AVERAGES OF EXPONENTS IN FACTORING INTEGERS[1]

IVAN NIVEN

For any positive integer $m>1$, say $m=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_r^{\alpha_r}$, define $h(m)=\min(\alpha_1,\alpha_2,\cdots,\alpha_r)$ and $H(m)=\max(\alpha_1,\alpha_2,\cdots,\alpha_r)$. For convenience take $h(1)=1$ and $H(1)=1$. We prove that

$$(1) \qquad \lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}h(j)=1,$$

$$(2) \qquad \lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}H(j)=1+\sum_{k=2}^{\infty}\{1-\zeta(k)^{-1}\},$$

where $\zeta(k)$ is the zeta function. The infinite series on the right side of (2) converges to 0.7 approximately.

We prove something more than (1). Paul Erdös suggested to me in correspondence that it was likely that

$$(3) \qquad \sum_{j=1}^{n}h(j)=n+c\sqrt{n}+o(\sqrt{n}).$$

His conjecture turned out to be correct, with $c=\zeta(3/2)/\zeta(3)$. Formula (3) with this value of c is proved in §1, and of course this also establishes (1). We prove (2) in §2. In §3 we discuss normal order of $h(n)$ and $H(n)$.

1. Proof of (3). Let S be the set of squares of the natural numbers, and let $S(n)$ denote the number of elements of S that do not exceed n. Let T be the set of positive integers m such that $h(m)\geqq 2$. Thus $T\supset S$, and any element m of T not in S can be written uniquely in the form

$$(4) \qquad m=k^2 q_1 q_2 \cdots q_t, \quad (q_1 q_2 \cdots q_t)\,|\,k, \quad t\geqq 1,$$

where $q_1, q_2, \cdots, q_t$ are distinct primes. Now fix $q_1 q_2 \cdots q_t$ and consider the number of elements of T that are $\leqq n$ and have the form (4) for some k. This is the same as the number of squares $\leqq n/(q_1 q_2 \cdots q_t)$ that are divisible by $q_1 q_2 \cdots q_t$. Now for any real $x>0$ the number of positive squares $\leqq x$ that are divisible by $q_1 q_2 \cdots q_t$ is $S(x/(q_1 q_2 \cdots q_t)^2)$. Hence the number of elements of T that are $\leqq n$ and have the form (4) is

(5) $$S(n/(q_1q_2 \cdots q_t)^3).$$

Also we observe that

(6) $$\sqrt{x} - 1 < [\sqrt{x}] = S(x) \leqq \sqrt{x}.$$

Thus if we sum the terms (5) over all subsets $q_1, q_2, \cdots, q_t$ of $p_1, p_2, \cdots, p_r$ where p_r is the rth prime and $p_{r+1} > n$, we see that

(7) $$T(n) = \sum S(n/(p_1^{\beta_1} p_2^{\beta_2} \cdots p_r^{\beta_r}))$$

where the sum is over the 2^r terms with each $\beta_i = 0$ or 3. This gives

$$T(n) \leqq \sum S\left(\frac{n}{p_1^{\beta_1} p_1^{\beta_2} \cdots p_r^{\beta_r}}\right) \leqq \sqrt{n} \cdot \prod_{i=1}^{r} (1 + p_i^{-3/2})$$

by use of (6). But also for $s > 1$ we have $\zeta(s) = \prod (1 - p^{-s})^{-1}$ where the product is over all primes, and so

(8) $$T(n)/\sqrt{n} \leqq \zeta(3/2)/\zeta(3).$$

On the other hand if we choose $N > (p_1p_2 \cdots p_r)^3$ we can rewrite (7) as an inequality with n replaced by N,

$$T(N) > S(N/(p_1^{\beta_1} p_2^{\beta_2} \cdots p_r^{\beta_r})) > \sqrt{N} \prod_{i=1}^{r} (1 + p_i^{-3/2}) - 2^r$$

by use of (6). It follows that

$$T(N)/\sqrt{N} > \zeta(3/2)/\zeta(2) \prod_{i > r} (1 + p_i^{-3/2})^{-1} - 2^r/N.$$

By taking r large enough, then N large enough, we see that this with (8) implies

$$\lim_{n \to \infty} T(n)/\sqrt{n} = \zeta(3/2)/\zeta(3),$$

(9)

$$T(n) = \sqrt{n}\zeta(3/2)/\zeta(3) + o(\sqrt{n}).$$

Next let S_3 denote the set of cubes of the natural numbers, and let T_3 denote the positive integers m such that $h(m) \geqq 3$. Then $S_3 \subset T_3$, and any member m of T_3 not in S_3 can be written uniquely in the form

(10) $$m = k^3 q_1^{\alpha_1} q_2^{\alpha_2} \cdots q_t^{\alpha_t}, (q_1q_2 \cdots q_t)\mid k, t \geqq 1, \alpha_i = 1 \text{ or } 2,$$

where $q_1, q_2, \cdots, q_t$ are distinct primes. Now fix $q_1^{\alpha_1} q_1^{\alpha_2} \cdots q_t^{\alpha_t}$ and consider the number of integers $\leqq n$ that belong to T_3 and have the form (10) for some k. The number of such integers is, by an argument analogous to that leading to (5),

(11) $$S_3(n/(q_1^{3+\alpha_1} q_2^{3+\alpha_2} \cdots q_t^{3+\alpha_t})).$$

Then the analog to (7) is

(12) $$T_3(n) = S_3(n/(p_1^{\beta_1} p_2^{\beta_2} \cdots p_r^{\beta_r}))$$

where the sum is over the 3^r terms with each $\beta_i = 0, 4$, or 5, and r is chosen so that $p_{r+1} > n$. Now $S_3(x) \leq \sqrt[3]{x}$ so that (12) leads to

$$T_3(n) \leq \sum \left\{ \frac{n}{p_1^{\beta_1} p_2^{\beta_2} \cdots p_r^{\beta_r}} \right\}^{1/3}$$

$$< n^{1/3} \prod_p (1 + p^{-4/3}) \prod_p (1 + p^{-5/3})$$

$$= n^{1/3}\zeta(4/3)\zeta(5/3)\{\zeta(8/3)\zeta(10/3)\}^{-1},$$

(13) $$T_3(n) = O(n^{1/3}).$$

A similar argument leads to $T_k(n) = O(n^{1/k})$ for any integer $k \geq 3$, where T_k denotes the positive integers m such that $h(m) \geq k$. However, we need this only for $k = 3$ in this paper.

Now we examine the positive integers $h(1), h(2), \cdots, h(n)$. The number of these that exceed 1 is $T(n)$; the number that exceed 2 is $T_3(n)$. Also we note that

$$\max\{h(1), h(2), \cdots, h(n)\} = [\log_2 n].$$

It follows that

$$n + T(n) \leq h(1) + h(2) + \cdots + h(n) \leq n + T(n) + T_3(n) \cdot \log_2 n.$$

This with (9) and (13) establishes (3) with $c = \zeta(3/2)/\zeta(3)$.

2. **Proof of** (2). Let Q_k be the set of k-power free integers, i.e. the set of positive integers m such that $H(m) \leq k-1$. Now the number of integers m satisfying $1 \leq m \leq n$ and $H(m) = k-1$ is $Q_k(n) - Q_{k-1}(n)$. Also for $n \geq 2$ the maximum of $H(1), H(2), \cdots, H(n)$ is $[\log_2 n]$. Thus we see that

$$\sum_{i=1}^n H(i) = \sum_{k=2}^{j+1} (k-1)\{Q_k(n) - Q_{k-1}(n)\}, \quad j = [\log_2 n].$$

But $Q_{j+1}(n) = n$, so this can be written as

(14) $$\sum_{i=1}^n H(i) = jn - \sum_{k=2}^j Q_k(n), \quad j = [\log_2 n].$$

Next we prove that if r satisfies $p_r^k \leq n \leq p_{r+1}^k$ then

$$Q_k(n) = \sum (-1)^{\alpha_1 + \alpha_2 + \cdots + \alpha_r} [n/(p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r})] \tag{15}$$

where the sum contains 2^r terms obtained by taking each $\alpha_i = 0$ or k. This can be established by interpreting $[n/s]$ as the number of integers $\leq n$ that are divisible by s. Thus the right side of (15) can be interpreted as the number of integers from 1 to n, first with those divisible by p_1^k deleted, then with those divisible by p_2^k deleted, $\cdots$, then with those divisible by p_r^k deleted, then with those divisible by $p_1^k p_2^k$ counted back in, and so on. Hence (15) follows by use of the inclusion-exclusion principle.

Next, equation (15) can be written as

$$Q_k(n) = \sum \mu(d)[n/d^k], \tag{16}$$

where the sum is over all positive divisors d of $p_1 p_2 \cdots p_r$. In this sum any term for which $d^k > n$ has $[n/d^k] = 0$ so we can take the sum in (16) over all positive integers d satisfying $d^k \leq n$. It is well known that

$$\zeta(k)^{-1} = \prod_p (1 - p^{-k}) = \sum_{d=1}^{\infty} \mu(d)/d^k.$$

This with (16) gives

$$n\zeta(k)^{-1} - Q_k(n) = \sum_{d=1}^{\infty} \frac{\mu(d) \cdot n}{d^k} - \sum_{d^k \leq n} \mu(d) \left[\frac{n}{d^k} \right]$$

$$= \sum_{d^k \leq n} \mu(d) \left\{ \frac{n}{d^k} - \left[\frac{n}{d^k} \right] \right\} + \sum_{d^k > n} \frac{\mu(d) \cdot n}{d^k},$$

$$\left| n\zeta(k)^{-1} - Q_k(n) \right| = \sum_{d^k \leq n} \left\{ \frac{n}{d^k} - \left[\frac{n}{d^k} \right] \right\} + \sum_{d^k > n} \frac{n}{d^k}. \tag{17}$$

Consider the two sums on the right side of (17). The first is less than $n^{1/k}$. As to the second sum, its first term is less than 1, and all other terms are bounded by

$$n \int_{n^{1/k}}^{\infty} x^{-k} dx = n^{1/k}/(k-1) < n^{1/k}$$

since $k \geq 2$. Hence (17) can be written

$$\left| n\zeta(k)^{-1} - Q_k(n) \right| \leq n^{1/k} + 1 + n^{1/k} < 3n^{1/k}.$$

This with (14) implies

$$\left| n^{-1} \sum_{i=1}^{n} H(i) - 1 - \sum_{k=2}^{j} \left\{ 1 - \varsigma(k)^{-1} \right\} \right|$$

$$= \left| j - n^{-1} \sum_{k=2}^{j} Q_k(n) - 1 - \sum_{k=2}^{j} \left\{ 1 - \varsigma(k)^{-1} \right\} \right|$$

$$= n^{-1} \left| \sum_{k=2}^{j} \left\{ n\varsigma(k)^{-1} - Q_k(n) \right\} \right|$$

$$\leqq n^{-1} \sum_{k=2}^{j} \left| n\varsigma(k)^{-1} - Q_k(n) \right|$$

$$\leqq n^{-1} \sum_{k=2}^{j} 3n^{1/k} \leqq n^{-1}(3n^{1/2}) \log_2 n.$$

As n tends to infinity so does $j = [\log_2 n]$, and hence we have (2).

3. **The normal order of** $h(n)$ **and** $H(n)$. We use the definition of normal order given in [1, p. 356]. It is clear from equation (1) that $h(n)$ has normal order 1, since $h(n) = 1$ for almost all positive integers n. On the other hand $H(n)$ has no normal order, because it is an integer-valued function whose average order, by equation (2), lies between 1 and 2.

REFERENCE

1. G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*, 4th ed., Oxford Univ. Press, London, 1960.

UNIVERSITY OF OREGON