Eleventh International Workshop on Algebraic and Combinatorial Coding Theory
June 16-22, 2008, Pamporovo, Bulgaria *pp. 273-279*

Burst-error correcting codes and lattice paths

Ulrich Tamm tamm@ieee.org
German Language Department of Business Informatics,
Marmara University, Istanbul, TURKEY and
Department of Mathematics, University of Bielefeld, GERMANY

1 Introduction

At the 3rd Waterloo Conference on Combinatorics [16, pp. 341-342], Berlekamp presented the following combinatorial problem. The problem will be illustrated with the following example also due to Berlekamp in [16].

$$\begin{array}{c|cccccc}
8 & & & & & & 1 \\
7 & & & & & 1 & 1 \\
6 & & & & & 1 & 2 \\
5 & & & & 1 & 1 & 3 \\
4 & & & 1 & 1 & 2 & 7 \\
3 & & & 1 & 2 & 5 & 19 \\
2 & & 1 & 1 & 3 & 9 & 37 \\
1 & 1 & 1 & 2 & 7 & 23 & 99 \\
0 & 1 & 2 & 5 & 19 & 66 & 293 \\
\hline
 & 0 & 1 & 2 & 3 & 4 & 5
\end{array} \tag{1}$$

Berlekamp defines an array to be unitary if any square submatrix whose upper left corner falls on the boundary of the array has a determinant equal to 1. For instance, in the array above

$$\det \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 5 \\ 1 & 3 & 9 \end{pmatrix} = 1$$

The problem then he states as follows: "... A periodic quasilinear boundary represents the best staircase approximation to a straight line of rational slope. ... Exact formulas are known for the values of the numbers in the unitary arrays generated by periodic quasilinear boundaries of slopes $1/n$ or n, but no such formulas are known (to me) for the values in the arrays with boundaries of slopes m/n where $1 < m < n$. The simplest such case is slope 2/3" – this is shown above in (1).

This problem arose already in Berlekamp's paper [4], where the numbers in the array above reduced modulo 2 were suggested as a convolution code As Berlekamp pointed out [6], the density of ones in these codes is very low, which was regarded as a disadvantage that time, "...But in the past decades we have seen great popularity of parity-check codes and of turbo codes, both of which are of low density (and hard to design in any highly structured way, compared with, say, RS block codes). The key is to find a criteria other than constraint length. Anyway, I think the topic of structured generator sequences for convolutional codes merits attention again, although my '63 paper seems to be perhaps the only one that attempts to initiate work in that area...".

The case slope 2/3 and 3/2, which yield codes of rates 2/5 and 3/5, respectively, was studied in further detail later on by Berlekamp in [5], where he derived some formulas for special parameters and then stated: "The patterns are clear but I know no explanation. Why does the formula apply to an individual entry, then to sums of pairs of entries from different rows, and then to the *negative* of an entry?". This question had been answered in our papers [13] (without being aware of the reference [5] at that time) and [15]. In Section IV we shall give the generating function for the entries in the array (1). The methods using lattice path enumeration are presented in Section II and III. They apply to further lattice path models, which is the topic of actual research, as briefly mentioned in Section V.

2 Lattice path enumeration

Carlitz, Roselle, and Scoville [8] later presented a fast algorithm for the computation of the number of such lattice paths by getting rid of the determinant calculation. They showed that the entries in this array enumerate the lattice paths from the beginning of the row to the top of the column which determine the respective entry, where these paths are not allowed to cross the boundary given by the 1's. For instance, in the array (1) above the positions of the 1's are below the boundary determined by $u_0 = 2$, $u_1 = 3$, $u_2 = 5$, $u_3 = 6$, $u_4 = 8$, $u_5 = 9$, etc.

A path here is a sequence of pairs (s_i, t_i), $i = 0, 1, \dots$ of nonnegative integers where (s_i, t_i) is either $(s_{i-1} + 1, t_{i-1})$ or $(s_{i-1}, t_{i-1} + 1)$. So, a particle following such a path can move either one step to the right, i. e. $s_i = s_{i-1} + 1$, or one step upwards, i. e. $t_i = t_{i-1} + 1$ in each time unit i. We shall assume that a path starts in the origin $(0, 0)$. There is a one–to–one correspondence between a $\{0, 1\}$- sequence x^m and a path with m steps: a 0 in the sequence x^m corresponds to a step upwards, a 1 to a step to the right in the corresponding path. The (infinite) path determined by this boundary hence corresponds to the periodic, binary sequence

$$001010010100101001...$$

Observe that the positions of the 1's in this sequence are at $v_{i+1} = u_i + i$ for all $i = 0, 1, 2, \ldots$. This holds, because there is exactly one step to the right after each u_i steps upwards in the boundary lattice path.

The rows in the array (1) above behave periodically in the sense that every third row has the same entries, which are only shifted according to the boundary. Because of this fact only two further sequences have to be considered in order to analyze Berlekamp's problem for slope $\frac{2}{3}$, namely the sequences

$$01010010100101001... \quad \text{and} \quad 01001010010100101...,$$

since the paths corresponding to these sequences characterize all possible boundaries arising in the array (1).

In terms of these sequences, Berlekamp's problem was analyzed in [15] by studying the size of the downsets $N(y^m)$ of the initial segments y^m of these three sequences in the so called pushing order (cf. [3] and [11]), which played a central role in Ahlswede's and Khachatrian's solution of the Erdös–Ko–Rado Problem [1, 2].

3 Gessels's probabilistic method

We shall consider paths in an integer lattice from the origin $(0, 0)$ to the point (n, u_n), which never touch any of the points (i, u_i), $i = 0, 1, \ldots, n-1$. In [9] Gessel introduced a general probabilistic method to determine the number of such paths, denoted by f_n , which he studied for the case that the subsequence $(u_i)_{i=1,2\ldots}$ is periodic. For period length 2 the elements of the sequence $(u_i)_{m=0,1,2,\ldots}$ are on the 2 lines (for $i = 0, 1, 2, \ldots$)

$$u_{2i} = s + ci \text{ and } u_{2i+1} = s + \mu + ci, \tag{2}$$

Gessel's probabilistic method is as follows. A particle starts at the origin $(0, 0)$ and successively moves with probability p one unit to the right and with probability $q = 1 - p$ one unit up. The particle stops if it touches one of the points (i, u_i). The probability that the particle stops at (n, u_n) is $p^n q^{u_n} \cdot f_n$. Setting

$$f(t) = \sum_{n=0}^{\infty} f_n t^n = \sum_{n=0}^{\infty} f_{2n} t^{2n} + \sum_{n=0}^{\infty} f_{2n+1} t^{2n+1} = g(t^2) + t \cdot h(t^2)$$

the probability that the particle eventually stops is

$$q^{u_0} g(p^2 q^c) + p q^{u_1} h(p^2 q^c)$$

If p is sufficiently small, the particle will touch the boundary $(i, u_i)_{i=0,1,\cdots}$ with probability 1. So for small p and with $t = pq^{c/2}$ we have

$$q(t)^{u_0} g(t^2) + p(t) q(t)^{u_1} h(t^2) = 1$$

For p sufficiently small one may invert $t = p(1-p)^{c/2}$ to express p as a power series in t, namely $p = p(t)$. Then changing t to $-t$ and denoting $p(-t)$ by $\overline{p}(t)$ and similarly $q(-t)$ by $\overline{q}(t)$ yields the system of equations

$$q^s \cdot g(t^2) + p \cdot q^{s+\mu} \cdot h(t^2) = 1,$$

$$\overline{q}^s \cdot g(t^2) + \overline{p} \cdot \overline{q}^{s+\mu} \cdot h(t^2) = 1 \tag{3}$$

which for $g(t^2)$ and $h(t^2)$ yield the solutions

$$g(t^2) = \frac{p^{-1} q^{-s-\mu} - \overline{p}^{-1} \overline{q}^{-s-\mu}}{p^{-1} q^{-\mu} - \overline{p}^{-1} \overline{q}^{-\mu}} = \frac{q^{c/2-\mu-s} + \overline{q}^{c/2-\mu-s}}{q^{c/2-\mu} + \overline{q}^{c/2-\mu}} \tag{4}$$

and

$$h(t^2) = \frac{q^{-s} - \overline{q}^{-s}}{t \cdot (q^{\mu-c/2} + \overline{q}^{\mu-c/2})} \tag{5}$$

By Lagrange inversion (cf. e.g. [12]) for any α we have

$$q^{-\alpha} = \sum_{n=0}^{\infty} \frac{\alpha}{(c/2+1)n+\alpha} \binom{(c/2+1)n+\alpha}{n} \cdot t^n \tag{6}$$

The following identities were derived in [9] and [13]. Since we are going to look at several random walks in parallel, we shall write the parameters determining the restrictions as superscripts. So, $g^{(s,c,\mu)}$ and $h^{(s,c,\mu)}$ are the generating functions (4) and (5) for even and odd n, respectively, for the random walk of a particle starting at the origin and first touching the boundary $(i, u_i)_{i=0,1,\dots}$ determined by the parameters s, c, and μ as defined under (2) in the lattice point (n, u_n).

Theorem [9, 13] a) Let c be an odd positive integer, $s = 1$ and $\mu = \frac{c-1}{2}$. Then

$$h^{(1,c,\frac{c-1}{2})}(t^2) = \frac{q^{-1/2} - \overline{q}^{-1/2}}{t} = \sum_{n=0}^{\infty} \frac{1}{(c+2)n+\mu+2} \binom{(c+2)n+\mu+2}{2n+1} t^{2n}.$$

b) For $0 \le \mu < \frac{c}{2}$ it is

$$g^{(s,c,\mu)}(t^2) + g^{(s,c,c-\mu)}(t^2) = q^{-s} + \overline{q}^{-s} = \sum_{n=0}^{\infty} \frac{2s}{(c+2)n+s} \binom{(c+2)n+s}{2n} t^{2n}$$

and

$$g^{(s,c,c-\mu)}(t^2) - g^{(s,c,\mu)}(t^2) = t^2 \cdot h^{(s,c,\mu)}(t^2) \cdot h^{(c-2\mu,c,\mu)}(t^2).$$

c) Let $s+\mu=c$ with $s\geq\mu$, then

$$h^{(s,c,c-s)}(t^2)+h^{(c-s,c,s)}(t^2)=\frac{1}{t^2}\cdot(p+\overline{p})=\sum_{n=1}^{\infty}\frac{2}{(c+2)n-1}\binom{(c+2)n-1}{2n}\cdot t^{2(n-1)}$$

In the special case c odd, $s=\frac{c+1}{2}$ and $\mu=\frac{c-1}{2}$ we have

$$h^{(\frac{c+1}{2},c,\frac{c-1}{2})}(t^2)-h^{(\frac{c-1}{2},c,\frac{c+1}{2})}(t^2)=\left(g^{(\frac{c+1}{2},c,\frac{c-1}{2})}(t^2)\right)^2,$$

where

$$g^{(\frac{c+1}{2},c,\frac{c-1}{2})}(t^2)=\frac{1}{t}\cdot(\overline{q}^{\frac{1}{2}}-q^{\frac{1}{2}})=\sum_{n=0}^{\infty}\frac{1}{(c+2)n+\frac{c+1}{2}}\binom{(c+2)n+\frac{c+1}{2}}{2n+1}\cdot t^{2n}.$$

d)

$$\left(g^{(s,c,\mu)}(t^2)+g^{(s,c,c-\mu)}(t^2)\right)\cdot h^{(s,c,\mu)}(t^2)=h^{(2s,c,\mu)}(t^2).$$

e)

$$g^{(c-2\mu,c,\mu)}(t^2)\cdot g^{(\mu,c,c-\mu)}(t^2)=g^{(c-\mu,c,\mu)}(t^2).$$

f) For $s_1+\mu_1+\mu_2=c$ we have

$$g^{(s_1,c,\mu_1)}(t^2)\cdot h^{(s_2,c,\mu_2)}(t^2)=h^{(s_2,c,s_1+\mu_2)}(t^2).$$

Especially, for odd c

$$g^{(1,c,\frac{c-1}{2})}(t^2)\cdot h^{(1,c,\frac{c-1}{2})}(t^2)=h^{(1,c,\frac{c+1}{2})}(t^2).$$

4 Solution of Berlekamp's problem

Now we are able to explain the entries from Berlekamp's example array for slope $\frac{2}{3}$. We have to inspect the parameter choices $(s=1,\mu=1)$, $(s=1,\mu=2)$, and $(s=2,\mu=1)$. By application of the previous theorem, the generating functions for these parameters (after mapping $t^2\to x$) look as follows.

Corollary [15]:

$$g^{(1,3,1)}(x)=\sum_{n=0}^{\infty}\frac{1}{5n+1}\binom{5n+1}{2n}x^n-\frac{x}{2}\cdot[h^{(1,3,1)}(x)]^2=1+2x+23x^2+377x^3+\ldots,$$

$$g^{(1,3,2)}(x)=\sum_{n=0}^{\infty}\frac{1}{5n+1}\binom{5n+1}{2n}x^n+\frac{x}{2}\cdot[h^{(1,3,1)}(x)]^2=1+3x+37x^2+624x^3+\ldots,$$

$$g^{(2,3,1)}(x) = \sum_{n=0}^{\infty} \frac{1}{5n+2}\binom{5n+2}{2n+1}x^n = 1 + 5x + 66x^2 + 1156x^3 + \ldots,$$

$$h^{(1,3,1)}(x) = \sum_{n=0}^{\infty} \frac{1}{5n+3}\binom{5n+3}{2n+1}x^n = 1 + 7x + 99x^2 + 1768x^3 + \ldots,$$

$$h^{(1,3,2)}(x) = \sum_{n=1}^{\infty} \frac{1}{5n-1}\binom{5n-1}{2n}x^{n-1} - \frac{1}{2}[g^{(2,3,1)}(x)]^2 = 1 + 9x + 136x^2 + \ldots,$$

$$h^{(2,3,1)}(x) = \sum_{n=1}^{\infty} \frac{1}{5n-1}\binom{5n-1}{2n}x^{n-1} + \frac{1}{2}\cdot[g^{(2,3,1)}(x)]^2 = 2 + 19x + 293x^2 + 5332x^3 + \ldots.$$

Using the results in the above theorem, in [15] we also derived the generating functions for the array in Berlekamp's problem with slope $\frac{2}{5}$.

5 Concluding remarks

1) There is a one–to–one correspondence between s–ary regular trees and ballot–type $\{0,1\}$–sequences $x^{sn} = (x_1, \ldots, x_{sn})$ of weight (= number of 1's) $wt(x^{sn}) = n$ fulfilling the condition $wt(x_1, \ldots, x_i) \geq \frac{i}{s}$ for all $i = 1, \ldots, sn-1$. This correspondence can be exploited to store regular trees, by assigning to them as codewords the ballot – type sequence. The codes thus obtained form a prefix code, cf. [10].

2) Probably most interesting, and indeed the topic of actual research, is the fact that the formulae from the above theorem also arise in the enumeration of a different type of lattice paths. Here, the boundary not allowed to be crossed is obtained by repeatedly moving s steps upwards, and t steps to the right. This model was seemingly first studied in [7]. Again we analyzed the case $s = 2, t = 3$ and $s = 3, t = 2$. Here the formulae from the above theorem enumerate the number of paths to any point on the boundary. Interestingly, all the six formulae have a natural interpretation, whereas in the analysis of Berlekamp's array only four of them really come into play. However, for further periodic slopes, the analysis is more difficult.

References

[1] R. Ahlswede, L. H. Khachatrian, The complete nontrivial–intersection theorem for systems of finite sets, *J. Combin. Theory* Ser. A 76, 1996, 121-138.

[2] R. Ahlswede, L. H. Khachatrian, The complete intersection theorem for systems of finite sets, *Europ. J. Combin.* 18, 1997, 125-136.

[3] R. Ahlswede, Z. Zhang, On maximal shadows of members in left – compressed sets, *Discr. Appl. Math.* 95, 1999, 3-9.

[4] E. R. Berlekamp, A class of convolutional codes, *Inform. Control* 6, 1963, 1-13.

[5] E. R. Berlekamp, Unimodular arrays, *Comp. Math. Appl.* 39, 2000, 77-88.

[6] E. R. Berlekamp, private communication.

[7] R. J. Chapman, T. Y. Chow, A. Khetan, D. P. Moulton, R. J. Waters, Simple formulas for lattice paths avoiding certain periodic staircase boundaries, Preprint, 2007.

[8] L. Carlitz, D. P. Roselle, R. A. Scoville, Some remarks on ballot – type sequences, *J. Combin. Theory* 11, 1971, 258-271.

[9] I. Gessel, A probabilistic method for lattice path enumeration, *J. Stat. Plann. Infer.* 14, 1986, 49-58.

[10] K. Kobayashi, H. Morita, M. Hoshi, Enumerative coding for k–ary trees, *Proc. 19th Symp. Inform. Theory Appl. (SITA96)*, Hakone, Japan, 1996, 377-379.

[11] R. P. Stanley, Weyl groups, the hard Lefschetz theorem, and the Sperner property, *SIAM J. Algebr. Discr. Meth.*, 1980, 168-184.

[12] R.P. Stanley, *Enumerative Combinatorics 2*, Cambridge, 1999.

[13] U. Tamm, Lattice paths not touching a given boundary, *J. Stat. Plann. Infer.* 105, 2002, 433-448.

[14] U. Tamm, On a problem of Berlekamp, *Proc. Intern. Symp. Inform. Theory*, Yokohama, Japan, 2003, 41.

[15] U. Tamm, Size of downsets in the pushing order and a problem of Berlekamp, *Discr. Appl. Math.*, to appear.

[16] W. Tutte (ed.), *Recent Progress in Combinatorics*, Proc. 3rd Waterloo Conf. Combin., 1968.