Estimation of the Partition Number: After Hardy and Ramanujan

LI Wenwei

December 28, 2016

liwenwei@ustc.edu
School of Mathematical Science
University of Science and Technology of China

Abstract

The number of conjugate classes of permutations of order n is the same as the partition number $p(n)$. There are already several practical formulae to calculate $p(n)$. But they are either inconvenient for ordinary people (not majored in math) who do not want do write programs, or unsatisfying in accuracy. In this paper, some elementary approximation formulae with high accuracy for $p(n)$ will be presented. These estimation formulae are revised from Hardy-Ramanujan's asymptotic formula and they can be used to obtain the approximate value of $p(n)$ by a pocket calculator without programming function.

Key Words: Partition number, Estimation formula, functional approximation, Accuracy,

AMS2000 Subject Classification: 05A17, 11P81, 65D15

Contents

1 Introduction

An integer solution of the equation

$$s_1 + s_2 + \cdots + s_q = n \quad (1 \leqslant s_1 \leqslant s_2 \leqslant \cdots \leqslant s_q,\ q \geqslant 1), \tag{1.1}$$

(where s_1, s_2, $\cdots$, s_q are unknowns) is called a *partition* of an positive integer n. The number of all the partitions of n is denoted by $p(n)$. $p(n)$ is also called the *partition number* or the *partition function*. For a definite q, the number of solutions of Equation (1.1) is usually denoted by $P_q(n)$ or $p(n,q)$.

Partitions are tightly connected with the permutation groups and Latin squares. There is a 1-1 correspondence between conjugate classes of permutations of order n and the partitions of n. All the members in a conjugate class of permutations share the same cycle structure. A cycle structure of a permutation of order n can be considered as a partition of n if we admit the cycle of length 1 and keep the ones in a cycle structure.

There are already a lot of literatures on many aspects of $p(n)$. Euler, Hardy, Ramanujan, Rademacher, Newman, Erdős, Andrews, Berndt and Ono have made great contribution to this subject. Some important literatures may be found in [1], or in the references of [21], [5], [4] and [17].

Some important results about $p(n)$ are mentioned (or can be found) in [4], [21], [5] or [17]. In recent years, a very important result dues to Ken Ono and his team who connected the partition function with the modular form and found the principles of the congruence property of $p(n)$ that may even be considered as the revealing of the nature of numbers (refer [2], [10], [6], [8] and [7]).

There are already several formulae to calculate $p(n)$.

In reference [12] (page 53, 57) or [15], we may find the generation function of $p(n)$ obtained by Euler:

$$F(x) = \sum_{n=0}^{\infty} p(n)x^n = \frac{1}{1-x}\frac{1}{1-x^2}\frac{1}{1-x^3}\cdots\frac{1}{1-x^i}\cdots\cdots = \prod_{i=1}^{\infty}\left(1-x^i\right)^{-1}, \qquad (1.2)$$

and a formula

$$p(n) = \frac{1}{2\pi i}\oint_C \frac{F(x)}{x^{n+1}}\mathrm{d}x, \qquad (1.3)$$

where C is a contour around the original point. Of course, we seldom use Equation (1.3) to compute the value of $p(n)$ in practical.

There is a recursion for $p(n)$,

$$\begin{aligned} p(n) &= p(n-1) + p(n-2) - p(n-5) - p(n-7) + \cdots + \\ &\quad (-1)^{k-1} p\left(n - \frac{3k^2 \pm k}{2}\right) + \cdots\cdots \\ &= \sum_{k=1}^{k_1} (-1)^{k-1} p\left(n - \frac{3k^2+k}{2}\right) + \sum_{k=1}^{k_2} (-1)^{k-1} p\left(n - \frac{3k^2-k}{2}\right), \end{aligned} \qquad (1.4)$$

(Refer [12], page 55), where

$$k_1 = \left\lfloor \frac{\sqrt{24n+1}-1}{6} \right\rfloor, \; k_2 = \left\lfloor \frac{\sqrt{24n+1}+1}{6} \right\rfloor, \qquad (1.5)$$

and assume that $p(0) = 1$. Here $\lfloor x \rfloor$ stands for the maximum integer that will not exceed the real number x.

Equation (1.4) is much better for computing the value of $p(n)$. We can gain the exact value of $p(n)$ efficiently with a program based on it. But it is not convenient for people who do not want to write programs.

Further more, if we want to calculate $p(n)$ by Equation (1.4) by a small program written in C or some other general computer Language, it is usually necessary to decide the size of the space in memory to store the results beforehand, which means we should know the approximate value of $p(n)$ before the calculation started, (actually, here it is sufficient to know $\left\lceil \frac{\log_2 p(n) + 1}{8} \right\rceil$, where $\lceil x \rceil$ stands for the minimum integer that is greater than or equal to the real number x.), [1] otherwise we have to do some extra work for overflow

[1] Obviously, the datatypes already defined in the C language itself are not suitable.

If we use the Dynamic Memory Allocation method, this problem is solved at the price of the program being more complicated. Actually, in a lot of cases, we can not decide the approximate size of the result, this method is the best choice available.

If we can use maple, maximal, axiom or some other computer algebra systems, there is no need to consider this problem. But it is not always an option, especially when the function to do this job is part of a big program written in a compile language while mixing programming of an interpretative language and a compile language is nearly unavailable in most cases (with very few exception, such as mixing programming C and matlab).

handling and consequently change the size of the space in memory to store the value of the variable that stands for $p(n)$.

In references [12], we can find the approximation of the asymptotic order of the partition number $p(n)$,

$$p(n) \sim \exp\left(\sqrt{\frac{2}{3}}\pi n^{1/2}\right).$$

From it we obtained some estimation formulae of this form $a \cdot \exp\left(\sqrt{\frac{2}{3}}\pi n^{1/2}\right) + b \cdot g(x)$ with small errors when n ranges in a short interval, but the accuracy are not so satisfying when n increases.

The analysis of $p(n)$ by contour integral with Equation (1.3) (refer [12], page 57) resulted a very good estimation of $p(n)$,

$$p(n) = \sum_{q=1}^{\lfloor \alpha\sqrt{n} \rfloor} A_q(n) \cdot \phi_q(n) + O(n^{-1/2}), \tag{1.6}$$

(or equivalently $p(n) \approx \sum_{q=1}^{\lfloor \alpha\sqrt{n} \rfloor} A_q(n) \cdot \phi_q(n)$) called the *Hardy-Ramanujan formula* (refer [11] or [16]), that 6 terms of this formula contain an error of 0.004 when $n = 100$, while 8 terms of this formula contain an error of 0.004 when $n = 200$. Here α is an arbitrary constant,

$$\phi_q(n) = \frac{\sqrt{q}}{2\pi\sqrt{2}} \cdot \frac{\mathrm{d}}{\mathrm{d}n}\left(\frac{\exp\left(\frac{\pi}{q}\sqrt{\frac{2}{3}\left(n - \frac{1}{24}\right)}\right)}{\sqrt{n - \frac{1}{24}}}\right),$$

$$A_q(n) = \sum_{\substack{0 < p < q \\ (p,q) = 1}} \omega_{p,q} \cdot \exp\left(\frac{-2np\pi i}{q}\right)$$

(while p runs through the non-negative integers that are prime to q and less than q),

$$\omega_{p,q} = \begin{cases} \left(\frac{-q}{p}\right) \exp\left[-\left\{\frac{1}{4}(2 - pq - p) + \frac{1}{4}\left(q - \frac{1}{q}\right)(2p - p' + p^2 p')\right\}\pi i\right], & p \text{ is odd}, \\ \left(\frac{-p}{q}\right) \exp\left[-\left\{\frac{1}{4}(q - 1) + \frac{1}{12}\left(q - \frac{1}{q}\right)(2p - p' + p^2 p')\right\}\pi i\right], & q \text{ is odd}, \end{cases}$$

is a certain $24q$-th root of unity,

$$\left(\frac{a}{b}\right) = \begin{cases} 1, & \text{if } a \text{ is a quadratic residue modulo } b \text{ and } a \not\equiv 0 \pmod{b}, \\ -1, & \text{if } a \text{ is a quadratic non-residue modulo } b, \\ 0, & \text{if } a \equiv 0 \pmod{b}, \end{cases}$$

is the Legendre symbol and b is an odd prime, [2] and p' is any positive integer such that $q \,|\, (1+pp')$. When n is very large, $p(n)$ is the integer nearest to $\sum_{q=1}^{\lfloor \alpha\sqrt{n} \rfloor} A_q(n) \cdot \phi_q(n)$.

In [12] or [16], can we find a convergent series for $p(n)$ modified from Equation (1.6) by Rademacher in 1937,

$$p(n) = \sum_{q=1}^{\infty} A_q(n) \cdot \psi_q(n), \tag{1.7}$$

where $A_q(n)$ is the same as mentioned above and

$$\psi_q(n) = \frac{\sqrt{q}}{\pi\sqrt{2}} \frac{\mathrm{d}}{\mathrm{d}n} \left(\frac{\sinh\left(\frac{\pi}{q}\sqrt{\frac{2}{3}\left(n-\frac{1}{24}\right)}\right)}{\sqrt{n-\frac{1}{24}}} \right).$$

Equation (1.6) or Equation (1.7) are valuable in theory and can be used to calculate the value of $p(n)$ with very high accuracy. But they are not convenient for engineers or other ordinary people (not familiar with any computer algebra system softwares) because they are too complicated, and they contain some special functions that most people (not majored in mathematics) do not know. As a result, it is very difficult for them to use these two formulae to calculate $p(n)$ on a pocket science calculator without programming function (even the recursion formula Equation (1.4) will be better for small n).

In references [21] or [3], we may find the famous asymptotic formula for $p(n)$,

$$p(n) \sim \frac{1}{4n\sqrt{3}} \exp\left(\sqrt{\frac{2}{3}}\pi n^{1/2}\right), \tag{1.8}$$

obtained by Godfrey Harold Hardy and Srinivasa Ramanujan in 1918 in the famous paper [11]. (Two different proofs can be found in [9] and [14]. The evaluation of the constant was shown in [13].) This formula may be called the *Hardy-Ramanujan's asymptotic formula* in this paper. This asymptotic formula is with great importance in theory. Equation (1.8) is much more convenient than formulae Equation (1.6) and Equation (1.7) for ordinary people not majored in mathematics.

Let

$$R_{\mathrm{h}}(n) = \frac{1}{4n\sqrt{3}} \exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right). \tag{1.9}$$

be the asymptotic function by Hardy and Ramanujan.

[2] An integer a will be called a quadratic residue modulo another integer b if there is an integer c such that $c^2 \equiv a \pmod{b}$. Otherwise, a will be called a quadratic non-residue modulo b.

n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err
1	87.67%	16	11.60%	40	7.34%	220	3.05%	520	1.97%
2	35.76%	17	12.03%	50	6.54%	240	2.92%	540	1.93%
3	36.35%	18	10.91%	60	5.95%	260	2.80%	560	1.90%
4	22.00%	19	11.25%	70	5.50%	280	2.70%	580	1.86%
5	27.74%	20	10.43%	80	5.13%	300	2.60%	600	1.83%
6	17.11%	21	10.53%	90	4.83%	320	2.52%	640	1.77%
7	21.78%	22	9.96%	100	4.57%	340	2.44%	680	1.72%
8	16.08%	23	10.05%	110	4.35%	360	2.37%	720	1.67%
9	17.50%	24	9.49%	120	4.16%	380	2.31%	760	1.63%
10	14.53%	25	9.56%	130	3.99%	400	2.25%	800	1.58%
11	16.02%	26	9.16%	140	3.84%	420	2.20%	840	1.55%
12	12.91%	27	9.15%	150	3.71%	440	2.14%	880	1.51%
13	14.22%	28	8.82%	160	3.59%	460	2.10%	920	1.48%
14	12.50%	29	8.81%	180	3.38%	480	2.05%	960	1.44%
15	12.80%	30	8.50%	200	3.20%	500	2.01%	1000	1.42%

Table 1: The relative error of $R_{\text{h}}(n)$ to $p(n)$ when $n \leqslant 1000$.

By the figure in reference [18], this asymptotic formula fits $p(n)$ very well when n is huge.

But when n is small, the relative error of $R_{\text{h}}(n)$ to $p(n)$ is not so satisfying as shown in Table 1 (when $n \leqslant 1000$) on page 6. When $n \leqslant 25$, the relative error is greater than 9%; when $25 < n \leqslant 220$, the relative error is greater than 3%; when $n \leqslant 500$, the relative error is greater than 2%; when $n \leqslant 1000$, the relative error is greater than 1.4%. Considering that $p(n)$ is an integer and $R_{\text{h}}(n)$ is definitely not, the round approximation of $R_{\text{h}}(n)$ may be a little more accurate, but that does not help.

Although Equation (1.6) is not so accurate when n is small, it provides some important clue for a practical formula for small n.

2 The Main idea for estimating $p(n)$

Since $p(n) \sim R_{\text{h}}(n)$, i.e., $\lim\limits_{n\to\infty} \dfrac{R_{\text{h}}(n)}{p(n)} = 1$, we believe that an approximate formula with better accuracy may be in this form

$$p(n) \approx \frac{1}{4\sqrt{3}(n+C_2)} \exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n+C_1}\right). \tag{2.1}$$

Where C_1 or C_2 may be a constant or an function of n that increases slowly than n.

There are many different ways to modify $R_{\text{h}}(n)$, e.g. we could also construct a function $p_1(n)$ to estimate $R_{\text{h}}(n) - p(n)$, then $R_{\text{h}}(n) - p_1(n)$ may reach a better accuracy when estimating $p(n)$, or we can estimate the value of $\frac{R_\mathrm{h}(n)}{p(n)}$ by a function $f_1(n)$ then estimate $p(n)$ by $\frac{R_\mathrm{h}(n)}{f_1(n)}$, etc. The problem is that the accuracy of $R_\mathrm{h}(n)-p_1(n)$ is not so satisfying if we do not use the idea shown in Equation (2.1), because the shape of the figure of $\ln\left(R_\mathrm{h}(n)-p(n)\right)$ is nearly the same as the shape of the figure of $\ln\left(p(n)\right)$, at least we can not tell the difference of the shapes by our eyes as shown on Figure 6.1 and Figure 6.2 (on page 28), though they are different in theory. We will discuss the details in subsection 5.

As we can not determine C_1 and C_2 at the same time because of technique problems, [3] we may decide C_1 first then determine C_2, the main reason is that $\frac{1}{(n+C_2)}$ and $\frac{1}{n}$ differs very little when n is very huge, at least we believe that the difference is much less that the difference of $\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n+C_1}\right)$ and $\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)$.[4] So, when $n \gg 1$, we believe

$$p(n) \doteq \frac{1}{4\sqrt{3}n}\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n+C_1}\right),$$

hence $4\sqrt{3}n \times p(n) \doteq \exp\left(\pi\sqrt{\frac{2}{3}(n+C_1)}\right)$, then

$$C_1\left(n\right) \doteq \frac{3}{2}\cdot\frac{\left(\ln\left(4n\sqrt{3}p(n)\right)\right)^2}{\pi^2} - n. \tag{2.2}$$

If we point the data $\left(n,\ \frac{3}{2}\cdot\frac{\left(\ln\left(4n\sqrt{3}p(n)\right)\right)^2}{\pi^2}\right)$ $(n = 20k+100,\ k = 1,\ 2,\ \cdots,\ 395)$ in the coordinate system, we will find that they lies in a straight line, as shown in the Figure 2.1 on page 8, which means that the Hardy-Ramanujan's asymptotic formula is close to perfect. Here every tiny cycle stands for a data point.

3 Fit $C_1(n)$

If we point the data $(n,\ C_1(n))$, i.e., $\left(n,\ \frac{3}{2}\cdot\frac{\left(\ln\left(4n\sqrt{3}p(n)\right)\right)^2}{\pi^2} - n\right)$ $(n = 20k+100,\ k = 1,\ 2,\ \cdots,\ 395)$ in the coordinate system, we will get the Figure 2.2 on page 8. Here the points when $n \leqslant 120$ are not shown on Figure 2.2, partly because the deduction above is

[3] Usually, we will get the value of C_1 and/or C_2 from a number of pairs of $(n, p(n))$ by the least square method, not from two pairs of $(n, p(n))$ only. Many software can get efficiently the undetermined coefficients (by the least square method) by solving a system of (incompatible) linear equations, while it is very difficult to "solve" a system of tens or hundreds of transcendental equations that are incompatible.

[4] It is not difficult to know that $\frac{1}{(n+\delta)} \approx \frac{1}{n}\left(1-\frac{\delta}{n}\right)$, $\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n+\delta}\right) \approx \exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)\left(1+\frac{\pi}{\sqrt{6}}\frac{\delta}{\sqrt{n}}\right)$, when $\delta \ll n$. Obviously, $\frac{\delta}{n} \ll \frac{\pi}{\sqrt{6}}\frac{\delta}{\sqrt{n}}$ (when $\max\{\delta, 1\} \ll n$).

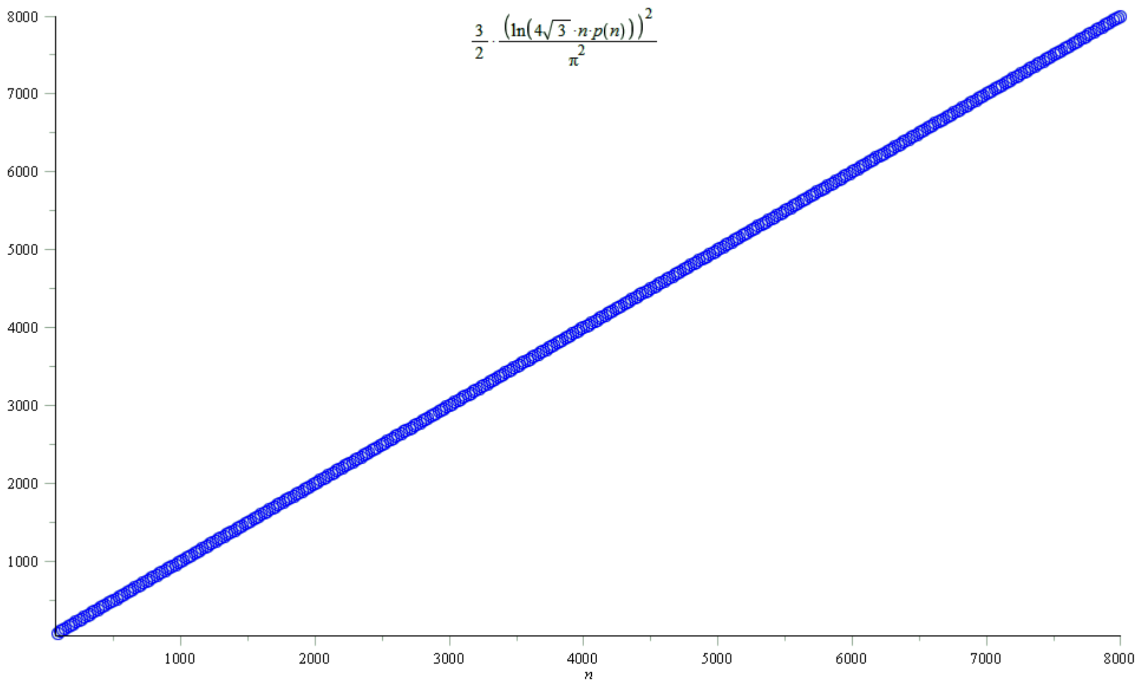

Figure 2.1: The graph of the data $\left(n,\ \frac{3}{2} \cdot \frac{\left(\ln\left(4n\sqrt{3}p(n)\right)\right)^2}{\pi^2}\right)$.

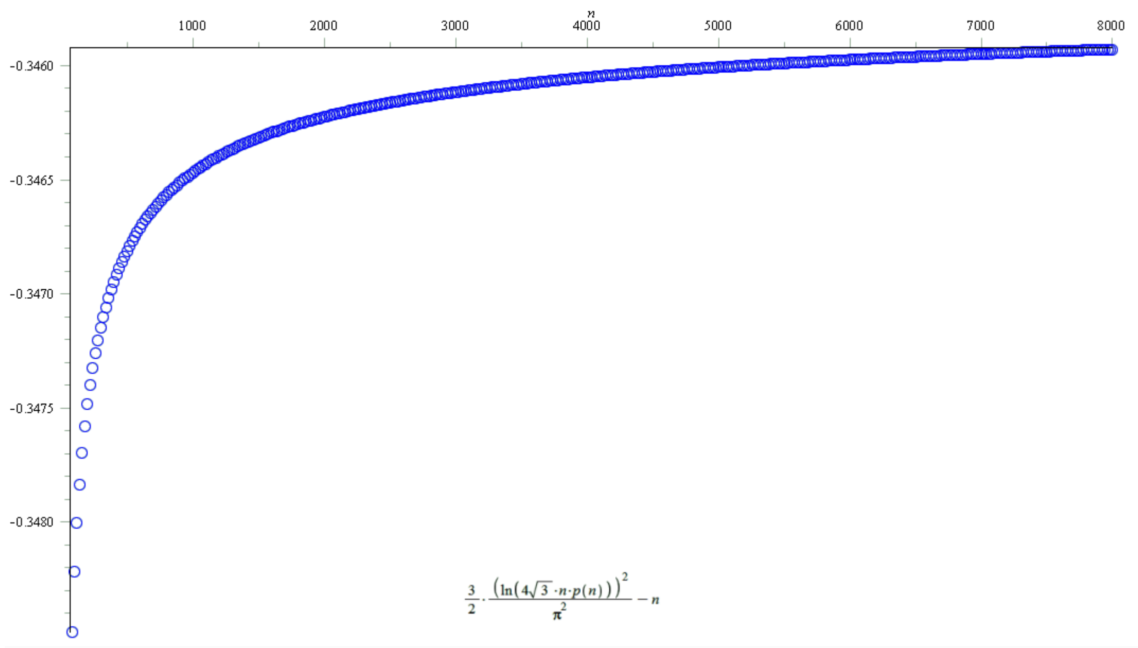

Figure 2.2: The graph of the data $(n,\, C_1(n))$ $(n \geqslant 120)$.

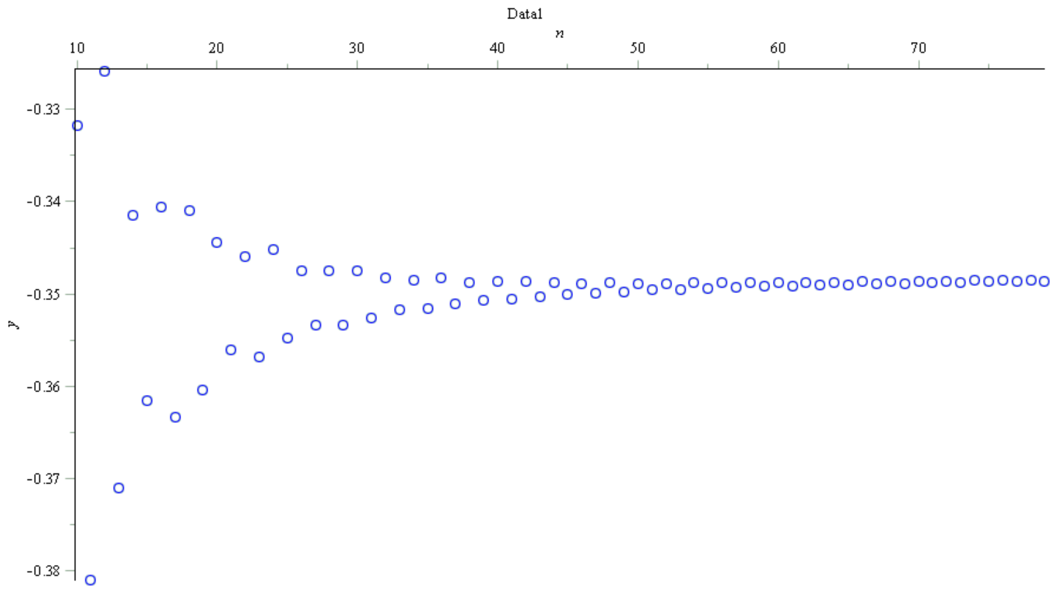

Figure 2.3: The graph of the data $\left(n,\ \frac{3}{2}\cdot\frac{\left(\ln\left(4n\sqrt{3}p(n)\right)\right)^2}{\pi^2}\right)$ $(n \leqslant 80)$.

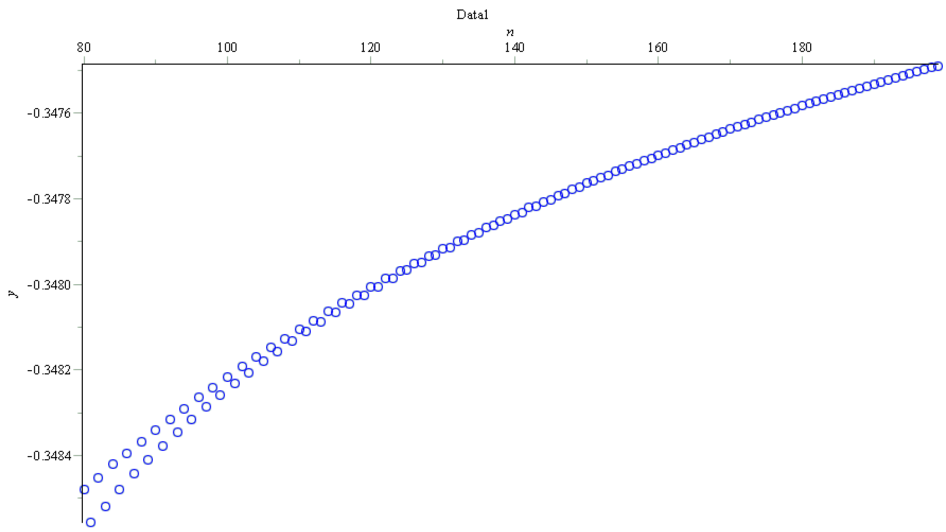

Figure 2.4: The graph of the data $(n,\ C_1\left(n\right))$ $(80 \leqslant n \leqslant 200)$.

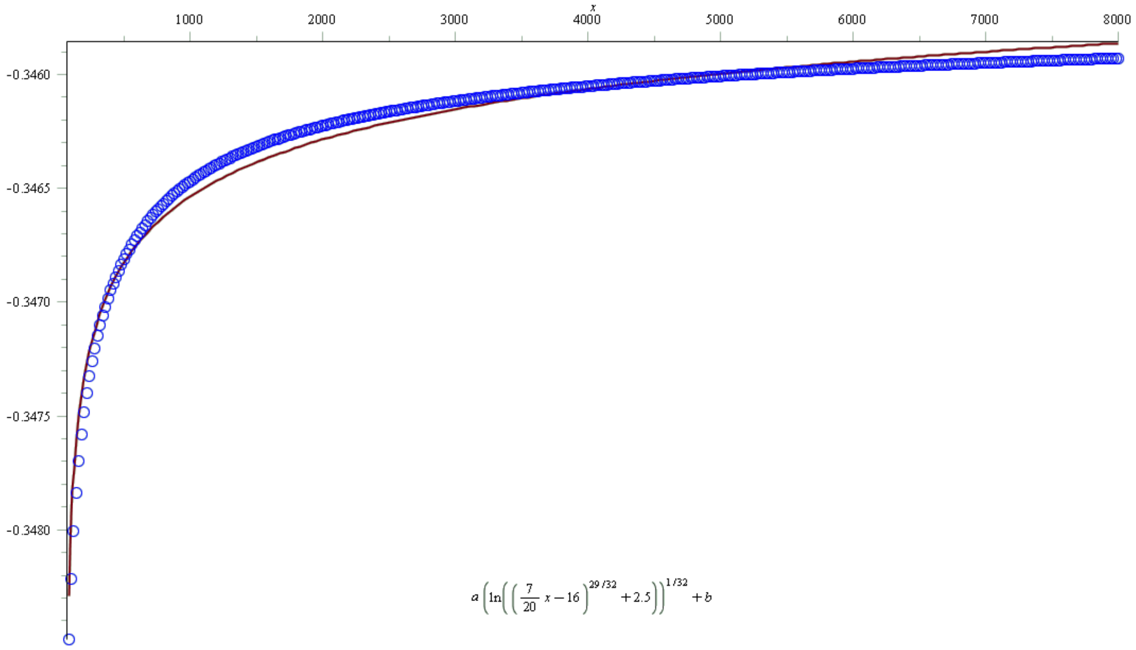

Figure 3.1: The graph of a bad fitting curve of the data $(n, C_1(n))$

based on $n \gg 1$, the main reason is that the points obviously do not lie in a curve when $n \leqslant 120$, as shown on Figure 2.3 and Figure 2.4 (on page 9).

Figure 2.2 looks like a logarithmic curve or a hyperbola. The author has tried hundreds of functions (by a small program written in MAPLE) like

$$a \cdot (\ln(x^{e_1} + c_1))^{e_2} + b,$$

where e_1, e_2 and c_1 are given constants while a, b are undermined coefficients to be decided by the least square method. But none of them fits the data very well. A function

$$y = a \cdot \left(\ln \left(\left(\frac{7}{20} \cdot x - 16 \right)^{29/32} + 2.5 \right) \right)^{1/32} + b,$$

where $a = 0.06656839293$ and $b =$ -0.4166945066, may fit the data better, but it is not as good as we expect, as shown on Figure 3.1 on page 10.

A hyperbola like $y = \dfrac{a}{x} + b$ does not fit the data very well, either, so we consider this type of functions

$$y = \frac{a}{(x + c_2)^{e_2}} + b, \tag{3.1}$$

where a, b, c_2 and e_2 are undetermined constants. This seems much better. For technique reason, we can not decide all the undetermined coefficients a, b, c_2, e_2 at the same time.

[5] These undetermined coefficients may be obtained in this way:

- A1. Give c_2 and e_2 initial values, such as $c_2 = 2.5$, $e_2 = 0.5$ (or some other values);
- A2. Fit the data $(n,\ C_1(n))$ by the least square method with Equation (3.1) and get the values of a and b, then get the average error of the fitting function for the values of c_2, e_2, a, b; [6]
- A3. Reevaluate e_2 and a. Plot the points of the data $(\ln(n + c_2), \ln(b - C_1(n)))$ $(n = 20k + 100,\ k = 1,\ 2,\ \cdots,\ 395)$ in the coordinate system with the values of b and c_2 just found, [7] fit the data by the least square method with

  $$y = e_1 \cdot x + a_1$$

  and find the values of a_1 and e_1, [8] then reevaluate e_2 and a by

  $$e_2 = -e_1,\ a = -\exp(a_1);$$

- A4. Reevaluate c_2. Plot the points of the data $\left(n, \left(\dfrac{a}{C_1(n) - b}\right)^{1/e_2}\right)$ $(n = 20k + 100,\ k = 1,\ 2,\ \cdots,\ 395)$ in the coordinate system with the value of b and the new

---

[5] Because most computer algebra system (CAS) could not solve system of incompatible nonlinear equations in the least square method, or the time-consumption is unacceptable.

[6] E.g., if $c_2 = 2.5$, $e_2 = 0.5$, then $a = -0.02635983935$, $b = -0.3456348045$.

If we plot the figure of Equation Equation (3.1) with the value of c_2, e_2, a, b, and compare the figure with Figure 2.2 on page 8, we will get a graph nearly the same as Figure 2.2 (although there should be a little different, but we can not distinguish the difference by our eyes).

Here we use the the square root of the mean square deviation

$$s = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - f(x_i))^2}$$

to measure the average error of the fitting function $y = f(x)$ to the original data $(x_i,\ y_i)$ $(i = 1,\ 2,\ \cdots,\ m)$. The average error of the fitting function for the values of c_2, e_2, a, b mentioned above is $1.074574171 \times 10^{-5}$, which seems to be very tiny.

[7] Such as shown in Figure 3.3 on page 13 when $c_2 = 2.5$ and $b = -0.3456348045$.

The purpose of this step is to obtain more accurate values of e_2 and a. Since $C_1(n) = \dfrac{a}{(n + c_2)^{e_2}} + b$, then $b - C_1(n) = \dfrac{-a}{(n + c_2)^{e_2}}$, (considering that $a < 0$), $\ln(b - C_1(n)) = \ln(-a) + e_2 \cdot \ln(n + c_2)$, so the figure of data $(\ln(n + c_2), \ln(b - C_1(n)))$ will be some points on a straight line if the previous assumption is correct and meanwhile the values of b and c_2 are proper.

[8] E.g., if $c_2 = 2.5$, $b = -0.3456365954$, then $a_1 = -3.626380777$, $e_1 = -0.5012314726$.

After reevaluation $e_2 = 0.5012314726$, $a = -0.02661232627$.

values of a and e_2, [9] fit the data by the least square method with

$$y = x + c_1$$

and find the value of c_1, then reevaluate c_2 by $c_2 = c_1$. [10]

- A5. goto step 2 until a fitting function with the least average error is obtained.

Actually, only a few times of repeating the steps form A2 to A4, we will obtain a very good fitting function, as shown on Figure 3.2 on page 13.

There are some explanations about the steps above:

- (1). In step A4, we did not plot the points of the data $\left(n, \left(\frac{a}{C_1(n) - b}\right)^{1/e_2} - n\right)$ because the shape of the figure is not a horizontal line as shown on Figure 3.5 on page 15 (the points in the right hand side are not so smooth because only 10 significance digits are kept in the process, if more significance digits are calculated, it will be better). Actually, it is a little complicated. But it will not help us to obtain better values of the undetermined in Equation (3.1) if we fit the data $\left(n, \left(\frac{a}{C_1(n) - b}\right)^{1/e_2} - n\right)$ with a more accurate fitting function.

- (2). In step A3, if we do not reevaluate a, the fitting parameters will not converge in general (even if we computing more significant figures in the process), or we can not continue the iterations steps at all since imaginary numbers appear.

- (3). If we started with a different initial value of c_2 and keep the initial value of e_2, such as $c_2 = 15$, after repeating 78 times of the steps from A2 to A4, we will find a fitting function

$$y = \frac{-0.02593608938}{(x + 3.272445238)^{0.4962730054}} - 0.3456286681, \tag{3.2}$$

with a minimal average error $9.109686836 \times 10^{-8}$.
If we started with some different initial values of both c_2 and e_2, such as $c_2 = 15$ and $e_2 = 0.7$, (from Figure 2.2 on page 8, we will find that e_2 should be less that

[9] Such as shown on Figure 3.4 on page 15 when $b = -0.3456365954$, $e_2 = 0.5012314726$ and $a = -0.02661232627$.

The main idea of this step: since $C_1(n) = \frac{a}{(n + c_2)^{e_2}} + b$, then $n + c_2 = \left(\frac{a}{C_1(n) - b}\right)^{1/e_2}$, hence the figure of data $\left(n, \cdot \left(\frac{a}{C_1(n) - b}\right)^{1/e_2}\right)$ will be some points on a straight line.

[10] E.g., for the values of b, e_2 and a mentioned before, after reevaluation $c_2 = 4.871833842$.

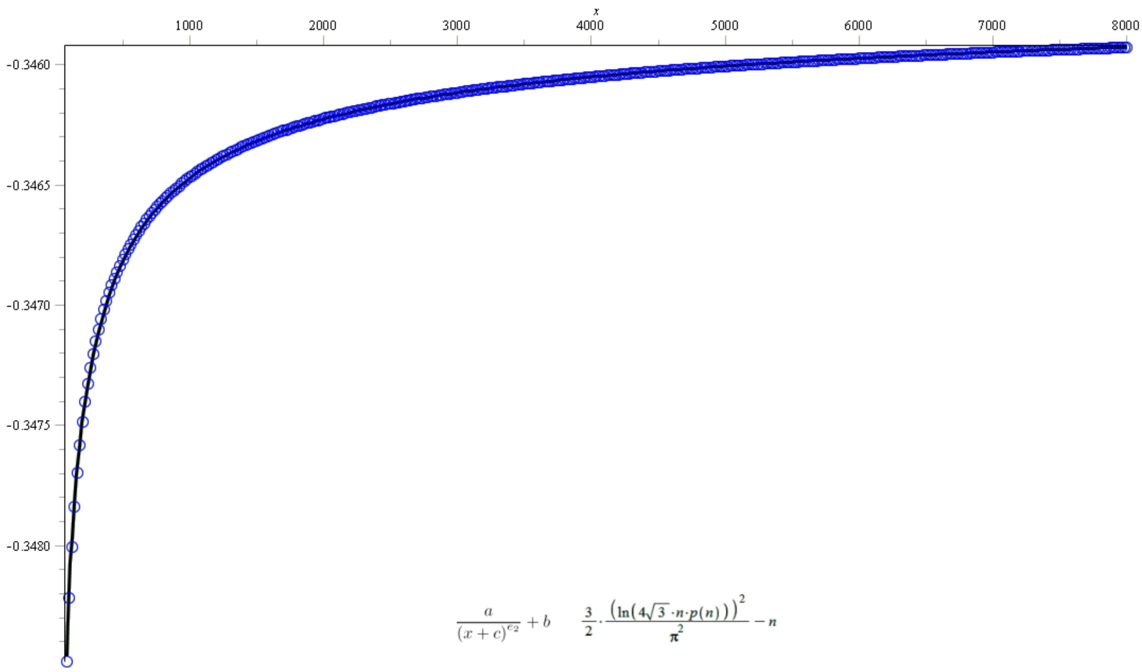

Figure 3.2: The graph of a good fitting curve of the data $(n,\, C_1\,(n))$

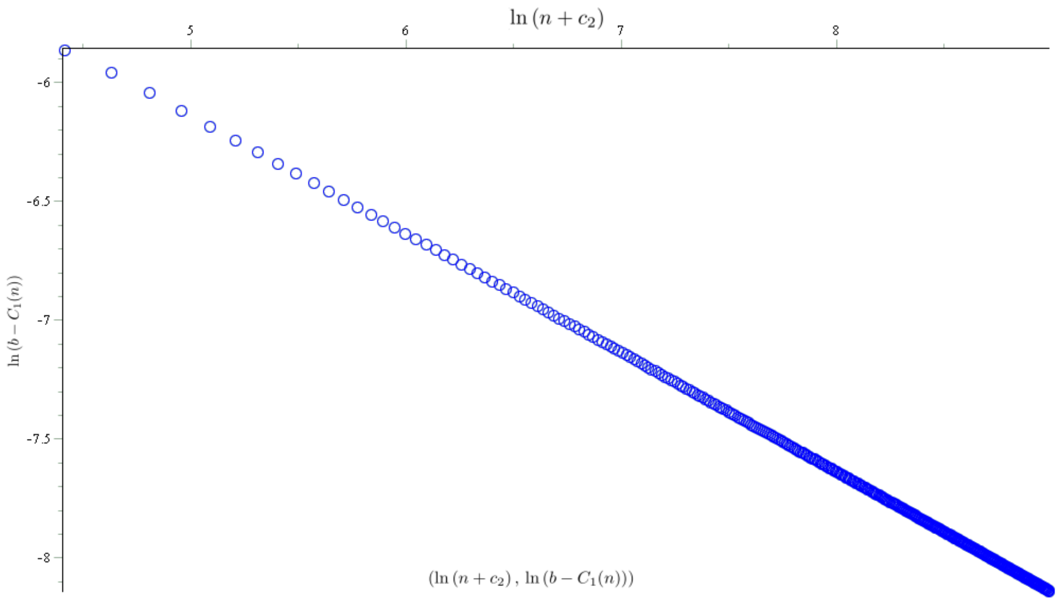

Figure 3.3: The graph of the data $(\ln\,(n+c_2)\,,\,\ln\,(b-C_1(n)))$

1.0), we will get the similar result. After repeating 125 times of the steps from A2 to A4, we will find a fitting function

$$y = \frac{-0.02593617719}{(x + 3.273513225)^{0.4962727258}} - 0.3456286655, \tag{3.3}$$

with a minimal average error $9.105941452 \times 10^{-8}$. After that, e_2 and c_2 will decrease slowly and slowly, and the average error will increase little by little if we continue the steps from A2 to A4.

As concerned to the errors in computing, the valid value of the undermined a, b, c_2 and e_2 should be -0.0259361, -0.34562866, 3.273, 0.49627, the average absolute error of the fitting function of $C_1(n)$ is about 9.1×10^{-8}. [11]

Considering that Equation (2.2) is an approximate formula, we may believe that the best value of e_2 is 0.5, since we prefer a simple exponent. Then it will be more convenient to obtain a, b and c_2.

Below e_2 is believed to be 1/2, which means that the fitting function of $C_1(n)$ is

$$y = \frac{a}{\sqrt{x + c_2}} + b. \tag{3.5}$$

When e_2 is fixed to be 1/2, if we use the iteration method described above but keep the value of e_2 in step A3, i.e., substitute step A3 by

A3'. Reevaluate a by [12]

$$a = -\exp\left(\frac{1}{395}\sum_{k=1}^{395}\left(\ln\left(b - C_1(20k + 100)\right) - e_2 \cdot \ln\left(20k + 100 + c_2\right)\right)\right);$$

(that means we evaluate a twice in every loop) the sequence of fitting functions of $C_1(n)$ will diverge. But we will obtain a converged sequence of the determinants if n ranges from 120 to 6000, (i.e., consider only the data $(n, p(n))$ when $n = 20k + 100$, $k = 1, 2, \cdots, 295$). The fitting function of $C_1(n)$ obtained in this way is

$$y = \frac{-0.02650620466}{\sqrt{x + 4.855479108}} - 0.3456326154, \tag{3.6}$$

[11] Actually, for the initial value $c_2 = 2.5$, $e_2 = 0.5$, after repeating 41 times of the steps from A2 to A4, we will find a fitting function

$$y = \frac{-0.02594609078}{(x + 3.320623832)^{0.4963284361}} - 0.3456286995, \tag{3.4}$$

with a minimal average error $9.010349470 \times 10^{-8}$. After a few times more of iteration, a result with similar coefficients will be found but with a little more error.

[12] or equivalently, Plot the points of the data $(\ln(n + c_2), \ln(b - C_1(n)))$ $(n = 20k + 100$, $k = 1, 2, \cdots, 395)$ in the coordinate system with the values of b, e_2 and c_2 just found, fit the data by the least square method with $y = e_2 \cdot x + a_1$ and find the values of a_1, then reevaluate a by $a = -\exp(a_1)$;

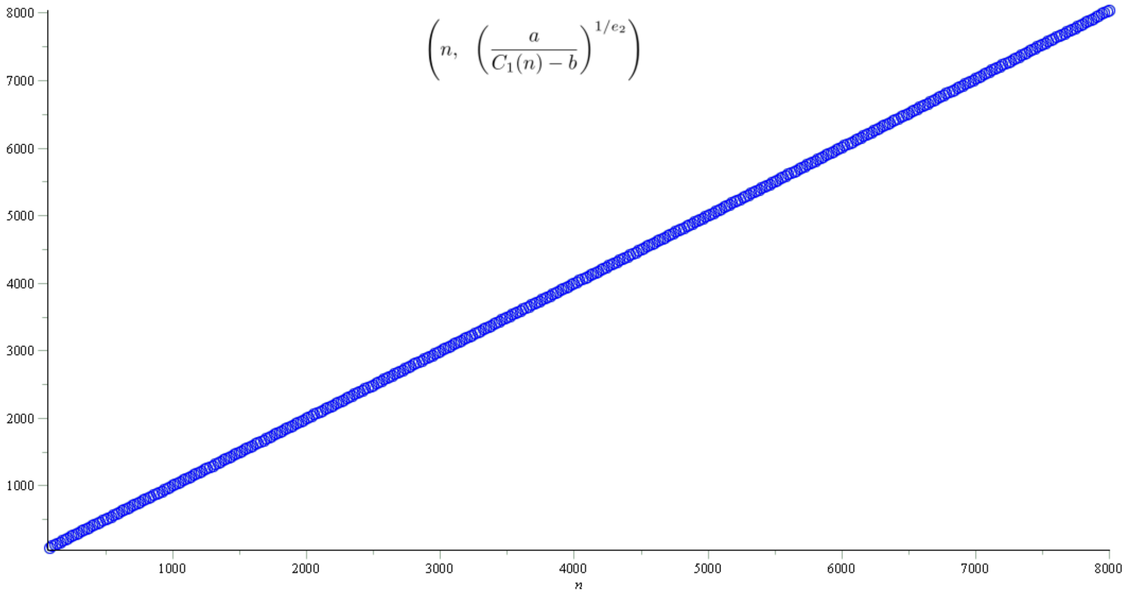

Figure 3.4: The graph of the data $\left(n, \cdot \left(\frac{a}{C_1(n)-b}\right)^{1/e_2}\right)$

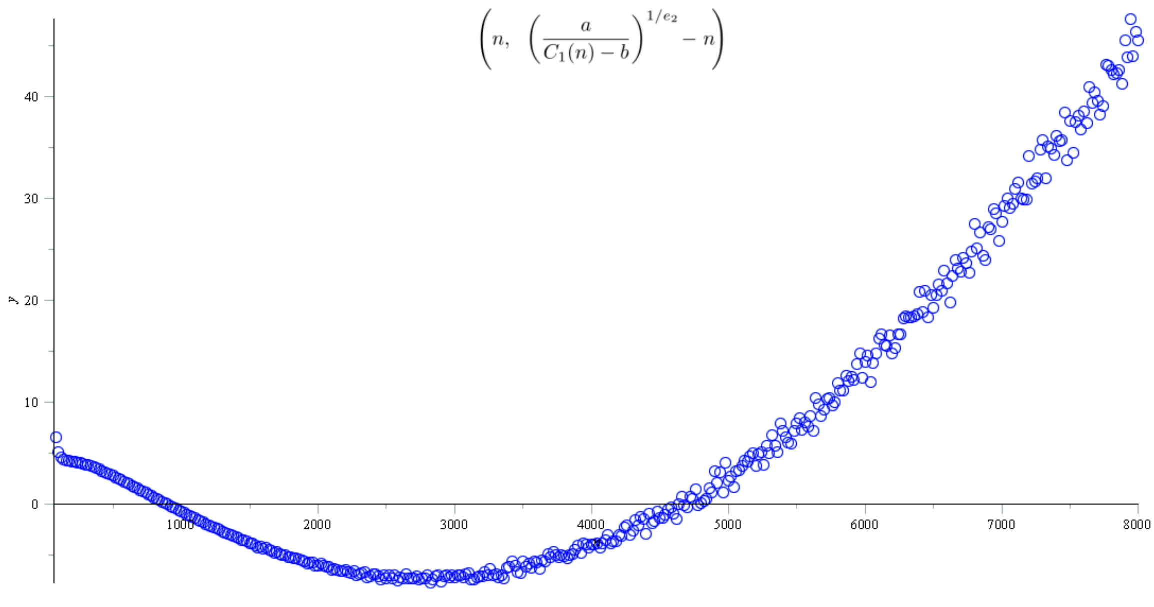

Figure 3.5: The graph of the data $\left(n, \cdot \left(\frac{a}{C_1(n)-b}\right)^{1/e_2} - n\right)$

with the minimal average error 2.374935895×10^{-7}. [13]

For the fixed value 1/2 of e_2, if we continue use the iteration method described above but ignore step 3, which means we reevaluate a only once in every loop, we will meet the same situation. The sequence of fitting functions of $C_1(n)$ will diverge if n ranges from 120 to 8000 (or 6000) even if we calculate more significance digits (such as 18 significance digits) in the process, but it will converge if n ranges from 120 to 4000. The fitting function of $C_1(n)$ obtained in this way is

$$y = \frac{-0.02647712648}{\sqrt{x+4.55083607}} - 0.345633305, \tag{3.7}$$

with the minimal average error 1.993012726×10^{-7} when the initial value of c_2 is 10 (iterated 4 times). But after more times of iteration, for several initial values of c_2 (such as 5, 10, 15, etc), the fitting functions converge to

$$y = \frac{-0.0268\cdots}{\sqrt{x+4.888\cdots}} - 0.345632760\cdots, \tag{3.8}$$

with the average error $2.68\cdots\times10^{-7}$.

Unlike the previous method, by the results mentioned above and some other results not mentioned here, the sequence of fitting functions of $C_1(n)$ usually converges to a function which is obviously different from the one with the minimal average error.

In order to get a fitting function with errors as tiny as possible, we can design another algorithm.

By the results described above, we known that c_2 is probably between 3 and 5, so we can find the fitting function of $C_1(n)$ and the average error for some values of c_2 in the possible range, then choose the one with minimal average error. To be cautious, we test the value of c_2 in the interval [0.5, 15]. The main steps are as below:

- (1) Initial c_a, c_b, c_0, s_0, D_t, a_0, b_0. Let c_a = 0.5, c_b = 15, c_0 = 0, s_0 = 1, a_0 = 0, b_0 = 0, D_t = 8, s_t = 0.1,.

[13] If we use the value of c_1 already found above, such as c_2= 3.273513225 in Equation (3.3), the fitting function is

$$y = \frac{-0.02640970103}{\sqrt{x+3.273513225}} - 0.3456340228,$$

with an average error 7.404647856×10^{-7}, which is about 3 times than that above.

If we choose c_2= 3.320623832 in Equation (3.4), the fitting function is

$$y = \frac{-0.02641281526}{\sqrt{x+3.320623832}} - 0.3456339736,$$

with an average error 7.205944166×10^{-7}.

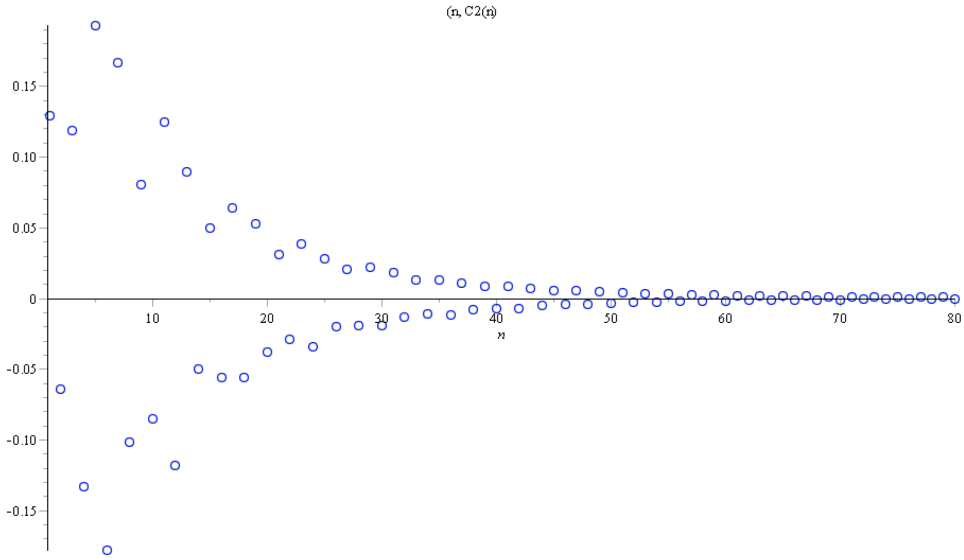

Figure 3.6: The graph of the data $(n,\, C_2(n))$

- (2) for c_2 from c_a to c_b by s_t do
  Fit the data $(n,\, C_1(n))$ by the least square method with Equation (3.5) and get the values of a and b, then get the average error s_1 of the fitting function for the values of c_2, a, b;
  if $s_1 < s_0$, then let $c_0 = c_2$, $s_0 = s_1$, $a_0 = a$, $b_0 = b$; end if;
  end do
- (3) If $D_\mathrm{t} > 1$, then set $D_\mathrm{t} \leftarrow D_\mathrm{t} - 1$, $c_\mathrm{a} \leftarrow c_0 - 5s_\mathrm{t}$, $c_\mathrm{b} \leftarrow c_0 + 5s_\mathrm{t}$;
  set $s_\mathrm{t} \leftarrow s_\mathrm{t}/10$; goto step (2);
  else, terminate the process.
  end if;

Here the symbol "$x \leftarrow y$" means that the variable x is evaluated by a value y; in step (1), $D_\mathrm{t} = 8$ means that we will get 8 significance digits of the value of c_2.

If n ranges from 120 to 8000, we can get a fitting function of $C_1(n)$,

$$y = \frac{-0.02651010067}{\sqrt{x + 4.8444724}} - 0.3456324524, \tag{3.9}$$

with a minimal average error $2.446731760 \times 10^{-7}$.

If n ranges from 120 to 6000, the fitting function of $C_1(n)$ is,

$$y = \frac{-0.02649625326}{\sqrt{x + 4.7152127}} - 0.3456327903, \tag{3.10}$$

with a minimal average error $2.279396699 \times 10^{-7}$.

Below, Equation (3.9) will be used to estimate $C_1(n)$, i.e.,

$$C_1(n) \doteq \frac{-0.02651010067}{\sqrt{n + 4.8444724}} - 0.3456324524. \tag{3.11}$$

4 Fit $C_2(n)$

By Equation (2.1) and Equation (3.11), we have

$$C_2(n) \doteq \frac{\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n+C_1(n)}\right)}{4\sqrt{3}p(n)} - n. \tag{4.1}$$

If we point out the data $(n,\, C_2(n))$ $(1 \leqslant n \leqslant 80)$ on the coordinate system as shown on Figure 3.6 on page 17, we will immediately know than $C_2(n)$ can not be fit by a simple function. From the Figure 3.6 (or the value of $C_2(n)$ calculated by a small program), it is clear that $C_2(n)$ is very small when n >40, at least much less than n, so there is no need to fit $C_2(n)$ when $n > 40$.

When n is odd, the points of $(n,\, C_2(n))$ are above the horizontal-axis, it is not difficult to separate them into two parts and fit them by two cubic curves, as shown on Figure 4.1 and Figure 4.2. The two fitting functions are

$$\begin{aligned} y = &-1.548835311 \times 10^{-6} \times x^3 + 1.880663805 \times 10^{-4} \times x^2 - \\ &0.008334098201 \times x + 0.1399798428, \\ y = &-5.416501948 \times 10^{-6} \times x^3 + 5.728510889 \times 10^{-4} \times x^2 - \\ &0.02125835759 \times x + 0.2882706948. \end{aligned}$$

For the points of $(n,\, C_2(n))$ under the horizontal-axis (when n is even), we have to separate them into at least 4 parts so as to fit them smoothly, two or three parts are not convenient.

As a result, we have to fit $C_2(n)$ by a hybrid function with at least 6 pieces, or fit $p(n)$ by a piecewise-defined function with 7 pieces, which is very complicated. This seems to contradict with our purpose at the beginning of this section.

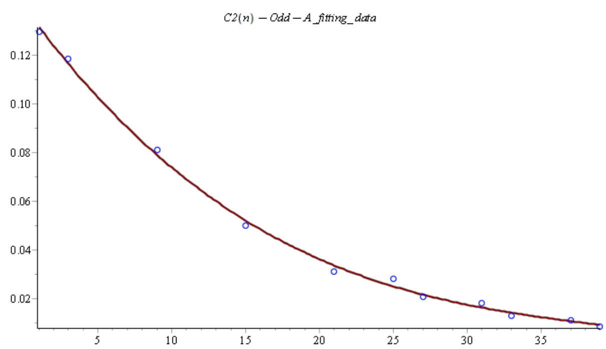

Figure 4.1: Fit $(n,\, C_2(n))$, the odd, Part A

From Figure 4.1 on page 18 we found that the value of $C_2(n)$ are much less than n when $n \geqslant 15$, so the error will be very tiny if we omit $C_2(n)$. Hence we can calculate $p(n)$ directly by

$$R_{\mathrm{h1}}(n) = \frac{1}{4\sqrt{3}n}\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n + \frac{a_1}{\sqrt{n+c_1}} + b_1}\right), \tag{4.2}$$

where $a_1 = -0.02651010067$, $b_1 = -0.3456324524$ and $c_1 = 4.8444724$.

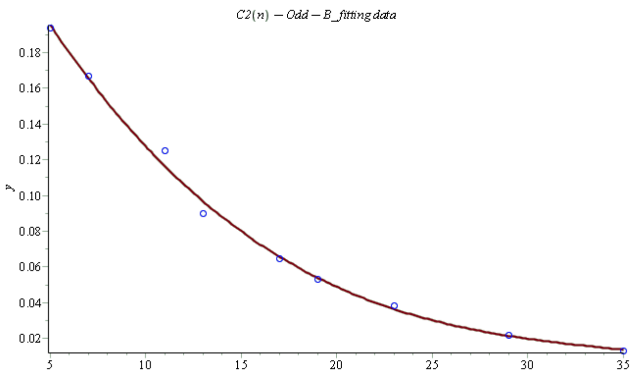

Figure 4.2: Fit $(n, C_2(n))$, the odd, Part B

The error of Equation (4.2) to $p(n)$ is shown on Table 2 on page 20. The accuracy is better than Equation (1.8). The relative error is less than 6×10^{-7} when $n \geqslant 100$, less than 1‰ when $n \geqslant 26$, less than 1% when $n \geqslant 11$, although this fitting function is obtained when $n \geqslant 120$. When $1000 \leqslant n \leqslant 3000$, the relative error is less than 1×10^{-8}. When $3000 \leqslant n \leqslant 10000$, the relative error is less than 5.3×10^{-9}, as shown on Figure 4.3 on page 19. But the relative error is not so satisfying when $n \leqslant 7$, especially when $n = 1$.

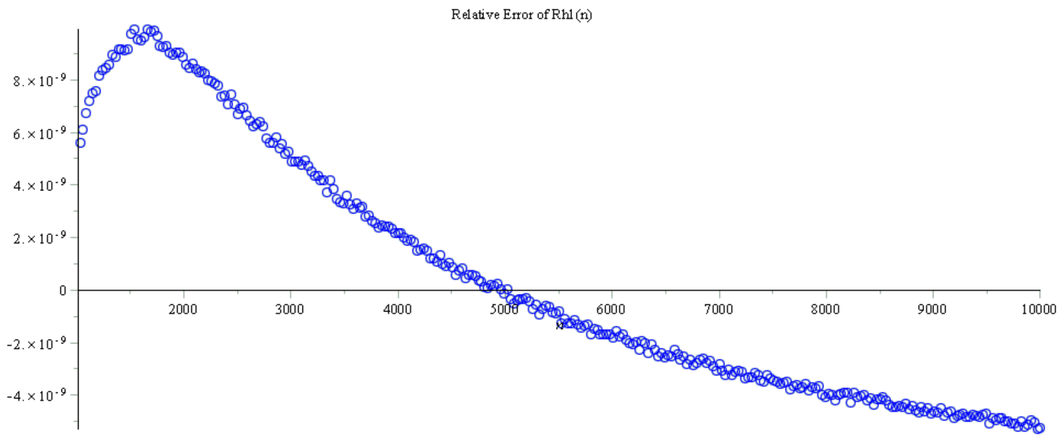

Figure 4.3: The Relative Error of $R_{\mathrm{h1}}(n)$ when $1000 \leqslant n \leqslant 10000$

Consider that $p(n)$ is an integer, if we take the round approximation of Equation (4.2),

$$R'_{\mathrm{h1}}(n) = \left\lfloor \frac{1}{4\sqrt{3}n} \exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n + \frac{a_1}{\sqrt{n+c_1}} + b_1}\right) + \frac{1}{2} \right\rfloor, \tag{4.3}$$

(we may call it *Hardy-Ramanujan's revised estimation formula* 1), it will solve perfectly the relative error problem when $n < 11$, as shown on Table 3 on page 20, although the relative error will increase very little for some n, which is negligible. (The average relative error is less than 2×10^{-8} when $n \geqslant 200$.) Take an example, when $n = 100$, $R_{\mathrm{h2}}(100) = 190569177$, $p(100) = 190569292$, the difference is 115; when $n = 200$, $R_{\mathrm{h2}}(200) = 3972999059745$, $p(200) = 3972999029388$, the difference is 30357. Although the errors are much greater than the error 0.004 of Hardy-Ramanujan formula with 6 terms ($n = 100$) or 8 terms ($n = 200$) (refer [11] or [16]), it contains only one term of elementary functions, and is convenient for a junior middle school student to calculate the value of $p(n)$ with high accuracy.

n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err
1	12.97%	16	-0.35%	40	-0.02%	220	-2.18E-08	520	-1.26E-08
2	-3.22%	17	0.38%	50	-0.01%	240	-3.11E-08	540	-3.00E-10
3	3.96%	18	-0.31%	60	-2.89E-05	260	-6.04E-08	560	2.00E-09
4	-3.32%	19	0.28%	70	-1.07E-05	280	-6.41E-08	580	3.00E-09
5	3.87%	20	-0.19%	80	-4.40E-06	300	-6.11E-08	600	-1.40E-09
6	-2.96%	21	0.15%	90	-1.87E-06	320	-6.48E-08	640	8.00E-09
7	2.38%	22	-0.13%	100	-5.96E-07	340	-3.59E-08	680	6.00E-09
8	-1.27%	23	0.17%	110	-1.06E-07	360	-3.31E-08	720	2.30E-08
9	0.90%	24	-0.14%	120	7.20E-08	380	-4.08E-08	760	6.00E-09
10	-0.85%	25	0.11%	130	1.35E-07	400	-2.21E-08	800	2.00E-09
11	1.13%	26	-0.08%	140	1.34E-07	420	-3.56E-08	840	2.10E-08
12	-0.98%	27	0.08%	150	1.16E-07	440	-1.59E-08	880	1.90E-08
13	0.69%	28	-0.07%	160	9.10E-08	460	-1.13E-08	920	2.60E-08
14	-0.35%	29	0.08%	180	4.40E-08	480	-1.52E-08	960	2.10E-08
15	0.33%	30	-0.06%	200	9.00E-09	500	-9.90E-09	1000	2.80E-08

Table 2: The relative error of $R_{\mathrm{h1}}(n)$ to $p(n)$ when $n \leqslant 1000$.

n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err
1	0	16	-0.43%	40	-1.87E-04	220	-2.76E-08	520	-2.42E-08
2	0	17	0.34%	50	-6.37E-05	240	-3.67E-08	540	-1.28E-08
3	0	18	-0.26%	60	-2.90E-05	260	-6.16E-08	560	-7.60E-09
4	0	19	0.20%	70	-1.08E-05	280	-6.93E-08	580	-3.19E-09
5	0	20	-0.16%	80	-4.43E-06	300	-6.77E-08	600	-1.21E-08
6	0	21	0.13%	90	-1.87E-06	320	-7.21E-08	640	1.96E-09
7	0	22	-0.10%	100	-6.03E-07	340	-3.74E-08	680	-8.54E-09
8	0	23	0.16%	110	-1.05E-07	360	-3.86E-08	720	1.30E-08
9	0	24	-0.13%	120	6.61E-08	380	-4.71E-08	760	-2.10E-09
10	0	25	0.10%	130	1.34E-07	400	-2.96E-08	800	-1.24E-08
11	1.79%	26	-8.21E-04	140	1.31E-07	420	-3.80E-08	840	1.25E-08
12	-1.30%	27	6.64E-04	150	1.07E-07	440	-2.31E-08	880	6.78E-09
13	0.99%	28	-8.07E-04	160	8.98E-08	460	-1.72E-08	920	1.52E-08
14	0	29	6.57E-04	180	3.94E-08	480	-2.31E-08	960	1.47E-08
15	0.57%	30	-7.14E-04	200	7.64E-09	500	-2.08E-08	1000	2.11E-08

Table 3: The relative error of $R'_{\mathrm{h1}}(n)$ to $p(n)$ when $n \leqslant 1000$.

5 Estimation $p(n)$ by Some Other Methods

In the previous subsection, we assume that $C_1(n) \doteq \frac{3}{2} \cdot \frac{\left(\ln\left(4n\sqrt{3}p(n)\right)\right)^2}{\pi^2} - n$, then fit the data $\left(n,\ \frac{3}{2} \cdot \frac{\left(\ln\left(4n\sqrt{3}p(n)\right)\right)^2}{\pi^2}\right)$ $(n = 20k + 100,\ k = 1,\ 2,\ \cdots,\ 395)$, and estimate $p(n)$ by $R_{\mathrm{h2}}(n) = \left\lfloor \frac{1}{4\sqrt{3}n} \exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n + C_1(n)}\right) + \frac{1}{2} \right\rfloor$.

If we assume that $p(n) \doteq \frac{1}{4\sqrt{3}(n + C_2)} \exp\left(\pi\sqrt{\frac{2}{3}n}\right)$, then

$$C_2(n) \doteq \frac{1}{4\sqrt{3}p(n)} \exp\left(\pi\sqrt{\frac{2}{3}n}\right) - n,$$

we wonder whether we can fit the data $\left(n,\ \frac{\exp\left(\pi\sqrt{\frac{2}{3}n}\right)}{4\sqrt{3}p(n)} - n\right)$ $(n = 20k + 100,\ k = 1,\ 2,\ \cdots,\ 395)$ by a function C_2 and estimate $p(n)$ by $\left\lfloor \frac{1}{4\sqrt{3}(n + C_2)} \exp\left(\pi\sqrt{\frac{2}{3}n}\right) + \frac{1}{2} \right\rfloor$?

The data $\left(n,\ \frac{\exp\left(\pi\sqrt{\frac{2}{3}n}\right)}{4\sqrt{3}p(n)} - n\right)$ $(n = 20k + 100,\ k = 1,\ 2,\ \cdots,\ 395)$ are shown on Figure 5.1 on page 24 (together with a fitting function). It is not difficult to know that a function in this form

$$y = a_1 \times (x + c_1)^{e_1} + b_1$$

will fit the points very well, and $e_1 = 0.5$ will be very satisfying. By the same method to fit $C_1(n)$, we can obtain a fitting function

$$y = 0.4432884566 \times \sqrt{x + 0.274078} + 0.1325096085$$

to fit $C_2(n)$ with an average error 3.65×10^{-6}.

Hence we can calculate $p(n)$ by

$$R_{\mathrm{h2}}(n) = \frac{\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)}{4\sqrt{3}\left(n + a_2\sqrt{n + c_2} + b_2\right)}, \tag{5.1}$$

where $a_2 = 0.4432884566$, $b_2 = 0.1325096085$ and $c_2 = 0.274078$, when n is not so small.

n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err
1	14.93%	16	-0.36%	40	-0.02%	220	3.90E-08	520	-3.27E-08
2	-3.06%	17	0.37%	50	-6.80E-05	240	3.00E-08	540	-5.00E-09
3	3.96%	18	-0.31%	60	-3.16E-05	260	2.90E-08	560	-7.80E-09
4	-3.34%	19	0.27%	70	-1.24E-05	280	8.00E-09	580	3.00E-09
5	3.84%	20	-0.19%	80	-5.48E-06	300	-2.50E-09	600	-9.00E-10
6	-2.99%	21	0.14%	90	-2.55E-06	320	9.00E-09	640	-3.50E-09
7	2.36%	22	-0.13%	100	-1.03E-06	340	2.00E-09	680	-2.84E-08
8	-1.29%	23	0.16%	110	-3.70E-07	360	-3.00E-10	720	-1.80E-09
9	0.88%	24	-0.14%	120	-1.01E-07	380	-5.60E-09	760	1.00E-08
10	-0.87%	25	0.11%	130	4.10E-08	400	3.00E-09	800	5.00E-09
11	1.12%	26	-0.08%	140	1.04E-07	420	-1.01E-08	840	1.70E-08
12	-0.99%	27	0.07%	150	1.15E-07	440	-1.48E-08	880	-3.49E-08
13	0.68%	28	-0.07%	160	1.21E-07	460	-9.40E-09	920	-1.67E-08
14	-0.36%	29	0.07%	180	8.70E-08	480	-1.93E-08	960	1.00E-08
15	0.33%	30	-0.06%	200	7.40E-08	500	-1.62E-08	1000	1.80E-08

Table 4: The relative error of $R_{\mathrm{h2}}(n)$ to $p(n)$ when $n \leqslant 1000$.

n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err
1	0	16	-0.43%	40	-1.87E-04	220	-2.76E-08	520	-2.42E-08
2	0	17	0.34%	50	-6.37E-05	240	-3.67E-08	540	-1.28E-08
3	0	18	-0.26%	60	-2.90E-05	260	-6.16E-08	560	-7.60E-09
4	0	19	0.20%	70	-1.08E-05	280	-6.93E-08	580	-3.19E-09
5	0	20	-0.16%	80	-4.43E-06	300	-6.77E-08	600	-1.21E-08
6	0	21	0.13%	90	-1.87E-06	320	-7.21E-08	640	1.96E-09
7	0	22	-0.10%	100	-6.03E-07	340	-3.74E-08	680	-8.54E-09
8	0	23	0.16%	110	-1.05E-07	360	-3.86E-08	720	1.30E-08
9	0	24	-0.13%	120	6.61E-08	380	-4.71E-08	760	-2.10E-09
10	0	25	0.10%	130	1.34E-07	400	-2.96E-08	800	-1.24E-08
11	1.79%	26	-8.21E-04	140	1.31E-07	420	-3.80E-08	840	1.25E-08
12	-1.30%	27	6.64E-04	150	1.07E-07	440	-2.31E-08	880	6.78E-09
13	0.99%	28	-8.07E-04	160	8.98E-08	460	-1.72E-08	920	1.52E-08
14	0	29	6.57E-04	180	3.94E-08	480	-2.31E-08	960	1.47E-08
15	0.57%	30	-7.14E-04	200	7.64E-09	500	-2.08E-08	1000	2.11E-08

Table 5: The relative error of $R'_{\mathrm{h2}}(n)$ to $p(n)$ when $n \leqslant 1000$.

The error of Equation (5.1) to $p(n)$ is shown on Table 4 on page 22. The accuracy is much better than Equation (1.8). Compared with Table 2 (page 20), the accuracy are almost the same when $n \leqslant 1000$. When $1500 \leqslant n \leqslant 10000$, the relative error is obviously less than that of Equation (4.2), as shown on Figure 5.2 on page 24 (compared with Figure 5.1 on page 24). Which means that $R_{\mathrm{h}2}(n)$ is more accurate than $R_{\mathrm{h}1}(n)$. (If we change the range of n of the data points, the accuracy of the fitting function obtained may not be so good.)

Consider that $p(n)$ is an integer, we can take the round approximation of Equation (5.1),

$$R'_{\mathrm{h}2}(n) = \left\lfloor \frac{\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)}{4\sqrt{3}\,(n + a_2\sqrt{n + c_2} + b_2)} + \frac{1}{2} \right\rfloor, \tag{5.2}$$

for small values of n. We may call it *Hardy-Ramanujan's revised estimation formula* 2 The error of Equation (5.2) to $p(n)$ is shown on Table 5 (on page 22) when $n \leqslant 1000$.

At the beginning of section 2, some other methods to estimate $p(n)$ are mentioned, such as estimating the value of $\dfrac{R_{\mathrm{h}}(n)}{p(n)}$ by a function $f_1(n)$, then estimate $p(n)$ by $\dfrac{R_{\mathrm{h}}(n)}{f_1(n)}$. The data $\left(n,\ \dfrac{R_{\mathrm{h}}(n)}{f_1(n)}\right)$ $(n = 20k + 100,\ k = 1,\ 2,\ \cdots,\ 395)$ are shown on Figure 5.4 on page 26 (together with a fitting function). It is not difficult to find out that a function

$$y = 1 + \frac{1}{\sqrt{a_3 x + b_3}},$$

where $a_3 = 5.062307637$ and $b_3 = -75.65700620$, will fit the data very well, as shown on the figure, with an average error 1.41×10^{-4}. (because the data $\left(n,\ \left(\frac{R_{\mathrm{h}}(n)}{f_1(n)} - 1\right)^{-2}\right)$ lies exactly on a straight line $y = a_3 x + b_3$, as shown on Figure 5.5 on page 26)

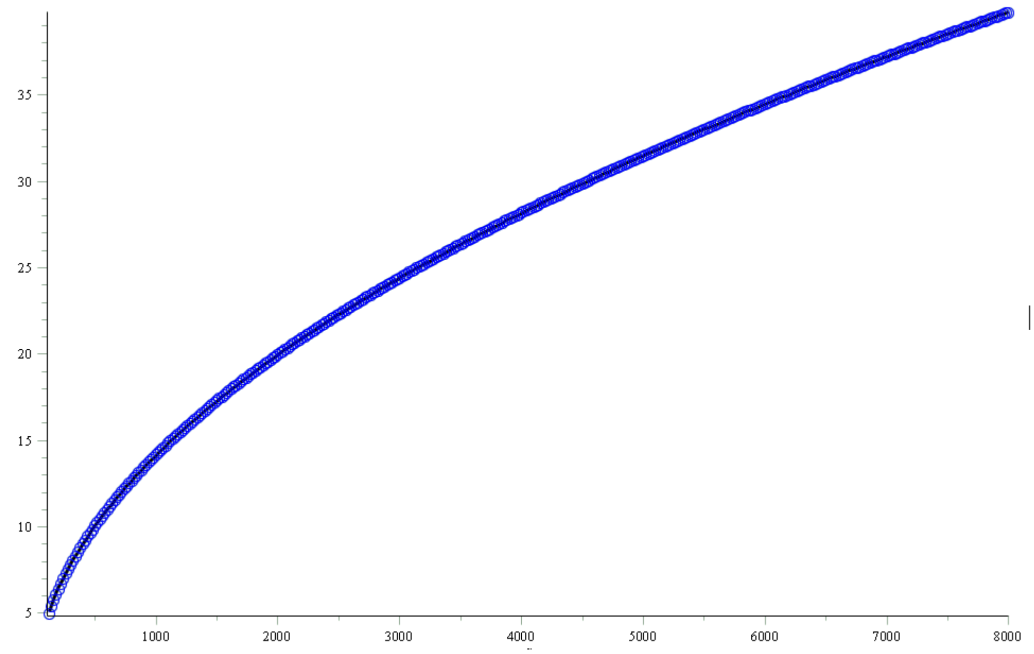

Figure 5.1: The graph of the data $\left(n,\ \frac{\exp\left(\pi\sqrt{\frac{2}{3}n}\right)}{4\sqrt{3}p(n)} - n\right)$

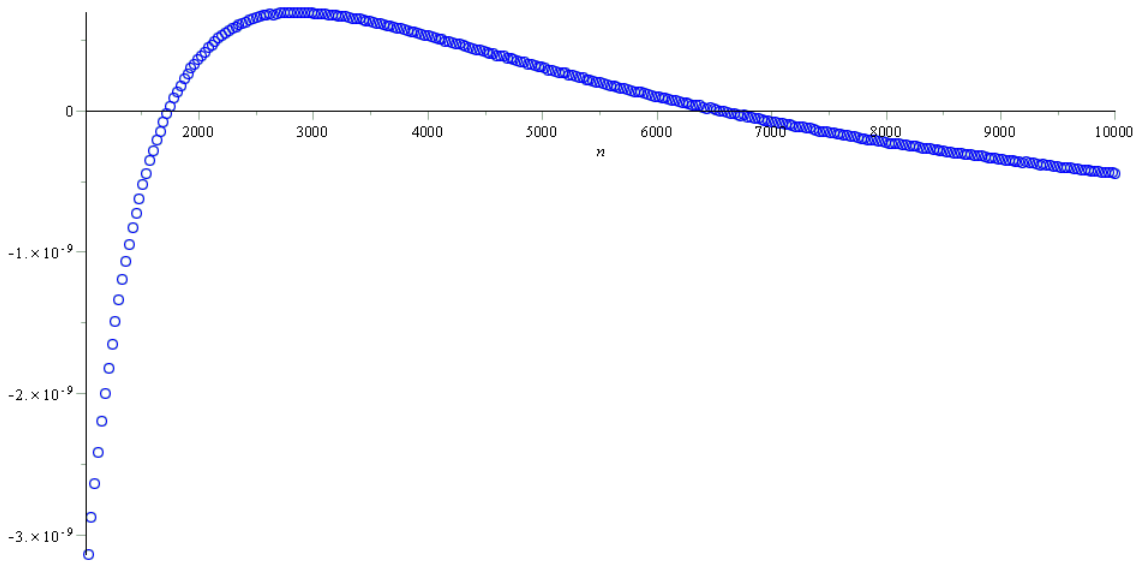

Figure 5.2: The Relative Error of $R_{\mathrm{h2}}(n)$ when $1000 \leqslant n \leqslant 10000$

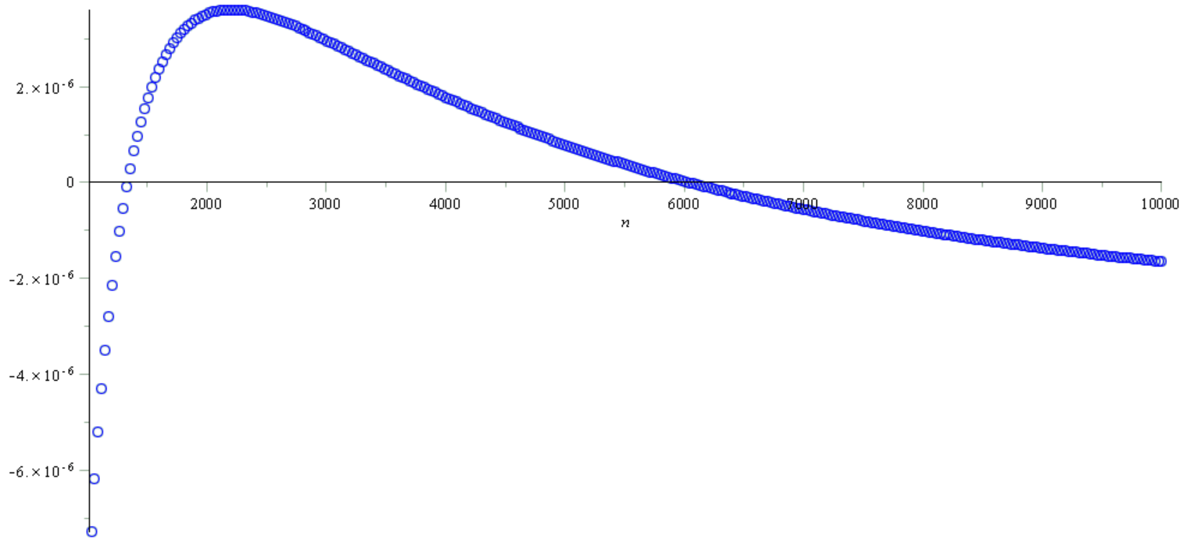

Figure 5.3: The Relative Error of $R_3(n)$ when $1000 \leqslant n \leqslant 10000$

So we have another fitting function for $p(n)$,

$$R_{d3}(n) = \frac{R_{\mathrm{h}}(n)}{1 + \dfrac{1}{\sqrt{a_3 x + b_3}}}.$$

However, this formula does not fit $p(n)$ very well when n is small. When $n \leqslant 14$, the value of $R_{d3}(n)$ is an imaginary number. Unfortunately, when $n > 1000$, the error of $R_{d3}(n)$ to $p(n)$ is about 1000 times of the error of $R_{\mathrm{h}2}(n)$, as shown on Figure 5.3 on page 25.

Actually, $R_{\mathrm{h}2}(n)$ is in the form $\dfrac{R_{\mathrm{h}}(n)}{f_1(n)}$, since $\dfrac{\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)}{4\sqrt{3}\,(n + a_2\sqrt{n + c_2} + b_2)} = \dfrac{\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)}{4\sqrt{3}n}$ $\dfrac{n}{n + a_2\sqrt{n + c_2} + b_2} = R_{\mathrm{h}}(n)\,\dfrac{1}{1 + \frac{a_2}{n}\sqrt{n + c_2} + \frac{b_2}{n}}$. As $1 + \frac{a_2}{n}\sqrt{n + c_2} + \frac{b_2}{n}$ fits $\dfrac{R_{\mathrm{h}}(n)}{p(n)}$ with very little error, $1 + \dfrac{1}{\sqrt{a_3 x + b3}}$ will not reach that accuracy.

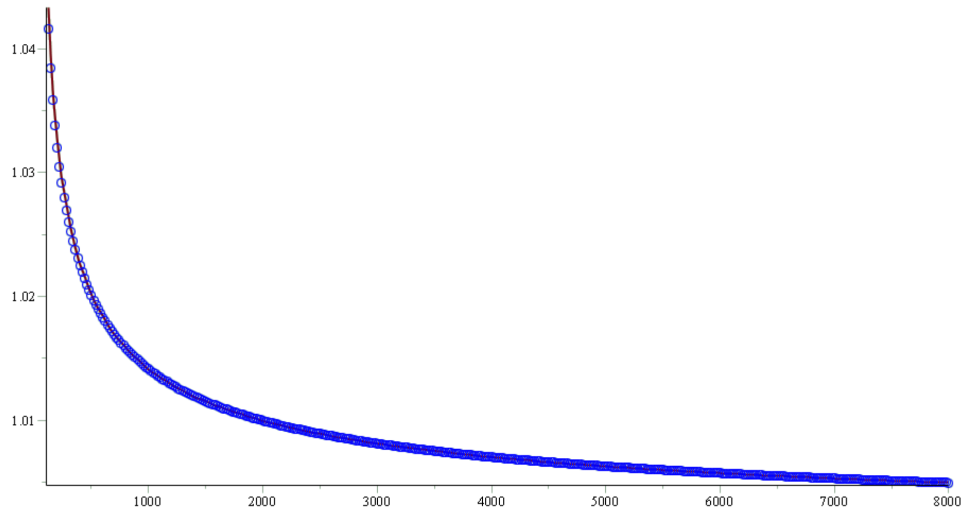

Figure 5.4: The graph of the data $\left(n,\ \dfrac{R_{\mathrm{h}}(n)}{p(n)}\right)$ and the fitting function

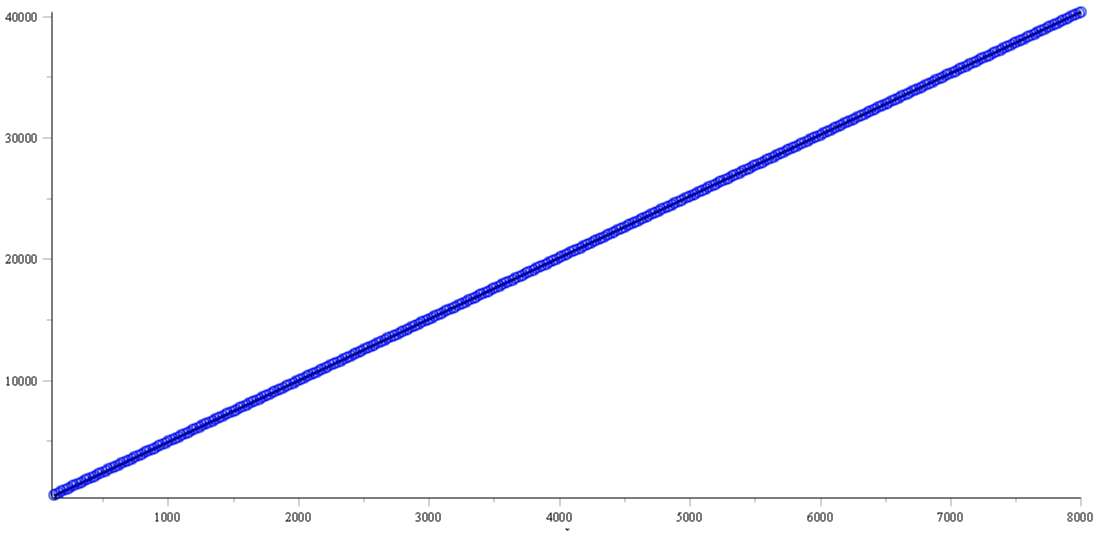

Figure 5.5: The data $\left(n,\ \left(\frac{R_{\mathrm{h}}(n)}{f_1(n)} - 1\right)^{-2}\right)$ and the fitting function

6 Estimate $p(n)$ by Fitting $R_{\mathrm{h}}(n) - p(n)$

The main idea to estimate $R_{\mathrm{h}}(n) - p(n)$ is similar to that introduced in the next subsection. By the same idea of the deduction of (6.1) on page 30, we wander whether we can fit $R_{\mathrm{h}}(n) - p(n)$ by $\dfrac{\pi}{12\sqrt{2C_3(n)}} \exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)$, where $C_3(n)$ is a cubic function, or equivalently, fit $\left(\dfrac{\pi \exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)}{12\sqrt{2}\,(R_{\mathrm{h}}(n) - p(n))}\right)^2$ by a cubic function $C_3(n)$, from the data with the data $(n,\, p(n))$ $(n = 20k + 60,\ k = 1,\, 2,\, \cdots,\, 397)$. If we do it, we will have a result

$$C_3(n) = a_1 n^3 + b_1 n^2 + c_1 n + d_1,$$

where

$$\begin{aligned} a_1 &= 8.383485427, \\ b_1 &= 130.0792015, \\ c_1 &= -1.197477259 \times 10^5, \\ d_1 &= 4.188653689 \times 10^7. \end{aligned}$$

Here c_1 and d_1 are very huge, which suggests that this result may not be so satisfying. As a sequence, if we fit $p(n)$ by

$$F_3(n) = R_{\mathrm{h}}(n) - \frac{\pi \exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)}{12\sqrt{2C_3(n)}},$$

the relative error differs very little with the relative error of $R_{\mathrm{h}}(n)$ to $p(n)$ when $n < 50$, and the relative error is not satisfying when $n < 280$, as shown in Table 6 on page 29.

If we fit $\left(\dfrac{\pi}{12\sqrt{2}\,(R_{\mathrm{h}}(n) - p(n))} \exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)\right)^2$ by a function like

$$C_3(n) = a_2 n^3 + b_2 n^{2.5} + c_2 n^2 + d_2 n^{1.5} + e_2 n + f_2 n^{0.5} + g_2,$$

the result are even worse, as imaginary number appeared (as concerned to the data mentioned in this section. If we fit less data, the imaginary problem might be avoid).

So we have consider a different method.

In the previous sub-subsection, we obtained the asymptotic order of $p(n) - p(n-1)$, and revised it to fit $h(n)$. Since $R_{\mathrm{h}}(n)$ is always greater than $p(n)$, we may guess that there is a t_0 such that $R_{\mathrm{h}}(n - t_0)$ is closer to $p(n)$ than $R_{\mathrm{h}}(n)$. Then we can revise the asymptotic order of $R_{\mathrm{h}}(n) - R_{\mathrm{h}}(n - t_0)$ and use the revised formula to fit $R_{\mathrm{h}}(n) - p(n)$.

By the algorithm mentioned on page 17, we can obtain the value $t_0 \doteq 0.3594143172$.

When $n \gg 1$,

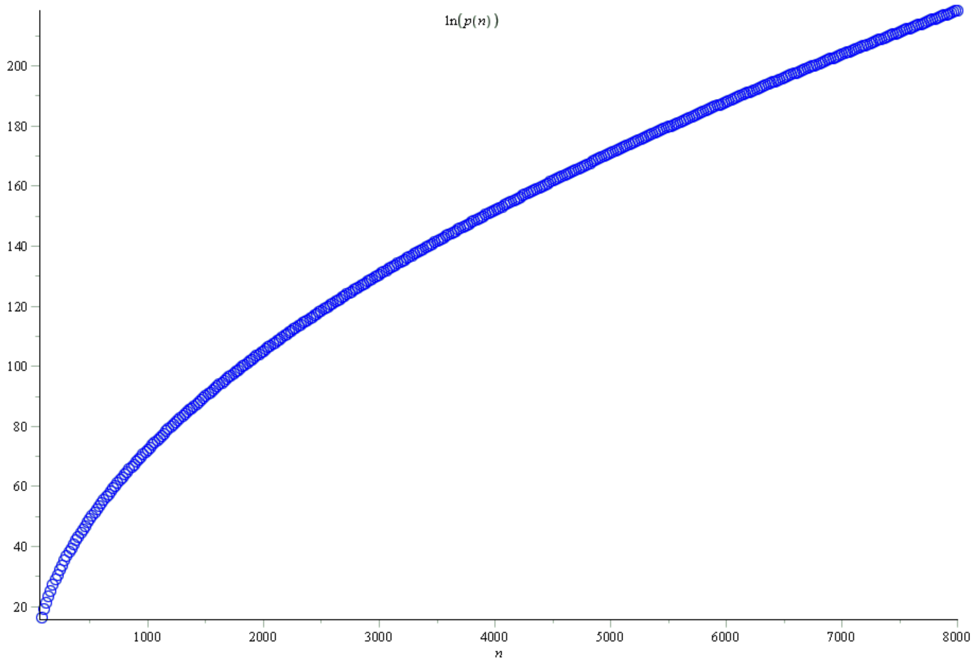

Figure 6.1: The graph of the data $(n, \ln(p(n)))$

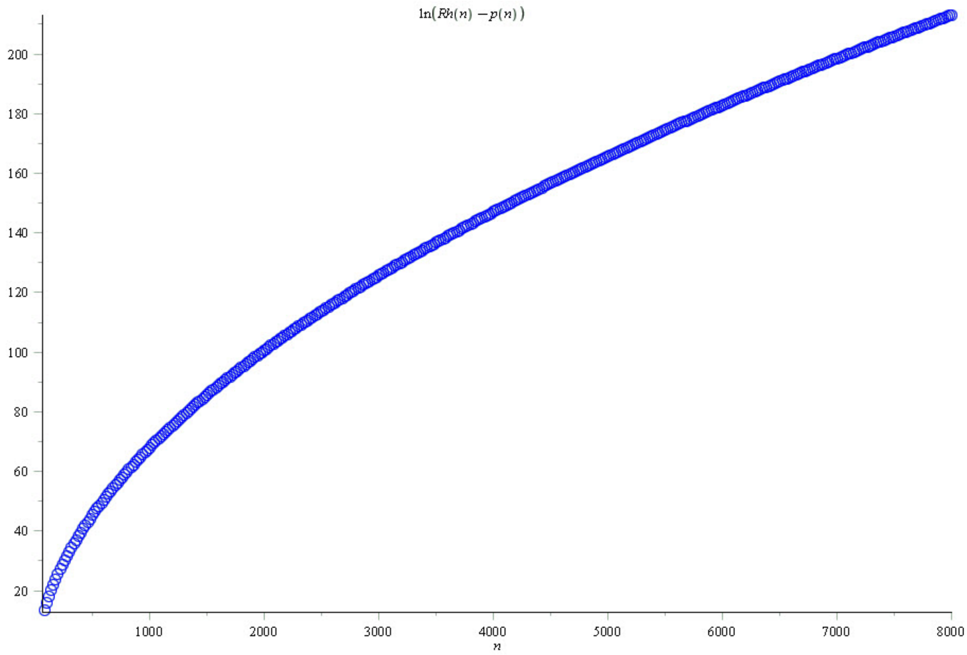

Figure 6.2: The graph of the data $(n, \ln(R_{\mathrm{h}}(n) - p(n)))$

n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err
1	14.93%	16	-0.36%	40	-0.02%	220	3.90E-08	520	-3.27E-08
2	-3.06%	17	0.37%	50	-6.80E-05	240	3.00E-08	540	-5.00E-09
3	3.96%	18	-0.31%	60	-3.16E-05	260	2.90E-08	560	-7.80E-09
4	-3.34%	19	0.27%	70	-1.24E-05	280	8.00E-09	580	3.00E-09
5	3.84%	20	-0.19%	80	-5.48E-06	300	-2.50E-09	600	-9.00E-10
6	-2.99%	21	0.14%	90	-2.55E-06	320	9.00E-09	640	-3.50E-09
7	2.36%	22	-0.13%	100	-1.03E-06	340	2.00E-09	680	-2.84E-08
8	-1.29%	23	0.16%	110	-3.70E-07	360	-3.00E-10	720	-1.80E-09
9	0.88%	24	-0.14%	120	-1.01E-07	380	-5.60E-09	760	1.00E-08
10	-0.87%	25	0.11%	130	4.10E-08	400	3.00E-09	800	5.00E-09
11	1.12%	26	-0.08%	140	1.04E-07	420	-1.01E-08	840	1.70E-08
12	-0.99%	27	0.07%	150	1.15E-07	440	-1.48E-08	880	-3.49E-08
13	0.68%	28	-0.07%	160	1.21E-07	460	-9.40E-09	920	-1.67E-08
14	-0.36%	29	0.07%	180	8.70E-08	480	-1.93E-08	960	1.00E-08
15	0.33%	30	-0.06%	200	7.40E-08	500	-1.62E-08	1000	1.80E-08

Table 6: The relative error of $F_3(n)$ to $p(n)$ when $n \leqslant 1000$.

$$r(n) = R_{\mathrm{h}}(n) - R_{\mathrm{h}}(n-t) = \frac{1}{4\sqrt{3}n}\exp\left(\sqrt{\tfrac{2}{3}}\pi\sqrt{n}\right) - \frac{1}{4\sqrt{3}(n-t)}\exp\left(\sqrt{\tfrac{2}{3}}\pi\sqrt{n-t}\right)$$

$$= \frac{1}{4\sqrt{3}}\exp\left(\sqrt{\tfrac{2}{3}}\pi\sqrt{n-t}\right)\left(\frac{\exp\left(\pi\sqrt{\tfrac{2}{3}}\left(\sqrt{n}-\sqrt{n-t}\right)\right)}{n} - \frac{1}{(n-t)}\right)$$

$$= \frac{1}{4\sqrt{3}}\exp\left(\sqrt{\tfrac{2}{3}}\pi\sqrt{n-t}\right)\left(\frac{\exp\left(\pi\sqrt{\tfrac{2}{3}}\frac{n-(n-t)}{\left(\sqrt{n}+\sqrt{n-t}\right)}\right)}{n} - \frac{1}{(n-t)}\right)$$

$$= \frac{1}{4\sqrt{3}}\exp\left(\sqrt{\tfrac{2}{3}}\pi\sqrt{n-t}\right)\left(\frac{\exp\left(\frac{t\pi\sqrt{2/3}}{\sqrt{n}+\sqrt{n-t}}\right)}{n} - \frac{1}{(n-t)}\right)$$

$$\sim \frac{1}{4\sqrt{3}}\exp\left(\sqrt{\tfrac{2}{3}}\pi\sqrt{n-t}\right)\left(\frac{\exp\left(\frac{t\pi\sqrt{2/3}}{2\sqrt{n-t/2}}\right)}{n} - \frac{1}{(n-t)}\right)$$

$$= \frac{1}{4\sqrt{3}}\exp\left(\sqrt{\tfrac{2}{3}}\pi\sqrt{n-t}\right)\left(\frac{\exp\left(\frac{t\pi}{\sqrt{6(n-t/2)}}\right)}{n}-\frac{1}{(n-t)}\right)$$

$$\sim \frac{1}{4\sqrt{3}}\exp\left(\sqrt{\tfrac{2}{3}}\pi\sqrt{n-t}\right)\left(\frac{1+\frac{t\pi}{\sqrt{6(n-t/2)}}}{n}-\frac{1}{(n-t)}\right)$$

$$= \frac{1}{4\sqrt{3}}\exp\left(\sqrt{\tfrac{2}{3}}\pi\sqrt{n-t}\right)\left(\frac{(n-t)\left(1+\frac{t\pi}{\sqrt{6(n-t/2)}}\right)-n}{n(n-t)}\right)$$

$$(\mathrm{e}^x \approx 1+x, \text{ when } x \ll 1.\ \frac{t\pi}{\sqrt{6n}} \ll 1, \text{ when } n \gg 1.)$$

$$= \frac{1}{4\sqrt{3}}\exp\left(\sqrt{\tfrac{2}{3}}\pi\sqrt{n-t}\right)\left(\frac{\frac{t\pi\sqrt{n+t/2}}{\sqrt{6}}-t+\frac{t^2\pi}{\sqrt{6(n-t/2)}}}{n(n-t)}\right)$$

$$\sim \frac{1}{4\sqrt{3}}\exp\left(\sqrt{\tfrac{2}{3}}\pi\sqrt{n-t}\right)\left(\frac{\frac{t\pi\sqrt{n+t/2}}{\sqrt{6}}}{n(n-t)}\right)$$

$$= \frac{1}{4\sqrt{3}}\exp\left(\sqrt{\tfrac{2}{3}}\pi\sqrt{n}\right)\left(\frac{t\pi}{\sqrt{6(n-t/2)}(n-t)}\right)$$

$$= \frac{t\pi\exp\left(\sqrt{\tfrac{2}{3}}\pi\sqrt{n-t}\right)}{12\sqrt{2}(n-t)\sqrt{(n-t/2)}}$$

$$\sim \frac{t\pi}{12\sqrt{2}\sqrt{n^3}}\exp\left(\sqrt{\tfrac{2}{3}}\pi\sqrt{n}\right).$$

As

$$r(n) \sim \frac{t\pi}{12\sqrt{2}(n-t)\sqrt{(n-t/2)}}\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n-t}\right) \sim \frac{t\pi}{12\sqrt{2}\sqrt{n^3}}\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right), \quad (6.1)$$

so we may consider to fit $R_\mathrm{h}(n)-p(n)$ by $\dfrac{\sqrt{2}t_0\pi\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n-t_0}\right)}{24C_4(n)}$, where

$$C_4(n) = a_2(x-t_0)^{1.5}+b_2(x-t_0)+c_2(x-t_0)^{0.5}+d_2.$$

n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err
1	1216.3%	16	-0.47%	40	-3.21E-04	220	-1.22E-06	520	6.00E-09
2	-20.86%	17	0.28%	50	-1.45E-04	240	-8.92E-07	540	4.10E-08
3	-2.09%	18	-0.40%	60	-8.10E-05	260	-6.50E-07	560	4.20E-08
4	-6.06%	19	0.20%	70	-4.59E-05	280	-4.94E-07	580	5.60E-08
5	2.14%	20	-0.26%	80	-2.94E-05	300	-3.71E-07	600	5.40E-08
6	-4.02%	21	0.08%	90	-2.02E-05	320	-2.60E-07	640	5.40E-08
7	1.60%	22	-0.19%	100	-1.44E-05	340	-1.92E-07	680	3.00E-08
8	-1.83%	23	0.12%	110	-1.07E-05	360	-1.35E-07	720	5.60E-08
9	0.47%	24	-0.19%	120	-8.21E-06	380	-9.63E-08	760	6.60E-08
10	-1.19%	25	0.07%	130	-6.44E-06	400	-5.29E-08	800	5.90E-08
11	0.86%	26	-0.11%	140	-5.14E-06	420	-3.87E-08	840	6.80E-08
12	-1.21%	27	0.04%	150	-4.17E-06	440	-2.24E-08	880	1.40E-08
13	0.50%	28	-0.10%	160	-3.41E-06	460	-4.00E-10	920	2.90E-08
14	-0.51%	29	0.05%	180	-2.37E-06	480	2.00E-09	960	5.40E-08
15	0.20%	30	-0.09%	200	-1.67E-06	500	1.50E-08	1000	5.90E-08

Table 7: The relative error of $R_{\text{h3}}(n)$ to $p(n)$ when $n \leqslant 1000$.

n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err
1	1200%	16	-0.43%	40	-3.21E-04	220	-1.22E-06	520	2.87E-08
2	0	17	0.34%	50	-1.47E-04	240	-8.99E-07	540	3.49E-08
3	0	18	-0.52%	60	-8.07E-05	260	-6.67E-07	560	4.01E-08
4	0	19	0.20%	70	-4.60E-05	280	-4.97E-07	580	4.35E-08
5	0	20	-0.32%	80	-2.94E-05	300	-3.70E-07	600	4.60E-08
6	0	21	0.13%	90	-2.02E-05	320	-2.74E-07	640	4.97E-08
7	0	22	-0.20%	100	-1.44E-05	340	-2.01E-07	680	5.10E-08
8	0	23	0.08%	110	-1.07E-05	360	-1.44E-07	720	5.08E-08
9	0	24	-0.19%	120	-8.21E-06	380	-1.01E-07	760	4.96E-08
10	0	25	0.05%	130	-6.44E-06	400	-6.67E-08	800	4.80E-08
11	0	26	-0.12%	140	-5.14E-06	420	-4.02E-08	840	4.60E-08
12	-1.30%	27	0.03%	150	-4.17E-06	440	-1.87E-08	880	4.42E-08
13	0.99%	28	-0.11%	160	-3.42E-06	460	-2.60E-09	920	4.19E-08
14	-0.74%	29	0.05%	180	-2.37E-06	480	1.05E-08	960	3.97E-08
15	0	30	-0.09%	200	-1.69E-06	500	2.08E-08	1000	3.75E-08

Table 8: The relative error of $R'_{\text{h3}}(n)$ to $p(n)$ when $n \leqslant 1000$.

When $t_0 \doteq 0.3594143172$, [14]it is not difficult to find out that

$$\begin{aligned} a_2 &= 1.039888529, \\ b_2 &= -0.3305606395, \\ c_2 &= 0.6134039843, \\ d_2 &= -0.8582793693, \end{aligned}$$

from the data $(n,\, p(n))$ ($n = 20k + 60$, $k = 1, 2, \cdots, 397$). Here none of the coefficients is very huge, which seems better than the previous method. As a matter of fact, if we estimate $p(n)$ by

$$R_{\mathrm{h}3}(n) = R_{\mathrm{h}}(n) - \frac{\sqrt{2}t_0\pi\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n-t_0}\right)}{24C_4(n)}, \tag{6.2}$$

the relative error is very small even when $n < 10$ (except the cases when $n = 1$ or 2) as shown on Table 7 on page 31. This is the first time to obtain an estimation formula of $p(n)$ which can reach a good accuracy without getting round approximation even when $n < 10$. This formula will be called *Hardy-Ramanujan's revised estimation formula* 3.

Further more, if we get the round value of $R_{\mathrm{h}3}(n)$,

$$R'_{\mathrm{h}3}(n) = \left\lfloor R_{\mathrm{h}}(n) - \frac{\sqrt{2}t_0\pi\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n-t_0}\right)}{24C_4(n)} + \frac{1}{2}\right\rfloor, \tag{6.3}$$

the relative error to error is even less, especially when $n = 15$ or $1 < n < 12$ it reaches 0, as shown on Table 8 on page 31. The relative error is less than 3×10^{-9} when $2500 < n < 10000$, as shown on Figure 6.1 on page 28.

Now that we can fit $R_{\mathrm{h}}(n) - p(n)$ by $\dfrac{\sqrt{2}t_0\pi\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n-t_0}\right)}{24C_4(n)}$, where $C_4(n) = a_2(x - t_0)^{1.5} + b_2(x - t_0) + c_2(x - t_0)^{0.5} + d_2$, maybe we can also fit $R_{\mathrm{h}}(n) - p(n)$ by $\dfrac{\pi}{12\sqrt{2}C_5(n)}\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)$ directly, where

$$C_5(n) = a_3n^{1.5} + b_3n + c_3n^{0.5} + d_3,$$

or equivalently, to fit $\dfrac{\pi}{12\sqrt{2}\left(R_{\mathrm{h}}(n) - p(n)\right)}\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)$ by a function $C_5(n)$ in the form mentioned above.

[14] In[19] (or [20]) or some other papers, there is a theoretic value $\dfrac{1}{24}$.

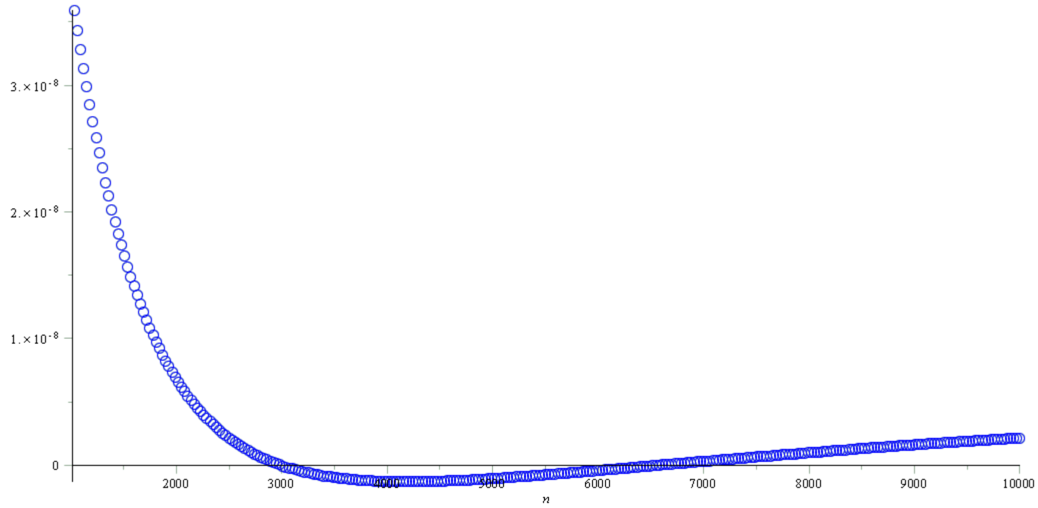

Figure 6.3: The Relative Error of $R_{\mathrm{h}3}(n)$ when $1000 \leqslant n \leqslant 10000$

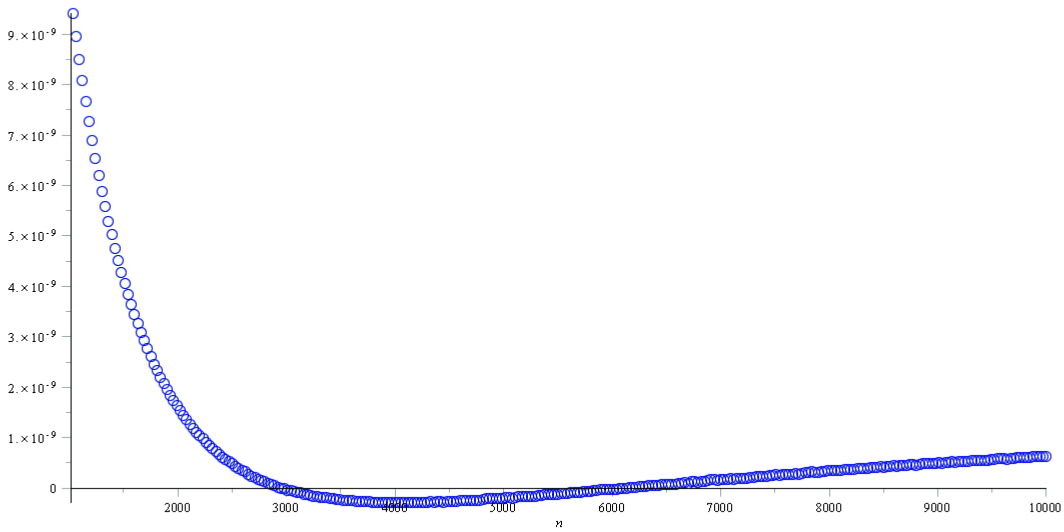

Figure 6.4: The Relative Error of $R_{\mathrm{h}4}(n)$ when $1000 \leqslant n \leqslant 10000$

We can easily obtain the unknown coefficients in the above equation by the least square method.

$$
\begin{aligned}
a_3 &= 2.893270736, \\
b_3 &= 0.4164546941, \\
c_3 &= -0.08501098214, \\
d_3 &= -0.4621004962.
\end{aligned}
$$

Again, none of the coefficients is very huge. As a result, the relative error of

$$R_{\mathrm{h}4}(n) = R_{\mathrm{h}}(n) - \frac{\pi \exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)}{12\sqrt{2}C_5(n)}, \tag{6.4}$$

to $p(n)$ is also very small when $n < 10$ (even in the cases when $n = 1$ or 2) as shown on Table 9 on page 35. This is the first time to obtain an estimation formula of $p(n)$ which can reach a good accuracy even when $n < 10$.

Further more, if we get the round value of $R_{\mathrm{h}4}(n)$,

$$R'_{\mathrm{h}4}(n) = \left\lfloor R_{\mathrm{h}}(n) - \frac{\pi \exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)}{12\sqrt{2}C_5(n)} + \frac{1}{2} \right\rfloor, \tag{6.5}$$

the relative error to error is even less, especially when $n = 15$ or $1 < n < 12$ it reaches 0, as shown on Table 10 on page 35. The relative error is less than 1×10^{-9} when $2500 < n < 10000$, as shown on Figure 6.2 on page 28. That is much better than $R_{\mathrm{h}4}(n)$ and $R'_{\mathrm{h}4}(n)$, besides, it is more simple. This formula will be called *Hardy-Ramanujan's revised estimation formula* 4.

n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err
1	5.4E-05	16	-0.38%	40	-0.02%	220	-2.16E-07	520	-7.00E-09
2	-5.53%	17	0.35%	50	-0.01%	240	-1.45E-07	540	2.10E-08
3	2.88%	18	-0.34%	60	-4.39E-05	260	-9.14E-08	560	1.80E-08
4	-3.88%	19	0.26%	70	-2.06E-05	280	-7.38E-08	580	2.90E-08
5	3.48%	20	-0.21%	80	-1.13E-05	300	-5.43E-08	600	2.50E-08
6	-3.21%	21	0.13%	90	-6.84E-06	320	-2.17E-08	640	2.10E-08
7	2.19%	22	-0.15%	100	-4.25E-06	340	-1.32E-08	680	-5.10E-09
8	-1.41%	23	0.15%	110	-2.84E-06	360	-3.70E-09	720	2.00E-08
9	0.79%	24	-0.15%	120	-2.02E-06	380	-5.00E-10	760	3.10E-08
10	-0.94%	25	0.10%	130	-1.48E-06	400	1.40E-08	800	2.50E-08
11	1.06%	26	-0.09%	140	-1.11E-06	420	7.00E-09	840	3.40E-08
12	-1.04%	27	0.07%	150	-8.66E-07	440	5.00E-09	880	-1.78E-08
13	0.64%	28	-0.08%	160	-6.78E-07	460	1.30E-08	920	-1.40E-09
14	-0.40%	29	0.07%	180	-4.51E-07	480	4.00E-09	960	2.40E-08
15	0.30%	30	-0.07%	200	-2.94E-07	500	9.00E-09	1000	3.10E-08

Table 9: The relative error of $R_{\mathrm{h4}}(n)$ to $p(n)$ when $n \leqslant 1000$.

n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err
1	0	16	-0.43%	40	-0.02%	220	-2.16E-07	520	1.49E-08
2	0	17	0.34%	50	-8.81E-05	240	-1.53E-07	540	1.55E-08
3	0	18	-0.26%	60	-4.35E-05	260	-1.08E-07	560	1.60E-08
4	0	19	0.20%	70	-2.05E-05	280	-7.63E-08	580	1.62E-08
5	0	20	-0.16%	80	-1.13E-05	300	-5.28E-08	600	1.63E-08
6	0	21	0.13%	90	-6.83E-06	320	-3.54E-08	640	1.61E-08
7	0	22	-0.10%	100	-4.26E-06	340	-2.24E-08	680	1.57E-08
8	0	23	0.16%	110	-2.84E-06	360	-1.25E-08	720	1.51E-08
9	0	24	-0.13%	120	-2.02E-06	380	-5.07E-09	760	1.44E-08
10	0	25	0.10%	130	-1.47E-06	400	6.22E-10	800	1.36E-08
11	1.79%	26	-0.08%	140	-1.12E-06	420	4.94E-09	840	1.28E-08
12	-1.30%	27	0.07%	150	-8.70E-07	440	8.22E-09	880	1.21E-08
13	0.99%	28	-0.08%	160	-6.89E-07	460	1.07E-08	920	1.13E-08
14	-0.74%	29	0.07%	180	-4.53E-07	480	1.25E-08	960	1.06E-08
15	0.57%	30	-0.07%	200	-3.10E-07	500	1.39E-08	1000	9.91E-09

Table 10: The relative error of $R'_{\mathrm{h4}}(n)$ to $p(n)$ when $n \leqslant 1000$.

7 Estimate $p(n)$ When $n \leqslant 100$

Until now, all the estimation function generated for $p(n)$ are with very good accuracy when n is greater than 100, but they are not so accurate when $n < 50$. Although $R'_{\mathrm{h2}}(n)$ and $R'_{\mathrm{h4}}(n)$ are better than others, the relative error are still greater than 1‰ for some values of n.

On the other hand, in subsections 3 and 4, when $n < 100$, it is nearly impossible to fit $C_1(n) \doteq \frac{3}{2} \cdot \frac{\left(\ln\left(4n\sqrt{3}p(n)\right)\right)^2}{\pi^2} - n$ or $C_2(n) \doteq \frac{\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n + C_1(n)}\right)}{4\sqrt{3}p(n)} - n$ by a simple piecewise function with less than 4 pieces and with high accuracy, as shown on Figure 2.3, Figure 2.4 (on page 9) and Figure 3.6 (on page 17), since the points do not lie on less than 4 smooth simple curves.

Can we reach a better accuracy when estimating $p(n)$ by a formula not too complicated?

In subsection 5, we fit the data $\left(n,\ \frac{\exp\left(\pi\sqrt{\frac{2}{3}n}\right)}{4\sqrt{3}p(n)} - n\right)$ $(n = 20k + 100,\ k = 1,\ 2,\ \cdots,\ 395)$ by a function and obtained a very good estimation of $p(n)$ when $n > 50$. So we wander whether we can fit the data $\left(n,\ \frac{\exp\left(\pi\sqrt{\frac{2}{3}n}\right)}{4\sqrt{3}p(n)} - n\right)$ $(n = 3,\ 4,\ \cdots,\ 100)$ by a piecewise function (with 2 pieces) so as to get a better estimation of $p(n)$ when $n \leqslant 100$?

The figure of the points of the data $\left(n,\ \frac{\exp\left(\pi\sqrt{\frac{2}{3}n}\right)}{4\sqrt{3}p(n)} - n\right)$ $(n = 3,\ 4,\ \cdots,\ 100)$ are shown on Figure 7.1 (on page 37). It is not difficult to find that the even points (where n is even) lie roughly on a smooth curve, so are the odd points. If we try to fit them respectively, we will have the fitting function below:

$$C'_2(n) = \begin{cases} 0.4527092482 \times \sqrt{n + 4.35278} - 0.05498719946, & n = 3, 5, 7, \cdots, 99; \\ 0.4412187317 \times \sqrt{n - 2.01699} + 0.2102618735, & n = 4, 6, 8 \cdots, 100. \end{cases} \tag{7.1}$$

Hence we can calculate $p(n)$ by

$$R_{\mathrm{h0}}(n) = \frac{\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)}{4\sqrt{3}\left(n + C'_2(n)\right)}, \quad 1 \leqslant n \leqslant 100. \tag{7.2}$$

Consider that $p(n)$ is an integer, we can take the round approximation of Equation (7.2),

$$R'_{h0}(n) = \left\lfloor \frac{\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n}\right)}{4\sqrt{3}\,(n + C'_2(n))} + \frac{1}{2} \right\rfloor, \quad 1 \leqslant n \leqslant 100. \tag{7.3}$$

The relative error of $R_{h0}(n)$ (or $R'_{h0}(n)$) to $p(n)$ are shown on Table 11 (or Table 12) on page 38. Compared with Table 5 on page 22, we will find that when $n \geqslant 80$, $R'_{h2}(n)$ is more accurate than $R'_{h0}(n)$; when $n \leqslant 50$, $R'_{h0}(n)$ is obviously better.

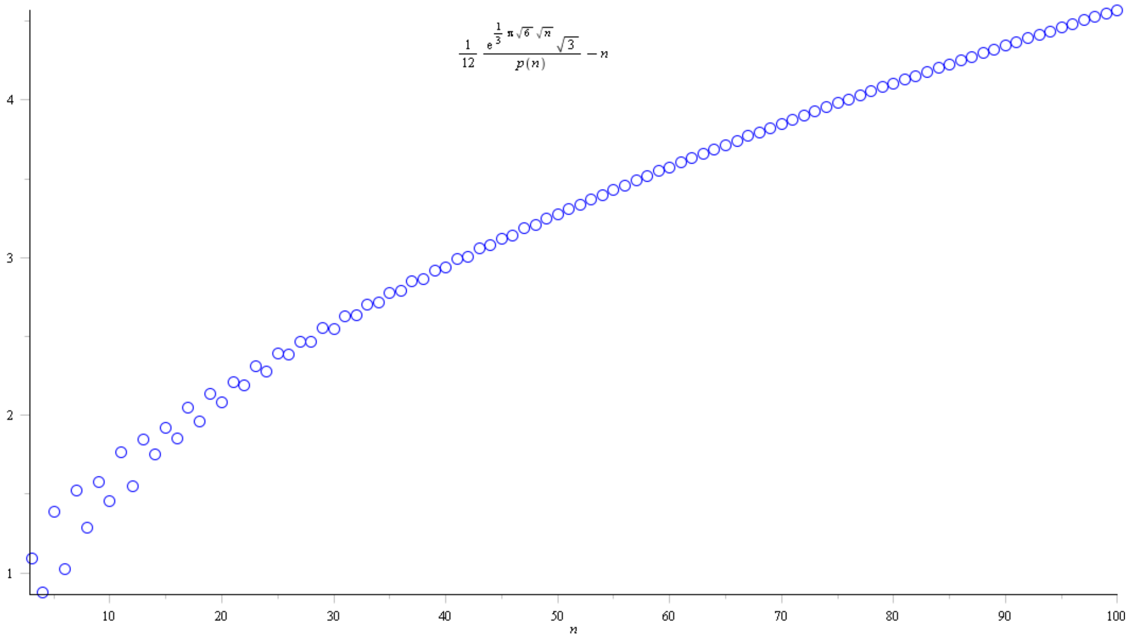

Figure 7.1: The graph of the data $\left(n, \dfrac{\exp\left(\sqrt{\frac{2}{3}}\pi\sqrt{n + C_1(n)}\right)}{4\sqrt{3}p(n)} - n\right)$ when $n \leqslant 100$

8 Summary

In this paper, we have presented several practical estimation formulae with high accuracy to calculated $p(n)$. When $n \leqslant 80$, we can use $R'_{h0}(n)$ (Equation (7.3)) , with a relative error less than 0.004%; when $n > 80$, we can use $R'_{h2}(n)$ (Equation (5.2)).

Equations (4.3), (6.3) and (6.5) are also very accurate although they are not as good as Equations (5.2).

n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err
1	-5.81%	21	-6.05E-04	41	-1.04E-04	61	-1.94E-06	81	2.44E-05
2	-	22	3.96E-04	42	1.54E-04	62	2.57E-05	82	-3.79E-05
3	-1.97%	23	-1.09E-05	43	-9.20E-05	63	1.72E-06	83	2.51E-05
4	1.00%	24	-9.13E-06	44	1.55E-04	64	1.58E-05	84	-4.18E-05
5	0.90%	25	-2.40E-04	45	-9.44E-05	65	8.04E-06	85	2.52E-05
6	-0.91%	26	4.12E-04	46	1.36E-04	66	6.33E-06	86	-4.48E-05
7	0.64%	27	-3.55E-04	47	-6.15E-05	67	1.17E-05	87	2.50E-05
8	-0.03%	28	2.98E-04	48	1.05E-04	68	2.90E-08	88	-4.79E-05
9	-0.23%	29	-1.64E-04	49	-5.30E-05	69	1.40E-05	89	2.49E-05
10	-0.03%	30	1.92E-04	50	1.02E-04	70	-7.08E-06	90	-5.08E-05
11	0.34%	31	-1.84E-04	51	-4.78E-05	71	1.78E-05	91	2.44E-05
12	-0.40%	32	2.86E-04	52	8.40E-05	72	-1.43E-05	92	-5.31E-05
13	0.12%	33	-2.46E-04	53	-3.03E-05	73	1.98E-05	93	2.37E-05
14	0.08%	34	2.78E-04	54	6.66E-05	74	-1.91E-05	94	-5.54E-05
15	-0.10%	35	-1.52E-04	55	-2.27E-05	75	2.11E-05	95	2.31E-05
16	-1.91E-04	36	1.84E-04	56	5.79E-05	76	-2.47E-05	96	-5.75E-05
17	4.63E-04	37	-1.47E-04	57	-1.82E-05	77	2.30E-05	97	2.22E-05
18	-4.89E-04	38	2.30E-04	58	4.62E-05	78	-2.98E-05	98	-5.92E-05
19	1.82E-04	39	-1.52E-04	59	-7.01E-06	79	2.40E-05	99	2.11E-05
20	1.96E-04	40	1.88E-04	60	3.21E-05	80	-3.39E-05	100	-6.09E-05

Table 11: The relative error of $R_{\mathrm{h0}}(n)$ to $p(n)$ when $n \leqslant 100$.

n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err	n	Rel-Err
1	0	21	0	41	-1.12E-04	61	-1.78E-06	81	2.44E-05
2	0	22	0	42	1.50E-04	62	2.54E-05	82	-3.79E-05
3	0	23	0	43	-9.48E-05	63	1.99E-06	83	2.52E-05
4	0	24	0	44	1.60E-04	64	1.55E-05	84	-4.18E-05
5	0	25	0	45	-8.98E-05	65	7.95E-06	85	2.52E-05
6	0	26	4.11E-04	46	1.33E-04	66	6.46E-06	86	-4.48E-05
7	0	27	-3.32E-04	47	-6.41E-05	67	1.16E-05	87	2.50E-05
8	0	28	2.69E-04	48	1.09E-04	68	0	88	-4.79E-05
9	0	29	-2.19E-04	49	-5.19E-05	69	1.41E-05	89	2.49E-05
10	0	30	1.78E-04	50	1.03E-04	70	-7.09E-06	90	-5.08E-05
11	0	31	-1.46E-04	51	-4.58E-05	71	1.79E-05	91	2.44E-05
12	0	32	2.40E-04	52	8.52E-05	72	-1.43E-05	92	-5.31E-05
13	0	33	-1.97E-04	53	-3.03E-05	73	1.97E-05	93	2.37E-05
14	0	34	2.44E-04	54	6.73E-05	74	-1.92E-05	94	-5.54E-05
15	0	35	-1.34E-04	55	-2.22E-05	75	2.11E-05	95	2.31E-05
16	0	36	1.67E-04	56	5.88E-05	76	-2.47E-05	96	-5.75E-05
17	0	37	-1.39E-04	57	-1.79E-05	77	2.30E-05	97	2.22E-05
18	0	38	2.31E-04	58	4.61E-05	78	-2.98E-05	98	-5.92E-05
19	0	39	-1.60E-04	59	-7.21E-06	79	2.40E-05	99	2.11E-05
20	0	40	1.87E-04	60	3.21E-05	80	-3.39E-05	100	-6.09E-05

Table 12: The relative error of $R'_{\mathrm{h0}}(n)$ to $p(n)$ when $n \leqslant 100$.

Acknowledgements

The author would like to express the gratitude to Prof. *Ian M. Wanless* from Monash University in Australia and Prof. Jack H. Koolen from USTC for their valuable advice.

References

[1] Anonymous, *“Bibliography on Partitions" in “Mathematical BBS"*, Internet: http://felix.unife.it/Root/d-Mathematics/d-Number-theory/b-Partitions (accessed September 12, 2016), December 2007.

[2] ______, *Fractal Structure to Partition Function: Hidden structure in partition function* , Internet: www.aimath.org/news/partition (accessed Noverber 13, 2015), December 2013.

[3] Tom M. Apostol, *Functions of Number Theory, Additive Number Theory: Unrestricted Partitions*, NIST Digital Library of Mathematical Functions (DLMF) (Frank W. J. Olver, Daniel W. Lozier, and Ronald F. Boisvert, eds.), August 2015, http://dlmf.nist.gov/27.14, Release 1.0.10 of 2015-08-07, (accessed November 12, 2015).

[4] Chris Berg, Christian Stump, Lahiru Kariyawasam, Garrett Filipek, and Chase Reuter, *Integer Partitions* , http://www.findstat.org/IntegerPartitions (accessed September 8, 2016), Oct. 2015.

[5] Jan Hendrik Bruinier, Amanda Folsom, Zachary A. Kent, and Ken Ono, *Recent Work on the Partition Function*, Proceedings of the Legacy of Ramanujan Conference, New Delhi, (Bruce C. Berndt and Dipendra Prasad, eds.), Ramanujan Mathematical Society Lecture Notes, vol. 20, Ramanujan Mathematical Society (RMS), 2013, pp. 139 – 151.

[6] Frank Calegari, *A Remark on a Theorem of Folsom, Kent, and Ono*, Internet: http://www.math.uchicago.edu/~fcale/Files/FKO.pdf (accessed November 13, 2015), Jan. 2011.

[7] Carol Clark, *New math theories reveal the nature of numbers: Finite formula found for partition numbers* , Internet: https://www.eurekalert.org/pub_releases/2011-01/eu-nmt011911.php (accessed September 8, 2016), Jan. 2011.

[8] ______, *New theories reveal the nature of numbers* , Internet: http://www.emory.edu/ EMORY_REPORT/stories/2011/01/research_ ken_ono_ math_ partition_ numbers.html (accessed September 8, 2016), Jan. 2011.

[9] Pál Erdős, *The Evaluation of the Constant in the Formula for the Number of Partitions of n*, Annals of Mathematics. Second Series **43** (1942), no. 3, 437 – 450 (English), MSC: 11P82 11P81, Zbl: 0061.07905.

[10] Amanda Folsom, Zachary A. Kent, and Ken Ono, *l-adic properties of the partition function*, Advances in Mathematics **229** (2012), no. 03, 1586 – 1609.

[11] Godfrey Harold Hardy and Srinivasa R. Ramanujan, *Asymptotic Formulae in Combinatory Analysis*, Proceedings of the London Mathematical Society, s2 **XVII** (1918), 75 – 115.

[12] Marshall Hall, Jr., *A survey of combinatorial analysis*, Some aspects of analysis and probability (Irving Kaplansky and etc, eds.), Surveys in applied mathematics, vol. IV, John Wiley and Sons, Inc. [New York] and Chapman and Hall, Limited [London], 1958, pp. 35 – 104.

[13] Donald J. Newman, *The evaluation of the constant in the formula for the number of partitions of n*, American Journal of Mathematics **73** (1951), no. 3, 599 – 601.

[14] ———, *A simplified proof of the partition formula*, The Michigan Mathematical Journal **9** (1962), no. 3, 283 – 287, MR0142529, ZMI 0105.26701.

[15] Ken Ono, *Arithmetic of the Partition Function*, Special Functions 2000: Current Perspective and Future Directions (Joaquin Bustoz, Mourad E. H. Ismail, and Sergei K. Suslov, eds.), NATO Science Series, Series II: Mathematics, Physics and Chemistry, vol. 30, Kluwer Academic Publishers, 2001, pp. 243 – 253.

[16] Hans Rademacher, *A Convergent Series for the Partition Function p(n)*, Proceedings of the National Academy of Sciences of the United States of America **23** (1937), no. 2, 78 – 84, PMC1076871.

[17] Neil James Alexander Sloane, *A000041, number of partitions of n (the partition numbers). (Formerly M0663 N0244) , in "The On-Line Encyclopedia of Integer Sequences (OEIS ®)"*, Internet: http://oeis.org/A000041 (accessed September 20, 2015).

[18] ———, *"A002865 as a graph" in "The On-Line Encyclopedia of Integer Sequences (OEIS ®)"*, Internet: http://oeis.org/A002865/graph (accessed October 10, 2015).

[19] József Sándor, Dragoslav S. Mitrinovic, and Borislav Crstici, *Unrestricted partitions of an integer*, Handbook of Number Theory, Volume 1 (Michiel Hazewinkel. etc., ed.), Mathematics and Its Applications, vol. 351, Kluwer Academic Publishers Group, Dordrecht, 1995, § XIV.1.

[20] ———, *Unrestricted partitions of an integer*, Handbook of Number Theory I, Springer Netherlands, 2006, Section XIV.1, p. 491.

[21] Eric W. Weisstein, *"Partition Function P." From MathWorld – A Wolfram Web Rdesource*, Internet: http://mathworld.wolfram.com/PartitionFunctionP.html (accessed September 20, 2015), 1999-2015.