare to make decisions at the right moment and accept all the consequences that follow.

4 In the Name of Euler: Retracing the Origins of e

From the previous two chapters, we have seen $\frac{1}{e}$ appear in both combinatorics and decision-making mathematics in a way that is both coincidental and almost fateful. This naturally leads us to ask: where does the constant e come from, and why does it appear in such a unified form across different areas of mathematics? We will now turn to the work of Euler, and see how he systematically uncovered the nature and significance of e starting from the derivative of the logarithmic function.

The 18th century was an era of rapid development in calculus, and Euler was also engaged in its study. One day, as he looked at the table below, he wondered: what is the antiderivative of x^{-1}?

$f(x)$	x^5	x^4	x^3	x^2	x	?
$f'(x)$	$5x^4$	$4x^3$	$3x^2$	$2x$	1	x^{-1}

At the current stage of mathematical development, there was no known primitive function whose derivative is $1/x$. Euler found this extremely unsatisfactory, so he was determined to find out what its primitive function was. He then attempted the following method:

Let $F_n(x)$ be an antiderivative of $\frac{1}{x}$.

$$F_n'(x) = \frac{1}{x}.$$

Let $y = F_n(x)$. Its inverse function is given by

$$x = F_n^{-1}(y).$$

$$F_n'(x) = \frac{dy}{dx}$$

$$\left(F_n^{-1}(y)\right)' = \frac{dx}{dy}$$

$$\left(F_n^{-1}(y)\right)' = \frac{1}{F_n'(x)} = \frac{1}{1/x} = x = F_n^{-1}(y)$$

Therefore, the derivative of the inverse function $F_n^{-1}(y)$ is itself.

$f(x)$	1^x	1.5^x	2^x	2.5^x	3^x	3.5^x
$f'(x)$	0	0.41×1.5^x	0.69×2^x	0.92×2.5^x	1.10×3^x	1.25×3.5^x

This is a table of exponential functions.

From the table, Euler observes that for any exponential function $y = a^x$, its derivative can be expressed as

$$\frac{dy}{dx} = C_a \cdot a^x,$$

where C_a is a constant related to the base a. Euler notices that when $a \in (2.5, 3)$, there must exist a base for which the constant $C_a = 1$, so that the function is equal to its own derivative — that is what Euler was trying to find.

Since the desired function is an exponential function, it must satisfy

$$F_n^{-1}(0) = 1$$

Since all higher-order derivatives of $F_n^{-1}(y)$ are equal to itself:

$$\left(F_n^{-1}\right)^{(k)}(y) = F_n^{-1}(y) \quad (\forall k \geq 1)$$

Therefore, the values at $y = 0$ are

$$\left(F_n^{-1}\right)^{(k)}(0) = \left(F_n^{-1}\right)^{(k-1)}(0) = \cdots = \left(F_n^{-1}\right)'(0) = F_n^{-1}(0) = 1.$$

Euler then uses the Maclaurin series formula:

$$F_n^{-1}(y) = \sum_{k=0}^{\infty} \frac{\left(F_n^{-1}\right)^{(k)}(0)}{k!} y^k,$$

which gives

$$F_n^{-1}(y) = \frac{\left(F_n^{-1}\right)^{(0)}(0)}{0!}y^0 + \frac{\left(F_n^{-1}\right)^{(1)}(0)}{1!}y^1 + \frac{\left(F_n^{-1}\right)^{(2)}(0)}{2!}y^2 + \cdots + \frac{\left(F_n^{-1}\right)^{(n)}(0)}{n!}y^n + \cdots$$
$$= \frac{1}{1}\cdot 1 + \frac{1}{1}\cdot y + \frac{1}{2!}\cdot y^2 + \frac{1}{3!}\cdot y^3 + \cdots + \frac{1}{n!}\cdot y^n + \cdots$$
$$= 1 + y + \frac{y^2}{2!} + \frac{y^3}{3!} + \cdots + \frac{y^n}{n!} + \cdots.$$

The Maclaurin series gives an infinite polynomial that approximates and eventually equals the original function. Since this function is the exponential function, Euler used the letter e (the first letter of 'exponential') to represent it; some also say e honors Euler himself. Thus, the function $F_n^{-1}(y)$ is denoted e^y.

Euler then substitutes $y = 1$ into $F_n^{-1}(y) = e^y$:

$$F_n^{-1}(1) = e^1 = e = 1+1+\frac{1^2}{2!}+\frac{1^3}{3!}+\frac{1^4}{4!}+\cdots+\frac{1^n}{n!}+\cdots \approx 2.718281828\cdots$$

We have

$$F_n'(x) = \frac{1}{x}.$$

Then

$$F_n'(ax) = \frac{1}{ax}\cdot a = \frac{1}{x},$$

so

$$F_n'(ax) = F_n'(x).$$

If two functions have the same derivative, they differ only by a constant. Thus,

$$F_n(ax) = F_n(x) + C$$
$$F_n(ax) - F_n(x) = C$$

Euler realized that this is essentially characteristic of the logarithmic function, since the logarithm follows the identity:

$$\log(ax) - \log(x) = \log(a) + \log(x) - \log(x) = \log(a) = C$$

Thus, we define $F_n(x) = \log_?(x)$. Since $F_n^{-1}(1) = e$, we infer that

$$F_n(e) = \log_?(e) = 1 \implies ? = e.$$

Therefore,

$$F_n(x) = \log_e(x) = \ln(x),$$

$$F_n^{-1}(x) = e^x.$$

In conclusion,

$$(\ln x)' = \frac{1}{x}.$$

5 Proving e Exists: The Limit Definition

In the previous chapter, we traced Euler's discovery of the constant e through the lens of calculus. Starting from the derivative of the logarithmic function, he derived the infinite series

$$e = \sum_{k=0}^{\infty} \frac{1}{k!} = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots,$$

which converges to approximately $2.71828\ldots$. Yet this was not the first appearance of e in mathematical history. Decades earlier, Jacob Bernoulli had investigated the problem of compound interest and encountered the limit

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n.$$

He observed that this expression is increasing yet bounded; it approaches a finite value between 2 and 3. Bernoulli, however, neither proved the existence of this limit nor established a connection between it and the series later derived by Euler. In this chapter, we fill this gap: we first prove that the sequence $a_n = \left(1 + \frac{1}{n}\right)^n$ is monotonically increasing and bounded above, and thus convergent. We then show that its limit coincides with the sum of the series $\sum_{k=0}^{\infty} \frac{1}{k!}$. This establishes the equivalence of the two classic definitions of e and gives us a deeper understanding of why e appears so naturally in mathematics.

Let $x_n = \left(1 + \frac{1}{n}\right)^n$. We aim to prove that $\{x_n\}$ is a monotonically increasing sequence and bounded above. By the Monotone Convergence Theorem, this will guarantee the existence of its limit.

Firstly, we consider the case when $n \to +\infty$.

By the binomial expansion,

$$x_n = \binom{n}{0}\left(\frac{1}{n}\right)^0 + \binom{n}{1}\left(\frac{1}{n}\right)^1 + \binom{n}{2}\left(\frac{1}{n}\right)^2 + \cdots + \binom{n}{n}\left(\frac{1}{n}\right)^n$$

$$
\begin{aligned}
x_n &= 1 + \frac{n}{1!} \cdot \frac{1}{n} + \frac{n(n-1)}{2!} \cdot \frac{1}{n^2} + \frac{n(n-1)(n-2)}{3!} \cdot \frac{1}{n^3} + \dots + \frac{n(n-1)\cdots(n-n+1)}{n!} \cdot \frac{1}{n^n} \\
&= 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n}\right) + \frac{1}{3!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) + \dots \\
&\quad + \frac{1}{n!}\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{n-1}{n}\right).
\end{aligned}
$$

$$
\begin{aligned}
x_{n+1} &= \left(1 + \frac{1}{n+1}\right)^{n+1} \\
&= 1 + 1 + \frac{1}{2!}\left(1 - \frac{1}{n+1}\right) + \frac{1}{3!}\left(1 - \frac{1}{n+1}\right)\left(1 - \frac{2}{n+1}\right) + \dots \\
&\quad + \frac{1}{n!}\left(1 - \frac{1}{n+1}\right)\left(1 - \frac{2}{n+1}\right)\cdots\left(1 - \frac{n-1}{n+1}\right) \\
&\quad + \frac{1}{(n+1)!}\left(1 - \frac{1}{n+1}\right)\left(1 - \frac{2}{n+1}\right)\cdots\left(1 - \frac{n}{n+1}\right).
\end{aligned}
$$

Because $\frac{1}{n} > \frac{1}{n+1}$, we have $1 - \frac{1}{n} < 1 - \frac{1}{n+1}$. Hence, all terms of x_{n+1} from the third term onward exceed those of x_n. We conclude $x_{n+1} > x_n$, so $\{x_n\}$ is monotonically increasing.

Since $1 - \frac{1}{n} < 1$, we have $\frac{1}{n!}\left(1 - \frac{1}{n}\right) < \frac{1}{n!}$. Therefore,

$$
x_n < 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!}.
$$

For $n \geq 2$, $n! > 2^{n-1}$ implies $\frac{1}{n!} < \frac{1}{2^{n-1}}$. Thus,

$$
x_n < 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^{n-1}}.
$$

We know that for a geometric series,

$$
1 + \frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^{n-1}} = 2 - \frac{1}{2^{n-1}}.
$$

Therefore,

$$
x_n < 1 + \left(2 - \frac{1}{2^{n-1}}\right) = 3 - \frac{1}{2^{n-1}} < 3,
$$

so the sequence $\{x_n\}$ is bounded above.

When $n \to -\infty$, let $n = -(t+1)$. Then $t \to +\infty$.

$$\begin{aligned}
\lim_{n\to-\infty}\left(1+\frac{1}{n}\right)^{n} &= \lim_{t\to+\infty}\left(1+\frac{1}{-(t+1)}\right)^{-(t+1)} \\
&= \lim_{t\to+\infty}\left(\frac{t}{t+1}\right)^{-(t+1)} \\
&= \lim_{t\to+\infty}\left(\frac{t+1}{t}\right)^{t+1} \\
&= \lim_{t\to+\infty}\left(1+\frac{1}{t}\right)^{t}\cdot\left(1+\frac{1}{t}\right) \\
&= \lim_{t\to+\infty}\left(1+\frac{1}{t}\right)^{t} \approx 2.718281828\ldots = e.
\end{aligned}$$

In conclusion, we have shown that

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n} = e \approx 2.718281828\ldots,$$

which rigorously establishes the existence of the constant e. This result also demonstrates the equivalence of two fundamental definitions of e: the limit of the sequence $\left(1+\frac{1}{n}\right)^{n}$ and the Maclaurin series $\sum_{k=0}^{\infty}\frac{1}{k!}$, as both converge to the same positive real number.

6 The Transcendence of e: A Number Beyond Algebra

In the previous two chapters, we saw that $e \approx 2.71828$, yet these results are merely analytical and computational: we know e's approximate value but not its place in the number system hierarchy.

In 1737, Euler first proved e irrational—its first unique property. Irrationality is common, though; $\sqrt{2}$ is also irrational. A deeper question arises: is e a root of any integer-coefficient polynomial? If not, e is transcendental, one of the most "mysterious" real numbers.
It took over a century to answer this: in 1873, Hermite used sophisticated analysis to prove e's transcendence.

Proof: The Transcendence of e
To prove that e is transcendental, we assume for contradiction that there exists a non-zero polynomial with integer coefficients

$$c_m x^m + c_{m-1}x^{m-1} + \cdots + c_0 = 0$$

such that $x = e$ is a root.

Let $f(x)$ be an arbitrary polynomial, and consider the integral

$$\int_0^b f(x)e^{-x}\,dx.$$

We apply the integration by parts formula:

$$\int u\,dv = uv - \int v\,du,$$

Let $u = f(x)$, $dv = e^{-x}dx$, so that $du = f'(x)dx$, $v = -e^{-x}$.

First integration by parts:

$$\int_0^b f(x)e^{-x}dx = -\left[f(x)e^{-x}\right]_0^b + \int_0^b e^{-x}f'(x)dx.$$

Second integration by parts:

$$= -\left[f(x)e^{-x}\right]_0^b - \left[f'(x)e^{-x}\right]_0^b + \int_0^b f''(x)e^{-x}dx.$$

Third integration by parts:

$$= -\left[f(x)e^{-x}\right]_0^b - \left[f'(x)e^{-x}\right]_0^b - \left[f''(x)e^{-x}\right]_0^b + \int_0^b f'''(x)e^{-x}dx.$$

Continuing this process for $n+1$ times (where $n = \deg f$):

$$\int_0^b f(x)e^{-x}dx = -\left[f(x)e^{-x}\right]_0^b - \left[f'(x)e^{-x}\right]_0^b - \cdots - \left[f^{(n)}(x)e^{-x}\right]_0^b + \int_0^b f^{(n+1)}(x)e^{-x}dx.$$

Since $f(x)$ is a polynomial of degree n, we have $f^{(n+1)}(x) = 0$. Therefore,

$$\int_0^b f(x)e^{-x}dx = -\left[f(x)e^{-x}\right]_0^b - \left[f'(x)e^{-x}\right]_0^b - \cdots - \left[f^{(n)}(x)e^{-x}\right]_0^b.$$

Define

$$F(x) = f(x) + f'(x) + \cdots + f^{(n)}(x).$$

Then

$$\int_0^b f(x)e^{-x}dx = -\left[F(x)e^{-x}\right]_0^b = -\left(F(b)e^{-b} - F(0)\right) = -F(b)e^{-b} + F(0).$$

Multiplying both sides by e^b, we obtain

$$e^b \int_0^b f(x)e^{-x}dx = -F(b) + F(0)e^b,$$

$$F(b) = F(0)e^b - e^b \int_0^b f(x)e^{-x}dx.$$

If e is a root of the integer-coefficient polynomial

$$\sum_{k=0}^{m} c_k x^k = 0 \quad (c_0, c_1, \ldots, c_m \in \mathbb{Z}),$$

then for $k = 0, 1, \ldots, m$, we have

$$\sum_{k=0}^{m} c_k F(k) = \sum_{k=0}^{m} c_k \left[F(0)e^k - e^k \int_0^k f(x)e^{-x}dx \right].$$

$$\sum_{k=0}^{m} c_k F(k) = \sum_{k=0}^{m} c_k F(0)e^k - \sum_{k=0}^{m} c_k e^k \int_0^k f(x)e^{-x}dx.$$

Since $\sum_{k=0}^{m} c_k e^k = 0$, we obtain

$$\sum_{k=0}^{m} c_k F(k) = -\sum_{k=0}^{m} c_k e^k \int_0^k f(x)e^{-x}dx.$$

$$\left| \sum_{k=0}^{m} c_k F(k) \right| = \left| -\sum_{k=0}^{m} c_k e^k \int_0^k f(x)e^{-x}dx \right| = \left| \sum_{k=0}^{m} c_k e^k \int_0^k f(x)e^{-x}dx \right|.$$

By the triangle inequality $|X + Y + \cdots| \leq |X| + |Y| + \cdots$, we get

$$\left| \sum_{k=0}^{m} c_k e^k \int_0^k f(x)e^{-x}dx \right| \leq \sum_{k=0}^{m} \left| c_k e^k \int_0^k |f(x)|e^{-x}dx \right|. \qquad (1)$$

Let

$$f(x) = \frac{1}{(p-1)!} x^{p-1}(x-1)^p(x-2)^p \cdots (x-m)^p, \quad x \in [0, m],$$

where $p > m$.

When $x \in [0, m]$, we have $|x| \leq m, |x-1| \leq m, \ldots, |x-m| \leq m$, therefore,

$$|f(x)| \leq \frac{1}{(p-1)!} \cdot m^{p-1} \cdot \underbrace{m^p \cdot m^p \cdots m^p}_{m \text{ terms}} = \frac{1}{(p-1)!} m^{(p-1)+mp} = \frac{1}{(p-1)!} m^{(m+1)p-1}.$$

Since the highest degree term of $f(x)$ is $(m+1)p-1$, we have

$$F(x) = f(x) + f'(x) + f''(x) + \cdots + f^{((m+1)p-1)}(x).$$

Therefore,

$$(1) \leq \sum_{k=0}^{m} \left| c_k e^k \cdot \frac{m^{(m+1)p-1}}{(p-1)!} \int_0^k e^{-x} dx \right|.$$

$$\int_0^k e^{-x} dx = \left[-e^{-x} \right]_0^k = -e^{-k} + e^0 = 1 - e^{-k}.$$

Since $k \geq 0$, we have $e^{-k} \geq 0$, therefore,

$$1 - e^{-k} \leq 1 \implies \int_0^k e^{-x} dx \leq 1.$$

$$(1) \leq \sum_{k=0}^{m} \left| c_k e^k \cdot \frac{m^{(m+1)p-1}}{(p-1)!} \int_0^k e^{-x} dx \right| \leq \sum_{k=0}^{m} \left| c_k e^m \cdot \frac{m^{(m+1)p-1}}{(p-1)!} \right|.$$

Since factorial growth is much faster than exponential growth, as $p \to +\infty$,

$$\frac{m^{(m+1)p-1}}{(p-1)!} \to 0.$$

Therefore, for sufficiently large p,

$$\sum_{k=0}^{m} \left| c_k e^m \cdot \frac{m^{(m+1)p-1}}{(p-1)!} \right| < 1,$$

which implies

$$\left| \sum_{k=0}^{m} c_k F(k) \right| < 1.$$

Now we prove:

$$\left|\sum_{k=0}^{m} c_k F(k)\right| \geq 1.$$

$$f(x)=\frac{1}{(p-1)!}x^{p-1}(x-1)^p(x-2)^p\cdots(x-m)^p, \quad x\in[0,m],\ p>m.$$

When $x=1,2,\ldots,m$, we have $f(x)=0$.

$$\begin{aligned} f'(x)=\frac{1}{(p-1)!}&\left\{(p-1)x^{p-2}\cdot(x-1)^p(x-2)^p\cdots(x-m)^p\right.\\ &\left.+x^{p-1}\cdot[(x-1)^p(x-2)^p\cdots(x-m)^p]'\right\}. \end{aligned}$$

Applying the product rule, we have

$[(x-1)^p(x-2)^p\cdots(x-m)^p]' = [(x-1)^p]'\cdot(x-2)^p\cdots(x-m)^p + (x-1)^p\cdot[(x-2)^p\cdots(x-m)^p]'$

$= p(x-1)^{p-1}\cdot(x-2)^p\cdots(x-m)^p + (x-1)^p\cdot[(x-2)^p\cdots(x-m)^p]'$.

When $x=1$, $(x-1)^p\cdot[(x-2)^p\cdots(x-m)^p]'=0$, so $f'(x)=0$.

In fact, for any $x=1,2,\ldots,m$, by the product rule $(uv)'=u'v+uv'$, we obtain

$$f(x)=f'(x)=f''(x)=\cdots=f^{(p-1)}(x)=0. \tag{2}$$

For any x^k $(k\in\mathbb{Z}^+,\ k>p)$:

$$(x^k)'=kx^{k-1}, \quad (x^k)''=k(k-1)x^{k-2},$$

$$(x^k)^{(p)}=k(k-1)\cdots(k-p+1)x^{k-p}=\frac{k!}{(k-p)!}x^{k-p},$$

where the coefficients of $(x^k)^{(p)}$ are integers and multiples of p. (Because $\frac{k!}{(k-p)!}=p!\binom{k}{p}$, and the binomial coefficient $\binom{k}{p}$ is an integer.) This long expression can be written as a sum:

$$x^{p-1}(x-1)^p(x-2)^p\cdots(x-m)^p=\sum_{k=0}^{mp+p-1}a_k x^k,$$

where $a_k\in\mathbb{Z}$.

$$f^{(p)}(x)=\frac{1}{(p-1)!}\sum_{k=0}^{mp+p-1}a_k\cdot\left(x^k\right)^{(p)}=\frac{1}{(p-1)!}\sum_{k=0}^{mp+p-1}a_k\left(x^k\right)^{(p)}$$

We know that the coefficients of $\left(x^k\right)^{(p)}$ are integers and multiples of p. Therefore, the coefficients of $f^{(p)}(x)$ are integers and multiples of p.

By the same reasoning, the coefficients of $f^{(p)}(x),\ f^{(p+1)}(x),\ \ldots,\ f^{(mp+p-1)}(x)$ are all integers and multiples of p. (3)

According to Leibniz formula:

$$\left(u(x)v(x)\right)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} u^{(k)}(x)v^{(n-k)}(x)$$

Let $u(x) = x^{p-1}$ and $v(x) = (x-1)^p(x-2)^p \cdots (x-m)^p$, so that

$$f(x) = \frac{1}{(p-1)!}u(x)v(x).$$

$$\begin{aligned}
\left(u(x)v(x)\right)^{(p-1)} &= \sum_{k=0}^{p-1} \binom{p-1}{k} u^{(k)}(x)v^{(p-1-k)}(x) \\
&= \sum_{k=0}^{p-2} \binom{p-1}{k} u^{(k)}(x)v^{(p-1-k)}(x) + \binom{p-1}{p-1} u^{(p-1)}(x)v^{(0)}(x) \\
&= \sum_{k=0}^{p-2} \binom{p-1}{k} \left(x^{p-1}\right)^{(k)} v^{(p-1-k)}(x) + \left(x^{p-1}\right)^{(p-1)} v(x) \\
&= \sum_{k=0}^{p-2} \binom{p-1}{k} (p-1)(p-2)\cdots(p-k)x^{p-1-k}v^{(p-1-k)}(x) + (p-1)!v(x).
\end{aligned}$$

$$\left(u(0)v(0)\right)^{(p-1)} = \sum_{k=0}^{p-2} \binom{p-1}{k} (p-1)(p-2)\cdots(p-k)\cdot(0)^{p-1-k}\cdot v^{(p-1-k)}(0) + (p-1)!v(0).$$

$$\left(u(0)v(0)\right)^{(p-1)} = 0 + (p-1)!v(0).$$

Since

$$f^{(p-1)}(x) = \frac{1}{(p-1)!}\left(u(x)v(x)\right)^{(p-1)},$$

we have

$$f^{(p-1)}(0) = \frac{1}{(p-1)!}\cdot(p-1)!\cdot v(0) = v(0) = (0-1)^p(0-2)^p\cdots(0-m)^p$$

$$= \big[(-1)(-2)\cdots(-m)\big]^p = \big[(-1)^m m!\big]^p. \qquad (4)$$

$$\sum_{k=0}^{m} C_k F(k) = C_0 F(0) + C_1 F(1) + C_2 F(2) + \cdots + C_m F(m).$$

Since

$$F(k) = f(k) + f'(k) + \cdots + f^{(p-1)}(k) + f^{(p)}(k) + \cdots + f^{(m+p-1)}(k),$$

by (2),we have:

$$f(k) + f'(k) + \cdots + f^{(p-1)}(k) = 0 \quad \text{for } k = 1, 2, \ldots, m,$$

Therefore,

$$F(k) = f^{(p)}(k) + \cdots + f^{(m+p-1)}(k) \quad \text{for } k = 1, 2, \ldots, m.$$

$$\begin{aligned}
\sum_{k=0}^{m} C_k F(k) &= C_0 F(0) + \sum_{k=1}^{m} C_k F(k) \\
&= C_0 F(0) + \sum_{k=1}^{m} C_k \big(f^{(p)}(k) + \cdots + f^{(m+p-1)}(k)\big) \\
&= C_0 F(0) + 0 \quad (\text{by (3), } f^{(p)}(k), \ldots, f^{(m+p-1)}(k) \text{ are multiples of } p) \\
&\equiv C_0 F(0) \pmod{p}.
\end{aligned}$$

By (4), $F(0) = \big[(-1)^m m!\big]^p$, hence

$$\sum_{k=0}^{m} C_k F(k) \equiv C_0 \big[(-1)^m m!\big]^p \pmod{p}.$$

Since $m < p$, $p \nmid \big[(-1)^m m!\big]^p$. Choose p sufficiently large such that $p > C_0$, so $p \nmid C_0$. Hence,

$$\sum_{k=0}^{m} C_k F(k) \not\equiv 0 \pmod{p}.$$

This implies that the remainder of $\sum_{k=0}^{m} C_k F(k)$ modulo p is at least 1, which implies:

$$\left| \sum_{k=0}^{m} C_k F(k) \right| \geq 1.$$

Since we have derived both

$$\left|\sum_{k=0}^{m} C_k F(k)\right| < 1 \quad \text{and} \quad \left|\sum_{k=0}^{m} C_k F(k)\right| \geq 1,$$

a contradiction arises. Hence, e cannot be algebraic, so the constant e is transcendental.

7 Conclusion

The constant e appears in many different "faces": as a limit, a series, a probability, an optimal strategy, and a transcendental number. These are not coincidences, but reflections of a deeper unifying structure in mathematics.

What begins as a simple question about a number gradually reveals a broader picture: mathematics is not a collection of isolated ideas, but a connected whole. The constant e is one example of how the same idea can appear in different forms.

By following these connections, we can better understand not only e, but also it offers a glimpse into the shared and cumulative nature of mathematical thought.