ON THE MINIMAL NUMBER OF TERMS OF THE SQUARE OF A POLYNOMIAL.

BY A. RÉNYI IN BUDAPEST

L. Rédei proposed the following problem: Let the number of terms of a polynomial be given. Find the minimal number of terms of its square.

Let us denote the minimal number of terms of the square of a polynomial containing \(n\) terms, by \(Q(n)\), and let us have

\[
q(n) = \frac{Q(n)}{n}
\]

In May, 1945 L. Kalmár, L. Rédei and the author discussed this problem and found the following result:

\[
\lim \inf \ q(n) = 0 \quad (1)
\]

Starting from this result I found recently the following stronger theorem:

\[
\lim_{n \to \infty} \frac{q(1) + q(2) + \ldots + q(n)}{n} = 0 \quad (2)
\]

In what follows this theorem shall be proved.

I would at this occasion express my thanks to L. Kalmár and L. Rédei for having kindly agreed to my publishing and using in this paper the results of our joint work.

It may be mentioned, that the problem is far from being exhausted. It seems likely, that

\[
\lim_{n \to \infty} q(n) = 0
\]

holds also, but the proof of this conjecture would require further refinements of our method. Corresponding problems arise in connection with the 3th, 4th etc. powers of polynomials. We hope to return to these problems at an other occasion.
We begin by proving a number of Lemmas.

**Lemma A.** \( q (n \cdot m) \leq q (n) \cdot q (m) \)

**Proof:** If \( R_n (x) \) and \( R_m (x) \) are the polynomials of \( n \) resp. \( m \) terms, for which the minimal numbers \( Q (n) \) resp. \( Q (m) \) of the terms of the squares, are reached, and if the degree of \( R_n (x) \) is \( k_n \), the polynomial

\[
R_n (x) \cdot R_m (x^{\ast n + 1})
\]

consists of \( n \cdot m \) terms and its square consists of at most \( Q (n) \cdot Q (m) \) terms, which proves Lemma A.

**Lemma B.** \( q (n) \leq 1 + \frac{1}{n} \leq \frac{3}{2}; \quad n \geq 2. \)

**Proof:** Let us have

\[
\sqrt{1 + \frac{1}{x}} = 1 + x + a_2 x^2 + \ldots + a_n x^n + \ldots \quad (3)
\]

The sum of the first \( n \) terms of this expansion shall be denoted by \( S_n (x) \). \( S_n (x) \) is a polynomial of degree \( n - 1 \), and we have

\[
\sqrt{1 + \frac{1}{x}} = S_n (x) + x^n r (x) . \quad (4)
\]

By squaring both sides of (4) we obtain

\[
S^2_n (x) = 1 + 2x - x^n [2 S_n (x) r (x) + x^n r^2 (x)] \quad (5)
\]

It can be seen from (5) that \( S^2_n (x) \) does not contain the terms with exponents \( 2, 3, \ldots n - 1 \) \( \). Thus \( S^2_n (x) \) contains only \( n + 1 \) terms, viz. the terms with exponents \( 0, 1, n, n + 1, \ldots 2n - 2 \), which proves Lemma B.

**Lemma C.** \( q (2n + 1) \leq 1. \)

We start again from the expansion of \( \sqrt{1 + 2x} \). If \( \lambda = \sqrt{\frac{a_{n-1}}{a_{n+1}} (a_{n-1}} \) and \( a_{n+1} \) have evidently the same sign) we have in

\[
\sqrt{1 + \frac{1}{x}} = 1 + b_1 x + b_2 x^2 + \ldots + b_{n-1} x^{n-1} + b_n x^n + b_{n+1} x^{n+1} + \ldots \quad (6)
\]

1 These polynomials were used by Legendre for solving the congruence \( x^2 \equiv 8r + 1 \mod 2^m \). See e. g. I. V. Uspensky & M. A. Heaslet Elementary Number Theory Newyork, 1939. p. 311.
\[ b_{n-1} = b_{n+1} \]  

Let us construct the following symmetrical polynomial containing \(2n+1\) terms:

\[ P_{2n+1}(x) = 1 + b_1x + \ldots + b_{n-1}x^{n-1} + b_nx^n + b_{n-1}x^{n+1} + \ldots + b_1x^{2n-1} + x^2n. \]

Owing to (7) the first \(n + 2\) terms of \(P_{2n+1}(x)\) are identical with the terms of \(S_{n+2}(\lambda x)\). But, as it has been shown in the course of the proof of Lemma B, \(S_{n+1}(\lambda x)\) does not contain the terms with exponents 2, 3, \ldots \(n + 1\), and therefore these terms cannot occur in \(P_{2n+1}(x)\) either. \(P_{2n+1}(x)\) being symmetrical, the terms with exponents \(4n-2\), 
\(4n-3\), \(4n-4\), \ldots \(3n-1\) do not occur in \(P_{2n+1}(x)\) either. Thus \(P_{2n+1}(x)\) contains only the terms 0, 1, \(n+2\), \(n+3\), \ldots 3n-2, 4n-1, 4n; i.e. it contains \(2n+1\) terms, which proves Lemma C.

**Lemma D.**

\[ q(4n+1) \equiv \frac{28}{29}; \quad n \geq 7. \]

Let \(P_5(x)\) and \(S_n(x)\) have the meanings as defined above. Let us consider the polynomial

\[ A_{4n+1}(x) = P_5(x) \cdot S_n(x^4) \]

By carrying out the multiplication all terms with exponents from 0 to \(4n\) occur. Simple calculation shows that all coefficients are different from 0, i.e. \(A_{4n+1}(x)\) consists exactly of \(4n+1\) terms. \(P_5(x)\) contains the five terms with exponents 0, 1, 4, 7, 8. \(S_n^2(x^4)\) contains the \(n + 1\) terms with exponents 4k; \(k = 0, 1, n, n + 1, \ldots 2n-2\). It follows that \(A_{4n+1}^2(x)\) contains the terms with the exponents 0, 1, 4, 5, 7, 8, 11, 12, further terms with exponents from 4n to 8n, excepting the terms with exponents 4n+3, 8n-3, and excepting those with exponents 4+4k+2; \(k = 0, 1, \ldots n-1\). Thus \(A_{4n+1}^2(x)\) consists of

\[ 8 + 4n + 1 - n - 2 = 3n + 7 \]

terms.

It follows

\[ q(4n+1) \equiv \frac{3n + 7}{4n + 1} \equiv \frac{28}{29} \quad \text{if } n \geq 7. \]

Thus Lemma D has been proved.²

² At first the problem was formulated by L. Rédei as follows: Find a polynomial the square of which has a less number of terms than the polynomial itself. \(A_{4n}(x)\) is, in fact, a polynomial of this kind.
Lemma E. \[ q(n) \equiv C \cdot q(V(n)) \]

where \( V(n) \) denotes the number of different prime divisors of \( n \). \( C \) is an absolute constant, \( q \) an absolute constant with \( 0 < q < 1 \).

**Proof:** Let us consider a number \( n \) with \( V(n) = r \)

\[ n = p_1^{a_1} \cdot p_2^{a_2} \cdots p_r^{a_r}. \]

First we single out the different prime factors of \( n \): \( \pi_1, \pi_2, \ldots \pi_n \) which are of the form \( 4k + 1; k \geq 7 \). The prime factors of \( n \) having the form \( 4k + 3; k \geq 2 \), can be coupled in pairs:

\[ (\tau_1, \tau_3), (\tau_3 \tau_4), \ldots (\tau_{2^{r-1}}, \tau_{2v}), \]

leaving aside the last one if their number is odd. Each product \( \tau_{2^{r-1}} \tau_{2v} \) is of the form \( 4k + 1; k \geq 7 \). We have according to Lemma D

\[ q(\pi_j) \equiv \frac{28}{29}; \quad j = 1, 2, \ldots u; \quad q(\tau_{2^{i-1}} \tau_{2i}) \equiv \frac{28}{29}; \quad i = 1, 2, \ldots v. \]

We have further

\[ u + 2v \equiv r - 6 \quad (9) \]

because only the primes 2, 3, 5, 13, 17, and if necessary one more prime of the form \( 4k + 3 \) have been left aside. It follows from (9) a fortiori

\[ u + v \equiv \frac{r - 6}{2} = \frac{V(n)}{2} - 3. \quad (10) \]

The product of all factors which have not been selected shall be denoted by \( m \). We have

\[ u = \pi_1 \pi_2 \ldots \pi_u \cdot (\tau_1 \tau_3) \cdot (\tau_3 \tau_4) \ldots (\tau_{2^{r-1}} \tau_{2v}) \cdot m. \quad (11) \]

According to Lemma A, and with respect to (10) it follows

\[ q(n) \equiv \left( \frac{28}{29} \right)^{u+v} \cdot q(m) \equiv \frac{3}{2} \left( \frac{28}{29} \right)^{\frac{V(n)}{2} - 3} = \frac{3}{2} \left( \frac{29}{28} \right)^{3} \left( \left| \frac{28}{29} \right| \right)^{V(n)}. \quad (12) \]

Thus Lemma E has been proved with

\[ C = \frac{3}{2} \left( \frac{29}{28} \right)^{3} ; q = \left| \frac{28}{29} \right| < 1. \]
Now we turn to the proof of the Theorem:

\[
\lim_{N \to \infty} \frac{q(1) + q(2) + q(3) + \ldots + q(N)}{N} = 0.
\]

Proof: According to a theorem of Hardy and Ramanujan, the number of integers \( n \) not exceeding \( N \), for which

\[
V(n) < \frac{1}{2} \log\log N
\]

holds, is \( o(N) \).

If \( V(n) > \frac{\log\log N}{2} \), applying Lemma E, we have

\[
q(n) \leq C \cdot \frac{\log\log N}{2};
\]

if \( V(n) < \frac{\log\log N}{2} \) we obtain by use of Lemma B

\[
q(n) \leq \frac{3}{2}.
\]

It follows, from the theorem of Hardy and Ramanujan, that

\[
\frac{q(1) + q(2) + \ldots + q(N)}{N} \leq \frac{3}{2} \frac{o(N)}{N} + C \cdot \frac{\log\log N}{2}.
\]

Thus we have

\[
\lim_{N \to \infty} \frac{q(1) + q(2) + \ldots + q(N)}{N} = 0
\]

and our Theorem is therewith proved.

It may be mentioned that we used only polynomials with real coefficients. It is an open question, whether it would be possible to obtain better results by using polynomials with complex coefficients.

\footnote{Quart. Journal of Math. 48, 1920, p. 77. An elementary proof has been given by P. Turán, Journal of the London Math. Soc. 9, Part 4, p. 274—276. The theorem is much sharper in its original form, but we do not need more.}