Let $a_1 < \ldots < a_l \leqslant x$ be the sequence of integers satisfying (29). From (29) we obtain by a simple computation that for every $r$, $1 \leqslant r \leqslant l$

$$\log_2 a_r - 2c_{14}(\log_2 a_r)^{1/2} < \nu(a_r) < \log_2 a_r + 2c_{14}(\log_2 a_r)^{1/2}. \tag{30}$$

Denote as before by $d^+(a_r)$ the number of $a$'s dividing $a_r$. To prove (28) it will suffice to show that for every $r$

$$d^+(a_r) < \exp\big(c_{14}(\log_2 x)^{1/2}\log_3 x\big). \tag{31}$$

Denote by $p_1 < \ldots < p_{\nu(a_r)}$ the prime factors of $a_r$. Assume $a_t | a_r$. If $\nu(a_t) \leqslant k_0$ then by (30) there are clearly fewer than $\nu(a_r)^{k_0+1} \leqslant (\log_2 x)^{k_0+2}$ choices for $a_t$, thus these can be ignored. If $\nu(a_t) > k_0$, let $p_s$ be the greatest prime factor of $a_t$. Since $a_t$ and $a_r$ both satisfy (29) and (30) a simple computation shows that

$$s - 3c_{14}(\log_2 a_r)^{1/2} \leqslant \nu(a_t) \leqslant s. \tag{32}$$

Thus by an easy argument and simple computation

$$d^+(a_r) \leqslant (\log_2 x)^{k_0+2} + \sum_{s=k_0+1}^{\nu(a_r)} \sum_{s-3c_{15}(\log_2 a_r)^{1/2}}^{s} \binom{s}{u}$$
$$< (\log_2 x)^{k_0+2} + \nu(a_r)\big(\nu(a_r)\big)^{4c_{15}(\log_2 a_r)^{1/2}}$$
$$< \nu(a_r)^{5c_{15}(\log_2 a_r)^{1/2}} < \exp\big(c_{16}(\log_2 x)^{1/2}\log_3 x\big).$$

Thus (31) is proved (with $c_{16} = c_{14}$).

### References


[1] F. Behrend, *On sequences of numbers not divisible one by another*, J. London Math. Soc. 10 (1935), pp. 42-44.

[2] A. S. Besicovitch, *On the density of certain sequences*, Math. Ann. 110 (1934), pp. 336-341.

[3] H. Davenport and P. Erdös, *On sequences of positive integers*, Acta Arith. 2 (1936), pp. 147-151.

[4] P. Erdös, *Note on sequences of integers no one of which is divisible by any other*, J. London Math. Soc. 10 (1935), pp. 126-128.

[5] — *On the distribution function of additive functions*, Ann. Math. 47 (1946), pp. 1-20.

[6] J. Kubilius, *Probabilistic methods in the theory of numbers*, Translation of Math. Monographs, Amer. Math. Soc. 1964, vol. 11.


*Reçu par la Rédaction le 2. 7. 1965*

ACTA ARITHMETICA
XI (1966)

# On sums of roots of unity

**(Solution of two problems of R. M. Robinson)**

by

A. Schinzel (Warszawa)

*To Professor Viggo Brun on his 80th birthday*

R. M. Robinson ([4]) proposed the following problem:

"How can we tell whether a given cyclotomic integer can be expressed as a sum of a prescribed number of roots of unity?"

An answer to this problem follows as Corollary 1 from the theorem below.

Theorem 1. *Let $\sum_{i=1}^{k} a_i\zeta_N^{\alpha_i} = \vartheta$, where the $a_i$ are rational integers, $\zeta_N = e^{2\pi i/N}$. Suppose that $\vartheta$ is an algebraic integer of degree $d$ and that $(N, \alpha_1, \alpha_2, \ldots, \alpha_k) = 1$. Then either there is a non-empty set $I \subset \{1, 2, \ldots, k\}$ such that*

$$\sum_{i\in I} a_i\zeta_N^{\alpha_i} = 0$$

*or*

$$N < d(2\log d + 200k^2\log 2k)^{20k^2}.$$

Corollary 1. *An algebraic integer of degree $d$ is a sum of $k$ roots of unity only if it is a sum of $k$ roots of unity of common degree less than $d(2\log d + 200k^2\log 2k)^{20k^2}$.*

Corollary 2. *An algebraic integer $\neq 0$ is a sum of $k$ roots of unity in infinitely many ways if and only if it is a sum of $k-2$ roots of unity.*

Corollary 3. *If $1+\sum_{i=1}^{k}\zeta_N^{\alpha_i} = 0$, and $(N, \alpha_1, \ldots, \alpha_k) = 1$ then either there is a non-empty set $I \subset \{1, 2, \ldots, k\}$ such that $\sum_{i\in I}\zeta_N^{\alpha_i} = 0$ or $N < (200\, k^2\log 2k)^{20k^2}$.*

The proofs of Theorem 1, Corollary 1 and 2 are given later, Corollary 3 follows immediately from the theorem and is stated with the purpose of asking the question how much the inequality for $N$ can be improved.

There is a statement in the literature ([2], p. 228) from which it would follow that $(200\,k^2\log 2k)^{20k^2}$ can be replaced by $k+2$. This is true for $k<5$ but false for $k=5$ as the following example due to Robinson shows

$$1+\zeta_{30}+\zeta_{30}^{7}+\zeta_{30}^{13}+\zeta_{30}^{19}+\zeta_{30}^{20}=0.$$

Robinson made a conjecture ([4], § 4) about the numbers $\sqrt{5}\cos(\pi/M)+i\sin(\pi/M)$. I prove this conjecture as the following

THEOREM 2. *The number* $\sqrt{5}\cos(\pi/M)+i\sin(\pi/M)$ *is a sum of three roots of unity if and only if* $M=2,\ 3,\ 5,\ 10$, *or* $30$.

According to Robinson two algebraic integers $\xi$ and $\eta$ are equivalent if for a suitable conjugate $\xi'$ of $\xi$, $\eta/\xi'$ is a root of unity. Theorem 2 implies

COROLLARY 4. (Conjecture 3 from [4]). *The numbers* $1+2i\cos(\pi/M)$ *and* $\sqrt{5}\cos(\pi/M)+i\sin(\pi/M)$ *are equivalent only for* $M=2,\ 10$ *or* $30$.

COROLLARY 5. *There exist infinitely many inequivalent cyclotomic integers which lie with all their conjugates in the circle* $|z|<3$ *and are not sums of three roots of unity.*

The last corollary, which follows immediately from the fact that the numbers $\sqrt{5}\cos(\pi/M)+i\sin(\pi/M)$ for different $M$ have different absolute values, disproves a conjecture made by Robinson at Boulder 1959 (cf. [4], § 4). An analogous conjecture for the circle $|z|<2$ is still unproved (l. c. Conjecture 1).

I conclude this introduction by expressing my thanks to Professor Robinson who let me have his manuscript before publication, to Professor Davenport who kindly supplied the proof of Lemma 2 and to Dr. A. Białynicki-Birula and Professor D. J. Lewis who discussed the subject with me and read my manuscript.

In the sequel $Q$ denotes the rational field, $[K_2:K_1]$ the degree of a field $K_2$ over a field $K_1$, and $|K|=[K:Q]$. The empty sums are 0, the empty products 1.

LEMMA 1. *For all positive integers* $h$ *and* $N\geqslant 3$ *there exists an integer* $D$ *satisfying the conditions*

$$(1)\qquad 1\leqslant D\leqslant(\log N)^{20h},$$

$$(2)\qquad (iD+1,N)=1\quad \text{for}\quad i=1,2,\ldots,h.$$

Proof. For $h=1$ we can take $D=q-1$, where $q$ is the least prime not dividing $N$. Since in that case $\sum_{p\leqslant D}\log p\leqslant\log N$, we get from [5], Theorem 10

$$D\leqslant 100\quad\text{or}\quad 0.84D\leqslant\log N.$$

On the other hand $D\leqslant N$, which implies $D\leqslant(\log N)^{20}$ for all $N\geqslant 3$.



Therefore we can assume $h\geqslant 2$. Since $D=N$ satisfies the condition (2) we can assume further $N>(\log N)^{20h}$, which implies

$$(3)\qquad \log N>110h,\qquad \log\log N>5.$$

Let $A$ be the product of all primes not exceeding $10h$, and let $p_1<p_2<\ldots<p_r$ be all the other primes dividing $N$. Let $P(A,X,p_1,\ldots,p_r)$ be the number of all integers $x$ satisfying the conditions

$$1\leqslant x\leqslant X,\qquad x\equiv 0\ (\mathrm{mod}\ A),$$

$$ix+1\not\equiv 0\ (\mathrm{mod}\ p_j)\qquad (1\leqslant i\leqslant h,\ 1\leqslant j\leqslant r).$$

The second condition above is equivalent to $h$ conditions of the form $x\not\equiv a_{ij}\ (\mathrm{mod}\ p_j)$. Thus by Brun's method ([1], cf. [6], Lemma 7) for any given sequence of integers $r=r_0\geqslant r_1\geqslant\ldots\geqslant r_t=1$ we have

$$(4)\qquad P(A,X,p_1,\ldots,p_r)>\frac{E}{A}X-R,$$

where

$$E=1-h\sum_{a=1}^{r}\frac{1}{p_a}+h^2\sum_{a=1}^{r}\sum_{\substack{a_1\leqslant r\\ a_1<a}}\frac{1}{p_ap_{a_1}}-h^3\sum_{a=1}^{r}\sum_{\substack{a_1\leqslant r\\ a_1<a}}\sum_{\substack{\beta_1\leqslant r\\ \beta_1<a_1}}\frac{1}{p_ap_{a_1}p_{\beta_1}}+$$
$$+\ldots+\sum_{a=1}^{r}\sum_{\substack{a_1\leqslant r_1\\ a_1<a}}\sum_{\substack{\beta_1\leqslant r_1\\ \beta_1<a_1}}\cdots\sum_{\substack{a_{t-1}\leqslant r_{t-1}\\ a_{t-1}<\beta_{t-2}}}\sum_{\substack{\beta_{t-1}\leqslant r_{t-1}\\ \beta_{t-1}<a_{t-1}}}\sum_{\substack{a_t\leqslant r_t\\ a_t<\beta_{t-1}}}\frac{1}{p_ap_{a_1}\cdots p_{a_t}}$$

and

$$(5)\qquad R\leqslant(1+hr)\prod_{n=1}^{t}(1+hr_n)^2.$$

Denote by $r_n$ $(1\leqslant n\leqslant t)$ the least positive integer such that

$$\pi_n=\prod_{r_n<s\leqslant r_{n-1}}\left(1-\frac{h}{p_s}\right)\geqslant\frac{1}{1.3}$$

and choose $t$ so that

$$\pi_t=\prod_{s\leqslant r_{t-1}}\left(1-\frac{h}{p_s}\right)\geqslant\frac{1}{1.3}.$$

It follows hence (cf. [6], formulae (18) and (32))

$$(6)\qquad \pi_n\leqslant\frac{10}{9}\cdot\frac{1}{1.3}=\frac{1}{1.17}<\frac{8}{9}$$

and

$$(7)\qquad E>0.5\prod_{s=1}^{r}\left(1-\frac{h}{p_s}\right).$$

We shall show that

$$(8)\qquad \log\prod_{s=1}^{r}\Big(1-\frac{h}{p_s}\Big) > -\frac{h\log\log N}{e\log eh} > -0.2h\log\log N.$$

Indeed, since $p_1 > 10h$ we have by [5] (formula at the bottom of p. 87)

$$\sum_{s=1}^{r}\frac{1}{p_s^2}\leqslant\frac{2.04}{10h\log 10h}.$$

Hence

$$(9)\qquad \log\prod_{s=1}^{r}\Big(1-\frac{h}{p_s}\Big)+\log\prod_{s=1}^{r}\Big(1-\frac{1}{p_s}\Big)^{-h}\geqslant-\sum_{s=1}^{r}\sum_{m=2}^{\infty}\frac{1}{m}\Big(\frac{h}{p_s}\Big)^m$$

$$\geqslant-\frac{1}{2}\sum_{s=1}^{r}\Big(\frac{h}{p_s}\Big)^2\frac{1}{1-h/p_s}\geqslant-\frac{5}{9}h^2\sum_{s=1}^{r}\frac{1}{p_s^2}\geqslant-\frac{0.2h}{\log 10h}.$$

On the other hand, by [5], Theorem 15

$$(10)\qquad \frac{A}{\varphi(A)}\prod_{s=1}^{r}\Big(1-\frac{1}{p_s}\Big)^{-1}=\frac{AN}{\varphi(AN)}< e^C\log\log AN+\frac{2.51}{\log\log AN}.$$

Since by [5], Theorem 9, and by (3)

$$(11)\qquad \log A<11h<0.1\log N$$

we get

$$e^C\log\log AN+\frac{2.51}{\log\log AN}<e^C\log\log N+\frac{e^C}{10}+\frac{2.51}{5}<e^C(\log\log N+0.4).$$

Further by [5], Theorem 8

$$\frac{A}{\varphi(A)}>e^C\log 10h\Big(1-\frac{1}{2\log^2 10h}\Big)>e^C(\log h+2.1).$$

Since by (3) $\log\log N>\log 10h$ we get from (9), (10) and the last two inequalities

$$(12)\qquad \log\prod_{s=1}^{r}\Big(1-\frac{h}{p_s}\Big)>-h\Big(\log(\log\log N+0.4)-\log(\log h+2.1)+\frac{0.2}{\log 10h}\Big)$$

$$>-h(\log\log\log N-\log\log eh).$$

Clearly, $\log x-\log a=1+\log(x/ae)\leqslant x/ae$. Thus (12) implies (8). Now by (6) and (8)

$$(t-1)\log 1.17\leqslant\frac{h\log\log N}{e\log eh}<\frac{h\log\log N}{e\log(h+1)},$$



hence

$$(13)\qquad (2t+1)\log(h+1)<3\log(h+1)+\frac{2h\log\log N}{e\log 1.17}$$

$$<3\log(h+1)+4.7h\log\log N.$$

This inequality permits to estimate $R$. The estimation of $R$ given in [6] is not quite correct and not applicable under the present circumstances. Since $p_s$ is certainly greater than the $s$th prime, we have by [5], Corollary to Theorem 3, $p_s>s\log s$. Hence

$$\log\pi_n=\sum_{r_{n-1}\geqslant s>r_n}\log\Big(1-\frac{h}{p_s}\Big)>-\frac{10}{9}\sum_{r_{n-1}\geqslant s>r_n}\frac{h}{p_s}$$

$$>-\frac{10}{9}h\int_{r_n}^{r_{n-1}}\frac{dt}{t\log t}=-\frac{10}{9}h\log\frac{\log r_{n-1}}{\log r_n}.$$

It follows by (6)

$$\frac{\log r_n}{\log r_{n-1}}<\Big(\frac{1}{1.17}\Big)^{\frac{9}{10h}}<\Big(1+\frac{9}{10h}\log 1.17\Big)^{-1}\leqslant(1+0.141h^{-1})^{-1},$$

and by induction

$$(14)\qquad \frac{\log r_n}{\log r}<(1+0.141h^{-1})^{-n}\qquad(1\leqslant n\leqslant t-1).$$

On the other hand

$$\log N\geqslant\sum_{s=1}^{r}\log p_s>r\log 10h\geqslant r\log 20,$$

thus $\log r<\log\log N-1$.

It follows from (5), (13) and (14) that

$$\log R\leqslant(2t+1)\log(h+1)+\log r+2\sum_{n=1}^{t-1}\log r_n$$

$$<3\log(h+1)+4.7h\log\log N+(\log\log N-1)\Big(2\sum_{n=0}^{\infty}(1+0.141h^{-1})^{-n}-1\Big)$$

$$<3\log(h+1)+4.7h\log\log N+(\log\log N-1)(14.2h+1)$$

$$<19.4h\log\log N-11h-1.$$

Since by (11) $\log A<11h$, we have

$$(15)\qquad \log R<19.4h\log\log N-\log A-1.$$

It follows from (7), (8) and (15) that

$$\log\left(\frac{E}{A}(\log N)^{20h}\right) > \log R$$

thus by (4)

$$P\left(A, (\log N)^{20h}, p_1, \ldots, p_r\right) > 0$$

and by the definition of $P$ there exists an integer $D$ satisfying (1) and (2), q.e.d.

Lemma 2. *Let $f_j(x_1, \ldots, x_n)$ $(1 \leqslant j \leqslant n)$ be polynomials of degrees $m_1, m_2, \ldots, m_n$ respectively, with coefficients in a number field $K$. If*

$$f_j(\xi_1, \ldots, \xi_n) = 0 \quad (1 \leqslant j \leqslant n)$$

*and*

(16) $$\frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)}(\xi_1, \ldots, \xi_n) \neq 0$$

*then*

$$[K(\xi_1, \ldots, \xi_n): K] \leqslant m_1 m_2 \ldots m_n.$$

Proof (due to H. Davenport). Let $\varphi_1(x_1, \ldots, x_n), \ldots, \varphi_n(x_1, \ldots, x_n)$ be complete polynomials of degrees $m_1, \ldots, m_n$ respectively, with arbitrary complex coefficients which differ by less than $\varepsilon$ in absolute value from the corresponding coefficients of $f_1, \ldots, f_n$. By Bezout's theorem, the equations $\varphi_1 = 0, \ldots, \varphi_n = 0$ have exactly $m_1 m_2 \ldots m_n$ distinct solutions for "general" values of all the coefficients. We shall prove that one of these solutions tends to $\xi_1, \ldots, \xi_n$ as $\varepsilon \to 0$.

This will suffice to prove the result. Indeed, the equations $f_j(x_1, \ldots, x_n) = 0$ $(j = 1, \ldots, n)$ define a union of algebraic varieties over $K$. If the point $(\xi_1, \ldots, \xi_n)$ were on a variety of positive dimension, defined by the equations $g_i(x_1, \ldots, x_n) = 0$ $(i = 1, \ldots, N)$, where $g_i = f_i$ for $i \leqslant n$, then by a known theorem ([3], p. 84) the rank of the matrix

$$\left[\frac{\partial g_i}{\partial x_j}(\xi_1, \ldots, \xi_n)\right]$$

would be less than $n$, contrary to (16). Hence $(\xi_1, \ldots, \xi_n)$ is an isolated point, and therefore the $\xi_i$ are algebraic over $K$. Now consider the points $(\xi_1^{(\nu)}, \ldots, \xi_n^{(\nu)})$ which are algebraically conjugate to $(\xi_1, \ldots, \xi_n)$ over $K$. These are distinct and their number is $[K(\xi_1, \ldots, \xi_n): K]$. Also each of them satisfies the equations $f_j = 0$ and the condition $\dfrac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)} \neq 0$.

Hence it will follow from the result stated above that near each of them there is one of the solutions of $\varphi_1 = 0, \ldots, \varphi_n = 0$ and so their number is at most $m_1 m_2 \ldots m_n$.



The value of $\varphi_j(\xi_1, \ldots, \xi_n)$, or of any derivative of $\varphi_j(x_1, \ldots, x_n)$ at $(\xi_1, \ldots, \xi_n)$, differs from the corresponding value for $f_j(\xi_1, \ldots, \xi_n)$ by an amount that is $O(\varepsilon)$. Hence

$$\varphi_j(\xi_1+\eta_1, \ldots, \xi_n+\eta_n) = \varepsilon_j + \sum_{i=1}^{n}(\lambda_{ij}+\varepsilon_{ij})\eta_i + \sum_{i_1=1}^{n}\sum_{i_2=1}^{n}(\lambda_{i_1 i_2 j}+\varepsilon_{i_1 i_2 j})\eta_{i_1}\eta_{i_2} + \ldots,$$

where all $\varepsilon_j, \varepsilon_{i_1 j}, \ldots$ are $O(\varepsilon)$ and where the numbers $\lambda_{ij}, \lambda_{i_1 i_2 j}, \ldots$ are partial derivatives of $f_j$ at $(\xi_1, \ldots, \xi_n)$ and so are independent of $\varepsilon$. Also

$$\det \lambda_{ij} = \frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)}(\xi_1, \ldots, \xi_n) \neq 0.$$

It follows from the well known process for the inversion of power series (e.g. by iteration) that the equations

$$\varphi_j(\xi_1+\eta_1, \ldots, \xi_n+\eta_n) = 0 \quad \text{for} \quad j = 1, \ldots, n$$

have a solution with $\eta_1, \ldots, \eta_n = O(\varepsilon)$. Hence the result.

Remark. The above proof fails if $K$ has characteristic $\neq 0$. However, Mr. Swinnerton-Dyer tells me that the lemma is still valid and can be proved by using Weil's theory of intersections.

Proof of Theorem 1. The theorem clearly holds for $N < 3$. Assume that $N \geqslant 3$,

(17) $$\sum_{i=1}^{k} a_i \zeta_N^{a_i} = \vartheta, \quad |Q(\vartheta)| = d, \quad (N, a_1, \ldots, a_k) = 1.$$

Let $D$ be an integer whose existence is ensured by Lemma 1 for $h = k-1$. Among the numbers $a_i$ let there be exactly $n$ that are distinct mod $N_1 = N/(N, D)$. By a suitable permutation of the terms in (17) we can achieve that $a_{s_1}, a_{s_2}, \ldots, a_{s_n}$ are all distinct mod $N_1$, $0 = s_0 < s_1 < \ldots < s_n = k$ and

(18) $$a_i \equiv a_{s_\nu} \bmod N_1 \quad \text{if} \quad s_{\nu-1} < i \leqslant s_\nu \quad (1 \leqslant \nu \leqslant n).$$

Let us choose numbers $\gamma_\nu$, such that

(19) $$\gamma_\nu \equiv a_{s_\nu} \pmod{N_1}, \quad (\gamma_\nu, N) = (a_{s_\nu}, N_1) \quad (1 \leqslant \nu \leqslant n).$$

It follows from elementary congruence considerations that such choice is possible.

We write equation (17) in the form

(20) $$\sum_{\nu=1}^{n} \zeta_N^{\gamma_\nu} S_\nu = \vartheta,$$

where

$$S_\nu = \sum_{i=s_{\nu-1}+1}^{s} a_i \zeta_N^{a_i-\gamma_\nu} \quad (1 \leqslant \nu \leqslant n).$$

By (18) and (19)

$$S_\nu \in Q(\zeta_D) \quad (1 \leqslant \nu \leqslant n).$$

By (2) $(N, jD-D+1) = 1$ thus $\zeta_N^{jD-D+1}$ is for each positive $j \leqslant k$ a conjugate of $\zeta_N$. Clearly

$$\zeta_N^{(a_i-\gamma_\nu)(jD-D+1)} = \zeta_N^{a_i-\gamma_\nu} \quad (s_{\nu-1} < i \leqslant s_\nu).$$

Substituting $\zeta_N^{jD-D+1}$ for $\zeta_N$ in (20) we get

$$\sum_{\nu=1}^{n} \zeta_N^{\gamma_\nu(jD-D+1)} S_\nu = \vartheta_j \quad (1 \leqslant j \leqslant n),$$

where $\vartheta_j$ is a suitable conjugate of $\vartheta$. Since $Q(\vartheta)$ is an Abelian field, $\vartheta_j \in Q(\vartheta)$.

In Lemma 2 we take:

$$f_j(x_1, \ldots, x_n) = \sum_{\nu=1}^{n} x_\nu^{jD-D+1} S_\nu - \vartheta_j \quad (1 \leqslant j \leqslant n),$$

$$K = Q(\zeta_D, \vartheta), \qquad \xi_\nu = \zeta_N^{\gamma_\nu} \quad (1 \leqslant \nu \leqslant n).$$

Hence

$$(21) \qquad \frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)}(\xi_1, \ldots, \xi_n) = \prod_{j=1}^{n}(jD-D+1) \prod_{\nu=1}^{n} S_\nu \prod_{1 \leqslant \nu'' < \nu' \leqslant n} (\zeta_N^{\gamma_{\nu'}D} - \zeta_N^{\gamma_{\nu''}D}).$$

If $S_\nu = 0$ for some $\nu \leqslant n$ then

$$\sum_{i=s_{\nu+1}+1}^{s_\nu} a_i \zeta_N^{a_i} = 0$$

and the theorem holds with $I = \{s_{\nu-1}+1, \ldots, s_\nu\}$.

If $S_\nu \neq 0$ for all $\nu \leqslant n$, then by (21) and the choice of $\gamma_\nu$ we have

$$\frac{\partial(f_1, \ldots, f_n)}{\partial(x_1, \ldots, x_n)}(\xi_1, \ldots, \xi_n) \neq 0.$$

Therefore, by Lemma 2

$$(22) \qquad |Q(\zeta_N^{\gamma_1}, \zeta_N^{\gamma_2}, \ldots, \zeta_N^{\gamma_n})| \leqslant |Q(\zeta_D, \vartheta)| \prod_{j=0}^{n-1} (jD+1) < n!\, D^n d \leqslant k!\, D^k d.$$



On the other hand by (18) and (19)

$$(N, \gamma_\nu) = (N_1, a_{s_\nu}) = (N_1, a_{s_{\nu-1}+1}, \ldots, a_{s_\nu}),$$

hence

$$(N, \gamma_1, \ldots, \gamma_n) = (N_1, a_1, \ldots, a_k) = 1$$

and

$$|Q(\zeta_N^{\gamma_1}, \ldots, \zeta_N^{\gamma_n})| = \varphi(N).$$

It follows now from (22) and (1) (applied with $h = k-1$)

$$(23) \qquad \varphi(N) \leqslant k!(\log N)^{20k(k-1)} d.$$

If $N < (200k^2 \log 2k)^{20k^2}$ the theorem certainly holds.

If $N \geqslant (200k^2\log 2k)^{20k^2} > 10^{42}$, it follows from [5], Theorem 15, that

$$(24) \qquad \varphi(N) > \frac{N}{\log N}.$$

Also, if $N \geqslant (200k^2\log 2k)^{20k^2}$

$$(25) \qquad k! < (\log N)^k.$$

It follows from (23), (24) and (25) that

$$N(\log N)^{-20k^2} \leqslant d.$$

Taking $N_0 = d(2\log d + 200k^2\log 2k)^{20k^2}$ we find that

$$N_0(\log N_0)^{-20k^2} = d\left(\frac{2\log d + 200k^2\log 2k}{\log d + 20k^2\log(2\log d + 200k^2\log 2k)}\right)^{20k^2} > d,$$

because $200k^2\log 2k > 20k^2\log 400k^2\log 2k$.

Since the function $N(\log N)^{-20k^2}$ is increasing for $N > e^{20k^2}$ it follows that $N < N_0$. The proof is complete.

Proof of Corollary 1. Assume that

$$\vartheta = \sum_{i=1}^{k} \zeta_N^{a_i}.$$

Let $I$ be a set contained in $\{1, 2, \ldots, k\}$ saturated with respect to the property that $\sum_{i\in I} \zeta_N^{a_i} = 0$. We have $\vartheta = \sum_{i \notin I} \zeta_N^{a_i}$ and by the choice of $I$ and Theorem 1

$$\frac{N}{(N, \underset{i\notin I}{GCD}\, a_i)} < d(2\log d + 200k^2\log 2k)^{20(k-\varkappa)^2},$$

where $\varkappa$ is the number of elements in $I$. If $\varkappa = 0$ we have the desired conclusion, if $\varkappa > 0$ then $\varkappa \geqslant 2$ and

$$\vartheta = \begin{cases} \sum_{i \notin I} \zeta_N^{a_i} + \sum_{j=1}^{\varkappa/2} 1 + \sum_{j=1}^{\varkappa/2} (-1), & \varkappa \text{ even}, \\ \sum_{i \notin I} \zeta_N^{a_i} + \zeta_3 + \zeta_3^{-1} + \sum_{j=1}^{(\varkappa-1)/2} 1 + \sum_{j=1}^{(\varkappa-3)/2} (-1), & \varkappa \text{ odd} \geqslant 3. \end{cases}$$

The least common degree of all $k$ roots of unity occuring in the above representation of $\vartheta$ does not exceed

$$6d(2\log d + 200k^2 \log 2k)^{20(k-\varkappa)^2} < d(2\log d + 200k^2 \log 2k)^{20k^2},$$

which completes the proof.

Proof of Corollary 2. The sufficiency of the condition is immediate since

$$\sum_{i=1}^{k-2} \zeta_N^{a_i} = \sum_{i=1}^{k-2} \zeta_N^{a_i} + \zeta_M - \zeta_M,$$

where $M$ is arbitrary. On the other hand, if $\vartheta$ has infinitely many representations as the sum of $k$ roots of unity, then there must be among them a representation

$$\vartheta = \sum_{i=1}^{k} \zeta_N^{a_i}, \quad (N, a_1, \ldots, a_k) = 1$$

not satisfying the inequality

$$N < d(2\log d + 200k^2 \log 2k)^{20k^2}.$$

By Theorem 1 there is a non-empty set $I \subset \{1, 2, \ldots, k\}$ such that $\sum_{i \in I} \zeta_N^{a_i} = 0$ and denoting by $\varkappa$ the number of elements in $I$ we have $k > \varkappa \geqslant 2$. Since

$$1 = \begin{cases} \sum_{j=1}^{\varkappa/2} 1 + \sum_{j=1}^{(\varkappa-2)/2} (-1), & \varkappa \text{ even}, \\ \zeta_6 + \zeta_6^{-1} + \sum_{j=1}^{(\varkappa-3)/2} 1 + \sum_{j=1}^{(\varkappa-3)/2} (-1), & \varkappa \text{ odd} \geqslant 3 \end{cases}$$

we can replace one of the $k-\varkappa$ terms in the sum $\sum_{i \notin I} \zeta_N^{a_i} = \vartheta$ by a sum of $\varkappa - 1$ roots of unity, thus obtaining a representation of $\vartheta$ as the sum of $k-2$ roots of unity.

Proof of Theorem 2. Suppose that

$$(26) \qquad \sqrt{5}\cos\frac{\pi}{M} + i\sin\frac{\pi}{M} = \zeta_{m_1}^{a_1} + \zeta_{m_2}^{a_2} + \zeta_{m_3}^{a_3}, \quad \text{where} \quad (a_i, m_i) = 1.$$



Put $N = 5[2M, m_1, m_2, m_3]$, $\alpha = \dfrac{N}{2M}$, $\beta = \dfrac{Na_1}{m_1}$, $\gamma = \dfrac{Na_2}{m_2}$, $\delta = \dfrac{Na_3}{m_3}$.

Then

$$(27) \qquad (\alpha, \beta, \gamma, \delta) = 5.$$

Since $\frac{1}{2}(\sqrt{5}-1) = \zeta_5 + \zeta_5^{-1} = \zeta_N^{N/5} + \zeta_N^{-N/5}$ (26) can be written in the form

$$(28) \qquad (\zeta_N^{N/5} + \zeta_N^{-N/5})(\zeta_N^{\alpha} + \zeta_N^{-\alpha}) + \zeta_N^{\alpha} = \zeta_N^{\beta} + \zeta_N^{\gamma} + \zeta_N^{\delta}.$$

Now we distinguish two cases according as $3 \mid N$ and $3 \nmid N$. In the first case at least one of the numbers $\pm\frac{1}{3}N + 1$ is relatively prime to $N$. Hence one of the numbers $\zeta_3^{\pm 1}\zeta_N$ is conjugate to $\zeta_N$. Denote it for simplicity by $\varrho\zeta_N$ and substitute for $\zeta_N$ into (28). Since $\varrho^{N/5} = 1$, we get

$$(29) \qquad (\zeta_N^{N/5} + \zeta_N^{-N/5})(\varrho^{\alpha}\zeta_N^{\alpha} + \varrho^{-\alpha}\zeta_N^{-\alpha}) + \varrho^{\alpha}\zeta_N^{\alpha} = \varrho^{\beta}\zeta_N^{\beta} + \varrho^{\gamma}\zeta_N^{\gamma} + \varrho^{\delta}\zeta_N^{\delta}.$$

By taking complex conjugates of (28) and (29) and substituting afterwards

$$y = \zeta_N^{\beta}, \quad z = \zeta_N^{\gamma}, \quad t = \zeta_N^{\delta};$$

$$A = \tfrac{1}{2}(\sqrt{5}+1)\zeta_{2M} + \tfrac{1}{2}(\sqrt{5}-1)\zeta_{2M}^{-1}, \quad B = \tfrac{1}{2}(\sqrt{5}-1)\zeta_{2M} + \tfrac{1}{2}(\sqrt{5}+1)\zeta_{2M}^{-1},$$

$$(30) \qquad C = \tfrac{1}{2}(\sqrt{5}+1)\varrho^{\alpha}\zeta_{2M} + \tfrac{1}{2}(\sqrt{5}-1)\varrho^{-\alpha}\zeta_{2M}^{-1},$$

$$D = \tfrac{1}{2}(\sqrt{5}-1)\varrho^{\alpha}\zeta_{2M} + \tfrac{1}{2}(\sqrt{5}+1)\varrho^{-\alpha}\zeta_{2M}^{-1}$$

we get the following system of equations

$$(31) \qquad A = y + z + t,$$

$$(32) \qquad B = y^{-1} + z^{-1} + t^{-1},$$

$$(33) \qquad C = \varrho^{\beta}y + \varrho^{\gamma}z + \varrho^{\delta}t,$$

$$(34) \qquad D = \varrho^{-\beta}y^{-1} + \varrho^{-\gamma}z^{-1} + \varrho^{-\delta}t^{-1}.$$

If $\beta \equiv \gamma \equiv \delta \pmod 3$ it follows from (31) and (33) that $C = \varrho^{\beta}A$. Hence by (30)

$$(35) \qquad \tfrac{1}{2}(\sqrt{5}+1)(\varrho^{\alpha} - \varrho^{\beta})\zeta_{2M} + \tfrac{1}{2}(\sqrt{5}-1)(\varrho^{-\alpha} - \varrho^{\beta})\zeta_{2M}^{-1} = 0.$$

The coefficients of $\zeta_{2M}$ and $\zeta_{2M}^{-1}$ do not both vanish, since that would give $\alpha \equiv \beta \equiv 0 \pmod 3$ and $\alpha \equiv \beta \equiv \gamma \equiv \delta \equiv 0 \pmod 3$ contrary to (27). Thus they have different absolute values, and (35) is impossible.

Consider now the case when exactly two among the numbers $\beta, \gamma, \delta$ are congruent mod 3, e.g. $\beta \equiv \gamma \not\equiv \delta \pmod 3$. Eliminating $y, z$, and $t$ from the equations (31) to (34) we get

$$(36) \qquad (C - \varrho^{\beta}A)(D - \varrho^{-\beta}B) = |\varrho^{\delta} - \varrho^{\beta}|^2 = 3.$$

Substituting the values for $A, B, C, D$ from (30) we obtain

$$(37)\qquad (\varrho^\alpha-\varrho^\beta)(\varrho^\alpha-\varrho^{-\beta})\zeta_M+\tfrac{1}{2}(3+\sqrt{5})|\varrho^\alpha-\varrho^\beta|^2+\tfrac{1}{2}(3-\sqrt{5})|\varrho^{-\alpha}-\varrho^\beta|^2-3+ +(\varrho^{-\alpha}-\varrho^\beta)(\varrho^{-\alpha}-\varrho^{-\beta})\zeta_M^{-1}=0.$$

If $\beta\equiv\pm\alpha \pmod 3$, we get $\frac{1}{2}(3\mp\sqrt{5})|\varrho^{\mp\alpha}-\varrho^\beta|^2-3=0$, which is mpossible. Hence $\beta\not\equiv\pm\alpha \pmod 3$ and (37) takes the form

$$(38)\qquad 3\zeta_M+6+3\zeta_M^{-1}=0,\qquad \text{if}\quad \beta\not\equiv 0 \pmod 3;$$

$$(39)\qquad -3\varrho^\alpha\zeta_M+6-3\varrho^{-\alpha}\zeta_M^{-1}=0,\qquad \text{if}\quad \beta\equiv 0\not\equiv\alpha \pmod 3.$$

It follows from (38) that $\zeta_M=-1$, $M=2$ and from (39) $\varrho^\alpha\zeta_M=1$, $M=3$.

Consider next the case when $\beta,\gamma,\delta$ are all different mod 3. We can assume without lost of generality that $\beta\equiv 0 \pmod 3$, $\gamma\equiv 1 \pmod 3$, $\delta\equiv 2 \pmod 3$.

If $\alpha\equiv 0 \pmod 3$, then $C=A$ and it follows from (31) and (33) that

$$A-y=z+t=\varrho z+\varrho^2 t,$$

hence $t=\varrho z$ and

$$(40)\qquad A=y-\varrho^2 z,\qquad B=y^{-1}-\varrho z^{-1}.$$

Since $y$ and $z$ are roots of unity, $|y-\varrho^2 z|\leqslant 2$. On the other hand by (30)

$$|A|=|\sqrt{5}\cos(\pi/M)+i\sin(\pi/M)|=\sqrt{5-4\sin^2(\pi/M)}.$$

It follows that

$$5-4\sin^2(\pi/M)\leqslant 4,\qquad |\sin(\pi/M)|\geqslant\tfrac{1}{2},$$

and $6\geqslant M>1$. Further, by (40)

$$-\varrho^2 yz=\frac{A}{B}=\frac{\sqrt{5}\cos(\pi/M)+i\sin(\pi/M)}{\sqrt{5}\cos(\pi/M)-i\sin(\pi/M)}.$$

It can easily be verified that for $M=3$, 4 or 6 the quotient on the right hand side is not an algebraic integer, hence the only possible values for $M$ here are $M=2$ or 5.

If $\alpha\not\equiv 0 \pmod 3$, then eliminating $y$, $z$ and $t$ from (31) to (34) we get

$$A^3-C^3=3yzt(AB-CD)\quad\text{and}\quad (AB-CD)^2-\tfrac{1}{9}(A^3-C^3)(B^3-D^3)=0.$$

The substitution of the values for $A, B, C, D$ from (30) gives

$$-3\varrho^{-\alpha}\zeta_M^2+3\varrho^\alpha\zeta_M-3+3\varrho^{-\alpha}\zeta_M^{-1}-3\varrho^\alpha\zeta_M^{-2}=0.$$



Hence

$$(\varrho^\alpha\zeta_M)^4-(\varrho^\alpha\zeta_M)^3+(\varrho^\alpha\zeta_M)^2-(\varrho^\alpha\zeta_M)+1=0,\qquad \varrho^\alpha\zeta_M=\zeta_{10}^\varepsilon,$$

where $(\varepsilon, 10)=1$ and $\zeta_M=\varrho^{-\alpha}\zeta_{10}^\varepsilon$. This gives $M=30$.

It remains to consider the case when $3\nmid N$. In this case $\zeta_N^3$ is a conjugate of $\zeta_N$ and substituting it for $\zeta_N$ in the equation (28) we get

$$(41)\qquad (\zeta_N^{3N/5}+\zeta_N^{-3N/5})(\zeta_N^{3\alpha}+\zeta_N^{-3\alpha})+\zeta_N^{3\alpha}=\zeta_N^{3\beta}+\zeta_N^{3\gamma}+\zeta_N^{3\delta}.$$

Now,

$$\zeta_N^{3N/5}+\zeta_N^{-3N/5}=\tfrac{1}{2}(-\sqrt{5}-1).$$

By taking the complex conjugate of (41) and substituting afterwards

$$(42)\qquad E=\tfrac{1}{2}(-\sqrt{5}+1)\zeta_{2M}^3+\tfrac{1}{2}(\sqrt{5}-1)\zeta_{2M}^{-3},$$
$$F=\tfrac{1}{2}(-\sqrt{5}-1)\zeta_{2M}^3+\tfrac{1}{2}(-\sqrt{5}+1)\zeta_{2M}^{-3}$$

we get the following system of equations

$$A=y+z+t,$$
$$B=y^{-1}+z^{-1}+t^{-1},$$
$$E=y^3+z^3+t^3,$$
$$F=y^{-3}+z^{-3}+t^{-3}.$$

Eliminating $y$, $z$ and $t$ we obtain

$$A^3-E=3yzt(AB-1)\quad\text{and}\quad (AB-1)^2-\tfrac{1}{9}(A^3-E)(B^3-F)=0.$$

The substitution of the values for $A, B, E, F$ from (30) and (42) gives

$$-\zeta_M^3-\zeta_M^2-\zeta_M^{-2}-\zeta_M^{-3}=0.$$

Hence

$$\zeta_M^6+\zeta_M^5+\zeta_M+1=(\zeta_M+1)(\zeta_M^5+1)=0,$$

$$\zeta_M=-1\ \text{or}\ \zeta_M^5=-1,\quad\text{and}\quad M=2\ \text{or}\ M=10.$$

This completes the proof that the only values $M$ for which $\eta_M=\sqrt{5}\cos(\pi/M)+i\sin(\pi/M)$ can be a sum of three roots of unity are 2, 3, 5, 10, or 30. On the other hand, it is easy to verify that

$$\eta_2=1+\zeta_2+\zeta_4,\qquad \eta_3=\zeta_5+\zeta_5^{-1}+\zeta_6,\qquad \eta_5=\zeta_6+\zeta_6^{-1}+\zeta_{10},$$
$$\eta_{10}=\zeta_{20}+\zeta_{20}^3+\zeta_{20}^{-3},\qquad \eta_{30}=\zeta_{12}^{-1}+\zeta_{20}^{-1}+\zeta_{60}^{11}.$$

Proof of Corollary 4. Since $1+2i\cos(\pi/M)=1+i(\zeta_{2M}+\zeta_{2M}^{-1})$, any number equivalent to $1+2i\cos(\pi/M)$ is a sum of three roots of unity. It follows by Theorem 2 that the numbers $\xi_M=1+2i\cos(\pi/M)$ and $\eta_M=\sqrt{5}\cos(\pi/M)+i\sin(\pi/M)$ can be equivalent only for $M=2$, 3, 5, 10, or 30.

If the numbers $\xi_3$ and $\eta_3$ or $\xi_5$ and $\eta_5$ were equivalent then since $\xi_3 = 1+i$ and $\eta_5 = 1+\zeta_{10}$, $\eta_3$ or $\xi_5$ would be a sum of two roots of unity. However if $\vartheta \neq 0$ is such a sum and $\bar{\vartheta}$ is its complex conjugate, then $\vartheta/\bar{\vartheta}$ is a root of unity. Since neither of the numbers $\eta_3/\bar{\eta}_3$ and $\xi_5/\bar{\xi}_5$ is an algebraic integer, the proof is complete.

Added in proof. 1. H. B. Mann has proved in Mathematika 12 (1965), pp. 107-117, that under the assumptions of Corollary 3, $N$ divides the product of all primes $\leqslant k+1$. This leads to a much better estimation of $N$ than that stated in the corollary. Mann's method could also be used to solve both Robinsons's problems considered in this paper.

2. In connection with Lemma 1 the question arises how much inequality (1) can be improved. Y. Wang has proved by Brun's method in a manuscript kindly placed at my disposal that for $N > N_0(h)$ one can replace $(\log N)^{20h}$ by $c(h) \times (\log N)^{4h+3}$. According to H. Halberstam (written communication), there is a possibility of reducing the exponent $4h+3$ to $2h+1$ by Selberg's method.

## References

[1] V. Brun, *Le crible d'Eratosthène et le théorème de Goldbach*, Norsk Videnskaps Selskabs Skrifter, Kristiania 1920.

[2] R. D. Carmichael, *Introduction to the Theory of Groups of Finite Order*, New York 1937.

[3] W. Hodge and D. Pedoe, *Methods of Algebraic Geometry II*, Cambridge 1952.

[4] R. M. Robinson, *Some conjectures about cyclotomic integers*, Math. Comp. 19 (1965), pp. 210-217.

[5] J. B. Rosser and L. Schoenfeld, *Approximate formulas for some functions of prime numbers*, Illinois J. Math. 6 (1962), pp. 64-89.

[6] A. Schinzel and Y. Wang, *A note on some properties of the functions* $\varphi(n)$, $\sigma(n)$ and $\theta(n)$, Ann. Polon. Math. 4 (1958), pp. 201-213.

*Reçu par la Rédaction le 9. 7. 1965*

---

ACTA ARITHMETICA
XI (1966)

# A refinement of a theorem of Schur on primes in arithmetic progressions

by

J. Wójcik (Warszawa)

I. Schur ([1]) has given a purely algebraic proof of the following special case of Dirichlet's theorem on arithmetic progression.

Let $l^2 \equiv 1 \bmod m$. If the arithmetic progression $mz+l$ contains a prime $> \frac{1}{2}\varphi(m)$, then it contains infinitely many primes.

In this paper by a refinement of Schur's method we prove

THEOREM. *Let $l^2 \equiv 1 \bmod m$. If the arithmetic progression $mz+l$ contains a prime, then it contains infinitely many primes.*

Let $Q$ be the rational field, $\zeta_m$ a primitive $m$th root of unity,

$$h(x) = \begin{cases} x+x^l & \text{if} \quad 2l \not\equiv m+2 \bmod 2m, \\ x^2 & \text{if} \quad 2l \equiv m+2 \bmod 2m, \end{cases}$$

$K = Q(h(\zeta_m))$.

Let $r$ be the degree of $K$, $N$ denote the norm from $K$ to $Q$.

LEMMA 1. *Let $a$ be any integral generating element of $K$, $a_1, \ldots, a_r$ $(a_1 = a)$ all its conjugates,*

$$G(x, y) = \prod_{i=1}^{r} (x - a_i y), \qquad d \text{ the discriminant of } G.$$

*If $q$ is a prime, $x, y$ rational integers, $q \mid G(x, y)$, $q \nmid mdy$, then $q$ is of the form $mz+1$ or $mz+l$.*

Proof. $a = \chi(h(\zeta_m))$, where $\chi$ is a polynomial with rational coefficients and since $a$ is a generating element of $K$

(1) $$\chi(h(\zeta_m^{s_1})) = \chi(h(\zeta_m^{s_2})),$$

where

(2) $$(s_1, m) = (s_2, m) = 1$$

implies

(3) $$h(\zeta_m^{s_1}) = h(\zeta_m^{s_2}).$$