

DENSITIES OF LENGTH SPECTRAL PARTITIONS OF NATURAL NUMBERS

BENJAMIN GARCIA MORALES, WAI YAN PONG

ABSTRACT. The notion of length spectrum for natural numbers was introduced by Pong in [5]. In this article, we answer the question of how often one can recover a random number from its length spectrum. We also include a quick deduction of a result of LeVeque in [4] on the average order of the size of length spectra.

1. LENGTH SPECTRA

Let m be a natural number. The *length spectrum* or simply the *spectrum* of m , denoted by $\text{lspec}(m)$, is the set of lengths of sequences of consecutive natural numbers that sum to m . For example, the number 9 can be written as sum of consecutive natural numbers in three different ways: 9 , $4 + 5$ and $2 + 3 + 4$. Hence, $\text{lspec}(9) = \{1, 2, 3\}$. It was shown in [5, Theorem 1.2] that $\text{lspec}(m)$ is

$$(1.1) \quad \{k: k \mid m, k \text{ odd}, k^2 < 2m\} \cup \{2m/k: k \mid m, k \text{ odd}, k^2 > 2m\}.$$

Two numbers are *spectral equivalent* if they have the same spectrum. We write $L(m)$ for the spectral class of m (the equivalence class of m under spectral equivalence). An algorithm for computing $L(m)$ from $\text{lspec}(m)$ was given in [5] and since $\text{lspec}(m)$ is computable from m via (1.1), so $L(m)$ is computable from m . For instance, $L(1) = \{2^k: k \geq 0\}$, $L(9) = \{9\}$ and $L(175) = \{175, 245\}$. Surprisingly, the sizes of these spectral classes are all that possible:

Theorem 1 (Theorem 4.6 [5]). *A spectral class has either 1, 2 or infinitely many numbers.*

In other words, X_1, X_2, X_ω , where X_k is the set of numbers with spectral class of size k , form a partition of \mathbb{N} . Clearly, the chance of recovering a number in X_k from its spectrum is 1-in- k (with 1-in- ω understood to be 0). We go one step further and ask:

What is the chance of guessing a randomly chosen number from its spectrum?

To be precise, we fix a notion of probability. For $X \subseteq \mathbb{N}$, let $X(n)$ denote the set of elements of X not exceeding n . The *natural density* of X is the limit, if exists, $\delta(X) := \lim_{n \rightarrow \infty} |X(n)|/n$. We regard $\delta(X)$ as the probability¹ of a random number being in X . If each $\delta(X_k)$ exists ($k = 1, 2, \omega$), then the answer to our question would be

$$1 \cdot \delta(X_1) + \frac{1}{2} \cdot \delta(X_2) + 0 \cdot \delta(X_\omega).$$

Date: September 8, 2022.

¹Natural density can be extended to a finitely additive probability measure on the power set of \mathbb{N} [3, Theorem 3].

2. SPECTRAL CLASSES

For $A, B \subseteq \mathbb{N}$, let AB denote the set $\{ab: a \in A, b \in B\}$. We write aB for $\{a\}B$ and A^2 for AA . We use A_0 and A_1 to denote the set of even and odd elements of A , respectively. A spectrum S is

- *unmixed* if $|S_0| = 0$;
- *balanced* if $|S_0| = |S_1|$; and
- *lopsided* if it is neither unmixed nor balanced.

For convenience, let us call an odd factor k of m *small* (resp. *large*) if $k^2 < 2m$ (resp. $k^2 > 2m$). It is clear that every number must have at least half of its odd factors being small. So according to (1.1), we always have $|S_1| \geq |S_0|$. Thus, a spectrum S is lopsided if and only if $|S_1| > |S_0| > 0$. Again, it follows readily from (1.1) that a balanced spectrum $S = \text{lspec}(m)$ must be of the form $S_1 \cup 2^{\alpha+1}S_1$ where α is the exponent of 2 in the prime factorization of m . Moreover, $m_1 \leq 2^\alpha m_1 = m_0/2$ where $m_i = \max S_i$ ($i = 0, 1$) and it follows from (1.1) that factors of m_1 are in S_1 . Because of that we call a balanced spectrum S *non-excessive* if S_1 is precisely the set of factors of m_1 , otherwise we call S *excessive*. The *exceptional set* of S is defined to be

$$E(S) = \{a \in S_1^2: a > m_0, F_{<a}(m_1 a) = S_1\}$$

where $F_{<b}(c)$ denotes the set of factors of c which are strictly less than b .

Example 2. A few examples should clarify these notions:

- The spectrum of 1, namely $\{1\}$, is an unmixed spectrum. The spectrum of 6, $\{1, 3\}$, is also unmixed.
- The spectrum of 9 is $\{1, 2, 3\}$ which is a lopsided spectrum.
- The spectrum of 3 is $\{1, 2\} = \{1\} \cup 2\{1\}$ which is a balanced spectrum. It is non-excessive with an empty exceptional set.
- The spectrum of 21 is $\{1, 2, 3, 6\} = \{1, 3\} \cup 2\{1, 3\}$ which is again non-excessive. Its exceptional set is $\{9\}$.
- The spectrum of $75 = \{1, 2, 3, 5, 6, 10\}$ is excessive since 3 is not a factor of 5. Its exceptional set is $\{15\}$.
- The spectrum of $175 = \{1, 2, 5, 7, 10, 14\}$ is again excessive. Its exceptional set is $\{25, 35\}$.

We need several results from [5] for our analysis:

Theorem 3 (Theorem 2.2 [5]). *The set of numbers with an unmixed spectrum is $\{2^\alpha k: \alpha \geq 0, k \text{ odd}, 2^{\alpha+1} > k\}$.*

Theorem 4 (Theorem 3.7 [5]). *Suppose $S = \text{lspec}(m)$ is balanced, then either*

- (1) $L(m) = \frac{m_0}{2}E(S)$, if S is excessive; or
- (2) $L(m) = \frac{m_0}{2}(P(S) \cup E(S))$, if S is non-excessive, where $P(S)$ is the set of primes exceeding m_0 .

Theorem 5 (Theorem 4.6 [5]). *Let $m \in \mathbb{N}$ and $S = \text{lspec}(m)$ then*

- $m \in X_1$ if and only if S is lopsided or excessive and $|E(S)| = 1$.
- $m \in X_2$ if and only if S is excessive and $|E(S)| = 2$.
- $m \in X_\omega$ if and only if S is either unmixed or non-excessive.

Theorem 6 (Theorem 4.4 [5]). *A balanced spectrum S is non-excessive if and only if $E(S) = \emptyset$ or $E(S) = \{q^{\mu+1}\}$ where q is the largest prime factor of m_1 and $\mu \geq 1$ is the exponent of q in the prime factorization of m_1 .*

Proposition 1 (Proposition 4.5, 4.9 [5]). *Let S be a balanced spectrum. Then $|E(S)| \leq 2$. Moreover, if $|E(S)| = 2$, then the two elements of $E(S)$ are of the form $p^{\gamma+1} < p^\varepsilon q^\beta$ with $\gamma, \varepsilon, \beta \geq 1$, p, q the two largest primes in S_1 and $p^\gamma \in S_1$.*

3. THE DENSITY OF X_2

For the rest of this article, p and q always stand for odd primes. As usual, $[x]$ denotes the largest integer not exceeding x and $\pi(x)$ denotes the number of primes not exceeding x .

We start by estimating $|X_2(n)|$. Let $m \in X_2(n)$, by Theorem 5 $S = \text{lspec}(m)$ is excessive and $|E(S)| = 2$. By Theorem 4 and Proposition 1, $L(m)$ is of the form $\{\ell p^{\gamma+1}, \ell p^\varepsilon q^\beta\}$ where $\ell = m_0/2$ and $\ell p^{\gamma+1} < \ell p^\varepsilon q^\beta$. In particular, $\ell p^{\gamma+1} \leq m$. Since $p^\gamma \in S_1$, $p^\gamma \leq m_1 \leq \ell$ and since $p^{\gamma+1} \in E(S)$, $2\ell = m_0 < p^{\gamma+1}$. Thus, the smaller of the two elements of $L(m)$ must belong to the following set

$$Y(n) := \{\ell p^{\gamma+1} \leq n : \ell, \gamma \geq 1, 2p^\gamma \leq 2\ell \leq p^{\gamma+1}\} \\ = \left\{ \ell p^{\gamma+1} : \ell, \gamma \geq 1, p^\gamma \leq \ell \leq \min \left\{ \frac{p^{\gamma+1}}{2}, \frac{n}{p^{\gamma+1}} \right\} \right\}.$$

Note that $p^3 \leq p^{2\gamma+1} \leq \ell p^{\gamma+1} \leq n$. So, $3 \leq p \leq \sqrt[3]{n}$ and for each such p , the possible values of γ are $1, 2, \dots, (\lceil \log_p n \rceil - 1)/2$. Therefore, the number of pairs (p, γ) such that $\ell p^{\gamma+1} \in Y(n)$ for some ℓ is at most $\pi(\sqrt[3]{n})(\lceil \log_p n \rceil - 1)/2 \leq \pi(\sqrt[3]{n})(\log_3 n)/2$ which is in $O(\sqrt[3]{n})$ by Chebyshev's upper estimate of $\pi(x)$ [6, Ch. 1 §7 Corollary 2]. For each possible pair of (p, γ) , the number of possible ℓ 's is bounded above by $\min\{p^{\gamma+1}/2, n/p^{\gamma+1}\} \leq \sqrt{(p^{\gamma+1}/2)(n/p^{\gamma+1})} = \sqrt{n/2}$. Thus, we conclude that

$$(3.1) \quad |Y(n)| = O(\sqrt{n})O(\sqrt[3]{n}) = O(n^{5/6}).$$

Since $X_2(n) \subseteq \bigcup \{L(m) : m \in X_2(n)\}$ and each such $L(m)$ has size 2,

$$|X_2(n)| \leq \left| \bigcup \{L(m) : m \in X_2(n)\} \right| \leq 2|Y(n)|.$$

Therefore, it follows from (3.1) that $\delta(X_2) = 0$.

4. THE DENSITY OF X_1

Since \mathbb{N} is the disjoint union of X_1, X_2, X_ω and $\delta(X_2) = 0$, if $\delta(X_\omega)$ exists then so does $\delta(X_1)$, moreover $\delta(X_1) = 1 - \delta(X_\omega)$. By Theorem 5, X_ω is the disjoint union of the set of numbers with an unmixed spectrum (UM) and the set of numbers with a non-excessive spectrum (NE). According to Theorem 3,

$$\text{UM}(n) = \left\{ 2^\alpha k : 0 \leq \alpha \leq \log_2 n, k \text{ odd}, k \leq \min \left\{ 2^{\alpha+1}, n/2^\alpha \right\} \right\}.$$

Note that

$$\min\{2^{\alpha+1}, n/2^\alpha\} = \begin{cases} 2^{\alpha+1} & 0 \leq \alpha \leq \log_2 \sqrt{n/2} \\ n/2^\alpha & \log_2 \sqrt{n/2} < \alpha \leq \log_2 n \end{cases}$$

and k is odd. Therefore, there are at most

$$(4.1) \quad \sum_{0 \leq \alpha \leq \log_2 \sqrt{n/2}} \frac{2^{\alpha+1}}{2} < 2^{\log_2 \sqrt{n/2} + 1} = \sqrt{2n}$$

elements of $\text{UM}(n)$ with $0 \leq \alpha \leq \log_2 \sqrt{n/2}$. Likewise, the number of elements of $\text{UM}(n)$ with $\log_2 \sqrt{n/2} < \alpha \leq \log_2 n$ is at most

$$\begin{aligned} & \sum_{\log_2 \sqrt{n/2} < \alpha \leq \log_2 n} \left\lfloor \frac{\frac{n}{2^\alpha} + 1}{2} \right\rfloor < \sum_{\log_2 \sqrt{n/2} < \alpha \leq \log_2 n} \frac{n}{2^{\alpha+1}} + \frac{1}{2} \\ & < \sum_{0 \leq \alpha \leq \log_2 \sqrt{2n}} \frac{\sqrt{2n}}{2^{\alpha+1}} + \frac{1}{2}(\log_2 \sqrt{2n} + 1) \leq \sqrt{2n} + \frac{1}{2}(\log_2 \sqrt{2n} + 1). \end{aligned}$$

Thus,

$$|\text{UM}(n)| \leq 2\sqrt{2n} + \frac{1}{2}(\log_2 \sqrt{2n} + 1) = O(\sqrt{n})$$

and so $\delta(\text{UM}) = 0$. Next we compute $\delta(\text{NE})$. First note that

Lemma 1. *The set NE is the disjoint union of the following two sets*

$$\begin{aligned} G &= \{\ell p : \ell \geq 1, p > 2\ell\} \text{ and} \\ E &= \{\ell q^{2\mu+1} : \ell, \mu \geq 1, q > 2\ell\}. \end{aligned}$$

Proof. First, G and E are disjoint since the power of the largest prime factor of any element of G is 1 and that of any element of E is at least 3.

Suppose $m \in \text{NE}$. Let $S = \text{lspc}(m)$. By Theorem 4 and 6, either m is of the form $(m_0/2)p$ where $p > m_0$ and so $m \in G$ (with $\ell = m_0/2$) or m is of the form $(m_0/2)q^{\mu+1}$ where q^μ ($\mu \geq 1$) is the power of q in the prime factorization of m_1 and $E(S) = \{q^{\mu+1}\}$. By writing $m_0/2 (= 2^\alpha m_1)$ as ℓq^μ where $(\ell, q) = 1$, we see that m is of the form $\ell q^{2\mu+1}$. Since $q^{\mu+1} \in E(S)$, in particular $q^{\mu+1} > m_0 = 2\ell q^\mu$ and so $q > 2\ell$. This shows that $\text{NE} \subseteq G \cup E$.

Conversely, suppose $m = \ell p \in G$. Let $\ell = 2^\alpha k$ with $\alpha \geq 0$ and k odd, then since $p > 2\ell = 2^{\alpha+1}k$, we have

$$(4.2) \quad k^2 < 2m = 2^{\alpha+1}kp < p^2.$$

Note that the factors of kp are precisely the factors of k together with their p multiples. It follows (1.1) and from the inequalities in (4.2) that the set of factors of k is exactly S_1 . Therefore, $\text{lspc}(m)$ is non-excessive. Now suppose $m = \ell q^{\mu+1} \in E$. Again we write ℓ as $2^\alpha k$ with $\alpha \geq 0$ and k odd, since $q > 2\ell = 2^{\alpha+1}k$, we have

$$(4.3) \quad k^2 q^{2\mu} < 2m = 2^{\alpha+1}kq^{2\mu+1} < q^{2(\mu+1)}.$$

The factors of $kq^{2\mu+1}$ are precisely the factors of kq^μ and their $q^{\mu+1}$ multiples. So again it follows from (1.1) and the inequalities in (4.3) that $\text{lspc}(m)$ is non-excessive. This shows that $G \cup E \subseteq \text{NE}$ and hence concludes the proof. \square

The set E , like X_2 , is also sparsely distributed in \mathbb{N} . If $\ell q^{2\mu+1} \in E(n)$, then by the same analysis for X_2 , we conclude that the number of possible pairs of (q, μ) is bounded above by $\pi(\sqrt[3]{n})(\log_3 n)/2 = O(\sqrt[3]{n})$ and since $\ell < q/2 \leq \sqrt[3]{n}/2$, therefore $|E(n)| = O(n^{2/3})$. Thus, $\delta(E) = 0$.

Now, we estimate $|G(n)|$. Since $G(n) = \{\ell p : 1 \leq \ell \leq \min\{n/p, p/2\}\}$,

$$(4.4) \quad |G(n)| = \sum_{3 \leq p \leq n} \left[\min \left\{ \frac{n}{p}, \frac{p}{2} \right\} \right] = \sum_{3 \leq p < \sqrt{2n}} \frac{p-1}{2} + \sum_{\sqrt{2n} < p \leq n} \left[\frac{n}{p} \right]$$

Again, by Chebyshev's upper estimate of $\pi(x)$,

$$(4.5) \quad \sum_{3 \leq p < \sqrt{2n}} \frac{p-1}{2} \leq \sqrt{\frac{n}{2}} \pi(\sqrt{2n}) = O\left(\frac{n}{\log(n)}\right).$$

We conclude from Mertens' second theorem ([1, Lemma 4.10]), $\sum_{p \leq x} 1/p \sim \log \log x$, that

$$(4.6) \quad \sum_{\sqrt{2n} < p \leq n} \frac{n}{p} \sim n(\log \log n - \log \log \sqrt{2n}) \sim n \log 2.$$

Moreover,

$$(4.7) \quad \sum_{\sqrt{2n} < p \leq n} \frac{n}{p} - \left[\frac{n}{p} \right] \leq \sum_{\sqrt{2n} < p \leq n} 1 = O\left(\frac{n}{\log n}\right).$$

The asymptotic relation in (4.6) and the bound in (4.7) together imply the second term in the sum in Equation (4.4) is asymptotic to $n \log 2$. Putting all these together, we have $\delta(\text{NE}) = \delta(G) + \delta(E) = \delta(G) = \log 2$ and $\delta(X_1) = 1 - \delta(\text{NE}) = 1 - \log 2$. Thus, the probability of guessing a randomly chosen number from its spectrum is

$$\begin{aligned} \delta(X_1) + \frac{\delta(X_2)}{2} &= \delta(X_1) = 1 - \log 2 = \frac{1}{2} - \frac{1}{3} + \frac{1}{4} - \dots \\ &\approx 0.306852819440055. \end{aligned}$$

5. THE AVERAGE ORDER OF LENGTH SPECTRA

It was first proved by LeVeque that [4, Theorem 2]²

$$(5.1) \quad \frac{1}{n} \sum_{m=1}^n |\text{lSpec}(m)| = \frac{\log n}{2} + \frac{2\gamma + \log 2 - 1}{2} + O\left(\frac{1}{\sqrt{n}}\right)$$

where γ is the Euler–Mascheroni constant. We give a quick deduction of this result from Dirichlet’s well-known estimation [1, Corollary 3.32]:

$$(5.2) \quad \sum_{m \leq n} \tau(m) = n \log n + (2\gamma - 1)n + O(\sqrt{n}),$$

where $\tau(m)$ is the number of factors of m . Note that $|\text{lSpec}(m)|$, the number of odd factors of m , can be conveniently expressed as $\tau(m) - \tau(m/2)$ with $\tau(m/2) = 0$ when m is odd. Thus,

$$\begin{aligned} \sum_{m \leq n} |\text{lSpec}(m)| &= \sum_{m \leq n} \tau(m) - \sum_{m \leq n} \tau(m/2) = \sum_{m \leq n} \tau(m) - \sum_{k \leq n/2} \tau(k) \\ &= n \log n + (2\gamma - 1)n - \frac{n}{2} \log \frac{n}{2} - (2\gamma - 1)\frac{n}{2} + O(\sqrt{n}) \\ &= \frac{n}{2} \log n + \frac{n}{2}(2\gamma + \log 2 - 1) + O(\sqrt{n}). \end{aligned}$$

Now, one obtains Equation (5.1) immediately by dividing n on both sides.

6. SOME EXPLICIT BOUNDS

In this last section we provide some loose but nonetheless explicit bounds of the sizes of various sets that appear in this article. The tools involved are some explicit estimations of $\pi(x)$ and $\sum_{p \leq x} 1/p$.

Corollary 5.2 in [2] gives several explicit bounds of $\pi(x)$. One of them is that for $x > 1$,

$$(6.1) \quad \pi(x) \leq (1.2551) \frac{x}{\log(x)}.$$

It follows directly from (6.1) and our estimations on $|X_2(n)|$ and $|E(n)|$ that

$$(6.2) \quad |X_2(n)| \leq 2\sqrt{\frac{n \log_3 n}{2}} \pi(\sqrt[3]{n}) \leq \frac{3.7653}{\sqrt{2} \log 3} n^{5/6} < 2.4235n^{5/6},$$

and that

$$(6.3) \quad |E(n)| \leq \frac{\sqrt[3]{n} \log_3 n}{2} \pi(\sqrt[3]{n}) \leq \frac{3.7653}{4 \log 3} n^{2/3} < 0.8569n^{2/3}.$$

²In [4], $|\text{lSpec}(m)|$, the number of representations of m as sum of consecutive natural numbers, is denoted by $\gamma(m)$.

From Equation (4.4), we have

$$\begin{aligned} \sum_{3 \leq p \leq \sqrt{2n}} \frac{p}{2n} + \sum_{\sqrt{2n} < p \leq n} \frac{1}{p} - \log 2 - \frac{\pi(n)}{n} &\leq \frac{|G(n)|}{n} - \log 2 \\ &\leq \sum_{3 \leq p \leq \sqrt{2n}} \frac{p}{2n} + \sum_{\sqrt{2n} < p \leq n} \frac{1}{p} - \log 2. \end{aligned}$$

Thus,

$$(6.4) \quad \left| \frac{|G(n)|}{n} - \log 2 \right| \leq \sum_{p \leq \sqrt{2n}} \frac{p}{2n} + \left| \sum_{\sqrt{2n} < p \leq n} \frac{1}{p} - \log 2 \right| + \frac{\pi(n)}{n}.$$

Let $c_1 \approx 0.261497$ be the Meissel-Mertens constant. Theorem 1.10 in [7] states that the constant in the Landau symbol in Merten’s second theorem

$$\sum_{p \leq x} \frac{1}{p} = \log \log(x) + c_1 + O\left(\frac{1}{\log x}\right) \quad (x \geq 2)$$

can be chosen $\leq 2(1 + \log 4) < 5$. Therefore,

$$(6.5) \quad \left| \sum_{\sqrt{2n} < p \leq n} \frac{1}{p} - \log 2 \right| \leq \frac{15}{\log n} + \log \log 2n - \log \log n \leq \frac{15 + \log(2)}{\log n}.$$

Finally, it follows from (6.1),(6.4) and the trivial bound $\sum_{p \leq x} p \leq x\pi(x)$ that

$$(6.6) \quad \begin{aligned} \left| \frac{|G(n)|}{n} - \log 2 \right| &\leq \frac{\sqrt{2n}}{2n} \pi(\sqrt{2n}) + \frac{15 + \log(2)}{\log n} + \frac{\pi(n)}{n} \\ &\leq \frac{3(1.2551)}{\log n} + \frac{15 + \log(2)}{\log n} \leq \frac{20}{\log n}. \end{aligned}$$

Acknowledgment. We would like to thank Gordan Savin and A. J. Hildebrand for several useful discussions and email exchanges on this question. We also thank the referee for encouraging us to produce some explicit bounds of our estimations.

REFERENCES

- [1] P. T. Bateman, and H. G. Diamond, *Analytic number theory. An introductory course.* (World Scientific Publishing., Hackensack, NJ, 2004).
- [2] P. Dusart, Explicit estimates of some functions over primes, *Ramanujan J.* **45**(1) (2018), 227–251.
- [3] J. Kadane, and A. O’Hagan, Using finitely additive probability: uniform distributions on the natural numbers, *J. Amer. Statist. Assoc.* **90**(430) (1995) 626–631.
- [4] W. J. LeVeque, On representations as a sum of consecutive integers, *Canadian J. Math.* **2** (1950) 399–405.
- [5] W.Y. Pong, Length Spectra of Natural Numbers, *International Journal of Number Theory.* **5**(6) (2009) 1089-1102.

- [6] G. Tenenbaum, and M. Mendés France, *The prime numbers and their distribution*, Student Mathematical Library, **6**. AMS, Providence, RI, 2000.
- [7] G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, third edition, Graduate Studies in Mathematics, **163**, American Mathematical Society (2015)

CALIFORNIA STATE UNIVERSITY DOMINGUEZ HILLS

Email address: wpong@csudh.edu