ON THE DISTRIBUTION OF DIVISORS OF INTEGERS IN THE RESIDUE CLASSES (MOD d).

By

P. ERDÖS (Haifa)

Throughout this paper k and l will denote integers satisfying $0<l<k$, $(l, k)=1$. Denote by $f(x; k, l)$ the number of integers $n<x$ which have a divisor t satisfying $t\equiv l \pmod k$ and $F(x; k)$ denotes the number of integers $n<x$ which have a divisor $\equiv l \pmod k$ for every l. Clearly $F(x; k)\leqq f(x; k, l)$. It is easy to prove that for fixed k

$$F(x; k)=x+o(x).$$

We obtain very much more difficult questions if k tends to infinity together with x. We are going to prove the following [1].

Theorem 1. *Let $\varepsilon>0$ be fixed but arbitrary, $k<2^{(1-\varepsilon)\log\log x}$ Then uniformly in k.*

$$F(x; k)=x+o(x). \tag{1}$$

In other words if $k<2^{(1-\varepsilon)\log\log x}$ then almost all numbers have a divisor in every residue class $l \pmod k$. A well known theorem of Hardy—Ramanujan [3] states that for almost all numbers $n<x$

$$d(n)<2^{(1+o(1))\log\log x} \tag{2}$$

(2) easily implies that $F(x, k)=o(x)$ if $k>2^{(1+\varepsilon)\log\log x}$. Thus in some sense our Theorem is best possible. But in fact we will outline the proof of the following stronger.

Theorem 2. *Let $k>2^{(1+\varepsilon)\log\log x}$. Then uniformly in k and l,*

$$f(x; k, l)=\frac{x}{l}+o(x).$$

It seems likely that the following stronger result also holds. Assume that $k=\log\log x+c(\log\log x)^{1/2}$.

Then

$$F(x;k)=(1+o(1))\,x\frac{1}{(2\pi)^{1/2}}\int_{c}^{\infty}e^{-y^2/2}\,dy. \tag{3}$$

I hope to return to (3) in a subsequent paper.

I was lead to these questions in connection with the following problem of Sivasankaranarayana Pillai: Denote by $Q(x)$ the number of integers $n<x$ which have no divisor of the form $p(kp+1)$. Pillai proved

$$Q(x)<cx/\log\log\log x.$$

Using Theorem 1 we will outline the proof of

Theorem 3. *Let C be Euler's constant. We have*

$$Q(x)=(1+o(1))\,\frac{e^{-C}x}{\log 2\log\log x}.$$

Denote by $d(n;k,l)$ the number of divisors of n which are $\equiv l \pmod k$. We will outline the proof of

Theorem 4. *Let $k<2^{1/2(1-\varepsilon)\log\log x}$. Then for every $\eta>0$, we have for every l_1 and l_2, for all but $o(x)$ integers $n<x$*

$$1-\eta<d(n;k,l_1)/d(n;k,l_2)<1+\eta.$$

Throughout this paper $c, c_1,\ldots$ will denote positive absolute constants, $\eta, \eta_1,\ldots$ will denote small but fixed positive numbers which usually will depend on ε and on the choice of previous η's. $S, S(t)\ldots$ will denote sets of integers $n<x$ and $N(S)$ will denote the number of integers in the set S, p and q will denote primes. $\nu(n)$ denotes the number of distinct prime factors of n.

It will suffice to prove Theorem 1 for the k satisfying

$$2^{(1-\varepsilon)\log\log x-1}<k\leqq 2^{(1-\varepsilon)\log\log x}. \tag{4}$$

To see this observe that every $k_1<2^{(1-\varepsilon)\log\log x-1}$ has a multiple satisfying (4) and if we prove Theorem 1 for k it follows for all $k_1 \mid k$. Henceforth we will always assume that k satisfies (4) (i.e. in the proof of Theorem 1).

The principal tools needed for the proof of Theorem 1 is a recent result of Rényi and myself [2], a theorem of Walfisz [6] on the distribution of primes in arithmetical progressions and a

theorem of Hardy—Ramanujan—Turán [3] [4], we also will need Brun's method.

Rényi and I [2] proved that if G is an Abelian group of m elements and if we choose $t>(1+\eta_1)\log m/\log 2$ distinct elements $a_1,\dots, a_t$ of G, then for all but $o\left(\binom{m}{t}\right)$ of these choices, all elements of G can be written in the form

$$\prod_{i=1}^{t} a_i^{\delta_i},\quad \delta_i=0 \text{ or } 1.$$

This theorem immediately implies

Lemma 1. *Let $t>(1+\eta_1)\log\varphi(k)/\log 2$, $(\eta_1<\varepsilon)$. Then for all but $o\left(\binom{\varphi(k)}{t}\right)$ choices of the distinct residues $l_1,\dots, l_t$ every l (mod k) is of the form*

$$\prod_{i=1}^{t} l_i^{\delta_i},\qquad \delta_i=0 \text{ or } 1.$$

Demote by I_x the interval

$$(e^{(\log x)^{\eta_2}},\ x^{1/(\log\log x)^3})$$

where

$$1-\eta_2>(1+\eta_1)(1-\varepsilon). \tag{5}$$

Denote by $\pi(y, d, l)$ the number of primes $p<y$, $p\equiv l$ (mod d). Let $A>0$ be any number, assume $d<(\log y)^A$. A theorem of Siegel—Walfisz [6] states

$$\pi(y, d, l)=\frac{1}{\varphi(d)}\int_2^y \frac{dz}{\log z}+O(ye^{-c(\log y)^{1/2}}). \tag{6}$$

The error term is uniform in y, d and l but may depend on A

Lemma 2. *For every l and k*

$$\sum=\sum_{p\equiv l(\mathrm{mod}\ k)}{}' \frac{1}{p}=\left(1+O\left(\frac{1}{\log x}\right)\right)\sum_{p\equiv l(\mathrm{mod}\ k)}{}' \frac{1}{p} \tag{7}$$

the dash in the summation indicates that p is in I_x. The error term is uniform in k and l.

It is easy to see that if k satisfies (4) and y is in I_x then (6)

is satisfied. (7) then follows by elementary and straightforward estimations which we supress. It would be easy to give an explicit asymptotic formula for Σ, but we do not need this.

Lemma 3. *Let $t<2\log\log x$, $l_1,\dots,l_t$ distinct residues* mod k. *Then*

$$\sum'' \frac{1}{p_1\cdots p_t}=\left(1+O\left(\frac{1}{(\log x)^{1/2}}\right)\right)\Sigma^t$$

in $\sum''$ p_i runs through the primes of I_x satisfying $p_i\equiv l_i \pmod k$. The error term is uniform in k and the l's.

Lemma 3 follows immediately from Lemma 2 by a simple computation (k satisfies (4)!).

Put

$$f(n)=\prod_{p^\alpha \| n}{}' p^\alpha$$

where $p^\alpha \| n$ means that $p^\alpha \mid n$ but $p^{\alpha+1} \nmid n$ and the dash indicates that p runs through the primes of I_x.

Lemma 4. *For all but $o(x)$ integers $n\leqslant x$ we have*

$$\nu(f(n))=(1+o(1))(1-\eta_2)\log\log x.$$

Lemma 4 follows easily by the method of Turán [4]. From the well known theorem of Mertens $\sum_{p<y} 1/p=\log\log y+O(1)$ we obtain

$$\sum_{p \text{ in } I_x} 1/p=(1-\eta_2+o(1))\log\log x. \tag{8}$$

From (8) we obtain following Turán [4].

$$\sum_{n=1}^{x} (\nu(f(n))-(1-\eta_2)\log\log x)^2=o(x(\log\log x^2)). \tag{9}$$

(9) immediately implies Lemma 4 by the inequality of Tchebicheff [4].

Lemma 5. *For all but $o(x)$ integers $n<x$ $f(n)$ is squarefree.*

The number of integers $n<x$ for which $f(n)$ is not squarefree is clearly less than (in $\sum' k>\exp((\log x)^{\eta_2})$

$$\sum_{p \text{ in } I_x} x/p^2<x\sum' 1/k^2=o(x).$$

Lemma 6. *The number of integers $n<x$ for which $f(n)$ has two prime factors $p\equiv q \pmod k$, (k satisfies (4)) is $o(x)$.*

The number of integers $n<x$ which have two prime factors $p\equiv q \pmod k$ in I_x is clearly less than

$$x \sum_{p \text{ in } I_x} \frac{1}{p} \sum_{\substack{q \text{ in } I_x \\ q\equiv p \pmod k}} \frac{1}{q} = A(x). \tag{10}$$

From (6) we obtain by a simple computation

$$\frac{1}{x} A(x) < \sum_{p \text{ in } I_x} \frac{1}{p} \frac{c_1 \log\log x}{\varphi(k)} < c_2 (\log\log x)^2 / \varphi(k) = o(1). \tag{11}$$

(10) and (11) proves Lemma 6.

Lemma 7. *Let $B<x^{1/\log\log x}$. Denote by $S^{(B)}$ the set of integers $n<x$ for wich $f(n)=B$. Then*

$$N(S^{(B)}) = (1+o(1)) \frac{x}{B} \prod_{p \text{ in } I_x} \left(1-\frac{1}{p}\right). \tag{12}$$

The integers $n<x$ for which $f(n)=B$ are of the form

$$yB,\ y<x/B,\ y \not\equiv 0 \pmod p,\ p \text{ in } I_x. \tag{13}$$

It easily follows from Brun's method [5] that the number of integers y satisfying (13) is given by (12), which completes the proof of Lemma 7.

Now we are ready to prove Theorem 1. By Lemmas 4,5 and 6 it suffices to consider those integers $n<x$ for which $f(n)$ is squarefree, all prime factors of $f(n)$ are incongruent mod k, and for which

$$(1-\eta_2-\eta_3)\log\log x < \nu(f(n)) < (1-\eta_2+\eta_3)\log\log x \tag{14}$$

where

$$1-\eta_2-\eta_3 > (1+\eta_1)(1-\varepsilon). \tag{15}$$

By (5), (15) can be satisfied if η_3 is small enough. Henceforth it is understood that n satisfies these conditions. Denote by $S(t)$ the set of integers n for which $\nu(f(n))=t$. The proof of Theorem 1 will be complete if we show that for every t (satisfying (14)) all but $o(N(S(t)))$ integers have a divisor $\equiv l \pmod k$ for every l, where the error term $o(N(S(t)))$ is uniform in t. To show this denote by $S(t; l_1,\ldots, l_t)(l_i \not\equiv l_j \pmod k)$ the set of integers n for which

$$v(f(n)) = \prod_{i=1}^{t} p_i,\ p_i \equiv l_i \pmod{k}.$$

A set $S(t; l_1, \ldots, l_t)$ will be called good if $n \in S(t; l_1, \ldots, l_t)$ implies that n has a divisor $\equiv \pmod{k}$ for every l. Clearly a set $S(t; l_1, \ldots, l_t)$ is good if (and only if) every $l \pmod{k}$ can be written in the form

$$\prod_{i=1}^{t} l_i^{\delta_i},\ \delta_i = 0 \text{ or } 1.$$

The number of sets $S(t; l_1, \ldots, l_t)$ is clearly $\binom{\varphi(k)}{t}$. (15) and Lemma 1 implies that all but $o\left(\binom{\varphi(k)}{t}\right)$ of these classes is good (the error term is uniform in k). Hence the proof of Theorem 1 will be complete if we prove that all the $N(S(t; l_1, \ldots, l_t))$ are asymptotically equal [1]. In fact we shall show

$$N(S(t; l_1, \ldots, l_t)) = (1+o(1))\, x \sum^{t} \prod_{p \text{ in } I_x} \left(1 - \frac{1}{p}\right) \tag{16}$$

where $\sum$ is defined by (7) and the error term is uniform in t and the l's. Clearly

$$N(S(t; l_1, \ldots, l_t)) = \sum'' N(S^{(p_1 \cdots p_t)}) \tag{17}$$

where $\sum''$ is defined as in Lemma 3 (i. e. p_i runs through the primes of I_x satisfying $p_i \equiv l_i \pmod{k}$). By the definition of I_x and t we evidently have

$$\prod_{i=1}^{t} p_i < (x^{1/(\log\log x)^3})^t < x^{1/\log\log x}. \tag{18}$$

Thus from (17), (18), Lemmas 7 and 3

$$N(S(t; l_1, \ldots, l_t)) = \sum'' N(S^{(p_1 \cdots p_t)}) =$$

$$(1+o(1))\, x \prod_{p \text{ in } I_x} \left(1 - \frac{1}{p}\right) \sum'' \frac{1}{p_1 \cdots p_t} = (1+o(1))\, x \sum^{t} \prod_{p \text{ in } I_x} \left(1 - \frac{1}{p}\right)$$

which proves (16) and hence the proof of Theorem 1 is complete.

The proof of Theorem 4 follows the same lines as that of Theorem 1. The only difference is that we need here the follow-

(1) Since clearly $N(S(t)) = \sum_{l_1, \ldots, l_t} N(S(t; l_1, \ldots, l_t))$.

ing theorem of Rényi and myself [2]. Let G be an Abelian group of m elements and let $t>(2+\eta)\log m/\log 2$. Then for all but $o\left(\binom{m}{t}\right)$ choices of t distinct elements $a_1,\dots,a_t$ of G, all elements of G can be written in $(1+o(1))2^t/m$ ways in the form

$$\prod_{i=1}^{t} a_i^{\delta_i}, \qquad \delta_i = 0 \text{ or } 1.$$

This theorem immediately implies

Lemma 1'. *Let $t>(2+\eta)\log\varphi(k)/\log 2$. Then for all but $o\left(\binom{\varphi(k)}{t}\right)$ choices of the distinct residues $l_1,\dots,l_t$ for every l (mod k) the number of solutions of*

$$\prod_{i=1}^{t} l_i^{\delta_i} \equiv l \pmod{k},\ \delta_i = 0 \text{ or } 1, \text{ is } (1+o(1))2^t/\varphi(k).$$

The proof of Theorem 4 now proceeds as the proof of Theorem 1.

It is possible that Theorem 4 holds for all $k<2^{(1-\varepsilon)\log\log x}$ instead of $2^{(1/2-\varepsilon)\log\log x}$. This would depend on the corresponding improvement of our theorem with Rényi.

Now we outline the proof of Theorem 2. The term $\frac{x}{l}$ clearly comes from considering the multiples of l, i.e. the numbers $u\,l$, $u<\frac{x}{l}$. Thus our proof will be complete if we can show that if $k>2^{(1+\varepsilon)\log\log x}$ and $N(k,l)$ denotes the number of integers $n<x$ which have at least one divisor in the progression

$$l+d\,k,\ 1\leqslant d\leqslant\frac{x-l}{k} \tag{19}$$

then

$$N(k,t)=o(x) \tag{20}$$

uniformly in l and k.

We will only give a brief indication of the proof of (20). Denote by $F(n)$ the number of prime factors of n where multiple factors are counted multiply. A well known theorem of Hardy and Ramanujan [3] states that the number of integers $n<x$ for which $F(n)>(1+\eta)\log\log x$ is $o(x)$. Put $[(1+\eta)\log\log x]=T$.

Then $\pi_u(x)$ denotes the number of integers $n<x$ with $\nu(n)=u$

$$N(k,l)=\sum_{d\leqslant\frac{x-l}{k}}\ \sum_{u\leqslant T-\nu(l+dk)}\pi_u(x/l+kd)+o(x). \tag{21}$$

(21) follows from the above quoted result of Hardy—Ramanujan and from

$$F((l+kd)s)\geqslant\nu(s)+\nu(l+dk).$$

Hardy and Ramanujan [3] proved

$$\pi_u(y)<\frac{c_1 y}{\log y}\,\frac{(\log\log y+c_2)^{u-1}}{(u-1)!} \tag{22}$$

and it is not hard to prove using their ideas that

$$\sum_{d\leqslant\frac{x-l}{k}}{}'\ \frac{1}{l+dk}<\frac{1}{k}\,\frac{c_3(\log\log x+c_4)^{v-1}}{(v-1)!} \tag{23}$$

where the dash indicates that the summation is extended over the d for which $\nu(l+dk)=v$.

Using (21), (22) and (23) we can obtain (20) by long but elementary and fairly straightforward computations. This completes the outline of the proof of Theorem 2.

Finally we outline the proof of Theorem 3. First we prove that for every $\eta>0$ if $x>x_0(\eta)$

$$Q(x)<(1+\eta)\frac{e^{-c}x}{\log^2\log x}. \tag{24}$$

Lemma 8. *Denote by $l(n)$ the least prime factor of n and by $N(p,x)$ the number of integers $n<x$ with $l(n)=p$. Let $p<\log x$. then*

$$N(p,x)=(1+o(1))\frac{x}{p}\prod_{q<p}\left(1-\frac{1}{q}\right)$$

Lemma 8 follows easily by the sieve of Eratosthenes.

Lemma 9. *Let $y<\log x$, $y\to\infty$ as $x\to\infty$. Then the number of integers $n<x$ satisfying $l(n)>y$ equals*

$$(1+o(1))\,x\,e^{-c}/\log y).$$

Lemma 9 follows easily by the sieve of Eratosthenes and from the well known result of Mertens

$$\prod_{p\leqslant y}\left(1-\frac{1}{p}\right)=(1+o(1))\,e^{-c}/\log y.$$

Lemma 10. *Let $p<2^{(1-\eta_1)\log\log x}$. Then for all but $o(N(p,x)/\log\log x)$ integers $n<x$ satisfying $l(n)=p$, n has a divisor $\equiv 1 \pmod p$. The o is uniform in p.*

Lemma 10 is the crucial lemma. It does not explicitely follows from Theorem 1 (even with $o(N,(p,x))$ replacing $o(N(p,x)/\log\log x)$ but the method of Theorem 1 gives it without much difficulty. For the proof of Lemma 10 we need an error term in the theorem of Rényi and myself [2] but this is easy to accomplish. In fact we obtain that all but $o(N(p,x)/\log\log x)$ integers satisfying $l(n)=p$ have a divisor $\equiv l \pmod p$ for every l. The error term $o(N(p,x)/\log\log x)$ can be very much improved and this I plan to investigate in a separate paper.

Now we prove (24). We split the integers $n<x$ into two classes. In the first class are the integers all whose prime factors are greater than $2^{(1-\eta_1)\log\log x}$. By Lemma 9 the number of integers of the first class is

$$(1+o(1))\,x\,e^{-c}/(1-\eta_1)\log 2\log\log x< \\ <\left(1+\frac{\eta}{2}\right)x\,e^{-c}/\log 2\log\log x \tag{25}$$

if $\eta_1=\eta_1(\eta)$ is sufficiently small. By Lemma 10 all but

$$\sum_{p} o(N(p,x)/\log\log x),\quad p\leqslant 2^{(1-\eta_1)\log\log x} \tag{26}$$

integers of the second class have a divisor of the form $p(kp+1)$. Now clearly $\sum_p N(p,x)=x$, hence from (26) all but $o\left(\frac{x}{\log\log x}\right)$ integers of the second class have a divisor of the form $p(kp+1)$, this together with (25) implies (24).

To complete the proof of Theorem 3 we have to show that for every $\eta>0$ if $x>x_0(\eta)$ if

$$Q(x)>(1-\eta)\frac{e^{-c}x}{\log 2\log\log}. \tag{27}$$

To show (27) observe that Lemma 9 implies that the number of integers $n<x$ for which $l(n)>2^{(1+\eta_1)\log\log x}$ is

$$(28) \qquad \begin{aligned} &(1+o(1))xe^{-c}/(1+\eta_1)\log 2\log\log x> \\ &>\left(1-\frac{\eta}{2}\right)xe^{-c}/\log 2\log\log x \end{aligned}$$

if $\eta_1=\eta_1(\eta)$ is sufficiently small.

It can be shown by the method used in the proof of Theorem 2 that only $o(x/\log\log x)$ integers $n<x$ satisfying $l(n)>2^{(1+\eta_1)\log\log x}$ have a divisor of the form $p(kp+1)$, the details are quite complicated and we do not give them. This together with (28) proves (27) and thus the proof of Theorem 3 is complete.

It seems likely that for all but $o(x)$ integers $n<x$ there is a k satisfying (4) for which n has no divisor $\equiv -1 \pmod{k}$, but this I have not been able to prove.

REFERENCES

[1] Theorem 1 is stated without proof in *Erdös, P.*: Applications of probability to analysis and number theory, J. London Math. Soc. 39 (1964), 692 — 696.

[2] *Erdös, P.* and *Rényi, A.*: Journal d' Analyse Math. 14 (1965), 127—138.

[3] *Hardy, G. H.* and *Ramanujan, S.*: The normal number of prime factors of a number n, Quarterly J. of Math. 48 (1917) 76—92, see also S. Ramanujan, Collected Papers, Cambridge Univ. Press (1927), 262—275, see also M. Kac, Note on the distribution of values of the arithmetic function $d(m)$, Bull. Amer. Math. Soc. 47 (1941), 815—817.

[4] See [3], *Turán, P.*: on a theorem of Hardy and Ramanujan, J. London Math. Soc. 9 (1934), 274—276, see also G. H. Hardy and E. M. Wright, Theory of numbers, Third edition Clarendon Press Oxford 1954, 356—358.

[5] Lemma 7 does not require full strength of Brun's method. It will suffice to follow E. Landau, Vorlesungen über Zahlentheorie, *S. Hirzel*, Leipzig 1927, Vol. 1 71—78 see also P. Erdös, On a problem of Chowla and some related problems, Proc. Cambridge Phil. Soc. 32 (1936), 530—540.

[6] *Walfisz, A.*: Zur additiven Zahlentheorie II Math. Zeitschrift. 40 (1936) 592—607, see also Prachar, Primzahlverteilung, Springer 1957, pages 149 and 310.

(Received January 28, 1965)

Department of Mathematics
Technion, Haifa, Israel.