CYCLONOMIAL NUMBER SYSTEMS AND THE RANKING OF LEXICOGRAPHICALLY ORDERED CONSTANT-SUM CODES

C. Roos, A. Snijders and A.J. van Zanten

Delft University of Technology
Faculty of Technical Mathematics and Informatics
Mekelweg 4
2628 CD Delft
The Netherlands

Abstract

Cyclonomial coefficients are defined as a generalization of binomial coefficients. It is proved that each natural number can be expressed, in a unique way, as the sum of cyclonomial coefficients, satisfying certain conditions. This cyclonomial number system generalizes the well-known binomial number system. It appears that this system is the appropriate number system to index the words of the lexicographically ordered code $L^q(n,k)$. *This code consists of all words of length* n *over an alphabet of* q *symbols, such that the sum of the digits is constant. It provides efficient algorithms for the conversion of such a codeword to its index, and vice versa.*

Index Terms - binomial number system, cyclonomial coefficients, cyclonomial number system, lexicographic constant-sum code, ranking problem, index system.

AMS class 05A15, 69G10, 69G11

1 Introduction

Bionomial coefficients have been generalized in many ways. In Section 2 of this paper we introduce a generalization $\binom{n}{k}_q$, for any integer $q \geq 2$, which we call a *cyclonomial coefficient*. It appears hat these coefficients share a number of properties with binomial coefficients (the $q = 2$-case). Among other things there exists, for any $k \geq 1$ and for any $q \geq 2$, a *cyclonomial number system*, which generalizes the concept of *binomial number system* (cf. [2]). More specifically, it is proved in Section 3 that, for fixed values of

k and q, any natural number can, in a unique way, be written as the sum of k cyclonomial coefficients $\binom{b_k}{k}_q, \binom{b_{k-1}}{k-1}_q, \cdots, \binom{b_1}{1}_q$, with $b_k \geq b_{k-1} \geq \cdots \geq b_1 \geq 0$, and such that at most $q-1$ consecutive numbers b_i are equal.
In Section 4 we present an application of the cyclonomial number system. It is shown that this system is the appropriate tool for indexing the words of a *lexicographic constant-sum code* over an alphabet of q elements.
We define such a code $L^q(n,k)$ as the lexicographically ordered list of all words over $R := \{0, 1, \cdots, q-1\}$, of length n, and such that the sum of the digits of each word is equal to k. In [1] a method is developed to calculate the index values of the words of a more general lexicographically ordered list. It appears that in the special case of $L^q(n,k)$ the *index-value* or *rank* of a codeword can quite naturally be expressed as a number in the cylonomial number system for that particular value of q. Hence, the cyclonomials number system is the appropriate number system to solve the *ranking problem* of $L^q(n,k)$, i.e. it provides us with efficient algorithms to convert a codeword to its index, and vice versa. We remark that such a ranking problem exists for any ordered list of combinatorial objects, and that inherent to this problem one can pose the question of an appropriate number system (cf. refs. [3,4,5,6]). The solution presented in Section 4 generalizes the solution for $q = 2$, given in [5], in terms of the binomial number system.
Finally, in Section 5, we demonstrate by an example an efficient way of converting a codeword to its index, and vice versa, using the cyclonomial number system.

2 Cyclonomial coefficients

Let n, k and q be natural numbers with $q \geq 2$, and let $V^q(n,k)$ be the set of words

$$w := w_{n-1}w_{n-2}\cdots w_0, \tag{1}$$

with $w_i \in \{0, 1, \cdots, q-1\}$, $0 \leq i \leq n-1$, and such that

$$\sum_{i=0}^{n-1} w_i = k. \tag{2}$$

The words w usally are called codewords of length n and weight k over an alphabet of q elements. The set $V^q(n,k)$ can be referred to as the complete q-ary constant-sum code of length n and weight k. Furthermore we define the number

$$\binom{n}{k}_q := |V^q(n,k)| \tag{3}$$

as the total number of words in $V^q(n,k)$. One could also say that $\binom{n}{k}_q$ is equal to the total number of compositions of k into n parts the size of which do not exceed $q-1$.
In the binary case this number is just the normal binomial coefficient, i.e. $\binom{n}{k}_2 \equiv \binom{n}{k}$. Therefore, it will be obvious that the numbers $\binom{n}{k}_q$ have properties which are generalizations of well-known relations for binomial coefficients. In the first place we mention the symmetry property

$$\binom{n}{k}_q = \binom{n}{\lambda n - k}_q, \tag{4}$$

where we introduce the abbreviation

$$\lambda := q - 1. \tag{5}$$

Equality (4) follows immediately from the one-to-one correspondence between the codewords w_{n-1}, w_{n-2}, $\cdots$, w_0 and $\lambda - w_{n-1}$, $\lambda - w_{n-2}$, $\cdots$, $\lambda - w_0$.
Next we give the generating function for the numbers $\binom{n}{k}_q$, which follows from their definition

$$(1 + x + \cdots + x^{q-1})^n = \sum_{k=0}^{\lambda n} \binom{n}{k}_q x^k. \tag{6}$$

Because of the occurence of the cyclotomic polynomial $1 + x + \cdots + x^{q-1}$ in the *lhs* of (6) we call $\binom{n}{k}_q$ *a cyclonomial coefficient of order* q. We can easily derive, either from (3) or from (6), the special cases

$$\binom{n}{0}_q = 1, \quad \binom{n}{1}_q = n, \tag{7}$$

and

$$\binom{1}{k}_q = \begin{cases} 1, & 0 \leq k \leq q-1, \\ 0, & q-1 < k. \end{cases} \tag{8}$$

It is also obvious from the definition of $V^q(n,k)$ that for $q > n$

$$\binom{n}{k}_q = \binom{k+n-1}{k}, \tag{9}$$

since for those q-values we have to do with compositions of k into n parts, where the size of the parts is not restricted by the size of the alphabet. Another simple consequence of (6), which follows by putting $x = 1$, is

$$\sum_{k=0}^{\lambda n} \binom{n}{k}_q = q^n. \tag{10}$$

A slightly less trivial property is obtained by applying (6) to the identity

$$(1 + x + \cdots + x^{q-1})^{n+m} = (1 + x + \cdots + x^{q-1})^n (1 + x + \cdots + x^{q-1})^m.$$

This yields

$$\sum_{i+j=k} \binom{n}{i}_q \binom{m}{j}_q = \binom{n+m}{k}_q. \tag{11}$$

By substituting $m = 1$ in (11) and using (8) we get

$$\sum_{i=k-\lambda}^{k} \binom{n}{i}_q = \binom{n+1}{k}_q. \tag{12}$$

This last equality, which clearly generalizes the "basic" addition theorem for binomial coefficients, enables us to construct a Pascal triangle for the numbers $\binom{n}{k}_q$, for any $q \geq 2$. In Figure 2.1 this has been done for $q = 3$, and $0 \leq n \leq 7$, $0 \leq k \leq 14$.

$k\backslash n$	0	1	2	3	4	5	6	7
0	1	1	1	1	1	1	1	1
1	0	1	2	3	4	5	6	7
2	0	1	3	6	10	15	21	28
3	0	0	2	7	16	30	50	77
4	0	0	1	6	19	45	90	161
5	0	0	0	3	16	51	126	266
6	0	0	0	1	10	45	141	357
7	0	0	0	0	4	30	126	393
8	0	0	0	0	1	15	90	357
9	0	0	0	0	0	5	50	266
10	0	0	0	0	0	1	21	161
11	0	0	0	0	0	0	6	77
12	0	0	0	0	0	0	1	28
13	0	0	0	0	0	0	0	7
14	0	0	0	0	0	0	0	1

Fig. 2.1

A second tool for the computation of cyclonomial coefficients is a relation which expresses them in terms of binomial coefficients. Writing

$$(1+x+\cdots+x^{q-1})^n = \left(\frac{1-x^q}{1-x}\right)^n = \frac{1}{(1-x)^n}\sum_{i=0}^{n}(-1)^i\binom{n}{i}x^{iq} =$$

$$\sum_{t=0}^{\infty}\sum_{i=0}^{n}(-1)^i\binom{n}{i}\binom{n-1+t}{n-1}x^{t+iq},$$

and applying (6) provides us with

$$\binom{n}{k}_q = \sum_{i=0}^{n}(-1)^i\binom{n}{i}\binom{n-1+k-iq}{n-1}. \tag{13}$$

3 The cyclonomial number system

In this section k and q, and consequently $\lambda := q-1$, are fixed, but otherwise arbitrary integers, with $k \geq 1$ and $q \geq 2$. We shall show that each natural number can in a unique way be represented as a sum of k cyclonomial coefficients of order q, or stated equivalently, that for any such k and q there exists a number system to represent the natural numbers. The existence of these systems is based on the following theorem.

Theorem 3.1
Let $k \geq 1$ and $q \geq 2$. Then any natural number N can uniquely be written as

$$N = \binom{b_k}{k}_q + \binom{b_{k-1}}{k-1} + \cdots + \binom{b_1}{1}_1,$$

with $b_k \geq b_{k-1} \geq \cdots \geq b_1 \geq 0$, and such that at most λ consecutive numbers b_i are equal.

Proof

A. First we establish the existence of an expression as stated in the Theorem. We choose b_k as large as possible such that $\binom{b_k}{k}_q \leq N$. Next we choose b_{k-1} as large as possible such that $\binom{b_{k-1}}{k-1}_1 \leq N - \binom{b_k}{k}_q$. Continue with choosing all numbers b_l, $k \geq l \geq 2$, as large as possible and such that $\binom{b_l}{l}_q \leq N - \sum_{i=l+1}^{k} \binom{b_i}{i}_q$. Finally we put $b_1 := N - \sum_{i=2}^{k} \binom{b_i}{i}_q$, and we end up with the stated expression for N.
Now suppose that $b_{k-1} > b_k$. Then we can write

$$N \geq \binom{b_k}{k}_q + \binom{b_{k-1}}{k-1}_q \geq \binom{b_k}{k}_q + \binom{b_k+1}{k-1}_q =$$

$$\binom{b_k}{k}_q + \binom{b_k}{k-1}_1 + \binom{b_k}{k-2}_q + \cdots + \binom{b_k}{k-\lambda+1}_q + \binom{b_k}{k-\lambda}_q =$$

$$\binom{b_k+1}{k}_q + \binom{b_k}{k-\lambda}_q \geq \binom{b_k+1}{k}_q,$$

where we applied twice the addition formula (12). It is clear that the last inequality contradicts the assumption that b_k is the largest integer with $\binom{b_k}{k} \leq N$.
Hence, we may conclude that $b_k \geq b_{k-1}$. In the same way it can be proved that $b_{k-1} \geq b_{k-2} \geq \cdots \geq b_1$.

Next, suppose there is an index l, $1 \leq l \leq k - \lambda$, such that $b_l = b_{l+1} = \cdots = b_{l+\lambda}$. Because of (12), with k replaced by l, we then have

$$N - \sum_{i=l+q}^{k} \binom{b_i}{i}_q \geq \sum_{j=0}^{\lambda} \binom{b_{l+j}}{l+j}_q = \binom{b_{l+\lambda}+1}{l+\lambda}_q,$$

contradicting the assumption that $b_{l+\lambda}$ is the largest integer satisfying $\binom{b_{l+\lambda}}{l+\lambda}_q \leq N - \sum_{i=l+q}^{k} \binom{b_i}{i}_q$. So there can be at most λ consecutive numbers b_i equal to each other.

B. The second part of the proof concerns the uniqueness of the representation.

Let $N = \sum_{i=1}^{k} \binom{c_i}{i}_q$ be another representation for some integer $N \geq 0$, such that the c_i satisfy the same requirements as the b_i, $1 \leq i \leq k$. Let l be the highest index with $b_i = c_i$ for $l+1 \leq i \leq k$, and $b_l \neq c_l$. Without restriction of the generality we assume that $b_l < c_l$ and define $c := c_l$. Furthermore we define unique integers a and b, with $1 \leq b \leq \lambda$, satisfying $l = a\lambda + b$. In order to come to a contradiction we write

$$M := N - \sum_{i=l+1}^{k} \binom{b_i}{i}_q = \sum_{i=1}^{l} \binom{b_i}{i}_q \leq$$

$$\sum_{j=1}^{a}\sum_{t=1}^{\lambda} \binom{c-j}{l-j\lambda+t}_q + \sum_{t=1}^{b} \binom{c-a-1}{t}_q <$$

$$\sum_{j=1}^{a}\sum_{t=1}^{\lambda} \binom{c-j}{l-j\lambda+t}_q + \binom{c-a}{b}_q.$$

For the first inequality we made use of $b_1 \leq b_2 \leq \cdots \leq b_l < c_l = c$, and of the property that at most λ consecutive numbers b_i can be equal, and for the second one we applied (12). By successively applying (12) again a number of times we can write for the *rhs* of the last inequality

$$\sum_{j=1}^{a-1}\sum_{t=1}^{\lambda} \binom{c-j}{l-j\lambda+t}_q + \sum_{t=0}^{\lambda} \binom{c-a}{b+t}_q =$$

$$\sum_{j=1}^{a-1}\sum_{t=1}^{\lambda} \binom{c-j}{l-j\lambda+t}_q + \binom{c-a+1}{b+\lambda}_q = \cdots =$$

$$\sum_{j=1}^{a-2}\sum_{t=1}^{\lambda} \binom{c-j}{l-j\lambda+t}_q + \binom{c-a+2}{b+2\lambda}_q = \cdots = \binom{c}{b+a\lambda}_q = \binom{c}{l}_q.$$

We conclude that $M < \binom{c}{l}_q$. On the other hand it follows from the definition of M that $M \geq \binom{c_l}{l}_q = \binom{c}{l}_q$. So we have a contradiction, and therefore there is no index l with $b_l \neq c_l$. This proves the uniqueness of the representation derived in part A. □

The number system based on the contents of Theorem 2.1 is called *the cyclonomial number system* (for those particular values of k and q), and we shall use the notation

$$N = (b_k b_{k-1} \cdots b_1)_q. \tag{14}$$

For $q = 2$ we obtain the well-known binomial number system (cf. [2]). As an example we present in Fig. 3.1 the representation of the integers $0, 1, \cdots, 15$ in the cyclonomial number system for $k = 5$ and $q = 3$.

$0 = (21100)_3$	$8 = (32211)_3$
$1 = (22100)_3$	$9 = (33100)_3$
$2 = (22110)_3$	$10 = (33110)_3$
$3 = (31100)_3$	$11 = (33200)_3$
$4 = (32100)_3$	$12 = (33210)_3$
$5 = (32110)_3$	$13 = (33211)_3$
$6 = (32200)_3$	$14 = (33220)_3$
$7 = (32210)_3$	$15 = (33221)_3$

Fig. 3.1

4 The lexicographic constant-sum code $L^q(n,k)$

We take the codewords of the code $V^q(n,k)$ of Section 2 and arrange them in *lexicographic order*. If both $w = w_{n-1}w_{n-2}\cdots w_0$ and $v = v_{n-1}v_{n-2}\cdots v_0$ are words of $V^q(n,k)$, then one says that w is lexicographically less than v if and only if, for some $l \geq 1$, one has $w_i = v_i$ for all $i < l$, and $w_l < v_l$. The ordered list of codewords we get in this way is called *the lexicograhic constant-sum code* $L^q(n,k)$, over the alphabet $\{0,1,\cdots,q-1\}$. Each word in this list is given an *index* or *rank*, which ranges from 0 to $\binom{n}{k}_q - 1$. As an example we present in Fig. 4.1 the code $L^3(4,5)$.

$\underline{L^3(4,5)}$

index	codeword	index	codeword
0	0122	8	1220
1	0212	9	2012
2	0221	10	2021
3	1022	11	2102
4	1112	12	2111
5	1121	13	2120
6	1202	14	2201
7	1211	15	2210

Fig. 4.1

In order to determine the index of a given codeword $w := w_{n-1}w_{n-2}\cdots w_0$ of $L^q(n,k)$, i.e. to solve the ranking problem of that code (cf. Section 1), we introduce the vector

$$\underline{b} = (n-1,\cdots,n-1,n-2,\cdots,n-2,\cdots,0,\cdots,0), \tag{15}$$

where each integer i occurs w_i times for $n-1 \geq i \geq 0$. Since $\sum_{i=0}^{n-1} w_i = k$ the vector $\underline{b}$ has precisely k components and we write

$$\underline{b} = (b_k, b_{k-1}, \cdots, b_1), \tag{16}$$

where $b_k, b_{k-1}, \cdots, b_1$ are determined by (15). The componets of $\underline{b}$ indicate the positions in w where the k "units" of its weight are located. For this reason $\underline{b}$ is called the *position vector* of w.

We are now ready to formulate the theorem which solves the ranking problem of $L^q(n,k)$.

Theorem 4.1
Let $w = w_{n-1}w_{n-2}\cdots w_0$ be a codeword of $L^q(n,k)$. Then the index of w is given by

$$ind(w) = (b_k b_{k-1}\cdots b_1)_q,$$

where the numbers b_k, b_{k-1}, b_1 are the components of the position vector $\underline{b}$ of w.

Proof
We prove the correctness of the expression for $ind(w)$ by counting the total number of codewords preceding w. For our convenience we introduce the partial sums

$$a_i := \sum_{j=0}^{i} w_j, \quad 0 \le i \le n-1.$$

First we consider the words v with $v_{n-1} = w_{n-1} - t$ for some fixed $t > 0$. The number of these words is equal to $\binom{n-1}{t+a_{n-2}}_q$.
Hence, the number of words v with $v_{n-1} < w_{n-1}$ equals $\sum_{t=1}^{w_{n-1}} \binom{n-1}{t+a_{n-2}}_q$.
In the same way it appears that the number of words v with $v_{n-1} = w_{n-1}$ and $v_{n-2} < w_{n-2}$ is equal to $\sum_{t=1}^{w_{n-2}} \binom{n-2}{t+a_{n-3}}_q$, etc. So we find that the total number of words preceding w is equal to

$$\sum_{j=1}^{n-1}\sum_{t=1}^{w_j} \binom{j}{t+a_{j-1}}_q.$$

We remark that the second sum has to be interpreted as 0 if $w_j = 0$. It is obvious that the set of integers $\{t + a_{j-1} \mid 1 \le j \le n-1, 1 \le t \le w_j\}$ is just the set $\{1, 2, \cdots, k\}$. If we arrange the non-vanishing cyclonomial coefficients in the above expression according to decreasing values of $t + a_{j-1}$, the variable j in these coefficients runs through the set of values as indicated in the *rhs* of (15) from left to right. Hence, applying (16), the index of w is equal to

$$\binom{b_k}{k}_q + \binom{b_{k-1}}{k-1}_q + \cdots + \binom{b_1}{1}_q = (b_k b_{k-1}\cdots b_1)_q.$$

The contents of Theorem 4.1 enable us to compute the index of a given codeword of $L^q(n,k)$. The inverse problem, i.e. computing the codeword corresponding to a given values of the index, can be performed by expressing this value in the cyclonomial number system by means of the method discussed in the proof of Theorem 3.1. From the digits $b_k, b_{k-1}, \cdots, b_1$ the codeword can be constructed immediately. If there are p consecutive digits equal to i, then we have $w_i = p$ for $0 \le i \le n-1$.

We remark that in the expression for $ind(w)$ in Theorem 4.1 to each digit w_i there corresponds a sum of w_i cyclonomial coefficients.
This sum can be considered as a weight assigned to that particular digit. In this sense we can speak of a *weighting system* for $L^q(n,k)$.

5 Example

We take $w = 1202 \in L^3(4,5)$. According to (15) and (16), its position vector is equal to $\underline{b} = (3,2,2,0,0)$. Hence, Theorem 4.1 and the table in Fig. 2.1 give

$$ind(w) = \binom{3}{5}_3 + \binom{2}{4}_3 + \binom{2}{3}_3 + \binom{0}{2}_3 + \binom{0}{1}_3 = 3+1+2+0+0 = 6.$$

Fig. 4.1 shows that this is the correct value.

Conversely, we want to know the codeword $v \in L^3(4,5)$ with $ind(v) = 12$. To this end we write 12 in the cyclonomial number system for $q = 3$ and $k = 5$. First we determine the largest integer b_5 such that $\binom{b_5}{5}_3 \leq 12$. From the table in Fig. 2.1 we conclude that $b_5 = 3$. Next we determine the largest integer b_4 such that $\binom{b_4}{4}_3 \leq 12 - \binom{3}{5}_3 = 12 - 3 = 9$. We find $b_4 = 3$. Continuing in this way yields $12 = (33210)_3$. Hence, the position vector of v is equal to $\underline{b} = (3,3,2,1,0)$. Since there are two components of $\underline{b}$ equal to 3, we may conclude that $v_3 = 2$. Furthermore there is one component equal to 2, so $v_2 = 1$. Analogously we find $v_1 = 1$ and $v_0 = 1$. Hence, $v = 2111$. Again Fig. 4.1 shows that this answer is correct.

References

[1] Clements, C.F., *More on the Generalized Macaulay Theorem*, Discrete Math. 1 (1971) 247-255.

[2] Knuth, D.E., *Fundamental Algorithms*, Reading Mass., Addison-Wesley, 1969.

[3] Reingold, E.M., Nievergelt, J. and Deo, N., *Combinatorial Algorithms: Theory en Practice*, Englewood Cliffs, N.J., Prentice-Hall, 1977.

[4] Wilf, H.S., *Combinatorial Algorithms: An Update*, CBMS-NSF Regional Conference Series in Applied Mathematics, Society for Industrial Applied Mathematics, Philadelphia, Penn., 1989.

[5] Zanten, A.J. van, *Index Systems and Separability of Constant Weight Gray Codes*, IEEE Trans. of Information Theory, 37 (1991) 1229-1233.

[6] Zanten, A.J. van, *The Ranking Problem of a Gray Code for Compositions*, Ars Combinatoria, 41 (1995) 257-268.