CANCELLATION IN A SHORT EXPONENTIAL SUM

G. MOLTENI

Abstract. Let \( q \) be an odd integer, let \( \tau \) be the order of 2 modulo \( q \) and let \( \xi \) be a primitive \( q \)th root of unity. Upper bounds for \( \sum_{k=1}^{\tau} \xi^{2^k} \) are proved in terms of the parameters \( \mu \) and \( \nu \) when \( q \) diverges along sequences \( S_{\mu,\nu} \) for which the quotient \( \tau / \log_2 q \) belongs to the interval \( [\mu, \nu] \), with \( 1 \leq \mu \) and \( \nu \) close enough to 1.


1. Introduction and results

Notation. We denote by \( \lfloor x \rfloor \) the integer part of \( x \), by \( \# A \) the cardinality of a set \( A \) and by \( \zeta(x) \) the value of the Riemann zeta function at \( x \). Moreover, several constants appear in this paper: in those inequalities where their numerical value appears explicitly, it is always rounded up or down in such a way to produce a correct statement.

Let \( q \) be an odd integer, let \( \tau \) be the order of 2 modulo \( q \) and let \( \xi \) be a primitive \( q \)th root of unity. In this paper we deal with bounds for the sum

\[
(1) \quad s(\xi) := \sum_{r=1}^{\tau} \xi^{2^r}.
\]

This problem and its generalizations appear in many different contexts and are the subject of an intense research: for example see [2, 3, 4, 5, 8, 11, 12, 18] and the bibliography cited therein. Roughly speaking, upper bounds of type \( s(\xi) \ll \tau^{1-\delta} \) for some positive and explicit constant \( \delta \) have been proved whenever \( \log \tau \gg \log q \), i.e. when the sum contains sufficiently many terms with respect to the order of \( \xi \). A considerably smaller cancellation is expected when the condition \( \log \tau \gg \log q \) is violated. In fact, it can be proved that

\[
\max_{\xi: \xi^q = 1 \atop \xi \text{ primitive}} \{ |s(\xi)| \} \geq 0.3 \tau
\]

when \( q \) diverges along suitable sequences. In his study of the Linnik’s problem about the representability of even integers as a sum of two primes and \( N \)

2010 Mathematics Subject Classification. Primary 11 L 07.
powers of 2, Gallagher (see Lemma 3 of [6]) proved that there exists a positive constant \( c \) such that
\[
|s(\xi)| \leq \tau - c
\]
(see also Thm. 1 in [10] and Eq. (4.2) in [11]). In [6] the values of \( c \) and \( N \) are not explicitly given. More recently, H. Li, J. Liu., M. Liu and T. Wang [13, 14, 15, 16] have proved that \( N \leq 1906 \) (\( N \leq 200 \) under GRH), and an essential ingredient of [13, 14, 15] is an explicit version of the argument of Gallagher, see Lemma 4 in [14]. This lemma implies (2) with any \( c < \sin^2(\pi/8) = 0.146 \ldots \). In this paper we are concerned with the behavior of \( s(\xi) \) when \( q \) diverges along a sequence for which the quotient \( \tau/L \) belongs to the range \([\mu, \nu]\) with \( 1 \leq \mu < \nu \) and \( \nu - 1 \) small enough, where \( L \) denotes the integer part of \( \log_2 q \). The interest for such a type of results comes from the fact that, according to the previous discussion, along these sequences we should get the smallest cancellation for \( s(\xi) \). A simple and typical example is the sequence \( q = 2^n - 1 \) for which \( \tau = L + 1 \). The following examples are less trivial.

**Example 1.** Let \( m \) be an odd integer not of the form \( 2^n - 1 \). Denote by \( \tau_m \) the order of 2 modulo \( m \). Let \( q := (2^{\tau_m} - 1)/m \) and finally let \( \tau_q \) be the order of 2 modulo \( q \). The number \( \tau_q \) divides \( \tau_m \) and is equal to \( \tau_m \) when \( m^2 < 2^{\tau_m} \). Indeed, let \( s := (2^{\tau_q} - 1)/q \). The equality
\[
1 + mq = 2^{\tau_m} = (2^{\tau_q})^{\tau_m/\tau_q} = (1 + sq)^{\tau_m/\tau_q}
\]
shows that if \( \tau_q \neq \tau_m \) we must have \( mq > q^2 \) implying that \( m^2 \geq 2^{\tau_m} \). Since \( m^2 < 2^{\varphi(m)} \) for every \( m > 5 \), the previous criterion shows that \( \tau_q = \tau_m = \varphi(m) \) whenever 2 is a primitive root modulo \( m \) and \( m > 5 \). These facts suggest the following construction: 2 is a primitive root modulo \( 3^k \) for every \( k \), therefore we take \( m = 3^k \) (for \( k > 1 \) and \( q := (2^{\varphi(m)} - 1)/m \). Then \( \tau_q = \tau_m = \varphi(m) = 2m/3 \) and \( q = (2^{\tau_q} - 1)/(3\tau_q/2) \) implying that for these numbers we have
\[
\tau_q = \log_2 q + \log_2 \log_2 q + O(1) = L + \log_2 L + O(1)
\]
as \( k \) diverges.

**Example 2.** For every couple of positive integers \( m, n \) with \( m > 1 \), let \( q := \frac{2^{mn} - 1}{2^n - 1} \). The order \( \tau_q \) of 2 modulo \( q \) is \( mn \) (since the congruence \( 2^{mn} = 1 \mod q \) implies that \( \tau_q \) divides \( mn \) and the congruence \( 2^{\tau_q} = 1 \mod q \) implies that \( (2^{\tau_q} - 1)(2^n - 1) \geq 2^{mn} - 1 \), so that \( \tau_q \) must be greater than \( (m - 1)n \)). Moreover, the inequalities \( 2^{(m-1)n} < q < 2^{(m-1)n+1} \) prove that \( L = (m - 1)n \), hence for such numbers we have
\[
\tau_q = mn = \frac{m}{m-1} L.
\]
Thus, when $n \to \infty$ and $m$ is fixed these numbers define a sequence for which $	au_q \sim \nu \mathcal{L}$ holds with $
u = m/(m - 1)$.

**Example 3.** Let $q, m, n$ and $	au$ be as in the previous example, but this time let $n$ be fixed while $m$ diverges. For such numbers we have

$$
\tau_q = mn = \mathcal{L} + n,
$$

so that along this sequence $	au_q \sim \mathcal{L}$ again, but this time the difference $	au_q - \mathcal{L}$ is constant.

A possible attack to our problem is via the Vinogradov’s method (see [1]). Very roughly speaking, this method provides a set of technics allowing one to obtain upper bounds for exponential sums via the study of the cardinality of sets of representations of integers as sum of numbers taken in a suitably chosen and fixed set. In this sense it is not surprising that one can deduce bounds for (1) from bounds for the number of representations of an integer as sum of powers of two. The following theorem represents an explicit and simple realization of this idea; here $\mathcal{U}(\ell, k)$ denotes the number of representations of $\ell$ as sum of $k$ powers of two (see Section 2):

**Theorem 1.** Let $\varrho$ be a constant such that

$$
\max_\ell \left\{ \mathcal{U}(\ell, k) \right\} \leq (\varrho k)^k \cdot k^{O(1)}
$$

holds for $k$ large enough. Let $\mu, \nu$ be positive constants with $1 \leq \mu < \nu$. Let

$$
\mu,\nu(\varrho, x) := -\frac{\nu - 1/2}{\mu} x^2 \log 2 + x \log(x/\varrho)
$$

and let $c_{\mu,\nu}(\varrho) := \max_{x \geq 1} \left\{ \mu,\nu(\varrho, x) \right\}$, which exists and is positive when $\frac{\nu - 1/2}{\mu} < \lambda := (e\varrho \log 2)^{-1}$. Suppose that $q$ diverges along a sequence $S$ for which $\tau/\mathcal{L} \in [\mu, \nu]$ with $\frac{\nu - 1/2}{\mu} < \lambda$ and where $\mathcal{L} = \lceil \log_2 q \rceil$. Then

$$
\max_{\xi \xi^{\ell} \equiv 1} \{|s(\xi)|\} \leq \tau - c_{\mu,\nu}(\varrho) + o_{\varrho,\mu,\nu}(1).
$$

It is evident that this proposition is useful only if we have an explicit value for $\varrho$ and Section 2 is devoted to the proof of the following fact.

**Theorem 2.** For every $k$, $\max_\ell \left\{ \mathcal{U}(\ell, k) \right\} \leq 2.62 \cdot (0.646661 k)^k \cdot k^{1/2}$.

In order to appreciate this result, we mention here also a second result that we will prove in the same section.

**Theorem 3.** For every $k$, $\max_\ell \left\{ \mathcal{U}(\ell, k) \right\} \gg (0.644591 k)^k$.

Theorem 2 shows that $0.646661$ is an admissible value for $\varrho$ in Theorem 1. This value gives $\lambda = (e\varrho \log 2)^{-1} > 0.8$ so that the bound in (3) applies to all sequences in Examples 1–3. For the sequence with $m = 2$ in Example 2 our theorem predicts the cancellation $c_{2,2}(\varrho) \geq 0.1809$ which is already better than that one previously known; the cancellation predicted by Theorem 1 for
every sequence in Example 2 becomes as better as greater is \( m \) and reaches its best (largest) result \( c_{1,1}(\rho) \) in the limit \( m \to \infty \). Besides, \( c_{1,1}(\rho) \) is also the cancellation which is predicted for each sequence with \( \tau / \mathcal{L} \to 1 \) (for example the sequences in Examples 1 and 3); since \( c_{1,1}(\rho) \geq 1.7465 \), we deduce that for these sequences
\[
\max_{\xi: \xi \equiv 1 \pmod{\xi_{\text{primitive}}}} \{|s(\xi)|\} \leq \tau - 1.7465 + o(1),
\]
which is a sharp improvement on all previously known bounds. This cancellation is the strongest we can recover from Theorem 1.

The argument proving Theorem 1 produces an explicit bound for the little-o term in (4) and in the more general (3) whenever the behavior of \( \tau \) with respect to \( \mathcal{L} \) is explicitly known. For example, using the full strength of Theorem 2 this approach shows that if \( q = 2^n - 1 \) then
\[
\max_{\xi: \xi \equiv 1 \pmod{\xi_{\text{primitive}}}} \{|s(\xi)|\} \leq n - 1.7465 + 33.5 \frac{\log n}{n}.
\]
This bound is non-trivial for every \( n > 87 \). We do not give here the details of its proof, the interested reader will be able to produce himself the necessary computations by following the argument in Section 3.

A final remark about the Linnik’s problem. It is possible that our result and the arguments in [14, 15] produce a bound for \( N \) lower than 1906, but we believe that such improvement will not overcome the best results (\( N \leq 8 \) unconditionally and \( N \leq 7 \) under GRH) that Heath-Brown and Puchta [9] and Pintz and Ruzsa [17] have obtained with different approaches that do not involve bounds for (1).

The paper is organized as follows: in Section 2 we first prove some facts mainly of combinatorial flavor about the representations of integers as sum of 2-powers, then we prove Theorems 2 and 3. The proof of Theorem 1 is given in Section 3.

2. Combinatorial tools

Given two positive integers \( k \) and \( \ell \), we call \( k \)-\emph{representation of} \( \ell \) a string \((n_1, \ldots, n_k)\) of non-negative integers such that \( \sum_{j=1}^{k} 2^{n_j} = \ell \), where strings differing by the order are considered as distinct. Moreover, we denote by \( U(\ell, k) \) the number of \( k \)-representations of \( \ell \):
\[
U(\ell, k) := \#\{(n_1, \ldots, n_k) \in \mathbb{N}^k : \sum_{j=1}^{k} 2^{n_j} = \ell\}.
\]
Let \( \sigma(\ell) \) be the Hamming weight of \( \ell \), i.e. the number of 1’s appearing in the binary representation of \( \ell \), so that \( \ell = \sum_{j=1}^{\sigma(\ell)} 2^{m_j} \), with \( m_1 < m_2 < \cdots < \).
For every fixed $k$, the behavior of $\mathcal{U}(\ell, k)$ in dependence on $\ell$ reveals a very chaotic pattern but it appears more regular when is considered along sequences of integers having the same Hamming weight. This fact suggests the introduction of the quantity

$$\mathcal{W}(\sigma, k) := \max_{\ell: \sigma(\ell) = \sigma} \{ \mathcal{U}(\ell, k) \}.$$ 

A manifestation of the greater regularity of $\mathcal{W}(\sigma, k)$ is the following circumstance: the Hamming weight is sub-additive, meaning that $\sigma(\ell_1 + \ell_2) \leq \sigma(\ell_1) + \sigma(\ell_2)$ for every couple of integers $\ell_1$ and $\ell_2$, so that $\mathcal{W}(\sigma, k) = 0$ when $\sigma$ is greater than $k$. Moreover, nothing is lost by studying $\mathcal{W}(\sigma, k)$ in place of $\mathcal{U}(\ell, k)$ because it is evident that $\max_{\sigma} \{ \mathcal{W}(\sigma, k) \} = \max_{\ell} \{ \mathcal{U}(\ell, k) \}$.

Let us consider the simpler case where also $k$ is a power of two, $\ell = 2^w$ say, so that a $k$-representation of $\ell$ is actually a solution of $2^w = 2^{n_1} + \cdots + 2^{n_k}$. The following proposition shows an important relation between the parameters $w$, $k$ and the set $\{ n_j \}_{j=1}^k$.

**Lemma 1.** Let $(n_1, \ldots, n_k)$ be a $k$-representation of $2^w$. Then $\min \{ n_j \} \geq w - k + 1$.

Note that the inequality is sharp, since $(0, 0, 1, \ldots, k - 2)$ is a $k$-representation of $2^{k-1}$.

**Proof.** The string $(n_1 - \min \{ n_j \}, \ldots, n_k - \min \{ n_j \})$ is a $k$-representation of $2^{w-\min \{ n_j \}}$, therefore, without loss of generality, we can assume that $0 = n_1 \leq n_2 \leq \cdots \leq n_k$: under these assumptions we have to prove that $w \leq k - 1$.

The claim is evident for $k = 1$ and 2, thus we suppose $k \geq 3$. The existence of an upper bound for $w$ becomes clear if we consider $\sum_{j=1}^k 2^{n_j}$ as an addition of binary digits, thus let $\bar{w}$ be this maximal value and let $0 = n_1 \leq n_2 \leq \cdots \leq n_k$ be a sequence producing it. The special sequence $(0, 0, 1, \ldots, k - 2)$ shows that $\bar{w} \geq k - 1$. The congruence $0 = 2^{\bar{w}} - \sum_{j=1}^k 2^{n_j} \equiv \{ j : n_j = 0 \} \pmod{2}$ shows that the number of indexes $j$ with $n_j = 0$ must be even, so that certainly $n_2 = n_1 = 0$. Let $r$ be such that $n_{2r} = 0$ and $n_{2r+1} \geq 1$. If $r > 1$ the sum of the powers associated with the new sequence

$$(0, 0, 1, \ldots, 1, n_{2r+1}, n_{2r+2}, \ldots, n_k, \bar{w}, \bar{w}+1, \bar{w}+2, \ldots, \bar{w}+r-2)$$

is $2^{\bar{w}+r-1}$, contradicting the maximality of the original sequence $n_1, \ldots, n_k$, so that $r = 1$ implying that $n_3 \geq 1$. The case $n_3 > 1$ is impossible, since otherwise we would have both $2^{\bar{w}} = 0 \pmod{2^2}$ (because we know that $\bar{w} \geq k - 1$ and we are assuming $k \geq 3$) and $2^{\bar{w}} = 2^{n_1} + 2^{n_2} = 2 \pmod{2^2}$. Hence $n_3 = 1$, thus proving the claim if $k = 3$. Suppose $k \geq 4$, then the congruence $0 = 2^{\bar{w}} = \sum_{j=1}^k 2^{n_j} \pmod{2^2}$ shows that $\{ j : j \geq 4, n_j = 1 \}$ is even. In particular, if
Let \( n_4 = 1 \) then also \( n_5 = 1 \). Let \( r \) be such that \( n_{2r+1} = 1 \) and \( n_{2r+2} > 1 \). If \( r > 1 \) the sum of powers associated with the new sequence

\[
(0, 0, 1, 2, \ldots, 2, n_{2r+2}, n_{2r+3}, \ldots, n_k, \bar{w}, \bar{w} + 1, \bar{w} + 2, \ldots, \bar{w} + r - 2)
\]

is \( 2^{w+r-1} \), contradicting the maximality of the original sequence \( n_1, \ldots, n_k \), so that \( r = 1 \) implying that \( n_4 \geq 2 \). If \( n_4 > 2 \) we have both \( 2^w = 0 \) (mod \( 2^3 \)) and \( 2^w = 2^{n_1} + 2^{n_2} + 2^{n_3} = 4 \) (mod \( 2^5 \)): the contradiction proves that \( n_4 = 2 \). Iterating the argument we prove that \( n_j = j - 2 \) for every \( j \geq 2 \), so that \( \bar{w} \) is exactly \( k - 1 \).

Adding 1 to each element of a \( k \)-representation of \( 2^w \) we get a \( k \)-representation of \( 2^{w+1} \), thus proving that \( U(2^w, k) \leq U(2^{w+1}, k) \). Vice versa, the lower-bound for \( \min\{n_j\} \) in Lemma 1 implies that we can subtract 1 from each element of every \( k \)-representation of \( 2^{w+1} \) whenever \( w \geq k - 1 \), obtaining in this way a \( k \)-representation of \( 2^w \). In other words, we have obtained that

\[
U(2, k) \leq U(2^2, k) \leq \cdots \leq U(2^{k-1}, k) = U(2^k, k) = U(2^{k+1}, k) = \ldots
\]

proving that the quantity \( W(k) := W(1, k) = U(2^{k-1}, k) \) represents the maximum number of \( k \)-representations that a power of 2 can have.

It is evident that a relation among the general function \( W(\sigma, k) \) and the special function \( W(k) \) must exist, because it is intuitively clear that every \( k \)-representation of an integer \( \ell \) is made of representations of its \( \sigma(\ell) \) nonzero binary digits. This idea is clarified by the next formula (7), that we now prove. We need a second lemma.

**Lemma 2.** Let \( \{m_j\}, \{n_j\} \) be finite sets of integers not necessarily distinct, with \( \sum_j 2^{m_j} = \sum_j 2^{n_j} \) and \( m_1 < m_j \) for every \( j \neq 1 \). Then there is a set \( S \subseteq \{n_j\} \) such that \( \sum_{j \in S} 2^{n_j} = 2^{m_1} \).

**Proof.** Without loss of generality we can suppose that \( n_1 \leq n_2 \leq \ldots \). The number \( n_1 \) cannot be strictly greater than \( m_1 \), otherwise the congruence \( \sum_j 2^{m_j} = \sum_j 2^{n_j} \) (mod \( 2^{m_1+1} \)) is false (here we use the hypothesis \( m_1 < m_j \) for every \( j \neq 1 \)). If \( n_1 = m_1 \) the claim holds with \( S = \{n_1\} \), therefore suppose \( n_1 < m_1 \). Suppose now that \( 2^{n_1} + 2^{n_2} > 2^{m_1} \), then \( 2^{n_1} + 2^{n_2} > 2^{m_1} \geq 2^{n_1+1} \) so that \( n_2 \) is strictly larger than \( n_1 \), but this is impossible because it contradicts the congruence \( \sum_j 2^{m_j} = \sum_j 2^{n_j} \) (mod \( 2^{m_1+1} \)). Hence \( 2^{n_1} + 2^{n_2} \leq 2^{m_1} \). If the equality holds we have the claim with \( S = \{n_1, n_2\} \). Suppose \( 2^{n_1} + 2^{n_2} < 2^{m_1} \). Then \( n_1 = n_2 \) (otherwise the congruence \( \sum_j 2^{m_j} = \sum_j 2^{n_j} \) (mod \( 2^{m_1+1} \)) is false) and \( m_1 \geq n_1 + 2 \). Consider the sum \( 2^{n_1} + 2^{n_2} + 2^{n_3} \). If this sum is greater than \( 2^{m_1} \) we have \( 2^{n_1} + 2^{n_2} + 2^{n_3} > 2^{m_1} \geq 2^{n_1+2} \) giving \( n_3 > n_1 + 1 \) which is impossible because the congruence \( \sum_j 2^{m_j} = \sum_j 2^{n_j} \) (mod \( 2^{m_1+2} \)) would be false, hence \( 2^{n_1} + 2^{n_2} + 2^{n_3} \leq 2^{m_1} \). If the equality holds here we take \( S = \{n_1, n_2, n_3\} \) and
the proof terminates, otherwise we repeat the previous steps. The argument terminates after a finite number of steps, because $\sum_j 2^{m_j} = \sum_j 2^{m_j} \geq 2^{m_1}$. □

Let $\ell$ be an arbitrary positive integer. Iterating Lemma 2, we see that every $k$-representation of $\ell$ can be decomposed as union of representations of its $\sigma(\ell)$ nonzero digits appearing in its binary representation. Note that the orders $k_1, \ldots, k_{\sigma(\ell)}$ of these representations satisfy the restriction $k_1 + \cdots + k_{\sigma(\ell)} = k$, that by definition there are $W(k_1)$ representations of order $k_1$ for the first digit, $W(k_2)$ representations for the second, and so on for every nonzero digit, and that these representations can be permutated in $k!/k_1! \cdots k_{\sigma(\ell)}!$ ways, at most; it follows that

\begin{equation}
(6) \quad U(\ell, k) \leq \sum_{k_1, \ldots, k_{\sigma(\ell)} \geq 1, k_1 + \cdots + k_{\sigma(\ell)} = k} W(k_1) \cdots W(k_{\sigma(\ell)}) \cdot \frac{k!}{k_1! \cdots k_{\sigma(\ell)}!}.
\end{equation}

The strict inequality can hold in (6), because different permutations of the representations of the nonzero digits can give the same $k$-representation of $\ell$: this happens iff there are two nonzero digits in $\ell$ admitting some representation with common integers. By Lemma 1 the representations of the binary digits in $\ell$ do not have common integers whenever the nonzero digits are separated by gaps of length $k - 1$, at least. In other words, if $\sum_j 2^{m_j}$ is the binary representation of $\ell$ and $m_j - m_{j-1} \geq k$ for every $j$ (with $m_0 := 0$), then the representations of the nonzero digits do not overlap and (6) holds as equality. Since for every integer $\sigma$ there exist (infinitely many) integers $\ell$ with $\sigma(\ell) = \sigma$ and whose binary nonzero digits have gaps of length $k - 1$ at least, we conclude that the quantities $W(\sigma, k)$ and $W(k_j)$ are related by the formula

\begin{equation}
(7) \quad W(\sigma, k) = \sum_{k_1, \ldots, k_{\sigma} \geq 1, k_1 + \cdots + k_{\sigma} = k} W(k_1) \cdots W(k_{\sigma}) \cdot \frac{k!}{k_1! \cdots k_{\sigma}!}.
\end{equation}

Denoting by $L_\sigma(x)$ the formal series $\sum_{k=1}^{+\infty} \frac{W(\sigma, k)}{k!} x^k$, the previous identity can be stated simply by saying that $L_\sigma(x) = (L_1(x))^\sigma$. The identity $L_\sigma(x) = L_1(x) L_{\sigma-1}(x)$ immediately gives the formula

\[ W(\sigma, k) = \sum_{n=1}^{k-1} W(n) \cdot W(\sigma - 1, k - n) \cdot \frac{k!}{n!(k-n)!} \]

which is particularly useful in order to compute $W(\sigma, k)$ iteratively from a given set of values for $W(k)$. For example, we have the following table:
Table 1: Values of \( W(\sigma, k) \) for \( k \leq 10 \).

<table>
<thead>
<tr>
<th>( \sigma \setminus k )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>13</td>
<td>75</td>
<td>525</td>
<td>4347</td>
<td>41245</td>
<td>441675</td>
<td>5259885</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>6</td>
<td>30</td>
<td>190</td>
<td>1470</td>
<td>13230</td>
<td>135982</td>
<td>1565694</td>
<td>19959390</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>36</td>
<td>270</td>
<td>2340</td>
<td>23310</td>
<td>260820</td>
<td>3242862</td>
<td>44292420</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td>240</td>
<td>2520</td>
<td>28560</td>
<td>355320</td>
<td>4823280</td>
<td>71147160</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>120</td>
<td>1800</td>
<td>25200</td>
<td>361200</td>
<td>5481000</td>
<td>88565400</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>720</td>
<td>15120</td>
<td>3175200</td>
<td>67737600</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>5040</td>
<td>1451520</td>
<td>39916800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>40320</td>
<td>141120</td>
<td>67737600</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>362880</td>
<td>16329600</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>3628800</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

These values suggest that both \( W(k) \) and \( \max_{\sigma} \{ W(\sigma, k) \} \) grow as \( c^k k! \) for suitable constants \( c \); we have not been able to prove an asymptotic result, nevertheless the next section provides tight upper and lower bounds of that form. A final remark: the value of \( W(k) \) can be computed by hand only for the smallest \( k \), but also a computer can be of little help if the computation is done in the naive way, i.e. by searching all \( k \)-representations of \( 2^k-1 \). In [7] a recursive formula allowing one to compute \( W(k) \) very efficiently is discussed.

Remark. All the numbers \( W(k) \) appearing in Table 1 are odd, a fact which is quite surprising because we do not see any simple or combinatorial explanation for it. Actually, more is true and the congruence \( W(k) = 4 + (-1)^k \mod 8 \) for \( k \geq 3 \) is proved in [7].

2.1. The lower bound: proof of Theorem 3. For every \( k \geq 2 \), the \( k!/2 \) permutations of the string \((0, 0, 1, \ldots, k-2)\) are \( k \)-representations of \( 2^{k-1} \), so that \( W(k) \geq k!/2 \) for every \( k \). This simple lower bound immediately produces a lower bound for \( \mathcal{W}(\sigma, k) \) of the type considered in Theorem 3. In fact, introducing it in (7) we have

\[
\frac{W(\sigma, k)}{k!} \geq \sum_{k_1, \ldots, k_\sigma \geq 1 \atop k_1 + \cdots + k_\sigma = k} \frac{1}{2^\sigma} = \frac{1}{2^\sigma} \binom{k-1}{\sigma-1},
\]

where we have used the combinatorial identity \( \sum_{k'_1 + \cdots + k'_\ell = c} \binom{c}{\ell} = \binom{b+c-1}{b-1} \). This result and the simple inequality \( \max_{\sigma} \{ W(\sigma, k) \} \geq \frac{1}{k} \sum_{\sigma=1}^k W(\sigma, k) \) gives the bound \( \max_{\sigma} \{ W(\sigma, k) \} \geq (3/2)^{k-1}(k-1)! \) that by Stirling becomes

\[
\text{(8)} \quad \max_{\sigma} \{ W(\sigma, k) \} \gg (0.5518k)^k.
\]
We consider this lower bound as the trivial one: aim of the present section is to improve it up to the result given in Theorem 3. The corollary following the next lemma improves the lower bound for $W(k)$.

**Lemma 3.** for every $k$ we have $W(k) \geq \sum_{j}^{*} \binom{k}{2j+1} W(k-2j+1)$, where the sum $\sum_{j}^{*}$ is restricted to the positive integers $j$ with $2^j < k$.

**Proof.** We fix a positive integer $j$ with $2^j < k$. Let $(n_1, \ldots, n_{k-2^j+1})$ be a $(k-2^j+1)$-representation of $2^{k-j}$. We notice that the number of these $(k-2^j+1)$-representations is $W(k-2^j+1)$ (by (5)) and that each $n_i$ is strictly lower than $k-j$. The $\binom{k}{2j-1}$ strings that we obtain by joining $2^j-1$ times the number $k-j$ in all possible positions to the string $(n_1, \ldots, n_{k-2^j+1})$ are $k$-representations of $2^k$. Since every $n_i$ is strictly lower than $k-j$, each representation that we generate in this way is completely characterized by the position where the numbers $k-j$ appear. In particular, they are distinct. Let $K_j$ denote the set of representations of $2^k$ that we obtain using the previous construction: we have just proved that $\sharp K_j = \binom{k}{2j-1} W(k-2^j+1)$. Every representation in $K_j$ contains the exponent $k-j$ and no exponent of greater value, therefore the representations in $K_j$ and $K_{j'}$ are distinct when $j \neq j'$, and the claim is proved. \hfill $\square$

**Corollary.** For every $\varepsilon > 0$ we have $W(k) \gg \varepsilon (\eta - \varepsilon)^k k!$, where $\eta$ is the solution of $\sum_{j=1}^{\infty} \frac{\eta^{1-2^j}}{(2^j-1)!} = 1$. In particular, $W(k) \geq 0.3316 \cdot (1.1305)^k k! \quad \forall k$.

**Proof.** Let $F_\infty(x) := \sum_{j=1}^{+\infty} \frac{x^{1-2^j}}{(2^j-1)!}$ and $F_n(x) := \sum_{j=1}^{n} \frac{x^{1-2^j}}{(2^j-1)!}$ for every $n > 1$. Functions $F_\infty$, $F_n$ decrease in $\mathbb{R}^+$ with $F_\infty(x) > F_n(x)$, $F_n(1) > 1$ and $F_\infty(2) < \sqrt{e} - 1 < 1$. Hence there exist a unique solution $\eta$ of $F_\infty(x) = 1$ and a unique solution $\eta_n$ of $F_n(x) = 1$ for every $n$, with $\eta, \eta_n \in (1, 2)$ and $\eta > \eta_n$. Moreover, $|F'_\infty(x)| = \sum_{j=1}^{+\infty} x^{-2^j}/(2^j-2)! > 1/4$ for $x \in (1, 2)$. This lower-bound and the equality $|F_\infty(\eta) - 1| = |F'_\infty(\gamma)||\eta - \eta|$ for a suitable $\gamma \in (\eta_n, \eta) \subset (1, 2)$ imply that

$$|\eta - \eta_n| \leq 4|F_\infty(\eta_n) - 1| = 4|F_\infty(\eta_n) - F_n(\eta_n)| = 4 \sum_{j=n+1}^{+\infty} \frac{\eta_n^{1-2^j}}{(2^j-1)!} \\ \\
\leq 4 \sum_{m=2^n}^{+\infty} \frac{\eta_n^{m-n}}{m!} \leq 4 \frac{\eta_n^{2^n} e^{1/\eta_n}}{(2^n)!} \leq \frac{4e}{(2^n)!}$$

thus proving that $\eta_n$ tends to $\eta$. Let $\varepsilon > 0$ be arbitrarily fixed, let $n = n(\varepsilon)$ be an integer such that $\eta_n \geq \eta - \varepsilon$ and let $\alpha > 0$ be chosen in such a way that
\( W(k) \geq \alpha \eta_n^k \cdot k! \) holds for \( k \leq 2^n \). By Lemma 3 we know that
\[
W(k) \geq \sum_{j=1}^{n} \left( \begin{array}{c} k \\ 2j-1 \end{array} \right) W(k-2^j+1)
\]
whenever \( k > 2^n \). By induction on \( k \), in order to prove that \( W(k) \geq \alpha \eta_n^k \cdot k! \) for every value of \( k \) it is sufficient that
\[
\sum_{j=1}^{n} \left( \begin{array}{c} k \\ 2j-1 \end{array} \right) \alpha \eta_n^{k-2^j+1} \cdot (k-2^j+1)! \geq \alpha \eta_n^k \cdot k!.
\]
This inequality can be written as \( F_n(\eta_n) \geq 1 \) and is evidently satisfied by the definition of \( \eta_n \), thus the first claim is proved. The second claim follows using this argument with \( n = 3 \) and the known values of \( W(k) \) for \( k \leq 8 \). \( \square \)

**Remark.** Using the bound \( |\eta_n - \eta| \leq 4e/(2^n)! \) it is possible to compute \( \eta \) with arbitrarily large precision: \( \eta = 1.13055033 \ldots \).

For its frequent use in the following part of this section it is convenient to introduce the symbols \( \alpha \) and \( \beta \) to denote the constants 0.3316 and 1.1305, respectively; with this notation, the previous corollary says that \( W(k) \geq \alpha \beta^k k! \).

This bound improves considerably the bound \( W(k) \geq k!/2 \) for large values of \( k \), nevertheless it badly underestimates \( W(k) \) for small values of \( k \). Since also these terms affect the final result, in order to recover a lower bound for \( W(\sigma,k) \) from (7) we split the range for \( k \) in two sets, \( k \leq \tilde{k} \) and \( k > \tilde{k} \), where \( \tilde{k} \) is a parameter \( \geq 3 \) that we will choose later, and we use the true value of \( W(k) \) in the first set and the bound \( W(k) \geq \alpha \beta^k k! \) in the second one. Decomposing (7) according to the number of variables whose index is \( \leq \tilde{k} \), we obtain
\[
\frac{W(\sigma,k)}{k!} = \sum_{h=0}^{\sigma} \sum_{k_1,\ldots,k_h \geq 1 \atop k_1 + \cdots + k_h = k} \prod_{j=1}^{\sigma} \frac{W(k_j)}{k_j!} = \sum_{h=0}^{\sigma} \left( \begin{array}{c} \sigma \\ h \end{array} \right) \sum_{1 \leq k_1,\ldots,k_h \leq k \atop k_1 + \cdots + k_h = h} \prod_{j=1}^{\sigma} \frac{W(k_j)}{k_j!}.
\]

We set \( w_j := W(j)/(j!\alpha \beta^j) \) for \( j \leq \tilde{k} \) so that we can bound \( W(k_j)/k_j! \) by \( w_j \alpha \beta^{k_j} \) when \( k_j \leq \tilde{k} \) and by \( \alpha \beta^{k_j} \) for \( k_j > \tilde{k} \), obtaining
\[
\frac{W(\sigma,k)}{k!} \geq \alpha^\sigma \beta^k \sum_{h=0}^{\sigma} \left( \begin{array}{c} \sigma \\ h \end{array} \right) \sum_{k_1,\ldots,k_h \geq 1 \atop k_1 + \cdots + k_h = k} \sum_{k_{h+1},\ldots,k_w \geq \tilde{k} \atop k_{h+1} + \cdots + k_w = k-w} w_{k_1} \cdots w_{k_h}
\]
\[
= \alpha^\sigma \beta^k \sum_{h=0}^{\sigma} \left( \begin{array}{c} \sigma \\ h \end{array} \right) \sum_{w=0}^{\tilde{k}} \left[ \sum_{k_{h+1},\ldots,k_w \geq \tilde{k} \atop k_{h+1} + \cdots + k_w = k-w} 1 \right] \left[ \sum_{1 \leq k_1,\ldots,k_h \leq \tilde{k} \atop k_1 + \cdots + k_h = k-w} w_{k_1} \cdots w_{k_h} \right].
\]
The third sum is evaluated by using the identity $\sum_{k_1,\ldots,k_\ell \geq 0} 1 = \binom{b+c-1}{b-1}$, while the last sum admits an alternative representation: for every $i \in 1, \ldots, \bar{k}$ let $a_i := \sharp\{j : k_j = i\}$, then

$$\sum_{1 \leq k_1, \ldots, k_h \leq \bar{k}} w_{k_1} \cdots w_{k_h} = \sum_{a_1 + \cdots + a_h = h, a_1 + 2a_2 + 3a_3 + \cdots + ka_\bar{k} = w} \frac{h!w_1^{a_1} \cdots w_k^{a_k}}{a_1!a_2! \cdots a_k!}$$

$$= \sum_{a_1, \ldots, a_{\bar{k}-2} \geq 0} h!w_1^{a_1} \cdots w_{\bar{k}-2}^{a_{\bar{k}-2}} w_{k-1}^{A} w_k^{B}$$

where $A := \bar{k}h - w - \sum_{i=1}^{\bar{k}-2}(\bar{k} - i)a_i$, $B := w - (\bar{k} - 1)h + \sum_{i=1}^{\bar{k}-2}(\bar{k} - i - 1)a_i$ and the symbol $\sum^*$ means that the sum is restricted to those $a_1, \ldots, a_{\bar{k}-2}$ such that $A, B \geq 0$. In this way we get the following lower bound for $W(\sigma, k)$:

$$W(\sigma, k) \geq \frac{\alpha \beta^k}{k!} \sum_{h=0}^{\sigma} \sum_{w=h}^{\bar{k}h} \sum^* \left( \frac{\sigma}{h} \binom{k - w - \bar{k}(\sigma - h) - 1}{\sigma - h - 1} \right) \frac{h!w_1^{a_1} \cdots w_{\bar{k}-2}^{a_{\bar{k}-2}} w_k^{A}}{a_1!a_2! \cdots a_{\bar{k}-2}!A!B!}$$

The previous multiple sum is quite intricate; we bound it from below simply with one of its terms, i.e.

$$W(\sigma, k) \geq \frac{\alpha \beta^k}{k!} \left( \frac{\sigma}{h} \binom{k - w - \bar{k}(\sigma - h) - 1}{\sigma - h - 1} \right) \frac{h!w_1^{a_1} \cdots w_{\bar{k}-2}^{a_{\bar{k}-2}} w_k^{A}}{a_1!a_2! \cdots a_{\bar{k}-2}!A!B!}$$

where $\sigma, h, w$ and $a_i$ for every $i$ can be arbitrary chosen but must be taken in such a way that the constraints $h \in (0, \sigma)$, $w \in (h, \bar{k}h)$, $a_i \geq 0$ and $A, B \geq 0$ be satisfied. Our aim is now to determine a convenient set of values for these parameters in such a way to pick up a value as large as possible for the R.H.S. of (9). We simplify a little bit the discussion by setting $\sigma = \lfloor uk \rfloor$, $h = \lfloor vk \rfloor$, $w = \lfloor zk \rfloor$ and $a_i = \lfloor s_i k \rfloor$ for every $i$, with $u, v, z$ and $s_i$ as new parameters, independent of $k$: in this way the dependence on $k$ appears only to the exponent and by Stirling we deduce that

$$W(\sigma, k) \geq \frac{\alpha^u \beta^v}{k!k^{O(1)}} \left[ u^w (1 - z - \bar{k}(u - v))^{1 - z - \bar{k}(u - v)} \right]^{u-v}(1 - z - (\bar{k} + 1)(u - v))^{1 - z - (\bar{k} + 1)(u - v)}$$

$$\cdot \frac{w_1^{a_1} \cdots w_{\bar{k}-2}^{a_{\bar{k}-2}} w_k^{A}}{s_1^{a_1} \cdots s_{\bar{k}-2}^{a_{\bar{k}-2}} A!B!}$$
where
\[
\mathcal{A} := \bar{k}v - z - \sum_{i=1}^{\bar{k}-2} (\bar{k} - i) s_i, \quad \mathcal{B} := z - (\bar{k} - 1)v + \sum_{i=1}^{\bar{k}-2} (\bar{k} - i - 1) s_i.
\]

The stationary points of the function of \(u, v, z, \{s_i\}_{i=1}^{\bar{k}-2}\) to the R.H.S. of (10) are solutions of the system
\[
\begin{align*}
(1 - z - (\bar{k} + 1)(u - v))^k (u - v)^2 &= \alpha u (1 - z - (\bar{k} + 1)(u - v))^{k+1} \\
(1 - z - (\bar{k} + 1)(u - v))^{k+1} (\mathcal{A}/w_{k-1})^k &= (1 - z - (\bar{k} + 1)(u - v)) (\mathcal{B}/w_{k-1})^k, \\
(1 - z - (\bar{k} + 1)(u - v)) (\mathcal{B}/w_{k-1}) = (1 - z - (\bar{k} + 1)(u - v)) (\mathcal{B}/w_k), \\
w_i (\mathcal{A}/w_{k-1})^{k-i} &= s_i (\mathcal{B}/w_k)^{k-1-i} \quad \forall i = 1, \ldots, \bar{k} - 2
\end{align*}
\]
and can be explicitly found. Let \(x, y, u, v\) and \(r_i\) for every \(i \leq \bar{k} - 2\) be a new set of variables related to the previous ones by: \(x = u - v, u = u' x, z = 1 - k x - y x\) and \(s_i = x r_i\), and let
\[
\begin{align*}
A := \mathcal{A}/x &= -1/x + \bar{k} u' + y - \sum_{i=1}^{k-2} (\bar{k} - i) r_i, \\
B := \mathcal{B}/x &= 1/x - (\bar{k} - 1) u' - y - 1 + \sum_{i=1}^{k-2} (\bar{k} - i - 1) r_i.
\end{align*}
\]
After simple computations the system yields
\[
\begin{align*}
u' &= \frac{y^k}{\alpha (y-1)^{k+1}}, \\
r_i &= \frac{w_i y^{k-1}}{(y-1)^{k+1}}, \quad \forall i = 1, \ldots, \bar{k} - 2, \quad \begin{cases} A = \frac{w_{k-1} y}{(y-1)^{2}} \\ B = \frac{w_k}{(y-1)}. \end{cases}
\end{align*}
\]
These relations and the equality \(A + B = u' - 1 - \sum_{i=1}^{k-2} r_i\), give a closed equation for \(y\):
\[
(y - 1)^{k+1} - \frac{1}{\alpha} y^k + \sum_{i=1}^{\bar{k}} w_i y^{k-i} (y - 1)^i = 0
\]
that admits a unique solution \(y > 1\). With this solution we can compute \(u', A, B\) and each \(r_i\) by (11) and \(x\) by the identity \(1/x = -A + \bar{k} u' + y - \sum_{i=1}^{\bar{k}-2} (\bar{k} - i) r_i\), hence also \(u, v, z\) and each \(s_i\) are determined. At last, we obtain from (10) a bound of the type \(\mathcal{W}(\sigma, k) \geq (c(\bar{k}) k) k \cdot k^{O(1)}\). We have computed \(c(\bar{k})\) with \(\bar{k} = 3, \ldots, 1500\); apparently \(c(\bar{k})\) steady grows with \(\bar{k}\), with \(c(3) = 0.641134\) and \(c(1500) = 0.644591\). The constant \(c(1500)\) yields the claim in Theorem 3.

**Remark.** The use of the exact value of \(\mathcal{W}(k)\) for small \(k\) is fundamental: if we use the inequality \(\mathcal{W}(k)/k! \geq \alpha \beta^k\) for every \(k\), our argument becomes much simpler but produces only the lower bound \(\max_{\sigma} \{\mathcal{W}(\sigma, k)\} \gg (0.5537 k)^k\) which is a very modest improvement on the trivial bound (8).
The upper bound: proof of Theorem 2. Table 1 suggests the validity of some inequalities among the values of \( W(\sigma, k) \); one of these says that 
\[ 2W(k) \leq W(2, k) \]
and another one that 
\[ 24(W(k) - kW(k - 1)) \leq W(4, k) \]
for every \( k \geq 4 \). Both inequalities are true and can be proved with similar arguments. Moreover, both can be used to prove upper-bounds for \( W(k) \), but the result we obtain from the second inequality is stronger, so we prove here only the second one.

**Lemma 4.** The inequality 
\[ 24(W(k) - kW(k - 1)) \leq W(4, k) \]
holds for every \( k \geq 4 \).

**Proof.** We need the following general fact which is a variation of Lemma 2: let \( S \) be a finite set of nonnegative integers, suppose that \( \sum n \in S 2^n \) is a 2-power, \( 2^q \) say, and that \( S \) contains two integers at least, then there exists \( S' \subset S \) such that \( \sum n \in S' 2^n = 2^{q-1} \). In fact, let \( S_0 \) be any proper and non-empty subset of \( S \). Exchanging \( S_0 \) with \( S_0^c \) if necessary, we can assume that \( \sum n \in S_0 2^n \geq 2^{q-1} \). If the equality holds here we have done, thus we assume that \( \sum n \in S_0 2^n > 2^{q-1} \). Let \( n' \) be the smallest integer in \( S_0 \). The set \( S_0 \) does not coincide with \( \{n'\} \), otherwise \( n' \) is equal to \( q \) (because \( 2^n < \sum n \in S 2^n = 2^q \) and \( 2^{q-1} < \sum n \in S_0 2^n = 2^{q'} \)) implying that \( S_0^c = \emptyset \) (because \( \sum n \in S_0 2^n = 2^{q'} = 2^q = \sum n \in S 2^n \)), against our assumption. Let \( S_1 := S_0 \setminus \{n'\} \); we have just proved that \( S_1 \) is not empty. If \( \sum n \in S_1 2^n < 2^{q-1} \) we have

\[ 2^{q-1} < \sum n \in S_0 2^n = 2^{q'} + \sum n \in S_1 2^n < 2^{q'} + 2^{q-1} \]

implying that \( 2^{q-1} < \sum n \in S_0 2^{n-n'} < 1 + 2^{q-1} \) which is evidently impossible, hence \( \sum n \in S_1 2^n \geq 2^{q-1} \). If the equality holds here the claim is proved, otherwise we repeat the argument with \( S_1 \) in place of \( S_0 \). The argument terminates after a finite number of steps because the definition of \( S_1 \) implies that \( \sum n \in S_1 2^n < \sum n \in S_0 2^n \).

Let \( (n_1, \ldots, n_k) \) be a \( k \)-representation of \( 2^{k-1} \). The previous remark shows that there exists a subset \( R \subset \{1, \ldots, k\} \) such that \( \sum j \in R 2^{n_j} = 2^{k-2} = \sum j \in R^c 2^{n_j} \). Let us assume that both \( R \) and \( R^c \) contain two integers, at least. Then we can iterate the decomposition of \( R \) as union of \( R_0, R_1 \), and of \( R^c \) as union of \( R_2, R_3 \), say, such that \( \sum j \in R_i 2^{n_j} = 2^{k-3} \) for \( i = 0, 1, 2, 3 \). Let \( \pi \) be a permutation of \( \{0, 1, 2, 3\} \) and consider the new string

\[ (n_1^{\pi_1}, \ldots, n_k^{\pi_k}), \text{ with } n_j^{\pi} := n_j + \pi(i) \text{ if } j \in R_i. \]

The sum \( s_k := \sum_{j=1}^k 2^{n_j^{\pi}} = 2^{4k-3} + 2^{3k-3} + 2^{2k-3} + 2^{k-3} \) is independent of \( \pi \) and each string \( (n_1^{\pi}, \ldots, n_k^{\pi}) \) is a \( k \)-representation of \( s_k \). These \( k \)-representations are distinct. In fact, let \( (n_1, \ldots, n_k) \) and \( (m_1, \ldots, m_k) \) be two \( k \)-representations of \( 2^{k-1} \) and let \( \pi, \pi' \) be two permutations. For \( i = 0, 1, 2, 3 \) let \( A_i := \{ j : n_j^{\pi} \in \]

CANCELLATION IN A SHORT EXPONENTIAL SUM 13
Proof. We decompose the sum defining \([ik, (i + 1)k]\) and \(A'_i := \{ j : m_j'' \in [ik, (i + 1)k) \} \). Suppose \((n^*_1, \ldots, n^*_k) = (m^*_1, \ldots, m^*_k)\), then \(A_i = A'_i\) for every \(i\). It follows that for every \(j\) we have \(n_j = n_j'' - ik = m_j'' - ik = m_j\) (where \(i = i(j)\) denotes the index \(i\) such that \(j \in A_i \equiv A'_i\)), i.e. the original \(k\)-representations \((n_1, \ldots, n_k)\) and \((m_1, \ldots, m_k)\) are equal. Under this hypothesis, the proof of the equality \(\pi = \pi'\) is immediate.

Let \(B(k)\) be the number of \(k\)-representations of \(2^{k-1}\) we have considered here, i.e. for which both \(R\) and \(R^{e}\) contain two numbers at least. The argument we have just discussed proves that \(s_k\) admits at least \(24B(k)\) different \(k\)-representations. The Hamming weight of \(s_k\) is four, hence we have proved that \(24B(k) \leq W(4, k)\). In order to terminate the proof we verify now that \(B(k) = W(k) - kW(k - 1)\). The number \(W(k) - B(k)\) counts the \(k\)-representations of \(2^{k-1}\) containing \(k - 2\). When \(k \geq 3\) the number \(k - 2\) appears in these representations only once, therefore these representations are exactly those ones we obtain adding \(k - 2\) to any \((k - 1)\)-representation of \(2^{k-2}\). The claim follows because there are \(k\) possible places for \(k - 2\) in any \((k - 1)\)-representation and \(W(k - 1)\) representations of \(2^{k-2}\) as sum of \((k - 1)\) powers of 2 (by (5)). \(\Box\)

For the proof of Theorem 2 we need to study the sequence

\[
\chi_\sigma(b, k) := \sum_{k_1, \ldots, k_\sigma \geq 1 \atop k_1 + \cdots + k_\sigma = k} \left( \frac{k}{k_1 \cdots k_\sigma} \right)^b
\]

and the constant \(\chi_\sigma(b) := \sup_k \{ \chi_\sigma(b, k) \} \). When \(b > 1\) the sequence converges to \(\sigma \zeta(b)^{\sigma - 1}\) as \(k\) diverges, therefore \(\chi_\sigma(b)\) is certainly finite. For our application we need an accurate determination of the value of \(\chi_\sigma(b)\), thus the next lemma not only proves the convergence but also provides an explicit inequality.

**Lemma 5.** Let \(b > 1\) and let \(c = c(b)\) be the positive constant such that \((1 - (\sigma - 1)c)^{-b} = 1 + 2(\sigma - 1)bc\). Then, for \(\sigma > 1\) we have

\[
\left| \frac{\chi_\sigma(b, k)}{\sigma} - \zeta(b)^{\sigma - 1} \right| \leq \frac{(\sigma - 1)\zeta(b)^{\sigma - 2}}{(b - 1)(ck)^{b - 1}} + \frac{2b\chi_{\sigma - 1}(b)}{k} \left[ \int_{\sigma - 1}^{(\sigma - 1)c/k} w^{1-b} dw + \frac{1}{(\sigma - 1)^{b - 1}} \right]
\]

\[
+ \frac{c^{-b} \zeta(b)^{\sigma - 1}}{\sigma} \sum_{m=2}^{\sigma} \left( \frac{\sigma}{m} \right) \frac{\zeta(b)^{-1}}{(b - 1)(ck)^{(b - 1)}} m^{-1}.
\]

**Proof.** We decompose the sum defining \(\chi_\sigma(b, k)\) according to the number \(m\) of variables which are “large”, where “large” here means greater than \(ck\). The
symmetry of the sum allows us to write this decomposition as

$$\chi_\sigma(b, k) = \sum_{m=0}^{\sigma} \binom{\sigma}{m} S_m \quad \text{with} \quad S_m := \sum_{\substack{k_1, \ldots, k_\sigma \geq 0 \atop \sum k_m = b}} \left( \frac{k}{k_1 \cdots k_\sigma} \right)^b.$$  

The term $S_0$ is empty because a simple argument proves that the constant $c$ is lower than $1/\sigma$; the following argument will show that the main term comes from $S_1$ and that the other terms contribute only at lower orders. The term $S_1$ is

$$S_1 = \sum_{1 \leq k_2, \ldots, k_\sigma \leq c k \atop k_1 > c k} \frac{1}{k_2 \cdots k_\sigma} \left[ \left( 1 - \frac{k_2 + \cdots + k_\sigma}{k} \right)^{-b} - 1 \right].$$

The parameters $k_2, \ldots, k_\sigma$ appearing in this sum are small with respect to $k$, while $k_1 = k - (k_2 + \cdots + k_\sigma)$ is large, so we write $S_1$ as:

$$S_1 = \sum_{k_2, \ldots, k_\sigma = k}^{c k} \frac{1}{k_2 \cdots k_\sigma} + \sum_{k_2, \ldots, k_\sigma = 1}^{c k} \frac{1}{k_2 \cdots k_\sigma} \left[ \left( 1 - \frac{k_2 + \cdots + k_\sigma}{k} \right)^{-b} - 1 \right].$$

The first sum is the $\sigma - 1$ power of the sum $\sum_{w=1}^{c k} w^{-b}$ converging to $\zeta(b)$. Using the upper bound $\sum_{w > c k} w^{-b} \leq \int_{c k}^{+\infty} w^{-b} dw$ we get

$$\left| \sum_{k_2, \ldots, k_\sigma = 1}^{c k} \frac{1}{k_2 \cdots k_\sigma} - \zeta(b)^{-1} \right| = \left| \left[ \sum_{w=1}^{c k} \frac{1}{w^b} \right]^{\sigma-1} - \zeta(b)^{\sigma-1} \right| \leq \frac{(\sigma - 1)\zeta(b)^{-2 - (b - 1)}(c k)^{-1}}{(b - 1)(c k)^{-1}}.$$  

In (13) every $k_j$ with $j \geq 2$ is lower than $c k$, hence $k_2 + \cdots + k_\sigma \leq (\sigma - 1)c$ so that by convexity we have $\left( 1 - \frac{k_2 + \cdots + k_\sigma}{k} \right)^{-b} - 1 \leq 2b k_2 + \cdots + k_\sigma$. We deduce that

$$\sum_{k_2, \ldots, k_\sigma = 1}^{c k} \frac{1}{k_2 \cdots k_\sigma} \left[ \left( 1 - \frac{k_2 + \cdots + k_\sigma}{k} \right)^{-b} - 1 \right] \leq \frac{2b}{k} \sum_{k_2, \ldots, k_\sigma = 1}^{c k} \frac{k_2 + \cdots + k_\sigma}{k_2 \cdots k_\sigma}.$$  

The R.H.S. here tends to 0 for every $b > 1$, but in different ways for $b \in (1, 2)$, $b = 2$ and $b > 2$. We bound it simply via the integral test $\sum_{w=\sigma-1}^{(\sigma-1)c k} \frac{1}{w^{b-1}} \leq \frac{1}{(\sigma - 1)^{b-1}} + \int_{\sigma-1}^{(\sigma-1)c k} w^{-b} dw$. The terms $S_m$ with $m \geq 2$ can be bounded by

$$\sum_{m=2}^{\sigma} \binom{\sigma}{m} S_m.$$
using $k_1$ to control the growth of the numerator and by splitting the sum over the large variables $(k_2, \ldots, k_m)$ and that one over the small variables $(k_{m+1}, \ldots, k_{\sigma})$:

$$S_m = \sum_{k_1, k_2, \ldots, k_m > ck} \left[ \frac{k}{k_1 \cdots k_{\sigma}} \right]^b \leq \frac{1}{c^b} \sum_{k_2, \ldots, k_m > ck} \left[ \frac{1}{k_2 \cdots k_{\sigma}} \right]^b$$

(16) $$= \frac{1}{c^b} \left[ \sum_{w > ck} \frac{1}{w} \right] \left[ \sum_{w \leq ck} \frac{1}{w} \right] \leq \frac{c^{-b} \zeta(b)^{\sigma-m}}{(b-1)^m(ck)^{(m-1)(b-1)}}.$$ The lemma follows by collecting the results in (12)–(16). 

**Remark.** Lemma 5 defines $c$ as the number such that $(1 - (\sigma - 1)c)^{-b} = 1 + 2b(\sigma - 1)c$; this choice is evidently arbitrary and $c$ could be defined by $(1 - (\sigma - 1)c)^{-b} = 1 + rb(\sigma - 1)c$ for any $r \geq 2$. Actually, the new parameter $r$ can be used to improve the final bound; for sake of simplicity we have not mentioned this fact in Lemma 5.

The importance of the sequence $\chi_{\alpha}(b, k)$ for our problem comes from the following result.

**Lemma 6.** Let $b > 1$, $k_0 \geq 4$ and $\gamma(b) := \max\{y_0(b), y(b)\}$, where $y_0(b)$ is the greatest value of $(\mathcal{W}(k)/k!)^{1/(k-1)}$ for $k < k_0$, and $y(b)$ is the positive solution of

$$y^3 - \frac{k^b}{(k_0 - 1)^b} y^2 - \frac{X_0(b)}{24} = 0.$$ Then $\mathcal{W}(k) \leq \gamma(b)^{k-1}k! \cdot k^{-b}$ for every $k$.

**Proof.** The definition of $\gamma(b)$ immediately implies the claim for $k < k_0$. Suppose $k \geq k_0$. By (7) and Lemma 4 we have

$$\mathcal{W}(k) \leq k\mathcal{W}(k-1) + \frac{1}{24} \mathcal{W}(4, k) = k\mathcal{W}(k-1) + \frac{k!}{24} \sum_{k_1, k_2, k_3, k_4 \geq 1} \frac{\mathcal{W}(k_j)}{k_j!}$$

that by induction gives

$$\frac{\mathcal{W}(k)}{k!} \leq \frac{\gamma(b)^{k-2}}{(k-1)^b} + \frac{\gamma(b)^{k-4}}{24} \sum_{k_1, k_2, k_3, k_4 \geq 1} \frac{1}{(k_1 k_2 k_3 k_4)^b}.$$ The claim is proved if $\gamma(b)$ satisfies

$$\frac{\gamma(b)^{k-2}}{(k-1)^b} + \frac{\gamma(b)^{k-4}}{24} \sum_{k_1, k_2, k_3, k_4 \geq 1} \frac{1}{(k_1 k_2 k_3 k_4)^k} \leq \gamma(b)^{k-1}k^{-b}.$$
Since we are assuming \( k \geq k_0 \), simple computations prove that this inequality holds whenever \( \gamma(b) \geq y(b) \).

We show now how Lemmas 5-6 and some numerical computations allow us to find a convenient upper bound for the growth of \( \mathcal{W}(k) \). We have at our disposal the values of \( \mathcal{W}(k) \) for \( k \leq 1500 \); the value of \( y_0(b) \) is only marginally influenced by the choice of \( k_0 \) while \( y(b) \) decreases with \( k_0 \), hence in all our computations we set \( k_0 = 1501 \). For a given \( b > 1 \) we compute \( y_0(b) \) and \( \chi_4(b, k) \) for small values of \( k \) (for \( k \leq 1000 \), say). In this range we identify the element \( \chi_4(b, \tilde{k}) \) having the greatest value. In our numerical computations \( \chi_4(b, k) \) decreases for \( k > \tilde{k} \) and \( \chi_4(b, \tilde{k}) \) is greater than \( 4 \zeta(b)^3 \); with these evidences it is natural to guess the equality \( \chi_4(b) = \chi_4(b, \tilde{k}) \). We can prove the correctness of this guess in two steps: first by using Lemma 5 to compute an index \( \tilde{k} \) such that \( |\chi_4(b, k) - 4 \zeta(b)^3| \leq |\chi_4(b, \tilde{k}) - 4 \zeta(b)^3| \) when \( k > \tilde{k} \), then by verifying that \( \chi_4(b, k) \leq \chi_4(b, \tilde{k}) \) for every \( k < \tilde{k} \) with a direct numerical computation. We notice that Lemma 5 produces \( \tilde{k} \) only if we already know \( \chi_4(b) \) hence a descent process is triggered here: to compute \( \chi_4(b) \) we use Lemma 5 needing \( \chi_2(b) \), and to compute \( \chi_2(b) \) we employ again Lemma 5. At this point the process terminates because \( \chi_1(b) \) is equal to 1 for every \( b \). Having determined \( \chi_4(b) \), we can compute \( y(b) \). The parameter \( y_0(b) \) grows with \( b \) while the numerical computations show that \( y(b) \) decreases with \( b \), therefore we can repeat this process several times adjusting the value for \( b \) until the difference between \( y_0(b) \) and \( y(b) \) is small enough. With \( k_0 = 1501 \) and \( b = 1.6056 \) this algorithm produces the following values: \( \chi_4(b) = \chi_4(b, 71) \leq 49.95 \) (and \( \chi_2(b) = \chi_2(b, 25) \leq 4.72, \chi_3(b) = \chi_3(b, 46) \leq 16.33 \)), \( y_0(b) \leq 1.71186, y(b) \leq 1.71154 \) proving that \( \gamma(b) \leq 1.71186 \), i.e., that \( \mathcal{W}(k) \ll (0.62976 k)^k \).

It is interesting to remark that the upper bound for the growth of \( \mathcal{W}(k) \) we found here is strictly lower than the lower bound we found in Theorem 2 for \( \max_\sigma \{ \mathcal{W}(\sigma, k) \} : 0.62976 \) instead of 0.64459. This is not surprising because \( \max_\sigma \{ \mathcal{W}(\sigma, k) \} \) is evidently larger than \( \mathcal{W}(1, k) = \mathcal{W}(k) \); nevertheless, being able to prove it here means that our argument producing upper and lower bounds for \( \mathcal{W}(k) \) and \( \max_\sigma \{ \mathcal{W}(\sigma, k) \} \) is sufficiently precise. Moreover it has an interest in itself: it proves that not only \( \max_\sigma \{ \mathcal{W}(\sigma, k) \} \) is greater than \( \mathcal{W}(k) \) but also that it grows exponentially faster than \( \mathcal{W}(k) \).

Now we describe the strategy for the proof of Theorem 2. For every \( b > 1 \), from (7) and Lemma 6 we get

\[
\frac{\mathcal{W}(\sigma, k)}{k!} \leq \gamma(b)^{k-\sigma} \sum_{k_1, \ldots, k_\sigma \geq 1} \left[ \frac{1}{k_1 \cdots k_\sigma} \right]^{b} \leq \gamma(b)^{k-\sigma} \zeta(b)^\sigma
\]
proving that \( \frac{\max_{\sigma} \{W(\sigma, k)\}}{k!} \leq \max\{\gamma(b), \zeta(b)\}^k \). A slightly more complicated argument produces a better bound. As we done in Section 2.1, we decompose (7) according to the number or variables whose index is \( \leq k \), where \( \bar{k} \) is a parameter that we will set later:

\[
\frac{W(\sigma, k)}{k!} = \sum_{h=0}^{\sigma} \binom{\sigma}{h} \sum_{1 \leq k_1, \ldots, k_h \leq k} \prod_{j=1}^{\sigma} \frac{W(k_j)}{k_j!}.
\]

We bound this sum from above by eliminating the constraint \( k_1 + \cdots + k_{\sigma} = k \); moreover, we introduce the quantities \( w_j := W(j)/(j! \gamma(b)^{j-1}) \) for \( j \leq \bar{k} \), so that each \( W(k_j)/k_j! \) can be estimated by \( w_{k_j} \gamma(b)^{k_j-1} \) when \( k_j \leq \bar{k} \) and by \( \gamma(b)^{k_j-1} \bar{k}_j^{-b} \) for \( k_j > \bar{k} \), obtaining

\[
\frac{W(\sigma, k)}{k!} \leq \gamma(b)^{k - \sigma} \sum_{h=0}^{\sigma} \binom{\sigma}{h} \sum_{k_1, \ldots, k_{\sigma} \geq 1} \left( \frac{k_1}{k_1 + \cdots + k_{\sigma}} \right)^{k - \sigma} \frac{w_{k_1} \cdots w_{k_{\sigma}}}{(k_{\sigma+1} \cdots k_{\sigma})^b}.
\]

\[
= \gamma(b)^{k - \sigma} \sum_{h=0}^{\sigma} \binom{\sigma}{h} \left[ \sum_{j=1}^{k} w_j \right]^h \left[ \zeta(b) - \sum_{j=1}^{k} j^{-b} \right]^{\sigma - h}
\]

\[
= \gamma(b)^{k - \sigma} \left[ \zeta(b) - \sum_{j=1}^{k} (j^{-b} - w_j) \right]^{\sigma},
\]

proving that

\[
(17) \quad \frac{\max_{\sigma} \{W(\sigma, k)\}}{k!} \leq \max\{\gamma(b), \psi(b)\}^k
\]

where for convenience we have set \( \psi(b) := \zeta(b) - \sum_{j=1}^{\bar{k}} (j^{-b} - w_j) \). The new upper bound improves the previous one because \( \psi(b) < \zeta(b) \), since each \( w_j \) is lower than \( j^{-b} \) (by Lemma 6); for the same reason in these computations we choose \( \bar{k} \) as large as possible: \( \bar{k} = 1500 \).

In order to bound \( \max_{\sigma} \{W(\sigma, k)\} \) from above we must find the smallest value for \( \max\{\gamma(b), \psi(b)\} \). We know that \( \gamma(1.6056) \leq 1.71185 \), but \( \psi(1.6056) > 1.8 \) so a different (larger) \( b \) must be chosen. Proceeding as we have shown before we arrive to the (almost) optimal choice \( b = 1.6578 \), giving: \( y_0(b) \leq 1.75781 \), \( x_4(b) \leq 44.32 \), \( y(b) \leq 1.66746 \), \( \gamma(b) \leq 1.75761 \) and \( \psi(b) \leq 1.75772 \) so that (17) yields the bound \( \max_{\sigma} \{W(\sigma, k)\} \leq \psi(b)^{k!} \). We obtain the claim in Theorem 2 by using the explicit inequality \( k! \leq (k/e)^k \sqrt{2\pi k} e^{1/12k} \).
3. Proof of Theorem 1

Now we can prove Theorem 1. Let $\varrho, h_{\mu,\nu}(\varrho, x)$ and $c_{\mu,\nu}(\varrho)$ be defined as in that theorem and suppose that $q$ diverges along a sequence $S$ for which $\tau/\mathcal{L} \in [\mu, \nu]$. Let $\mathcal{T}(q; k)$ be the number of solutions of the congruence

$$\sum_{i=1}^{k} 2^{n_i} = \sum_{j=1}^{k} 2^{m_j} \pmod{q},$$

with $0 \leq n_i, m_j < \tau$ for every $i$ and $j$. The congruence means that

$$\sum_{i=1}^{k} 2^{n_i} = \sum_{j=1}^{k} 2^{m_j} + qw$$

for some $w \in \mathbb{Z}$: suppose $w \geq 0$, then $w < k2^\tau/2q$, so that there are $k2^\tau/2q \cdot \tau^k$ possible choices for the values of the set of parameters $w$ and $m_j$s; for every such choice there are $\leq (\varrho k)^k \cdot k^{O(1)}$ solutions for $n_1, \ldots, n_k$, hence we have $(\varrho k)^k 2^\tau \cdot k^{O(1)}/2q$ solutions, at most. If $w < 0$ we obtain the same bound by moving $w$ to the L.H.S. in (18), hence we have proved that

$$\mathcal{T}(q; k) \leq \frac{2^\tau}{q} (\varrho k \tau)^k \cdot k^{O(1)}. \tag{19}$$

The second inequality of Lemma 3.1 in [11] says that

$$\max_{\xi: \xi^q=1} \{ |s(\xi)| \} \leq q^{1/2k^2} \mathcal{T}(q; k)^{1/k^2} \tau^{1-2/k},$$

for every integer $k \geq 1$, so that (19) gives

$$\max_{\xi: \xi^q=1} \{ |s(\xi)| \} \leq q^{1/2k^2} \left( \frac{2^\tau}{q} (\varrho k \tau)^k \cdot k^{O(1)} \right)^{1/k^2} \tau^{1-2/k}. \tag{20}$$

By hypothesis, for $q \in S$ we have $\tau/\mathcal{L} \leq \nu$ so that $2^\tau/q \leq q^{\nu-1}$, hence

$$\max_{\xi: \xi^q=1} \{ |s(\xi)| \} \leq \tau \exp \left( (\nu - \frac{1}{2}) \frac{\log q}{k^2} + \frac{\log(\varrho k / \tau)}{k} + O\left( \frac{\log^2 k}{k^2} \right) \right).$$

In this inequality we set $k = \lceil \tau/x \rceil$ for a constant $x \geq 1$ that we will choose later, obtaining that

$$\max_{\xi: \xi^q=1} \{ |s(\xi)| \} \leq \tau \exp \left( (\nu - \frac{1}{2}) x^2 \frac{\mathcal{L}}{\tau^2} \log 2 - x \frac{\log(x/\varrho)}{\tau} + O_{x,\nu}\left( \frac{\log(\mathcal{L}/\tau)}{\tau^2} \right) \right).$$
By hypothesis, for $q \in S$ we have also $\mathcal{L}/\tau \leq \mu^{-1}$ so that we deduce from the previous inequality that

$$
\max_{\xi: \xi \neq 1 \text{ primitive}} \{|s(\xi)|\} \leq \tau - h_{\mu,\nu}(q, x) + o_{x,\mu,\nu}(1).
$$

The proof concludes by choosing for $x$ the value $x_{\mu,\nu}$ for which $h_{\mu,\nu}(q, x_{\mu,\nu}) = c_{\mu,\nu}(q)$ in the previous inequality.

**Remark.** At the web page [www.mat.unimi.it/users/molteni/research/cancellation/paper.html](http://www.mat.unimi.it/users/molteni/research/cancellation/paper.html) we have collected both the data files and the macros written with the PARIgp [19] programming language which are necessary for the computations contained in this paper.

**Acknowledgements.** The author thanks A. Languasco, A. Giorgilli and A. Perelli for many interesting remarks and suggestions about the subject of this paper.

**References**


Dipartimento di Matematica, Università di Milano, via Saldini 50, I-20133 Milano, Italy
*E-mail address: giuseppe.molteni1@unimi.it*