ON RAMANUJAN'S Q–FUNCTION†

Philippe Flajolet, Peter J. Grabner,
Peter Kirschenhofer and Helmut Prodinger

Dedicated to D. E. Knuth, on the occasion of the 30th Anniversary of his first analysis of an algorithm in 1962.

Abstract. This study provides a detailed analysis of a function which Knuth discovered to play a central rôle in the analysis of hashing with linear probing. The function, named after Knuth $Q(n)$, is related to several of Ramanujan's investigations. It surfaces in the analysis of a variety of algorithms and discrete probability problems including hashing, the birthday paradox, random mapping statistics, the "rho" method for integer factorization, union–find algorithms, optimum caching, and the study of memory conflicts.

A process related to the complex asymptotic methods of singularity analysis and saddle point integrals permits to precisely quantify the behaviour of the $Q(n)$ function. In this way, tight bounds are derived. They answer a question of Knuth (*The Art of Computer Programming*, Vol. 1, 1968, [Ex. 1.2.11.3.13]), itself a rephrasing of earlier questions of Ramanujan in 1911–1913.

— August 25, 1993 —

1. Introduction

In the 1911 issue of the Journal of the Indian Mathematical Society, Ramanujan [Ra11] poses the following problem: *Show that*

$$\frac{1}{2}e^n = 1 + \frac{n}{1!} + \frac{n^2}{2!} + \cdots + \frac{n^n}{n!}\theta, \qquad \textit{where } \theta \textit{ lies between } \tfrac{1}{2} \textit{ and } \tfrac{1}{3}. \tag{1.1}$$

A solution was then outlined in [Ra12]. Later in his first letter to Hardy dated January 16, 1913 (see [Ra62, p. xxvi], [Be69, p. 181] and [Ha78]), Ramanujan makes a stronger assertion, namely that

$$\theta = \frac{1}{3} + \frac{4}{135(n+k)} \qquad \textit{where } k \textit{ lies between } \tfrac{8}{45} \textit{ and } \tfrac{2}{21}. \tag{1.2}$$

---

†Research of the four authors was supported by the French Austrian scientific cooperation programme. Work of P. Flajolet was supported in part by the ESPRIT III Basic Research Action Programme of the E.C. under contract ALCOM II (#7141). Work of P. Grabner was supported by the Austrian Science Foundation, Project Nr. P8274-PHY

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TeX

It is immediately clear nowadays, on probabilistic grounds, that the left hand side member in (1.1) is approximated by the sum of the first n terms on the right. In fact, the sum $1+\cdots+n^{n-1}/(n-1)!$ multiplied by e^{-n} represents the probability that a random variable with a Poisson distribution of mean n be less than n. But by the Gaussian approximation of Poisson laws of large mean, this probability is close to $\frac{1}{2}$ and the argument supplemented by elementary real estimates readily shows that $\theta \equiv \theta(n) = O(1)$. Thus Ramanujan's assertions appear as asymptotic refinements of a basic probabilistic observation.

Berndt's scholarly edition of Ramanujan's Notebooks contains a complete bibliography on the original problem which is closely related to Entries 47 and 48 of Chapter 12 (p. 179–184 of [Be89]) as well as to Entry 6 of Chapter 13 (p. 193–195 of [Be89]). Berndt qualifies the problem as *"an ultimately famous problem"*. In his partial answer [Ra12] and in the notebooks, Ramanujan provides in essence a way of constructing an asymptotic expansion for $\theta \equiv \theta(n)$. ¿From his analysis there results that $\theta(\infty) = \frac{1}{3}$ while we have $\theta(0) = \frac{1}{2}$. This, translated in terms of $k \equiv k(n)$ says that $k(\infty) = \frac{2}{21}$ and $k(0) = \frac{8}{45}$, and supported by a few initial value computations, must have naturally led Ramanujan to his assertion (1.2).

A solution to the weaker inequality (1.1) was given by G. Szegő in 1928 [Sz28], and almost simultaneously Watson [Wa29] wrote a paper where he proved (1.1) and adds regarding (1.2): *"I shall also give reasons, which seem to me fairly convincing, for* believing *that k lies between* 8/45 *and* 2/21*."* Our purpose here is to finally provide a complete proof of Ramanujan's assertion (1.2).

In 1962, which is exactly 50 years after Ramanujan's original note [Ra12] and 75 years after Ramanujan's birth, Knuth conducted his first average–case analysis of an algorithm, namely the analysis of *hashing with linear probing* (see the footnote on page 529 of [Kn73]). That analysis is given in full in Volume 3 of *The Art of Computer Programming* [Kn73] and, in Volume 1, Knuth uses it as an illustration of asymptotic analysis techniques [Kn68, p.113]. A key rôle in the analysis is played by a function closely related to $\theta(n)$, the $Q(n)$–function, and in Exercise 1.2.11.3.13, Knuth asks for a final solution to Ramanujan's problem (1.2).

The variant form used by Knuth introduces the two functions

$$\begin{aligned} Q(n) &= 1 + \frac{n-1}{n} + \frac{(n-1)(n-2)}{n^2} + \cdots \\ R(n) &= 1 + \frac{n}{n+1} + \frac{n^2}{(n+1)(n+2)} + \cdots, \end{aligned} \tag{1.3}$$

and one finds easily

$$Q(n) + R(n) = n!e^n/n^n.$$

Thus $Q(n)$ and $R(n)$ are closely related, and the asymptotics of one of them determines the other. The relation of Knuth's definition (1.3) to Ramanujan's problem should also be clear, considering the equality

$$\frac{n^n}{n!}Q(n) = \frac{n^{n-1}}{(n-1)!} + \frac{n^{n-2}}{(n-2)!} + \cdots + 1, \tag{1.4}$$

which entails

$$\frac{n^n}{n!}Q(n) = \frac{1}{2}e^n - \theta\frac{n^n}{n!} \qquad \text{and} \qquad \theta(n) = \frac{1}{2}(R(n) - Q(n)).$$

In this way, Ramanujan's problem can be rephrased as: *"Show that*

$$R(n) - Q(n) = \frac{2}{3} + \frac{8}{135(n+k)}$$

where $k \equiv k(n)$ lies between $\frac{2}{21}$ and $\frac{8}{45}$". Following Knuth, this is the form that we shall take as our starting point, setting $D(n) = R(n) - Q(n)$, so that $D(n) = 2\theta(n)$.

The approaches followed by Ramanujan himself and later authors all make use of *real* integral representations derived from

$$Q(n) = \int_0^\infty e^{-x}\,(1 + \frac{x}{n})^n\,dx, \tag{1.5}$$

and proceed using the Laplace method for the asymptotic evaluation of integrals [DB81]. ¿From this representation, $Q(n)$ is up to normalization an incomplete gamma function, a ${}_1F_1$–hypergeometric, and accordingly, it admits an explicit continued fraction representation of the Gauss type, as already noticed by Ramanujan (Entry 47 in Chapter 12 of [Be89]).

As a historical fantasy, it may be of interest to observe that, quite possibly, Ramanujan was led to his conjecture by considerations related to the distribution of prime divisors in integers. The Erdős-Kac Theorem asserts that, for integers, the 'number-of-divisors' function is asymptotically Gaussian distributed over large ranges. A form of this theorem does appear in Ramanujan's notes, under a Poisson formulation: *The number of integers less than n with numbers of prime divisors at most k is asymptotically*

$$\frac{x}{\log x}\left[1 + \frac{\log\log x}{1!} + \frac{(\log\log x)^2}{2!} + \cdots + \frac{(\log\log x)^k}{k!}\right].$$

Letting k vary with n in the "interesting" region $k \approx \log\log x$ naturally leads to questions like (1.1).

It should also be said at this stage that, apart from sentimental value, the $Q(n)$ function appears in a number of problems in discrete probability and the analysis of algorithms:

(1) $1 + Q(n)$ is the expected number of persons it takes in order to find two having the same birth date (when the year has n days!). This is the classical birthday problem which is of interest in random allocations and general hashing algorithms (see Section 4.1 of [Kn73] or [FGT92]). For instance, on earth, it takes on average only 24.61658 persons to find two with the same birth date.

(2) The analysis of hashing with linear probing, when the table is full, is expressed simply in terms of $Q(n)$. The cost of a successful search is about $\frac{1}{2}\sqrt{\frac{\pi n}{2}}$, as results from the asymptotic analysis of $Q(n)$ (see [Kn73, p. 530] and [VF90, p. 509–511]).

(3) Random mapping statistics involve $Q(n)$ in several places, in relation to their cycle structure. There is a vast literature on this topic, see [Kn69, p. 8] for the basic results, [KP89] for related problems, and [FO90a] for a

recent survey of random mapping statistics. The corresponding analyses are also relevant to the study of random number generators [Kn69, Sec. 3.1]. It is starting from there that Pollard conceived one of his integer factorization algorithms, the well known "rho method", which itself permitted in its time the discovery of the factorization of the eighth Fermat number $F_8 = 2^{2^8} + 1$.

(4) Analysis of Union–Find algorithms under the random spanning tree model essentially depends on $Q(n)$ [KS78].

(5) Analysis of optimum caching [Kn85] and the study of memory conflicts or deadlocks [KR75,CR92] involve the function $Q(n)$.

The paper is organized as follows. Section 2 develops a complex integral representation based on a generating function of the $Q(n)$ from which yet another proof of its asymptotic expansion follows. Our approach is in fact a hybrid of singularity analysis and saddle point in the following sense: It starts with an integral representation based on an expansion essentially dictated by the singularity analysis method (see Eq. (2.4), and the proof of Proposition 2); then a suitable change of variable is introduced that causes the integration contour to pass through a saddle point (see Eq. (2.5),(3.1),(3.2) and the developments at the beginning of Section 3). In Section 3, effective error bounds for the intervening integrals are developed. The various lemmas and estimates are then woven together in Section 4, where the proof of the main inequality is completed. Section 5 offers a few hindsights.

2. Generating Functions, Asymptotics, and an Integral Representation

An important function in combinatorial analysis is the function $y(z)$ defined implicitly by the equation

$$y(z) = ze^{y(z)}, \tag{2.1}$$

with $y(z) = z + z^2 + 3z^3/2 + \cdots$. By the Lagrange inversion formula, we have[1]

Proposition 1. *The Taylor coefficients of $y(z) = ze^{y(z)}$ and its powers are given by*

$$[z^n]y(z) = \frac{n^{n-1}}{n!} \qquad \text{and} \qquad [z^n]y^k(z) = k\frac{n^{n-k-1}}{(n-k)!}. \tag{2.2}$$

Furthermore, a generating function of $Q(n)$ is expressible in terms of $y(z)$:

$$\sum_{n=1}^{\infty} Q(n)n^{n-1}\frac{z^n}{n!} = \log\frac{1}{1-y(z)}\,. \tag{2.3}$$

The well known expansions (2.2) go back to Legendre and Eisenstein. The companion facts from combinatorics are classical (see [GJ83] or [FO90a]): $y(z)$ is the exponential generating function (EGF) of rooted labelled trees and the function

[1] We let $[z^n]f(z)$ denote the coefficient of z^n in the Taylor expansion of $f(z)$.

appearing in (2.3), $L(z) = \log(1 - y(z))^{-1}$, is the EGF of functional graphs (mappings) that are connected. Thus, $Q(n)/n$ also represents the probability that a random mapping be connected, a result already known to Rényi and Harris.

Equation (2.3) has already appeared in the literature [KS78,FO90a,KP89,VF90]. It constitutes the starting point of a transparent analysis of the asymptotics of $Q(n)$. In this way, we rederive through generating functions the asymptotic expansion of $Q(n)$ known from the works of Ramanujan, Watson, and Knuth [Be89,Wa29,Kn68].

Theorem 1. [Ramanujan, Watson, Knuth] *The quantities $Q(n), R(n)$ admit full asymptotic expansions in descending powers of $\sqrt{n}$:*

$$Q(n) \sim \sqrt{\frac{\pi n}{2}} - \frac{1}{3} + \frac{1}{12}\sqrt{\frac{\pi}{2n}} - \frac{4}{135n} + \cdots$$

$$R(n) \sim \sqrt{\frac{\pi n}{2}} + \frac{1}{3} + \frac{1}{12}\sqrt{\frac{\pi}{2n}} + \frac{4}{135n} + \cdots .$$

Proof. We sketch here the proof based on singularity analysis (see [VF90,KP89,FO90a] for related developments). The implicitly defined function $y(z)$ has a singularity of the square–root type at $z = 1/e$, and

$$y(z) = 1 - \varepsilon + \frac{1}{3}\varepsilon^2 - \frac{11}{72}\varepsilon^3 + \frac{43}{540}\varepsilon^4 - \cdots, \qquad \varepsilon = \sqrt{2 - 2ez},$$

as $z \to e^{-1}$. This induces a logarithmic singularity for $L(z)$ at $z = 1/e$,

$$L(z) = \log\frac{1}{\varepsilon} + \frac{1}{3}\varepsilon - \frac{7}{72}\varepsilon^2 + \frac{133}{3240}\varepsilon^3 + \cdots = \frac{1}{2}\log(1 - ez)^{-1} + \cdots .$$

By singularity analysis [FO90b], the asymptotic equivalent of $L(z)$ transfers to the coefficients, which gives $[z^n]L(z) \sim \frac{1}{2}e^n n^{-1}$, and $Q(n) \sim \sqrt{\pi n/2}$. The full expansion of $L(z)$ in powers of $(1 - ez)^{1/2}$ also translates into a full asymptotic expansion of $Q(n)$ in powers of $n^{1/2}$. The developments for $R(n)$ are entirely similar. □

In preparation for the more detailed treatment involving bounds for $D(n)$, we next need to make fully explicit the various expansions and representations related to its generating function.

We have $Q(n) + R(n) = n!e^n/n^n$ so that a generating function of $D(n)$ is

$$\sum_{n=1}^{\infty} D(n)n^{n-1}\frac{z^n}{n!} = \log\frac{(1-y)^2}{2(1-ez)}.$$

Appeal to Cauchy's integral formula for coefficients of analytic functions. When applied to (2.3), it gives

$$D(n) = \frac{n!}{n^{n-1}}\frac{1}{2\pi i}\oint_{\mathcal{C}} \log\frac{(1-y(z))^2}{2(1-ez)}\, d\omega \qquad \text{with} \qquad d\omega = \frac{dz}{z^{n+1}}, \tag{2.4}$$

where $\mathcal{C}$ is a sufficiently small countour surrounding the origin.

A key step is now to allow y to be taken as the independent variable (this is precisely inspired by the analytic proof of the Lagrange inversion theorem) so that $z = ye^{-y}$. The result is an alternative form of Eq. (2.4).

Proposition 2. *Quantity $D(n) = R(n) - Q(n)$ satisfies*

(2.5)

$$D(n) = \frac{n!}{n^{n-1}} \frac{1}{2\pi i} \oint_{\mathcal{C}} \log \frac{(1-y)^2}{2(1-ye^{1-y})}\, d\omega \quad \textit{with } d\omega = \frac{dz}{z^{n+1}} = (1-y)e^{ny}\frac{dy}{y^{n+1}},$$

where $\mathcal{C}$ is a small contour around the origin in either the z–plane (the first form of $d\omega$ is used with y implicitly defined by $y = ze^y$) or in the y–plane (the second form is used with y being the independent variable).

In order to make use of (2.5), we first observe that, from (2.2), the coefficients of $(y-1)^k$ form an asymptotic scale:

$$\begin{gathered} \frac{n!}{n^{n-1}}[z^n](y-1) = 1, \quad \frac{n!}{n^{n-1}}[z^n](y-1)^2 = -\frac{2}{n}, \\ \frac{n!}{n^{n-1}}[z^n](y-1)^3 = -\frac{3}{n} + \frac{6}{n^2}, \quad \frac{n!}{n^{n-1}}[z^n](y-1)^4 = \frac{20}{n^2} - \frac{24}{n^3}, \\ \frac{n!}{n^{n-1}}[z^n](y-1)^5 = \frac{15}{n^2} - \frac{130}{n^3} + \frac{120}{n^4}, \quad \ldots \end{gathered} \tag{2.6}$$

with $n!n^{-n-1}[z^n](y-1)^k$ being $O(n^{-\lfloor k/2 \rfloor})$ and comprising $\lceil k/2 \rceil$ terms.

If we expand the integrand of (2.5) in powers of $(y-1)$ and use the *first* form of $d\omega$, $d\omega = dz/z^{n+1}$, then through (2.6) we get an asymptotic expansion for $D(n)$:

$$D(n) \sim \sum_{k \geq 1} c_k \cdot [z^n](y-1)^k \tag{2.7}$$

where ($\delta = y - 1$):

$$\begin{aligned} &\log \frac{\delta^2}{2(1-(1+\delta)e^{-\delta})} = \sum_{k\geq 1} c_k \delta^k \\ &= \frac{2}{3}\delta - \frac{1}{36}\delta^2 - \frac{1}{810}\delta^3 + \frac{1}{12960}\delta^4 \\ &\quad + \frac{1}{68040}\delta^5 + \frac{1}{12247200}\delta^6 - \frac{1}{6123600}\delta^7 - \frac{13}{1175731200}\delta^8 + \cdots . \end{aligned} \tag{2.8}$$

By (2.6), Eq. (2.7) itself transforms into a standard expansion,

$$D(n) \sim \sum_{k\geq 0} \frac{d_k}{n^k}, \tag{2.9}$$

whose first few terms are

$$D(n) \sim \frac{2}{3} + \frac{8}{135n} - \frac{16}{2835n^2} - \frac{32}{8505n^3} + \frac{1794}{12629925n^4} + \cdots .$$

Analytically, the complete process leading to (2.9) is again justified by the singularity analysis theorems. It is quite parallel to the proof of Theorem 1 (only $y - 1$ is used instead of the asymptotically related ε) and refinements of it are going to provide the required bounds.

3. Expansions with Effective Error Bounds

The main device employed here follows the outline given in the equations (2.7)–(2.9). We use a *terminating* form of (2.8) inside our main integral representation,

$$\frac{n^{n-1}}{n!}D(n) = \frac{1}{2\pi i}\oint_{\mathcal{C}}\left(\sum_{k=0}^{K} c_k(y-1)^k\right)d\omega + \frac{1}{2\pi i}\oint_{\mathcal{C}} R_K(y-1)\,d\omega, \tag{3.1}$$

where the c_k are defined by (2.8) and

$$R_K(y-1) = \log\frac{(1-y)^2}{2(1-ye^{1-y})} - \sum_{k=1}^{K} c_k(y-1)^k = \sum_{k\geq K+1} c_k(y-1)^k. \tag{3.2}$$

The integral with the finite sum in (3.1) is known *exactly* from Section 2 and Eq. (2.6); our aim is to derive constructive bounds for the integral containing the remainder.

To estimate the remainder integral, we make use the *second* form of $d\omega$, namely $d\omega = (1-y)e^{ny}dy/y^{n+1}$ together with a contour *in the y–plane* whose choice is dictated by a *saddle point heuristic.* The quantity $e^{ny}y^{-n}$ has a saddle point at $y=1$ with axis perpendicular to the real line; accordingly, we take $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$ where

$$\mathcal{C}_1 = \{1 + i\tau \;/\; -1 \leq \tau \leq 1\} \quad\text{and}\quad \mathcal{C}_2 = \{y \;/\; |y| = \sqrt{2},\ \Re(y) \leq 1\}. \tag{3.3}$$

Our proof, that will eventually fix $K = 10$, consists of two simple phases.

(a) The integral of the remainder term along $\mathcal{C}_2$ is small since there $y^{-n} = O(2^{-n/2})$; in addition, it can be effectively bounded.
(b) Along $\mathcal{C}_1$, quantity R_K is small since it was obtained by subtracting from a function the first K terms of its locally convergent Taylor expansion; in addition, constructive bounds can be derived.

We thus proceed with this programme, starting with estimates of the c_k coefficients, continuing with the estimates along $\mathcal{C}_2$ then $\mathcal{C}_1$, which we give for a general K. To take care of recurring factors of the form $n!/n^n$, we appeal to a weak form of Stirling's formula valid for all $n \geq 1$,

$$n! < \frac{11}{10}n^n e^{-n}\sqrt{2\pi n}. \tag{3.4}$$

Lemma 1. *For all $k \geq 1$, we have*

$$|c_k| < \frac{10.96714833}{\pi^k}. \tag{3.5}$$

Proof. ¿From Eq. (2.8) We estimate the coefficients of $\log f(z)$ where

$$f(z) = \frac{z^2}{2(1-(z+1)e^{-z})}$$

Figure 1. The transform by $f(z)$ of the square of side 2π centered at the origin together with a blow up of the picture near the origin (lower right).

---

by means of Cauchy's formula:

$$c_k = \frac{1}{2\pi i} \oint_{\mathcal{Q}} \log f(z) \frac{dz}{z^{k+1}}, \tag{3.6}$$

where $\mathcal{Q}$ is any small enough simple contour encircling the origin. We propose to take here as contour $\mathcal{Q}$ the boundary of the square $|\Re z| \leq \pi$, $|\Im z| \leq \pi$.

First, elementary computations show that there are no zeros nor poles of $f(z)$ within $\mathcal{Q}$. The graph of Figure 1 which represents the image of $\mathcal{Q}$ by $f(z)$ confirms this via the principle of the argument since $f(\mathcal{Q})$ has winding number 0 with respect to the origin. Thus $\log f(z)$ is analytic in and on $\mathcal{Q}$ and the Cauchy integral (3.6) for c_k can be evaluated by taking $\mathcal{Q}$ as the integration contour.

By considering the four sides of the square $|\Re z| \leq \pi$, $|\Im z| \leq \pi$ and using trivial estimates we obtain:

$$\frac{\pi^2}{2(1 + (2\pi + 1)e^{\pi})} \leq |f(z)| \leq \frac{\pi^2}{1 - (\pi + 1)e^{-\pi}}.$$

We next have to consider the argument of $f(z)$. For this purpose we investigate the argument of the denominator of $f(z)$. We restrict our investigation to the part of the square with $\Im z > 0$. By examining the behaviour of the sign of the real and imaginary part of $g(z) = 1 - (z+1)e^{-z}$ on the three lines $\Re z = \pi, 0 \leq \Im z \leq \pi$, $-\pi \leq \Re z \leq \pi, \Im z = \pi$ and $\Re z = -\pi, 0 \leq \Im z \leq \pi$, we find that $g(z)$ does not enter the third quadrant. Thus we have $-\frac{\pi}{2} < \arg g(z) < \pi$. On the other hand we have $0 \leq \arg z^2 \leq 2\pi$. Combining the two bounds yields

$$|\arg f(z)| \leq \frac{5}{2}\pi.$$

Thus we obtain

$$|\log f(z)| \leq \left(\log^2 \frac{\pi^2}{2((2\pi+1)e^{\pi}+1)} + \frac{25}{4}\pi^2\right)^{\frac{1}{2}} = 8.613578\ldots.$$

Bounding the integral (3.6) trivially, we arrive at

$$|c_k| \leq \frac{1}{2\pi}\left(\frac{8.613578}{\pi^{k+1}}\right)(8\pi),$$

which is equivalent to (3.5). □

Numerical computations suggest that the c_k decrease roughly like 7^{-k}, so that the estimate of Lemma 1 is not too tight. It is however amply sufficient for our purposes. With it, we next bound the remainder integral along $\mathcal{C}_2$.

Lemma 2. *We have*

$$\left|\frac{n!}{n^{n-1}}\frac{1}{2\pi i}\int_{\mathcal{C}_2} R_K(y-1)d\omega\right| \leq H(K)n^{\frac{3}{2}}2^{-\frac{n}{2}}, \tag{3.7}$$

where

$$H(K) = 181.7306404\left(\frac{1+\sqrt{2}}{\pi}\right)^{K}.$$

Proof. Estimate (3.2) by the triangle inequality applying the bounds (3.5) of Lemma 1 for the coefficients c_k; use Stirling's formula (3.4) to eliminate the factorials. This gives the upper bound

$$\left(\frac{11}{10}\frac{ne^{-n}\sqrt{2\pi n}}{2\pi}\right)\left(M\frac{\frac{1+\sqrt{2}}{\pi}}{1-\frac{1+\sqrt{2}}{\pi}}\right)\left(\frac{1+\sqrt{2}}{\pi}\right)^{k}\left(\frac{(1+\sqrt{2})e^{n}}{\sqrt{2}}2^{-n/2}\right)\left(2\pi\sqrt{2}\frac{3}{4}\right),$$

with $M = 10.96714833$ being the constant appearing in Lemma 1. All reductions done, this provides the stated bound. □

It remains to evaluate the remainder integrals along $\mathcal{C}_1$.

Lemma 3. *For all $n \geq 1$, we have the bound*

$$\left|\frac{n!}{n^{n-1}}\frac{1}{2\pi i}\oint_{\mathcal{C}_1} R_K(y-1)\,d\omega\right| \leq G(K)\frac{1}{n^{K/2}} \tag{3.8}$$

where

$$G(K) = \frac{11}{10}\frac{2^{K+3}}{\sqrt{2\pi}}\mu_K\Gamma\left(\frac{K+3}{2}\right) \qquad \text{and} \qquad \mu_K = \max_{y\in\mathcal{C}_1}\left|\frac{R_K(y-1)}{(y-1)^K}\right| \tag{3.9}$$

Proof. Setting $y = 1+i\tau$, and bounding the integral in (3.8), we find

$$\left|\frac{n!}{n^{n-1}}\frac{1}{2\pi i}\int_{\mathcal{C}_1} R_K(y-1)\,d\omega\right| \leq \frac{11}{10}\frac{n\sqrt{2\pi n}}{2\pi}\mu_K J(K), \tag{3.10}$$

with μ_K defined in (3.9) and

$$\begin{aligned} J(K) &= \int_{-1}^{+1} |\tau|^{K+2} \frac{d\tau}{|1+i\tau|^{n+1}} \\ &< 2\int_0^1 \tau^{K+2} \frac{d\tau}{(1+\tau^2)^{n/2}}. \end{aligned} \tag{3.11}$$

Since, in the given range, we have $1 + u \geq e^{u/2}$, we find

$$J(K) < 2\int_0^{\infty} \tau^{K+2} e^{-n\tau^2/4}\, d\tau$$

and a change of variable shows that the right hand side coincides with a gamma function, namely

$$J(K) < 2^{K+3}\Gamma\left(\frac{K+3}{2}\right) n^{-(K+3)/2}. \tag{3.12}$$

Using (3.12) inside (3.10) yields the statement of the lemma. □

4. Ultimate inequalities and the main theorem

We are finally in a position to combine the effective bounds provided by Lemmas 1, 2 and 3 and derive the main result of the paper.

Theorem 2. *With the quantity $\theta \equiv \theta(n)$ being defined by*

$$\frac{1}{2}e^n = 1 + \frac{n}{1!} + \frac{n^2}{2!} + \cdots + \frac{n^n}{n!}\theta,$$

one has, for all integers $n \geq 0$,

$$\theta = \frac{1}{3} + \frac{4}{135(n+k)},$$

where $k \equiv k(n)$ lies between $\frac{8}{45}$ and $\frac{2}{21}$.

Proof. The proof uses the decomposition (3.1) together with the estimates derived in the last section. We fix the value $K = 10$. The remainder integral containing R_K is estimated along $\mathcal{C}_2$ and $\mathcal{C}_1$. Then we consider the main terms.

(i). The remainder integral along $\mathcal{C}_2$ is estimated by Lemma 2. For $K = 10$, we get $H(10) = 13.05227701$ so that we find

$$\begin{aligned} \left| \frac{n!}{n^{n-1}} \frac{1}{2\pi i} \int_{\mathcal{C}_2} R_{10}(y-1)\, d\omega \right| &< 13.06 n^{3/2} 2^{-n/2} \\ &< 10^{-7}\frac{1}{n^3} \qquad \text{for } n \geq n_0 = 116. \end{aligned} \tag{4.1}$$

(Numerical studies show that much better could be done, but the effect on n_0 is marginal.)

(ii). The remainder integral along $\mathcal{C}_1$ is estimated by Lemma 3. This requires estimating the quantity μ_K appearing in the bound (3.8). The smallness of the μ_K's is naturally related to the exponential decay of the coefficients c_K since $\mu_K \approx c_K$.

¿From Lemma 1, we find with $M = 10.96714833$,

$$\mu_K \leq M \sum_{j>K} \frac{1}{\pi^j} = \frac{M}{\pi - 1}\left(\frac{1}{\pi}\right)^K,$$

and in particular, for $K = 10$:

$$\mu_{10} < 0.00005468 \qquad \text{so that} \qquad G(10) = 56.59398.$$

(From numerical experiments, we expect a slightly better bound to hold: $\mu_{10} < 10^{-7}$—but once more the effect is marginal.)

(iii). We finally consider the first 10 $(= K)$ terms in the expansion (3.1) which contribute a polynomial of degree 9 in $\frac{1}{n}$ to $D(n)$. Define

$$\begin{aligned} D_{10}(n) &= \frac{n!}{n^{n-1}} \frac{1}{2\pi i} \oint_{\mathcal{C}} \sum_{k=0}^{10} c_k (y-1)^k \, d\omega \\ &= \frac{2}{3} + \frac{8}{135\,n} - \frac{16}{2835\,n^2} - \frac{32}{8505\,n^3} + \frac{17984}{12629925\,n^4} + \frac{13159709}{9699782400\,n^5} \\ &\quad - \frac{977069}{1039262400\,n^6} - \frac{36669961}{28291032000\,n^7} + \frac{117191}{56582064\,n^8} - \frac{479}{561330\,n^9}. \end{aligned}$$

Equation (3.1) and the bounds found for $K = 10$ yield

$$D_{10}(n) - \Delta(n) \leq D(n) \leq D_{10}(n) + \Delta(n) \tag{4.2}$$

where

$$\Delta(n) = \frac{10^{-7}}{n^3} + \frac{57}{n^5},$$

under the sole condition that $n \geq n_0 = 116$.

The two basic inequalities

$$\begin{aligned} D_{10}(n) - \Delta(n) &\geq \frac{2}{3} + \frac{8}{135(n + \frac{8}{45})} \qquad \text{for } n \geq n_1 = 24 \\ D_{10}(n) + \Delta(n) &\leq \frac{2}{3} + \frac{8}{135(n + \frac{2}{21})} \qquad \text{for } n \geq n_2 = 116, \end{aligned}$$

are verified by normalizing the involved rational fractions and studying the variations of the numerator polynomials that are of degrees 8 and 7 respectively. Thus, from (4.2), the main assertion of the theorem is established for

$$n \geq \max(n_0, n_1, n_2) = 116.$$

To finish the proof, it suffices to calculate $k(n)$ for $n \leq 115$ by computer and to verify the inequality in these cases (see Fig. 2 for a display). ☐

Figure 2. A plot of the first 120 values of $k(n)$ confirms that they all lie inside the interval $[\frac{2}{21}, \frac{8}{45}]$.

Proof and Computation. The symbolic manipulations needed by our proof were performed with the help of the computer algebra system Maple. Computer algebra intervened at most stages in the construction of the proof: in performing various expansions like (2.6),(2.8),(2.9), and in obtaining the various bounds of Sections 3 and 4. As soon as enough terms have been gathered in the asymptotic expansion of $Q(n)$, the truth of the conjecture in its weaker asymptotic form, for "large enough" n, is immediate. The difficulty in obtaining a complete proof of Theorem 2 is then twofold: (i). the proof of asymptotic estimates should be suitably transformed in order to yield inequalities valid for large enough n; (ii). the range of excluded values of n should be small enough that the remaining cases, here $n < 116$, be amenable to exhaustive verification. The second of these conditions has necessitated a rather delicate tuning process that could be done rather painlessly with the help of a few hours of interaction with our computer algebra system. A hand carved proof along our lines, though perhaps conceivable, would have involved a rather formidable calculational effort.

5. Some Conclusions

It may be of interest, at last, to reflect on the various alternatives that offer themselves in order to estimate asymptotically sequences like $Q(n)$ or $D(n)$.

1. *Laplace Method.* The Laplace method for integrals, based on the integral representation (1.5) was the starting point of earlier approaches. As Szegö and Watson show, it can be made "constructive" (instead of providing only O-bounds) but its operation becomes then somewhat intricate.

2. *Singularity Analysis.* This is the method that gave us here the expansion of Theorem 1. It is based on the fact that the implicitly defined function $y(z)$ has an algebraic singularity of the $\sqrt{\ }$–type, from which the singularity types of the generating functions associated with $Q(n)$ or $D(n)$ follow. The method can also accommodate constructive bounds on a function's coefficients [FO90b]. Consequently it might be applicable to derive Theorem 2, although, in this case, bounding $y(z)$ in the appropriate region would probably prove unwieldy.

3. *Darboux's Method.* Darboux's method also leads to a full asymptotic expansion by a route very similar to singularity analysis. However it does not have the capacity to provide bounds since it is intrinsically based on a non–constructive lemma on Fourier series.

4. *Saddle Point.* This is in essence the route that we took, after a suitable change of variable. Its application in the case of implicitly defined functions and Lagrange series is also to be traced in Darboux's works, an interesting combinatorial application occurring in [MM78]. By this method, we were able to reduce the problem to the task of finding simple bounds for elementary functions on circles and line segments. Interestingly enough, when considering the conformal mapping defined by $y^{(-1)}(z)$, it appears that the induced contour in the z–plane closely resembles the type of "Hankel" contour used in the z–plane under the singularity analysis approach. This establishes a perhaps unexpected relation between two seemingly unrelated methods—singularity analysis and saddle point—at least in the context of implicitly defined functions and Lagrange series.

References

[Be89] B. Berndt, *Ramanujan's Notebooks, Part II*, Springer, Berlin, 1989.

[CR92] K. Compton and C. Ravishankar, *Mean Deadlock Time in a Multiprocessing System*, Preprint (1992).

[DB81] N. G. de Bruijn, *Asymptotic Methods in Analysis*, Dover, 1981.

[FGT92] P. Flajolet, D. Gardy and L. Thimonier, *Birthday paradox, coupon collectors, caching algorithms, and self–organizing search*, Discrete Applied Mathematics **42** (1992), In press.

[FO90a] P. Flajolet and A.M. Odlyzko, *Random Mapping Statistics*, Lect. Notes in Comp. Sc. **434** (1990), 329–354. In *Advances in Cryptology*, Proceedings of Eurocrypt'89, J-J. Quisquater and J. Vandewalle Editors..

[FO90b] P. Flajolet and A.M. Odlyzko, *Singularity Analysis of Generating Functions*, SIAM J. Disc. Math. **3** (1990), 216–240.

[GJ83] I.P. Goulden and D.M. Jackson, *Combinatorial Enumeration*, J. Wiley, New York, 1983.

[Ha78] G. H. Hardy, *Ramanujan: Twelve Lectures on Subjects Suggested by his Life and Work*, Chelsea Publishing Company, New York, 1978.

[He77] P. Henrici, *Applied and Computational Complex Analysis*, J. Wiley, New York, 1977.

[Kn68] D.E. Knuth, *The Art of Computer Programming*, vol. 1, Addison Wesley, Reading, 1968.

[Kn69] D.E. Knuth, *The Art of Computer Programming*, vol. 2, Addison Wesley, Reading, 1969.

[Kn73] D.E. Knuth, *The Art of Computer Programming*, vol. 3, Addison Wesley, Reading, 1973.

[Kn85] D.E. Knuth, *An analysis of optimal caching*, J. of Algorithms (1985), 181–199.

[KP89] D. E. Knuth and B. Pittel, *A recurrence related to trees*, Proc. Amer. Math. Soc. **105** (1989), 335–349.

[KR75] D. E. Knuth and G. S. Rao, *Activity in an Interleaved Memory*, IEEE Transactions on Computers **C-24** (1975), 943–944.

[KS78] D. E. Knuth and A. Schönhage, *The expected linearity of a simple equivalence algorithm*, Theoret. Comp. Science **6** (1978), 281–315.

[MM78] A. Meir and J.W. Moon, *On the altitude of nodes in random trees*, Canadian Journal of Mathematics **30** (1978), 997–1015.

[Ra11] S. Ramanujan, *Question 294*, J. Indian Math. Soc. **3** (1911), 128.

[Ra12] S. Ramanujan, *On question 294*, J. Indian Math. Soc. **4** (1912), 151–152.

[Ra62] S. Ramanujan, *Collected Papers*, Chelsea Publishing Company, New York, 1962.

[Sz28] G. Szegő, *Über einige von S. Ramanujan gestellte Aufgaben*, J. London Math. Soc. **3** (1928), 225–232.

[VF90] J. S. Vitter and P. Flajolet, *Analysis of Algorithms and Data Structures*, In Handbook of Theoretical Computer Science. Vol A: Algorithms and Complexity (J. van Leeuwen Ed.), Chapter 9 (1990), 431–524.

[Wa29] G.N. Watson, *Theorems Stated by Ramanujan (V): Approximations Connected with* e^x, Proc. London Math. Soc. (2) **29** (1929), 293–308.

P. Flajolet
Algorithms Project
INRIA, Rocquencourt
78153 Le Chesnay (France)

P.J. Grabner
Technische Universität Graz
Steyrergasse 30
A–8010 Graz (Austria)

P. Kirschenhofer & H. Prodinger
Technische Universität Wien
Wiedner Hauptstrasse 8–10
A–1040 Vienna (Austria)