

Counting words with vector spaces

Carlos Segovia *

December 17, 2013

Abstract

The sequence 2, 5, 15, 51, 187, . . . with the form $(2^n + 1)(2^{n-1} + 1)/3$ has two interpretations in terms of the dimension of the universal embedding of the symplectic polar space and the density of a language with four letters. This article presents a way to relate this two approaches.

Introduction

It is interesting that the sequence 2, 5, 15, 51, 187, 715, . . . which writes as $g(n) := (2^n + 1)(2^{n-1} + 1)/3$ follows different approaches. To my knowledge, see the page [5] sequence A007581, this number represents:

1. the dimension of the universal embedding of the symplectic polar space, denoted by udim , see [1]; or
2. the number of isomorphic classes of regular four folding coverings of a graph with respect to the identity automorphism, see [2]; or
3. the density of a language with four letters, see [4]; or
4. the rank of the \mathbb{Z}_2^n -cobordism category in dimension $1 + 1$, see [6].

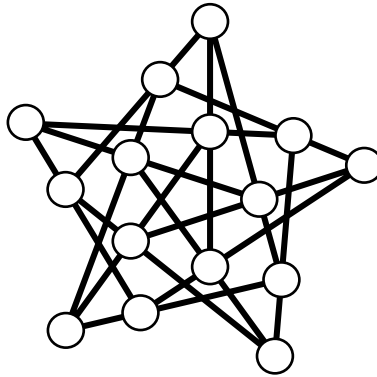
The approach (1) was called the Brower's conjecture. First, A. E. Brower has shown that $\text{udim} \geq g(n) = (2^n + 1)(2^{n-1} + 1)/3$. Later, P. Lee in [3], introduce a set \mathcal{N}^n and he prove that $|\mathcal{N}^n| = g(n)$. Subsequently, he shows that $\text{udim} \leq |\mathcal{N}^n|$ which gives the Brower's conjecture. The proof of the last inequality uses a stratification of the set \mathcal{N}^n in 7 cases. We use this idea in this article in order to give a new proof of the Brower's ex-conjecture using the approach (3).

The author proves in [7], the equivalence between (3) and (4). This article provides a form to relate (1) and (3). It rest to give the relation with approach (2). The study of this problem comes from personal interests of the author, in order to find the graphs associated to the universal embedding of the symplectic polar space, for example for $n = 2$ this gives the *Cremona-Richmond configuration* of figure 1.1.

*Mathematisches Institut, Universität Heidelberg, Deutschland. csegovia@mathi.uni-heidelberg.de

1 The binary dual polar space

The dimension of the universal embedding of the symplectic polar space takes into account a \mathbb{Z}_2 -vector space of dimension $2n$ with a symplectic form ω . Consider the geometry with lines of three elements defined as follows. The points are the maximal totally isotropic subspaces of dimension n , i.e. $\omega(V) = 0$ for V a subspace. The lines are given by the totally isotropic subspaces of dimension $n - 1$. Denote X and \mathcal{L} the sets of points and lines respectively. We consider the linear map $\sigma : \mathbb{Z}_2\mathcal{L} \rightarrow \mathbb{Z}_2X$ sending each line to the sum of its three elements. The dimension of the universal embedding of the symplectic polar space is the dimension of the module $\mathbb{Z}_2X/\sigma(\mathbb{Z}_2\mathcal{L})$. For example for $n = 1$ we have $X = \{(0, 1), (1, 0), (1, 1)\}$ with only one line. For $n = 2$ the geometry gives the Cremona-Richmond configuration as follows



(1.1)

A	0001	B	0001	C	0001	D	0010	E	0010
	0010		1000		1010		0100		0101
F	0100	G	0100	H	0101	I	0101	J	0110
	1000		1010		1000		1010		1001
K	0011	L	0011	M	0110	N	0111	O	0111
	1100		1101		1011		1011		1001

You can verify that the dimension of the universal embedding of the symplectic polar space with $n = 2$ is 5. For this find 5 points and fill the circles of the last figure where the lines are given by $ABC, AKL, DAE, DGF, EIH, JDM, EON, BHF, JBO, CGI, CMN, FNK, MHL, GOL$ and JIK . You need 5 points to realize yourself how to fill the crossword.

Now we resume some work of P. Li from [3]. Let n be a fixed integer with $n \geq 3$ and let Γ be the graph associated to the geometry of points of lines (X, \mathcal{L}) . Fix a point $x_0 \in X$, and let Γ_k ($0 \leq k \leq n$) denote the set of points at distance k from x_0 . Then $y \in \Gamma_k$ if and only if $\dim(y \cap x_0) = n - k$. We have that every line in \mathcal{L} contains two elements from Γ_k and one from Γ_{k-1} , for some $1 \leq k \leq n$. Moreover, for two points $p, q \in \Gamma_k$, p and q lie in the same connected component of Γ_k if and only if $p \cap x_0 = q \cap x_0$. Thus the connected components of Γ_k are in one-to-one correspondence with the $n - k$ -subspaces of x_0 . We write x_1, \dots, x_n

for the standart basis of \mathbb{Z}_2^n . For any vector $v = a_1x_1 + \dots + a_nx_n \in \mathbb{Z}_2^n$, define its support $\text{supp}(v) = \{i : a_i \neq 0\}$ and its weight $\text{wt}(v) = |\text{supp}(v)|$ and for any nonzero vector $v \in \mathbb{Z}_2^n$, set $\alpha(v) = \min \text{supp}(v)$ and $\beta(v) = \max \text{supp}(v)$. We define a total ordering on the vectors of \mathbb{Z}_2^n as follows: $a_1x_1 + \dots + a_nx_n \succ b_1x_1 + \dots + b_nx_n$ if there is some $i \in \{1, \dots, n\}$ such that $a_j = b_j$ for all $j < i$ and $(a_i, b_i) = (1, 0)$. For counting the subspaces P. Li introduce the set \mathcal{N}^n given by the collection of all subspaces of \mathbb{Z}_2^n whose reduced echelon basis $v_1 \succ \dots \succ v_k$ (where k the dimension of the subspace) satisfies all of the following conditions:

- (N1) $\text{wt}(v_i) \leq 2$ for every $i \in \{1, \dots, k\}$;
- (N2) if $v_i \succ v_j$ (i.e., $i < j$) and $\text{wt}(v_i) = \text{wt}(v_j) = 2$, then $\beta(v_i) \leq \beta(v_j)$;
- (N3) if $v_i \succ v_j \succ v_k$, $\text{wt}(v_i) = \text{wt}(v_j) = \text{wt}(v_k) = 2$, and $\beta(v_i) = \beta(v_j) < \beta(v_k)$, then $\alpha(v_k) > \beta(v_i)$;
- (N4) there do not exist $v_i \succ v_j \succ v_k \succ v_l$ such that $\text{wt}(v_i) = \text{wt}(v_j) = \text{wt}(v_k) = \text{wt}(v_l) = 2$ and $\beta(v_i) = \beta(v_j) = \beta(v_k) < \beta(v_l)$.

The importance of the set \mathcal{N}^n is the following, see [3] for the proof.

Theorem 1.1. *The dimension of universal embedding of the symplectic polar space does not exceed the cardinality of \mathcal{N}^n .*

The proof of the Brower's conjecture is done with the following result (we give a new proof in section 3).

Proposition 1.2. *We have the identity $|\mathcal{N}^n| = g(n) = (2^n + 1)(2^{n-1} + 1)/3$.*

2 The density of a language with four letters

The density of a language with four letters is defined as follows: take the number of words of length n made with letters 1,2,3,4 with the property that numbered from left to right each letter satisfies $0 < a_i \leq \max_{j \leq i} \{a_j\} + 1$. Thus we can dismiss the first letter which is always 1. For example, for $n = 2$ there are two words 1 and 2, for $n = 3$ the words are 11, 12, 21, 22, 23, while for $n = 4$ we have 15 words

111	112	121	122	123
211	212	213	221	222
223	231	232	233	234.

We can construct the next stage $n = 4$ by considering 7 cases as follows:

Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
1111	111 2	211 2	21 22	21 32	234 2	11 22
1121	11 23	231 2	23 22	23 32	234 3	12 33
1211	121 3	121 2	12 22	12 32		21 33
1221	12 23	221 2	22 22	22 32		22 33
1231	12 34	231 3	23 23	23 33		
2111	211 3					
2121	21 23					
2131	21 34					
2211	221 3					
2221	22 23					
2231	22 34					
2311	231 4					
2321	23 24					
2331	23 34					
2341	23 44					

Let L_n denote the set of words of length n with the hypothesis defined before. We have operations

$$E_i : L_n \longrightarrow L_{n-1},$$

which are given by erasing the i -letter from left to right. We can define the cases recursively starting with the following values

Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
111	11 2	21 2	2 22	2 32	\emptyset	1 22
121	12 3					2 33
211	21 3					
221	22 3					
231	23 4					

Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
1 1	1 2	\emptyset	\emptyset	\emptyset	\emptyset	2 2
2 1	2 3					

We define by $L_n(i)$ the set of words of length n of the case i . For a word $a = a_1 a_2 \cdots a_{n-1} a_n$ the cases are totally characterized by the following description:

Case 1. $a_n = 1$;

Case 2. $a_n = \max_{j < n} (a_j) + 1$ or $a_n = 4$;

Case 3. $a_{n-1} = 1$ and $E_{n-1}(a) \notin L_{n-1}(1), L_{n-1}(2)$;

Case 4. $a_{n-1} = 2$ and $E_{n-1}(a) \notin L_{n-1}(1), L_{n-1}(2)$;

Case 5. $a_{n-1} = 3$ and $E_{n-1}(a) \notin L_{n-1}(1), L_{n-1}(2)$;

Case 6. $a_{n-1} = 4$ and $E_{n-1}(a) \notin L_{n-1}(1), L_{n-1}(2)$; and

Case 7. $a \notin L_n(1), L_n(2), L_n(3), L_n(4), L_n(5), L_n(6)$.

These descriptions imply automatically the following two results.

Proposition 2.1. $|L_n(1)| = g(n-1)$ and $|L_n(2)| = g(n-1)$

Proposition 2.2. $|L_n(3)| = g(n-1) - 2g(n-2)$, $|L_n(4)| = g(n-1) - 2g(n-2)$,
 $|L_n(5)| = g(n-1) - 2g(n-2)$

We give a more detail treatment for the cases 6 and 7.

Proposition 2.3. $|L_n(6)| + |L_n(7)| = |L_{n-1}(3)| + |L_{n-1}(4)| + |L_{n-1}(5)| + |L_{n-1}(6)| + |L_{n-1}(7)| + 1$

Proof. We fix an integer $n > 3$. For $a \in L_n(6)$ by definition $E_{n-1}(a) \notin L_{n-1}(1), L_{n-1}(2)$. Thus a has at least a letter 3 between the position 1 to $n-2$. Therefore, by the assignment $E_{n-1}(a)$ we recover all the elements of $L_{n-1}(5) \cup L_{n-1}(6) \cup L_{n-1}(3) \cup L_{n-1}(4) \cup L_{n-1}(7)$ which has a letter 3 in the position from 1 to $n-2$. Now we take $a \in L_n(7)$. Since $a \notin L_n(1), L_n(2)$, then the last letter of a is not 1 or 4, and consequently $E_{n-1}(a) \in L_{n-1}(2)$. As a consequence, if the last letter of a is 2, then there are only letters 1 in the position 1 to $n-2$. Moreover, the position $n-1$ is a 2 since $a \notin L_n(2)$. Thus the element is of the form 11..122. Now, we take that the last letter of a is 3, since $E_{n-1}(a) \in L_{n-1}(2)$, then there are only 1 or 2 between the positions 1 to $n-2$ and the position $n-1$ is a 3 since $a \notin L_n(2)$. Thus with the assignment $E_{n-1}E_n(a)2$ we recover every word in $L_{n-1}(5) \cup L_{n-1}(6) \cup L_{n-1}(3) \cup L_{n-1}(4) \cup L_{n-1}(7)$ which does not have a letter 3 in the position from 1 to $n-2$. The sum of all the elements ends the proof of this proposition. \square

Consequently, by induction we can conclude the following result.

Theorem 2.4. *There is the identity*

$$|L_n| = 2g(n-1) + 4(g(n-1) - 2g(n-2)) + 1 = g(n). \quad (2.1)$$

3 Main constructions

In this section we define a bijection between the set \mathcal{N}^n from section 1 and the set L_n from section 2. We write x_1, \dots, x_n for the standard basis of \mathbb{Z}_2^n . For any vector $v = a_1x_1 + \dots + a_nx_n \in \mathbb{Z}_2^n$, we recall that its support is $\text{supp}(v) = \{i : a_i \neq 0\}$ and its weight $\text{wt}(v) = |\text{supp}(v)|$ and for any nonzero vector $v \in \mathbb{Z}_2^n$, $\alpha(v) = \min \text{supp}(v)$ and $\beta(v) = \max \text{supp}(v)$. For any subspace $V \leq \mathbb{Z}_2^n$, we define $\text{supp}(V) = \bigcup_{v \in V} \text{supp}(v)$. We can stratified the element of \mathcal{N}^n in 7 cases which we described below. In addition, for each case we give an example of the inductive step.

Case 1. $n \notin \text{supp}(V)$.

$$\begin{array}{cccccc} 1 & 0 & 0 & 1 & 0 & 0 & & 1 & 0 & 0 & 1 & 0 & \emptyset & & 1 & 0 & 0 & 1 & 0 \\ & 1 & 0 & 0 & 1 & 0 & \longrightarrow & 1 & 0 & 0 & 1 & \emptyset & = & & 1 & 0 & 0 & 1 \\ & & 1 & 0 & 1 & 0 & & & 1 & 0 & 1 & \emptyset & & & & 1 & 0 & 1 \end{array}$$

(b) $\text{wt}(v_j) = 2$ and $\beta(v_j) = t$ for t unique j .

$$\begin{array}{cccccccc} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ & 1 & 0 & 0 & \mathbf{1} & 0 & 0 & & 1 & 0 & 0 & 0 & \mathbf{1} \\ & & 1 & 0 & 0 & 0 & & & & 1 & 0 & 0 & \\ & & & & \mathbf{1} & 1 & & & & & \mathbf{1} & 0 & \end{array} \longrightarrow$$

(c) $\text{wt}(v_s) = 2$ and $\beta(v_s) = t$ for (exactly) two different values of s .

$$\begin{array}{cccccccc} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ & 1 & 0 & \mathbf{1} & 0 & 0 & 0 & & 1 & 0 & 0 & 0 & \mathbf{1} \\ & & 1 & 0 & 0 & 0 & 0 & & & 1 & 0 & 0 & 0 \\ & & & & 1 & 1 & & & & & & & \end{array} \longrightarrow$$

We note that we have described some operations, which correspond exactly to the erase operations E_i for words introduced in section 2. The bijection from L_n to \mathcal{N}^n is defined only in the initial values. The use of the operations E_i and the ones, for vector spaces exemplified before, constructs inductively the bijection. For $n = 1$ the assignments are

$$\begin{array}{l} 1 \mapsto 0 \\ 2 \mapsto 1 \end{array}$$

and for $n = 2$, they are

$$\begin{array}{l} 11 \mapsto 0 \ 0 \\ 12 \mapsto 0 \ 1 \\ 21 \mapsto 1 \ 0 \\ 22 \mapsto 1 \ 1 \\ 23 \mapsto \begin{array}{l} 1 \ 1 \\ 0 \ 1 \end{array} \end{array}$$

Finally, for $n = 3$ we divide the assignments for each case

$$\begin{array}{l} \text{Case 1} \\ 111 \mapsto 0 \ 0 \ 0 \mapsto 00 \\ 121 \mapsto 0 \ 1 \ 0 \mapsto 10 \\ 211 \mapsto 1 \ 0 \ 0 \mapsto 01 \\ 221 \mapsto 1 \ 1 \ 0 \mapsto 11 \\ \\ 231 \mapsto \begin{array}{l} 1 \ 1 \ 0 \\ 0 \ 1 \ 0 \end{array} \mapsto \begin{array}{l} 1 \ 1 \\ 0 \ 1 \end{array} \end{array}$$

$$\begin{array}{lclcl}
& \text{Case 2} & & & \\
112 & \mapsto & 0 & 0 & 1 & \mapsto & 00 \\
123 & \mapsto & 0 & 1 & 1 & \mapsto & 01 \\
& & & & 1 & & \\
213 & \mapsto & 1 & 0 & 1 & \mapsto & 10 \\
& & & & 1 & & \\
223 & \mapsto & 1 & 1 & 1 & \leftrightarrow & 1 & 1 & 0 & \mapsto & 11 \\
& & & & 1 & & 1 & & 1 & & \\
234 & \mapsto & 1 & 1 & 0 & \mapsto & 1 & 1 \\
& & 0 & 1 & 0 & & 0 & 1 \\
& & & & 1 & & &
\end{array}
,$$

where we note that the case $\begin{smallmatrix} 1 & 1 & 1 \\ & & 1 \end{smallmatrix}$ has to be chance to $\begin{smallmatrix} 1 & 1 & 0 \\ & & 1 \end{smallmatrix}$ in order to have a subspace inside \mathcal{N}^n .

$$\begin{array}{lclcl}
\text{Case 3} & 212 & \mapsto & 1 & \emptyset & 1 & \mapsto & 11 \\
\text{Case 4} & 222 & \mapsto & 1 & \begin{array}{c} \emptyset \\ | \\ \hline 1 \end{array} & 1 & \mapsto & 11 \\
& & & & \hline & \emptyset & & \\
\text{Case 5} & 232 & \mapsto & 1 & \begin{array}{c} \emptyset \\ | \\ \hline 1 \end{array} & 1 & \mapsto & 11 \\
& & & & \hline & \emptyset & & \\
\text{Case 6} & \emptyset & & & & & & \\
\text{Case 7} & 122 & \mapsto & 0 & 1 & 1 & \mapsto & 11 \\
233 & \mapsto & 1 & 1 & 1 & \leftrightarrow & 1 & 0 & 0 & \mapsto & 11 \\
& & & 1 & 1 & & 1 & 1 & & &
\end{array}
.$$

References

- [1] A. Blokhuis and A.E. Brouwer, *The Universal Embedding Dimension of the Binary Symplectic Dual Polar Space*, Discrete Mathematics, 2003, 264, 3-11.
- [2] Sungpyo Hong and Jin Ho Kwak, *Regular Fourfold Coverings with Respect to the Identity Automorphism*, Journal of Graph Theory, 1993, 17, 621-627.
- [3] Paul Li, *On the Universal Embedding of the $Sp_{2n}(2)$ Dual Polar Space*, Journal of Combinatorial Theory, Series A 94, 100-117 (2001).
- [4] Nelma Moreira and Rogério Reis, *On the Density of Languages Representing Finite Set Partitions*. Journal of Integer Sequences, 2005, 8, 1-11. 2nd Edition, 1994.
- [5] The On-Line Encyclopedia of Integer Sequences, <http://oeis.org>.
- [6] Carlos Segovia, *The classifying space of the 1+1 dimensional G -cobordism category*, <http://arxiv.org/abs/1211.2144>.
- [7] Carlos Segovia, *Numerical computations in cobordism categories*, <http://arxiv.org/abs/1307.2850>.