Séminaire Delange-Pisot-Poitou.
Théorie des nombres

HEINI HALBERSTAM

The distribution of sequences in arithmetic progressions

Séminaire Delange-Pisot-Poitou. Théorie des nombres, tome 13, n° 2 (1971-1972), exp. n° 25, p. 1-5

<http://www.numdam.org/item?id=SDPP_1971-1972__13_2_A10_0>

Séminaire DELANGE-PISOT-POITOU
(Théorie des Nombres)
13e année, 1971/72, n° 25, 5 p.

12 juin 1972

THE DISTRIBUTION OF SEQUENCES IN ARITHMETIC PROGRESSIONS

by Heini HALBERSTAM

In 1968, I gave a talk [3] on this theme at a number theory meeting in Rome, when I suggested various possible lines of enquiry that seemed to me to spring from certain results about the distribution of primes due to BOMBIERI and to DAVENPORT and myself. My purpose to-day is to indicate the progress along these lines which has occurred since then.

1. Let $\Lambda(n)$ be von Mangoldt's function. As usual, write

$$\psi(x\ ;\ q\ ,\ h) = \sum_{n\leq x, n\equiv h \bmod q} \Lambda(n)\ ,$$

$$E(x\ ;\ q\ ,\ h) = \psi(x\ ;\ q\ ,\ h) - (x/\phi(q))\ ,$$

and

$$E(x\ ,\ q) = \max_{(h,q)=1} |E(x\ ;\ q\ ,\ h)|\ .$$

In its latest form, due to H. L. MONTGOMERY ([4], chap. 15), the theorem of Bombieri-Vinogradov on the distribution of primes in arithmetic progressions asserts (I do not give the most general formulation) that, for any $A > 0$,

(1) $$\sum_{q\leq x^{1/2}\log^{-A-13} x} E(x\ ,\ q) \ll x/\log^A x\ ,\quad x > x_0(A)\ ;$$

and it is easy to see that a result only insignificantly weaker than (1) follows from (see [4], chap. 15)

(2) $$\sum_{q\leq Q} qE^2(x\ ,\ q) \ll x^{3/2}\, Q \log^{14} x\ ,\quad x^{1/2}\log^{-c} x \leq Q \leq x^{1/2}\ ,\quad c > 0\ ,\quad x > x_0(c)\ .$$

(Conversely, (2) with $Q = x^{1/2}\log^{-A-13} x$ follows easily from (1)). The form (2), apart from having applications, is in a suitable form for comparison with ([1])

(3) $$\sum_{q\leq Q}\ \sum_{h=1,(h,q)=1} E^2(x\ ;\ q\ ,\ h) = Qx\log Q\{1 + O(\frac{1}{\log Q})\} + O(\frac{x^2}{\log^A x})\ ,$$
$$Q \leq x\ ,\quad A > 0\ ,\quad x > x_0(A)\ .$$

This result stems from a series of papers by BARBAN, DAVENPORT-HALBERSTAM and GALLAGHER, and MONTGOMERY ([4], chap. 17) was the first to give an asymptotic formula ; its present sharp form was computed by CROFT [1], a student of mine, following a suggestion of MONTGOMERY.

Of the relations (2) and (3), (2) is the deeper but does not imply (3) which has

---

([1]) Assuming the grand Riemann hypothesis, the error is $x^{3/2}\log^B x$, for any fixed $B > 0$.

a longer range of validity for Q . Actually, using an old result of TURÁN, it follows from the generalised RH that

(4) $$\sum_{q\leq Q} q \sum_{h=1,(h,q)=1} E^2(x ; q , h) \ll Q^2 x \log^4 x ,$$

which, for $Q \leq x^{1/2}$, is a substantially stronger result than (2) (and therefore (1) too). Curiously enough, if we restrict q in (4) to prime values only, then (4) (with $Q = x^{1/2}$) is true, without use of any unproved hypothesis, even with $\log x$ in place of $\log^4 x$; this follows directly from the large sieve, and so one is a little surprised to find (4) lying apparantly so deep. Inequality (4) would imply conversely that none of $\zeta(s)$, $L(x , \chi)$ has a zero ρ with $\rho > 3/4$.

2. These results give rise to a whole class of tantalising speculations. First of all, since $\psi(x ; q , h)$ counts the primes from a set of about x/q elements, one would expect on probabilistic grounds that

$$E(x ; q , h) \approx (x/q)^{(1/2)+\varepsilon} \log x ;$$

[GRH ===> only $x^{1/2} \log^2 x$] accordingly, one may conjecture that (cf. (1))

$$\sum_{x^{1/2}<q\leq Q} E(x , q) \ll x^{1/2} Q^{1/2} \log x , \quad Q < x ,$$

or, to put it in a weaker but more likely form,

(5) $$\sum_{q<x^{1-\varepsilon}}{}' E(x , q) \ll x/\log^A x .$$

The same probabilistic considerations support also (4).

Although (5), if true, lies very deep indeed, any result intermediate in quality between (1) and (5) would also have important consequences. Writing such a result in the form

(6) $$\sum_{q\leq x^{\alpha}/\log^B x} E(x , q) \ll x/\log^A x , \quad B = B(A) > 0 , \quad 1/2 \leq \alpha < 1 ,$$

it is well-known that the validity of (6) with $\alpha = 0.54$ would imply that [2] $n = p + P_2$ $(n > n_0)$ is possible ; and my student, J. W. PORTER [6], has proved that the validity of (6) with $\alpha = 1 - (1/k)$ would imply that

(7) $$\sum_{a<p\leq x} d_k(p - a)$$

$$\sim \frac{1}{(k - 1)!} \prod_{p\nmid a}\left(1 + \frac{p^{k-1} - (p - 1)^{k-1}}{p^{k-1}(p - 1)}\right) \prod_{p|a} (1 - (1/p))^{k-1} x \log^{k-2} x ,$$

for $k = 3,4 , \ldots$ [(7) follows with $k = 2$ from (1)]. R. C. VAUGHAN [6] has provided some independent support for (7) by showing, without using any hypothesis, that (7) is true for almost all a , and that the number of possible exceptions is surprisingly small.

---

(2) P_2 denotes a number that is a prime or the product of two primes.

3. This regularity on the average of the distribution of primes in arithmetic progressions, as typified by (1) and (3), may be a property shared by other integer sequences. Let us return for a moment to probabilistic considerations. Let Ω be the space of integer sequences ω, and impose on Ω a probability measure in the manner pioneered by ERDŐS and RENYI, and described in detail in Chapter 4 of "Sequences", by writing

$$\mu\{\omega : n \in \omega\} = \alpha_n \qquad (n = 1, 2, 3, \ldots),$$

where $\{\alpha_n\}$ is a decreasing sequence of positive numbers between 0 and 1 tending to 0 as $n \to \infty$; suppose also that the natural extension of α_n to a function α_x on $[1, +\infty[$ is continuous. This procedure throws into relief all integer sequences with essentially the same rate of growth, for it turns out in the resulting probability space that, with probability 1,

$$\omega(x) = \sum_{n\leq x, n\in\omega} 1 \sim \sum_{n\leq x} \alpha_n = M(x) \quad (x \to \infty),$$

say. ELLIOTT [2] has proved that <u>for any assigned sequence</u> $\{\alpha_n\}$ <u>of basic probabilities satisfying the conditions set out above, and any</u> $A > 1$,

$$(8) \quad \sum_{q\leq M(x)/\log^A x} q \max_{(h,q)=1} \{\sum_{n\leq x, n\in\omega, n\equiv h \bmod q} 1 - M(x)/q\}^2 \ll M^2(x)/\log^{A-2} x$$

<u>with probability</u> 1.

For example, with $\alpha_n = 1/\log n$ $(n = 3, \ldots)$, we obtain a statement about prime-like sequences, and therefore some support for conjecture (5) (in a form analogous to (2)) ; and when

$$\alpha_n = c/\sqrt{\log n} \ (n = 3, \ldots) \ (c = 1/\sqrt{2} \prod_{p\equiv 3 \bmod 4} (1 - (1/p^2))^{-1/2},$$

we obtain sequences distributed rather like the numbers that are sums of two squares.

We may expect such results for specific sequences to be extremely difficult to establish ; but intermediate results, e. g. results analogous to (1) and (2), may well be accessible if required. WOLKE [9] studied such questions recently in connection with multiplicative functions and he has found some interesting results which I shall now describe briefly.

WOLKE set himself the task of deriving asymptotic formulae for sums of the type

$$\sum_{n\leq x} d(n) f(n+a), \quad f \text{ multiplicative.}$$

Such sums go back to RAMANUJAN, INGHAM $(f(n) = d(n))$, HOOLEY $(f(n) = d_3(n))$ and LINNIK $(f(n) = d_k(n)$, $k = 2, 3, \ldots)$; HOOLEY's method was intricate and LINNIK used his highly complicated dispersion method. WOLKE has shown, on the contrary, that if one knows enough about $\sum_{n\leq x} f(n)$, and if one has a Bombieri-like estimate of

(9) $\sum_{q\leq x^{1/2-\varepsilon}} \max_{(h,q)=1} \left|\sum_{n\leq x, n\equiv h \bmod q} f(n) - \frac{1}{\phi(q)} \sum_{n\leq x,(n,q)=1} f(n)\right|$

($\varepsilon > 0$, arbitrarily small),

then asymptotic formulae of the desired sort are readily attainable. To be precise, his Bombieri-like result is as follows : if

$$\text{(i)} \qquad 0 \leq f(p^a) \leq c_1 a^{c_2} \quad (a = 1 , 2 , \ldots)$$

$$\text{(ii)} \qquad \sum_{p<x} |f(p) - \tau| \ll x^{1-\eta} \quad (\tau = \tau(f) > 5/2 , \quad \eta = \eta(f) > 0) ,$$

then there exists a $D = D(f) > 0$ and, for each $\varepsilon > 0$ a number $A = A(f , \varepsilon) > 0$, such that

$$(10) \quad \sum_{q\leq x^{\frac{1}{2}-\varepsilon}} \max_{(h,q)=1} \left| \sum_{\substack{n\leq x\\ n\equiv h \bmod q}} f(n) - \frac{D}{\phi(q)} \prod_{p|q} \left(1 + \frac{f(p)}{p} + \frac{f(p^2)}{p^2} + \ldots\right) x \log^{\tau-1} x \right|$$
$$\ll x \log^{\tau-A} x .$$

Conditions (i) and (ii) ensure, by virtue of the work of Wirsing, adequate information about $\sum_{n\leq x} f(n)$.

WOLKE's method is based on the large sieve, but uses much else besides ; it is rather complicated and should repay further study.

4. If we take $f(n) = \mu^2(n)$ above we find that WOLKE's result (10) is not applicable because $\tau = 1$ here ; but, in any case, the condition $(h , q) = 1$ is not natural here and one would expect to go higher with q than $x^{(1/2)-\varepsilon}$. Indeed, ORR [5] has proved that if

$$Q(x ; q , h) = \sum_{n\leq x, n\equiv h \bmod q} \mu^2(n) ,$$

and

$$F(q , h) = \prod_{p\nmid q} \left(1 - \frac{1}{p^2}\right) \frac{\phi(q)}{q(h ; q)\, \phi(q/(h , q))},$$

then, for any $A > 0$,

$$(11) \qquad \sum_{q\leq x^{2/3}/\log^{A+1} x} \max_{\mu^2((h,q)) = 1} |Q(x ; q , h) - xF(q , h)| \ll \frac{x}{\log^A x} .$$

Here $2/3$ appears to take the place $1/2$ in BOMBIERI, and although ORR's argument takes little advantage of averaging, it does seem hard to improve upon. It would be worthwhile to do so, for one could then improve the classical result of ERDÖS-PRACHAR on the least squarefree number in an arithmetic progression.

Just as (11) is an analogue of (1), so there is a partial analogue of (3) for squarefree numbers. Writing

$$S(x , Q) = \sum_{q\leq Q} \sum_{\substack{h=1\\ \mu^2((h,q))=1}}^{q} \{Q(x , q , h) - xF(q , h) ,$$

it is curious that one cannot readily obtain, as in the prime case, an asymptotic formula. What is know is that

$$S(x , Q) \ll \begin{cases} xQ , & 1 \leq Q \leq x^{5/8} \quad (^3) \\ x^{(3/2)+\varepsilon} , & x^{5/8} < Q \leq x^{2/3} \\ x^{1/2} Q^{3/2} \log^3 x , & x^{2/3} < Q \leq x . \end{cases}$$

The first result is due to ORR and WARLIMONT [8], the last two to CROFT. ORR's paper gives many interesting applications.

On probabilistic grounds one would expect the difference $Q(x ; q , h) - xF(q , h)$ (as MONTGOMERY printed out to me in a letter) to be about $(x/q)^{(1/4)+\varepsilon}$ **for large q**, and so CROFT's estimate for large Q is about right. However, reaching an asymtotic formula still presents some technical difficulties.

BIBLIOGRAPHY

[1] CROFT (M. J.), in preparation.

[2] ELLIOTT (P. D. T. A.). - Some applications of a theorem of Raikov to number theory, J. of Number Theory, t. 2, 1970, p. 22-25.

[3] ELLIOTT (P. D. T. A.) and HALBERSTAM (H.). - A conjecture in prime number theory, Symposia Mathematica, Vol. 4, p. 59-72. - London, Academic Press, 1970 (Istituto nazionale di alta Mathematica).

[4] MONTGOMERY (H. L.). - Topics in multiplicative number theory. - Berlin Springer-Verlag, 1971 (Lecture Notes in Mathematics, 227).

[5] ORR (R. C.). - Remainder estimates for squarefree integers in arithmetic progression, J. of Number Theory, t. 3, 1971, p. 474-497.

[6] PORTER (J. W.).-The generalized Titchmarsh-Linnik divisor problem, Proc. London math. Soc., 3rd Series, t. 24, 1972, p. 15-26.

[7] VAUGHAN (R. C.), to appear in J. London math. Soc.

[8] WARLIMONT (R.). - On squarefree numbers in arithmetic progressions, Monatsh. Math., t. 73, 1969, p. 433-448.

[9] WOLKE (D.). - Einige Anwendungen des grossen Siebes auf zahlentheoretische Funktionen. Habilitationsschrift. - Marburg/Lahn, Philipps-Universität Marburg/Lahn, 1970.

(Texte reçu le 30 juin 1972)

Heini HALBERSTAM
Mathematical Department
University of Nottingham
NOTTINGHAM (Grande-Bretagne)

---

(3) This estimate is not available for $x^{1/3} < Q < x^{1/3} \log^{10/3} x$; but WARLIMONT shows also that $S(x , Q) \ll x^{1+\varepsilon} Q$ always, and that $S(x , x^{\alpha}) \ll x^{3/2(2\alpha+1)+\varepsilon}$ if $1/2 \leq \alpha \leq 1$. Note that WARLIMONT uses $(h , q) = 1$.