JOURNAL OF COMBINATORIAL THEORY, Series A 69, 149–158 (1995)

Note

The Number of Distinct Part Sizes in a Random Integer Partition*

WILLIAM M. Y. GOH AND ERIC SCHMUTZ

M.C.S. Department, Drexel University, Philadelphia, Pennsylvania 19104

Communicated by Andrew Odlyzko

Received June 21, 1993

We prove a central limit theorem for the number of different part sizes in a random integer partition. If λ is one of the $P(n)$ partitions of the integer n, let $\mathbf{D}_n(\lambda)$ be the number of distinct part sizes that λ has. (Each part size counts once, even though there may be many parts of a given size.) For any fixed x,

$$\frac{\#\{\lambda : \mathbf{D}_n(\lambda) \le A_n + xB_n\}}{P(n)} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$$

as $n \to \infty$, where $A_n = (\sqrt{6}/\pi)n^{1/2}$ and $B_n = (\sqrt{6}/2\pi - \sqrt{54}/\pi^3)^{1/2} n^{1/4}$.

1. INTRODUCTION

A partition of n is a multiset of positive integers whose sum is n. The summands, i.e., the elements of the multiset, are called parts. Let $\mathscr{P}_n$ be the set of all partitions of n, and let $P(n) = |\mathscr{P}_n|$. Put the uniform probability measure on $\mathscr{P}_n$; $m_n(\{\lambda\}) = 1/P(n)$ for all $\lambda \in \mathscr{P}_n$. Then any real valued function $\mathbf{X}_n$ on $\mathscr{P}_n$ can be regarded as a random variable. If $\mathbf{X}_n$ $(n = 1, 2, 3, \dots)$ is a sequence of functions that arises naturally in combinatorics, then it is often reasonable to ask questions about the asymptotic distribution of values as $n \to \infty$. Erdős and Lehner [3] were apparently the first to study random integer partitions in this way. Subsequent work by a number of authors provides considerable information about the structure of a "typical" partition. (See, for example, Fristedt [4] and Szalay and Turán [6].)

*Research supported by the National Science Foundation.

0097-3165/95 $6.00

If λ is a partition of the integer n (written $\lambda \vdash n$), let $\mathbf{N}_n(\lambda)$ and $\mathbf{D}_n(\lambda)$ denote the number of parts that λ has, counted respectively with and without multiplicity. For example, if $\lambda = \{(1^{(n-2)}, 2\}$ consists of $n - 2$ parts of size 1 and 1 part of size 2, then $\mathbf{N}_n(\lambda) = n - 1$ and $\mathbf{D}_n(\lambda) = 2$. In this paper we expand on Wilf's observation that, for most partitions of n, $\mathbf{N}_n$ is much larger than $\mathbf{D}_n$.

Let $\mathbf{N}_n^{(r)}(\lambda)$ be the size of the rth largest part that λ has. It it well known that $\mathbf{N}_n^{(1)}$ and $\mathbf{N}_n$ are identically distributed. Erdős and Lehner proved that, for any fixed x,

$$\lim_{n\to\infty}\left(\frac{\#\{\lambda \vdash n: \mathbf{N}_n^{(1)}(\lambda) \le (\sqrt{6n}/\pi)\log(\sqrt{6n}/\pi) + x(\sqrt{6n}/\pi)\}}{P(n)}\right)$$
$$= \exp[-e^{-x}].$$

(Later Fristedt strengthened this result by determining the joint distribution of the d largest parts [4], i.e., the asymptotic distribution of the random vector $(\mathbf{N}_n^{(1)}, \mathbf{N}_n^{(2)}, \dots, \mathbf{N}_n^{(d)})$ for any fixed d.) Thus a typical partition of n has $\Theta(\sqrt{n}\log n)$ parts. Clearly this cannot be true for $\mathbf{D}_n$. In fact, the inequality $\sum_{i=1}^{\mathbf{D}_n} i \le n$ implies that $\mathbf{D}_n \le \sqrt{2n}$ for *all* partitions of n. Wilf proved, as an example in [7], that the expected value of $\mathbf{D}_n$ is $(\sqrt{6n}/\pi)(1 + o(1))$. Our goal here is to extend Wilf's result with the following central limit theorem:

THEOREM. *For any fixed x,*

$$\frac{\#\{\lambda \vdash n: \mathbf{D}_n(\lambda) \le A_n + xB_n\}}{P(n)} \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}\,dt$$

as $n \to \infty$, where $A_n = (\sqrt{6}/\pi)n^{1/2}$ and $B_n = (\sqrt{6}/2\pi - \sqrt{54}/\pi^3)^{1/2}n^{1/4}$.

Some different, but closely related results can be found in [5].

Proof. Let $p_{n,k}$ be the fraction of partitions for which there are exactly k different part sizes, i.e.,

$$p_{n,k} := m_n(\{\lambda: \mathbf{D}_n(\lambda) = k\}).$$

It is not difficult to verify (see [7] for an interesting proof) that

$$\sum_{k\ge 0} p_{n,k}y^k = \frac{1}{P(n)}COEFF_{w^n}\left\{\prod_{m\ge 1}\left(1 + \frac{yw^m}{1-w^m}\right)\right\}.$$

Let $\mu_n(r) := \Sigma_k k^r p_{n,k}$, and let $M_n(t) := \Sigma_{r \geq 0} \mu_n(r) t^r / r!$. Then

$$M_n(t) = \frac{1}{P(n)} COEFF_{w^n}\left\{ \prod_{m \geq 1} \left(1 + \frac{e^t w^m}{1 - w^m} \right) \right\}.$$

This is the "moment generating function." By a variant of the continuity theorem [2], it suffices to prove that, for any real number t,

$$\lim_{n \to \infty} M_n(t/B_n) e^{-tA_n/B_n} = e^{t^2/2}.$$

2. Major Contribution

Let $z_n := (e^{t/B_n} - 1)$, and let $f(w) := \prod_{m \geq 1} 1/(1 - w^m)$. Then, for a suitable contour C_n (to be specified later),

$$\begin{aligned} M_n(t/B_n) &= \frac{1}{2\pi i P(n)} \int_{C_n} \prod_{m \geq 1} \left(\frac{1}{1 - w^m} + \frac{z_n w^m}{1 - w^m} \right) \frac{dw}{w^{n+1}} \\ &= \frac{1}{2\pi i P(n)} \int_{C_n} f(w) \prod_{m \geq 1} (1 + z_n w^m) \frac{dw}{w^{n+1}}. \end{aligned}$$

Setting $w = e^{-2\pi v}$, we get

$$M_n(t/B_n) = \frac{-i}{P(n)} \int_{\gamma_n} f(e^{-2\pi v}) \prod_{m \geq 1} \left(1 + z_n e^{-2\pi m v}\right) e^{2\pi n v} \, dv.$$

This is a convenient point at which to specify the contour of integration. Shown schematically in Fig. 1 is the trace of γ_n in the complex v-plane. It consists of a circular arc γ_n', and a vertical segment γ_n''. The circular arc has radius r_n (defined in a moment), and extends from $\theta = -1/\log^2 n$ to $\theta = +1/\log^2 n \overset{\text{def}}{=\!:} \theta_n$. In order to define r_n, first let $\Delta_n := \Sigma_{l \geq 1} (-1)^{l-1} z_n^l / l^2$. Then let $r_n := (1/\sqrt{24n})(1 + (6/\pi^2)\Delta_n)^{1/2}$. The reason for this choice of r_n will become clear later; for now, simply note r_n is small and positive. With these definitions we can give γ_n explicitly:

$$\gamma_n(x) = \begin{cases} r_n e^{i(x - \theta_n)}, & x \in [0, 2\theta_n] \\ r_n e^{i\theta_n} + i(x - 2\theta_n), & x \in [2\theta_n, 1 - 2r_n \sin\theta_n + 2\theta_n] \end{cases}.$$

Note that, because $w = e^{-2\pi v}$, the vertical part γ_n'' of γ_n corresponds to a circular arc C_n'' in the complex w-plane. For the first part of the integration, it is convenient to work in the v-plane, but for the second part

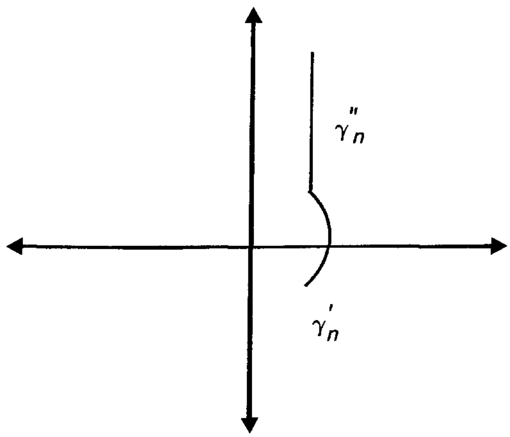

FIG. 1. The trace of γ_n.

it is more convenient to remain in the w-plane. Thus $M_n(t/B_n) = T_1 + T_2$, where

$$T_1 = \frac{-i}{P(n)} \int_{\gamma_n'} f(e^{-2\pi v}) \prod_{m \geq 1} (1 + z_n e^{-2\pi m v}) e^{2\pi n v} \, dv,$$

and

$$T_2 = \frac{1}{2\pi i P(n)} \int_{C_n''} \prod_{m \geq 1} \left(\frac{1}{1 - w^m} + \frac{z_n w^m}{1 - w^m} \right) \frac{dw}{w^{n+1}}.$$

In order to estimate T_1, we replace the integrand with a close approximation that is easier to integrate. The product in the integrand is approximated with the aid of the following lemma:

LEMMA 1. *For v on γ_n', we have*

$$\prod_{m \geq 1} (1 + z_n e^{-2\pi m v}) = e^{\Delta n / 2\pi v} (1 + z_n)^{-1/2} (1 + O(z_n v)),$$

where the constant implicit in the O can be chosen uniformly for all v on γ_n'.

Proof. It suffices to show that

$$\log\left(\prod_{m \geq 1} (1 + z_n e^{-2\pi m v}) \right) = \frac{\Delta_n}{2\pi v} - \frac{1}{2} \log(1 + z_n) + O(z_n v).$$

This can be verified using the Euler–Maclaurin formula, but a more appealing proof rests on the fact that, for s not an integral multiple of $2\pi i$,

$$\frac{1}{e^s - 1} = \frac{-1}{2} + \frac{i}{2}\cot\frac{is}{2} = \frac{-1}{2} + \frac{1}{s} + \sum_{k=1}^{\infty} \frac{2s}{s^2 + 4\pi^2 k^2}.$$

We have

$$\begin{aligned}
&\log\Bigl(\prod_{m\ge 1}\bigl(1 + z_n e^{-2\pi m v}\bigr)\Bigr)\\
&\quad= \sum_{m\ge1}\sum_{l\ge1} \frac{(-1)^{l-1} z_n^l}{l} e^{-2\pi m v l}\\
&\quad= \sum_{l\ge1} \frac{(-1)^{l-1} z_n^l}{l}\,\frac{e^{-2\pi v l}}{1 - e^{-2\pi v l}}\\
&\quad= \sum_{l\ge1} \frac{(-1)^{l-1} z_n^l}{l}\left(\frac{-1}{2} + \frac{1}{2\pi v l} + \sum_{k=1}^{\infty} \frac{4\pi v l}{(2\pi v l)^2 + 4\pi^2 k^2}\right)\\
&\quad= -\frac{1}{2}\log(1 + z_n) + \frac{\Delta_n}{2\pi v}\\
&\qquad+ v z_n\left\{\sum_{l\ge1}(-1)^{l-1} z_n^{l-1} \sum_{k=1}^{\infty} \frac{4\pi}{(2\pi v l)^2 + 4\pi^2 k^2}\right\}.
\end{aligned}$$

We need only show that the expression inside the large braces is $O(1)$. But if $v = a + ib$ is on γ_n', then $a > b$ and

$$\begin{aligned}
&\left|\sum_{l\ge1}(-1)^{l-1} z_n^{l-1} \sum_{k=1}^{\infty} \frac{4\pi}{(2\pi v l)^2 + 4\pi^2 k^2}\right|\\
&\quad< \sum_{l\ge1} |z_n|^{l-1} \sum_{k=1}^{\infty} \frac{4\pi}{\sqrt{\bigl(4\pi^2 l^2(a^2 - b^2) + 4\pi^2 k^2\bigr)^2 + 64\pi^4 l^4 a^2 b^2}}\\
&\quad< \sum_{l\ge1} |z_n|^{l-1} \sum_{k=1}^{\infty} \frac{4\pi}{4\pi^2 k^2} < 1. \quad \blacksquare
\end{aligned}$$

The other part of the integrand is f, which can be approximated with the help of the well known "transformation formula" (see, for example, [1]):

$$f(e^{-2\pi v}) = \sqrt{v}\, e^{\pi/12v - \pi v/12} f(e^{-2\pi/v}).$$

Hence

$$\frac{-i}{P(n)}\int_{\gamma_n'} f(e^{-2\pi v}) \prod_{m\geq 1}\left(1+z_n e^{-2\pi m v}\right)\cdot e^{2\pi n v}\,dv = I'_{\mathrm{maj}} + I'_{\mathrm{err}},$$

where

$$I'_{\mathrm{maj}} = \frac{-i}{P(n)}\int_{\gamma_n'} \sqrt{v}\, e^{\pi/(12v)} e^{\Delta n/2\pi v}(1+z_n)^{-1/2} e^{2\pi n v}\,dv,$$

and

$$I'_{\mathrm{err}} := \frac{-i}{P(n)}\int_{\gamma_n'} \sqrt{v}\, e^{\pi/(12v)} e^{\Delta n/(2\pi v)}(1+z_n)^{-1/2} e^{2\pi n v} \times\left(e^{-\pi v/12} f(e^{-2\pi/v})(1+O(z_n v)) - 1\right)dv.$$

It will later be shown that I'_{err} is negligible. For now we concentrate on the major contribution I'_{maj}. On γ_n', we have $v = r_n e^{i\theta}$, and therefore

$$I'_{\mathrm{maj}} = \frac{r_n^{3/2}}{P(n)\sqrt{1+z_n}}\int_{-\theta_n}^{\theta_n} e^{3i\theta/2}\exp\left[\left(\frac{\pi}{12}+\frac{\Delta_n}{2\pi}\right)\frac{1}{r_n e^{i\theta}} + 2\pi n r_n e^{i\theta}\right]d\theta.$$

Because of the way r_n was defined, this is equal to

$$\frac{r_n^{3/2}}{P(n)\sqrt{1+z_n}}\int_{-\theta_n}^{\theta_n} e^{3i\theta/2}\exp\left[\pi\sqrt{\frac{2n}{3}}\left(1+\frac{6\Delta_n}{\pi^2}\right)^{1/2}\cos\theta\right]d\theta.$$

Since $\sin(3\theta/2)$ is odd, this is equal to

$$\frac{r_n^{3/2}}{P(n)\sqrt{1+z_n}}\int_{-\theta_n}^{\theta_n} \cos\left(\frac{3\theta}{2}\right)\exp\left[\pi\sqrt{\frac{2n}{3}}\left(1+\frac{6\Delta_n}{\pi^2}\right)^{1/2}\cos\theta\right]d\theta. \quad (1)$$

To simplify notation, let $Q_n := \pi\sqrt{2n/3}\,(1+(6/\pi^2)\Delta_n)^{1/2}$. The integral is easily estimated by the method of Laplace, the conclusion being that

$$I'_{\mathrm{maj}} \sim \frac{\sqrt{2\pi}\, r_n^{3/2} e^{Q_n}}{P(n)\sqrt{(1+z_n)Q_n}}.$$

Recall that $z_n = e^{t/B_n} - 1 \to 0$ as $n\to\infty$, and $\Delta_n = \sum_{l\geq 1}(-1)^{l-1} z_n^l / l^2$.

Hence

$$Q_n = \pi\sqrt{2n/3} + \frac{\sqrt{6n}\,t}{\pi B_n} + t^2/2 + O(n^{-1/4}),$$

and therefore

$$I'_{\mathrm{maj}} \sim \frac{\sqrt{2\pi}\, r_n^{3/2} \exp\left[\pi\sqrt{2n/3} + \sqrt{6n}\,t/(\pi B_n) + t^2/2\right]}{P(n)\left(\pi\sqrt{2n/3}\right)^{1/2}}.$$

Recall that $r_n = (1/\sqrt{24n})(1 + (6/\pi^2)\Delta_n)^{1/2}$. Combining this with the well-known [1] fact that $P(n) \sim e^{\pi\sqrt{2n/3}}/4\sqrt{3}\,n$, we get the formula we want:

$$e^{-tA_n/B_n} I'_{\mathrm{maj}} \to e^{t^2/2} \qquad \text{as } n \to \infty.$$

3. Error Terms

It remains to be shown that I'_{err} and T_2 are negligible. To estimate I'_{err}, note that $f(u) = 1 + O(u)$ as $u \to 0$. Since $r_n = \Theta(1/\sqrt{n})$, it follows that

$$f(e^{-2\pi/v}) = 1 + O(e^{-k\sqrt{n}})$$

for some fixed k. Hence

$$\left(e^{-\pi v/12} f(e^{-2\pi/v})(1 + O(z_n v)) - 1\right) = O(v)$$

uniformly for v on γ'_n, and therefore

$$|I'_{\mathrm{err}}| \ll \frac{1}{P(n)\sqrt{1+z_n}} \int_{\gamma'_n} |v|^{3/2} |e^{\pi/(12v) + \Delta_n/(2\pi v)} e^{2\pi n v}|\,|dv|.$$

Since $z_n \to 0$, this is

$$\ll \frac{n^{-5/4}}{P(n)} \int_{-\theta_n}^{\theta_n} \left|\exp\left[\left(\frac{\pi}{12} + \frac{\Delta_n}{2\pi}\right) r_n^{-1} e^{-i\theta} + 2\pi n r_n e^{i\theta}\right]\right| d\theta.$$

As before, with the estimation of I'_{maj}, the choice of r_n implies that this

$$\ll \frac{n^{-5/4}}{P(n)} \int_{-\theta_n}^{\theta_n} \exp\left[\pi\sqrt{2n/3}\left(1 + \frac{6\Delta_n}{\pi^2}\right)^{1/2} \cos\theta\right] d\theta.$$

Comparing this with Eq. (1), we see immediately that it is $o(I'_{\mathrm{maj}})$.

Finally, we must show that T_2 is negligible. Our argument is motivated by a paper of Wright [8]. The integrand is

$$J_n(w) := \prod_{m \geq 1} (1 + z_n w^m) \cdot \prod_{m \geq 1} \frac{1}{1 - w^m}.$$

Then

$$\begin{aligned}\log J_n(w) &= \sum_{m \geq 1} \left(\sum_{l \geq 1} \frac{(-1)^{l-1} z_n^l w^{ml}}{l} + \sum_{l \geq 1} \frac{w^{ml}}{l} \right) \\ &= \sum_{l \geq 1} \frac{\left(1 + (-1)^{l-1} z_n^l\right)}{l} \frac{w^l}{1 - w^l} \\ &= \frac{w(1 + z_n)}{1 - w} + \sum_{l \geq 2} \frac{\left(1 + (-1)^{l-1} z_n^l\right)}{l} \frac{w^l}{1 - w^l}.\end{aligned}$$

Since $z_n \to 0$, we have $(1 + (-1)^{l-1} z_n^l) > 0$, and therefore

$$\begin{aligned}|\log J_n(w)| &\leq \frac{|w|(1 + z_n)}{|1 - w|} + \sum_{l \geq 2} \frac{\left(1 + (-1)^{l-1} z_n^l\right)}{l} \frac{|w|^l}{1 - |w|^l} \\ &= \log J_n(|w|) - \left(\frac{|w|(1 + z_n)}{1 - |w|} - \frac{|w|(1 + z_n)}{|1 - w|} \right).\end{aligned}$$

Let $v_n = r_n \cos\theta$. For w on C_n'', we have

$$\begin{aligned}\frac{1 - |w|}{|1 - w|} &\leq \frac{1 - |e^{-2\pi v_n}|}{|1 - e^{-2\pi v_n}|} \\ &= \frac{1 - e^{-2\pi r_n \cos\theta_n}}{\left|1 - \exp\left[-2\pi r_n e^{i\theta_n}\right]\right|} = \frac{2\pi r_n \cos\theta_n + O\left(r_n^2\right)}{\left|2\pi r_n e^{i\theta_n} + O\left(r_n^2\right)\right|} \\ &= \cos\theta_n + O(r_n).\end{aligned}$$

Since $r_n = O(1/\sqrt{n})$, it follows that

$$1 - \frac{1 - |w|}{|1 - w|} > \frac{\theta_n^2}{3} = \frac{1}{3 \log^4 n}$$

for n sufficiently large. But then we have

$$\left(\frac{|w|(1+z_n)}{1-|w|}-\frac{|w|(1+z_n)}{|1-w|}\right) \geq \frac{|e^{-2\pi v_n}|(1+z_n)}{1-|e^{-2\pi v_n}|}\left(1-\frac{1-|e^{-2\pi v_n}|}{|1-e^{-2\pi v_n}|}\right)$$
$$\geq \frac{\sqrt{n}}{6\log^4 n},$$

and therefore

$$|T_2| \leq \frac{1}{2\pi P(n)}\int_{C_n''}\frac{|J_n(w)|\,|dw|}{|w|^{n+1}} \leq \frac{e^{-\sqrt{n}/(6\log^4 n)}}{2\pi P(n)}\int_{C_n''} J_n(|w|)\frac{|dw|}{|w|^{n+1}}$$

On C_n'', we have $|w| = e^{-2\pi r_n \cos\theta_n}$. As in Lemma 1, one finds that

$$\prod_{m\geq 1}(1+z_n|w|^m) = \exp\left[\frac{\Delta_n}{2\pi r_n\cos\theta_n}-\frac{\log(1+z_n)}{2}+O(r_n z_n\cos\theta_n)\right].$$

Then, using the transformation formula again, we get

$$J_n(|w|) \sim \sqrt{r_n\cos\theta_n}\exp\left[\frac{\pi}{12 r_n\cos\theta_n}+\frac{\Delta_n}{2\pi r_n\cos\theta_n}\right](1+z_n)^{-1/2}$$
$$\ll n^{-1/4}\exp\left[\frac{\pi}{12}\sqrt{24n}+\frac{\Delta_n}{2\pi}\sqrt{24n}\right].$$

Thus

$$T_2 \ll \frac{e^{-\sqrt{n}/6\log^4 n}}{P(n)}n^{-1/4}\exp\left[\frac{\pi}{12}\sqrt{24n}+\frac{\Delta_n}{2\pi}\sqrt{24n}+\frac{2\pi n}{\sqrt{24n}}\right].$$

Using again the formula $P(n) \sim e^{\pi\sqrt{2n/3}}/4\sqrt{3}n$, we get the estimate we need: $T_2 e^{-tA_n/B_n} = o(1)$. ∎

ACKNOWLEDGMENT

Rodney Canfield suggested several references, and also made some helpful criticisms of an early draft of this paper.

References

1. T. Apostal, "Modular Forms and Dirichlet Series in Number Theory," Springer-Verlag, Berlin/New York, 1976.
2. J. Curtiss, A note on the theory of moment generating functions, *Ann. of Math. Statist.* **13** (1942), 430–433.
3. P. Erdős and J. Lehner, Distribution of the number of summands in the partitions of a positive integer, *Duke Math. J.* **8** (1941), 335–345.
4. B. Fristedt, The structure of random partitions of large integers, *Trans. Amer. Math. Soc.*, to appear.
5. L. B. Richmond, The moments of partitions, I, *Acta Arith.* **26**, 411–425.
6. M. Szalay and P. Turán, On some problems of the statistical theory of partitions with application to characters of the symmetric group, *Acta Math. Sci. Hungar.* **29** (1977) and **32** (1978).
7. H. Wilf, Three problems in combinatorial asymptotics, *J. Combin. Theory Ser. A* **35** (1983), 199–207.
8. E. M. Wright, Asymptotic partition formulae. I. Plane partitions, *Quart. J. Math. Oxford Ser.* **2** (1931), 177–189.