Available online at www.sciencedirect.com
ScienceDirect
European Journal of Combinatorics 29 (2008) 443–456
European Journal of Combinatorics
www.elsevier.com/locate/ejc

Counting vertices and cubes in median graphs of circular split systems

Y.B. Choe [a], K.T. Huber [b], J.H. Koolen [c], Y.S. Kwon [d], V. Moulton [b]

[a] *Combinatorial and Computational Mathematics Center, POSTECH, Pohang, South Korea*
[b] *School of Computing Sciences, University of East Anglia, Norwich, NR4 7TJ, United Kingdom*
[c] *Department of Mathematics, POSTECH, Pohang, South Korea*
[d] *Department of Mathematics, Yeungnam University, Gyeongsan, South Korea*

Received 23 December 2005; accepted 10 February 2007
Available online 3 March 2007

Abstract

Median graphs are a natural generalisation of trees and hypercubes that are closely related to distributive lattices and graph retracts. In the past decade, they have become of increasing interest to the biological community, where, amongst other things, they are applied to the study of evolutionary relationships within populations.

Two simple measures of complexity for a median graph are the number of vertices and the number of maximal induced subcubes. These numbers can be useful in biological applications, and they are also of purely mathematical interest. However, they can be hard to compute in general. Here we present some special families of median graphs where it is possible to find formulae and recursions for these numbers.

1. Introduction

Let $\mathbb{N}$ and $\mathbb{N}_0$ denote the set of positive and non-negative integers, respectively, and, for $n \in \mathbb{N}$, let $[n]$ denote the set $\{1, 2, \ldots, n\}$. We call a bipartition of $[n]$ a *split*, and in case S is a split consisting of parts A, B, we put $S = \{A, B\}$. Given a set of splits or *split system* $\mathcal{S}$ on $[n]$, the *median graph* associated to $\mathcal{S}$ (sometimes also called the *Buneman graph* [9]) is the subgraph of the $|\mathcal{S}|$-dimensional hypercube with vertex set consisting of those maps ϕ mapping $\mathcal{S}$ to the

E-mail addresses: ybchoe@postech.ac.kr (Y.B. Choe), katharina.huber@cmp.uea.ac.uk (K.T. Huber), koolen@postech.ac.kr (J.H. Koolen), ysookwon@ynu.ac.kr (Y.S. Kwon), vincent.moulton@cmp.uea.ac.uk (V. Moulton).

0195-6698/$ - see front matter
doi:10.1016/j.ejc.2007.02.003

power-set of $[n]$ which, for $S, S' \in \mathcal{S}$, satisfy

$$\phi(S) \in S \quad \text{and} \quad \phi(S) \cap \phi(S') \neq \emptyset,$$

and edge set consisting of those pairs of vertices $\{\phi, \phi'\}$ with

$$|\{S \in \mathcal{S} : \phi(S) \neq \phi'(S)\}| = 1.$$

Median graphs are a natural generalisation of trees and hypercubes. They are closely related to distributive lattices and graph retracts [1,2,14], and have been characterised in various ways (see e.g. [18] for a survey). Moreover, in the past decade, they have become of increasing interest to the biological community, where, amongst other things, they are applied to the study of evolutionary relationships within populations [4].

Two simple measures of the complexity of a median graph are the number of vertices and the number of maximal induced subcubes. These numbers can be useful to compute for biological applications [16], and they are also of purely mathematical interest — see e.g., [6,9,18] and [5], respectively. However, in general these numbers can be hard to compute — see e.g. [5,20] where the number of vertices of the median graph associated to the system of all possible splits of $[n]$ is presented for $n \leq 7$.

Here we will consider split systems where it is possible to provide formulae and recursions for the number of vertices and maximal subcubes in the associated median graph. In particular, for $n \geq 3$, let C_n be the n-cycle with vertex set $[n]$, and edge set consisting of those pairs $\{i, j\} \subsetneq [n]$ with $i - j \equiv \pm 1 \pmod{n}$. Then, for $n \geq 2$, define the *full circular split system* $\mathcal{S}(n)$ on $[n]$ by letting $\mathcal{S}(2)$ consist of the split $\{\{1\}, \{2\}\}$ and, for $n \geq 3$, letting $\mathcal{S}(n)$ consist of all splits of $[n]$ that are induced by removing two edges from C_n and taking the split of $[n]$ corresponding to the two connected components. In addition, for $1 \leq m \leq \lfloor n/2 \rfloor$, we define the split subsystem $\mathcal{S}(n, m)$ of $\mathcal{S}(n)$ to be the split system consisting of the splits $S = \{A, B\}$ in $\mathcal{S}(n)$ with $\min\{|A|, |B|\} = m$.

The split systems $\mathcal{S}(n)$ and $\mathcal{S}(n, m)$ are both special examples of *circular split systems* [3]. Besides median graphs, the closely related *split networks* are commonly associated to circular split systems in biological applications [7], and formulae have been derived for the number of vertices and edges of such networks [12]. In this paper, we present formulae and recursions for the number of vertices and maximal induced subcubes in the median graphs associated to $\mathcal{S}(n)$ and $\mathcal{S}(n, m)$.

We now present a brief summary of our main results. In Section 2, we show that the number of vertices in the median graph of $\mathcal{S}(n)$ equals 2^{n-1} (see Eq. (2)). In addition, we show for $1 \leq p \leq \lfloor \frac{n}{2} \rfloor$ that the number of maximal induced p-dimensional subcubes in this graph is

$$\frac{n}{n-2p} \sum_{j=0}^{p-1} 2^j \binom{p-1}{j} \binom{n-2p}{j+1}$$

in case $p \neq \frac{n}{2}$ and 1 in case $p = \frac{n}{2}$ (see Eq. (4)) from which the total number of maximal induced subcubes can easily be computed. In Section 3, we show that the number v_m^q of vertices and the number c_m^q of maximal induced subcubes in the median network associated to $\mathcal{S}(2m + q, m)$, $q \in \mathbb{N}_0$ and $m \geq 1$ is

$$v_m^q = 1 + \sum_{i \geq 0} \frac{2m+q}{2i+1} \binom{m-qi-1}{2i} 2^{m-2i-qi-1}$$

and

$$c_m^q = \sum_{i \geq 0} \frac{2m+q}{2i+1} \binom{m-qi-1}{2i},$$

respectively (see Eqs. (6) and (7)). Then, in Section 4, we provide recursive formulae for these numbers. In particular, we show that v_m^q can be computed recursively using

$$v_m^q = \begin{cases} 1 & \text{if } m = 0, \\ 1 + (2m+q)2^{m-1} & \text{for all } 1 \leq m \leq q+1, \\ 3v_{m-1}^q - \sum_{i=1}^{q} v_{m-i-1}^q & \text{for all } m > q+1 \end{cases} \tag{1}$$

(see Corollary 4.2), and the number c_m^q can be computed using

$$c_m^q = \begin{cases} 0 & \text{for } m = 0, \\ 2m+q & \text{for all } 1 \leq m \leq q+1, \\ 2c_{m-1}^q - c_{m-2}^q + c_{m-2-q}^q & \text{for all } m \geq q+2, \end{cases}$$

(Eq. (11)).

Intriguingly, in case $q = 1$, putting $w_m := v_m^1$, the recurrence relation in (1) becomes

$$w_m = 3w_{m-1} - w_{m-2}, \quad m \geq 2, \ (w_0 = 1 \text{ and } w_1 = 4).$$

In Section 5 we give a combinatorial proof for this formula, from which it will immediately follow that the sequence $\{w_m\}_{m \geq 0}$ is the bisection of the sequence of Lucas numbers $\{l_n\}_{n \geq 0}$. As an immediate consequence of this fact we obtain an affirmative answer to a question posed by B. Sturmfels concerning tight-spans.

2. The median graph associated to $\mathcal{S}(n)$

In this section we provide formulae for the number of vertices and maximal subcubes in the median graph associated to $\mathcal{S}(n)$, $n \geq 2$.

To do this, we will use the following key observations, that will be of use later on as well. Two splits $\{A, B\}$ and $\{C, D\}$ of $[n]$ are said to be *incompatible* if all of the intersections $A \cap C$, $A \cap D$, $B \cap C$, and $B \cap D$ are non-empty. In addition, a split system on $[n]$ is said to be incompatible if every pair of splits in this set is incompatible. For notational purposes, we also define split systems with cardinality 0 and 1 as being incompatible. From results presented in [9,17] it follows that if $\mathcal{S}$ is a split system on $[n]$, then

- the incompatible split subsystems of $\mathcal{S}$ are in one-to-one correspondence with the vertices of the median graph associated to $\mathcal{S}$.
- The maximal q-dimensional induced subcubes of the median graph associated to $\mathcal{S}$ are in one-to-one correspondence with the maximal incompatible subsystems in $\mathcal{S}$ with cardinality q, where $1 \leq q \leq |\mathcal{S}|$.

Therefore, in order to count vertices and maximal induced subcubes of the median graph associated to a given split system $\mathcal{S} \subseteq \mathcal{S}(n)$, we will count the incompatible and maximal incompatible subsystems of $\mathcal{S}$, respectively.

With these observations in hand it is now fairly routine to find a formula for the number of vertices in the median graph associated to $\mathcal{S}(n)$, or equivalently, for the number of incompatible

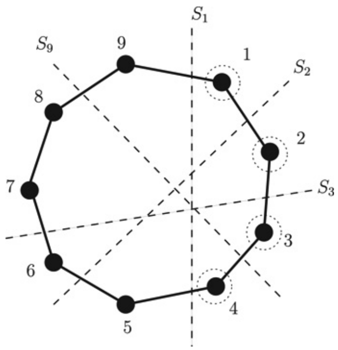

Fig. 1. An incompatible split subsystem of $\mathcal{S}(9)$ with cardinality 4, consisting of the splits S_1, S_2, S_3 and S_9. The cycle C_9 is depicted in bold lines, and the splits are depicted by dashed lines. The elements of the part A_1 of S_1, as defined in the next section, are highlighted by dashed circles.

split subsystems of $\mathcal{S}(n)$. Clearly, if $n = 2$ then $\mathcal{S}(n)$ contains two incompatible split subsystems, namely $\emptyset$ and $\mathcal{S}(2)$. So assume $n \geq 3$. By definition, each split in $\mathcal{S}(n)$ corresponds to choosing two distinct edges of the n-cycle C_n (which, by the way, implies that the cardinality of $\mathcal{S}(n)$ is $\binom{n}{2}$). Therefore, as can be easily verified, the incompatible split subsystems of $\mathcal{S}(n)$ with cardinality m, $1 \leq m \leq \lfloor \frac{n}{2} \rfloor$, are in one-to-one correspondence with the subsets of the edge set of C_n with cardinality $2m$. Hence, adding 1 to take into account the empty split subsystem, the number of incompatible split subsystems of $\mathcal{S}(n)$ equals

$$1 + \sum_{m=1}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2m} = 2^{n-1}, \tag{2}$$

as stated in the introduction.

Deriving a formula for the number of maximal incompatible split subsystems of $\mathcal{S}(n)$ is more complicated. Assume $n \geq 3$, noting that $\mathcal{S}(2)$ has one maximal incompatible split subsystem. For all $p \in \mathbb{N}$ with $p \leq \lfloor \frac{n}{2} \rfloor$, let $\mathcal{I}_{n,p}$ denote the set of maximal incompatible subsets of $\mathcal{S}(n)$ with cardinality p. For example, in Fig. 1 we depict a maximal incompatible split system in $\mathcal{I}_{9,4}$. Clearly, the number of maximal incompatible split subsystems of $\mathcal{S}(n)$ equals $\sum_{p \geq 1} |\mathcal{I}_{n,p}|$, and therefore from now on we will concentrate on finding a formula for $|\mathcal{I}_{n,p}|$.

We begin by translating the problem of finding the cardinality of $\mathcal{I}_{n,p}$ into that of finding the cardinality of another set defined as follows. For $i \in \mathbb{N}_0$, with $0 \leq i < p$, let $\mathcal{Z}^i_{n,p}$ denote the set of pairs $(\mathbf{z}, k)$ with $\mathbf{z}$ a string of length $2p$ (having terms $z_j \in \mathbb{N}$) and $k \in [n]$ that satisfy the following conditions:

(Z1) $z_1 = 1$ and $z_{p+1} > 1$;
(Z2) $\min(z_j, z_{j+p}) = 1$, for all $1 \leq j \leq p$;
(Z3) $|\{j \in [p] : z_j = 1 = z_{j+p}\}| = i$;
(Z4) $\sum_{j=1}^{2p} z_j = n$.

Proposition 2.1. *Suppose* $1 \leq p \leq \lfloor \frac{n}{2} \rfloor$. *If* $p = \frac{n}{2}$, *then* $|\mathcal{I}_{n,p}| = 1$, *else*

$$|\mathcal{I}_{n,p}| = \sum_{i=0}^{p-1} \frac{|\mathcal{Z}^i_{n,p}|}{p - i}.$$

Proof. Let p be as in the statement of the proposition. Clearly $|\mathcal{I}_{n,p}| = 1$ if $n = 2p$. So suppose from now on that $n > 2p$.

For $0 \le i \le p-1$, define a map γ that takes each element $(\mathbf{z}, k)$ in $\mathcal{Z}^i_{n,p}$ to the split system $\mathcal{S}(\mathbf{z}, k)$ of cardinality p consisting of the splits

$$S_j = \{X_j \dot\cup X_{j+1} \dot\cup \cdots \dot\cup X_{j+p-1}, X_{j+p} \dot\cup \cdots \dot\cup X_{j+2p-1}\}, \quad 1 \le j \le p,$$

(with addition of indices taken modulo $2p$) where, for $1 \le j \le 2p$,

$$X_j := \left\{ k + \sum_{\ell=1}^{j-1} z_\ell, k + \sum_{\ell=1}^{j-1} z_\ell + 1, \ldots, k + \sum_{\ell=1}^{j-1} z_\ell + z_j - 1 \right\},$$

(where addition is modulo n). By (Z4) it follows that $S_j \in \mathcal{S}(n)$, for all $1 \le j \le p$.

Now, observe that for all $1 \le j \le 2p$, by (Z2) and (Z1), we have $|X_j| = z_j \ge 1$. Hence, $X_j \ne \emptyset$, for all $1 \le j \le 2p$. It follows that the split system $\mathcal{S}(\mathbf{z}, k)$ is incompatible. Moreover, since a split $S \in \mathcal{S}(n) \setminus \mathcal{S}(\mathbf{z}, k)$ can only be incompatible with every element of $\mathcal{S}(\mathbf{z}, k)$ if there exists some $1 \le j \le 2p$ and non-empty disjoint subsets $A_k, B_k \in X_k$ with $A_k \dot\cup B_k = X_k$, $k \in \{j, j+p\}$, such that

$$S = \{A_j \dot\cup X_{j+1} \dot\cup \cdots \dot\cup X_{j+p-1} \dot\cup B_{j+p}, A_{j+p} \dot\cup X_{j+p} \dot\cup \cdots \dot\cup X_{j+2p-1} \dot\cup B_j\},$$

(Z2) implies that $\mathcal{S}(\mathbf{z}, k)$ is *maximal* incompatible.

In summary, the above observations imply that γ is a well-defined map from $\mathcal{Z}^i_{n,p}$ to $\mathcal{I}_{n,p}$. Moreover, by the characterisation of maximal incompatible split subsystems of $\mathcal{S}(n)$ given on [10, page 2], it follows that every element of $\mathcal{I}_{n,p}$ equals $\mathcal{S}(\mathbf{z}, k)$ for some $(\mathbf{z}, k) \in \mathcal{Z}^i_{n,p}$. Therefore, γ is a surjection. Thus, to complete the proof of the proposition it suffices to show that γ is $p - i$ to 1.

To see why this is the case, suppose $(\mathbf{z}, k) \in \mathcal{Z}^i_{n,p}$. Taking sums of string indices modulo $2p$ and using (Z3) and (Z2) let $z_{j_1} = z_1, z_{j_2}, \ldots, z_{j_{p-i}}$ denote the $p - i$ terms of $\mathbf{z}$ for which $z_{j_\ell} = 1 \ne z_{j_\ell + p}$. Then, for all $1 \le \ell \le p - i$, the pair $(\mathbf{z}_{j_\ell}, j_\ell)$ with $\mathbf{z}_{j_\ell} := z_{j_\ell} z_{j_\ell+1} \ldots z_{2p} z_1 \ldots z_{j_\ell - 1}$ is clearly contained in $\mathcal{Z}^i_{n,p}$. It is straightforward to check that $\mathcal{S}(\mathbf{z}, k) = \mathcal{S}(\mathbf{z}_{j_\ell}, k + z_1 + z_2 + \cdots + z_{j_\ell - 1})$, $\ell = 1, 2, \ldots, p - i$. Therefore γ is $p - i$ to 1, as required. ■

For $0 \le i \le p$, we now give a formula for the cardinality of $\mathcal{Z}^i_{n,p}$.

Lemma 2.2. *For* $0 \le i < p \le \lfloor \frac{n}{2} \rfloor$ *and* $p \ne \frac{n}{2}$

$$|\mathcal{Z}^i_{n,p}| = n \binom{p-1}{i} 2^{p-i-1} \binom{n-2p-1}{p-i-1}. \tag{3}$$

Proof. Suppose that $(\mathbf{z}, k)$ is some element of $\mathcal{Z}^i_{n,p}$. By definition k can clearly be chosen in n ways, and, for any fixed choice of k, by (Z3) there are $\binom{p-1}{i}$ ways to choose pairs of terms z_j, z_{j+p} with $z_j = z_{j+p} = 1$. Moreover, for any fixed choice of such a pair z_j, z_{j+p}, by (Z2) it follows that for all other pairs of terms $z_\ell, z_{\ell+p}$ with $\ell \in [p] \setminus \{j \in [p] : z_j = z_{j+p} = 1\}$ precisely one of z_ℓ and $z_{\ell+p}$ equals 1, and so there are 2^{p-i-1} ways in which to choose either z_ℓ or $z_{\ell+p}$ to equal 1.

Now, suppose that all of the above choices have been made. Then, $\mathbf{z}$ has $p - i$ terms z_j with $z_j \ge 2$ and the rest of the terms of $\mathbf{z}$ are all equal to 1. Therefore, by (Z4) the number of choices for the terms z_j of $\mathbf{z}$ with $z_j \geq 2$ equals the number of solutions to the equation $y_1 + y_2 + \cdots + y_{p-i} = n - p - i$, where $y_j \in \mathbb{N}$, $y_j \geq 2$. Using standard counting arguments, it is straightforward to see that there are $\binom{n-2p-1}{p-i-1}$ such solutions. The lemma now follows immediately. ■

As an immediate consequence of the last lemma and Proposition 2.1, for $1 \leq p \leq \lfloor \frac{n}{2} \rfloor$, $p \neq \frac{n}{2}$, by replacing the exponent $p - i - 1$ by j in (3), we obtain

$$|\mathcal{I}_{n,p}| = \frac{n}{n-2p} \sum_{j=0}^{p-1} 2^j \binom{p-1}{j} \binom{n-2p}{j+1}, \tag{4}$$

noting also that, in case $p = \frac{n}{2}$, $|\mathcal{I}_{n,p}| = 1$.

Note that we can use this expression to derive a recurrence relation for computing $|\mathcal{I}_{n,p}|$ as follows. For p and n as above, put $\iota_{n,p} := |\mathcal{I}_{n,p}|$, and let $b_{n,p} := \frac{n-2p}{n} \iota_{n,p}$. Then,

$$\begin{aligned}
b_{n,p} &= \sum_{j=0}^{p-1} \binom{p-1}{j} 2^j \binom{n-2p}{j+1} = \sum_{j=0}^{p-1} \left(\binom{p-2}{j} + \binom{p-2}{j-1} \right) 2^j \binom{n-2p}{j+1} \\
&= \sum_{j=0}^{p-2} \binom{p-2}{j} 2^j \binom{n-2-2(p-1)}{j+1} + 2 \sum_{j=0}^{p-2} \binom{p-2}{j} 2^j \binom{n-2p}{j+2} \\
&= b_{n-2,p-1} + 2 \sum_{j=0}^{p-2} \binom{p-2}{j} 2^j \left(\binom{n-2p-1}{j+1} + \binom{n-2p-1}{j+2} \right) \\
&= b_{n-2,p-1} + 2b_{n-3,p-1} + 2 \sum_{j=0}^{p-2} \binom{p-2}{j} 2^j \left(\binom{n-2p-2}{j+1} + \binom{n-2p-2}{j+2} \right) \\
&= \cdots = b_{n-2,p-1} + 2(b_{n-3,p-1} + b_{n-4,p-1} + \cdots + b_{2p-1,p-1}),
\end{aligned}$$

as $\binom{n-(n-2p+1)-2(p-1)}{j+2} = 0$. Therefore, for all $1 \leq p \leq \lfloor \frac{n}{2} \rfloor$ but $p \neq \frac{n}{2}$, we have $b_{n,p} - b_{n-1,p} = b_{n-2,p-1} + b_{n-3,p-1}$ where $b_{n,1} = n - 2$ for $n \geq 3$ and $b_{n,2} = (n-4)^2$ for $n \geq 5$. Since, for all $n \geq 4$, $\iota_{n,p} = 1$ if $p = \frac{n}{2}$, it follows that

$$\frac{n-2p}{n} \iota_{n,p} = \frac{n-2p-1}{n-1} \iota_{n-1,p} + \frac{n-2p}{n-2} \iota_{n-2,p-1} + \frac{n-2p-1}{n-3} \iota_{n-3,p-1},$$

with $\iota_{n,1} = n$ for $n \geq 3$ and $\iota_{n,2} = n(n-4)$ for $n \geq 5$ can be used to recursively compute the sequence $\{\iota_{n,p}\}_{n \geq 4}$.

3. The median graph associated to $\mathcal{S}(n, m)$

In this section we find formulae for the number of vertices and the number of maximal induced subcubes in the median graph associated to $\mathcal{S}(n, m)$, $n \geq 2$, that is, the number of incompatible split subsystems and the number of maximal incompatible split subsystems of $\mathcal{S}(n, m)$, respectively.

To compute these numbers, it is convenient to re-parametrise by putting $n := 2m + q$, for $q \in \mathbb{N}_0$, and considering the split system $\mathcal{S}(2m + q, m)$. We will also employ the following explicit description of $\mathcal{S}(2m + q, m)$: It consists of those splits $S_i := \{A_i, B_i\}$, $1 \leq i \leq 2m + q$

of $[n]$ with $A_i := \{i, i+1, \ldots, i+m-1\}$ and $B_i := [n] \setminus A_i$ (where addition is taken modulo $2m+q$).

Denote the set of non-empty incompatible split subsystems of $\mathcal{S}(2m+q, m)$ by $\mathcal{I}_m^q$, and put $v_m^q = |\mathcal{I}_m^q| + 1$ (for example, the split system depicted in Fig. 1 is an element of $\mathcal{I}_4^1$). The aim of the following collection of results is to determine a formula for v_m^q, $m \geq 2$ (noting that $v_1^q = |\mathcal{I}_1^q| + 1 = 3 + q$).

Let

$$\mathcal{I} := \{(\mathcal{S}, S_k) : \mathcal{S} \in \mathcal{I}_m^q \text{ and } S_k \in \mathcal{S}\}$$

and let $\mathcal{T}_m$ denote the set of ternary strings with length m and $\mathcal{T} = \mathcal{T}_m \times [2m+q]$. Define a map $\phi : \mathcal{I} \to \mathcal{T}$ which takes an element $(\mathcal{S}, S_k) \in \mathcal{I}$ to the pair $(\mathbf{v}, k) = \phi(\mathcal{S}, S_k)$, where the string $\mathbf{v}$ has terms v_j, $j = 1, \ldots, m$, given by

$$v_j = \begin{cases} 1 & \text{if } \ell := j+k-1 \in \{k, k+1, \ldots, k+m-1\} \text{ and } S_\ell \in \mathcal{S}, \\ 2 & \text{if } \ell := j+k-1 \in \{k, k+1, \ldots, k+m-1\} \text{ and } S_{\ell+m+q} \in \mathcal{S}, \\ 0 & \text{else,} \end{cases}$$

where addition is taken modulo $2m+q$. Note that, by definition, for any $(\mathcal{S}, S_k) \in \mathcal{I}$ the first digit of $\mathbf{v}$ in $\phi(\mathcal{S}, S_k) = (\mathbf{v}, k)$ always equals 1. To illustrate this definition, note that for the incompatible split system $\mathcal{S} \in \mathcal{I}_4^1$ depicted in Fig. 1, we have $\phi(\mathcal{S}, S_1) = (1112, 1)$ and $\phi(\mathcal{S}, S_3) = (1222, 3)$.

We begin by giving a characterisation of the pairs that are contained in $\phi(\mathcal{I})$. For $(\mathbf{v}, k) \in \mathcal{T}$, define the set $\mathcal{S}_{(\mathbf{v},k)}$ to be the split system

$$\bigcup_{i \in [m], v_i = 1} \{S \in \mathcal{S}(2m+q, m) : S = S_{\ell(i)}\} \cup \bigcup_{i \in [m], v_i = 2} \{S \in \mathcal{S}(2m+q, m) : S = S_{\ell(i)+m+q}\}, \tag{5}$$

where $\ell(i) := i + k - 1$, for all $i \in [m]$, and addition in the split indices is taken modulo $2m+q$. For example, $\mathcal{S}_{(1222,3)}$ is the split system depicted in Fig. 1.

Lemma 3.1. *A pair $(\mathbf{v}, k) \in \mathcal{T}$ is contained in $\phi(\mathcal{I})$ if and only if $\mathbf{v}$ starts with 1 and any substring of $\mathbf{v}$ of the form $200\cdots01$ contains q or more zeros.*

Proof. Let $(\mathbf{v}, k) = \phi(\mathcal{S}, S)$, for $(\mathcal{S}, S) \in \mathcal{I}$. Suppose there is a substring $\mathbf{a}$ of $\mathbf{v}$ of the form $200\cdots01$ containing p zeros, with $p < q$. Then there exists some $j \in [m]$ with $v_j = a_1 = 2$, $v_{j+p+1} = 1$ and $v_{j+i} = a_{i+1} = 0$, for all $1 \leq i \leq p$. Hence, $S_{j+m+q}, S_{j+p+1} \in \mathcal{S}$, and therefore S_{j+m+q} and S_{j+p+1} are incompatible. But this is a contradiction since $A_{j+p+1} \subsetneq B_{j+m+q}$.

Conversely, suppose that $(\mathbf{v}, k) \in \mathcal{T}$, that the first term of $\mathbf{v}$ is 1, and that any substring of $\mathbf{v}$ of the form $200\cdots01$ contains q or more zeros. Put $\mathcal{S} = \mathcal{S}_{(\mathbf{v},k)}$. We claim $(\mathcal{S}, S_k) \in \mathcal{I}$ which, since it is straightforward to verify that $(\mathbf{v}, k) = \phi(\mathcal{S}, S_k)$, will complete the proof of the lemma.

Since the first term of $\mathbf{v}$ is 1 it immediately follows that $S_k \in \mathcal{S}$. We now show that $\mathcal{S} \in \mathcal{I}_m^q$, from which it follows that $(\mathcal{S}, S_k) \in \mathcal{I}$, as claimed. If $|\mathcal{S}| = 1$, then $\mathcal{S}$ is incompatible by definition. Now suppose that $|\mathcal{S}| \geq 2$ and that S and S' are distinct splits in $\mathcal{S}$. We prove that S and S' are incompatible.

Since $S \neq S'$, there exist $i, j \in [m]$ distinct such that $S = S_{\tilde{i}}$ and $S' = S_{\tilde{j}}$ where either (a) $v_i = 1 = v_j$, in which case $\tilde{i} := i + k - 1$ and $\tilde{j} := j + k - 1$, (b) $v_i = 2 = v_j$, in which case $\tilde{i} := i + k + m + q - 1$ and $\tilde{j} := j + k + m + q - 1$, (c) $v_i = 1$ and $v_j = 2$, in which case $\tilde{i} := i + k - 1$ and $\tilde{j} := j + k + m + q - 1$, or (d) $v_i = 2$ and $v_j = 1$, in which case $\tilde{i} := i + k + m + q - 1$ and $\tilde{j} := j + k - 1$. For each of these 4 cases it is straight forward to show that S and S' are incompatible. Thus, $\mathcal{S} \in \mathcal{I}_m^q$, as required. ■

Now, for $\mathcal{S} \in \mathcal{I}_m^q$, let $\mathcal{I}_{\mathcal{S}}$ denote the subset $\{(\mathcal{S}, S) \in \mathcal{I} : S \in \mathcal{S}\}$ of $\mathcal{I}$. Clearly, $\mathcal{I} = \coprod_{\mathcal{S} \in \mathcal{I}_m^q} \mathcal{I}_{\mathcal{S}}$. Moreover, if $\mathcal{S}, \mathcal{S}' \in \mathcal{I}_m^q$ distinct, then it immediately follows by the definition of ϕ that $\phi(\mathcal{I}_{\mathcal{S}}) \cap \phi(\mathcal{I}_{\mathcal{S}'}) = \emptyset$. We now show that for any $\mathcal{S} \in \mathcal{I}_m^q$ there are some elements in $\phi(\mathcal{I}_{\mathcal{S}})$ with a particularly special form.

Lemma 3.2. *For all $(\mathcal{S}, S) \in \mathcal{I}$, there is some $(\mathbf{v}, k) \in \phi(\mathcal{I}_{\mathcal{S}})$ with $v_\ell = 1$ for $\ell := \max\{j \in [m] : v_j \neq 0\}$.*

Proof. Suppose $(\mathcal{S}, S_k) \in \mathcal{I}$ and put $(\mathbf{v}, k) := \phi(\mathcal{S}, S_k)$ and $\ell = \max\{j \in [m] : v_j \neq 0\}$. If $v_\ell = 1$ then $(\mathbf{v}, k)$ is of the required form. If $v_\ell \neq 1$ then $(\mathbf{v}, k)$ can be transformed iteratively into an element of $\phi(\mathcal{I}_{\mathcal{S}})$ having the required form by successively removing the last $t := m - \ell + 1$ digits from $\mathbf{v}$ and concatenating the length t string $100 \cdots 0$ onto the beginning of the resulting new string $\mathbf{v}' \in \mathcal{T}_m$. Note that in each iteration $(\mathbf{v}', k-t) = \phi(\mathcal{S}, S_{k-t})$ and so $(\mathbf{v}', k-t) \in \phi(\mathcal{I}_{\mathcal{S}})$. Also note that this process must clearly terminate since $v_1 = 1$ and m is finite. ■

Before continuing, we define a useful map $\tau : \mathcal{T}_m \to \mathbb{N}_0$ as follows. For $\mathbf{v} \in \mathcal{T}_m$, let $\mathbf{v}^*$ denote the string obtained by deleting all 0's from $\mathbf{v}$. In particular, $\mathbf{v}^*$ is a concatenation of blocks of 1's and 2's. We define $\tau(\mathbf{v})$ to be the number of occurrences of the string 21 in $\mathbf{v}^*$.

Now, let $\mathcal{P} \subseteq \phi(\mathcal{I})$ denote the set of pairs $(\mathbf{v}, k)$ where $\mathbf{v}$ satisfies the property given in Lemma 3.2, that is, $v_\ell = 1$ for $\ell := \max\{j \in [m] : v_j \neq 0\}$. Note that $\mathcal{P} \cap \phi(\mathcal{I}_{\mathcal{S}}) \neq \emptyset$ for all $\mathcal{S} \in \mathcal{I}_m^q$, and that if $(\mathbf{v}, k) \in \mathcal{P}$ then $v_1 = 1 = v_\ell$ where $\ell = \max\{j : v_j \neq 0\}$, and so there are $2\tau(\mathbf{v}) + 1$ blocks of 1's and 2's in $\mathbf{v}^*$.

Theorem 3.3. *Suppose $\mathcal{S} \in \mathcal{I}_m^q$. Then*

(i) *For all $(\mathbf{v}, k), (\mathbf{w}, r) \in \mathcal{P} \cap \phi(\mathcal{I}_{\mathcal{S}})$, we have $\tau(\mathbf{v}) = \tau(\mathbf{w})$.*
(ii) *If $(\mathbf{v}, k) \in \mathcal{P} \cap \phi(\mathcal{I}_{\mathcal{S}})$, then $|\mathcal{P} \cap \phi(\mathcal{I}_{\mathcal{S}})| = 2\tau(\mathbf{v}) + 1$.*

Proof. Let $\mathcal{S} \in \mathcal{I}_m^q$. Fix $(\mathbf{v}, k) \in \mathcal{P} \cap \phi(\mathcal{I}_{\mathcal{S}})$. Without loss of generality assume $k = 1$. Suppose $(\mathbf{w}, r) \in \mathcal{P} \cap \phi(\mathcal{I}_{\mathcal{S}})$ with $(\mathbf{v}, k) \neq (\mathbf{w}, r)$. Note that we cannot have $m + 1 \leq r \leq m + q + 1$ since in this case

$$A_r = \{r, r+1, \ldots, r+m-1\} \subsetneq \{m+1, m+2, \ldots, 2m+q\} = B_1$$

which implies that S_1 and S_r are not compatible. But since S_1 and S_r are both splits in $\mathcal{S}_{(\mathbf{w},r)}$ (as defined in (5)) and $\mathcal{S}_{(\mathbf{w},r)} = \mathcal{S} \in \mathcal{I}_m^q$, this is impossible.

Now, suppose $\mathbf{v}^* = a_1 a_2 \ldots a_t$ and $\mathbf{w}^* = b_1 b_2 \ldots b_t$ with $t = |\mathcal{S}|$. We claim that if $1 \leq r \leq m$ and $1 \leq r' \leq t$ is such that $a_{r'}$ corresponds to $v_r = 1$, then

$$\mathbf{w}^* = a_{r'} a_{r'+1} \cdots a_t \overline{a_1}\, \overline{a_2} \cdots \overline{a_{r'-1}},$$

and if $m + q + 2 \leq r \leq 2m + q$ and $1 \leq r' \leq t$ is such that $a_{r'}$ corresponds to $v_{r-m-q} = 2$, then

$$\mathbf{w}^* = \overline{a_{r'}}\, \overline{a_{r'+1}} \cdots \overline{a_t} a_1 a_2 \cdots a_{r'-1},$$

where for $x = 1, 2$, we define $\overline{x} := 3 - x$. Suppose $1 \leq r \leq m$. Then for all $r \leq \ell \leq m$, $w_{\ell-r+1} = v_\ell$. In addition, for all $1 \leq \ell \leq r-1$, if $v_\ell = 1$ then $w_{\ell+m-r+1} = 2$, and if $v_\ell = 2$ then

$w_{\ell+m+q-r+1} = 1$. Note that $w_l = 0$ for all other values of l. Similarly, if $m+2+q \le r \le 2m+q$ then $v_{r-m-q} = 2$. For all $r-m-q \le \ell \le m$, if $v_\ell = 2$ then $w_{\ell+m+q-r+1} = 1$, and if $v_\ell = 1$ then $w_{\ell+m-r+1} = 2$, and $w_{\ell+2m+q-r+1} = v_\ell$, for all $1 \le \ell \le r-m-q-1$. As before, $w_l = 0$ for all other values of l. This establishes the claim.

Now, note that $(\mathbf{w}, r) \in \mathcal{P} \cap \phi(\mathcal{I}_\mathcal{S})$ if and only if $b_1 = b_t = 1$. Thus, in view of the above claim it follows that $a_{r'} = 1$ and $a_{r'-1} = 2$ in case $1 \le r \le m$, and $a_{r'} = 2$ and $a_{r'-1} = 1$ in case $m+q+2 \le r \le 2m+q$. Since $a_1 = a_t = 1$, it is now straightforward to see that $\tau(\mathbf{v}) = \tau(\mathbf{w})$ and that $\mathcal{P} \cap \phi(\mathcal{I}_\mathcal{S})$ contains $2\tau(\mathbf{v})$ distinct elements that are not equal to $(\mathbf{v}, k)$. Therefore, (i) and (ii) hold. ■

For $i \in \mathbb{N}_0$, define $\mathcal{P}_i$ to be the subset of $\mathcal{P}$ consisting of those pairs $(\mathbf{v}, k)$ with $\tau(\mathbf{v}) = i$. Clearly, $\mathcal{P} = \coprod_{i\ge 0} \mathcal{P}_i$. The following corollary is a straightforward consequence of the previous theorem.

Corollary 3.4. *For $m \ge 2$, and $q \in \mathbb{N}_0$,*

$$|\mathcal{I}_m^q| = \sum_{i\ge 0} \frac{1}{2i+1} |\mathcal{P}_i|.$$

Hence, to find a formula for v_m^q, $m \ge 2$, it suffices to find one for $|\mathcal{P}_i|$, $i \ge 0$. To this end, for all $t \ge 2$, put

$$\mathcal{P}_{i,t} := \{(k, \mathbf{v}) \in \mathcal{P}_i : |\mathbf{v}^*| = t\},$$

noting that $\mathcal{P}_i = \dot\cup_{t\ge 1} \mathcal{P}_{i,t}$.

Lemma 3.5. *For all $i \in \mathbb{N}_0$ and all $t \ge 2$,*

$$|\mathcal{P}_{i,t}| = (2m+q) \binom{m-qi-1}{2i} \binom{m-qi-2i-1}{t-2i-1}.$$

Proof. Suppose $i \ge 0$, $t \ge 2$ and $(\mathbf{v}, k) \in \mathcal{P}_{i,t}$. Then, since $\mathbf{v}^*$ is the concatenation of $2i+1$ blocks of 1's and 2's, $v_1^* = 1$ (as $v_1 = 1$), and $|\mathbf{v}^*| = t$, there are $\binom{t-1}{2i}$ ways in which $\mathbf{v}^*$ could be made up from these blocks.

We now count the number of elements $(\mathbf{w}, k)$ in $\mathcal{P}_{i,t}$ with $\mathbf{w}^* = \mathbf{v}^*$ by constructing all $\mathbf{w}$ satisfying this equality starting from the string $\mathbf{v}^*$. In view of Lemma 3.1, we first place q consecutive 0's in front of each block of 1's in $\mathbf{v}^*$ that is preceded by a 2. In the resulting ternary string $\mathbf{v}'$, consider the substrings of the form $00\cdots01$ as one digit. Then there are $\binom{m-qi-1}{t-1}$ possible ways to add further 0's to $\mathbf{v}'$ to obtain a length m ternary string starting with 1, that is, a string $\mathbf{w}$ as required.

Now, since

$$\binom{t-1}{2i}\binom{m-qi-1}{t-1} = \binom{m-qi-1}{2i}\binom{m-qi-2i-1}{t-2i-1},$$

and there are $2m+q$ possible choices for k in the pair $(\mathbf{v}, k)$, the lemma now follows immediately. ■

The last lemma and the Binomial Theorem imply

$$|\mathcal{P}_i| = \sum_{t\ge 1} |\mathcal{P}_{i,t}| = (2m+q) \binom{m-qi-1}{2i} 2^{m-2i-qi-1},$$

for all $i \geq 0$, and so by Corollary 3.4 and the fact that $v_1^q = 3 + q$ (see above)

$$v_m^q = 1 + |\mathcal{I}_m^q| = 1 + \sum_{i \geq 0} \frac{1}{2i+1} |\mathcal{P}_i| = 1 + \sum_{i \geq 0} \frac{2m+q}{2i+1} \binom{m - qi - 1}{2i} 2^{m-2i-qi-1} \quad (6)$$

where 1 in the sum counts the empty split subsystem of $\mathcal{S}(2m+q, m)$, $m \geq 1$.

We conclude this section by deriving a formula for the number c_m^q of maximal incompatible subsystems of $\mathcal{I}_m^q$, $m \geq 2$, $q \in \mathbb{N}_0$ (noting that $c_1^q = |\mathcal{S}(2+q, 1)| = 2 + q$). First note that Lemma 3.1 immediately implies that $\mathcal{S}$ is a maximal incompatible subsystem of $\mathcal{S}(2m+q, m)$ if and only if for every $(\mathbf{v}, k) \in \mathcal{P} \cap \phi(\mathcal{I}_\mathcal{S})$ every 0 in the string $\mathbf{v}$ is contained in a substring of the form $20 \cdots 01$, and every such substring contains precisely q zeros. Therefore, for $i \geq 0$ fixed, $\mathcal{S}$ is a maximal incompatible subsystem of $\mathcal{S}(2m+q, m)$ if and only if for every $(\mathbf{v}, k) \in \mathcal{P}_i \cap \phi(\mathcal{I}_\mathcal{S})$ the associated string $\mathbf{v}^*$ has i occurrences of 21 and i occurrences of 12 in $m - qi - 1$ possible positions. Therefore, since k can take on $2m + q$ possible values, by Theorem 3.3(ii), for $m \geq 1$ we have

$$c_m^q = \sum_{i \geq 0} \frac{2m+q}{2i+1} \binom{m - qi - 1}{2i}. \quad (7)$$

4. Recurrence relations for v_m^q and c_m^q

In this section we use generating functions to obtain recurrence relations for the sequences $\{v_m^q\}_{m \geq 1}$ and $\{c_m^q\}_{m \geq 1}$, $q \in \mathbb{N}_0$.

Let

$$F_q(y) := \sum_{m \geq 1} v_m^q y^m$$

be the generating function associated to the sequence $\{v_m^q\}_{m \geq 1}$. To obtain a formula for $F_q(y)$, we put

$$f_q(x, y) := \sum_{m \geq 1} \left[1 + \sum_{i \geq 0} \frac{2m+q}{2i+1} \binom{m - qi - 1}{2i} 2^{m-2i-qi-1} x^{2i+1} \right] y^{m-1},$$

noting that, by (6), we have

$$F_q(y) = y f_q(1, y). \quad (8)$$

Theorem 4.1. *For $q \in \mathbb{N}_0$, we have*

$$f_q(x, y) = \frac{(q + 2 - 2qy)x}{1 - 4y + 4y^2 - x^2 y^{q+2}} + \frac{1}{1-y}.$$

Proof. Put $q = 2q'$. Then, with m replaced by $m + 1$,

$$f_q(x, y) = \sum_{m \geq 0} \left(1 + (2m + 2 + 2q') \sum_{i \geq 0} \frac{1}{2i+1} \binom{m - 2q'i}{2i} \frac{x^{2i+1}}{2^{(2q'+2)i}} \right) (2y)^m. \quad (9)$$

Differentiating $f_q(x, y)$ with respect to x yields

$$\frac{\partial}{\partial x} f_q(x, y) = \sum_{m \geq 0} \left((2m + 2 + 2q') \sum_{i \geq 0} \binom{m - q'2i}{2i} \frac{x^{2i}}{2^{(q'+1)(2i)}} \right) (2y)^m.$$

Now, consider the polynomials $A(x, y)$ and $B(x, y)$ defined by

$$A(x, y) = \sum_{m \geq 0} \left(\sum_{i \geq 0} \binom{m - q'2i}{2i} \left(\frac{x}{2^{(q'+1)}} \right)^{2i} \right) (2y)^m \quad \text{and}$$

$$B(x, y) = \sum_{m \geq 0} \left(\sum_{i \geq 0} \binom{m - q'i}{i} \left(\frac{x}{2^{(q'+1)}} \right)^{i} \right) (2y)^m.$$

Then since for all $j \geq 0$ the terms of the form ax^{2j+1} cancel and those of the form ax^{2j} get duplicated, it follows that

$$A(x, y) = \frac{1}{2}(B(x, y) + B(-x, y)).$$

We claim that $(1 - 2y - xy^{q'+1})B(x, y) = 1$. To see this, note that it suffices to show that the coefficients of $x^i y^m$ in the product $(1 - 2y - xy^{q'+1})B(x, y)$ cancel whenever $(i, m) \neq (0, 0)$. But this follows by using the identity $\binom{a+1}{b} = \binom{a}{b} + \binom{a}{b-1}$ to replace $\binom{m-q'i}{i}$ in the product $1 \cdot B(x, y)$. Consequently, we also have $(1 - 2y + xy^{q'+1})B(-x, y) = 1$, and so

$$B(x, y) = \frac{1}{1 - 2y - xy^{q'+1}} \quad \text{and} \quad B(-x, y) = \frac{1}{1 - 2y + xy^{q'+1}}.$$

A straightforward check shows that

$$\begin{aligned} \frac{\partial}{\partial x} f_q(x, y) &= \frac{\partial}{\partial y} A(x, y) \cdot (2y) + (2 + 2q')A(x, y) \\ &= y \frac{\partial}{\partial y}(B(x, y) + B(-x, y)) + (1 + q)(B(x, y) + B(-x, y)) \\ &= \frac{q' + 1 - 2q'y}{(1 - 2y - xy^{q'+1})^2} + \frac{q' + 1 - 2q'y}{(1 - 2y + xy^{q'+1})^2}. \end{aligned}$$

Hence,

$$f_q(x, y) = \int \frac{\partial}{\partial x} f_q(x, y) \mathrm{d}x = \frac{(q' + 1 - 2q'y)2x}{1 - 4y + 4y^2 - x^2 y^{2q'+2}} + h(y), \tag{10}$$

for some function $h(y)$. Since $f_q(0, y) = h(y)$, we obtain $h(y) = \frac{1}{1-y}$. The formula stated in the theorem follows. ■

Corollary 4.2. *Let $q \in \mathbb{N}_0$. Then the sequence $\{v_m^q\}_{m \geq 1}$ can be obtained by using*

$$v_m^q = \begin{cases} 1 & \text{if } m = 0, \\ 1 + (2m + q)2^{m-1} & \text{for all } 1 \leq m \leq q + 1, \\ 4v_{m-1}^q - 4v_{m-2}^q + v_{m-q-2}^q & \text{for all } m > q + 1; \end{cases}$$

or,

$$v_m^q = \begin{cases} 1 & \text{if } m = 0, \\ 1 + (2m+q)2^{m-1} & \text{for all } 1 \le m \le q+1, \\ 3v_{m-1}^q - \sum_{i=1}^{q} v_{m-i-1}^q & \text{for all } m > q+1. \end{cases}$$

Proof. The fact that $v_0^q = 1$ and that $v_m^q = 1 + (2m+q)2^{m-1}$ for all $1 \le m \le q+1$ both hold follows from (6). Now, in view of (8), Theorem 4.1 and the fact that 1 is a zero of both the numerator and the denominator of $yf_q(1, y)$, we have

$$\begin{aligned} F_q(y) &= yf_q(1, y) \\ &= \frac{(q+1)y - (2q+3)y^2 + y^3 + y^4 + \cdots + y^{q+2}}{1 - 4y + 4y^2 - y^{q+2}} \\ &= \frac{(q+1)y - qy^2 - (q-1)y^3 - (q-2)y^4 - \cdots - (q-(q-2))y^q - y^{q+1}}{1 - 3y + y^2 + y^3 + \cdots + y^{q+1}}. \end{aligned}$$

The corollary now follows immediately. ■

A recurrence relation can be obtained in a similar way for the sequence $\{c_m^q\}_{m\ge1}$ as follows. Let

$$G_q(z) := \sum_{m\ge1} c_m^q z^m$$

be the associated generating function, and put

$$g_q(x, z) = \sum_{m\ge1}\sum_{i\ge0} \left[\frac{2m+q}{2i+1} \binom{m-1-qi}{2i} x^{m-qi} \right] z^{m-1},$$

noting that by (7), we have $G_q(z) = zg_q(1, z)$. Using similar arguments to those used in the proof of Theorem 4.1, it can be shown that

$$g_q(x, z) = \frac{(q+2)x - qx^2z}{1 - 2xz + x^2z^2 - x^2z^{q+2}},$$

and, therefore, that

$$G_q(z) = zg_q(1, z) = \frac{(q+2)z - qz^2}{1 - 2z + z^2 - z^{q+2}}$$

is a formula for $G_q(z)$. Combined with (7), it immediately follows, that

$$c_m^q = \begin{cases} 0 & \text{for } m = 0, \\ 2m+q & \text{for all } 1 \le m \le q+1, \\ 2c_{m-1}^q - c_{m-2}^q + c_{m-2-q}^q & \text{for all } m \ge q+2, \end{cases} \tag{11}$$

can be used to recursively compute the terms of the sequence $\{c_m^q\}_{m\ge1}$ as stated in the introduction.

Remark 4.3. The recurrence relations in Corollary 4.2 and, (11) can also be obtained using Zeilberger's Algorithm (Section 5.8 in [15]). However, our proof provides recurrence relations for general $q \in \mathbb{N}_0$, whereas Zeilberger's algorithm only gives recurrence relations for q fixed.

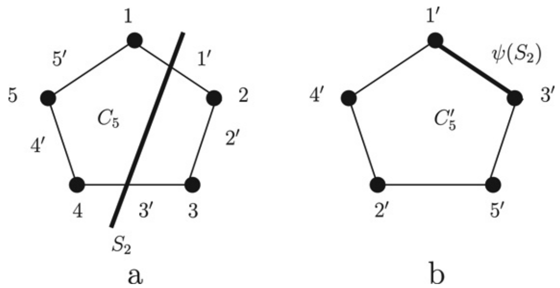

Fig. 2. For $m = 2$ the cycle C_5 is depicted in (a). For each $i \in \{1, \dots, 5\}$, the edge $\{i, i+1\}$ of C_5 is labelled by i'. The split S_2 is depicted by a bold straight line. The graph C_5' associated to C_5 is depicted in (b). The edge $\psi(S_2)$ of G is pictured in bold.

5. The sequence $\{w_m\}_{m \geq 0}$

As in Section 3, for $q \in \mathbb{N}_0$ and $m \geq 1$, let $\mathcal{I}_m^q$ denote the set of non-empty incompatible split subsystems of $\mathcal{S}(2m+q, m)$. In this section, we provide a combinatorial proof that recurrence relation in (1) given in Section 1 holds in case $q = 1$, i.e. putting $w_m := 1 + |\mathcal{I}_m^1|$, we show that

$$w_m = 3w_{m-1} - w_{m-2}, \quad m \geq 2, \ (w_0 = 1 \text{ and } w_1 = 4).$$

Suppose $m \geq 1$. In C_{2m+1}, label each edge $\{i, i+1\}$, $1 \leq i \leq 2m+1 \pmod{2m+1}$, by i'. We next associate a graph C_{2m+1}' to C_{2m+1} as follows. The vertex set $V(C_{2m+1}')$ of C_{2m+1}' comprises of the labels i', $1 \leq i \leq 2m+1$, and any two vertices $i', j' \in V(C_{2m+1}')$ are joined by an edge if $i - j \equiv \pm m \pmod{2m+1}$ (see Fig. 2 for an example for the case $m = 2$). It is straightforward to see that C_{2m+1}' is connected and that the degree of every vertex in C_{2m+1}' is 2. Since $|V(C_{2m+1}')| = 2m+1$, it follows that C_{2m+1}' is a $(2m+1)$-cycle. Now the map

$$\psi : \mathcal{S}(2m+1, m) \to E(C_{2m+1}') : S_{i+1} \mapsto \{i', (i+m)'\}$$

is clearly a bijection. Since any two distinct splits $S, S' \in \mathcal{S}(2m+1, m)$ are incompatible if and only if $\psi(S)$ and $\psi(S')$ do not have a vertex in common, it follows that w_m equals the number of matchings in C_{2m+1}'. Since C_{2m+1}' is a $(2m+1)$-cycle and the Lucas number l_n, $n \geq 1$, is the number of matchings in C_n [19, sequence id: A000204] it follows that $w_m = l_{2m+1}$. In particular, $\{w_m\}_{m \geq 1}$ is the bisection of $\{l_m\}_{m \geq 1}$.

Defining a maximal matching in the obvious way, similar arguments can be applied to show that, for all $m \geq 1$, c_m^1 equals the number of maximal matchings in C_{2m+1}.

We conclude by answering a question concerning tight-spans that was raised by B. Sturmfels. A split S on $[n]$ *separates* a pair of distinct elements $i, j \in [n]$ if i and j are contained in different parts of S. To any split system $\mathcal{S}$ of $[n]$ we associate the metric $d_{\mathcal{S}} : [n] \times [n] \to \mathbb{R}_{\geq 0}$, given by

$$d_{\mathcal{S}}(i, j) = |\{S \in \mathcal{S} : S \text{ separates } i \text{ and } j\}|, \quad i, j \in [n].$$

In addition, define the *tight-span*[1] associated to $\mathcal{S}$ to be the polytopal complex that is given by the union of the bounded faces of the polytope in $\mathbb{R}^{[n]}$ given by

$$P(d_{\mathcal{S}}) = \{f : [n] \to \mathbb{R} : f(i) + f(j) \geq d_{\mathcal{S}}(i, j), \text{ for all } i, j \in [n]\}.$$

[1] The tight-span can be defined for metrics in general — see e.g. [8,13] for more details.

Sturmfels asked whether the number of vertices of the tight-span associated to the split system $\mathcal{S}(2m+1, m)$ can be given in terms of Lucas numbers. Since by [11, Theorem 3.1] the median graph associated to $\mathcal{S}(2m+1, m)$ is isomorphic to the 1-skeleton of the tight-span associated to $\mathcal{S}(2m+1, m)$, it immediately follows by the above observations that this is indeed the case.

Acknowledgements

We thank B. Sturmfels for interesting discussions and for raising the question discussed in the last section concerning tight-spans at the Phylogenetic Combinatorics meeting in Uppsala in July 2004, and Mark Noy for pointing out the simple proof for counting vertices in the median graph of $\mathcal{S}(n)$ presented in Section 2. YBC, JHK, and YSK thank the Com^{2}MaC center and acknowledge the support received from the SRC/ERC program of MOST/KOSEF (R11-1999-054). VLM was supported in part by EPSRC grant EP/D068800/1. In addition, we thank the anonymous referees for helpful comments, in particular the referee who provided the proof using matchings given in the last section as an alternative to the proof that we originally presented.

References

[1] S.P. Avann, Metric ternary distributive semi-lattices, Proceedings of the American Mathematical Society 12 (1961) 407–414.
[2] H.-J. Bandelt, Retracts of hypercubes, Journal of Graph Theory 8 (1984) 501–510.
[3] H.-J. Bandelt, A. Dress, A canonical decomposition theory for metrics on a finite set, Advances in Mathematics 92 (1992) 47–105.
[4] H.-J. Bandelt, P. Forster, B.C. Sykes, M.B. Richards, Mitochondrial portraits of human populations using median networks, Genetics 141 (1994) 743–753.
[5] H.-J. Bandelt, M. Van de Vel, Superextensions and the depth of median graphs, Journal of Combinatorial Theory, Series A 57 (1991) 187–202.
[6] B. Brešar, S. Klavžar, R. Škrekovski, The cube polynomial and its derivatives: The case of median graphs, The Electronic Journal of Combinatorics 10 (2003) #R3.
[7] D. Bryant, V. Moulton, NeighborNet: An agglomerative method for the construction of phylogenetic networks, Molecular Biology and Evolution 21 (2004) 255–265.
[8] A. Dress, Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: A note on combinatorial properties of metric spaces, Advances in Mathematics 53 (1984) 321–402.
[9] A. Dress, M. Hendy, K.T. Huber, V. Moulton, On the number of vertices and edges of the Buneman graph, Annals of Combinatorics 1 (1997) 329–337.
[10] A. Dress, K.T. Huber, V. Moulton, An exceptional split geometry, Annals of Combinatorics 4 (2000) 1–11.
[11] A. Dress, K.T. Huber, V. Moulton, An explicit computation of the injective hull of certain finite metric spaces in terms of their associated Buneman complex, Advances in Mathematics 168 (2002) 1–28.
[12] A. Dress, D. Huson, Constructing splits graphs, IEEE/ACM Transactions in Computational Biology and Bioinformatics 1 (2004) 109–115.
[13] A. Dress, V. Moulton, W. Terhalle, T-theory: An overview, European Journal of Combinatorics 17 (1996) 161–175.
[14] D. Duffus, I. Rival, Graphs orientable as distributive lattices, Proceedings of the American Mathematical Society 88 (1983) 197–200.
[15] R.L. Graham, D.E. Knuth, O. Patashnik, Concrete Mathematics: A Foundation for Computer Science, 2nd edition, Addison-Wesley, Reading, MA, 1994.
[16] B. Holland, F. Delsuc, V. Moulton, Visualizing conflicting evolutionary hypotheses in large collections of trees using consensus networks, Systematic Biology 54 (2005) 66–76.
[17] K.T. Huber, J. Koolen, V. Moulton, On the structure of the tight-span of a totally split-decomposable metric, European Journal of Combinatorics 27 (3) (2006) 461–479.
[18] S. Klavžar, H.M. Mulder, R. Škrekovski, An Euler-type formula for median graphs, Discrete Mathematics 187 (1998) 255–258.
[19] The On-Line Encyclopedia of Integer Sequences, http://www.research.att.com/˜njas/sequences/.
[20] A. Verbeek, Superextensions of Topological Spaces, in: Math. Centre Tracts, vol. 41, Amsterdam, 1972.