LIMIT DISTRIBUTIONS OF SMALLEST GAP AND LARGEST REPEATED PART IN INTEGER PARTITIONS

STEPHAN WAGNER

ABSTRACT. We study two parameters in random integer partitions, namely the first gap and the last repeated part, that have been introduced by Grabner and Knopfmacher in a recent paper [Ramanujan J. 12/3, 439–454]. More generally, the first part that occurs at most r times and the last part that occurs at least r times are considered. For both parameters, we determine the limit distribution, which turn out to be the Rayleigh and Gumbel distribution respectively. This also generalises the well-known result by Erdős and Lehner on the distribution of the largest part in a random integer partition. Furthermore, extensions to general Λ-partitions as well as results on related parameters such as the length of the first gap are provided.

1. INTRODUCTION

The theory of partitions lies at the border between number theory and combinatorics and combines algebraic and analytic aspects. The classic asymptotic formula for the partition function $p(n)$ that was found by Hardy and Ramanujan [11] and further improved by Rademacher [16] arguably belongs to the most beautiful theorems in 20th century mathematics. Its main asymptotic term can be written as

$$p(n) = \frac{\exp\left(\pi\sqrt{\frac{2n}{3}}\right)}{4\sqrt{3}n}\left(1+O(n^{-1/2})\right). \tag{1}$$

See [1] and [2] for excellent expositions. This result has been generalised in various directions, most notably by Meinardus [14]: consider a nondecreasing sequence $\Lambda_1 \le \Lambda_2 \le \dots$ of positive integers with the property that $\Lambda_k \to \infty$. A (unrestricted) Λ-partition can be seen as a solution to the equation

$$\sum_{k=1}^{\infty} a_k\Lambda_k = n,$$

where the coefficients a_k have to be nonnegative integers; a_k is the multiplicity of the part Λ_k in a partition. Note that the case $\Lambda_k = k$ corresponds to ordinary partitions. Assume that the following conditions are satisfied:

2000 *Mathematics Subject Classification.* 05A17; 11P82, 60C05.

Key words and phrases. partition, smallest gap, largest repeated part, distribution, local limit theorem.

This material is based upon work supported financially by the National Research Foundation of South Africa under grant number 70560.

(1) The associated Dirichlet series $D(s) = \sum_{k=1}^{\infty} \Lambda_k^{-s}$ has a simple pole at $\alpha > 0$ with residue A and can otherwise be continued analytically to the strip $\sigma \geq -C_0$ ($C_0 > 0$); here we write $s = \sigma + it$ as usual.
(2) $D(s) = O(|t|^{C_1})$ for some fixed positive constant C_1, uniformly in the region $\sigma \geq -C_0$, as $t \to \infty$.
(3) The function

$$v(\tau) = \sum_{k=1}^{\infty} e^{-\Lambda_k \tau}$$

satisfies

$$\operatorname{Re} v(\tau) - v(y) \leq -C_2 y^{-\eta}$$

for $\tau = y + 2\pi i x$, $|\arg \tau| > \frac{\pi}{4}$, $|x| \leq \frac{1}{2}$ and sufficiently small y, where η and C_2 are positive constants.

Then it can be shown that the number of Λ-partitions is asymptotically given by

$$p_\Lambda(n) = Cn^\kappa \exp\left(n^{\alpha/(\alpha+1)}\left(1 + \frac{1}{\alpha}\right)(A\Gamma(\alpha+1)\zeta(\alpha+1))^{1/(\alpha+1)}\right)\left(1 + O(n^{-\kappa_1})\right), \tag{2}$$

where Γ and ζ denote the Gamma and zeta functions respectively, and C, κ, κ_1 are constants that depend on α and D; specifically,

$$C = e^{D'(0)}\left(2\pi(1+\alpha)\right)^{-1/2}(A\Gamma(\alpha+1)\zeta(\alpha+1))^{(1-2D(0))/(2\alpha+2)},$$
$$\kappa = \frac{D(0) - 1 - \frac{\alpha}{2}}{1+\alpha},$$
$$\kappa_1 = \frac{\alpha}{\alpha+1}\min\left(\frac{C_0}{\alpha} - \frac{\delta}{4}, \frac{1}{2} - \delta\right).$$

See [1] for details. A similar result exists for *restricted* Λ-partitions ($a_k \leq 1$, i.e., parts may not be repeated). In this case, the number of partitions, denoted $q_\Lambda(n)$, satisfies

$$q_\Lambda(n) = Bn^\lambda \exp\left(n^{\alpha/(\alpha+1)}\left(1 + \frac{1}{\alpha}\right)\left(A\Gamma(\alpha+1)\zeta(\alpha+1)(1-2^{-\alpha})\right)^{1/(\alpha+1)}\right) \left(1 + O(n^{-\kappa_1})\right), \tag{3}$$

where

$$B = 2^{D(0)}\left(2\pi(1+\alpha)\right)^{-1/2}\left(A\Gamma(\alpha+1)\zeta(\alpha+1)(1-2^{-\alpha})\right)^{1/(2\alpha+2)},$$
$$\lambda = -\frac{1+\frac{\alpha}{2}}{1+\alpha}.$$

An even stronger theorem holds for the bivariate generating function in which the second variable marks the length, see [12]. In particular, for $\Lambda_k = k$, the following asymptotic formula for the number of partitions into unequal parts results:

$$q(n) = \frac{\exp\left(\pi\sqrt{\frac{n}{3}}\right)}{4\sqrt[4]{3}n^{3/4}}\left(1 + O(n^{-1/2})\right).$$

Once that information about the number of partitions is available, it is natural to consider parameters of partitions. The first important example of a distributional result is due to Erdős and Lehner [5], who considered the length (number of parts) of a random partition. Since the length of a partition is the largest part of its conjugate, the maximum follows the same distribution law, which is an extreme value distribution (also known as Gumbel distribution) after appropriate rescaling. This will also be a corollary of one of the main results of this paper. Szekeres [17, 18] refined the result of Erdős and Lehner and also studied the joint distribution of length and maximum [20] (it turns out that, around their mean, the two are essentially independent of each other). The analogous problem for partitions into unequal parts is considered in [19].

Further interesting contributions to the distributional theory of integer partitions include those by Erdős and Szalay [6] on the largest block size (the size of the block induced by a number j in a partition is j times the multiplicity of j). This was further generalised by Fristedt [8]; Corteel et al. [4] showed that a randomly selected part of a randomly selected partition of n has multiplicity m with probability $\frac{1}{m(m+1)}$ as $n \to \infty$.

The number of distinct parts in a random partition (the number of parts with positive multiplicity) was shown to follow a normal law by Goh and Schmutz [9]; the Gaussian limit law also arises in the study of ascents in partitions [3]. The longest run (or largest ascent), on the other hand, was found to have a rather unusual limit distribution by Mutafchiev [15]; see also [21].

In an interesting recent paper by Grabner and Knopfmacher [10], several new partition parameters were brought forward, among them the smallest gap (the first positive integer that does not occur as a part) and the largest repeated part. The former was shown to be on average of order $n^{1/4}$ if random integer partitions of n are considered, while the latter is of average order $\sqrt{n}\log n$. Very precise asymptotic expansions for the mean can be found in [10].

The present paper is devoted to the limit distributions of the aforementioned partition statistics; indeed, we consider two slightly more general parameters. For a positive integer r (which is fixed throughout the paper), we study the distribution of the smallest integer whose multiplicity is less than r as well as that of the largest integer whose multiplicity is at least r; in particular, for $r = 1$, we obtain the smallest gap and the maximum respectively; $r = 2$ corresponds to the largest repeated part size in the latter case. Note that the first parameter exists for any partition, while the latter does not necessarily exist (it is possible that there is no repeated part). However, it turns out that a part that is repeated at least r times exists with probability tending to 1 as $n \to \infty$ in a random partition of n.

We find that both parameters have a limit distribution for any fixed r that is a Rayleigh distribution in the first and a Gumbel distribution in the second case. The two main theorems read as follows:

Theorem 1. *Let $G_{r,n}$ be the smallest integer in a random partition of n that occurs less than r times. Then the normalised random variable $n^{-1/4}G_{r,n}$ tends to a Rayleigh distribution with mean $\left(\frac{3}{2}\right)^{1/4} r^{-1/2}$ and variance $\frac{\sqrt{6}(4-\pi)}{2\pi r}$. The local limit theorem*

$$\mathbb{P}(G_{r,n} = hn^{1/4}) \sim \frac{\pi r h}{\sqrt{6}n^{1/4}} \exp\left(-\frac{\pi r h^2}{2\sqrt{6}}\right)$$

holds.

Theorem 2. *Let $R_{r,n}$ be the largest integer in a random partition of n that occurs at least r times (if such an integer exists; otherwise, define the value to be 0). Then the normalised random variable $n^{-1/2}\left(R_{r,n} - \frac{\sqrt{6n}}{2\pi r}\log n\right)$ tends to a Gumbel distribution with mean $\frac{\gamma - \log(\pi r/\sqrt{6})}{\pi r/\sqrt{6}}$ and variance $\frac{1}{r^2}$, where γ is the Euler-Mascheroni constant. The local limit theorem*

$$\mathbb{P}\left(R_{r,n} = \frac{\sqrt{6n}}{2\pi r}\log n + h\sqrt{n}\right) \sim \frac{1}{\sqrt{n}} \exp\left(-\frac{\pi r h}{\sqrt{6}} - \frac{\sqrt{6}}{\pi r}\exp\left(-\frac{\pi r h}{\sqrt{6}}\right)\right)$$

holds.

These two theorems are treated in Section 2 and 3 respectively; possible extensions and generalisations are discussed as well as related problems.

2. Smallest gaps

In order to prove Theorem 1, we first need the generating function for the number of those partitions for which k is the first integer that occurs with multiplicity less than r. It is easy to see that this generating function is given by

$$\begin{aligned}\prod_{j=1}^{k-1}\left(\sum_{m=r}^{\infty} z^{jm}\right) \cdot \left(\sum_{m=1}^{r-1} z^{mk}\right) \cdot \prod_{j=k+1}^{\infty}\left(\sum_{m=1}^{\infty} z^{jm}\right) &= \left(\prod_{j=1}^{k-1} z^{jr}\right)(1 - z^{kr})\prod_{j=1}^{\infty}(1 - z^j)^{-1} \\ &= z^{k(k-1)r/2}(1 - z^{kr})P(z),\end{aligned}$$

where $P(z)$ is the generating function for all partitions. Hence, if $p(n)$ denotes the number of partitions of n, then the coefficient of z^n in the above generating function (and thus the number of partitions of n for which k is the first integer that occurs less than r times) is

$$[z^n]z^{k(k-1)r/2}(1 - z^{kr})P(z) = p(n - k(k-1)r/2) - p(n - k(k+1)r/2). \tag{4}$$

Now we make use of the Hardy-Ramanujan-Rademacher formula for the partition function: applying (1), we find for $k = o(n^{3/8})$ that

$$\begin{aligned}
p(n-k(k-1)r/2) &= \frac{\exp\left(\pi\sqrt{\frac{2(n-k(k-1)r/2)}{3}}\right)}{4\sqrt{3}(n-k(k-1)r/2)}\left(1+O(n^{-1/2})\right) \\
&= \frac{\exp\left(\pi\sqrt{\frac{2n}{3}}\right)}{4\sqrt{3}n}\exp\left(\pi\sqrt{\frac{2n}{3}}\left(\sqrt{1-\frac{k(k-1)r}{2n}}-1\right)\right) \\
&\quad \left(1+O(k^2n^{-1}+n^{-1/2})\right) \\
&= p(n)\exp\left(\pi\sqrt{\frac{2n}{3}}\left(-\frac{k(k-1)r}{4n}+O(k^4n^{-2})\right)\right) \\
&\quad \left(1+O(k^2n^{-1}+n^{-1/2})\right) \\
&= p(n)\exp\left(-\frac{\pi k(k-1)r}{2\sqrt{6n}}\right)\left(1+O(k^4n^{-3/2}+k^2n^{-1}+n^{-1/2})\right).
\end{aligned}$$

Likewise,

$$p(n-k(k+1)r/2) = p(n)\exp\left(-\frac{\pi k(k+1)r}{2\sqrt{6n}}\right)\left(1+O(k^4n^{-3/2}+k^2n^{-1}+n^{-1/2})\right)$$

and thus

$$\begin{aligned}
&\frac{p(n-k(k-1)r/2)-p(n-k(k+1)r/2)}{p(n)} \\
&\quad = \exp\left(-\frac{\pi k^2 r}{2\sqrt{6n}}\right)\left(\exp\left(\frac{\pi k r}{2\sqrt{6n}}\right)-\exp\left(-\frac{\pi k r}{2\sqrt{6n}}\right)+O(k^4n^{-3/2}+k^2n^{-1}+n^{-1/2})\right) \\
&\quad = \frac{\pi k r}{\sqrt{6n}}\exp\left(-\frac{\pi k^2 r}{2\sqrt{6n}}\right)\left(1+O(k^3n^{-1}+kn^{-1/2}+k^{-1})\right).
\end{aligned}$$

In particular, if we set $k = hn^{1/4}$, we find

$$\frac{p(n-k(k-1)r/2)-p(n-k(k+1)r/2)}{p(n)} = \frac{\pi r h}{\sqrt{6}n^{1/4}}\exp\left(-\frac{\pi r h^2}{2\sqrt{6}}\right)\left(1+O(n^{-1/4})\right),$$

uniformly in h on compact subsets of $(0,\infty)$. This readily proves our first main theorem. ∎

Remark. The probability decreases rapidly for larger values of k: one has

$$\mathbb{P}(G_{r,n} \geq k) = \frac{p(n-k(k-1)r/2)}{p(n)} = O\left(\exp\left(-\frac{\pi k^2 r}{2\sqrt{6n}}\right)\right).$$

Theorem 1 generalises to *Meinardus'* scheme that was mentioned in the introduction: however, the error term in (2) is not sufficient to apply the same method for the proof.

Instead, we work directly with the generating function and apply the saddle point method to obtain the following result:

Theorem 3. *Let $G_{r,n}$ be the first element of the sequence Λ that occurs less than r times in a random Λ-partition of n; if the conditions of Meinardus' scheme are satisfied, then the normalised random variable $n^{-1/(\alpha+1)^2}G_{r,n}$ tends to a Weibull distribution whose density is given by*

$$f(x) = Krx^\alpha \exp\left(-\frac{Kr}{\alpha+1}x^{\alpha+1}\right)$$

for $x \geq 0$, where $K = A\left(A\Gamma(\alpha+1)\zeta(\alpha+1)\right)^{1/(\alpha+1)}$. The local limit theorem

$$\mathbb{P}(G_{r,n} = \Lambda_k) \sim \frac{Kr}{A}\Lambda_k n^{-1/(\alpha+1)} \exp\left(-\frac{Kr}{\alpha+1}\Lambda_k^{\alpha+1}n^{-1/(\alpha+1)}\right)$$

holds uniformly if Λ_k is restricted by $an^{1/(\alpha+1)^2} \leq \Lambda_k \leq bn^{1/(\alpha+1)^2}$, for any fixed a, b.

Of course, Theorem 1 is included as a corollary. However, it seemed sensible to present the somewhat easier proof for Theorem 1 before the general case is treated.

Remark. There is some ambiguity in the formulation of the local limit theorem, since some of the Λ_k may have the same value; in this case, think of several copies of the same number that are coloured differently.

Proof. Making use of the same argument as before, we find the generating function for the number of Λ-partitions with the property that Λ_k is the first element of the sequence Λ that occurs less than r times:

$$F(z) = \left(\prod_{j=1}^{k-1} z^{\Lambda_j r}\right)\left(1 - z^{\Lambda_k r}\right)\prod_{j=1}^{\infty}\left(1 - z^{\Lambda_j}\right)^{-1}.$$

Now set $z = e^{-\tau}$, $F(z) = f(\tau)$, and $\tau = y + 2\pi i x$ and apply the residue theorem to obtain

$$p_\Lambda(n) \cdot \mathbb{P}(G_{r,n} = \Lambda_k) = \int_{-1/2}^{1/2} f(y + 2\pi i x)e^{ny+2\pi i n x}\,dx \tag{5}$$

upon change of variables. Furthermore, take y to be the saddle point:

$$y = \left(A\Gamma(\alpha+1)\zeta(\alpha+1)\right)^{1/(\alpha+1)} n^{-1/(\alpha+1)} = \frac{K}{A}n^{-1/(\alpha+1)},$$

see Chapter 6 of [1], and set $\beta = 1 + \frac{\alpha}{2}\left(1 - \frac{\delta}{2}\right)$ for some $\delta > 0$. Only the part of the integral for which $|x| \leq y^\beta$ holds is asymptotically relevant; the rest is negligible and only yields an error term that is exponentially smaller than the main term. This is a consequence of the following lemma:

Lemma 4 ([1], Lemma 6.1)**.** *Let*

$$g(\tau) = G(z) = \prod_{j=1}^{\infty} \left(1 - z^{\Lambda_j}\right)^{-1}$$

be the ordinary generating function for Λ-partitions (where $z = e^{-\tau}$). Then there exist positive constants C_3, η_1 such that

$$g(y + 2\pi i x) = O\left(\exp\left(A\Gamma(\alpha)\zeta(\alpha+1)y^{-\alpha} - C_3 y^{-\eta_1}\right)\right)$$

holds uniformly in x with $y^\beta \le |x| \le \frac{1}{2}$ as $y \to 0$.

Estimating the additional factors in $f(y + 2\pi i x)$ in the trivial way, we now find that

$$\begin{aligned}\int_{y^\beta}^{\frac{1}{2}} f(y+2\pi i x)e^{ny+2\pi i n x}\,dx &= O\left(\exp\left(ny + A\Gamma(\alpha)\zeta(\alpha+1)y^{-\alpha} - C_3 y^{-\eta_1}\right)\right) \\ &= O\left(p_\Lambda(n)\exp\left(-C_4 n^{\eta_1/(\alpha+1)}\right)\right).\end{aligned}$$

Now assume that $\Lambda_k = O\left(n^{\min(1+\eta_1,\beta)/(\alpha+1)^2-\epsilon}\right)$ for some $\epsilon > 0$. By Ikehara's Tauberian theorem [13],

$$\sum_{j=1}^{k-1} \Lambda_j \sim \frac{A}{\alpha+1}\Lambda_k^{\alpha+1}.$$

It follows that for $|x| \le y^\beta$, one has

$$1 - \exp\left((y+2\pi i x)\Lambda_k r\right) = \Lambda_k r y\left(1 + O\left(\Lambda_k y + y^{\beta-1}\right)\right) \sim \Lambda_k r y$$

as well as

$$\exp\left(2\pi i \sum_{j=1}^{k-1} \Lambda_j r x\right) = 1 + O\left(n^{-\epsilon}\right),$$

both uniformly in x. Therefore, the right hand side in (5) can be written as

$$\left(\int_{-1/2}^{1/2} g(y+2\pi i x)e^{ny+2\pi i n x}\,dx\right) \cdot \Lambda_k r y \exp\left(-\sum_{j=1}^{k-1}\Lambda_j r y\right)(1+o(1)),$$

so that

$$p_\Lambda(n) \cdot \mathbb{P}(G_{r,n} = \Lambda_k) = p_\Lambda(n) \cdot \Lambda_k r y \cdot \exp\left(-\sum_{j=1}^{k-1}\Lambda_j r y\right)(1+o(1)).$$

Note that

$$\sum_{j=1}^{k-1}\Lambda_j r y \sim \frac{Ar}{\alpha+1}\Lambda_k^{\alpha+1} y = O\left(n^{\eta_1/(\alpha+1)-\epsilon(\alpha+1)}\right)$$

by our choice of Λ_k, which shows that the error term obtained from Lemma 4 is indeed smaller than the main term. Hence we end up with

$$\mathbb{P}(G_{r,n} = \Lambda_k) \sim \Lambda_k ry \cdot \exp\left(- \sum_{j=1}^{k-1} \Lambda_j ry \right).$$

Finally, if $an^{1/(\alpha+1)^2} \leq \Lambda_k \leq bn^{1/(\alpha+1)^2}$ for fixed a, b, then one has

$$\Lambda_k ry = \frac{Kr}{A} \Lambda_k n^{-1/(\alpha+1)}$$

and

$$\sum_{j=1}^{k-1} \Lambda_j ry = \frac{Ar}{\alpha+1} \Lambda_k^{\alpha+1} \cdot \frac{K}{A} n^{-1/(\alpha+1)} + o(1) = \frac{Kr}{\alpha+1} \Lambda_k^{\alpha+1} n^{-1/(\alpha+1)} + o(1),$$

which completes the proof of the theorem. ■

Example. If Λ is an arithmetic progression, then the limit law is still a Rayleigh distribution. If $\Lambda_k = k^m$ for some positive integer exponent m, then $\alpha = A = \frac{1}{m}$ (note that $D(s) = \zeta(ms)$ in this case), and one obtains a Weibull distribution with exponent $1 + \frac{1}{m}$. This can be further generalised to arbitrary exponents $m \geq 1$ by setting $\Lambda_k = \lfloor k^m \rfloor$.

Unlike the second parameter that we treat in this paper (the largest repeated part), the first gap also makes sense in the case of partitions into *unequal parts.* In this case, similar reasoning shows that the generating function for partitions into unequal parts whose first gap is k is given by

$$\prod_{j=1}^{k-1} z^j \cdot \prod_{j=k+1}^{\infty} (1 + z^j) = z^{k(k-1)/2} \prod_{j=k+1}^{\infty} (1 + z^j).$$

Interestingly, it turns out that the limit distribution is discrete in this case:

Theorem 5. *The first gap in a random partition of n into distinct summands asymptotically follows a geometric distribution: the probability that the first gap equals k tends to 2^{-k}.*

Proof. Note that the product

$$\prod_{j=k+1}^{\infty} (1 + z^j)$$

can also be interpreted as the generating function for the number of partitions into distinct parts $> k$. For our purposes, we need the coefficient of $z^{n-k(k-1)/2}$. The Dirichlet generating function $D(s) = \sum_{j=k+1}^{\infty} j^{-s} = \zeta(s) - \sum_{j=1}^{k} j^{-s}$ satisfies the conditions of Meinardus' scheme (generally, changing finitely many elements does not affect the validity), and the pole at 1 (with residue 1) is the same as for $\zeta(s)$, since the difference is only a finite sum of terms that represents an analytic function. Furthermore, $D(0) = -\frac{1}{2} - k$, so that (3) yields

$$\begin{aligned}
[z^n]z^{k(k-1)/2} \prod_{j=k+1}^{\infty} (1+z^j) &= [z^{n-k(k-1)/2}] \prod_{j=k+1}^{\infty} (1+z^j) \\
&= 2^{-k} \cdot \frac{\exp\left(\pi\sqrt{\frac{n-k(k-1)/2}{3}}\right)}{4\sqrt[4]{3}(n-k(k-1)/2)^{3/4}} \left(1+O(n^{-1/2})\right) \\
&= 2^{-k} \cdot \frac{\exp\left(\pi\sqrt{\frac{n}{3}}\right)}{4\sqrt[4]{3}n^{3/4}} \left(1+O(n^{-1/2})\right) \\
&= 2^{-k} q(n) \left(1+O(n^{-1/2})\right)
\end{aligned}$$

for any fixed k, which proves the theorem. ■

Remark. Intuitively, every fixed integer occurs with probability $\frac{1}{2}$ in a random partition of a large number into unequal parts; the geometric distribution of the first gap follows naturally.

This theorem generalises to Meinardus' scheme as well; the limit law is still geometric, and the proof is completely analogous:

Theorem 6. *If the conditions of Meinardus' scheme are satisfied, then the first gap in a restricted Λ-partition of n asymptotically follows a geometric distribution: the probability that the first gap equals Λ_k tends to 2^{-k}.*

Let us finally look at the length of the first gap: if k is the first number that is left out and $k+\ell$ is the first part larger than k that occurs in a certain partition, then we say that the length of the gap is ℓ (if such an ℓ exists). The following theorem provides information about the length of the first gap:

Theorem 7. *In a random partition of n, the length of the first gap equals 1 with probability $1+O(n^{-1/4})$. If partitions into unequal parts are considered, however, the length of the first gap asymptotically follows the same geometric distribution as the first gap itself, that is, the probability that the length of the first gap is ℓ tends to $2^{-\ell}$ as $n \to \infty$.*

The geometric distributions of the location of the first gap and its length are even asymptotically independent in the unrestricted case. Once again, the theorem can be generalised to Meinardus' scheme.

Theorem 8. *Suppose that the conditions of Meinardus' scheme are satisfied. The length of the first gap in a random unrestricted Λ-partition of n equals 1 with probability $1+O(n^{-1/(\alpha+1)^2+\epsilon})$ for any $\epsilon > 0$. For restricted Λ-partitions, the probability that the length of the first gap is ℓ tends to $2^{-\ell}$ as $n \to \infty$.*

Proof (Sketch). By an argument akin to the proof of Theorem 3, one finds that the probability for a gap of length at least two to occur at Λ_k (in other words, Λ_k is the first element of the sequence Λ that is not part of the partition, and Λ_{k+1} is not a part either) is $O\left(\Lambda_k^2 n^{-2/(\alpha+1)}\right)$. Since the probability that the first gap is greater than $n^{1/(\alpha+1)^2+\epsilon_1}$ is exponentially small for any fixed $\epsilon_1 > 0$ (by the same idea), it is sufficient to estimate the sum

$$n^{-2/(\alpha+1)} \sum_{\Lambda_k \le n^{1/(\alpha+1)^2+\epsilon_1}} \Lambda_k^2.$$

Applying Ikehara's Tauberian Theorem again, we see that this sum is

$$O\left(n^{-2/(\alpha+1)} \cdot n^{(1/(\alpha+1)^2+\epsilon_1)(2\alpha+1)}\right) = O\left(n^{-1/(\alpha+1)^2+\epsilon}\right)$$

for sufficiently small ϵ_1, which completes the proof of the first part.

For the second part, note that the generating function for restricted Λ-partitions whose first gap occurs at Λ_k and has length ℓ (i.e., the elements $\Lambda_k, \Lambda_{k+1}, \ldots, \Lambda_{k+\ell-1}$ are missing, but $\Lambda_{k+\ell}$ is not) is given by

$$\prod_{j=1}^{k-1} z^{\Lambda_j} \cdot z^{\Lambda_{k+\ell}} \cdot \prod_{j=k+\ell+1} (1 + z^{\Lambda_j}).$$

The same argument as in the proof of Theorem 5 (and Theorem 6) can be applied to show that the probability for a gap of length ℓ to occur at Λ_k tends to $2^{-k-\ell}$ for fixed k and ℓ, which proves the geometric distribution (and independence of position and length of the first gap). ■

Remark. The generating function for all partitions with the property that the first r-gap (i.e., the first sequence consisting of parts occurring with multiplicity $< r$) has length at least ℓ is given by

$$\sum_{k=1}^{\infty} (1 - z^{kr})(1 - z^{(k+1)r}) \cdots (1 - z^{(k+\ell-1)r}) z^{k(k-1)r/2} \prod_{j=1}^{\infty} (1 - z^j)^{-1}.$$

The sum can be simplified as follows: set $z^r = q$ and rewrite it as

$$\sum_{k=0}^{\infty} \prod_{j=k+1}^{k+\ell} (1 - q^j) q^{k(k+1)/2} = \prod_{j=1}^{\ell} (1 - q^j) \cdot \sum_{k=0}^{\infty} \frac{\prod_{j=\ell+1}^{k+\ell} (1 - q^j)}{\prod_{j=1}^{k} (1 - q^j)} q^{k(k+1)/2}.$$

This, however, is a special case of a well-known q-series (see [1, Corollary 2.7]); it is equal to the product

$$\prod_{j=1}^{\ell} (1 - q^j) \prod_{m=1}^{\infty} (1 - q^{2m+\ell})(1 + q^m) = \prod_{m=1}^{\infty} \frac{1 - q^{2m}}{1 - q^{2m+\ell-1}}$$

after a couple of simplifications. Therefore, our generating function can be written as the product

$$\prod_{m=1}^{\infty} \frac{1-z^{2mr}}{1-z^{(2m+\ell-1)r}} \prod_{j=1}^{\infty}(1-z^j)^{-1},$$

which can be interpreted as the generating function for a different kind of partitions. In particular, let us mention the special case that $r=1$ and ℓ is odd:

Proposition 9. *The number of partitions of n whose first gap has length at least ℓ, where ℓ is odd, is equal to the number of partitions of n that do not contain any of the parts $2, 4, \ldots, \ell-1$, as well as the number of partitions of n that contain each of the parts $1, 2, \ldots, \frac{\ell-1}{2}$ at most once.*

It would be interesting to see a combinatorial proof of this result. Furthermore, more precise asymptotics can be obtained from the product representation of the generating function:

Proposition 10. *The probability that the length of the first r-gap in a random partition of n is at least ℓ is asymptotically*

$$\Gamma\left(\frac{\ell+1}{2}\right)(2r)^{(\ell-1)/2}\zeta(2)^{(\ell-1)/4}n^{-(\ell-1)/4}.$$

Of course, “at least” can be replaced by “exactly” in the statement of this proposition.

3. Largest repeated parts

This problem is dual to the one that was considered in the previous section: How is the largest repeated part size in a random distribution distributed? More generally, for a fixed value of r, one may consider the distribution of the largest part whose multiplicity is at least r. Note that the case $r=1$ corresponds to the largest part in a partition, while $r=2$ is the original problem (largest repeated part). It should be noted that such a part does not have to exist (in such a case, we might simply define it to be 0 for the sake of convenience). However, it will turn out that this is almost surely not the case for fixed r as $n \to \infty$ (i.e., the probability of this event tends to 0). Furthermore, it is clear that this problem does not make sense for partitions into unequal parts, as opposed to the problem of the smallest gap.

The case $r=1$ has already been mentioned in the introduction: it is known that the largest part size follows a Gumbel (extreme value) distribution. The aim of this section is to show that this is also the limit distribution in the case of arbitrary r; in particular, the largest repeated part size asymptotically follows a Gumbel distribution.

More generally, we will prove the following theorem on Λ-partitions:

Theorem 11. *Let $R_{r,n}$ be the largest part in a random Λ-partition of n that occurs at least r times. Suppose that the conditions of Meinardus' scheme are satisfied. Set $M = (A\Gamma(\alpha+1)\zeta(\alpha+1))^{1/(\alpha+1)}$ and $y = Mn^{-1/(\alpha+1)}$; furthermore, assume that*

$$\left(\max(\tfrac{\alpha}{2}, \alpha - \eta_1) + \epsilon\right) \cdot \frac{1}{(\alpha+1)Mr} n^{1/(\alpha+1)} \log n \leq \Lambda_k \leq n^{(1+\alpha/2-\epsilon)/(\alpha+1)}$$

holds for some $\epsilon > 0$, where η_1 is taken as in Lemma 4. Then the following asymptotic formula for the probability $\mathbb{P}(R_{r,n} = \Lambda_k)$ holds:

$$\mathbb{P}(R_{r,n} = \Lambda_k) \sim e^{-ry\Lambda_k} \exp\left(- \sum_{j=k+1}^{\infty} e^{-ry\Lambda_j} \right).$$

Special cases (in particular, the case of ordinary unrestricted partitions, Theorem 2) will be discussed at the end of this section.

Proof. The approach is very similar to the one that was used in the proof of Theorem 3, and also follows essentially the lines of Szekeres [17, 18, 20]. As in the previous section, we start with the generating function. If we want to count partitions with the property that Λ_k is the largest part that occurs at least r times, then the corresponding generating function is given by

$$H(z) = \prod_{j=1}^{k-1} (1 - z^{\Lambda_j})^{-1} \cdot \frac{z^{r\Lambda_k}}{1 - z^{\Lambda_k}} \cdot \prod_{j=k+1}^{\infty} \frac{1 - z^{r\Lambda_j}}{1 - z^{\Lambda_j}} = z^{r\Lambda_k} \prod_{j=k+1}^{\infty} (1 - z^{r\Lambda_j}) \cdot G(z),$$

where $G(z)$ is the ordinary generating function for Λ-partitions. Now substitute $z = e^{-\tau}$, $h(\tau) = H(z)$, $g(\tau) = G(z)$, and apply the residue theorem to obtain the integral representation

$$p_\Lambda(n) \cdot \mathbb{P}(R_{r,n} = \Lambda_k) = \int_{-1/2}^{1/2} h(y + 2\pi i x) e^{ny + 2\pi i n x} \, dx \tag{6}$$

upon change of variables; y is taken as in the statement of the theorem (as in the proof of Theorem 3, this choice is based on the fact that y is a saddle point). We want to prove again that only the central part of the integral is asymptotically relevant. To this end, we need the following estimate:

Lemma 12. *Let $u > 0$ be a real number and k an integer; furthermore, we assume that $u\Lambda_k \geq C_5$ for some constant $C_5 > 0$. Then there exists a constant C_6 depending on C_5 such that the estimate*

$$\sum_{j=k+1}^{\infty} e^{-u\Lambda_j} \leq C_6 \Lambda_k^{\alpha} e^{-u\Lambda_k}$$

holds.

To prove this lemma, note that the number of elements of the sequence Λ inside the interval $[\ell\Lambda_k, (\ell+1)\Lambda_k]$ is $O((\ell\Lambda_k)^\alpha)$ (making use of Ikehara's Tauberian theorem once again). Therefore, we have

$$\sum_{j=k+1}^{\infty} e^{-u\Lambda_j} \ll \Lambda_k^\alpha \sum_{\ell=1}^{\infty} \ell^\alpha e^{-\ell u \Lambda_k} \ll \Lambda_k^\alpha e^{-u\Lambda_k},$$

proving the lemma. Now we can proceed with the proof of the main theorem. Set $\beta = 1 + \frac{\alpha}{2}\left(1 - \frac{\delta}{2}\right)$, where $0 < \delta < 4\epsilon/\alpha$. By the lemma, we have

$$\begin{aligned}\left| e^{-\tau r \Lambda_k} \prod_{j=k+1}^{\infty} (1 - e^{-\tau r \Lambda_j}) \right| &\le \prod_{j=k+1}^{\infty} (1 + e^{-yr\Lambda_j}) \le \exp\left(\sum_{j=k+1}^{\infty} e^{-yr\Lambda_j} \right) \\ &\le \exp\left(C_6 \Lambda_k^\alpha e^{-yr\Lambda_k} \right).\end{aligned}$$

Note that $y\Lambda_k \to \infty$ by our assumptions, so that the lemma applies for sufficiently large n. Furthermore,

$$\begin{aligned}\Lambda_k^\alpha e^{-yr\Lambda_k} &\ll n^{\alpha/(\alpha+1)} (\log n)^\alpha \exp\left(-\frac{\alpha - \eta_1 + \epsilon}{(\alpha+1)} \log n \right) \\ &\ll (\log n)^\alpha n^{(\eta_1 - \epsilon)/(\alpha+1)} \ll (\log n)^\alpha y^{-(\eta_1 - \epsilon)}\end{aligned}$$

by our assumptions on Λ_k. If we combine this with Lemma 4, we find that

$$\begin{aligned}\int_{y^\beta}^{\frac{1}{2}} h(y + 2\pi i x) e^{ny + 2\pi i n x}\, dx &= O\left(\exp\left(ny + A\Gamma(\alpha)\zeta(\alpha+1) y^{-\alpha} - C_7 y^{-\eta_1} \right) \right) \\ &= O\left(p_\Lambda(n) \exp\left(-C_8 n^{\eta_1/(\alpha+1)} \right) \right),\end{aligned}$$

showing once again that the outer parts of the integral are negligible, so that we can focus on the integral between $-y^\beta$ and y^β. For $|x| \le y^\beta$, one has

$$e^{-\tau r \Lambda_k} = e^{-yr\Lambda_k} \left(1 + O(\Lambda_k y^\beta) \right) = e^{-yr\Lambda_k} \left(1 + O\left(n^{(\delta\alpha/4 - \epsilon)/(\alpha+1)} \right) \right)$$

by our assumptions on Λ_k. Since we chose δ to be less than $4\epsilon/\alpha$, the exponent in the error term is indeed negative. The second factor in the generating function is estimated as follows:

$$
\begin{aligned}
\sum_{j=k+1}^{\infty} \log\left(1-e^{-\tau r\Lambda_j}\right) &= -\sum_{m=1}^{\infty}\frac{1}{m}\sum_{j=k+1}^{\infty} e^{-m\tau r\Lambda_j} \\
&= -\sum_{j=k+1}^{\infty} e^{-\tau r\Lambda_j} + O\left(\sum_{m=2}^{\infty}\sum_{j=k+1}^{\infty} e^{-myr\Lambda_j}\right) \\
&= -\sum_{j=k+1}^{\infty} e^{-yr\Lambda_j} + O\left(y^{\beta}\Lambda_k \sum_{j=k+1}^{\infty} e^{-yr\Lambda_j} + \sum_{j=k+1}^{\infty} e^{-2yr\Lambda_j}\right) \\
&= -\sum_{j=k+1}^{\infty} e^{-yr\Lambda_j} + O\left(y^{\beta}\Lambda_k^{\alpha+1}e^{-yr\Lambda_k} + \Lambda_k^{\alpha}e^{-2yr\Lambda_k}\right) \\
&= -\sum_{j=k+1}^{\infty} e^{-yr\Lambda_j} + O\Big(n^{-\beta/(\alpha+1)+1}(\log n)^{\alpha+1}\exp\left(-\tfrac{\alpha/2+\epsilon}{\alpha+1}\log n\right) \\
&\qquad + n^{\alpha/(\alpha+1)}(\log n)^{\alpha}\exp\left(-\tfrac{\alpha+2\epsilon}{\alpha+1}\log n\right)\Big) \\
&= -\sum_{j=k+1}^{\infty} e^{-yr\Lambda_j} + O\Big(n^{(\delta\alpha/4-\epsilon)/(\alpha+1)}(\log n)^{\alpha+1} + n^{-2\epsilon/(\alpha+1)}(\log n)^{\alpha}\Big),
\end{aligned}
$$

making use of Lemma 12 again. Therefore, we finally have

$$\prod_{j=1}^{\infty}\left(1-e^{-\tau r\Lambda_j}\right) = \exp\left(-\sum_{j=k+1}^{\infty} e^{-yr\Lambda_j}\right)\left(1+O\left(n^{(\delta\alpha/4-\epsilon)/(\alpha+1)}(\log n)^{\alpha+1}\right)\right).$$

The final step is analogous to Theorem 3 again:

$$
\begin{aligned}
p_{\Lambda}(n)\cdot\mathbb{P}(R_{r,n}=\Lambda_k) &= \left(\int_{-1/2}^{1/2} g(y+2\pi ix)e^{ny+2\pi inx}\,dx\right)\cdot e^{-ry\Lambda_k} \\
&\qquad \cdot\exp\left(-\sum_{j=k+1}^{\infty} e^{-yr\Lambda_j}\right)(1+o(1)),
\end{aligned}
$$

and thus

$$p_{\Lambda}(n)\cdot\mathbb{P}(R_{r,n}=\Lambda_k) = p_{\Lambda}(n)\cdot e^{-ry\Lambda_k}\exp\left(-\sum_{j=k+1}^{\infty} e^{-yr\Lambda_j}\right)(1+o(1)),$$

which proves the theorem. ∎

Example. In the case of ordinary unrestricted integer partitions ($\Lambda_k = k$), we have $y = \frac{\pi}{\sqrt{6n}}$ and

$$\sum_{j=k+1}^{\infty} e^{-ryj} = \frac{1}{ry}e^{-rky}(1+O(y)).$$

Hence, if we set $k = \frac{\sqrt{6n}}{2r\pi} \log n + h\sqrt{n} = \frac{1}{2ry} \log n + h\sqrt{n}$, then we find that

$$e^{-ryk} = n^{-1/2} \exp\left(-\frac{\pi r h}{\sqrt{6}}\right)$$

and finally

$$\mathbb{P}(G_{r,n} = k) \sim \frac{1}{\sqrt{n}} \exp\left(-\frac{\pi r h}{\sqrt{6}} - \frac{\sqrt{6}}{\pi r} \exp\left(-\frac{\pi r h}{\sqrt{6}}\right)\right),$$

proving Theorem 2. Similar calculations show that for $\Lambda_k = ak + b$ (a, b coprime), one has

$$\mathbb{P}(G_{r,n} = \Lambda_k) \sim \frac{1}{\sqrt{n}} \exp\left(-\frac{\pi r h}{\sqrt{6a}} - \frac{\sqrt{6}}{\pi r \sqrt{a}} \exp\left(-\frac{\pi r h}{\sqrt{6a}}\right)\right)$$

if $\Lambda_k = \frac{\sqrt{6an}}{2r\pi} \log n + h\sqrt{n}$. For $\Lambda_k = k^m$, things are slightly more complicated: one has to use the asymptotic formula

$$\sum_{j=k+1}^{\infty} e^{-ryj^m} \sim \frac{1}{mryk^{m-1}} e^{-ryk^m}.$$

Further, one finds that $M = \left(\frac{1}{m}\Gamma\left(1+\frac{1}{m}\right)\zeta\left(1+\frac{1}{m}\right)\right)^{m/(m+1)}$ in this case. If now

$$\Lambda_k = \frac{1}{(m+1)Mr} n^{m/(m+1)} \log n - \frac{m-1}{mMr} n^{m/(m+1)} \log\log n + hn^{m/(m+1)},$$

then the following limit law holds:

$$\mathbb{P}(G_{r,n} = \Lambda_k) \sim \frac{(\log n)^{(m-1)/m}}{n^{1/(m+1)}} \exp\left(-rMh - \frac{(m+1)^{(m-1)/m}}{m(Mr)^{1/m}} \exp(-rMh)\right).$$

The occurrence of a $\log\log n$ term in the last formula (that is not present in the case $m = 1$) is quite surprising.

Remark. It is interesting to note that the limit distributions that arise in the birthday problem and the coupon collector problem, which are somewhat similar to the problems considered in this paper and also duals of each other in a certain sense, are also the Rayleigh and the Gumbel distribution, respectively (see [7]).

Remark. The length of the smallest gap is mapped to the multiplicity of the largest repeated part (minus 1) upon conjugation of the Ferrers diagram. Therefore, Proposition 10 also implies that the probability for the multiplicity of the largest repeated part in an ordinary integer partition of n to be ℓ is asymptotically equal to

$$\Gamma\left(\frac{\ell}{2}\right) (2r)^{(\ell-2)/2} \zeta(2)^{(\ell-2)/4} n^{-(\ell-2)/4}.$$

References

[1] G. E. Andrews. *The theory of partitions*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1998.

[2] T. M. Apostol. *Modular functions and Dirichlet series in number theory*, volume 41 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990.

[3] C. Brennan, A. Knopfmacher, and S. Wagner. The distribution of ascents of size d or more in partitions of n. *Combin. Probab. Comput.*, 17(4):495–509, 2008.

[4] S. Corteel, B. Pittel, C. D. Savage, and H. S. Wilf. On the multiplicity of parts in a random partition. *Random Structures Algorithms*, 14(2):185–197, 1999.

[5] P. Erdős and J. Lehner. The distribution of the number of summands in the partitions of a positive integer. *Duke Math. J.*, 8:335–345, 1941.

[6] P. Erdős and M. Szalay. On the statistical theory of partitions. In *Topics in classical number theory, Vol. I, II (Budapest, 1981)*, volume 34 of *Colloq. Math. Soc. János Bolyai*, pages 397–450. North-Holland, Amsterdam, 1984.

[7] P. Flajolet and R. Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009.

[8] B. Fristedt. The structure of random partitions of large integers. *Trans. Amer. Math. Soc.*, 337(2):703–735, 1993.

[9] W. M. Y. Goh and E. Schmutz. The number of distinct part sizes in a random integer partition. *J. Combin. Theory Ser. A*, 69(1):149–158, 1995.

[10] P. J. Grabner and A. Knopfmacher. Analysis of some new partition statistics. *Ramanujan J.*, 12(3):439–454, 2006.

[11] G. H. Hardy and S. Ramanujan. Asymptotic formulæ in combinatory analysis [Proc. London Math. Soc. (2) **16** (1917), Records for 1 March 1917]. In *Collected papers of Srinivasa Ramanujan*, page 244. AMS Chelsea Publ., Providence, RI, 2000.

[12] H.-K. Hwang. Limit theorems for the number of summands in integer partitions. *J. Combin. Theory Ser. A*, 96(1):89–126, 2001.

[13] J. Korevaar. A century of complex Tauberian theory. *Bull. Amer. Math. Soc. (N.S.)*, 39(4):475–531 (electronic), 2002.

[14] G. Meinardus. Asymptotische Aussagen über Partitionen. *Math. Z.*, 59:388–398, 1954.

[15] L. R. Mutafchiev. On the maximal multiplicity of parts in a random integer partition. *Ramanujan J.*, 9(3):305–316, 2005.

[16] H. Rademacher. On the expansion of the partition function in a series. *Ann. of Math. (2)*, 44:416–422, 1943.

[17] G. Szekeres. An asymptotic formula in the theory of partitions. *Quart. J. Math., Oxford Ser. (2)*, 2:85–108, 1951.

[18] G. Szekeres. Some asymptotic formulae in the theory of partitions. II. *Quart. J. Math., Oxford Ser. (2)*, 4:96–111, 1953.

[19] G. Szekeres. Asymptotic distribution of the number and size of parts in unequal partitions. *Bull. Austral. Math. Soc.*, 36(1):89–97, 1987.

[20] G. Szekeres. Asymptotic distribution of partitions by number and size of parts. In *Number theory, Vol. I (Budapest, 1987)*, volume 51 of *Colloq. Math. Soc. János Bolyai*, pages 527–538. North-Holland, Amsterdam, 1990.

[21] S. Wagner. On the distribution of the longest run in number partitions. *Ramanujan J.*, 20(2):189–206, 2009.

Stephan Wagner, Department of Mathematical Sciences, Stellenbosch University, Private Bag X1, Matieland 7602, South Africa

E-mail address: swagner@sun.ac.za