THE STRUCTURE OF RANDOM PARTITIONS OF LARGE INTEGERS

BERT FRISTEDT

ABSTRACT. Random partitions of integers are treated in the case where all par-
titions of an integer are assumed to have the same probability. The focus is
on limit theorems as the number being partitioned approaches ∞. The lim-
iting probability distribution of the appropriately normalized number of parts
of some small size is exponential. The large parts are described by a particular
Markov chain. A central limit theorem and a law of large numbers holds for
the numbers of intermediate parts of certain sizes. The major tool is a simple
construction of random partitions that treats the number being partitioned as a
random variable. The same technique is useful when some restriction is placed
on partitions, such as the requirement that all parts must be distinct.

1. INTRODUCTION

A *partition* of a positive integer is a way of writing it as the sum of positive
integers without regard to order; the summands are called *parts*. Thus, there
are five partitions of the number 4: $1 + 1 + 1 + 1$, $2 + 1 + 1$, $2 + 2$, $3 + 1$, and
4. The number of partitions of the nonnegative integer n will be denoted by
$p(n)$; by definition $p(0) = 1$ and the one partition of 0 is the empty partition.
A formula for the generating function of the sequences $(p(0), p(1), p(2), \ldots)$
is well known:

$$(1.1) \qquad \sum_{n=0}^{\infty} p(n)q^n = \prod_{k=1}^{\infty} (1 - q^k)^{-1}.$$

Were one asked how many partitions of 4 have no part equal to 1, one would
answer "two"—namely, $2 + 2$ and 4. The same information could be given
in probabilistic language: "The probability is $2/5$ that a random partition of
4 has no part equal to 1." Alternatively, one could say that the probability
distribution of the number of parts of size 1 assigns measure (or probability)
$2/5$ to the value 0. When making such a direct translation between counting and
probabilistic statements, one is tacitly assuming, for the probabilistic setting,
that the partitions are equiprobable. The equiprobable probabilistic approach
is used in this paper. P_n will denote the probability measure which assigns
probability $1/p(n)$ to each of the partitions of n. Our goal is to study some
aspects of P_n as $n \to \infty$. Any asymptotic relation we obtain for P_n can be

easily transformed into an asymptotic relation about the number of partitions satisfying some conditions; multiplication at the appropriate place by $p(n)$ is all that is required.

The probabilistic structure of a random partition of an integer is quite complicated: the value of the largest part influences, in a somewhat complicated manner, the value of the second largest part and the number of parts equal to, say, 16 influences the number of parts equal to 13. In view of these intricacies a simpler one-parameter probabilistic model based on generating functions is introduced, one in which the number being partitioned is itself random. The results that we obtain for the simpler model are limit theorems for the generating function parameter approaching 1; they convert, with some work, into theorems for n approaching ∞. The tools for this conversion are described in §4.

Major results are given in §2. In §3 some consequences are developed and connections with some of the results of Szalay and Turán [7, 8] are explained. I thank the referee for indicating that such an explanation is appropriate. The companion paper [9] by the same authors is not directly related to this paper, but is included for completeness in the reference list.

The proofs of the results stated in §2 are given in §§5, 6, 7, and 8. One result says that the number of parts of different small sizes are asymptotically independent and exponentially distributed; it is proved in §5. There is also a theorem for the large parts—they behave, in the limit, like a particular Markov chain; the proof is in §6. Thus, the limiting probability distributions of the large parts are obtained—in particular, duplicating the result of Erdös and Lehner [3] giving the limiting probability distribution of the largest part. Also treated is a variation where each part is multiplied by its multiplicity and the larger of the products thus obtained are studied. The limiting probability distribution of the largest such product was previously obtained by Erdös and Szalay [4]. The relevant proof, generalizing the result of Erdös and Szalay, is in §7. The last two results in §2 concern the number of parts satisfying an inequality. One of these two theorems is about the number of parts greater than k_n, assumed to approach ∞ faster than $n^{1/2}$. The other is about the number of parts less than k_n when it is assumed that $k_n/n^{1/2} \to 0$. Section 8 contains the proofs. Each of §§5, 6, 7, and 8 begins with the appropriate proof without a formal announcement. An end of the proof symbol ∎ is used, however. Besides the theorems given in §2, there are other results in the paper either inside or outside the proofs of the major theorems. The proofs of these other results are given immediately after the statements and their ends are denoted by □.

The methods that work for random partitions also work for random distinct-part partitions. The results are given in §9 and comments on their proofs are made in §10. Requiring partitions to have distinct parts places a particular restriction on the partitions. Section 11 describes how the methodology of this paper can be extended to treat partitions satisfying some other restrictions.

Theorems will be stated in terms of certain functions on the space Λ consisting of all partitions of nonnegative integers (including the empty partition of 0). For $\lambda \in \Lambda$ and k a positive integer, $X_k(\lambda)$ will denote the number of parts equal to k in the partition λ. For t a positive integer and $\lambda \in \Lambda$, $Y_t(\lambda)$ will denote the tth largest part in λ; if the number of parts of λ is less than t, then $Y_t(\lambda) = 0$. For each λ, $Y_1(\lambda) \geq Y_2(\lambda) \geq \cdots$, where equality is possible

$$X_5 = X_6 = \cdots = 0$$

$X_4 = 1$

$X_3 = 2$

$X_2 = 0$

$X_1 = 3$

$Y_1 = 4$

$Y_2 = 3$

$Y_3 = 3$

$Y_4 = 1$
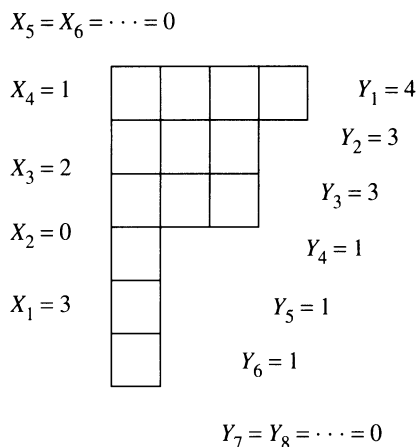
$Y_5 = 1$

$Y_6 = 1$

$$Y_7 = Y_8 = \cdots = 0$$

FIGURE 1.1. Ferrers diagram for the partition $4 + 3 + 3 + 1 + 1 + 1$ of 13.

since it is possible for some parts to be equal. Figure 1.1 uses what is called a Ferrers diagram to illustrate the partition $4 + 3 + 3 + 1 + 1 + 1$ of 13 and to identify the values of the functions X_k and Y_t for this partition. The symbols $|\lambda|$ and $N(\lambda)$, where λ is a partition, are alternative notations for the integer of which λ is a partition. Thus, if λ denotes the partition just described, then $N(\lambda) = |\lambda| = 13$.

We call the functions X_k, Y_t, and N random variables since they are functions on a space to which a probability measure has been attached. Actually, infinitely many probability measures have been described for Λ so when we speak, say, of the probability distribution of X_3 we must make clear which P_n is intended. Also, we use phrases such as *governed by* P_n and *when the underlying measure is* P_n in conjunction with probabilistic concepts such as variance and independence to indicate the relevant probability measure. When governed by P_n, the random variable N equals the constant n with probability 1. We will typically want to consider the sequence $(P_n : n = 1, 2, \ldots)$ of probability measures and with it, say, the corresponding sequence of probability distributions of X_3, and then analyze the limiting behavior as $n \to \infty$.

If we reflect the Ferrers diagram of a partition λ in a ray beginning at the upper left corner and slanting down and to the right at $45°$, we obtain a Ferrers diagram of a partition λ', called the dual of λ. It is worth noting that $Y_1(\lambda) = \sum_{k=1}^{\infty} X_k(\lambda')$. Since P_n is invariant under the bijection $\lambda \leftrightarrow \lambda'$, we conclude that the random variables Y_1 and $\sum_{k=1}^{\infty} X_k$ have the same probability distribution. Similarly, $Y_t - Y_{t+1}$ has the same probability distribution as X_t. More generally, results of the kind we will obtain have immediate corollaries that are consequences of duality. It will be left to the reader to formulate such corollaries.

Notice that the preceding paragraph does not apply to distinct-part partitions since the dual of a distinct-part partition is not necessarily a distinct-part partition.

The usual conventions that empty sums equal to 0 and empty products equal to 1 are in force. The symbol $O(a_n)$ or $O(b_q)$ will be used for a quantity that when divided by a_n or b_q, respectively, gives a quotient whose absolute value

remains bounded as $n \to \infty$ or $q \uparrow 1$. The symbol $o(a_n)$ or $o(b_q)$ will be used to indicate that the quotient approaches 0. We write $a_n \sim b_n$ to indicate that $a_n/b_n \to 1$ as $n \to \infty$. The symbols $\lfloor x \rfloor$ and $\lceil x \rceil$ denote, respectively, the floor and ceiling of the real number x, that is the largest integer no larger than x and the smallest integer no smaller than x. The ceiling of x equals the floor or is one larger than the floor according as x is or is not an integer.

2. THE MAJOR RESULTS. UNRESTRICTED PARTITIONS

Proofs for the results described here will be given in later sections.

As defined in §1, $X_k(\lambda)$ equals the number of parts of the partition λ that equal k, $k = 1, 2, 3, \ldots$. For each λ, $X_k(\lambda) = 0$ for all but finitely many k. The probability distribution of the sequence $(X_1, X_2, X_3, \ldots)$ of random variables depends on which of the underlying probability measures P_n is used. So, for example, when $n = 4$, the random variable X_2 equals $0, 1$, or 2 with probabilities $3/5$, $1/5$, and $1/5$, respectively, and the random sequence $(X_1, X_2, X_3, \ldots)$ equals one of $(4, 0, 0, 0, \ldots)$, $(2, 1, 0, 0, \ldots)$, $(0, 2, 0, 0, \ldots)$, $(1, 0, 1, 0, 0, \ldots)$, and $(0, 0, 0, 1, 0, 0, \ldots)$, each with probability $1/5$. The fact that, when governed by P_4, X_2 equals 0 with probability $3/5$ can be expressed precisely as

$$P_4(\{\lambda : X_2(\lambda) = 0\}) = 3/5,$$

but we would typically write $P_4(X_2 = 0) = 3/5$.

Our first result says that when $k X_k$ is multiplied by $\pi/(6n)^{1/2}$, the resulting random variable has a probability distribution close to the exponential distribution having expectation 1. Although the most interesting case is when k is a constant, the result also holds when k depends on n provided that k grows slowly compared to $n^{1/2}$.

Theorem 2.1. *If $k_n/n^{1/2} \to 0$ as $n \to \infty$, then, for each nonnegative real number v,*

$$\lim_{n \to \infty} P_n \left(\frac{\pi}{\sqrt{6n}} k_n X_{k_n} \le v \right) = 1 - e^{-v}.$$

Thus, the distribution function of the limiting probability distribution of X_{k_n}, appropriately normalized, is $1 - e^{-v}$. In preparation for the next theorem we examine a concept that gives another view of Theorem 2.1 and which generalizes in a natural manner. For a Borel set B let B^ε denote the set of points whose distance from B is less than ε. Consider two probability measures ζ and η on $\mathbb{R}$. The *Prohorov distance* between ζ and η is defined as

$$\inf\{\varepsilon > 0 : \zeta(B) \le \eta(B^\varepsilon) + \varepsilon \text{ for all Borel } B\}.$$

It can be shown that the collection of probability measures on $\mathbb{R}$ is a metric space with metric given by the Prohorov distance; thus, we speak of a *Prohorov metric*. In order to reformulate the conclusion of Theorem 2.1 we introduce Borel probability measures ξ_n and η on $\mathbb{R}$:

(2.1)
$$\xi_n(B) = P_n \left(\frac{\pi}{\sqrt{6n}} k_n X_{k_n} \in B \right);$$
$$\eta(B) = \int_B \left\{ \begin{matrix} e^{-v} & \text{if } v \ge 0 \\ 0 & \text{if } v < 0 \end{matrix} \right\} dv.$$

The conclusion of Theorem 2.1 is that the Prohorov distance between ξ_n and η approaches 0 as $n \to \infty$. The probability measure ξ_n is called the probability distribution of the random variable $(\pi/\sqrt{6n})k_n X_{k_n}$, when governed by P_n, or, alternatively, the probability measure induced by P_n via the random variable $(\pi/\sqrt{6n})k_n X_{k_n}$. Thus, the conclusion of Theorem 2.1 is that the Prohorov distance between η and the probability distribution of $(\pi/\sqrt{6n})k_n X_{k_n}$, when governed by P_n, approaches 0 as $n \to \infty$.

The theory described above carries over from $\mathbb{R}$ to general metric spaces. For our purposes the relevant metric spaces are $\mathbb{R}^d$ for various finite dimensions d with the l_1-metric—the distance between two members u and v of $\mathbb{R}^d$ equals $\sum_{i=1}^{d} |v_i - u_i|$. Suppose that $(d_n: n = 1, 2, \ldots)$ is a sequence of positive integers and that $(\xi_n: n = 1, 2, \ldots)$ and $(\nu_n: n = 1, 2, \ldots)$ are two sequences of probability measures such that, for each n, ξ_n, and ν_n are probability measures on $\mathbb{R}^{d_n}$. Since we have already established a metric for each $\mathbb{R}^d$, it is meaningful to ask whether the Prohorov distance between ξ_n and ν_n approaches 0 as $n \to \infty$. We will be giving affirmative answers to such questions for interesting measures ξ_n induced by P_n and measures ν_n that can be described quite explicitly.

One can ask about the joint probability distribution of, say, the number of 2's and the number of 5's in a random partition of a large integer. One feels that these two random variables should be more or less unrelated. In fact, as a consequence of the forthcoming Theorem 2.2, they are asymptotically independent:

$$\lim_{n \to \infty} P_n \left(\frac{\pi}{\sqrt{6n}} 2X_2 \le v_2, \frac{\pi}{\sqrt{6n}} 5X_5 \le v_5 \right) = (1 - e^{-v_2})(1 - e^{-v_5}).$$

The restriction on part size in Theorem 2.2 is more stringent than that in Theorem 2.1.

Theorem 2.2. *Suppose that $k_n/n^{1/4} \to 0$ as $n \to \infty$. For each n, let ξ_n be the probability measure on $\mathbb{R}^{k_n}$ induced by P_n via the random vector*

$$(\pi/\sqrt{6n})(1X_1, 2X_2, \ldots, k_n X_{k_n})$$

and let η_n denote the product measure on $\mathbb{R}^{k_n}$ of k_n copies of η defined by (2.1). Then, the Prohorov distance between ξ_n and η_n approaches 0 as $n \to \infty$.

The restrictions on the growth of k_n in the preceding two theorems indicate that these results only give information about the small parts of a random partition. We now turn to the large parts. As defined in §1, Y_t equals the tth largest part. The sequence $(Y_t: t = 1, 2, \ldots)$ is a sequence of random variables whose probability distribution depends on the underlying probability measure P_n. It is not difficult to show that, for each y, $P_n(Y_1 > y) \to 1$ as $n \to \infty$. Erdös and Lehner [3] did better by finding an appropriate normalization for Y_1 in order that there be a nontrivial limiting probability distribution. They showed that

$$(2.2) \qquad \lim_{n \to \infty} P_n \left(\frac{\pi}{\sqrt{6n}} Y_1 - \log \frac{\sqrt{6n}}{\pi} \le v \right) = e^{-e^{-v}}$$

for all real numbers v. According to the following generalization of their result, the same normalization works for Y_t. Of course, the limiting probability distribution depends on t.

Theorem 2.3. *As* $n \to \infty$,

$$P_n\left(\frac{\pi}{\sqrt{6n}}Y_t - \log\frac{\sqrt{6n}}{\pi} \le y\right) \to \int_{-\infty}^{y} \frac{\exp(-e^{-v} - tv)}{(t-1)!}\, dv.$$

To obtain the expectation of the limiting probability distribution in the preceding theorem we integrate the product of the variable v with the density (that is, the integrand) from $-\infty$ to ∞. For the second moment we use the product of v^2 with the density. The variance is obtained from the second moment by subtracting the square of the expectation. The result is that the expectation and variance equal

$$-\frac{\Gamma'(t)}{\Gamma(t)} = -\log t + o(t^{-1/2}) \qquad (t \to \infty),$$

and

$$\frac{\Gamma(t)\Gamma''(t) - [\Gamma'(t)]^2}{[\Gamma(t)]^2} \sim \frac{1}{t} \qquad (t \to \infty),$$

respectively, where Γ denotes the gamma function. (Notice that the variance goes to 0 as $t \to \infty$.) By subtracting the expectation from

$$\frac{\pi}{\sqrt{6n}}Y_t - \log\frac{\sqrt{6n}}{\pi}$$

and then dividing by the standard deviation (that is, the square root of the variance) and also making the corresponding changes in the limiting probability distribution we could rewrite the conclusion of Theorem 2.3 so that the limiting probability distribution would have expectation 0 and variance 1. There is no particular reason to do so for fixed t, but there is a reason to do so if the fixed t is replaced by t_n. The reason is that a probability distribution with a small variance like $1/t_n$ (for $t_n \to \infty$) is close to the probability distribution of a constant random variable. In such a situation we would like to spread out the probability distribution and identify a nontrivial probability distribution to which it is close. In the next theorem we do that and at the same time use the fact that when the limiting probability distribution in Theorem 2.3 is normalized to have expectation 0 and variance 1 it is, for large t, close to the normal distribution, that is, the probability distribution with density $(2\pi)^{-1/2}\exp(-v^2/2)$.

Theorem 2.4. *Let* $t_n \to \infty$ *as* $n \to \infty$ *sufficiently slowly that* $t_n/n^{1/4} \to 0$ *as* $n \to \infty$. *Then*

$$P_n\left(\sqrt{\frac{\pi^2 t_n}{6n}}Y_{t_n} - \sqrt{t_n}\log\frac{\sqrt{6n}}{\pi t_n} \le y\right) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-v^2/2}\, dv,$$

as $n \to \infty$.

Theorems 2.3 and 2.4 give information about individual large parts, but not about relations among different large parts. The next result gives such a relation in terms of a limiting Markov chain (the definition of which is not essential for reading this paper, as is also the case for the associated terms *initial probability distribution* and *transition density*) obtained when the normalization of Theorem 2.3 is used. The probability distribution in (2.2) is the appropriate initial probability distribution of the Markov chain; its density f_1 is given by

(2.3) $f_1(v) = \exp(-e^{-v} - v).$

It develops that the appropriate transition density f is given by

$$(2.4) \qquad f(u, v) = \begin{cases} \exp(e^{-u} - e^{-v} - v) & \text{if } v \leq u, \\ 0 & \text{if } v > u. \end{cases}$$

The density of the limiting probability distribution in the next result is a product of f_1 and $t - 1$ factors of f evaluated at different points.

Theorem 2.5. *Let f_1 and f be defined by (2.3) and (2.4). Then*

$$\lim_{n \to \infty} P_n \left(\frac{\pi}{\sqrt{6n}} Y_s - \log \frac{\sqrt{6n}}{\pi} \leq v_s, \; 1 \leq s \leq t \right)$$

$$= \int_{-\infty}^{v_1} \int_{-\infty}^{v_2} \cdots \int_{-\infty}^{v_t} f_1(u_1) \prod_{s=2}^{t} f(u_{s-1}, u_s) \, du_t \cdots du_2 \, du_1.$$

For understanding the limiting Markov chain the form

$$(2.5) \qquad f_1(v_1) \prod_{s=2}^{t} f(v_{s-1}, v_s),$$

used for the density in the preceding theorem is good, but, for calculational purposes, one can simplify the product (2.5) by combining the exponentials to obtain

$$\exp \left(-e^{-v_t} - \sum_{s=1}^{t} v_s \right) \quad \text{for } v_1 \geq v_2 \geq \cdots \geq v_t.$$

By integrating each of the variables $v_1, v_2, \ldots, v_{t-1}$ over all possible values, one obtains Theorem 2.3.

We would like to let t, of the preceding theorem, depend on n. It develops that we can do that provided that t does not grow too quickly compared to n. For the same reason that we had a normalization in Theorem 2.4 different from that in Theorem 2.3, we will use normalizations in Theorem 2.6 different from that used in Theorem 2.5.

Theorem 2.6. *Let $t_n \to \infty$ as $n \to \infty$ sufficiently slowly that $t_n/n^{1/4} \to 0$ as $n \to \infty$ and let f_1 and f be defined by (2.3) and (2.4). Then, as $n \to \infty$, the Prohorov distance between the probability measure having density*

$$\sqrt{t_n!} \exp \left(-e^{-w_{t_n}/\sqrt{t_n}} - \sum_{s=1}^{t_n} \frac{w_s}{\sqrt{s}} \right) \quad \text{for } \frac{w_1}{\sqrt{1}} - \log 1 \geq \cdots \geq \frac{w_{t_n}}{\sqrt{t_n}} - \log t_n,$$

and the probability measure induced by P_n via the random vector

$$\left(\sqrt{\frac{\pi^2 s}{6n}} Y_s - \sqrt{s} \log \frac{\sqrt{6n}}{\pi s} : 1 \leq s \leq t_n \right)$$

approaches 0.

An immediate consequence is the following corollary which says that, with high probability, the t_n largest parts are distinct.

Corollary. *For t_n as in the preceding theorem,*

$$\lim_{n \to \infty} P_n(Y_1 > Y_2 > \cdots > Y_{t_n}) = 1.$$

It may not be that Y_1 constitutes the largest contribution to n by a single part size; because some smaller part may occur with sufficient multiplicity that its product with its multiplicity is larger than Y_1. We will see that this possibility is typical. Let Z_1 denote the largest value of kX_k as k ranges over the positive integers. Erdös and Szalay [4, Theorem 1] found an appropriate normalization for Z_1 in order that there be a nontrivial limiting probability distribution. They showed that

$$(2.6) \qquad \lim_{n \to \infty} P_n \left(\frac{\pi}{\sqrt{6n}} Z_1 - \log \frac{\sqrt{6n}}{\pi} - \log\log\log n \le v \right) = e^{-e^{-v}}.$$

The following theorem is a generalization.

Theorem 2.7. *Let Z_t denote the tth largest member of the sequence $(kX_k : k = 1, 2, \ldots)$. Then*

$$\lim_{n \to \infty} P_n \left(\frac{\pi}{\sqrt{6n}} Z_t - \log \frac{\sqrt{6n}}{\pi} - \log\log\log n \le v \right) = \int_{-\infty}^{v} \frac{\exp(-e^{-u} - tu)}{(t-1)!} \, du.$$

The next theorem treats the largest t values of Z_s.

Theorem 2.8. *Let Z_s be as in the preceding theorem. Let f_1 and f be defined by (2.3) and (2.4). For fixed t,*

$$\lim_{n \to \infty} P_n \left(\frac{\pi}{\sqrt{6n}} Z_s - \log \frac{\sqrt{6n}}{\pi} - \log\log\log n \le v_s, \ 1 \le s \le t \right)$$

$$= \int_{-\infty}^{v_1} \int_{-\infty}^{v_2} \cdots \int_{-\infty}^{v_t} f_1(u_1) \prod_{s=2}^{t} f(u_{s-1}, u_s) \, du_t \cdots du_2 \, du_1.$$

Just as Theorem 2.3 is a corollary of Theorem 2.5, so is Theorem 2.7 a corollary of Theorem 2.8.

In §3 we will see how to use Theorem 2.5 to study the number of parts larger than some very large k_n. The next theorem here focuses on the number of parts larger than some moderately large k_n. It gives the asymptotic probability distribution of the number of parts, appropriately normalized, that are larger than a k_n that grows slightly faster than $n^{1/2}$.

Theorem 2.9. *Suppose that $k_n / n^{1/2} \to \infty$ and*

$$\frac{\pi k_n}{\sqrt{6n}} - \frac{\log n}{2} \to -\infty$$

as $n \to \infty$. Then

$$P_n \left(\frac{\sum_{k \ge k_n} X_k - \pi^{-1}(6n)^{1/2} \log(1 - e^{-\pi k_n / \sqrt{6n}})^{-1}}{\pi^{-1/2}(6n)^{1/4} e^{-\pi k_n / 2\sqrt{6n}}} \le x \right) \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} \, du$$

as $n \to \infty$.

From the preceding limit theorem one can easily obtain a law of large numbers. For it one can replace $\log(1 - e^{-\pi k_n \sqrt{6n}})^{-1}$ by $e^{-\pi k_n / \sqrt{6n}}$.

Corollary. *Let k_n be as in Theorem 2.9. Then, for every $\varepsilon > 0$,*

$$P_n \left(\left| \frac{\sum_{k \geq k_n} X_k}{\pi^{-1} \sqrt{6n} e^{-\pi k_n / \sqrt{6n}}} - 1 \right| > \varepsilon \right) \to 0$$

as $n \to \infty$.

Theorem 2.2 implies that if $k_n / n^{1/4} \to 0$, then, for n large, the probability distribution of $\pi(X_1 + X_2 + \cdots + X_{k_n}) / \sqrt{6n}$ is close to that of the sum of independent exponentially distributed random variables having means $1, 1/2, \ldots, 1/k_n$. It is straightforward to calculate the convolution of these exponential distributions (a calculation which also happens to be relevant for the study of the Yule stochastic process) in order to obtain the probability distribution of the sum. The result is that

$$P_n \left(\frac{\pi}{\sqrt{6n}} (X_1 + X_2 + \cdots + X_{k_n}) \leq x \right) - (1 - e^{-x})^{k_n} \to 0$$

uniformly in x as $n \to \infty$. By replacing x by $x + \log k_n$, we conclude that

$$P_n \left(\frac{\pi}{\sqrt{6n}} (X_1 + X_2 + \cdots + X_{k_n}) - \log k_n \leq x \right) - \left(1 - \frac{e^{-x}}{k_n} \right)^{k_n} \to 0.$$

Thus, for fixed k_n, $(1 - e^{-x}/k_n)^{k_n}$ is the limiting distribution function of the number of parts no larger than k_n, normalized by multiplication by $\pi/\sqrt{6n}$ and subtraction of $\log k_n$.

If, in addition to our assumption that $k_n / n^{1/4} \to 0$, we also assume that $k_n \to \infty$, we conclude that

$$P_n \left(\frac{\pi}{\sqrt{6n}} (X_1 + X_2 + \cdots + X_{k_n}) - \log k_n \leq x \right) - e^{-e^{-x}} \to 0$$

uniformly in x. Thus, the limiting distribution function in this case is $e^{-e^{-x}}$. The next theorem says that this conclusion is valid for somewhat larger k_n.

Theorem 2.10. *Suppose that $k_n / n^{1/2} \to 0$ and $k_n \to \infty$ as $n \to \infty$. Then*

$$P_n \left(\frac{\pi}{\sqrt{6n}} (X_1 + X_2 + \cdots + X_{k_n}) - \log k_n \leq x \right) \to e^{-e^{-x}}$$

as $n \to \infty$.

From this limit theorem we get the following law of large numbers.

Corollary. *Let k_n be as Theorem 2.10. Then, for every $\varepsilon > 0$,*

$$P_n \left(\left| \frac{\sum_{k \leq k_n} X_k}{(6n)^{1/2} (\log k_n)/\pi} - 1 \right| > \varepsilon \right) \to 0$$

as $n \to \infty$.

As mentioned earlier, the special case of Theorem 2.3 obtained by setting $t = 1$ was treated by Erdös and Lehner [3], but in their paper they actually state the dual result by describing the probability distribution of the number of parts in a random partition of a large integer. Their result (alternatively Theorem 2.3) tells us that the number of parts is, with high probability, close to $\sqrt{6n} \log n / 2\pi$. In case $k_n = n^{\tau/2}$ for some $\tau \in (0, 1)$, the preceding corollary

says that, with high probability, $\sum_{k \leq k_n} X_k$ is close to $\tau \sqrt{6n} \log n / 2\pi$. Hence, we see that the proportion of the total number of parts that are smaller than $n^{\tau/2}$ is approximately τ.

Neither Theorem 2.9 nor Theorem 2.10 treats the number of parts larger or smaller than a fixed multiple of $n^{1/2}$. The methods of this paper do not work in a straightforward manner for this situation.

The theorems described above all have a similar form. For each theorem there is, for each n, a function W_n from Λ, the set of all partitions, to some $\mathbb{R}^{d_n}$. Then P_n, which we regard as a probability measure on Λ, induces, via W_n, a probability measure ξ_n on $\mathbb{R}^{d_n}$; ξ_n is the probability distribution of the random vector W_n. Also, a probability measure ν_n on $\mathbb{R}^{d_n}$ is explicitly described. The conclusion of each theorem is the assertion that the Prohorov distance between ξ_n and ν_n goes to 0 as $n \to \infty$. The scheme for obtaining such conclusions will be described in §4. The specific details for particular theorems are given in later sections. But first we examine, in §3, some consequences of the theorems in this section.

3. CONSEQUENCES OF SOME THEOREMS

Since $\exp(-e^{-x}) \geq 1 - e^{-x}$, we conclude from Theorem 2.10 that

$$\lim_{n \to \infty} P_n \left(-\log A \leq \frac{\pi}{\sqrt{6n}} \sum_{k=1}^{k_n} X_k - \log k_n \leq A \right) \geq 1 - 2e^{-A}$$

for $A \geq \log 2$, whenever $k_n \to \infty$ sufficiently slowly that $k_n / n^{1/2} \to 0$. This assertion is very similar to that obtained by Szalay and Turán in [6, Theorem III]. Here are the differences. They do not use a limit, but rather assert the truth of an inequality for all sufficiently large n. They have $A(n)$ in place of A and let, with some restriction, $A(n) \to \infty$; whereas our argument depends on A being fixed. They require that k_n approach ∞ no faster than $13 \log n$ and have $1 - 8e^{-A}$, rather than $1 - 2e^{-A}$, on the right-hand side of their inequality.

Since the probability distribution of the largest part equals that of the total number of parts, Theorem 2.3 tells us that the distribution function $\exp(-e^{-x})$, obtained in Theorem 2.10 as the limiting distribution function of the appropriately normalized number of parts no larger than k_n, also arises as the limiting distribution function of the total number of parts appropriately normalized. Moreover, the two normalizations differ in an additive manner only. Thus, we conclude that the limiting variance of the total number of parts appropriately normalized is essentially due to the variance of the number of small parts. ($\pi^2/6$ equals the variance of the probability distribution whose distribution function is $\exp(-e^{-x})$, and Euler's constant (approximately 0.577) equals the expectation.)

From the corollary of Theorem 2.10 and the fact that the total number of parts can be approximated by $(6n)^{1/2}(\log n)/(2\pi)$ with probability approaching 1, we conclude that

$$P_n \left(\left| \frac{\sum_{k \geq n^{\tau/2}} X_k}{(6n)^{1/2}(\log n)/2\pi} - (1 - \tau) \right| > \varepsilon \right) \to 0$$

as $n \to \infty$. From Theorem 2.9 we get the same conclusion if ε is replaced by ε_n with $\varepsilon_n \to 0$ sufficiently slowly that $\varepsilon_n \log n \to \infty$. But this improved conclusion is not as sharp as that obtained at the beginning of §6 of [7] by

Szalay and Turán. The preceding paragraph describes why Theorem 2.9 cannot be expected to give a sharp result in this situation.

From Theorem 2.5 we conclude, for each v, that, as $n \to \infty$,

$$P_n \left(Y_t > \frac{\sqrt{6n}}{\pi} \left[\log \frac{\sqrt{6n}}{\pi} + v \right] \geq Y_{t+1} \right) \to \frac{(-e^{-v})^t}{t!} e^{-(e^{-v})},$$

with the convention that $Y_0 = \infty$. That is, the number of parts greater than

$$\frac{\sqrt{6n}}{\pi} \left(\log \frac{\sqrt{6n}}{\pi} + v \right)$$

has a limiting Poisson distribution with expectation e^{-v}.

Letting

$$k_n = \frac{\sqrt{6n}}{\pi} \left(\log \frac{\sqrt{6n}}{\pi} + v \right),$$

we see that $\pi k_n / \sqrt{6n} - \log n / 2 \to$ constant as $n \to \infty$. Thus the Poisson limit result mentioned in the preceding paragraph complements Theorem 2.9 where a limiting normal distribution is obtained in case $\pi k_n / \sqrt{6n} - \log n / 2 \to -\infty$.

In Theorems I and II of [6], Szalay and Turán treat the issues similar to those addressed by Theorems 2.9 and 2.10 of this paper. Here is a comparison of their conclusions for a specific example to which Theorem 2.9 also applies. They prove that, with probability approaching one as fast as $1 - c/n^{7/4}$ for some constant c, the number of parts greater than $\sqrt{6n}[\log n - 6 \log \log n]/2\pi$ equals

$$(3.1) \qquad \frac{\sqrt{6}}{\pi} \log^3 n + O(\log^2 n).$$

In this paper I have not obtained an estimate on the speed of convergence of probabilities. On the other hand, Theorem 2.9 gives a sharper conclusion than the bound in (3.1). It identifies the probability distribution as approximately normal with expectation $(\sqrt{6}/\pi) \log^3 n$ and standard deviation $[(\sqrt{6}/\pi) \log^3 n]^{1/2}$.

The papers [7, 8] of Szalay and Turán also apply for the number of parts greater than a multiple of $n^{1/2}$, whereas none of the theorems in §2 do. On the other hand, the hypotheses in Theorems 2.9 and 2.10 are less stringent at the extremes than are theirs.

4. UNRESTRICTED PARTITIONS OF A RANDOM INTEGER

Let $q \in (0, 1)$. For any partition λ let

$$(4.1) \qquad Q_q(\lambda) = q^{|\lambda|} \prod_{k=1}^{\infty} (1 - q^k),$$

where, as mentioned earlier, $|\lambda|$ denotes the nonnegative integer of which λ is a partition. If we sum $Q_q(\lambda)$ over all λ for which $|\lambda| = n$ we obtain $p(n)q^n \prod_{k=1}^{\infty}(1 - q^k)$. Then, if we sum over n we obtain 1, since, as mentioned at (1.1), $\prod(1 - q^k)^{-1}$ is the generating function of the sequence $(p(n)$:

$n = 0, 1, \dots$). Thus, we can view Q_q as the probability measure for an experiment in which a partition is chosen at random and in which the integer $N = \sum_{k=1}^{\infty} k X_k$ being partitioned is itself random. It will develop that, for the study of P_n there is a particularly useful Q_q—the one given by $q = q_n$, where

$$(4.2) \qquad q_n = \exp(-\pi/\sqrt{6n}).$$

Fix a sequence $(W_n : n = 1, 2, \dots)$ of random vectors as described in the last paragraph of the preceding section. The probability measure Q_{q_n} induces, via W_n, a probability measure ζ_n on $\mathbb{R}^{d_n}$. Our method of achieving the goal, mentioned at the end of §2, of proving that the Prohorov distance between ξ_n and ν_n approaches 0 will be to prove that both the Prohorov distance between ζ_n and ν_n and the Prohorov distance between ξ_n and ζ_n approach 0 as $n \to \infty$.

One reason this program can be successful is that the probability measures Q_q are easily understood. When Q_q is the underlying probability measure, the numbers of parts of various sizes are independent random variables. Here is the precise statement.

Proposition 4.1. *When governed by Q_q, the random variables X_k are independent and the probability distribution of X_k is geometric with mean $q^k/(1 - q^k)$. That is, for any sequence of nonnegative integers $(x_k : k = 1, 2, \dots)$,*

$$(4.3) \qquad Q_q(X_k = x_k \text{ for } k = 1, 2, \dots) = \prod_{k=1}^{\infty} (1 - q^k)(q^k)^{x_k}.$$

Proof. Insertion of $\lambda = \sum k x_k$ into (4.1) gives the result. $\square$

Remarks. The infinite product in the preceding proposition is 0 if and only if infinitely many of the x_k are positive. Probability distributions for a random vector composed of a finite number of the X_k can be obtained by summing (4.3). If, for instance, one sums over all values of $x_2, x_4, x_5, x_6, x_7, \dots$ one obtains the probability distribution of the random vector (X_1, X_3)

$$Q_q(X_1 = x_1 \text{ and } X_3 = x_3) = (1 - q)q^{x_1}(1 - q^3)q^{3x_3}.$$

In subsequent sections, Proposition 4.1 will be used to prove that the Prohorov distance between ζ_n and ν_n approaches 0 for an appropriate choice of the sequence $(\nu_n : n = 1, 2, \dots)$, a choice depending on which sequence $(W_n : n = 1, 2, \dots)$ of random variables is being studied. Now we turn to the second issue raised earlier in this section. Roughly speaking, we ask: Why should there be a close connection between the probability measures P_n and Q_{q_n}? More precisely, we look for a condition under which the Prohorov distance between ξ_n and ζ_n approaches 0. Such a condition will be given in Lemma 4.6. Skipping from here to the statement of that lemma and then to the last paragraph of this section is a reasonable plan for a first reading of this paper.

When governed by Q_q, the random variable N, which equals the number being partitioned, is not a constant. Its probability distribution can be obtained either from (4.1) by summing over all partitions λ for which $|\lambda|$ equals an arbitrary constant, say n, or from (4.3) by summing over all sequences $(x_k : k = 1, 2, \dots)$ for which $\sum k x_n = n$

$$(4.4) \qquad Q_q(N = n) = p(n)q^n \prod (1 - q^k).$$

For any two partitions λ_1 and λ_2 for which $N(\lambda_1) = n = N(\lambda_2)$, we have $P_n(\lambda_1) = P_n(\lambda_2)$ and $Q_q(\lambda_1) = Q_q(\lambda_2)$. This fact and the equality $P_n(N = n) = 1$ imply that for each n and q, the probability measure P_n is equal to the conditional probability measure obtained by conditioning Q_q by the event $\{\lambda : N(\lambda) = n\}$. That is, for any set A of partitions

$$P_n(A) = Q_q(A|N = n) = \frac{Q_q(A \cap \{\lambda : N(\lambda) = n\})}{Q_q(\{\lambda : N(\lambda) = n\})}.$$

Thus, the Prohorov distance between ξ_n and ζ_n is bounded above by
(4.5)
$$\sup\left\{Q_{q_n}(W_n^{-1}(B)) - \frac{Q_{q_n}(W_n^{-1}(B) \cap \{\lambda : N(\lambda) = n\})}{Q_{q_n}(N = n)} : B \text{ Borel, } B \subseteq \mathbb{R}^{d_n}\right\}.$$

This quantity is well defined even if $d_n = \infty$ and W_n is $\mathbb{R}^\infty$-valued.

The next result encompasses this generalization.

Lemma 4.2. *For each n, let q_n be defined by (4.2), let d_n be a positive integer or ∞, let W_n be an $\mathbb{R}^{d_n}$-valued function on the space Λ of all partitions, and let B_n be a Borel subset of $\mathbb{R}^{d_n}$. Suppose that, as $n \to \infty$, $Q_{q_n}(W_n^{-1}(B_n)) \to 1$ and*

(4.6)
$$\frac{Q_{q_n}(N = n|W_n = w_n)}{Q_{q_n}(N = n)} \to 1$$

uniformly for $w_n \in B_n$. Then the sequence defined at (4.5) converges to 0 as $n \to \infty$.

Proof. For any Borel subset B of $\mathbb{R}^{d_n}$ the quantity in (4.5) is bounded above by

$$Q_{q_n}(W^{-1}(B - B_n))$$
$$+ \sum_{w_n \in B \cap B_n} \left(Q_{q_n}(W_n^{-1}(w_n)) - \frac{Q_{q_n}(W_n^{-1}(w_n) \cap \{\lambda : N(\lambda) = n\})}{Q_{q_n}(N = n)}\right)$$
$$= \zeta_n(B - B_n) + \sum_{w_n \in B \cap B_n} \zeta_n(w_n)\left(1 - \frac{Q_{q_n}(N = n|W_n = w_n)}{Q_{q_n}(N = n)}\right)$$
$$\leq \zeta_n(\mathbb{R}^{d_n} - B_n) + \sum_{w_n \in B_n} \zeta_n(w_n)\left|1 - \frac{Q_{q_n}(N = n|W_n = w_n)}{Q_{q_n}(N = n)}\right|,$$

which does not depend on B and approaches 0 as $n \to \infty$. □

With a view to applying the preceding lemma we give attention to the denominator in (4.6). As a function of r, the generating function of N, when governed by Q_q, equals

$$\sum_{n=0}^\infty Q_q(N = n)r^n = \prod_{k=1}^\infty (1 - q^k) \sum_{n=0}^\infty p(n)q^n r^n = \prod_{k=1}^\infty \frac{1 - q^k}{1 - (qr)^k},$$

where we have both used the formula (4.4) for $Q_q(N = n)$ and replaced q by qr in the formula (1.1) for the generating function of the sequence $(p(n): n = 0, 1, \ldots)$. The expected value of N, when governed by Q_q, can be obtained either as the value of the derivative of the generating function at $r = 1$ or as the sum over k of the expectations of the random variables kX_k. The variance

of N is the sum of the variances of the random variables kX_k; alternatively, we can evaluate the second derivative at 1 and then add the expected value and subtract the square of the expected value. Here is a summary of the results.

Proposition 4.3. *Governed by* Q_q, *the random variable* N *has generating function given by the function of* r, $\prod_{k=1}^{\infty}(1-q^k)/(1-(qr)^k)$, *expectation equal to* $\sum_{k=1}^{\infty} kq^k/(1-q^k)$, *and variance equal to* $\sum_{k=1}^{\infty} k^2 q^k/[1-q^k]^2$.

Corollary 4.4. *Let* q_n *be defined by* (4.2). *As* $n \to \infty$, *the variance of* N, *when governed by* Q_{q_n}, *is asymptotic to* $(24^{1/2}/\pi)n^{3/2}$ *and the difference between* n *and the expected value of* N *equals* $o(n^{3/4})$.

Proof. We first consider general Q_q and let $q \uparrow 1$. Then, as a last step, we convert to a result for $n \to \infty$ by using (4.2). The variance given in Proposition 4.3 is an approximating sum for the integral

$$\log^{-3}(1/q) \int_0^\infty \frac{u^2 e^{-u}}{(1-e^{-u})^2}\, du,$$

which, by [5, Formula 3.423-3], equals $\pi^2/[3\log^3(1/q)]$. A similar argument for the expected value introduces an error of order $1/\log(1/q)$ when the sum is replaced by the integral

$$\log^{-2}(1/q) \int_0^\infty \frac{u e^{-u}}{1-e^{-u}}\, du,$$

which, by [5, Formula 3.411-7], equals $\pi^2/[6\log^2(1/q)]$. An appeal to (4.2) completes the proof. □

   As the reader has probably guessed, a portion of the preceding proof was constructed before the choice for q_n, given at (4.2), was made. The choice was then made so that the expected value of N, when governed by Q_{q_n}, would be asymptotic to n as $n \to \infty$.

   Let μ_n and σ_n, respectively, denote the expectation and standard deviation of the random variable N when governed by Q_{q_n}. In the proof of the following proposition we will show that the probability distribution of $(N-\mu_n)/\sigma_n$, when governed by Q_{q_n}, approaches the normal distribution as $n \to \infty$, that is, that, for each real v,

$$(4.7) \qquad Q_{q_n}\left(\frac{N-\mu_n}{\sigma_n} \le v\right) \to \int_{-\infty}^{v} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du$$

as $n \to \infty$. We want more—an asymptotic formula for

$$Q_{q_n}(N=n) = Q_{q_n}((N-\mu_n)/\sigma_n = (n-\mu_n)/\sigma_n).$$

Notice that, as a consequence of Corollary 4.4, $(n-\mu_n)/\sigma_n \to 0$; so, we might reasonably hope that an asymptotic formula for $Q_{q_n}(N=n)$ can be obtained by multiplying the value $1/\sqrt{2\pi}$ of the normal density at 0 by $1/\sigma_n$, the distance between possible values of $(N-\mu_n)/\sigma_n$. In view of Corollary 4.4, this multiplication would give $(96n^3)^{-1/4}$. Here is the desired result.

Lemma 4.5. *Let* q_n *be given by* (4.2). *Then, as* $n \to \infty$,

$$Q_{q_n}(N=n) \sim \frac{1}{\sqrt[4]{96n^3}}.$$

Proof. The proof has two parts. In the first we use characteristic functions (that is, Fourier-Stieltjes transforms of probability distributions) to prove (4.7). Then, in the second part we verify conditions that enable us to go, as described in the paragraph preceding the statement of the lemma, from the global result (4.7) to a local result. This local result at 0 is our goal. Let φ_n denote the characteristic function of the random variable $(N - \mu_n)/\sigma_n$, when governed by Q_{q_n}.

Part 1. The characteristic function of the normal distribution is

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} e^{i\theta x}\, dx = e^{-\theta^2/2}\,.$$

Thus, we want to show that $\log \varphi_n(\theta)$ converges pointwise to $-\theta^2/2$ as $n \to \infty$. From proposition 4.3 and a change of variables we see that

$$(4.8) \qquad \log \varphi_n(\theta) = -\frac{i\theta\mu_n}{\sigma_n} - \sum_{k=1}^{\infty} \log\left(1 + \frac{q_n^k(1 - e^{i\theta k/\sigma_n})}{1 - q_n^k}\right)\,.$$

We approximate $\log(1 + z)$ by $z - z^2/2$. The error we make in doing so is bounded by a constant multiple of

$$\frac{|\theta|^3}{\sigma_n^3} \sum_{k=1}^{\infty} \frac{q_n^{k3}k^3}{(1 - q_n^k)^3} \sim \frac{|\theta|^3}{\sigma_n^3 \log^4(1/q_n)} \int_0^{\infty} \frac{e^{-3u}u^3}{(1 - e^{-u})^3}\, du \to 0\,.$$

In each of the terms z and $-z^2/2$ the factor $1 - \exp(i\theta k/\sigma_n)$ appears to either the first or second power. We also plan to use a power series for it. An argument similar to the one just made for the logarithm shows that only the terms involving θ and θ^2 need be considered. Thus, in lieu of (4.8) we consider

$$(4.9) \qquad -\frac{i\theta\mu_n}{\sigma_n} + \sum_{k=1}^{\infty} \frac{q_n^k i\theta k}{(1 - q_n^k)\sigma_n} - \sum_{k=1}^{\infty} \frac{q_n^k \theta^2 k^2}{(1 - q_n^k)2\sigma_n^2} - \sum_{k=1}^{\infty} \frac{q_n^{2k} \theta^2 k^2}{(1 - q_n^k)^2 2\sigma_n^2}\,.$$

By Proposition 4.3 the sum of the first two terms equals 0. The sum of the last two terms is asymptotic to

$$-\frac{\theta^2}{2\sigma_n^2 \log^3(1/q_n)} \left(\int_0^{\infty} \frac{e^{-u}u^2}{1 - e^{-u}}\, du + \int_0^{\infty} \frac{e^{-2u}u^2}{(1 - e^{-u})^2}\, du\right)$$

$$\sim -\frac{3\theta^2}{2\pi^2} \int_0^{\infty} \frac{e^{-u}u^2}{(1 - e^{-u})^2}\, du\,,$$

which, by [5, Formula 3.423-3], equals $-\theta^2/2$.

Part 2. Chaganty and Sethuraman [2, Theorem 4.2] have shown that we can obtain the desired local limit assertion from the convergence in distribution established in Part 1 provided that we establish the existence of a function ρ and an integrable function φ^* such that $\rho(n) \to \infty$,

$$(4.10) \qquad \sup_{\rho(n) \le |\theta| \le \pi\sigma_n} |\varphi_n(\theta)| = o(1/\sigma_n)\,,$$

and

$$(4.11) \qquad \sup_n |\varphi_n(\theta)| I[|\theta| < \rho(n)] \le \varphi^*(\theta)\,,$$

where I denotes an indicator function. We take $\rho(n) = \pi\sigma_n^{1/3}$ and plan to choose φ^* as the proof develops. A straightforward calculation based on Proposition 4.3 gives

$$(4.12) \qquad \log|\varphi_n(\theta)| = -\frac{1}{2}\sum_{k=1}^{\infty}\log\left(1 + \frac{2q_n^k(1 - \cos(\theta k/\sigma_n))}{(1 - q_n^k)^2}\right).$$

Since all terms in (4.12) have the same sign we may throw away any subset of them as we look for upper bounds. Also, we may replace the denominators $(1 - q_n^k)^2$ by 1.

In order to prove (4.11) we consider θ satisfying $|\theta| < \rho(n)$ and restrict k in (4.12) to those k for which $\sigma_n^{2/3}/2 \le k \le \sigma_n^{2/3}$. By using the first nonzero terms in the power series for $1 - \cos$ and $\log$ we conclude that there are absolute constants b and c such that

$$\log|\varphi_n(\theta)| \le -\frac{1}{2}\sum_{k=\lceil\sigma_n^{2/3}/2\rceil}^{\lfloor\sigma_n^{2/3}\rfloor}\log(1 + b\sigma_n^{-2/3}\theta^2) \le -c\theta^2 \, ;$$

so we take $\varphi^*(\theta) = \exp(-c\theta^2)$ to satisfy (4.11).

To prove (4.10) we restrict k in (4.12) to $k \le \sigma_n^{2/3}$, and, further, to those k's that make the cosine negative so that we may replace it by 0 when obtaining an upper bound for (4.12). For $|\theta| \ge \rho(n)$, the number of such k's is greater than $(1/3)\sigma_n^{2/3}$ so we get, using the relationships among n, q_n, and σ_n,

$$\log|\varphi_n(\theta)| \le -\frac{1}{8}\sigma_n^{2/3}\log(1 + 2e^{-(\pi^2/3)^{1/3}})$$

for sufficiently large n. The inequality (4.10) follows. $\square$

The preceding lemma gives an asymptotic formula for the denominator in Lemma 4.2. In the proof of the following lemma we obtain, under an appropriate condition, the same asymptotic formula for the numerator in Lemma 4.2. This forthcoming lemma is the tool we will use in subsequent sections to prove that the Prohorov distance between ξ_n and ζ_n approaches 0 as $n \to \infty$.

Lemma 4.6. *For each n let d_n be a positive integer, W_n be a function from the space Λ of all partitions into $\mathbb{R}^{d_n}$, and K_n be a set of positive integers having the property that the values of the functions X_k, $k \in K_n$, determine the value of W_n. Let ξ_n and ζ_n denote the probability distributions of W_n when governed by P_n and Q_{q_n}, respectively. Suppose that*

$$(4.13) \qquad\qquad \sum_{k \in K_n}\frac{k^2 q_n^k}{(1 - q_n^k)^2} = o(n^{3/2})$$

as $n \to \infty$. Then the Prohorov distance between ξ_n and ζ_n converges to 0 as $n \to \infty$.

Proof. Suppose that (4.13) holds. The goal is to show that the sequence defined at (4.5) converges to 0; for then it will follow that the Prohorov distance between ξ_n and ζ_n approaches 0, since (4.5) is an upper bound for this Prohorov distance. It suffices to prove (4.5) in the particular case that $W_n = (X_k : k \in K_n)$.

The plan is to apply Lemma 4.2 to $W_n = (X_k : k \in K_n)$. Lemma 4.5 gives an asymptotic formula for the denominator of (4.6) (in Lemma 4.2). It remains for us to obtain the same asymptotic relation for the numerator of (4.6), uniformly for w_n belonging to appropriate Borel sets B_n for which $Q_{q_n}(W_n^{-1}(B_n)) \to 1$.

By (4.13) there exists a sequence $(b_n : n = 1, 2, \dots)$ such that

$$(4.14) \qquad \sum_{k \in K_n} \frac{k^2 q_n^k}{(1 - q_n^k)^2} = o(b_n^2),$$

and $b_n = o(n^{3/4})$. Let

$$B_n = \left\{ w_n = (x_{k,n} : k \in K_n) : \left| \sum_{k \in K_n} \frac{k q_n^k}{1 - q_n^k} - \sum_{k \in K_n} k x_{k,n} \right| \leq b_n \right\}.$$

Using the fact that sums of expectations equal expectations of sums, as described in the discussion preceding Proposition 4.3, we see that the expectation of $\sum_{k \in K_n} k X_k$, when governed by Q_{q_n}, equals

$$\sum_{k \in K_n} \frac{k q_n^k}{1 - q_n^k}.$$

Therefore, by Chebyshev's Inequality, the probability, under Q_{q_n}, that W_n belongs to the complement of B_n is bounded above by the variance of $\sum_{k \in K_n} k X_k$ divided by the square of b_n. The argument leading to Proposition 4.3 shows that the left side of (4.13) is the variance of $\sum_{k \in K_n} k X_k$, when governed by Q_{q_n}. Thus,

$$Q_{q_n}(W_n^{-1}(\mathbb{R}^{\mathrm{card}(K_n)} - B_n)) \leq b_n^{-2} \sum_{k \in K_n} \frac{k^2 q_n^k}{[1 - q_n^k]^2},$$

which, by (4.14), approaches 0 as $n \to \infty$.

Consider an arbitrary sequence $(w_n : n = 1, 2, \dots)$ where $w_n = (x_{k,n} : k \in K_n) \in B_n$ for each n. We can finish the proof by showing that (4.6) holds for this sequence. Because the random variables X_k are independent, we can remove the conditioning in the numerator in (4.6) by making a simple algebraic adjustment:

$$(4.15)$$

$$Q_{q_n}(N = n | W_n = (x_{k,n} : k \in K_n)) = Q_{q_n}\left(\sum_{k \notin K_n} k X_k = n - \sum_{k \in K_n} k x_{k,n} \right).$$

To find an asymptotic formula for the right side of (4.15) we will mimic Corollary 4.4 and the proof of Lemma 4.5. The left side of (4.13) is the difference between the variance of $\sum_k k X_k$ and the variance of $\sum_{k \notin K_n} k X_k$. Hence, by (4.13) and Corollary 4.4, the variances of these two random variables are asymptotic to each other. The difference between $n - \sum_{k \in K_n} k x_{k,n}$ and the mean of $\sum_{k \notin K_n} k X_k$ equals

$$(4.16) \qquad \left(n - \sum_{k=1}^{\infty} \frac{k q_n^k}{1 - q_n^k} \right) + \left(\sum_{k \in K_n} \frac{k q_n^k}{1 - q_n^k} - \sum_{k \in K_n} k x_{k,n} \right).$$

As in Corollary 4.4 we are interested in whether this sequence is $o(n^{3/4})$. By Corollary 4.4 the sequence of first terms has this property. In view of the definition of B_n and the fact that $b_n = o(n^{3/4})$, the sequence of second terms also has this property. Thus, if we can mimic the proof of Lemma 4.5, then we will be able to conclude that (4.6) holds, as desired.

When we try to mimic the proof of Lemma 4.5 we obtain a relation similar to that at (4.8), with some terms missing from the sum in (4.8) and φ_n and μ_n having modified meanings. When we get to a modified (4.9) the first two terms cancel each other as before and (4.13) shows that the missing summands in the last two summations do not contribute to a change in the asymptotic formulas. In Part 2 of the proof of Lemma 4.5, bounds, valid for certain $k \le \sigma_n^{2/3}$, were obtained, and then these bounds were multiplied by the number of terms, a multiple of $\sigma_n^{2/3}$. For the current setting the number of such terms is somewhat smaller, but, as a consequence of (4.13), still a multiple of $\sigma_n^{2/3}$. $\quad\square$

In the following four sections we prove the theorems stated in §2. Proposition 4.1 will be used to obtain the appropriate result for Q_{q_n} and then Lemma 4.6 will be used to convert that result into the desired result for P_n.

5. THE SMALL PARTS. PROOFS OF 2.1 AND 2.2

Only Theorem 2.2 will be proved here and the slightly easier argument for Theorem 2.1 will be left for the reader.

Consider a family $(v_{k,n}: 1 \le k \le k_n, 1 \le n)$ of nonnegative numbers satisfying the additional condition that $v_{k,n}$ is an integral multiple of $k\pi/\sqrt{6n}$ for each k and n. From Proposition 4.1 we see that

$$Q_{q_n}\left(\frac{\pi}{\sqrt{6n}}(1X_1, \ldots, k_nX_{k_n}) = (v_{1,n}, \ldots, v_{k_n,n})\right)$$
$$= \left(\prod_{k=1}^{k_n} e^{-v_{k,n}}\right)\left(\prod_{k=1}^{k_n}(1 - q_n^k)\right).$$

Using the definition (4.2) of q_n, we bound the second factor on the right above by

$$\prod_{k=1}^{k_n}(k\log(1/q_n)) = k_n!\left(\frac{\pi}{\sqrt{6n}}\right)^{k_n}.$$

It is bounded below by

$$\prod_{k=1}^{k_n}[(k\log(1/q_n)) - (k\log(1/q_n))^2/2]$$
$$= \left(\prod_{k=1}^{k_n}(k\log(1/q_n))\right)\left(\prod_{k=1}^{k_n}[1 - k\log(1/q_n)/2]\right)$$
$$\ge k_n!\left(\frac{\pi}{\sqrt{6n}}\right)^{k_n}\left(1 - \log(1/q_n)\sum_{k=1}^{k_n}(k/2)\right) \sim k_n!\left(\frac{\pi}{\sqrt{6n}}\right)^{k_n}.$$

Therefore,

$$Q_{q_n}\left(\frac{\pi}{\sqrt{6n}}(1X_1,\ldots,k_nX_{k_n})=(v_{1,n},\ldots,v_{k_n,n})\right)$$

(5.1)

$$\sim \left(\prod_{k=1}^{k_n} e^{-v_{k,n}}\right) k_n!\left(\frac{\pi}{\sqrt{6n}}\right)^{k_n},$$

uniformly in $((v_{1,n},\ldots,v_{k_n,n}):\ n=1,2,\ldots)$. Let ζ_n denote the probability distribution of $(\pi/\sqrt{6n})(1X_1,2X_2,\ldots,k_nX_{k_n})$ when governed by Q_{q_n}. We omit the straightforward details, based on the preceding uniform asymptotic relation, of showing that the Prohorov distance between ζ_n and η_n approaches 0 as $n\to\infty$.

Let ξ_n denote the probability distribution of $\pi(1X_1,\ldots,k_nX_{k_n})/\sqrt{6n}$, when governed by P_n. We want to show that the Prohorov distance between ξ_n and η_n approaches 0. We have already shown that the Prohorov distance between ζ_n and η_n converges to 0. We will finish the proof by using Lemma 4.6 to conclude that the Prohorov distance between ξ_n and ζ_n approaches 0. Let $K_n=\{1,2,\ldots,k_n\}$. Lemma 4.6 applies since the left side of (4.13) is asymptotic to $k_n/\log^2(1/q_n)=o(n^{5/4})$. ∎

The local limit relation (5.1) is actually stronger than the assertion that the Prohorov distance between ζ_n and η_n converges to 0. It is natural to ask whether there is such an improvement of the assertion that the Prohorov distance between ξ_n and η_n converges to 0. There is such an improvement when certain, not very stringent, assumptions are made on $(v_{k,n}:\ 1\le k\le k_n,\ 1\le n)$. Such improvements can be made using the methods of this paper. The major change involves adding a hypothesis to Lemma 4.6 so that conclusions about ratios of probabilities can be drawn. The reader who wants to see the nature of such a local limit type improvement of a convergence in Prohorov distance (that is, convergence in distribution) type result can consult the local limit version of (2.2) by Auluck, Chowla, and Gupta [1].

6. The large parts. Proofs of 2.3, 2.4, 2.5, and 2.6

Since Theorem 2.3 follows from Theorem 2.5 by a straightforward calculation, we only consider Theorems 2.4–2.6. We begin with a proof of Theorem 2.5, replacing t by t_n in anticipation of an eventual modification to handle Theorem 2.6.

Let ν_n be the probability measure on $\mathbb{R}^{t_n}$ having the density

(6.1)
$$f_1(v_1)\prod_{s=2}^{t_n} f(v_{s-1},v_s),$$

where

(6.2)
$$f_1(v)=\exp(-e^{-v}-v)$$

as in (2.3) and

(6.3)
$$f(u,v)=\begin{cases} \exp(e^{-u}-e^{-v}-v) & \text{if } v\le u,\\ 0 & \text{if } v>u \end{cases}$$

as in (2.4). Let ξ_n and ζ_n denote the probability distributions of the random vector

$$\frac{\pi}{\sqrt{6n}}(Y_1, Y_2, \ldots, Y_{t_n}) - \left[\log\frac{\sqrt{6n}}{\pi}\right](1, 1, \ldots, 1)$$

when governed by P_n and Q_{q_n}, respectively. As indicated earlier, our plan is to use Proposition 4.1 to prove that the Prohorov distance between ζ_n and ν_n approaches 0 and to use Lemma 4.6 to prove that the Prohorov distance between ξ_n and ζ_n approaches 0. We begin by identifying ζ_n.

Lemma 6.1. *Let* $v_{1,n} \geq v_{2,n} \geq \cdots \geq v_{t_n,n}$ *and suppose that the sum of each* $v_{s,n}$ *and* $\log((6n)^{1/2}/\pi)$ *is a nonnegative integral multiple of* $\pi/(6n)^{1/2}$. *Then, the* ζ_n-*measure of the point* $(v_{1,n}, v_{2,n}, \ldots, v_{t_n,n})$ *is given by*

$$(6.4)\qquad e^{-\sum_{s=1}^{t_n} v_{s,n}}\left(\frac{\pi}{\sqrt{6n}}\right)^{t_n}\prod_{k=1+(\sqrt{6n}/\pi)(v_{t_n,n}+\log(\sqrt{6n}/\pi))}^{\infty}(1-q_n^k).$$

Proof. The proof consists of two parts. The first is to identify as a Markov chain the sequence $(Y_1, Y_2, \ldots)$ when governed by Q_q. The second is to insert q_n for q, make the appropriate change of variables to accommodate $((6n)^{1/2}/\pi)$ and its logarithm, and then simplify.

Part 1. We define $X_0 = \infty$ in order to avoid the necessity for special treatment of trivial cases. Let y_1 be a nonnegative integer. The event that $Y_1 = y_1$ is the event that $X_{y_1} > 0$ an $X_k = 0$ for $k > y_1$. By Proposition 4.1 the Q_q-probability of this event equals

$$(6.5)\qquad q^{y_1}\prod_{k=1+y_1}^{\infty}(1-q^k).$$

Next, we turn to the conditional probability, under Q_q, that $Y_s = y_s$ given the nonnegative integral values $y_1 \geq y_2 \geq \cdots \geq y_{s-1}$ of $Y_1, Y_2, \ldots, Y_{s-1}$. Of course, this conditional probability equals 0 unless y_s is a nonnegative integer satisfying $y_s \leq y_{s-1}$.

If $y_s = y_{s-1}$, then this conditional probability is the conditional probability that $X_{y_{s-1}} > x$ given that $X_{y_{s-1}} \geq x$ for some particular $x > 0$. We can calculate this conditional probability explicitly since (cf. Proposition 4.1) $X_{y_{s-1}}$ is geometrically distributed. Upon replacing y_{s-1} by y_s, the result is

$$\frac{q^{y_s(x+1)}}{q^{y_s x}} = q^{y_s}.$$

Consider the other case—$y_s < y_{s-1}$. In this case we are also given that $X_{y_{s-1}} \geq x$ for some $x > 0$. Now, conditioned by this information, we want the Q_q-probability that $X_{y_{s-1}} = x$, $X_{y_s} > 0$, and $X_k = 0$ for k strictly between y_s and y_{s-1}. By Proposition 4.1 this conditional probability equals

$$(6.6)\qquad q^{y_s}\prod_{k=1+y_s}^{y_{s-1}}(1-q^k).$$

With the convention that an empty product equals 1, this expression is also the correct formula, obtained above, in case $y_s = y_{s-1}$. In either case it does not depend on y_u for $u < s - 1$. Therefore, the sequence $(Y_1, Y_2, \ldots)$, when

governed by Q_q, is a Markov chain. Its initial probability distribution is given by (6.5) and its transition probabilities are given by (6.6).

Multiplication of the appropriate conditional probabilities gives

$$(6.7) \qquad Q_q(Y_s = y_s, \ 1 \le s \le t) = q^{\sum_{s=1}^{t} y_s} \prod_{k=1+y_t}^{\infty} (1 - q^k).$$

Part 2. We replace q by $q_n = \exp(-\pi/(6n)^{1/2})$, t by t_n, and y_s by

$$((6n)^{1/2}/\pi)(v_{s,n} + \log((6n)^{1/2}/\pi)).$$

The result is formula (6.4). $\square$

For $v_{1,n} \ge v_{2,n} \ge \dots$, we rewrite (6.1) in order to compare it to (6.4):

$$(6.8) \qquad e^{-\sum_{s=1}^{t_n} v_{s,n}} \exp(-e^{-v_{t_n,n}}).$$

The leading exponential is common to both (6.4) and (6.8). The power of $(\pi/(6n)^{1/2})$ in (6.4) is to be expected in a situation where one is hoping to prove that a discrete probability distribution is close to one with a density; it is the product over s of distances between possible values of $v_{s,n}$ in (6.4). Were we able to prove that the infinite product in (6.4) is uniformly asymptotic to the double exponential in (6.8), it would follow that the Prohorov distance between ζ_n and ν_n converges to 0 as $n \to \infty$. We will not be able to do quite this much, but we now proceed to show this uniform asymptotic relation for $v_{t_n,n} \ge -g(n)$, where g is any function for which

$$(6.9) \qquad g(n) - \frac{\log n}{4} \to -\infty$$

as $n \to \infty$. That is, we will show that

$$(6.10) \qquad \prod_{k > (\sqrt{6n}/\pi)(v + \log(\sqrt{6n}/\pi))} (1 - q_n^k) \sim e^{-e^{-v}}$$

uniformly for $v \ge -g(n)$, provided that (6.9) holds.

To prove (6.10) we multiply by $\exp(e^{-v})$ and take logarithms. Our task becomes that of showing

$$(6.11) \qquad \sum_{k > (\sqrt{6n}/\pi)(v + \log(\sqrt{6n}/\pi))} \log(1 - q_n^k) + e^{-v} \to 0$$

uniformly for $v \ge -g(n)$. We use the power series for the logarithms. A straightforward estimation using (6.9) shows that the total contribution from all terms after the first in the series for the various logarithms goes to 0 uniformly. The first terms themselves in the various power series for the logarithms constitute a geometric series the sum of which is between

$$-e^{-v} \frac{(\pi/\sqrt{6n})e^{-\pi/\sqrt{6n}}}{1 - e^{-\pi/\sqrt{6n}}} \quad \text{and} \quad -e^{-v} \frac{\pi/\sqrt{6n}}{1 - e^{-\pi/\sqrt{6n}}}$$

each of whose sums with e^{-v} goes to 0 uniformly for $v \ge -g(n)$, again because of (6.9).

To complete the proof that the Prohorov distance between ζ_n and ν_n approaches 0, we only need show that

$$(6.12) \qquad \nu_n(\{(u_1, u_2, \dots, u_{t_n}) : u_{t_n} < -g(n)\}) \to 0$$

as $n \to \infty$. When t_n is fixed, as it is in Theorem 2.5, we can satisfy (6.12) (and simultaneously (6.9)) by taking $g(n) \to \infty$ sufficiently slowly, say, $g(n) = (\log n)/8$.

Before completing the proof of Theorem 2.5 by showing that the Prohorov distance between ξ_n and ζ_n approaches 0, let us turn our attention to t_n depending on n and Theorem 2.6. To show that the Prohorov distance between ζ_n and ν_n goes to 0 for the setting of Theorem 2.6 we must concern ourselves with two things: (i) showing that the linear changes of variable involved in going from Theorem 2.5 to Theorem 2.6 do not create any problems for the argument given above and (ii) showing that g may be chosen to satisfy both (6.9) and (6.12).

For the first of these tasks we note that for the setting of Theorem 2.5 we showed an infinite product to be asymptotic to a double exponential, a relation that is preserved under linear changes of variables in both expressions. Also, for the setting of Theorem 2.5 we went from a discrete situation to a continuous situation, moving no more than $\pi/\sqrt{6n}$ in each coordinate direction. For the setting of Theorem 2.6, these movements are different. They are of order $\sqrt{s/n}$ in the direction of the coordinate axis corresponding to Y_s. The sum over s from 1 to t_n is of order $t_n^{3/2}/n$ which approaches 0 since a hypothesis of Theorem 2.6 is that $t_n/n^{1/4} \to 0$.

In order to bring ourselves to the same place with respect to Theorem 2.6 as we are with respect to Theorem 2.5 we must show that g can be chosen to satisfy both (6.9) and (6.12). With no loss of generality we may assume that $t_n \to \infty$ and choose $g(n) = \log(2t_n)$. That (6.9) holds is a consequence of the hypothesis $t_n/n^{1/4} \to 0$. That (6.12) holds is a consequence of the calculation:

$$
\nu_n\{(u_1, u_2, \ldots, u_{t_n}) : u_{t_n} < -\log(2t_n)\}
$$

$$
= \frac{1}{(t_n - 1)!} \int_{-\infty}^{-\log(2t_n)} \exp(-e^{-u} - t_n u)\, du
$$

$$
= \frac{1}{(t_n - 1)!} \int_{2t_n}^{\infty} e^{-r} r^{t_n - 1}\, dr \leq \frac{2}{(t_n - 1)!} \int_{2t_n}^{\infty} e^{-r} r^{t_n - 1} \left(1 - \frac{t_n - 1}{r}\right)\, dr
$$

$$
= \frac{2}{(t_n - 1)!} e^{-2t_n} (2t_n)^{t_n - 1},
$$

which, with the use of Stirling's Formula, is seen to approach 0.

To finish the proofs of Theorems 2.5 and 2.6 we must show, for each of the two settings, that the Prohorov distance between ξ_n and ζ_n approaches 0. The preceding paragraph shows that when studying the limiting behavior of the t_n largest parts governed by Q_{q_n} we need only consider parts belonging to the set

$$
K_n = \left\{k : k \geq \frac{\sqrt{6n}}{\pi} \left(\log \frac{\sqrt{6n}}{\pi} - \log(2t_n)\right)\right\}.
$$

A straightforward calculation using $t_n/n^{1/4} \to 0$ shows that (4.13) is satisfied and, hence, that Lemma 4.6 applies. Thus, when studying the limiting behavior of the t_n largest parts governed by P_n we need only consider parts belonging to K_n, and the limiting behavior is the same as when Q_{q_n} is the governing measure. ∎

To obtain Theorem 2.4 from Theorem 2.6 we integrate the variables w_1, w_2, $\ldots$, w_{t_n-1} over all possible values to conclude that the Prohorov distance between the probability distribution of

$$\sqrt{\frac{\pi^2 t_n}{6n}} Y_{t_n} - \sqrt{t_n} \log \frac{\sqrt{6n}}{\pi t_n}$$

and the probability distribution having density

$$(6.13) \qquad \frac{t^{t-(1/2)}}{\Gamma(t)} e^{-te^{-w/\sqrt{t}} - w\sqrt{t}}$$

approaches 0, where we have replaced t_n by t in the formula for the density, a function of w. It remains for us to show that, as $t \to \infty$, the distribution with density (6.13) converges to the normal distribution. Converting to characteristic functions we see that our problem is to prove that the Fourier transform of (6.13), expressed in terms of a variable θ, converges to $\exp(-\theta^2/2)$. The Fourier transform of (6.13) is easily calculated by substituting a new variable for $t \exp(-w/\sqrt{t})$. The result is

$$t^{i\theta\sqrt{t}} \frac{\Gamma(t - i\theta\sqrt{t})}{\Gamma(t)}.$$

Stirling's asymptotic formula for the gamma function gives $-\theta^2/2$ as the limit of the logarithm of this transform. ∎

7. Large values of kX_k. Proofs of 2.7 and 2.8

We focus on Theorem 2.8 since Theorem 2.7 is an immediate consequence of it. It suffices to prove that

$$\lim_{n\to\infty} P\left(w_s < \frac{\pi}{\sqrt{6n}} Z_s - \log \frac{\sqrt{6n}}{\pi} - \log\log\log n \le v_s, \ 1 \le s \le t\right)$$

$$= \int_{w_1}^{v_1} \int_{w_2}^{v_2} \cdots \int_{w_t}^{v_t} f_1(u_1) \prod_{s=2}^{t} f(u_{s-1}, u_s)\, du_t \cdots du_2\, du_1,$$

whenever $v_1 > w_1 \ge v_2 > w_2 \ge \cdots \ge v_t > -\infty = w_t$. The integration can be performed [by using the formulas for f_1 and f given at (6.2) and (6.3)]: the result is

$$\exp(-e^{-v_t}) \prod_{s=1}^{t-1} (e^{-w_s} - e^{-v_s}).$$

Thus, by proving each of the following statements we will have completed the proof of Theorem 2.8:

$$(7.1) \qquad \prod_{k=1}^{\infty} Q_{q_n}\left(\frac{\pi}{\sqrt{6n}} kX_k - \log \frac{\sqrt{6n}}{\pi} - \log\log\log n \le v\right) \to \exp(-e^{-v});$$

$$(7.2) \qquad \sum_{k=1}^{\infty} Q_{q_n}\left(w < \frac{\pi}{\sqrt{6n}} kX_k - \log \frac{\sqrt{6n}}{\pi} - \log\log\log n \le v\right) \to e^{-w} - e^{-v};$$

$$Q_{q_n}\left(w_s < \frac{\pi}{\sqrt{6n}}Z_s - \log\frac{\sqrt{6n}}{\pi} - \log\log\log n \le v_s, \ 1 \le s \le t\right)$$

$$(7.3) \qquad \sim \prod_{k=1}^{\infty} Q_{q_n}\left(\frac{\pi}{\sqrt{6n}}kX_k - \log\frac{\sqrt{6n}}{\pi} - \log\log\log n \le v_t\right)$$

$$\times \prod_{s=1}^{t-1}\sum_{k=1}^{\infty} Q_{q_n}\left(w_s < \frac{\pi}{\sqrt{6n}}kX_k - \log\frac{\sqrt{6n}}{\pi} - \log\log\log n \le v_s\right) ;$$

$$P_n\left(w_s < \frac{\pi}{\sqrt{6n}}Z_s - \log\frac{\sqrt{6n}}{\pi} - \log\log\log n \le v_s, \ 1 \le s \le t\right)$$

$$(7.4) \qquad \sim Q_{q_n}\left(w_s < \frac{\pi}{\sqrt{6n}}Z_s - \log\frac{\sqrt{6n}}{\pi} - \log\log\log n \le v_s, \ 1 \le s \le t\right) .$$

The proofs of (7.1) and (7.2) will rely on the following lemma.

Lemma 7.1. *As* $n \to \infty$,

$$(7.5) \qquad \sum_{k=1}^{\infty} Q_{q_n}\left(\frac{\pi}{\sqrt{6n}}kX_k - \log\frac{\sqrt{6n}}{\pi} - \log\log\log n > v\right) \to e^{-v} .$$

Proof. Recall that $\lceil x \rceil$ denotes the least integer that is no smaller than x. Also, the notations

$$(7.6) \qquad\qquad\qquad a_n = \frac{\pi}{\sqrt{6n}},$$

$$(7.7) \qquad\qquad\qquad b_n = \log\frac{\sqrt{6n}}{\pi} + \log\log\log n,$$

$$c_n = b_n + v$$

will be used. By (4.2) and Proposition 4.1 the left side of (7.5) equals

$$(7.8) \qquad \sum_{k=1}^{\infty} e^{-ka_n\lceil c_n/ka_n \rceil} = \sum_{k=1}^{(\beta_n/a_n)-1} e^{-ka_n\lceil c_n/ka_n \rceil} + \sum_{k=\beta_n/a_n}^{\infty} e^{-ka_n\lceil c_n/ka_n \rceil} ,$$

where β_n is chosen so that it is an integral multiple of a_n and that it goes to ∞ as $n \to \infty$ slowly enough so that $\beta_n/\log\log n \to 0$. In the first summation on the right side of (7.8), there are less than (β_n/a_n) terms each of which is bounded by $\exp(-c_n)$. This summation must approach 0 as $n \to \infty$ since $(\beta_n/a_n)\exp(-c_n) \to 0$.

Breaking the last summation in (7.8) according to the value of $m = \lceil c_n/ka_n \rceil$, we get

$$\sum_{m=1}^{\lceil c_n/\beta_n \rceil}\sum_{k=\lceil c_n/ma_n \rceil}^{\lceil c_n/(m-1)a_n \rceil - 1} e^{-kma_n}$$

$$(7.9)$$

$$= \sum_{m=1}^{\lceil c_n/\beta_n \rceil} \frac{e^{-ma_n\lceil c_n/ma_n \rceil}}{1 - e^{-ma_n}}(1 - e^{ma_n(\lceil c_n/ma_n \rceil - \lceil c_n/(m-1)a_n \rceil)}) .$$

The exponent in the second exponential is bounded above by

$$(7.10) \qquad ma_n + mc_n \left(\frac{1}{m} - \frac{1}{m-1} \right).$$

The first term in (7.10) goes to 0 uniformly for $m \le \lceil c_n/\beta_n \rceil$ since $c_n a_n \to 0$. That the second term in (7.10) goes uniformly to $-\infty$ follows from the fact that $\beta_n \to \infty$. So, in evaluating the limit of (7.9) we may replace the second factor by 1. Also, since $c_n a_n \to 0$, we may replace the denominator in the first factor by ma_n. The numerator in the first factor in (7.9) is bounded above by $\exp(-c_n)$ and below by

$$\exp(-c_n) \exp(-ma_n) \ge \exp(-c_n) \exp(-c_n a_n) \sim \exp(-c_n).$$

Therefore, the limit of (7.9) as $n \to \infty$ is the same as the limit of

$$\frac{e^{-c_n}}{a_n} \sum_{m=1}^{\lceil c_n/\beta_n \rceil} \frac{1}{m} = e^{-v} \frac{e^{-b_n}}{a_n} (\log b_n + O(\log \log \log n)),$$

the limit of which is e^{-v}. $\square$

Notice that the summands in (7.8), being bounded above by $\exp(-c_n)$ approach 0 as $n \to \infty$, uniformly in k. Let us label this fact for future reference:

$$(7.11) \qquad \lim_{n \to \infty} \sup_{1 \le k < \infty} Q_{q_n} \left(\frac{\pi}{\sqrt{6n}} k X_k - \log \frac{\sqrt{6n}}{\pi} - \log \log \log n > v \right) = 0.$$

To prove (7.1) we, using the notation in (7.6) and (7.7), write the left side in that assertion as

$$\exp \left(\sum_{k=1}^{\infty} \log(1 - Q_{q_n}(a_n k X_k - b_n > v)) \right),$$

which is asymptotic to

$$\exp \left(-\sum_{k=1}^{\infty} Q_{q_n}(a_n k X_k - b_n > v) \right) \to \exp(-e^{-v})$$

because of (7.11), the fact that all terms in the summation have the same sign, and Lemma 7.1.

To prove (7.2) we apply Lemma 7.1 twice, once as it stands and once with w in lieu of v, and then subtract.

The right side of (7.3) is the sum over all sequences $(k_1, k_2, \ldots, k_{t-1})$ of

$$\prod_{k=1}^{\infty} Q_{q_n} \left(\frac{\pi}{\sqrt{6n}} k X_k - \log \frac{\sqrt{6n}}{\pi} - \log \log \log n \le v_t \right)$$

$$\times \prod_{s=1}^{t-1} Q_{q_n} \left(w_s < \frac{\pi}{\sqrt{6n}} k_s X_{k_s} - \log \frac{\sqrt{6n}}{\pi} - \log \log \log n \le v_s \right).$$

In view of (7.11) we see that we can do two things without changing the asymptotic behavior of the right side of (7.3): (i) only sum over those $(k_1, k_2, \ldots, k_{t-1})$ consisting of $t-1$ different values and (ii) in the first product eliminate the $t-1$ factors corresponding to k being equal to some k_s, $1 \le s \le t-1$.

When this has been done all the products are products of probabilities of independent events, and thus the products represent probabilities of intersections. The sum over $(k_1, k_2, \ldots, k_{t-1})$ is the sum of probabilities of disjoint events. The upshot is that when the right side of (7.3) is adjusted as described above, it becomes equal to the left side of (7.3). Therefore, (7.3) is true.

Lemma 4.6 is the natural tool for a proof of (7.4). The issue is what we should use for the set K_n that plays a role in that lemma, because for the problem under discussion the random variables X_k for all k apparently play a role. We let

$$K_n = \{k : k \geq n^{1/2}\beta_n\} \cup \{k : k \leq n^{1/2}/\beta_n\},$$

where $\beta_n \to \infty$ sufficiently slowly that $\beta_n^2/\log\log n \to 0$. A straightforward estimation shows that $\sum_{k \in K_n} k^2 e^{-ka_n} = o(n^{3/2})$. So, Lemma 4.6 applies for K_n as chosen. We turn to investigating the changes in the rest of our argument when we consider only $k \in K_n$ as opposed to considering all k.

The fact that only the second term on the right side of (7.8) has played a contributing role in the analysis of the limiting behavior of (7.8) means that, as $n \to \infty$, the Q_{q_n} probability approaches 1 that the largest t values of kX_k come from among those k that are greater than or equal to $n^{1/2}\beta_n$. Thus, when Q_{q_n} is the governing probability measure, it is correct to restrict consideration to $k \in K_n$ when considering the largest t values of kX_k.

In so far as P_n is concerned when $n \to \infty$, we will show the values of $k \notin K_n$ can be ignored when looking for the t largest values of kX_k, thus completing the proof. We have already shown that, when governed by Q_{q_n}, the t largest values of kX_k for $k \in K_n$ are, with probability approaching 1, greater than

$$\frac{\sqrt{6n}}{\pi}(\log n^{1/2} + \log\alpha_n) = \frac{\sqrt{6n}}{\pi}\log(n^{1/2}\alpha_n),$$

provided that α_n is chosen so that $\alpha_n/\log\log n \to 0$. By Lemma 4.6, the same is true when the probability measures P_n are used. We choose such a sequence α_n that also satisfies $\beta_n^2/\alpha_n \to 0$, which we can do since $\beta_n^2/\log\log n \to 0$.

It remains for us to show that the probability approaches 0 that for some $k \notin K_n$, kX_k is greater than $\sqrt{6n}\log(n^{1/2}\alpha_n)/\pi$. If a particular kX_k is greater than this quantity, then by removing from the partition of n all the parts of size k we obtain a partition of some integer $j < n - \sqrt{6n}\log(n^{1/2}\alpha_n)/\pi$ for which k divides $n - j$. To obtain an asymptotic formula for the total number such partitions we can use the Hardy-Ramanujan asymptotic formula $e^{2\pi\sqrt{j/6}}/4j\sqrt{3}$ and sum over those j that are less than

$$n - \frac{\sqrt{6n}}{\pi}\log(n^{1/2}\alpha_n)$$

and for which $n - j$ is a multiple of k. Replacing the sum by an integral and using an approximate antiderivative gives a bound on the sum; for some

positive finite constant c, the sum is no larger than

$$\frac{c}{k\sqrt{n}} \exp\left[\frac{2\pi}{\sqrt{6}}\left(n - \frac{\sqrt{6n}}{\pi}\log(n^{1/2}\alpha_n)\right)^{1/2}\right]$$

$$\leq \frac{c\beta_n}{n} \exp\left[2\pi\sqrt{\frac{n}{6}}\left(1 - \frac{1}{2\pi}\sqrt{\frac{6}{n}}\log(n^{1/2}\alpha_n)\right)\right] = \frac{c\beta_n}{n^{3/2}\alpha_n}e^{2\pi\sqrt{n/6}}.$$

We divide by $p(n)$, the number of partitions of n, to obtain an upper bound on the P_n-probability that a particular kX_k is among the largest t of such products and then multiply by $n^{1/2}\beta_n$, an upper bound for the number of $k \notin K_n$. The result of doing these operations, using the Hardy-Ramanujan asymptotic formula for $p(n)$, is $O(\beta_n^2/\alpha_n)$, which goes to 0 as $n \to \infty$, as desired. ∎

8. Parts $\geq$ or $\leq k_n$. Proofs of 2.9 and 2.10

For Theorem 2.9 there are four steps: (i) the use of Lemma 4.6 to conclude that Q_{q_n} can be replaced by P_n in the conclusions we reach for Q_{q_n} via the forthcoming steps (ii), (iii), and (iv), (ii) the calculation of the variance of $\sum_{k \geq k_n} X_k$ when governed by Q_{q_n}, (iii) the calculation of its expectation, and (iv) the verification that the sequence $(Q_{q_n}: n = 1, 2, \ldots)$ satisfies sufficient conditions for a standard central theorem to apply. For Theorem 2.10 we replace $\sum_{k \geq k_n} X_k$ by

$$\sum_{k_n^{1/2} < k \leq k_n} X_k$$

and, for it, engage in steps (i), (ii), and (iii) together with an alternative step (iv).

For step (i) we need only check (4.13) for $K_n = \{k: k \geq k_n\}$ when $k_n/n^{1/2} \to \infty$ and $K_n = \{k: k \leq k_n\}$ when $k_n/n^{1/2} \to 0$. We omit the easy estimates.

For step (ii) we use Proposition 4.1 and the known formula for the variance of a geometrically distributed random variable to conclude that, when $k_n/n^{1/2} \to \infty$, the variance of $\sum_{k \geq k_n} X_k$, when governed by Q_{q_n}, equals

$$\sum_{k \geq k_n} \frac{q_n^k}{[1 - q_n^k]^2} = \sum_{k \geq k_n} \frac{e^{-\pi k/\sqrt{6n}}}{[1 - e^{-\pi k/\sqrt{6n}}]^2} \sim \frac{\sqrt{6n}}{\pi} \int_{\pi k_n/\sqrt{6n}}^{\infty} \frac{e^{-y}}{[1 - e^{-y}]^2}\, dy$$

$$= \frac{\sqrt{6n}}{\pi}\left(\frac{e^{-\pi k_n/\sqrt{6n}}}{1 - e^{-\pi k_n/\sqrt{6n}}}\right) \sim \frac{\sqrt{6n}}{\pi}e^{-\pi k_n/\sqrt{6n}}.$$

Similarly, when $k_n/n^{1/2} \to 0$ and $k_n \to \infty$, the variance of

$$\sum_{k_n^{1/2} < k \leq k_n} X_k$$

equals

$$\sum_{k_n^{1/2} < k \leq k_n} \frac{q_n^k}{[1 - q_n^k]^2} \sim \sum_{k_n^{1/2} < k \leq k_n} \frac{6n}{\pi^2 k^2} = o(n).$$

For step (iii) we use Proposition 4.1 to calculate the expectation of $\sum_{k \geq k_n} X_k$, when governed by Q_{q_n}

$$
\sum_{k \geq k_n} \frac{q_n^k}{1 - q_n^k} = \sum_{k \geq k_n} \frac{e^{-\pi k/\sqrt{6n}}}{1 - e^{-\pi k/\sqrt{6n}}}
$$

$$
= \frac{\sqrt{6n}}{\pi} \int_{\pi k_n/\sqrt{6n}}^{\infty} \frac{e^{-y}}{1 - e^{-y}} \, dy + O\left(\frac{e^{-\pi k_n/\sqrt{6n}}}{1 - e^{-\pi k_n/\sqrt{6n}}} \right)
$$

$$
= \frac{\sqrt{6n}}{\pi} \log \frac{1}{1 - e^{-\pi k_n/\sqrt{6n}}} + O\left(\frac{e^{-\pi k_n/\sqrt{6n}}}{1 - e^{-\pi k_n/\sqrt{6n}}} \right) .
$$

The first term is the term that is subtracted in the numerator of Theorem 2.9; and the second term (the "big Oh" term) divided by the standard deviation approaches 0, and so can be ignored. The expectation of

$$
\text{(8.1)} \qquad\qquad \sum_{k_n^{1/2} < k \leq k_n} X_k
$$

equals

$$
\text{(8.2)} \qquad \sum_{k_n^{1/2} < k \leq k_n} \frac{q_n^k}{1 - q_n^k} \sim \sum_{k_n^{1/2} < k \leq k_n} \frac{\sqrt{6n}}{\pi k} = \frac{\sqrt{6n}}{\pi} \log k_n^{1/2} + o(n^{1/2}) .
$$

Turning to step (iv) for the setting of Theorem 2.9, there is one minor difficulty in applying a standard theorem giving conditions for convergence to the normal distribution of distributions of finite series of independent random variables; our series are infinite, being over all $k \geq k_n$. Of course, with probability 1 all but finitely many terms are 0. And, with probability close to 1 all terms but those in an appropriate fixed finite collection are 0. This last fact makes it clear that we can apply the standard theorems even though in our setting the individual series are infinite. In view of Proposition 4.1 and the Normal Convergence Criterion in §22.2 of [6], we can complete the proof of convergence to the normal distribution by showing that, for each positive ε,

$$
\text{(8.3)} \qquad \sum_{k \geq k_n} \sum_{x > \varepsilon n^{1/4} \exp(-n k_n/2\sqrt{6n})} x^2(1 - q_n^k)q_n^{kx} \to 0
$$

as $n \to \infty$. To do this one begins by discarding the factor $(1 - q_n^k)$, then interchanging the order of summation, and, after the resulting interior summation is completed, approximating the resulting single summation by an integral. Finally, (8.3) follows in a straightforward manner from the hypothesis $\pi k_n/\sqrt{6n} - \log n/2 \to -\infty$.

In the discussion leading to the statement of Theorem 2.10 in §2, a proof of that theorem was given for k_n approaching ∞ more slowly than $n^{1/4}$. Since we now, as we are proving Theorem 2.10 in full generality, are assuming that $k_n/n^{1/2} \to 0$, we already know the desired conclusion to hold with $k_n^{1/2}$ in lieu of k_n. So, beginning with this conclusion for $k_n^{1/2}$, we add $\pi/\sqrt{6n}$ multiplied by (8.1) and subtract this same factor multiplied by (8.2) at the left of the inequality in Theorem 2.10. The conclusion is not affected by these manipulations because step (i) of this proof shows that the variance of the difference that we

have inserted approaches 0 as $n \to \infty$. The above manipulations have resulted in an undesired term $o(1) = o(n^{1/2})/n^{1/2}$ from step (iii) that can be dropped without affecting the conclusion.   ∎

9. RESULTS FOR DISTINCT-PART PARTITIONS

A *distinct-part* partition of a positive integer is a way of writing it as the sum of distinct positive integers without regard to order. We denote the number of distinct-part partitions of n by $p_d(n)$. There are four distinct-part partitions of the number $6: 3 + 2 + 1$, $4 + 2$, $5 + 1$, and 6; thus, $p_d(6) = 4$. The empty partition of 0 is a distinct-part partition so $p_d(0) = 1$. The generating function for the sequence $(p_d(0), p_d(1), p_d(2), \dots)$ is well known:

$$(9.1) \qquad \sum_{n=0}^{\infty} p_d(n) q^n = \prod_{k=1}^{\infty} (1 + q^k).$$

The functions X_k and Y_t introduced in §1 will, in this and the next section, still be a focus of our attention. But they will have probability distributions different from those in earlier sections because the underlying probability distributions on Λ will be different. We let $P_{d,n}$ denote the probability measure that assigns probability $1/p_d(n)$ to each of the $p_d(n)$ distinct-part partitions of n (and probability 0 to any partition that is not a distinct-part partition of n).

For the remainder of this section we describe the results for the 'distinct-part' setting. In the following section we deal with the methods of proof. The first two theorems are concerned with the small parts.

Theorem 9.1. *Suppose that $k_n/n^{1/2} \to 0$ as $n \to \infty$. Then, as $n \to \infty$,*

$$P_{d,n}(X_{k_n} = 1) \to 1/2.$$

This theorem says that the presence or absence of any particular small part in a random distinct-part partition of a large integer is approximately determined by a coin flip. Can independent coin flips be used for different sizes of small parts? The answer, according to the following theorem, is 'yes' provided that a stricter interpretation of small is used than that used in the preceding theorem.

Theorem 9.2. *Suppose that $k_n/n^{1/4} \to 0$ as $n \to \infty$. For each n, let ξ_n be the probability measure on $\mathbb{R}^{k_n}$ induced by $P_{d,n}$ via the random vector $(X_1, X_2, \dots, X_{k_n})$ and let η_n denote the measure on $\mathbb{R}^{k_n}$ that assigns probability 2^{-k_n} to each point all of whose coordinates equal 1 or 0. Then, the Prohorov distance between ξ_n and η_n approaches 0 as $n \to \infty$.*

Under the slightly stronger assumption that $k_n \le n^{1/5}$, the preceding theorem can be proved in a rather straightforward manner from the lemma in §5 of [4]; in fact, the stronger assertion that

$$P_{d,n}(X_k = x_{k,n} \text{ for } k \le k_n) \sim (1/2)^{k_n}$$

can be proved from that lemma. This local limit theorem can also be proved using the methods of this paper even if only $k_n/n^{1/4} \to 0$ is assumed.

From Theorem 9.2 we conclude that the number of parts no larger than a fixed constant k_n has a distribution that converges to the binomial distribution on $\{0, 1, \dots, k_n\}$ with mean $k_n/2$. Also, we can conclude that if $k_n \to \infty$ sufficiently slowly that $k_n/n^{1/4} \to 0$, then the distribution, after normalization

by subtraction of $k_n/2$ and division by $\sqrt{k_n}/2$, is close to the normal distribution. The following theorem says that there is a normal limit even if k_n grows faster than $n^{1/4}$ provided that it grows more slowly than $n^{1/2}$. However, if k_n grows at least as fast as $n^{1/3}$, then the quantity that must be subtracted from the number of parts no larger than k_n is essentially smaller than $k_n/2$.

Theorem 9.3. *If* $k_n \to \infty$ *and* $k_n/n^{1/2} \to 0$ *as* $n \to \infty$*, then*

$$P_{d,n}\left(\frac{\sum_{k \leq k_n} X_k - \frac{\sqrt{12n}}{\pi}\log(2/(1 + \exp(-\pi k_n/\sqrt{12n})))}{\sqrt{k_n}/2} \leq x\right)$$

$$\to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-u^2/2}\,du$$

as $n \to \infty$.

Even though we cannot replace the logarithm by a simpler expression in the preceding theorem we can make such a substitution in the law of large numbers that follows from it.

Corollary. *Under the hypotheses of Theorem 9.3,*

$$\lim_{n \to \infty} P_{d,n}\left(\left|\frac{\sum_{k \leq k_n} X_k}{k_n/2} - 1\right| > \varepsilon\right) = 0$$

for each $\varepsilon > 0$.

For the large parts there are results very similar to Theorems 2.3–2.6. The only difference comes from a constant in the normalization. We state this semi-formally.

Theorem 9.4. *Each of Theorems* 2.3, 2.4, 2.5, *and* 2.6 *remains true if* P_n *is replaced by* $P_{d,n}$ *and every occurrence of* $6n$ *is replaced by* $12n$.

Just as we obtained, in §3, a limiting Poisson distribution for the number of very large parts in the case of unrestricted partitions, we conclude that, when governed by $P_{d,n}$, the number of parts greater than

$$\frac{\sqrt{12n}}{\pi}\left(\log\frac{\sqrt{12n}}{\pi} + v\right)$$

has a limiting Poisson distribution with expectation e^{-v}. In so far as an analog of Theorem 2.9 is concerned we have the following result which cannot be written by merely replacing $6n$ by $12n$ in Theorem 2.9.

Theorem 9.5. *Suppose that* $k_n/n^{1/2} \to \infty$ *and*

$$\frac{\pi k_n}{\sqrt{12n}} - \frac{\log n}{2} \to -\infty$$

as $n \to \infty$. *Then*

$$P_{d,n}\left(\frac{\sum_{k \geq k_n} X_k - \pi^{-1}(12n)^{1/2}\log(1 + e^{-\pi k_n/\sqrt{12n}})}{\pi^{-1/2}(12n)^{1/4}e^{-\pi k_n/2\sqrt{12n}}} \leq x\right) \to \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-u^2/2}\,du$$

as $n \to \infty$.

We will not formally state the law of large numbers that is an immediate consequence. It is similar to the corollary of Theorem 2.9; $12n$ replaces $6n$, and the hypotheses are those of Theorem 9.5.

10. METHODS. DISTINCT-PART PARTITIONS

The results described in the preceding section can be proved by methods very similar to those used in §§4, 5, 6, and 8. The details will be omitted, but the remainder of this section will be devoted to some comments on the proofs.

Let $q \in (0, 1)$. For any distinct-partition λ let

$$(10.1) \qquad Q_{d,q}(\lambda) = q^{|\lambda|} \prod_{k=1}^{\infty} (1 + q^k)^{-1}.$$

In view of (9.1), summation of (10.1) over all distinct-part λ for which $|\lambda| = n$ followed by summation over all n gives 1. Thus, $Q_{d,q}$ is a probability measure. The first of these two summations also gives the probability distribution of the random variable $N(\lambda) = |\lambda|$, when governed by $Q_{d,q}$:

$$Q_{d,q}(N = n) = p_d(n)q^n \prod_{k=1}^{\infty} (1 + q^k)^{-1}.$$

The joint probability distribution of the random variables $X_1, X_2, X_3, \ldots$ is also easily obtained:

$$Q_{d,q}(X_k = x_k \text{ for } k = 1, 2, \ldots) = \prod_{k=1}^{\infty} \frac{q^{kx_k}}{1 + q^k}$$

for all sequences $(x_1, x_2, \ldots)$ of 0's and 1's. Thus, the X_k are independent $\{1, 0\}$-valued random variables and $Q_{d,q}(X_k = 1) = q^k/(1 + q^k)$. The probability measure $P_{d,n}$, which is really the measure of interest, equals the measure obtained by conditioning the measure $Q_{d,q}$ by the event $N = n$.

Arguments similar to those leading to Proposition 4.3 give the generating function of N, when governed by $Q_{d,q}$. As a function of r it equals

$$\prod_{k=1}^{\infty} \frac{1 + (qr)^k}{1 + q^k}.$$

The expectation and variance, which are also easily obtained, equal

$$\sum_{k=1}^{\infty} \frac{kq^k}{(1 + q^k)} \quad \text{and} \quad \sum_{k=1}^{\infty} \frac{k^2q^k}{(1 + q^k)^2},$$

respectively. The proof of Corollary 4.4 can be mimicked for the present setting. As one does so, one sees that an appropriate choice for q_n is $\exp(-\pi/\sqrt{12n})$, rather than the value given in (4.2). For this choice of q_n, the variance of N, when governed by Q_{d,q_n}, is asymptotic to $(48^{1/2}/\pi)n^{3/2}$ and the difference between n and the expected value of N equals $o(n^{3/4})$.

The path leading to Lemma 4.6 can be followed for the current setting. The condition (4.13) should be replaced by

$$\sum_{k \in K_n} \frac{k^2 q_n^k}{(1 + q_n^k)^2} = o(n^{3/2}).$$

The methods of §§5, 6, and 8 carry over in a straightforward manner.

11. General part restrictions

Distinct-part partitions can be regarded as partitions satisfying the restriction that the multiplicity of each part size k is a member of $H_k = \{1, 0\}$. We can generalize by considering H_k to be an arbitrary nonempty set of nonnegative integers. We write $h = (H_1, H_2, \ldots)$ and assume that $0 \in H_k$ for all but finitely many k. Let $p_h(n)$ denote the number of partitions of n that satisfy the restriction that the number of parts of size k belongs to H_k. The generating function of p_h is given by

$$(11.1) \qquad \sum_{n=0}^{\infty} p_h(n) q^n = \prod_{k=1}^{\infty} \left(\sum_{w \in H_k} q^{kw} \right).$$

Provided that $p_h(n) > 0$, we set $P_{h,n}(\lambda) = 1/p_h(n)$ if λ is a partition of n for which the sequence of cardinalities of its parts belongs to h and $P_{h,n}(\lambda) = 0$ otherwise. The probability measure $P_{h,n}$ is the appropriate measure for the experiment of choosing, on an equiprobable basis, a partition of n for which, for each k, the number of parts equal to k is a member of H_k.

Even in this generality it is possible to arrange for a partition to be chosen at random so that the random variables X_k are independent random variables and, when the probability distribution is conditioned by the event that $N = n$, the measure $P_{h,n}$ is obtained. The appropriate probability measure $Q_{h,q}$ is given by

$$(11.2) \qquad Q_{h,q}(\lambda) = q^{|\lambda|} \prod_{k=1}^{\infty} \left(\sum_{w \in H_k} q^{kw} \right)^{-1}$$

for each partition λ having the property that the number of parts of size k belongs to H_k.

From (11.1) and (11.2) the following facts follow in a straightforward manner. The distribution of N, when governed by $Q_{h,q}$, is given by

$$Q_{h,q}(N = n) = p_h(n) q^n \prod_{k=1}^{\infty} \left(\sum_{w \in H_k} q^{kw} \right)^{-1},$$

and the corresponding generating function is given by

$$\sum_{n=0}^{\infty} Q_{h,q}(N = n) r^n = \prod_{k=1}^{\infty} \frac{\sum_{x \in H_k} (qr)^{kx}}{\sum_{w \in H_k} q^{kw}}.$$

Also,

$$Q_{h,q}(X_k = x_k \text{ for } k = 1, 2, \ldots) = \prod_{k=1}^{\infty} \frac{q^{kx_k}}{\sum_{w \in H_k} q^{kw}}$$

for every sequence $(x_1, x_2, \ldots)$ satisfying $x_k \in H_k$ for each k. The random variables X_k are thus independent and, for $x \in H_k$,

$$Q_{h,q}(X_k = x) = \frac{q^{kx}}{\sum_{w \in H_k} q^{kw}}.$$

If $p_h(n) > 0$, or, equivalently, $Q_{h,q}(N = n) > 0$, then $Q_{h,q}$ conditioned by the event that $N = n$ is the probability measure $P_{h,n}$.

REFERENCES

1. F. C. Auluck, S. Chowla, and H. Gupta, *On the maximum value of the number of partitions of n into k parts*, J. Indian Math. Soc. (N.S.) **6** (1942), 105–112.

2. Narasingha R. Chaganty and J. Sethuraman, *Strong large deviation and local limit theorems*, Ann. Probab. (to appear).

3. Paul Erdös and Joseph Lehner, *The distribution of the number of summands in the partitions of a positive integer*, Duke Math. J. **8** (1941), 335–345.

4. Paul Erdös and M. Szalay, *On the statistical theory of partitions*, Topics in Classical Number theory, vol. I (G. Halász, ed.), North-Holland, Amsterdam, 1984, pp. 397–450.

5. I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 4th ed., Academic Press, New York, 1965.

6. M. Loéve, *Probability theory* I, 4th ed., Springer-Verlag, New York, 1977.

7. M. Szalay and P. Turán, *On some problems of the statistical theory of partitions with application to characters of the symmetric group.* I, Acta Math. **29** (1977), 361–379.

8. ____, *On some problems of the statistical theory of partitions with application to characters of the symmetric group.* II, Acta Math. **29** (1977), 381–392.

9. ____, *On some problems of the statistical theory of partitions with application to characters of the symmetric group.* III, Acta Math. **32** (1978), 129–155.

SCHOOL OF MATHEMATICS, UNIVERSITY OF MINNESOTA, MINNEAPOLIS, MINNESOTA 55455
E-mail address: fristedt@math.umn.edu