ON THE SUM $\sum_{k=1}^{x} d(f(k))$

P. Erdös*.

1. Let $d(n)$ denote the number of divisors of a positive integer n, and let $f(k)$ be an irreducible polynomial of degree l with integral coefficients. We shall suppose for simplicity that $f(k) > 0$ for $k = 1, 2, \ldots$. In the present paper we prove the following result.

Theorem. *There exist positive constants* c_1 *and* c_2 *such that*

$$c_1 x \log x < \sum_{k=1}^{x} d(f(k)) < c_2 x \log x \tag{1}$$

for $x \geqslant 2$.

Throughout the paper x is supposed to be large, and $c_1, c_2, \ldots$ denote positive constants which are independent of x but may depend on the polynomial f.

The lower bound in (1) is not difficult to prove, and is in fact known†. It would not be hard to show that

$$\sum_{k=1}^{x} d_x(f(k)) = c_3 x \log x + o(x \log x), \tag{2}$$

where $d_x(n)$ denotes the number of divisors of n which do not exceed x. In §§4 and 5 we give a proof that the sum in (2) is greater than $c_1 x \log x$, from which the lower bound in (1) follows.

The upper bound in (1) is much more difficult to prove, since it is not easy to find an upper estimate for $d(f(k))$ in terms of $d_x(f(k))$. It is possible to do so if $l = 2$, and in this case Bellman and Shapiro‡ have proved that

$$\sum_{k=1}^{x} d(f(k)) = c_4 x \log x + o(x \log x). \tag{3}$$

Very likely (3) holds also if $l > 2$, but I cannot prove this.

The method used to prove the upper bound in (1), if combined with Brun's method, would enable one to prove that

$$\sum_{p \leqslant x} d(f(p)) = O(x).$$

This answers a question proposed in Bellman's paper (*loc. cit.*).

---

* Received 3 January, 1951; read 18 January, 1951.

† Bellman, *Duke Math. J.*, 17 (1950), 159–168.

‡ This result is unpublished.

2. We need several lemmas for the proof of the upper bound in (1).

Lemma 1. $$\sum_{k=1}^{x} \{d(f(k))\}^2 < x(\log x)^{c_5}.$$

This result is due to van der Corput*.

Lemma 2. *Let $k_1, k_2, \ldots, k_t$ be distinct positive integers, all less than x, and suppose that $t < x(\log x)^{-c_5}$. Then*

$$\sum_{i=1}^{t} d(f(k_i)) < x.$$

This follows at once from Lemma 1 by Schwarz's inequality:

$$\sum_{i=1}^{t} d(f(k_i)) \leqslant \left\{t \sum_{i=1}^{t} \{d(f(k_i))\}^2\right\}^{\frac{1}{2}} < x.$$

Let $\rho(a)$ denote the number of solutions of

$$f(k) \equiv 0 \pmod{a}, \quad 0 \leqslant k < a.$$

Let D denote the discriminant of the polynomial $f(k)$. If p is any prime, we use the notation $p^\sigma \| D$ to express the fact that p^σ is the greatest power of p which divides D.

Lemma 3. *We have*

(i) $\rho(ab) = \rho(a)\rho(b)$ *if* $(a, b) = 1$;

(ii) $\rho(p^\alpha) \leqslant l$ *if* $p \nmid D$;

(iii) $\rho(p^\alpha) = \rho(p^{2\sigma+1})$ *if* $p^\sigma \| D$ *and* $\alpha > 2\sigma$;

(iv) $\rho(p^\alpha) \leqslant c_6$ *always.*

The first two results are well known and trivial. The third result was given by Nagell†, who also showed that (iv) is valid with $c_6 = lD^2$. As it stands, with an unspecified c_6, (iv) follows from (ii) and (iii).

Lemma 4. *Suppose that $1 \leqslant u \leqslant x$. Let N denote the number of integers k satisfying*

$$f(k) \equiv 0 \pmod{u}, \quad 1 \leqslant k \leqslant x.$$

Then $$\frac{x}{2u}\rho(u) \leqslant N \leqslant \frac{2x}{u}\rho(u).$$

The proof is immediate, since obviously

$$\left[\frac{x}{u}\right]\rho(u) \leqslant N \leqslant \left(\left[\frac{x}{u}\right]+1\right)\rho(u).$$

* *Proc. K. Neder. Akad. van. Wet., Amsterdam*, 42 (1939), 547–553.

† "Généralisation d'un théorème de Tchebycheff", *Journal de Math.*, 8e série, Tome IV 1921.

LEMMA 5. *There exists c_7 such that the number of positive integers $k \leqslant x$ for which $f(k)$ is divisible by a prime power p^a, with $a > 1$ and*

$$p^a > (\log x)^{c_7}, \tag{4}$$

is $o\big(x(\log x)^{-c_5}\big)$.

By Lemma 4 and Lemma 3, the number of integers k in question is less than

$$2x \sum_{p,\,a} \frac{\rho(p^a)}{p^a} < 2c_6 x \sum_{p,\,a} p^{-a} < 4c_6 x \left\{ \sum_{p \leqslant (\log x)^{\frac{1}{2}c_7}} (\log x)^{-c_7} + \sum_{p > (\log x)^{\frac{1}{2}c_7}} p^{-2} \right\},$$

and the result follows on taking $c_7 > 2c_5$.

We define $\bar{x}$ by

$$\bar{x} = x^{(\log\log x)^{-2}}.$$

LEMMA 6. *Let $f(k)$ be factorized into prime powers as $f(k) = \Pi\, p^a$, for each positive integer k. Then the number of values of $k \leqslant x$ for which*

$$\prod_{p < \bar{x}} p^a \geqslant x^{\frac{1}{2}} \tag{5}$$

is $o\big(x(\log x)^{-c_5}\big)$.

Consider first those values of k for which there is one at least of the prime powers p^a occurring in the product (5) which satisfies $p^a \geqslant \bar{x}$. Each such prime power satisfies (4), and $a > 1$. Hence the number of values of k of this kind is $o\big(x(\log x)^{-c_5}\big)$ by Lemma 5.

There remain those integers k for which every prime power in (5) is less than $\bar{x}$. By (5), every such value of $f(k)$ has at least $\frac{1}{8}(\log\log x)^2$ distinct prime factors, whence

$$d\big(f(k)\big) > 2^{\frac{1}{8}(\log\log x)^2} > (\log x)^{2c_5}.$$

By Lemma 1, the number of such integers is $O\big(x(\log x)^{-3c_5}\big)$, and this completes the proof of the present lemma.

LEMMA 7. *We have*

$$\sum_{p \leqslant x} \rho(p) = \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right), \tag{6}$$

$$\sum_{p \leqslant x} \frac{\rho(p)}{p} = \log\log x + c_8 + o(1). \tag{7}$$

The first result follows from the prime ideal theorem*, which implies that

$$\sum_{N\mathfrak{p} \leqslant x} 1 = \frac{x}{\log x} + O\left(\frac{x}{\log^2 x}\right),$$

* See, for example, Landau, *Algebraische Zahlen*, 111.

where the summation is extended over prime ideals $\mathfrak{p}$ in the field $k(\theta)$ generated by a root θ of the equation $f(\theta)=0$, and N denotes the norm. We have $N\mathfrak{p}=p^f$, where p is a rational prime and f a positive integer. We may ignore rational primes p which divide D, since they contribute only $O(1)$ to the sum in (6) and to the sum last written above. If $f=1$, the same rational prime p arises as $N\mathfrak{p}$ for $\rho(p)$ different prime ideals $\mathfrak{p}$. When $f>1$, the same p arises from at most l prime ideals, and the corresponding part of the last sum is at most

$$l\Big(\sum_{p^2\leqslant x} 1+\sum_{p^3\leqslant x} 1+\ldots\Big),$$

which is $O(x^{\frac{1}{2}}\log x)$. This proves (6), and (7) follows from (6) by partial summation.

Lemma 8. *For sufficiently large y, we have*

$$\sum_{y<p<y^2}\frac{\rho(p)}{p}<1.$$

This is obvious from (7).

Lemma 9. *We have*

$$\prod_{p\leqslant x}\left\{1+\frac{\rho(p)}{p}+\frac{\rho(p^2)}{p^2}+\ldots\right\}<c_9\log x.$$

This follows from (7) on using (iv) of Lemma 3 for $\rho(p^a)$ when $a>1$. For the logarithm of the above product is

$$\sum_{p\leqslant x}\frac{\rho(p)}{p}+O(1),$$

whence the result.

3. We now come to the proof of the upper estimate in (1). By Lemmas 2, 5 and 6, we have

$$\sum_{k=1}^{x} d\big(f(k)\big)=\Sigma_1 d\big(f(k)\big)+O(x), \tag{8}$$

where in Σ_1 the variable of summation k is restricted to positive integers not exceeding x which satisfy the following two conditions:

$$f(k)\not\equiv 0 \pmod{p^a} \text{ if } a>1 \text{ and } p^a>(\log x)^{c_7}, \tag{9}$$

$$\prod_{p<x} p^a<x^{\frac{1}{2}}, \tag{10}$$

the last product being extended over the prime powers composing $f(k)$.

We can obviously restrict ourselves to values of k for which $f(k) > x$. If $f(k) = p_1^{\alpha_1} \dots p_s^{\alpha_s}$, we define j by

$$p_1^{\alpha_1} \dots p^{\alpha_j} \leqslant x < p_1^{\alpha_1} \dots p_{j+1}^{\alpha_{j+1}}. \tag{11}$$

Put

$$a_k = p_1^{\alpha_1} \dots p_j^{\alpha_j}, \quad b_k = f(k)/a_k. \tag{12}$$

Write

$$\Sigma_1 d\big(f(k)\big) = \Sigma_2 d\big(f(k)\big) + \Sigma_3 d\big(f(k)\big), \tag{13}$$

where in Σ_2 we take those values of k satisfying (9) and (10) for which $p_{j+1} \leqslant x^{1/32}$, and in Σ_3 we take those for which $p_{j+1} > x^{1/32}$.

First we estimate the sum Σ_3. Any prime factor of b_k is greater than $x^{1/32}$, and since $f(k) < x^{l+1}$ it follows that the total number of prime factors of b_k (multiple factors being counted multiply) is less than $32(l+1)$. Hence

$$d(b_k) < 2^{32(l+1)}. \tag{14}$$

Also, since $a_k \leqslant x$, we obviously have

$$d(a_k) \leqslant d_x\big(f(k)\big). \tag{15}$$

By (14) and (15),

$$\Sigma_3 d\big(f(k)\big) < 2^{32(l+1)} \sum_{k=1}^{x} d_x\big(f(k)\big).$$

The sum on the right here is the number of solutions of $f(k) \equiv 0 \pmod{u}$ in integers k and u satisfying $1 \leqslant k \leqslant x$, $1 \leqslant u \leqslant x$. By Lemma 4, this number is less than

$$2x \sum_{u=1}^{x} \frac{\rho(u)}{u}.$$

Using Lemma 3 (i) and Lemma 9, we obtain

$$\Sigma_3 d\big(f(k)\big) < c_{10}\, x \prod_{p \leqslant x} \left\{1 + \frac{\rho(p)}{p} + \frac{\rho(p^2)}{p^2} + \dots\right\} < c_{11}\, x \log x. \tag{16}$$

We have now to estimate the sum Σ_2. For each k in this sum we have $p_{j+1} \leqslant x^{1/32}$. We now prove that

$$p_{j+1}^{\alpha_{j+1}} \leqslant x^{1/2}. \tag{17}$$

In fact, if this were false there would be an exponent β such that $1 < \beta \leqslant \alpha_{j+1}$ for which $x^{1/32} < p_{j+1}^{\beta} \leqslant x^{1/2}$, and this would contradict (9).

It follows from (11), (12) and (17) that for the k in Σ_2,

$$a_k > \frac{x}{p_{j+1}^{\alpha_{j+1}}} > x^{1/2}. \tag{18}$$

In view of (10), we have now

$$\bar{x} \leqslant p_{j+1} \leqslant x^{1/32}. \tag{19}$$

We write

$$\Sigma_2 d\big(f(k)\big) = \sum_r \Sigma_2^{(r)} d\big(f(k)\big), \tag{20}$$

where in $\Sigma_2^{(r)}$ the prime factor p_{j+1} satisfies

$$x^{1/(r+1)} \leqslant p_{j+1} < x^{1/r}.$$

By (19), the values of r in question satisfy $32 \leqslant r \leqslant (\log\log x)^2$, since $\bar{x} = x^{(\log\log x)^{-2}}$. For any k in $\Sigma_2^{(r)}$, the total number of prime factors of b_k is less than $(l+1)(r+1)$, and it follows that

$$\Sigma_2^{(r)} d\big(f(k)\big) < 2^{(l+1)(r+1)} \Sigma_2^{(r)} d(a_k). \tag{21}$$

Since at least half of the divisors of a_k are greater than or equal to $\sqrt{a_k}$, it follows from (18) that

$$d(a_k) \leqslant 2d^+(a_k), \tag{22}$$

where $d^+(m)$ denotes the number of divisors of m that are $\geqslant x^{\frac{1}{2}}$. It follows from (9) and (10) that all the divisors of a_k that are $\geqslant x^{\frac{1}{2}}$ are included in a set of numbers $n_1^{(r)}, n_2^{(r)}, \ldots$ satisfying

(i) $x^{\frac{1}{2}} \leqslant n_j^{(r)} \leqslant x$,

(ii) if $p \mid n_j^{(r)}$ then $p < x^{1/r}$, and $\rho(p) > 0$,

(iii) if $p^\alpha \| n_j^{(r)}$ and $\alpha > 1$ then $p^\alpha \leqslant (\log x)^{c_7}$,

(iv) $\Pi' p^\alpha < x^{\frac{1}{4}}$,

where the last product is extended over $p^\alpha \| n_j^{(r)}, p < \bar{x}$. The sum $\Sigma_2^{(r)} d^+(a_k)$ does not exceed the number of solutions in k and j of $f(k) \equiv 0 \pmod{n_j^{(r)}}$. Hence, by (21), (22) and Lemma 4,

$$\Sigma_2^{(r)} d\big(f(k)\big) < 2^{(l+1)(r+1)+2}\, x \sum_j \frac{\rho(n_j^{(r)})}{n_j^{(r)}}. \tag{23}$$

We have now to estimate the last sum. Let $I_t^{(r)}$ denote the interval $(x^{1/(r2^{t+1})}, x^{1/(r2^t)})$, where $t = 0, 1, \ldots, Z$, and Z is the largest integer for which $r2^Z \leqslant (\log\log x)^2$. Any number $n_j^{(r)}$ must have at least $N_t^{(r)}$ prime factors in at least one of these intervals, where

$$N_t^{(r)} = \left[\frac{r(t+1)}{32}\right] + 1.$$

For a prime in one of these intervals is at least equal to $x^{\frac{1}{2}(\log\log x)^{-2}}$, and so can only divide $n_j^{(r)}$ to the first power, by condition (iii) above. Every prime greater than $\bar{x}$ comes in one of the intervals, and if the above statement were false we should have

$$n_j^{(r)} < (\Pi' p^a) \prod_t (x^{1/(r2^t)})^{r(t+1)/32}.$$

Since $\sum_{t=0}^{\infty} (t+1)/2^t = 4$, this gives $n_j^{(r)} < x^{\frac{1}{8}}$, contrary to (i) above.

We define s to be the least integer for which $n_j^{(r)}$ has at least $N_s^{(r)}$ prime factors in the interval $I_s^{(r)}$, and write

$$\sum_j \frac{\rho(n_j^{(r)})}{n_j^{(r)}} = \sum_s \sum_{j(s)} \frac{\rho(n_j^{(r)})}{n_j^{(r)}}, \tag{24}$$

where the inner sum on the right is extended over those values of j for which s has a prescribed value. We put $n_j^{(r)} = uv$, where u is composed entirely of the prime factors of $n_j^{(r)}$ in the interval $I_s^{(r)}$, and v of the other prime factors. As already observed, u is square-free. By the multiplicative property of ρ [Lemma 3 (i)], we have

$$\sum_{j(s)} \frac{\rho(n_j^{(r)})}{n_j^{(r)}} \leqslant \Big(\sum_u \frac{\rho(u)}{u}\Big)\Big(\sum_v \frac{\rho(v)}{v}\Big), \tag{25}$$

where the summation on the right is extended over all u and v which divide any $n_j^{(r)}$ for which s has the prescribed value. Since u is square-free and has at least $N = N_s^{(r)}$ prime factors, we have

$$\sum_u \frac{\rho(u)}{u} \leqslant \frac{1}{N!}\Big\{\sum_p \frac{\rho(p)}{p}\Big\}^N,$$

the summation being over primes p in $I_s^{(r)}$. By Lemma 8, it follows that

$$\sum_u \frac{\rho(u)}{u} < \frac{1}{N!}, \quad N = N_s^{(r)}. \tag{26}$$

For the sum over v, we use the simple estimate (Lemma 9)

$$\sum_v \frac{\rho(v)}{v} \leqslant \prod_{p \leqslant x} \Big\{1 + \frac{\rho(p)}{p} + \frac{\rho(p^2)}{p^2} + \dots\Big\} < c_9 \log x. \tag{27}$$

From (25), (26), (27),

$$\sum_{j(s)} \frac{\rho(n_j^{(r)})}{n_j^{(r)}} < \frac{c_9 \log x}{N_s^{(r)}!}. \tag{28}$$

By the definition of $N_s^{(r)}$ we have, since $r \geqslant 32$,

$$\Big[\frac{r}{32}\Big] < N_0^{(r)} < N_1^{(r)} < \dots.$$

Hence, by (24) and (28),

$$\sum \frac{\rho(n_j^{(r)})}{n_j^{(r)}} < c_9 \log x \sum_{s=0}^{\infty} \frac{1}{N_s^{(r)}!} < 2c_9 \log x \Big(\Big[\frac{r}{32}\Big]!\Big)^{-1}. \tag{29}$$

Finally, by (20), (23), (29),

$$\Sigma_2 d\big(f(k)\big) < 2c_9 x \log x \sum_{r=32}^{\infty} 2^{(l+1)(r+1)+2}\left(\left[\frac{r}{32}\right]!\right)^{-1}$$
$$< c_{12} x \log x. \tag{30}$$

The upper bound in (1) follows from (8), (13), (16) and (30).

4. To prove the lower bound in (1), we use another lemma.

Lemma 10. *For large y, we have*

$$\sum_{k=1}^{y} \rho(k) > c_{13} y.$$

Let $\zeta_K(s)$ denote the zeta-function of the field $K = k(\theta)$, so that

$$\zeta_K(s) = \sum_{\mathfrak{a}} (N\mathfrak{a})^{-s} = \prod_{\mathfrak{p}} \{1-(N\mathfrak{p})^{-s}\}^{-1},$$

where the sum is extended over all the ideals $\mathfrak{a}$ of K, and the product over the prime ideals of K. It is well known* that provided $p \nmid D$, the factorization of a rational prime p as

$$p = \mathfrak{p}_1 \mathfrak{p}_2 \ldots, \text{ where } N\mathfrak{p}_1 = p^{r_1}, \text{ etc.},$$

corresponds to the factorization

$$f(x) \equiv f_1(x) f_2(x) \ldots \pmod{p},$$

where $f_1(x), f_2(x), \ldots$ are irreducible (mod p), and are of degrees $r_1, r_2, \ldots$. Obviously p has at most l prime ideal factors, provided $p \nmid D$.

We split the product defining $\zeta_K(s)$ into three parts. The first part arises from the prime ideals $\mathfrak{p}$ for which $N\mathfrak{p} = p^g$ with $g > 1$. This is easily seen to be a regular function of $s = \sigma + it$ for $\sigma > \frac{1}{2}$. The second part arises from the prime ideals $\mathfrak{p}$ for which $p \mid D$, and is regular for $\sigma > 0$. The third part arises from the prime ideals $\mathfrak{p}$ for which $N\mathfrak{p} = p$ and $p \nmid D$. The number of such prime ideals corresponding to a given p is the number of linear polynomials among $f_1(x), f_2(x), \ldots$, and so is $\rho(p)$. Hence

$$\zeta_K(s) = \phi(s) \prod_{p \nmid D} (1-p^{-s})^{-\rho(p)},$$

where $\phi(s)$ is regular for $\sigma > \frac{1}{2}$.

Define $\rho'(n)$ by

$$\sum_n \rho'(n) n^{-s} = \prod_{p \nmid D} \left\{1 + \frac{\rho(p)}{p^s} + \frac{\rho(p^2)}{p^{2s}} + \ldots\right\}$$
$$= \prod_{p \nmid D} \left\{1 + \frac{\rho(p)}{p^s - 1}\right\},$$

* See Dedekind, *Gesammelte math. Werke*, I, 202–232.

since $\rho(p^{s}) = \rho(p)$ when $p \nmid D$. Since, for large p,

$$\log\left(1+\frac{\rho(p)}{p^{s}-1}\right) - \log\left(1-\frac{1}{p^{s}}\right)^{-\rho(p)} = O\left(\frac{\rho(p)}{p^{2\sigma}}\right),$$

it follows that

$$\sum_{p \nmid D} \log\left(1+\frac{\rho(p)}{p^{s}-1}\right) + \sum_{p \nmid D} \rho(p) \log\left(1-\frac{1}{p^{s}}\right)$$

is regular for $\sigma > \frac{1}{2}$. Taking exponentials, we obtain

$$\sum_{n} \rho'(n)\, n^{-s} = \zeta_K(s)\, \Psi(s),$$

where $\Psi(s)$ is regular for $\sigma > \frac{1}{2}$.

Write $\qquad \zeta_K(s) = \sum a_n n^{-s}, \quad \Psi(s) = \sum \alpha_n n^{-s}.$

Then $\qquad \sum_{n=1}^{y} \rho'(n) = \sum_{cd \leqslant y} \alpha_c a_d = \sum_{c=1}^{y} \alpha_c \sum_{d \leqslant y/c} a_d.$

It is well known* that

$$\sum_{n=1}^{z} a_n = c_{13} z + O(z^{1-\delta}),$$

where $\delta > 0$. Since $\sum \alpha_c c^{-1+\delta}$ converges absolutely, we easily deduce that

$$\sum_{n=1}^{y} \rho(n) > \sum_{n=1}^{y} \rho'(n) = c_{14} y + O(y^{1-\delta}).$$

This proves Lemma 10.

5. To prove the lower bound in (1), it suffices to prove that

$$\sum_{k=1}^{x} d_x\big(f(k)\big) > c_1 x \log x.$$

The sum on the left is the number of solutions of $f(k) \equiv 0 \pmod{y}$ with $1 \leqslant k \leqslant x$, $1 \leqslant y \leqslant x$. By Lemmas 4 and 10, and partial summation, it is greater than

$$\tfrac{1}{2} x \sum_{y=1}^{x} \frac{\rho(y)}{y} > \tfrac{1}{2} x \sum_{y=1}^{x} y^{-2} \sum_{k=1}^{y} \rho(k)$$

$$> \tfrac{1}{2} c_{13} x \sum_{y=1}^{x} y^{-1} > c_1 x \log x.$$

The University,
Aberdeen.

* Landau, *Algebraische Zahlen*, 131.