Discrete Mathematics 106/107 (1992) 181–188
North-Holland

Binomial coefficient codes over GF(2)

Persi Diaconis
Harvard University, Cambridge, MA 02138, USA

Ron Graham
AT & T Bell Laboratories, Murray Hill, NJ 07974, USA

Received 10 December 1991
Revised 5 February 1992

Abstract

Diaconis, P. and R. Graham, Binomial coefficient codes over GF(2) Discrete Mathematics 106/107 (1992) 181–188.

In this note we study codes over GF(2) which are generated for given d and r by binary vectors of the form $(\binom{0}{i}, \binom{1}{i}, \ldots, \binom{j}{i}, \ldots, \binom{2^r-1}{i}) \pmod 2$, $0 \leq i \leq d$. We describe the weight enumerators of these codes and the numbers of codewords of weights 1 and 2. These results can be used to obtain sharp bounds on the rates of convergence to uniformity for certain random walks on the n-cube $GF(2)^n$.

1. Introduction

For fixed $r \geq 0$ and $2^{r-1} < d \leq 2^r$, let W_i be the binary n-tuple defined by

$$W_i = (W_i(0), W_i(1), \ldots, W_i(2^r - 1)), \quad 0 \leq i < d,$$

where $W_i(j) \equiv \binom{j}{i} \pmod 2$. Define $\mathscr{C}_d$ to be the linear code (i.e., vector space over GF(2)) generated by the *words* W_i, $0 \leq i < d$. For $W \in \mathscr{C}_d$, let $|W|$, the *weight* of W, denote the number nonzero entries of W. Finally, N_k will denote the number of words of weight k in $\mathscr{C}_d$, and

$$D_d(t) := \sum_{k=0}^{n} N_k t^k$$

will denote the *weight enumerator* of $\mathscr{C}_d$ (for general coding theory references, see [3]). We call the $\mathscr{C}_d$ *binomial coefficient codes* for the obvious reason.

Correspondence to: R.L. Graham, Research, Information Sciences Division, AT & T Bell Laboratories, Room 2c-380, 600 Mountain Avenue, Murray Hill, NY 07974, USA.

As will be explained in Section 3, knowledge of the weight structure of $\mathscr{C}_d$ can be used to derive rather tight bounds on convergence rates of certain random walks of the n-cube. Our objective of this paper will be to point out several facts concerning $D_d(t)$.

2. The main results

Theorem 1. *The weight enumerators* $D_d(t)$ *are determined by the following recurrences*:

$$D_1(t) = 1 + t, \tag{1}$$

$$D_{2m}(t) = D_m(t)^2, \qquad m \geq 1, \tag{2}$$

$$D_{2m+1}(t) - D_{2m}(t) = (D_{m+1}(t) - D_m(t))^2, \qquad m \neq 2^s, \tag{3}$$

$$D_{2^s+1}(t) = (1 + t^2)^{2^s} + (2t)^{2^s}. \tag{4}$$

Proof. We first note the following modular relations between binomial coefficients, all of which follow from the fact (e.g., see [1]) that the power of 2 which divides $\binom{a+b}{a}$ is just the number of 'carries' occurring in the base 2 addition of a and b.

$$\begin{aligned} &\binom{2j}{2i} \equiv \binom{j}{i} \pmod 2, && \binom{2j+1}{2i} \equiv \binom{j}{i} \pmod 2, \\ &\binom{2j}{2i+1} \equiv 0 \pmod 2, && \binom{2j+1}{2i+1} \equiv \binom{j}{i} \pmod 2. \end{aligned} \tag{5}$$

The proofs of (2), (3) and (4) are recursive. To form the basic recursion, let V be the d by 2^r array formed by taking W_i as its ith row, $0 \leq i < d$. Let V_0 and V_1 be d by 2^{r-1} arrays formed from the even and odd columns of V, respectively. It follows from (5) and the definition of V that V_0 and V_1 have the following form (where all quantities are considered modulo 2):

$$V_0: \quad \begin{matrix} \binom{0}{0} & \binom{1}{0} & \cdots & \binom{j}{0} & \cdots & \binom{2^{r-1}-1}{0} \\ 0 & 0 & \cdots & 0 & \cdots & 0 \\ & & & \vdots & & \\ \binom{0}{i} & \binom{1}{i} & \cdots & \binom{j}{i} & \cdots & \binom{2^{r-1}-1}{j} \\ 0 & 0 & \cdots & 0 & \cdots & 0 \\ & & & \vdots & & \end{matrix}$$

$$V_1:\quad \begin{matrix} \binom{0}{0} & \binom{1}{0} & \cdots & \binom{j}{0} & \cdots & \binom{2^{r-1}-1}{0} \\ \binom{0}{0} & \binom{1}{0} & \cdots & \binom{j}{0} & \cdots & \binom{2^{r-1}-1}{0} \\ & & & \vdots & & \\ \binom{0}{i} & \binom{1}{i} & \cdots & \binom{j}{i} & \cdots & \binom{2^{r-1}-1}{i} \\ \binom{0}{i} & \binom{1}{i} & \cdots & \binom{j}{i} & \cdots & \binom{2^{r-1}-1}{i} \\ & & & \vdots & & \end{matrix}$$

The exact form of the last rows of V_0 and V_1 will depend on the parity of d and will determine the differences between parts (2), (3) and (4) of Theorem 1.

Proof of (2). Here, $d = 2m$. The last row of V is W_{2m-1} and the last rows of V_0 and V_1 are (mod 2):

$$V_0:\quad \begin{matrix} & & & \vdots & & \\ \binom{0}{m-1} & \binom{1}{m-1} & \cdots & \binom{j}{m-1} & \cdots & \binom{2^{r-1}-1}{m-1} \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{matrix}$$

$$V_1:\quad \begin{matrix} & & & \vdots & & \\ \binom{0}{m-1} & \binom{1}{m-1} & \cdots & \binom{j}{m-1} & \cdots & \binom{2^{r-1}-1}{m-1} \\ \binom{0}{m-1} & \binom{1}{m-1} & \cdots & \binom{j}{m-1} & \cdots & \binom{2^{r-1}-1}{m-1} \end{matrix}$$

The code C_{2m} is formed by taking sums of all possible subsets of rows of V. For the rows W_{2i} and W_{2i+1} there are four possibilities: take neither, take W_{2i} alone, take W_{2i+1} alone, and take both. Consider the effect of these choices on the corresponding pairs of rows in V_0 and V_1. In the first case (neither), $(00\cdots 0)$ is added to both V_0 and V_1. In the second case (W_{2i} alone), $(\binom{0}{i}\binom{1}{i}\cdots\binom{j}{i}\cdots\binom{2^{r-1}-1}{i})$ is added to both V_0 and V_1. In the third case (W_{2i+1} alone), $(00\cdots 0)$ is added to V_0 and $(\binom{0}{i}\cdots\binom{2^{r-1}-1}{i})$ is added to V_1. Finally, in the last case (both), this has the effect of adding $(\binom{0}{i}\cdots\binom{2^{r-1}-1}{i})$ to V_0 and $(00\cdots 0)$ to V_1. Thus, the four possibilities of adding or not adding the row $(\binom{0}{i}\cdots\binom{2^{r-1}-1}{i})$ to V_0 and V_1 each occur exactly once. Of course, this holds in general for each of the $m = d/2$ pairs of rows in V. Hence, in generating $\mathscr{C}_{2m}$ we are actually generating $\mathscr{C}_m$ independently in both V_0 and V_1. This immediately implies $D_{2m}(t) = D_m(t)^2$, which is (2).

Proof of (3). Here, $d = 2m + 1$ with $m \neq 2^s$. The last three rows of V_0 and V_1 are (mod 2):

$$
\begin{array}{lccccccc}
 & \binom{0}{m-1} & \binom{1}{m-1} & \cdots & \binom{j}{m-1} & \cdots & \binom{2^{r-1}-1}{m-1} \\
V_0: & 0 & 0 & \cdots & 0 & \cdots & 0 \\
 & \binom{0}{m} & \binom{1}{m} & \cdots & \binom{j}{m} & \cdots & \binom{2^{r-1}-1}{m} \\
 & & & & \vdots & & \\
 & \binom{0}{m-1} & \binom{1}{m-1} & \cdots & \binom{j}{m-1} & \cdots & \binom{2^{r-1}-1}{m-1} \\
V_1: & \binom{0}{m-1} & \binom{1}{m-1} & \cdots & \binom{j}{m-1} & \cdots & \binom{2^{r-1}-1}{m-1} \\
 & \binom{0}{m} & \binom{1}{m} & \cdots & \binom{j}{m} & \cdots & \binom{2^{r-1}-1}{m}
\end{array}
$$

In this case, each of V_0 and V_1 have a single repeated unpaired row. In each of V_0 and V_1 the codes generated by the first $2m$ rows are $\mathscr{C}_m$ as before. The last row is what changes $\mathscr{C}_{2m}$ (and so, $D_{2m}(t)$) into $\mathscr{C}_{2m+1}$ (and so, $D_{2m+1}(t)$). Thus,

$$D_{2m+1}(t) = D_{2m}(t) + (D_{m+1}(t) - D_m(t))^2$$

which is (3).

Proof of (4). In this case $d = 2^s + 1$. In going from 2^s to 2^{s+1}, the new row added has the form (mod 2):

$$
\begin{array}{lcccccccc}
W_{2^s} = & \binom{0}{2^s} & \binom{1}{2^s} & \cdots & \binom{2^s-1}{2^s} & \binom{2^s}{2^s} & \cdots & \binom{2^{s+1}}{2^s} \\
\equiv & 0 & 0 & \cdots & 0 & 1 & \cdots & 1
\end{array}
$$

by (5). The lengths of the words jump from 2^s to 2^{s+1}. Thus, the array V appears as

$$
\begin{array}{cc}
W_0 & W_0 \\
W_1 & W_1 \\
\vdots & \vdots \\
W_{2^s-1} & W_{2^s-1} \\
00\cdots 0 & 11\cdots 1
\end{array}
$$

with all vectors having length 2^s. Hence, a word consisting of any linear combination of the first 2^s rows has the form (Z, Z) where $Z \in \mathscr{C}_{2^s}$, therefore, the code generated by the first 2^s rows has weight enumerator $D_{2^s}(t^2)$. Adding the final row gives words of the form $(Z, \bar{Z})$ where $\bar{Z}$ is coordinate-wise complement of Z. Any such word has weight 2^s, and there are $|\mathscr{C}_{2^s}| = 2^{2^s}$ such words. Since all the rows of V are linearly independent then $D_{2^s}(t) = (1+t)^{2^s}$. Thus,

$$D_{2^s+1}(t) = (1+t^2)^{2^s} + 2^{2^s} t^{2^s}$$

which implies (4). □

It follows from Theorem 1 that $N_0 = 1$ for all d, and

$$N_1 = \begin{cases} 2^r & \text{if } d = 2^r, \\ 0 & \text{otherwise.} \end{cases}$$

In our next result, we describe the set S_2 of words in $\mathscr{C}_d$ of weight 2. By Theorem 1, if $d = 2^r$ then S_2 consists of all possible words of weight 2 and length 2^r, so that $|S_2| = \binom{2^r}{2}$.

Theorem 2. *Suppose $2^{r-1} < d < 2^r$ and that the binary expansion of d begins with s ones. Then the words in S_2 can be described as follows. Partition the 2^r coordinates into 2^s disjoint blocks each of length 2^{r-s}. Each word W in S_2 can be uniquely specified by selecting two of the 2^s blocks and an integer k, $0 \leq k < 2^{r-s}$. W then has a one in the kth position of each of the two selected blocks and zeros everywhere else. In particular, $|S_2| = \binom{2^s}{2}2^{r-s}$.*

Proof. Let us analyze the structure of the array $V = V(d)$ formed from the rows $W_0, W_1, \ldots, W_{d-1}$. Write $r = s + t$ so that d begins with s ones, then a zero, then $t-1$ following binary digits. In particular,

$$d \leq 2^{r-1} + 2^{r-2} + \cdots + 2^t + 2^{t-1} - 1.$$

The array V can be pictured as shown in Fig. 1.

The lower line L which defines the lower boundary of the array is above row $W_{2^{r-1}+\cdots+2^t+2^{t-1}}$. Hence, any subset sum of rows between L and $L' = 2^{r-1} + 2^{r-2} + \cdots + 2^t$ has the form $00\cdots0XX$ with X of length 2^{t-1} having even weight.

Now note the following:

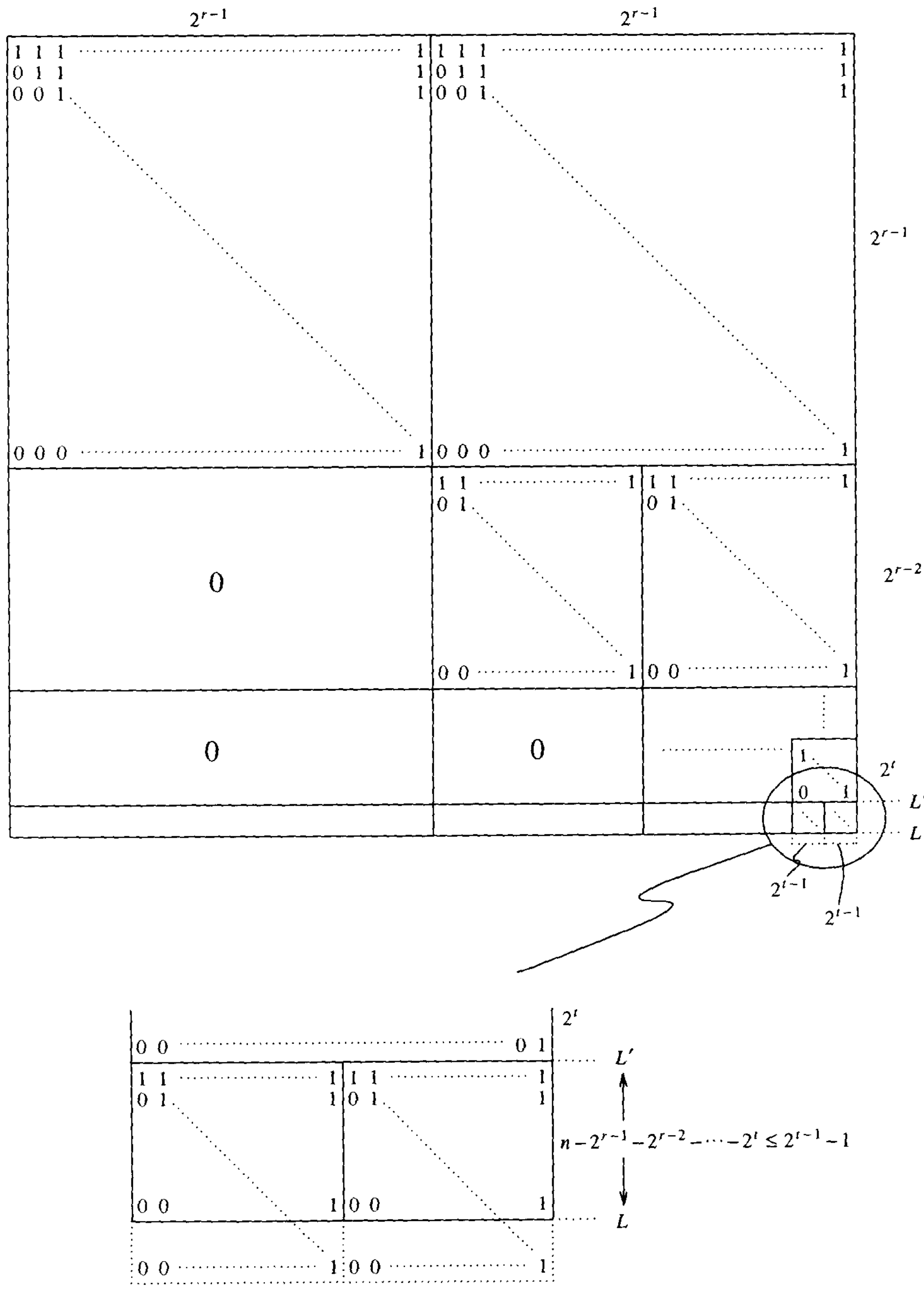
(i) The codes generated by the rows above L' are exactly the codes $\mathscr{C}_{2^r-2^t}$ (using ideas from the proof of Theorem 1 (2)). These are exactly the codes with the following property: for every $k = 0, 1, \ldots, 2^t - 1$, the sum of the entries in positions congruent to $k \pmod{2^t}$ is even;

(ii) The nonzero codes generated by rows below L' lie in the last block of positions and have at least four ones in each word.

Theorem 2 now follows at once from these remarks. □

3. Applications

We give a brief sketch of the problem which motivated our investigations here. Full details can be found in [1].



2^{r-1}
2^{r-1}
2^{r-1}
2^{r-2}
0
0
0
2^t
L'
L
2^{t-1}
2^{t-1}
2^t
L'
$n-2^{r-1}-2^{r-2}-\cdots-2^t \le 2^{t-1}-1$
L
2^{t-1}
2^{t-1}

Fig. 1.

We consider the random walk $X_n = AX_{n-1} + \varepsilon_n$ with $X_i \in \mathrm{GF}(2)^d$, A a fixed non-singular lower triangular d by d matrix over GF(2), and ε_n a random vector of disturbance terms. More specifically, take A to have ones on or just below the diagonal (and zero elsewhere), and the ε_n are independent and identically distributed vectors having common distribution

$$\Pr(\varepsilon_n = 0\} = 1 - \theta, \qquad \mathrm{P}(\varepsilon_n = e_1) = \theta$$

with $0 < \theta < 1$ and e_1 the vector with a single one in the first coordinate and zeros elsewhere. Then

$$\lim_{n\to\infty} \Pr(X_n = y) = 1/2^d$$

for any $y \in \mathrm{GF}(2)^d$. If $U(y) = 1/2^d$ denotes the uniform distribution and $Q_n(y) = \Pr\{X_n = y\}$ then the total variation distance between Q_n and U is given by

$$\|Q_n - U\| = \max_{B \subseteq \mathrm{GF}(2)^d} |\Pr\{X_n \in B\} - U(B)|.$$

A typical question in random walks is the estimation of the number of steps needed to force $\|Q_n - U\|$ to be close to 0 (so that X_n is 'close' to being random).

By employing techniques from Fourier analysis, it can be shown (see [1]) that $\|Q_n - U\|$ can be expressed in terms of the weight enumerator D_d of the code $\mathscr{C}_d$. More precisely, if $2^{r-1} < d \leqslant 2^r$ and $n = m \cdot 2^r$ then

$$4\,\|Q_n - U\|^2 \leqslant D_d((1-2\theta)^{2m}) - 1. \tag{6}$$

This explains our motivation for needing to know the structure of the very low weight words in $\mathscr{C}_d$. We conclude with a sharp bound obtained by this method in the particularly simple case that $d = 2^r$.

Theorem 3 [1]. *With $d = 2^r$ and*

$$n = \frac{d(\log d + c)}{2|\log|1-2\theta||},$$

we have

$$\|Q_n - U\| = 1 - 2\Phi(-\tfrac{1}{2}e^{-c/2}b(n/d)) + O(d^{-\frac{1}{2}}),$$

where

$$\Phi(x) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$$

and $b(n/d)$ is the bounded oscillating function given by

$$b(n/d) = (1-2\theta)^{-\{n/d\}}(1 - 4\{n/d\}\theta(1-\theta))^{\frac{1}{2}}$$

with $\{x\}$ denoting the fractional part of x.

4. Concluding remarks

The reason we restricted d to satisfy $2^{r-1} < d \leqslant 2^r$ is that for $d \leqslant 2^{r-1}$, the rows of length 2^r are just repetitions of rows from shorter codes. We have not investigated the structure of words of $\mathscr{C}_d$ of weight 3 (or more). We also have not looked at the behavior of $\mathscr{C}_d$ as a code. Finally, the same questions could be asked over GF(p) for a prime p, or even over $\mathbb{Z}/n\mathbb{Z}$ for general n.

References

[1] P. Diaconis and R.L. Graham, An affine walk on the hypercube, to appear.
[2] D.E. Knuth, The Art of Computer Programming, Vol. 1 (Addison-Wesley, Menlo Park, 2nd ed., 1973).
[3] J.H. van Lint, Introduction to Coding Theory (Springer, Berlin, 1982).