Size of downsets in the pushing order and a problem of Berlekamp

Ulrich Tamm

Department of Computer Science, University of Chemnitz, 09107 Chemnitz, Germany

Received 8 September 2003; received in revised form 26 July 2005; accepted 2 November 2006
Available online 28 August 2007

Dedicated to the memory of Levon Khachatrian

Abstract

The shifting technique is a useful tool in extremal set theory. It was successfully used and developed by Levon Khachatrian to obtain many significant results. The shifting operation also referred to as pushing gives rise to a partial order called pushing order. Here we consider the problem of determination of the size of special downsets in this order. For the analysis, the pushing order will be expressed isomorphically in terms of lattice paths and of majorization of sequences. In the case that the sequences under consideration are periodic the generating function for the numbers arising in an old combinatorial problem due to Berlekamp will be determined.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Pushing order; Majorization; Lattice paths

1. The pushing order and majorization

We are going to consider \( \{0, 1\} \)-sequences \( x^m = (x_1, \ldots, x_m) \) of length \( m \) and weight \( wt(x^m) = k \), i.e., \( (x_1, \ldots, x_m) \) consists of exactly \( k \) ones and \( m - k \) zeros.

These can equivalently be regarded as \( k \)-element subsets of the \( m \)-element set \( \{1, \ldots, m\} \), namely via \( x_i = 1 \) exactly if \( i \) is contained in the subset corresponding to \( x^m \).

On the set \( \binom{[m]}{k} \) of all sequences of length \( m \) and constant weight \( k \) we define an order relation \( \leq_p \) in the following way. Let \( \{v_1, \ldots, v_k\} \) and \( \{v'_1, \ldots, v'_k\} \) denote the sets of positions (in ascending order) of the 1’s in the sequences \( x^m \) and \( y^m \), respectively. Then \( x^m \leq_p y^m \) exactly if \( v_r \leq v'_r \) for all \( r = 1, \ldots, k \). So for all \( r \) the \( r \)th 1 in \( x^m \) is not allowed to occur later than the \( r \)th 1 in \( y^m \). This can be interpreted in such a way that \( x^m \) can be obtained by “pushing” the 1’s in \( y^m \) to the left (if \( x^m \) and \( y^m \) are written as row vectors).

Following Ahlswede and Zhang [3], the problem we are going to address is to determine the size of special downsets or shadows in this order, namely we are interested in the number

\[
N(y^m) = \left| \left\{ x^m \in \binom{[m]}{k} : x^m \leq_p y^m \right\} \right|
\]

preceding a given element \( y^m \in \binom{[m]}{k} \).

E-mail address: tamm@informatik.tu-chemnitz.de.

0166-218X/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.dam.2006.11.022
The pushing order plays an important role in the analysis of intersection theorems in extremal combinatorics, e.g. [1,2,6]. They can equivalently be defined in terms of majorization. Recall that a sequence \( a^m = (a_1, \ldots, a_m) \) of nonnegative real numbers is majorized by \( b^m = (b_1, \ldots, b_m) \), denoted \( a^m \preceq b^m \) if

\[
\begin{align*}
(i) \quad & \sum_{j=1}^{m} a_j = \sum_{j=1}^{m} b_j \quad \text{and} \quad (ii) \quad \sum_{j=1}^{i} a_j \leq \sum_{j=1}^{i} b_j \quad \text{for all } i = 1, 2, \ldots, m - 1.
\end{align*}
\]

i.e., all the partial sums formed by the initial segments of \( a^m \) are less or equal to the corresponding partial sums of \( b^m \).

Obviously, via majorization a partial order is defined on all sequences of length \( m \) over the nonnegative reals. Majorization on sequences of nonnegative integers also referred to as domination is a useful concept in the analysis of lattice paths, cf. [10]. On the set of binary sequences of length \( m \) this just yields the order obtained by pushing 1’s to the right. In the sequel we are more interested in left-pushing, \( \preceq_p \) defined above. In this case, \( x^m \preceq_p y^m \) exactly if \( y^m \) is majorized by \( x^m \), since by condition (i) the sequences \( x^m \) and \( y^m \) must have the same weight and by (ii) every 1 in \( x^m \) must precede its counterpart in \( y^m \).

A further isomorphic poset was studied by Proctor [11] and Stanley [12,13].

Let us mention that a second isomorphic order in terms of majorization is obtained by adding a final 1 to each sequence \( x^m \) and \( y^m \), i.e., introducing \( x_{m+1} = y_{m+1} = 1 \) and interpreting the sequences

\[
\begin{align*}
x^m &= (0, \ldots, 0, 1, 0, \ldots, 0, 1, \ldots, 0, \ldots, 0, 1) \\
y^m &= (0, \ldots, 1, 0, \ldots, 0, 1, \ldots, 0, \ldots, 0, 1)
\end{align*}
\]

as two partitions \( m+1 = \lambda_0 + \lambda_1 + \cdots + \lambda_k \) of the integer \( m+1 \) into \( k+1 \) positive integers, where each summand \( \lambda_j \) or \( \mu_j \) is just defined by the number of \( \mu_{j-1} \) or \( \lambda_{j-1} \), consecutive 0’s preceding the \( j \)th 1 in \( x^m \) and \( y^m \), respectively. Then \( x^m \preceq_p y^m \) exactly if \( (\mu_0, \ldots, \mu_k) \) is predecessor of \( (\lambda_0, \ldots, \lambda_k) \) in the majorization order \( \preceq \) (where now the elements are sequences of length \( k \) of positive integers, for several properties of this order cf. also [14, pp. 288–289]).

In order to attack our enumeration problem, the sequences will be represented as a path in the lattice of pairs of integers. A path here is a sequence of pairs \( (s_i, t_i), \ i = 0, 1, \ldots \) 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 \((0,0)\).

The one-to-one correspondence between the \( \{0,1\} \)-sequence \( x^m \) and a path with \( m \) steps is obtained as follows: 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 fact that \( x^m \preceq_p y^m \) in the lattice model translates to the property that the path obtained from \( x^m \) never crosses the path obtained from \( y^m \). So, the path corresponding to \( y^m \) is always above any other path obtained from a sequence in the set \( \{x^m \preceq y^m : x^m \in \binom{[m]}{k} \} \).

The size of downsets or even general intervals in the pushing order can be determined recursively for the equivalent lattice path problem, cf. [10]. Ahlswede and Zhang [3] gave asymptotic results for general sequences \( y^m \). In the case that the sequences \( y^m \) are periodic, one might even obtain exact results. For special such periodic sequences we shall derive the generating function for the number of predecessors of \( y^m \) in the pushing order. We shall see that our approach from [15] allows to determine the numbers which arose in an old combinatorial problem due to Berlekamp [17].

2. A problem by Berlekamp and Gessel’s probabilistic method

At the 3rd Waterloo Conference on Combinatorics [17, pp. 341–342], Berlekamp presented the following combinatorial problem. The problem will be illustrated with the following example also due to Berlekamp in [17].
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

\[
\begin{array}{ccccccc}
8 & 1 \\
7 & 1 & 1 \\
6 & 1 & 2 \\
5 & 1 & 1 & 3 \\
4 & 1 & 1 & 2 & 7 \\
3 & 1 & 2 & 5 & 19 \\
2 & 1 & 1 & 3 & 9 & 37 \\
1 & 1 & 2 & 7 & 23 & 99 \\
0 & 1 & 2 & 5 & 19 & 66 & 293 \\
0 & 1 & 2 & 3 & 4 & 5 \\
\end{array}
\]

(1)

This array was presented in [17] by flipping rows and columns.

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 \(\frac{2}{3}\) —this is shown above in (1), actually this array was presented in [17] by flipping rows and columns.

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. Many more examples of such arrays are collected and further studied in [5].

Carlitz, Roselle, and Scoville [7] later presented a fast algorithm for the computation of the number of such lattice paths, cf. also [10], 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

\[
u_0 = 2, u_1 = 3, u_2 = 5, u_3 = 6, u_4 = 8, u_5 = 9, \ldots
\]

The (infinite) path determined by this boundary corresponds to the periodic, binary sequence

\[
001010010100101001 \ldots
\]

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

\[
01010010100101001 \ldots \quad \text{and} \quad 01001010010100101 \ldots,
\]

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

In terms of the pushing order, Berlekamp’s problem can be analyzed by studying the size of the downsets \(N(y^m)\) of the initial segments \(y^m\) of these three sequences.

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 [8] 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 \(i = 0, 1, 2, \ldots\))

\[
u_{2i} = s + ci \quad \text{and} \quad u_{2i+1} = s + \mu + ci,
\]

(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
\[
f(t) = \sum_{n=0}^{\infty} f_nt^n = \sum_{n=0}^{\infty} f_{2n}t^{2n} + \sum_{n=0}^{\infty} f_{2n+1}t^{2n+1} = g(t^2) + t \cdot h(t^2)
\]
the probability that the particle eventually stops is
\[
q^{u_0} g(p^2q^c) + pq^{u_1} h(p^2q^c).
\]

If \(p\) is sufficiently small, the particle will touch the boundary \((i, u_i)\) 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
\[
g(t^2) = \frac{p^{-1}q^{-s-\mu} - \overline{p}^{-1}\overline{q}^{-s-\mu}}{p^{-1}q^{-\mu} - \overline{p}^{-1}\overline{q}^{-\mu}} = \frac{q^{c/2-\mu-s} + \overline{q}^{c/2-\mu-s}}{q^{c/2-\mu} + \overline{q}^{c/2-\mu}}
\]
and
\[
h(t^2) = \frac{q^{-s} - \overline{q}^{-s}}{t \cdot (q^{\mu-c/2} + \overline{q}^{\mu-c/2})}.
\]

By Lagrange inversion (cf. e.g., [14]) for any \(x\) we have
\[
q^{-x} = \sum_{n=0}^{\infty} \frac{x}{(c/2 + 1)n + x} \left( \frac{c/2 + 1 + 2n}{n} \right) t^n.
\]

The following identities were derived in [8,15]. 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 \((i, u_i)\) determined by the parameters \(s, c,\) and \(\mu\) as defined under (2) in the lattice point \((n, u_n)\).

**Theorem** (Gessel [8], Tamm [15]). (a) Let \(c\) be an odd positive integer, \(s = 1\) and \(\mu = \frac{c-1}{2}\). Then
\[
h^{(1,c-(c-1)/2)}(t^2) = q^{-1/2} - \overline{q}^{-1/2} = \sum_{n=0}^{\infty} \frac{1}{(c + 2)n + \mu + 2} \left( \frac{c + 2n + \mu + 2}{2n + 1} \right) t^{2n}.
\]

(b) For \(0 \leq \mu < (c/2)\) it is
\[
g^{(s,c,\mu)}(t^2) + g^{(s,c,-\mu)}(t^2) = q^{-s} + \overline{q}^{-s} = \sum_{n=0}^{\infty} \frac{2s}{(c + 2)n + s} \left( \frac{c + 2n + s}{2n} \right) t^{2n}
\]
and
\[
g^{(s,c,\mu)}(t^2) - g^{(s,c,\mu)}(t^2) = t^2 \cdot h^{(s,c,\mu)}(t^2) \cdot h^{(c-2\mu,c,\mu)}(t^2).
\]
(c) Let \( s + \mu = c \) with \( s \geq \mu \), then
\[
\begin{align*}
  h^{(s,c,c-s)}(t^2) + h^{(c-s,c,s)}(t^2) &= \frac{1}{t^2} \cdot (p + \overline{p}) \\
  &= \sum_{n=1}^{\infty} \frac{2}{(c+2)n - 1} \begin{pmatrix} (c+2)n - 1 \\ 2n \end{pmatrix} \cdot t^{2(n-1)}.
\end{align*}
\]

In the special case \( c \) odd, \( s = (c+1)/2 \) and \( \mu = (c-1)/2 \) we have
\[
\begin{align*}
  h^{((c+1)/2,c,(c-1)/2)}(t^2) - h^{((c-1)/2,c,(c+1)/2)}(t^2) &= (g^{((c+1)/2,c,(c-1)/2)}(t^2))^2,
\end{align*}
\]
where
\[
\begin{align*}
  g^{((c+1)/2,c,(c-1)/2)}(t^2) &= \frac{1}{t} \cdot (q^{1/2} - q^{1/2}) = \sum_{n=0}^{\infty} \frac{1}{(c + 2)n + (c + 1)/2} \begin{pmatrix} (c+2)n + \frac{c+1}{2} \\ 2n + 1 \end{pmatrix} \cdot t^{2n}.
\end{align*}
\]

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

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

(f) For \( s_1 + \mu_1 + \mu_2 = c \) we have
\[
\begin{align*}
  g^{(s_1,\mu_1)}(t^2) \cdot h^{(s_2,c,\mu_2)}(t^2) &= h^{(s_2,c,s_1+\mu_2)}(t^2).
\end{align*}
\]

Especially, for odd \( c \)
\[
\begin{align*}
  g^{(1,c,(c-1)/2)}(t^2) \cdot h^{(1,c,(c-1)/2)}(t^2) &= h^{(1,c,(c+1)/2)}(t^2).
\end{align*}
\]

3. Analysis for slopes \( \frac{2}{3} \) and \( \frac{2}{5} \)

For the analysis of Berlekamp’s problem, in case the entries are below a line of slope \( 2/c \), with \( c \geq 3 \) being an odd, positive integer, we have to determine the \( 2c \) generating functions \( g^{((c+1)/2,c,(c-1)/2)}(t^2), h^{((c+1)/2,c,(c-1)/2)}(t^2), \) and \( g^{(s,c,(c-1)/2)}(t^2), h^{(s,c,(c-1)/2)}(t^2) \) for \( s = 1, \ldots, (c-1)/2 \).

Let us first consider 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 1 (Tamm [15]).**

\[
\begin{align*}
  g^{(1,3,1)}(x) &= \sum_{n=0}^{\infty} \frac{1}{5n+1} \begin{pmatrix} 5n + 1 \\ 2n \end{pmatrix} x^n - \frac{x}{2} \cdot [h^{(1,3,1)}(x)]^2 \\
  &= 1 + 2x + 23x^2 + 377x^3 + \ldots,
\end{align*}
\]

\[
\begin{align*}
  g^{(1,3,2)}(x) &= \sum_{n=0}^{\infty} \frac{1}{5n+1} \begin{pmatrix} 5n + 1 \\ 2n \end{pmatrix} x^n + \frac{x}{2} \cdot [h^{(1,3,1)}(x)]^2 \\
  &= 1 + 3x + 37x^2 + 624x^3 + \ldots,
\end{align*}
\]

\[
\begin{align*}
  g^{(2,3,1)}(x) &= \sum_{n=0}^{\infty} \frac{1}{5n+2} \begin{pmatrix} 5n + 2 \\ 2n + 1 \end{pmatrix} x^n \\
  &= 1 + 5x + 66x^2 + 1156x^3 + \ldots,
\end{align*}
\]
Using the results in the above theorem it is also possible to derive the following identities, which just give the generating functions for the array in Berlekamp’s problem with slope \( \frac{2}{5} \).

**Corollary 2.**

\[
\begin{align*}
  g^{(1,3,1)}(x) &= \sum_{n=0}^{\infty} \frac{1}{5n+3} \left( \frac{5n+3}{2n+1} \right) x^n \\
  &= 1 + 7x + 99x^2 + 1768x^3 + \ldots, \\
  h^{(1,3,2)}(x) &= \sum_{n=1}^{\infty} \frac{1}{5n-1} \left( \frac{5n-1}{2n} \right) x^{n-1} - \frac{1}{2} [g^{(2,3,1)}(x)]^2 \\
  &= 1 + 9x + 136x^2 + \ldots, \\
  h^{(2,3,1)}(x) &= \sum_{n=1}^{\infty} \frac{1}{5n-1} \left( \frac{5n-1}{2n} \right) x^{n-1} + \frac{1}{2} [g^{(2,3,1)}(x)]^2 \\
  &= 2 + 19x + 293x^2 + 5332x^3 + \ldots.
\end{align*}
\]

Remarks. (1) Observe that a nice closed expression for the entries in Berlekamp’s array (1) only holds for the two generating functions \( g^{(2,3,1)} \) and \( h^{(1,3,1)} \). The theorem can obviously be also applied to analyze further arrays with a
periodic boundary of period length 2. For instance, the identities in Corollary 2 are the entries in the two-dimensional array from Berlekamp’s problem when the boundary is determined by a line of slope $\frac{2}{5}$. For slopes $2/c$, $c \geq 7$ odd integer, further identities similar to those in the theorem have to be derived.

(2) Computer observations strongly suggest that identities similar to those in Corollaries 1 and 2 exist for arrays with a periodic boundary of period length $d > 2$. However, in order to derive such identities, one has to find the solution of a more complex system of equations than (3).

(3) Further results related to a different combinatorial approach due to Carlitz, Roselle, and Scoville [7] in order to attack Berlekamp’s problem are presented in [15].

(4) 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 $wt(x_1, \ldots, x_i) \geq i/s$ for all $i = 1, \ldots, sn − 1$. All such sequences are obtained from the sequence

$$y^{sn} = (0, \ldots, 0, 1, 0, \ldots, 0, 1, \ldots, 0, \ldots, 0, 1)$$

by left-pushing of 1’s. 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. [9].

(5) After having presented my results at the 2003 IEEE Symposium on Information Theory [16], Berlekamp pointed out that he had published a further paper [5] on the subject. On pp. 86—87 of [5] several identities concerning sums of entries of the array (1) are presented followed by the remark “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 Theorem 2 and Proposition 6 of [15] (without being aware of the reference [5] at that time). The reason is that the generating functions $g(1,3,1)(x)$ and $g(1,3,2)(x)$ in Corollary 1 sum up to $2/(5n + 1) \left(\frac{5n+1}{2n}\right)x^n$. For the details see [15].

References