Can. J. Math., Vol. XXIX, No. 4, 1977, pp. 819–826

ON THE DISTRIBUTION MODULO 1 OF THE SEQUENCE $\alpha n^2 + \beta n$

WOLFGANG M. SCHMIDT

In memory of H. Heilbronn

1. Introduction. Dirichlet's Theorem says that for any real α and for $N \geqq 1$, there exists a natural $n \leqq N$ with

$$||\alpha n|| < N^{-1},$$

where $|| \quad ||$ denotes the distance to the nearest integer. Heilbronn [**2**], improving estimates of Vinogradov [**3**], showed that for α, N as above and for $\epsilon > 0$, there exists an $n \leqq N$ with

$$||\alpha n^2|| < c_1(\epsilon)N^{-(1/2)+\epsilon}.$$

Davenport [**1**], as part of a more general investigation, proved that for a quadratic polynomial $\alpha x^2 + \beta x$, for $N \geqq 1$ and $\epsilon > 0$, there is an $n \leqq N$ with

$$||\alpha n^2 + \beta n|| < c_2(\epsilon)N^{-(1/3)+\epsilon}.$$

The example $0.x^2 + 0.x + \frac{1}{2}$ shows that the restriction to polynomials with constant term zero is essential. An important feature of the results is that they are uniform in α, β and they are "localized", i.e. they specify n to lie in a given range. They imply but are not implied by non-localized results; e.g. Heilbronn's Theorem implies that there are infinitely many n with $||\alpha n^2|| < c_1(\epsilon)n^{-(1/2)+\epsilon}$.

THEOREM. *Suppose α, β are real, and $\epsilon > 0$. Given $N \geqq 1$, there exists an $n \leqq N$ with*

$$||\alpha n^2 + \beta n|| < c_3(\epsilon)N^{-(1/2)+\epsilon}.$$

This generalizes Heilbronn's Theorem and sharpens Davenport's estimate. Our point of departure from the standard arguments will be the estimate (7) for exponential sums.

2. A routine beginning. We start out with the Heilbronn-Davenport approach. Write $e(x) = e^{2\pi ix}$. Suppose $M > 2$, and let r be natural. According to Vinogradov ([**4**, Lemma 12], applied with $\beta = -\alpha = \frac{1}{2}M^{-1}$, $\Delta = M^{-1}$), there exists a real-valued periodic function

$$\psi(x) = \sum_{m=-\infty}^{\infty} c_m e(mx)$$

Received September 10, 1976. This research was written while the author was partially supported by NSF-MCS 75-08233-A01.

with

(i) $\psi(x) = 0$ if $||x|| \geqq M^{-1}$,

(ii) $c_0 = M^{-1}$,

(iii) $|c_m| \ll \min (M^{-1}, |m|^{-r-1}M^r)$

for $m \neq 0$, where the constant in $\ll$ depends only on r.

Put

$$M = N^{1/2-\epsilon}$$

and suppose N to be a large integer, say $N > c_4(\epsilon)$. Everything is fine if there exists an $n \leqq N$ with $||\alpha n^2 + \beta n|| < M^{-1}$. We therefore assume that there is no such n. Then by (i),

$$\sum_{n=1}^{N} \psi(\alpha n^2 + \beta n) = 0.$$

In view of (ii) we obtain

$$(1) \qquad \sum_{m \neq 0} c_m S_m = -NM^{-1}$$

with

$$S_m = \sum_{n=1}^{N} e(m(\alpha n^2 + \beta n)).$$

Putting $L = [M^{1+\epsilon}]$ where [ ] denotes the integer part, we have

$$(2) \qquad N^{1/2-\epsilon} < L \leqq N^{1/2-(\epsilon/2)}.$$

By (iii),

$$\sum_{|m|>L} |c_m| \ll M^r L^{-r} \ll M^{-\epsilon r} \ll M^{-2},$$

if we fix $r > 2\epsilon^{-1}$. Here and in the sequel, the constants in $\ll$ and in 'big O' depend only on ϵ. If N, whence M, is sufficiently large, we have

$$\sum_{|m|>L} |c_m S_m| \leqq N \sum_{|m|>L} |c_m| < \tfrac{1}{2}NM^{-1}.$$

Comparison with (1) yields

$$\left| \sum_{0<|m|\leqq L} c_m S_m \right| \geqq \tfrac{1}{2}NM^{-1},$$

whence

$$\sum_{0<|m|\leqq L} |S_m| \gg N$$

by (iii). Since $|S_{-m}| = |S_m|$, and by Cauchy's inequality,

$$(3) \qquad \sum_{m=1}^{L} |S_m|^2 \gg N^2 L^{-1}.$$

https://doi.org/10.4153/CJM-1977-084-8 Published online by Cambridge University Press

3. Exponential sums.

$$|S_m|^2 = \sum_{n_1=1}^{N} \sum_{n_2=1}^{N} e(m(n_2 - n_1)(\alpha(n_1 + n_2) + \beta)).$$

Putting $u = n_1 + n_2$, $v = n_2 - n_1$, we obtain

$$|S_m|^2 = \sum_{u,v} e(mv(\alpha u + \beta)),$$

where the sum is over integers u, v with $u \equiv v \pmod 2$ and

$$0 < u + v \leqq 2N, \quad 0 < u - v \leqq 2N.$$

The N summands with $v = 0$ give a contribution N. Since the substitution $v \to -v$ changes each summand into its complex conjugate, we have

$$|S_m|^2 = N + 2\mathrm{Re} \sum_{u,v} e(mv(\alpha u + \beta)),$$

where the sum is over u, v with $u \equiv v \pmod 2$ and

(4) $\quad 0 < v < N, \quad v < u \leqq 2N - v.$

For fixed v, the terms $mv(\alpha u + \beta)$ with integers $u \equiv v \pmod 2$ form an arithmetic progression with common difference $2mv\alpha$. The sum of $e(mv(\alpha u + \beta))$ over the terms of this arithmetic progression (which has length $N - v < N$ by (4)) is

$$\ll \min (N, ||2mv\alpha||^{-1}).$$

Summation over v in $0 < v < N$ yields the well known Weyl estimate for $|S_m|^2$, namely

$$|S_m|^2 = N + O\left(\sum_{v=1}^{N} \min (N, ||2mv\alpha||^{-1}) \right).$$

Since by (2), NL is small compared to N^2L^{-1}, our estimate (3) yields

$$\sum_{m=1}^{L} \sum_{v=1}^{N} \min (N, ||2vm\alpha||^{-1}) \gg N^2L^{-1}.$$

Observing that the number of divisors of an integer $\leqq 2LN$ is $\ll N^{\epsilon/3}$, we get

(5) $$\sum_{k=1}^{2LN} \min (N, ||k\alpha||^{-1}) \gg N^{2-(\epsilon/3)}L^{-1}.$$

We now reverse the roles of u, v above. The inequalities (4) may be re-written as

(6) $\quad 0 < u < 2N, \quad 0 < v \leqq \min (u - 1, 2N - u).$

For fixed u, the terms $mv(\alpha u + \beta)$ with $v \equiv u \pmod 2$ form an arithmetic progression with common difference $2m(\alpha u + \beta)$. The sum of $e(mv(\alpha u + \beta))$

https://doi.org/10.4153/CJM-1977-084-8 Published online by Cambridge University Press

over the terms of this arithmetic progression is

$$\ll \min (N, ||2m(\alpha u + \beta)||^{-1}).$$

Summation over u yields

$$(7) \qquad |S_m|^2 = N + O\left(\sum_{u=1}^{2N} \min (N, ||2m(\alpha u + \beta)||^{-1}) \right).$$

We now invoke (3) to obtain

$$(8) \qquad \sum_{m=1}^{2L} \sum_{u=1}^{2N} \min (N, ||m(\alpha u + \beta)||^{-1}) \gg N^2 L^{-1}.$$

4. A first application of Dirichlet's Theorem. We briefly return to Heilbronn's argument, i.e. to (5). By Dirichlet, there are coprime integers a, q with

$$(9) \qquad 1 \leqq q \leqq N^{2-(\epsilon/2)} L^{-1}, \quad |\alpha q - a| < N^{-2+(\epsilon/2)} L,$$

so that in particular, $|\alpha - (a/q)| < q^{-2}$. It is well known that for a block $\mathscr{B}$ of q consecutive integers,

$$\sum_{k \in \mathscr{B}} \min (N, ||k\alpha||^{-1}) \ll N + q \log q.$$

Dividing the range $1 \leqq k \leqq 2LN$ into $\leqq (2LN/q) + 1$ blocks of q or fewer consecutive integers, we have

$$(10) \qquad \sum_{k=1}^{2LN} \min (N, ||k\alpha||^{-1}) \ll (2LNq^{-1} + 1)(N + q \log q).$$

In view of (2) and (9), the only one of the four summands on the right hand side of (10) which could possibly be as large as $N^{2-(\epsilon/3)} L^{-1}$, is $2LN^2 q^{-1}$. Thus by (5) and (2),

$$(11) \qquad q \ll L^2 N^{\epsilon/3} \ll N^{1-(\epsilon/2)}.$$

5. Auxiliary lemmas.

LEMMA 1. *Suppose* $|\sigma| \leqq N^{-1}$. *Then*

$$\min (N, ||\rho + \sigma||^{-1}) \ll \min (N, ||\rho||^{-1}).$$

Proof. Obvious.

LEMMA 2. *Let* $\xi_1, \ldots, \xi_K$ *be reals with* $||\xi_i - \xi_j|| \geqq \rho > 0$ *for* $i \neq j$, *and with* $||\xi_1|| = \min (||\xi_1||, \ldots, ||\xi_K||)$. *Then*

$$\sum_{i=2}^{K} ||\xi_i||^{-1} \ll \rho^{-1} \log K.$$

Proof. We may suppose that $||\xi_1|| \leqq \ldots \leqq ||\xi_K||$. Then $||\xi_i|| \geqq (i - 1)\rho/2$ for $i = 2, 3, \ldots, K$. The lemma follows.

LEMMA 3. *Let ρ, σ be real, and K natural with $|K\rho| \leqq 1$. Write*

$$\delta = \min_{1\leqq j\leqq K} ||\rho j + \sigma||, \quad \Delta = \max_{1\leqq j\leqq K} ||\rho j + \sigma||.$$

Then

$$(12) \quad \sum_{j=1}^{K} ||\rho j + \sigma||^{-1} = \delta^{-1} + O(\Delta^{-1}K \log K).$$

Proof. Write the numbers $\rho j + \sigma$ with $j = 1, \ldots, K$ as $\xi_1, \ldots, \xi_K$, arranged such that $||\xi_1|| = \delta = \min (||\xi_1||, \ldots, ||\xi_K||)$. We have $||\xi_i - \xi_j|| \geqq |\rho|$ in view of $|K\rho| \leqq 1$. Thus if $|\rho| \geqq \Delta/2K$, the desired conclusion follows from Lemma 2. On the other hand, if $|\rho| < \Delta/2K$, then $||\rho j + \sigma|| > \frac{1}{2}\Delta$ $(j = 1, \ldots, K)$, and the sum in (12) is estimated by $2\Delta^{-1}K$.

LEMMA 4. *Suppose r, s are coprime,*

$$(13) \quad 1 \leqq s \leqq N \quad \text{and} \quad ||\xi s|| = |\xi s - r| < (3N)^{-1}.$$

Then

$$(14) \quad \sum_{j=1}^{s} \min (N, ||\xi j + \eta||^{-1}) \ll \min (N, s||\eta s||^{-1}) + s \log s.$$

Proof. Writing $\xi = (r/s) + \xi_0$, we have $|j\xi_0| \leqq |s\xi_0| = ||\xi s|| < (3N)^{-1}$. So by Lemma 1, our sum is

$$\ll \sum_{j=1}^{s} \min (N, ||(r/s)j + \eta||^{-1}) = \sum_{j=1}^{s} \min (N, ||(j/s) + \eta||^{-1}).$$

We now apply Lemma 3 with $K = s$, $\rho = 1/s$, $\sigma = \eta$, and obtain

$$\ll \min (N, \delta^{-1}) + \Delta^{-1}s \log s.$$

In our special situation, δ is the distance from η to the nearest integer multiple of $1/s$, or $\delta = s^{-1}||\eta s||$. If $s \geqq 2$, then $\Delta \geqq \frac{1}{4}$, and we are done. Of course we are also done when $s = 1$.

LEMMA 5. *Suppose r, s, N, ξ are as in Lemma 4. Then*

$$\sum_{u=1}^{2N} \min (N, ||\xi u + \eta||^{-1}) \ll (\log N) \min \left(\frac{N^2}{s}, \frac{N}{||\eta s||}, \frac{1}{||\xi s||}\right).$$

Proof. Write $u = sz + j$ $(j = 1, \ldots, s)$. The sum in question is

$$\leqq \sum_{z=0}^{[2N/s]} \sum_{j=1}^{s} \min (N, ||\xi j + \xi sz + \eta||^{-1}),$$

and is

$$\ll \sum_{z=0}^{[2N/s]} (\min (N, s||\xi s^2 z + \eta s||^{-1}) + s \log s)$$

by Lemma 4. Writing $\xi_1 = \xi s - r$ with $|\xi_1| = ||\xi s||$, we obtain

$$(15) \quad \ll N \log N + \sum_{z=0}^{[2N/s]} \min (N, s||\xi_1 s z + \eta s||^{-1}).$$

It is clear that this is

$$\ll N \log N + (N^2/s) \ll (\log N)(N^2/s),$$

which is the first of the desired estimates.

To get the other estimates we now apply Lemma 3 to the sum in (15), that is, we apply it with $K = [2N/s] + 1$, $\rho = \xi_1 s$, $\sigma = \eta s - \xi_1 s$. Observe that $|K\rho| \leqq (3N/s)\,|\xi_1 s| = 3N\,||\xi s|| < 1$. We obtain

$$\begin{aligned} &\ll N \log N + \min (N, s\,\delta^{-1}) + s\,\Delta^{-1} K \log K \\ &\ll N \log N + \Delta^{-1} N \log N \\ &\ll \Delta^{-1} N \log N. \end{aligned}$$

Here Δ is the maximum of $||\xi_1 s z + \eta s||$ for $z = 0, \ldots, [2N/s]$. Clearly $\Delta \geqq ||\eta s||$. But since $|N\xi_1| < 1/3$, we have also

$$|\xi_1 s[N/s]| = ||\xi_1 s[N/s]|| \leqq ||\xi_1 s[N/s] + \eta s|| + ||\eta s|| \leqq 2\Delta,$$

whence $\Delta \geqq \frac{1}{2}\,|\xi_1| s[N/s] \geqq \frac{1}{4}\,|\xi_1| N = \frac{1}{4}\,N||\xi s||$. We therefore obtain

$$\Delta^{-1} N \log N \ll (\log N) \min (N||\eta s||^{-1}, ||\xi s||^{-1}).$$

6. Making use of (8). Let d be a divisor of the integer q of § 4. Let Σ_d be the double sum in (8), but restricted to summands with $(m, q) = d$. Writing $q = dq_1$, $m = dm_1$,

$$\Sigma_d = \sum_{\substack{m_1=1 \\ (m_1, q_1)=1}}^{[2L/d]} \sum_{u=1}^{2N} \min (N, ||\alpha d m_1 u + \beta d m_1||^{-1}).$$

Since the number of divisors of q is $\ll N^{\epsilon/3}$, there will by (8) be a d with

$$(16) \quad \Sigma_d \gg N^{2-(\epsilon/3)} L^{-1}.$$

We now consider the inner sum in the definition of Σ_d. It is the type of sum considered in Lemma 5, with $\xi = \alpha d m_1$ and $\eta = \beta d m_1$. With $s = q_1$, $r = a m_1$ we have $(s, r) = 1$, and

$$s \leqq q \leqq N$$

by (11). Further

$$|\xi s - r| = |\alpha d m_1 s - a m_1| = |\alpha q - a| m_1 < N^{-2+(\epsilon/2)} L(2L) < (3N)^{-1}$$

by (9), (2), so that $||\xi s|| = |\xi s - r| < (3N)^{-1}$. The hypotheses of Lemma 5 are satisfied, and the inner sum in the definition of Σ_d is

$$\ll (\log N) \min \left(\frac{N^2}{q_1}, \frac{N}{||\beta q m_1||}, \frac{1}{||\alpha q m_1||} \right).$$

https://doi.org/10.4153/CJM-1977-084-8 Published online by Cambridge University Press

So by (16), and since $||\alpha q m_1|| = ||\alpha q|| m_1$,

$$(17) \quad \sum_{m_1=1}^{[2L/d]} \min\left(\frac{N^2}{q_1}, \frac{N}{||\beta q m_1||}, \frac{1}{||\alpha q|| m_1}\right) \gg N^{2-(\epsilon/2)} L^{-1}.$$

7. A second application of Dirichlet's Theorem. There are coprime t, w with

$$(18) \quad 1 \leqq t \leqq 4L \quad \text{and} \quad ||\beta q t|| = |\beta q t - w| < (4L)^{-1}.$$

LEMMA 6. *Suppose $m_1 \leqq 2L$ is not divisible by t. Then*

$$||\beta q m_1|| \geqq (2t)^{-1}.$$

Proof. Assuming that $m_1 \leqq 2L$ and that $||\beta q m_1|| < (2t)^{-1}$, we have $|\beta q m_1 - l| < (2t)^{-1}$ for some l. Combining this with $|\beta q t - w| < (4L)^{-1}$, we obtain

$$|lt - m_1 w| < (\tfrac{1}{2}) + (m_1/4L) \leqq 1,$$

so that $lt - m_1 w = 0$, and t is a divisor of $m_1 w$, hence of m_1.

We are going to apply the lemma to estimate the part of the sum (17) where m_1 is not divisible by t. Let $\mathfrak{C}$ be a block of $\leqq t/2$ consecutive integers $\leqq 2L$ which are not divisible by t. By the lemma we know that $||\beta q m_1|| \geqq (2t)^{-1}$ for $m_1 \in \mathfrak{C}$. On the other hand, if m_1, m_1' are distinct elements of $\mathfrak{C}$, we write $\beta q = (w/t) + \beta_0$ and note that $|\beta_0| < (4tL)^{-1}$, so that

$$\begin{aligned} ||\beta q m_1 - \beta q m_1'|| &\geqq ||(w/t)(m_1 - m_1')|| - |\beta_0| m_1 - m_1'| \\ &\geqq t^{-1} - (4tL)^{-1}(t/2) \geqq (2t)^{-1} = \rho, \end{aligned}$$

say. Lemma 2 yields

$$\sum_{m_1 \in \mathfrak{C}} ||\beta q m_1||^{-1} \ll 2t + \rho^{-1} \log t \ll t \log N.$$

We divide the integers m_1 in $1 \leqq m_1 \leqq 2L/d$ with $t \nmid m_1$ into $\leqq (4L/dt) + 1$ blocks of $\leqq t/2$ consecutive integers, and obtain

$$\sum_{\substack{m_1=1 \\ t \nmid m_1}}^{[2L/d]} ||\beta q m_1||^{-1} \ll (t \log N)((4L/dt) + 1) \ll L \log N.$$

It follows that the sum (17), restricted to m_1 with $t \nmid m_1$ is $\ll LN \log N$, and is smaller in magnitude than the right hand side of (17). We thus may restrict ourselves to m_1 of the form $m_1 = t m_2$, and we obtain

$$(19) \quad \sum_{m_2=1}^{[2L/dt]} \min\left(\frac{N^2}{q_1}, \frac{1}{||\alpha q|| t m_2}\right) \gg N^{2-(\epsilon/2)} L^{-1}.$$

In particular, $(2L/dt)(N^2/q_1) \gg N^{2-(\epsilon/2)} L^{-1}$, and putting $n = qt$, we have $n = q_1 dt \ll L^2 N^{\epsilon/2}$, whence

$$(20) \quad n = qt \leqq N$$

by (2). On the other hand, (19) yields

$$(||\alpha q||t)^{-1} \log N \gg N^{2-(\epsilon/2)}L^{-1},$$

and

$$||\alpha q||t \ll LN^{\epsilon-2} \ll N^{(\epsilon/2)-(3/2)}.$$

So

$$||\alpha n^2 + \beta n|| \leqq n||\alpha n|| + ||\beta n|| \leqq nt||\alpha q|| + ||\beta qt|| \ll N.N^{(\epsilon/2)-(3/2)} + L^{-1} \ll N^{-(1/2)+\epsilon},$$

by (2), (18).

References

1. H. Davenport, *On a theorem of Heilbronn*, Quart. J. Math. (2) *18* (1967), 337–344.

2. H. Heilbronn, *On the distribution of the sequence $n^2\theta$ (mod 1)*, Quart. J. Math. *19* (1948), 249–256.

3. I. M. Vinogradov. *Analytischer Beweis des Satzes über die Verteilung der Bruchteile eines ganzen Polynoms*, Bull. Acad. Sci. USSR (6) *21* (1927), 567–578.

4. ——— *The method of trigonometric sums in the theory of numbers* (1947), English transl. (Interscience, New York, 1954).

University of Colorado,
Boulder, Colorado

https://doi.org/10.4153/CJM-1977-084-8 Published online by Cambridge University Press