Available online at www.sciencedirect.com
ScienceDirect
European Journal of Combinatorics 28 (2007) 1814–1831
European Journal of Combinatorics
www.elsevier.com/locate/ejc

Linearly independent split systems

David Bryant[a,1], Andreas Dress[b,c]

[a] Department of Mathematics, University of Auckland, Private Bag 92019, Auckland, New Zealand
[b] Department of Combinatorics and Geometry, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, 320 Yue Yang Road, Shanghai 200031, China
[c] Max Planck Institute for Mathematics in the Sciences, Inselstrasse 22-26, D 04103 Leipzig, Germany

Received 21 December 2005; accepted 28 April 2006
Available online 11 July 2006

Abstract

An important procedure in the mathematics of phylogenetic analysis is to associate, to any collection of weighted splits, the metric given by the corresponding linear combination of split metrics. In this note, we study necessary and sufficient conditions for a collection of splits of a given finite set X to give rise to a linearly independent collection of split metrics. In addition, we study collections of splits called *affine split systems* induced by a configurations of lines and points in the plane. These systems not only satisfy the linear-independence condition, but also provide a $\mathbb{Z}$-basis of the $\mathbb{Z}$-lattice $\mathcal{D}_{\text{even}}(X \mid \mathbb{Z})$ consisting of all integer-valued symmetric maps $D : X \times X \to \mathbb{Z}$ defined on X that vanish on the diagonal and for which, in addition, $D(x, y) + D(y, z) + D(z, x) \equiv 0 \bmod 2$ holds for all $x, y, z \in X$. This $\mathbb{Z}$-lattice is generated by all split metrics considered as vectors in the real vectorspace $\mathcal{D}(X \mid \mathbb{R})$ consisting of all real-valued symmetric maps defined on X that vanish on the diagonal — and, hence, is also an $\mathbb{R}$-basis of that vectorspace.

1. Introduction

Let X denote a finite set of cardinality n. A *split* $S = \{A, B\}$ of X is a partition of X into two disjoint parts $A, B \in \mathcal{P}(X)$; the split is *proper* if both parts are non-empty, and *trivial* if one part contains exactly one element. The set of all splits is denoted by $\mathcal{S}(X)$, and the set of all proper splits is denoted by $\mathcal{S}^*(X)$. Any subset $\mathcal{S}$ of $\mathcal{S}(X)$ (or $\mathcal{S}^*(X)$) is called a *split system* (or a *proper split system*, respectively).

E-mail addresses: bryant@math.auckland.ac.nz (D. Bryant), dress@sibs.ac.cn, dress@mis.mpg.de (A. Dress).
[1] Tel.: +64 373 7599x88763; fax: +64 373 7457.

0195-6698/$ - see front matter
doi:10.1016/j.ejc.2006.04.007

Given a split $S = \{A, B\}$ in $\mathcal{S}(X)$ and an element x in X, let $S(x)$ denote the subset, A or B, in S that contains x, and let $\overline{S}(x)$ denote its complement $X - S(x)$, i.e. the subset, A or B, of X in S that does not contain x. The split S in $\mathcal{S}(X)$ is said to *separate* two elements x, y in X if $S(x) \neq S(y)$ holds. For any split $S \in \mathcal{S}(X)$, let d_S denote the associated *split metric*, i.e. the map

$$d_S : X \times X \to \mathbb{R} : (x, y) \mapsto \delta_{S(x),\overline{S}(y)} = \begin{cases} 1 & \text{if } S \text{ separates } x, y \\ 0 & \text{otherwise,} \end{cases}$$

where, for any two subsets U and V of X, we follow standard conventions and put $\delta_{U,V} := 1$ if $U = V$, and $\delta_{U,V} := 0$ otherwise.

Since $\delta_{S(x),\overline{S}(y)} = 1 - \delta_{S(x),S(y)} = \delta_{\overline{S}(x),S(y)}$ always holds, d_S can be considered as an element in the $\mathbb{R}$-vectorspace $\mathcal{D}(X \mid \mathbb{R})$ consisting of all real-valued symmetric maps $D : X \times X \to \mathbb{R}$ defined on X that vanish on the diagonal. Furthermore, as $\delta_{S(x),\overline{S}(y)} \in \mathbb{Z}$ and

$$d_S(x, y) + d_S(y, z) + d_S(z, x) \in \{0, 2\}$$

always holds, the map d_S can also be viewed as an element in the $\mathbb{Z}$-lattice $\mathcal{D}_{\text{even}}(X \mid \mathbb{Z})$ consisting of all integer-valued maps $D \in \mathcal{D}(X \mid \mathbb{R})$ for which, in addition,

$$D(x, y) + D(y, z) + D(z, x) \equiv 0 \mod 2$$

holds for all $x, y, z \in X$. This lattice is also called the *cut lattice* associated with X (cf. [4]).

A split system $\mathcal{S} \subseteq \mathcal{S}(X)$ is said to be $\mathcal{D}$-*independent* if the corresponding collection

$$d(\mathcal{S}) := \{d_S : S \in \mathcal{S}\}$$

of split metrics forms a linearly independent subset of the vector space $\mathcal{D}(X \mid \mathbb{R})$ in which case it must, of course, be a proper split system, that is, it must be contained in $\mathcal{S}^*(X)$. The mathematics of split metrics (also called *cut metrics*) is surprisingly rich and multi-faceted, with applications to many quite diverse fields — see Deza and Laurent [4] for a comprehensive survey. Our motivation here is in the application of split metrics to evolutionary biology. Some of the very first methods for constructing evolutionary trees used models based on split metrics [3], and the analysis of split metrics was also crucial in the development of phylogenetic network methods [1].

We will derive two distinct characterizations of split-system independence:

Theorem 1. *For any two splits $S = \{A, B\}$ and $S' = \{A', B'\}$ in $\mathcal{S}(X)$, define*

$$\mu(S, S') := |A \cap A'||B \cap B'| + |A \cap B'||B \cap A'|,$$

and, for any split system $\mathcal{S} \subseteq \mathcal{S}(X)$, put

$$\mathbf{M}(\mathcal{S}) := (\mu(S, S'))_{S,S' \in \mathcal{S}}.$$

Then, $\mathbf{M}(\mathcal{S})$ is a positive semi-definite symmetric matrix for every split system $\mathcal{S} \subseteq \mathcal{S}(X)$ whose rank coincides with the dimension of the subspace

$$\langle \mathcal{S} \rangle_{\mathbb{R}} := \langle d_S : S \in \mathcal{S} \rangle_{\mathbb{R}}$$

of $\mathbb{R}$-vectorspace $\mathcal{D}(X \mid \mathbb{R})$ generated by the collection $\{d_S : S \in \mathcal{S}\}$ of split metrics associated with $\mathcal{S}$. Hence, the following four assertions are equivalent:

(1) *$\mathcal{S}$ is $\mathcal{D}$-independent;*
(2) *$\mathbf{M}(\mathcal{S})$ is positive definite;*

(3) *the determinant* $\det(\mathcal{S}) := \det \mathbf{M}(\mathcal{S})$ *of* $\mathbf{M}(\mathcal{S})$ *does not vanish;*
(4) *this determinant is positive.*

Theorem 2. *For any element* $z \in X$ *and any two splits* $S, S' =\in \mathcal{S}(X)$*, put*

$$\mu_z(S, S') := \binom{|\overline{S}(z) \cap \overline{S'}(z)| + 1}{2}$$

and put

$$\mathbf{M}_z(\mathcal{S}) := (\mu_z(S, S'))_{S,S' \in \mathcal{S}}$$

for any split system $\mathcal{S} \subseteq \mathcal{S}(X)$*. Then, this matrix is also always a positive semi-definite symmetric matrix whose rank coincides, for every* $z \in Z$*, with the dimension of the* $\mathbb{R}$*-vectorspace* $\langle \mathcal{S} \rangle_{\mathbb{R}}$*. In particular, the following four assertions are equivalent:*

(1) $\mathcal{S}$ *is* $\mathcal{D}$*-independent;*
(2) *the determinant* $\det_z(\mathcal{S}) := \det \mathbf{M}_z(\mathcal{S})$ *does not vanish (or, equivalently, is positive) for all* $z \in X$*;*
(3) *the determinant* $\det_z(\mathcal{S})$ *does not vanish (or, equivalently, is positive) for at least one* $z \in X$*;*
(4) $\mathbf{M}_z(\mathcal{S})$ *is positive definite for at least one or, equivalently, for all,* $z \in X$*.*

The determinants $\det(\mathcal{S})$ and $\det_z(\mathcal{S})$ do not appear to be strongly related to each other in general, beyond the fact that one vanishes if and only if the other vanishes. Consider the case $|\mathcal{S}| = 1$. If $\mathcal{S}$ contains the single split $S = \{A, B\}$ and $z \in B$ holds, then only one of the two numbers

$$\det(\mathcal{S}) = \mu(S, S) = |A|(n - |A|)$$

and

$$\det_z(\mathcal{S}) = \mu_z(S, S) = \binom{|A| + 1}{2}$$

depends on n. However, there is a direct connection in the case when $\mathcal{S}$ has cardinality $\binom{n}{2}$; that is, when $\mathcal{S}$ has the largest cardinality a $\mathcal{D}$-independent split system can attain:

Theorem 3. *Let* $\mathcal{S} \subseteq \mathcal{S}(X)$ *be a split system with* $\#\mathcal{S} = \binom{n}{2}$*. Then*

$$\det(\mathcal{S}) = 2^{(n-1)(n-2)} \det_z(\mathcal{S})$$

holds for every $z \in X$*. One has* $\det(\mathcal{S}) = 2^{(n-1)(n-2)}$ *— or, equivalently,* $\det_z(\mathcal{S}) = 1$*— if and only if the* $\mathbb{Z}$*-sublattice*

$$\langle \mathcal{S} \rangle_{\mathbb{Z}} := \langle d_S : S \in \mathcal{S} \rangle_{\mathbb{Z}}$$

of $\mathbb{Z}$*-lattice* $\mathcal{D}_{\text{even}}(X \mid \mathbb{Z})$ *generated by the collection* $d(\mathcal{S})$ *of split metrics associated with* $\mathcal{S}$ *coincides with* $\mathcal{D}_{\text{even}}(X \mid \mathbb{Z})$*, in which case the set* $d(\mathcal{S})$ *is a* $\mathbb{Z}$*-basis of* $\mathcal{D}_{\text{even}}(X \mid \mathbb{Z})$*. In general,* $\det_z(\mathcal{S})$ *coincides with the square of the index* $(\mathcal{D}_{\text{even}}(X \mid \mathbb{Z}) : \langle \mathcal{S} \rangle_{\mathbb{Z}})$ *of* $\langle \mathcal{S} \rangle_{\mathbb{Z}}$ *in* $\mathcal{D}_{\text{even}}(X \mid \mathbb{Z})$ *whenever this index is finite, and vanishes otherwise.*

A split system $\mathcal{S} \subseteq \mathcal{S}(X)$ is *weakly compatible* [1] if it is proper and one of the three intersections $S_1(x) \cap \overline{S_2}(x) \cap \overline{S_3}(x)$, $\overline{S_1}(x) \cap S_2(x) \cap \overline{S_3}(x)$, and $\overline{S_1}(x) \cap \overline{S_2}(x) \cap S_3(x)$ is empty for any three splits S_1, S_2, S_3 in $\mathcal{S}$ and any $x \in X$. It was observed in [1] that every

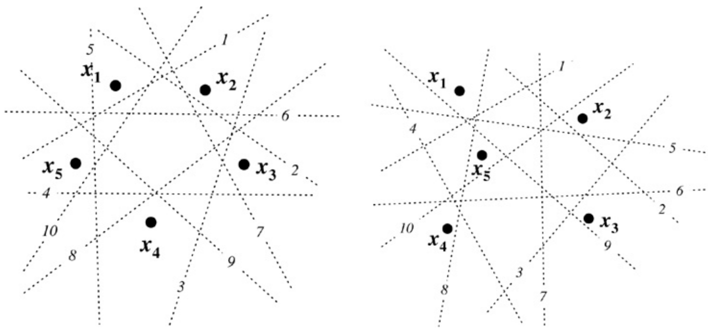

Fig. 1. Two configurations generating affine split systems on the set $X = \{x_1, x_2, \ldots, x_5\}$. Both configurations give rise to ten distinct splits, one for each line. In the left-hand configuration, Lines 1 to 5 give the splits separating out the single elements $x_1, x_2, \ldots, x_5$, respectively, while Lines 6 to 10 induce the splits separating out the pairs $\{x_1, x_2\}$, $\{x_2, x_3\}$, $\{x_3, x_4\}$, $\{x_4, x_5\}$, and $\{x_5, x_1\}$, respectively. In the right-hand configuration, Lines 1 to 4 induce the splits separating out the single elements $x_1, \ldots, x_4$, while Lines 5 to 10 separate out the pairs $\{x_1, x_2\}$, $\{x_3, x_4\}$, $\{x_2, x_3\}$, $\{x_4, x_1\}$, $\{x_4, x_5\}$, and $\{x_5, x_1\}$, respectively.

weakly compatible split system is $\mathcal{D}$-independent; so, $\det(\mathcal{S}) \geq 2^{(n-1)(n-2)}$ must hold for the determinant associated with any weakly compatible split system $\mathcal{S}$ of cardinality $\binom{n}{2}$. We will show here that, more specifically,

$$\det(\mathcal{S}) = 2^{(n-1)(n-2)}$$

holds whenever $\mathcal{S}$ is weakly compatible and has cardinality $\binom{n}{2}$. In fact, we will prove a more general result that refers to the following definition:

Definition 1. A split system $\mathcal{S} \subseteq \mathcal{S}(X)$ is *affine* if it is proper and there exists a map $\varphi : X \to \mathbb{E}^2$ from X to the Euclidian plane $\mathbb{E}^2$ and an $\mathcal{S}$-indexed family $(\ell_S)_{S \in \mathcal{S}}$ of straight lines in $\mathbb{E}^2$ so that none of these lines intersects the φ-image $\varphi(X)$ of X while ℓ_S intersects the line segment $[\varphi(x), \varphi(y)]$ for any $S \in \mathcal{S}$ and any two points $x, y \in X$ that are separated by S.

(i) A straight line $\ell \subset \mathbb{E}^2$ *induces* a given split S of X relative to the map φ if and only if ℓ intersects the straight line segment $[\varphi(x), \varphi(y)]$ for any two points $x, y \in X$ if and only if these two points are separated by S (in which case $\varphi(X) \cap \ell = \emptyset$ must, of course, hold).
(ii) We write $S = S(\varphi, \ell)$ in this case.
(iii) The collection of all proper splits of X arising in this way from a given map $\varphi : X \to \mathbb{E}^2$ will be denoted by $\mathcal{S}_\varphi$, i.e. we put

$$\mathcal{S}_\varphi := \mathcal{S}^*(X) \cap \{S(\varphi, \ell) : \ell \text{ a straight line in } \mathbb{E}^2, \varphi(X) \cap \ell = \emptyset\}.$$

(iv) Given any split system $\mathcal{S} \subseteq \mathcal{S}(X)$, a map $\varphi : X \to \mathbb{E}^2$ will be called an $\mathcal{S}$-map if $\mathcal{S} = \mathcal{S}_\varphi$ holds.

Point and line systems of this sort (or, more frequently, their geometric duals) arise in combinatorial and computational geometry (see Fig. 1 for two examples and, e.g., [6,9] for further discussion). Here are three of their well-known properties that are relevant in our context:

(S1) $\#\mathcal{S} \leq \binom{n}{2}$ holds for every affine split system $\mathcal{S}$. This follows from the bound on the number of cells generated by a configuration of n lines, first discussed by Steiner [10].
(S2) Every affine split system is contained in an affine split system of cardinality $\binom{n}{2}$.

(S3) The affine split systems $\mathcal{S} \subseteq \mathcal{S}(X)$ of cardinality $\binom{n}{2}$ are exactly the split systems of the form $\mathcal{S} = \mathcal{S}_\varphi$ for some map $\varphi : X \to \mathbb{E}^2$ that maps X injectively onto a set "in general position" in $\mathbb{E}^2$, i.e. onto a subset of $\mathbb{E}^2$ such that no straight line $\ell \subset \mathbb{E}^2$ contains more than two points from that set.

Regarding such split systems, we prove

Theorem 4. *The collection $\{d_S : S \in \mathcal{S}\}$ of split metrics associated with an affine split system $\mathcal{S}$ of maximal cardinality $\binom{n}{2}$ is a $\mathbb{Z}$-basis of $\mathcal{D}_{\text{even}}(X \mid \mathbb{Z})$. Equivalently,*

$$\det(\mathcal{S}) = 2^{(n-1)(n-2)} \tag{1}$$

holds for every affine split system $\mathcal{S}$ of maximal cardinality.

In particular, any affine split system is $\mathcal{D}$-independent.

In [1], it was shown that any weakly compatible split system $\mathcal{S} \subseteq \mathcal{S}^*(X)$ of cardinality $\binom{n}{2}$ is a *circular* split system, meaning (in the above terminology) that it is a maximal affine split system arising from an embedding $\varphi : X \to \mathbb{E}^2$ mapping X bijectively onto the vertices of a convex n-gon or, equivalently, that it is a maximal affine split system that contains all trivial splits of X. Thus, the identity (1) holds in particular (as already mentioned above) for weakly compatible splits of maximal cardinality.

2. $\mathcal{D}$-independence of split systems

Given any two finite sets U and V, we define a (real-valued) $U \times V$-matrix to be any real-valued matrix whose rows are indexed by the elements in U and whose columns are indexed by the elements in V (or, equivalently, any map from the set $U \times V$ into $\mathbb{R}$).

Let $\mathcal{S} \subseteq \mathcal{S}(X)$ be a split system of cardinality m and let $\mathbf{A} = \mathbf{A}(\mathcal{S})$ be the binary $\binom{X}{2} \times \mathcal{S}$-matrix defined by

$$\mathbf{A}_{\{x,y\},S} := \delta_{S(x),\overline{S}(y)} = d_S(x, y) \tag{2}$$

for all $\{x, y\} \in \binom{X}{2}$ and $S \in \mathcal{S}$. Then, the associated matrix $\mathbf{A}^T\mathbf{A}$ is a positive semi-definite matrix that is positive definite if and only if $\mathbf{A}^T\mathbf{A}$ is non-singular, if and only if $\mathbf{A}$ has rank m and, hence, if and only if $\mathcal{S}$ is $\mathcal{D}$-independent. So, all we need to observe is that the two matrices $\mathbf{M}(\mathcal{S})$ and $\mathbf{A}(\mathcal{S})^T\mathbf{A}(\mathcal{S})$ coincide for every split system $\mathcal{S} \subseteq \mathcal{S}(X)$ which follows, of course, immediately from our definitions: For any two splits $S = \{A, B\}$ and $S' = \{A', B'\}$ in $\mathcal{S}$, we have

$$\begin{aligned}
\left(\mathbf{A}(\mathcal{S})^T\mathbf{A}(\mathcal{S})\right)_{S,S'} &= \sum_{\{x,y\}\subseteq X} d_S(x, y)d_{S'}(x, y) \\
&= \#\left\{\{x, y\} \subseteq X :\ \text{both } S \text{ and } S' \text{ separate } x, y\right\} \\
&= \#(A \cap A')\#(B \cap B') + \#(A \cap B')\#(B \cap A') \\
&= \mu(S, S') \\
&= \mathbf{M}(\mathcal{S})_{S,S'}.
\end{aligned}$$

For the second characterization of independence (Theorem 2), we apply a version of the curiously ubiquitous *Farris Transform* [7,5]. This transform appears in many different guises and formulations across several fields (see, for example, the *covariance mapping* of [4]). Choose any element $z \in X$ and let $\mathcal{M}(X{:}\,z \mid \mathbb{R})$ denote the space consisting of all symmetric maps from $X \times X$ into $\mathbb{R}$ that vanish on all pairs $(x, y) \in X \times X$ with $z \in \{x, y\}$ (but not necessarily in the case $x = y \neq z$). Consider the map

$$\xi_z : \mathcal{D}(X \,|\, \mathbb{R}) \to \mathcal{M}(X{:}\,z \mid \mathbb{R})$$

from the space $\mathcal{D}(X \,|\, \mathbb{R})$ into $\mathcal{M}(X{:}\,z \mid \mathbb{R})$, which is defined by putting

$$\xi_z(D)(x, y) := \frac{1}{2}\left(D(x, z) + D(y, z) - D(x, y)\right)$$

for every map $D \in \mathcal{D}(X \,|\, \mathbb{R})$ and all $x, y \in X$. Clearly, ξ_z is a linear isomorphism from $\mathcal{D}(X \,|\, \mathbb{R})$ onto $\mathcal{M}(X{:}\,z \mid \mathbb{R})$ whose inverse ξ_z^{-1} maps any M in $\mathcal{M}(X{:}\,z \mid \mathbb{R})$ onto the symmetric map $\xi_z^{-1}(M) \in \mathcal{D}(X \,|\, \mathbb{R})$ defined by putting

$$\xi_z^{-1}(M)(x, y) := M(x, x) + M(y, y) - 2M(x, y)$$

for any map $M \in \mathcal{M}(X{:}\,z \mid \mathbb{R})$ and all $x, y \in X$.

Now, fix $z \in X$, put $X' := X - \{z\}$, and let $\binom{X'}{\leq 2}$ denote the set of all non-empty subsets $\{x', y'\} \subseteq X'$ of cardinality ≤ 2. Further, given any split system $\mathcal{S} \subseteq \mathcal{S}(X)$ of cardinality, say, m as above, let $\mathbf{B} = \mathbf{B}(\mathcal{S}, z)$ denote the binary $\binom{X'}{\leq 2} \times \mathcal{S}$-matrix defined by

$$\mathbf{B}_{\{x',y'\},S} := \xi_z(d_S)(x', y') \tag{3}$$

for all $x', y' \in X'$ and $S \in \mathcal{S}$.

Just as above, it is a consequence of well-known facts in linear algebra that the associated matrix $\mathbf{B}^T\mathbf{B}$ is always positive semi-definite, and that $\mathbf{B}^T\mathbf{B}$ is positive definite if and only if $\mathbf{B}^T\mathbf{B}$ is non-singular if and only if $\mathbf{B}$ has rank m if and only if the family $(\xi_z(d_S)_{S\in\mathcal{S}})$ is linearly independent and, hence (as ξ_z is an isomorphism from $\mathcal{D}(X \mid \mathbb{R})$ onto $\mathcal{M}(X{:}\,z \mid \mathbb{R})$), if and only if $\mathcal{S}$ is $\mathcal{D}$-independent. So, again, all we need to observe is that the two matrices $\mathbf{M}_z(\mathcal{S})$ and $\mathbf{B}(\mathcal{S})^T\mathbf{B}(\mathcal{S})$ coincide for every split system $\mathcal{S} \subseteq \mathcal{S}(X)$ which, we will see, follows from the following identity:

Lemma 5. *For any split S, we have*

$$\delta_{S(z),\overline{S}(x)}\,\delta_{S(z),\overline{S}(y)} = \frac{1}{2}(\delta_{S(z),\overline{S}(x)} + \delta_{S(z),\overline{S}(y)} - \delta_{S(x),\overline{S}(y)})$$

or, equivalently,

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

It follows that

$$\begin{aligned}\mathbf{B}_{\{x',y'\},S} &= \frac{1}{2}\left(d_S(x', z) + d_S(y', z) - d_S(x', y')\right)\\ &= d_S(z, x')\,d_S(z, y')\\ &= \begin{cases} 1 & \text{if } x', y' \in \overline{S}(z);\\ 0 & \text{otherwise}\end{cases}\end{aligned}$$

holds for all $x', y' \in X'$ and $S \in \mathcal{S}$ and, hence,

$$\begin{aligned}(\mathbf{B}^T\,\mathbf{B})_{S,S'} &= \sum_{x',y'\in X'} \mathbf{B}_{\{x',y'\},S}\,\mathbf{B}_{\{x',y'\},S'}\\ &= \sum_{x',y'\in \overline{S}(z)\cap\overline{S'}(z)} 1\\ &= \binom{|\overline{S}(z)\cap\overline{S'}(z)|}{2} + |\overline{S}(z)\cap\overline{S'}(z)|\\ &= \binom{|\overline{S}(z)\cap\overline{S'}(z)|+1}{2}\\ &= \mu_z(S,S') = \mathbf{M}_z(\mathcal{S})_{S,S'}\end{aligned}$$

for any two splits S, S' in $\mathcal{S}$, implying that also

$$\mathbf{B}(\mathcal{S},z)^T\,\mathbf{B}(\mathcal{S},z) = \mathbf{M}_z(\mathcal{S})$$

and

$$\det_z(\mathcal{S}) = \det \mathbf{M}_z(\mathcal{S}) = \det\Big(\mathbf{B}(\mathcal{S},z)^T\,\mathbf{B}(\mathcal{S},z)\Big)$$

holds. This establishes Theorem 2.

3. Split systems of cardinality $\binom{n}{2}$

Let us first note that, given any split system $\mathcal{S} \subseteq \mathcal{S}(X)$ and any element $z \in X$, one has

$$\mathbf{A}(\mathcal{S}) = \mathbf{\Xi}(\mathcal{S},z)\,\mathbf{B}(\mathcal{S},z) \tag{4}$$

for the square $\binom{X}{2} \times \binom{X-\{z\}}{\le 2}$-matrix $\mathbf{\Xi} = \mathbf{\Xi}(\mathcal{S},z)$ describing the map ξ_z in terms of the "canonical" bases of the vectorspaces $\mathcal{D}(X\,|\,\mathbb{R})$ and $\mathcal{M}(X{:}\,z\mid\mathbb{R})$ that is defined by

$$\begin{aligned}\mathbf{\Xi}_{\{x,y\},\{x',y'\}} &:= \delta_{x,x'}\delta_{x,y'} + \delta_{y,x'}\delta_{y,y'} - 2\delta_{\{x,y\},\{x',y'\}}\\ &= \begin{cases} 1 & \text{in case } x = x' = y',\\ 1 & \text{in case } y = x' = y',\\ -2 & \text{in case } \{x,y\} = \{x',y'\},\\ 0 & \text{otherwise}\end{cases}\end{aligned}$$

for all $\{x,y\} \in \binom{X}{2}$ and all $x', y' \in X' := X - \{z\}$. Indeed, using Lemma 5 for the last step in the following computation, it is easily seen that

$$\begin{aligned}(\mathbf{\Xi}\,\mathbf{B})_{\{x,y\},S} &= \sum_{x',y'\in X'} \mathbf{\Xi}_{\{x,y\},\{x',y'\}}\mathbf{B}_{\{x',y'\},S}\\ &= \sum_{x',y'\in X'} \delta_{S(z),\overline{S}(x')}\,\delta_{S(z),\overline{S}(y')}\,(\delta_{x,x'}\delta_{x,y'} + \delta_{y,x'}\delta_{y,y'})\\ &\quad - 2\sum_{x',y'\in X'} \delta_{S(z),\overline{S}(x')}\,\delta_{S(z),\overline{S}(y')}\delta_{\{x,y\},\{x',y'\}}\\ &= \delta_{S(z),\overline{S}(x)} + \delta_{S(z),\overline{S}(y)} - 2\,\delta_{S(z),\overline{S}(x)}\,\delta_{S(z),\overline{S}(y)}\\ &= \delta_{S(x),\overline{S}(y)}\end{aligned}$$

holds for all $\{x,y\} \in \binom{X}{2}$ and every split $S \in \mathcal{S}$.

Let us now consider the case $|\mathcal{S}| = \binom{n}{2}$, in which case both matrices $\mathbf{A}$ and $\mathbf{B}$ are square matrices, implying not only that

$$\det(\mathcal{S}) = \det(\mathbf{A}^T\,\mathbf{A}) = \det(\mathbf{A})^2$$

and

$$\det_z(\mathcal{S}) = \det(\mathbf{B}^T\,\mathbf{B}) = \det(\mathbf{B})^2,$$

but also — in view of (4) — that

$$\det(\mathbf{A}) = \det(\Xi)\det(\mathbf{B})$$

holds, provided that, for evaluating these three determinants, the indices are ordered consistently for all three matrices; otherwise, the left-hand side and the right-hand side may differ in their signs which, however, is irrelevant for the squares of these matrices considered in Theorem 3.

The rows and columns of Ξ can be ordered so that Ξ has the form

$$\Xi = \begin{bmatrix} -2\mathbf{I} & \mathbf{0} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}$$

consisting of

- one square $\binom{X'}{2} \times \binom{X'}{2}$-block whose rows and columns are indexed by the 2-subsets of X', which is just -2 times the identity matrix;
- one rectangular $\binom{X'}{2} \times X'$-block whose rows are indexed by the 2-subsets of X' and whose columns are indexed by the elements in X', which is simply the 0-matrix;
- one rectangular block $\mathbf{G}$ whose rows are indexed by the 2-subsets of X containing the element z and whose columns are indexed by the 2-subsets of X' whose explicit form, however, is fortunately of no concern; and
- one square block whose rows are indexed by the 2-subsets of X containing the element z and whose columns are indexed by the elements in X' which, upon identifying these two sets in the canonical way, is the identity matrix.

So, we see immediately that Ξ has determinant $\pm 2^{\binom{n-1}{2}}$.

The fact that $2^{(n-1)(n-2)}$ always divides $\det(\mathcal{S})$ when $\mathcal{S}$ has cardinality $\binom{n}{2}$ can also be derived by employing a more conceptual point of view.

Let $\mathcal{D}(X\,|\,\mathbb{Z})$ denote the $\mathbb{Z}$-lattice consisting of all integer-valued maps in $\mathcal{D}(X\,|\,\mathbb{R})$ and recall that $\mathcal{D}_{\text{even}}(X\,|\,\mathbb{Z})$ denotes its sublattice consisting of all maps $D \in \mathcal{D}(X\,|\,\mathbb{Z})$ for which the sum $D(x, y) + D(y, z) + D(z, x)$ is an even integer for all $x, y, z \in X$. For each 2-subset $\{x, y\}$ of X, let $\mathbf{D}_{\{x,y\}}$ denote the map from $X \times X$ into $\mathbb{Z}$ that maps the pairs (x, y) and (y, x) onto 1, and all other pairs in $X \times X$ onto 0, and note that $\mathcal{D}(X\,|\,\mathbb{Z})$ is freely generated (as a $\mathbb{Z}$-module) by the collection $\left\{\mathbf{D}_{\{x,y\}} : \{x, y\} \in \binom{X}{2}\right\}$ of all such maps. Further, we denote by $\mathbf{D}_x$, for every $x \in X$, the sum $\sum_{y \in X - \{x\}} \mathbf{D}_{\{x,y\}}$ and note the following:

(E1) $\mathbf{D}_x = d_{\{\{x\}, X - \{x\}\}} \in \mathcal{D}_{\text{even}}(X\,|\,\mathbb{Z})$ holds for every $x \in X$;

(E2) Given any element $z \in X$ and putting $X' := X - \{z\}$ as above, the subset

$$\mathcal{B}_z := \left\{2\,\mathbf{D}_{\{x,y\}} : \{x, y\} \in \binom{X'}{2}\right\} \cup \{\mathbf{D}_x : x \in X'\}$$

of $\mathcal{D}_{\text{even}}(X\,|\,\mathbb{Z})$ is actually a $\mathbb{Z}$-basis of $\mathcal{D}_{\text{even}}(X\,|\,\mathbb{Z})$ for every $z \in X$ (indeed, with z and X' as above, one has

$$D = \sum_{x \in X'} D(x,z)\,\mathbf{D}_x + \sum_{\{x,y\} \in \binom{X'}{2}} \frac{1}{2}\,(D(x,y) - D(x,z) - D(y,z))\ (2\mathbf{D}_{\{x,y\}})$$

for every $D \in \mathcal{D}_{\text{even}}(X\,|\,\mathbb{Z})$, and this representation of D as a $\mathbb{Z}$-linear combination of the maps in $\mathcal{B}_z$ is, of course, unique; see also [4]);

(E3) The subset

$$\mathcal{B}^z := \left\{\mathbf{D}_{\{x,y\}} : \{x,y\} \in \binom{X'}{2}\right\} \cup \{\mathbf{D}_x : x \in X'\}$$

of $\mathcal{D}(X\,|\,\mathbb{Z})$ is a $\mathbb{Z}$-basis of $\mathcal{D}(X\,|\,\mathbb{Z})$ for every $z \in X$ (indeed, one has

$$D = \sum_{x \in X'} D(x,z)\,\mathbf{D}_x + \sum_{\{x,y\} \in \binom{X'}{2}} (D(x,y) - D(x,z) - D(y,z))\ \mathbf{D}_{\{x,y\}}$$

for every $D \in \mathcal{D}(X\,|\,\mathbb{Z})$, and also this representation of D as a $\mathbb{Z}$-linear combination of the maps in $\mathcal{B}^z$ is, of course, unique).

Altogether, this implies that the factor module $\mathcal{D}(X \mid \mathbb{Z})/\mathcal{D}_{\text{even}}(X \mid \mathbb{Z})$ of $\mathcal{D}(X \mid \mathbb{Z})$ is an elementary abelian 2-group of rank $\binom{n-1}{2}$. Consequently, if $D_1, \ldots, D_{\binom{n}{2}}$ is any family of $\binom{n}{2}$ maps in $\mathcal{D}_{\text{even}}(X\,|\,\mathbb{Z})$, the power $2^{\binom{n-1}{2}}$ must divide the index of the sublattice generated by this family in $\mathcal{D}(X\,|\,\mathbb{Z})$ and, therefore, also the absolute value of the determinant

$$\det(D_1, \ldots, D_{\binom{n}{2}})$$

of the $\binom{X}{2} \times \{1, \ldots, \binom{n}{2}\}$-matrix $(D_1, \ldots, D_{\binom{n}{2}})$ whose entry at position $(\{x,y\}, i)$ is $D_i(x,y)$ while

$$|\det(D_1, \ldots, D_{\binom{n}{2}})| = 2^{\binom{n-1}{2}}$$

holds if and only if the family $D_1, \ldots, D_{\binom{n}{2}}$ is a $\mathbb{Z}$-basis of $\mathcal{D}_{\text{even}}(X\,|\,\mathbb{Z})$.

In particular, as $d_S \in \mathcal{D}_{\text{even}}(X\,|\,\mathbb{Z})$ holds for every split $S \in \mathcal{S}(X)$,

$$2^{\binom{n-1}{2}} \,\Big|\, \det(\mathbf{A}(\mathcal{S}))$$

must hold for every split system $\mathcal{S} \subseteq \mathcal{S}(X)$ of cardinality $\binom{n}{2}$, and

$$2^{\binom{n-1}{2}} = |\det(\mathbf{A}(\mathcal{S}))|$$

holds for a split system $\mathcal{S} \subseteq \mathcal{S}(X)$ of cardinality $\binom{n}{2}$ if and only if the split metrics $\{d_S : S \in \mathcal{S}\}$ form a $\mathbb{Z}$-basis of $\mathcal{D}_{\text{even}}(X\,|\,\mathbb{Z})$.

This argument actually shows that the following considerably stronger and more explicit assertion regarding the integer matrix $\mathbf{A}$ holds (see, for instance, [8] for the relevant definitions):

Theorem 6. *If $\mathcal{S} \subseteq \mathcal{S}(X)$ is any split system of cardinality $\binom{n}{2}$, then there are at most 2^{n-1} $\left(= 2^{\binom{n}{2}-\binom{n-1}{2}}\right)$ elementary divisors among the $2^{\binom{n}{2}}$ elementary divisors of the matrix* **A** *that are not divisible by 2.*

4. Affine split systems

Let us now consider an affine split system $\mathcal{S} \subseteq \mathcal{S}(X)$ of maximal cardinality $\binom{n}{2}$. The goal of this section is to show that, in this case, $\det(\mathcal{S}) = 2^{(n-1)(n-2)}$ holds.

We begin by studying split systems with a particular separation property, a property that is easily seen to hold for maximal affine split systems (see Lemma 9 below).

Definition 2. A split system $\mathcal{S} \subseteq \mathcal{S}(X)$ has the *pairwise separation property* if, for any two distinct points $x, y \in X$, there are two (possibly empty) subsets A, B of $X - \{x, y\}$ such that the four splits

$$\begin{aligned} S_1 &= \{A \cup \{x, y\}, B\} \\ S_2 &= \{A \cup \{x\}, B \cup \{y\}\} \\ S_3 &= \{A \cup \{y\}, B \cup \{x\}\} \\ S_4 &= \{A, B \cup \{x, y\}\} \end{aligned} \tag{5}$$

are all in $\mathcal{S}' := \mathcal{S} \cup \{\{\emptyset, X\}\}$.

Lemma 7. *Let $\mathcal{S} \subseteq \mathcal{S}(X)$ be a split system that satisfies the pairwise separation property. Then,*

$$\langle \mathcal{S} \rangle_{\mathbb{Z}} = \langle \mathcal{S}' \rangle_{\mathbb{Z}} = \mathcal{D}_{\text{even}}(X \,|\, \mathbb{Z})$$

holds.

Proof. Note first that $\langle \mathcal{S} \rangle_{\mathbb{Z}} = \langle \mathcal{S}' \rangle_{\mathbb{Z}}$ holds for any split system $\mathcal{S}$, in view of the fact that the split metric $d_{\{\emptyset, X\}}$ associated with the split $\{\emptyset, X\}$ is the 0-map.

Next, choose any two distinct points $x, y \in X$ and, using the pairwise separation property, a partition $\{A, B\}$ of $X - \{x, y\}$ such that the splits $S_1, \ldots, S_4$ defined in (5) are all in $\mathcal{S}'$, and note that

$$d_{S_2} + d_{S_3} - d_{S_1} - d_{S_4} = 2\,\mathbf{D}_{\{x,y\}}$$

always holds. So,

$$2\,\mathbf{D}_{\{x,y\}} \in \langle \mathcal{S} \rangle_{\mathbb{Z}}$$

holds for any two distinct points $x, y \in X$. Also, we have

$$d_{S_3} - d_{S_1} = \sum_{z \in A \cup \{y\}} \mathbf{D}_{\{x,z\}} - \sum_{z \in B} \mathbf{D}_{\{x,z\}} = \mathbf{D}_x - 2 \sum_{z \in B} \mathbf{D}_{\{x,z\}},$$

implying that, in view of the fact that has just been established that $2\,\mathbf{D}_{\{x,z\}} \in \langle \mathcal{S} \rangle_{\mathbb{Z}}$ holds for all $\{x, z\} \in \binom{X}{2}$, also

$$\mathbf{D}_x \in \langle \mathcal{S} \rangle_{\mathbb{Z}}$$

must hold for all $x \in X$. So, our claim $\langle \mathcal{S} \rangle_{\mathbb{Z}} = \mathcal{D}_{\text{even}}(X \mid \mathbb{Z})$ follows from Assertion (E2) in Section 3. $\square$

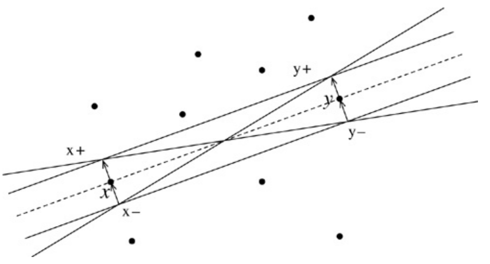

Fig. 2. Illustration of the pairwise separation property for affine splits: the four solid lines induce the splits $S_1, \ldots, S_4$ of (5).

Corollary 8. *If $\mathcal{S} \subseteq \mathcal{S}(X)$ is a split system of cardinality $\binom{n}{2}$ that satisfies the pairwise separation property, then*

$$\det(\mathcal{S}) = 2^{(n-1)(n-2)}$$

must hold.

Note that the pairwise separation property is not a *necessary* condition for a split system $\mathcal{S}$ of cardinality $\binom{n}{2}$ to have determinant $\det(\mathcal{S}) = 2^{(n-1)(n-2)}$. A counter-example will be provided in the last section.

As we will see now, we can apply these results to affine split systems:

Lemma 9. *Every affine split system $\mathcal{S} \subseteq \mathcal{S}^*(X)$ of cardinality $\binom{n}{2}$ satisfies the pairwise separation property.*

Proof. Choose an $\mathcal{S}$-map φ, i.e. a map $\varphi : X \to \mathbb{E}^2$ with $\mathcal{S} = \mathcal{S}_\varphi$, and recall that, according to (S3), any such map must be injective and that — denoting, for any two distinct points $x, y \in X$, the straight line connecting the two points $\varphi(x)$ and $\varphi(y)$ by $\overleftrightarrow{x : y}$ — we must have

$$\overleftrightarrow{x : y} \cap \varphi(X) = \{\varphi(x), \varphi(y)\}$$

for any two distinct points $x, y \in X$. Next, let

$$S(\varphi : x, y) := S\left(\varphi\,|_{X-\{x,y\}}, \overleftrightarrow{x : y}\right)$$

denote the associated split of $X - \{x, y\}$ induced by the straight line $\overleftrightarrow{x : y}$, choose $A, B \subseteq X - \{x, y\}$ with $S(\varphi : x, y) = \{A, B\}$, and put

$$\mathcal{S}_+(\varphi : x, y) := \{\{A \cup \{x\}, B \cup \{y\}\}, \{A \cup \{y\}, B \cup \{x\}\}\}$$

and

$$\mathcal{S}_-(\varphi : x, y) := \{\{A \cup \{x, y\}, B\}, \{A, B \cup \{x, y\}\}\}.$$

We claim that the four splits in $\mathcal{S}_+(\varphi : x, y) \cup \mathcal{S}_-(\varphi : x, y)$ are all in $\mathcal{S} = \mathcal{S}_\varphi$. To show this, choose a small non-zero vector ϵ orthogonal to the line $\overleftrightarrow{x : y}$ and note that we can choose ϵ sufficiently small so that the four lines connecting each of the two points $\varphi(x) \pm \epsilon$ with each of the two points $\varphi(y) \pm \epsilon$ all have an empty intersection with X and induce the same split $S(\varphi : x, y)$ on $X - \{x, y\}$. Therefore, these four lines induce the four required splits (Fig. 2). $\square$

Taken together, these facts establish Theorem 4. More specifically, the above arguments imply that, given any affine split system $\mathcal{S}$ of maximal cardinality $\binom{n}{2}$ and any $\mathcal{S}$-map $\varphi : X \to \mathbb{E}^2$, one has

$$\mathbf{D}_{\{x,y\}} = \sum_{S\in\mathcal{S}} \frac{1}{2}\,\eta(S : x, y)\; d_S$$

for all $\{x, y\} \in \binom{X}{2}$, where the coefficients $\eta(S : x, y)$ are defined, for all $S \in \mathcal{S}$ and all $\{x, y\} \in \binom{X}{2}$, by

$$\eta(S : x, y) = \eta(S : x, y)_\varphi := \begin{cases} 1 & \text{in case } S \in \mathcal{S}_+(\varphi : x, y),\\ -1 & \text{in case } S \in \mathcal{S}_-(\varphi : x, y),\\ 0 & \text{otherwise.} \end{cases}$$

Therefore, also

$$D = \sum_{\{x,y\}\in\binom{X}{2}} D(x, y)\mathbf{D}_{\{x,y\}} = \sum_{S\in\mathcal{S}} \left(\sum_{\{x,y\}\in\binom{X}{2}} \frac{1}{2}\,\eta(S : x, y)\, D(x, y) \right) d_S \tag{6}$$

holds for every map $D \in \mathcal{D}(X\,|\,\mathbb{R})$, implying in particular that

$$\sum_{\{x,y\}\in\binom{X}{2}} \eta(S : x, y)\, D(x, y)$$

must be an even integer for every split $S \in \mathcal{S}$ and every map $D \in \mathcal{D}_{\text{even}}(X\,|\,\mathbb{Z})$ — a fact that can, of course, also be derived easily by direct computation.

Note that $\eta(S : x, y) = 1$ (or $\eta(S : x, y) = -1$) holds for some $S \in \mathcal{S}$ and some $\{x, y\} \in \binom{X}{2}$ if and only if (i) the split S separates any two points $u, v \in X - \{x, y\}$ if and only if the straight line segment $[\varphi(u), \varphi(v)]$ intersects $\overleftrightarrow{x : y}$ and (ii) S separates the two points x and y (or does not separate these two points, respectively).

5. Discussion

In this section, we discuss some specific cases as well as some questions that arise naturally in this context, providing examples (or counter-examples) when available.

5.1. *Existence of maximal split systems with* $\det(\mathcal{S}) \neq 2^{(n-1)(n-2)}$

To begin with, it is easy to check that every split system $\mathcal{S} \subseteq \mathcal{S}(X)$ is affine in the case $n \leq 3$ and that a split system $\mathcal{S} \subseteq \mathcal{S}(X)$ is affine in the case $n = 4$ if and only if $\#\mathcal{S} \leq 6 = \binom{4}{2}$ holds. For every $n \geq 5$, there exist split systems $\mathcal{S} \subseteq \mathcal{S}(X)$ of cardinality $\binom{n}{2}$ with $\det(\mathcal{S}) > 2^{(n-1)(n-2)}$. Consider, for example, the system $\mathcal{S}_2 \subseteq \mathcal{S}(X)$ consisting of all “2-splits” of X, i.e. put

$$S_{uv} := \{\{u, v\}, X - \{u, v\}\}$$

for all $u, v \in X$ and consider the split system

$$\mathcal{S}_2 = \mathcal{S}_2(X) := \left\{ S_{xy} : \{x, y\} \in \binom{X}{2} \right\}.$$

It is easy to see that $\mathcal{S}_2$ is a $\mathcal{D}$-independent split system of cardinality $\binom{n}{2}$ for every n that is distinct from 2 and 4. More specifically, we have

Theorem 10. *One has*

$$\det(\mu_{S_{xy}, S_{x'y'}})_{\{x,y\},\{x',y'\} \in \binom{X}{2}} = (n-2)^2 \, (4-n)^{2n-2} \, 2^{(n-1)(n-2)} \tag{7}$$

for all $n \geq 2$ and, therefore,

$$\det(\mathcal{S}_2) = (n-2)^2 \, (4-n)^{2n-2} \, 2^{(n-1)(n-2)} \tag{8}$$

for all $n \geq 2$ except $n = 4$ (in which case, every 2-split is counted twice in the family $(S_{xy})_{\{x,y\} \in \binom{X}{2}}$).

In particular, there are split systems of cardinality $\binom{n}{2}$ with a non-vanishing determinant distinct from $2^{(n-1)(n-2)}$ for all $n \geq 5$.

Proof. Given a fixed element $z \in X$, put $X' := X - \{z\}$ as above, and consider the $\binom{X'}{\leq 2} \times \binom{X'}{\leq 2}$-matrix $\mathbf{B} = \mathbf{B}(2, z)$ defined in analogy to (3) by

$$\mathbf{B}_{\{x,y\},\{u,v\}} := \begin{cases} \mathbf{B}_{\{x,y\}, S_{uv}} & \text{if } u \neq v, \\ \mathbf{B}_{\{x,y\}, S_{uz}} & \text{otherwise.} \end{cases}$$

For convenience, we identify the subset $\{\{x\} : x \in X'\}$ of $\binom{X'}{\leq 2}$ consisting of all one-element subsets of X' with X' in the obvious way.

Clearly, $\mathbf{B}$ can be viewed as a 2×2 block matrix

$$\mathbf{B} = \begin{bmatrix} \mathbf{I} & \mathbf{H} \\ \mathbf{K} & \mathbf{V} \end{bmatrix}$$

consisting of

- the $\binom{X'}{2} \times \binom{X'}{2}$ identity matrix $\mathbf{I}$;
- the $\binom{X'}{2} \times X'$ matrix $\mathbf{H}$ defined by

  $$\mathbf{H}_{\{x,y\},u} = \begin{cases} 1 & \text{if } u \notin \{x, y\}, \\ 0 & \text{otherwise} \end{cases}$$

  for all $\{x, y\} \in \binom{X'}{2}$ and $u \in X'$;
- the $X' \times \binom{X'}{2}$ matrix $\mathbf{K}$ defined by

  $$\mathbf{K}_{u,\{x,y\}} = 1 - \mathbf{H}_{\{x,y\},u} = \begin{cases} 1 & \text{if } u \in \{x, y\}, \\ 0 & \text{otherwise} \end{cases}$$

  for all $u \in X'$ and $\{x, y\} \in \binom{X'}{2}$; and
- the $X' \times X'$ matrix $\mathbf{V}$ which has zeros on the diagonal and ones everywhere else.

Using elementary row manipulations, we can use the $\binom{X'}{2} \times \binom{X'}{2}$ identity submatrix to zero the entries of $\mathbf{K}$, replacing $\mathbf{V}$ by the matrix $(4-n)\mathbf{V}$. Using further row and column manipulations (or noting that the constant vectors form a one-dimensional eigenspace of $\mathbf{V}$ with eigenvalue $n-2$ while the vectors with vanishing total entry sum form an $(n-2)$-dimensional eigenspace of $\mathbf{V}$ with eigenvalue -1), we see immediately that

$$\det(\mathbf{V}) = (-1)^{n-2}(n-2)$$

holds. Thus, arguing as above in the proof of Theorem 3, we see that also

$$\det(\mu_{S_{xy},S_{x'y'}})_{\{x,y\},\{x',y'\}\in\binom{X}{2}} = (n-2)^2\,(4-n)^{2n-2}\,2^{(n-1)(n-2)}$$

holds for all $n \geq 2$, and

$$\begin{aligned}\det(\mathcal{S}_2) &= 2^{(n-1)(n-2)}\det(\mathbf{B}(2,z))^2\\ &= 2^{(n-1)(n-2)}\left((4-n)^{n-1}\,(-1)^{n-2}\,(n-2)\right)^2\\ &= (n-2)^2\,(n-4)^{2n-2}\,2^{(n-1)(n-2)}\end{aligned}$$

for all n distinct from 4, as claimed. □

Theorem 10 implies in particular that, for every natural number m, there exists a smallest number $n = n(m) \leq m+2$ such that a maximal independent split system $\mathcal{S}$ with $m^2 \mid \det(\mathcal{S})$ exists on a set of cardinality n. The fact that, for every set X of cardinality n, there exist only finitely many split systems $\mathcal{S} \subseteq \mathcal{S}(X)$ of cardinality $\binom{n}{2}$ implies, of course, that

$$\lim_{m\to\infty} n(m) = \infty$$

must hold. It might be interesting to determine the number $n = n(m)$, at least in the case that m is a prime number p and to check, for example, in particular whether $n(p) = p+2$ holds for every prime number p. We will see below that, for $p := 3$, $n(p) = p+2$ indeed holds as well as $n(q) > p+2 = 5$ for any prime q larger than $p = 3$, and that, in addition, $|\mathbf{B}(\mathcal{S},z)| \leq |\mathbf{B}(2,z)|$ and, therefore, $\det(\mathcal{S}) \leq \det(\mathcal{S}_2)$ also holds in the case $n := 5$ for every split system $\mathcal{S} \subset \mathcal{S}(\{1,2,3,4,5\})$ of cardinality 10, suggesting that these inequalities might hold for every $n \geq 5$.

There are, of course, further arithmetic invariants of the matrices $\mathbf{M}(\mathcal{S})$ considered as a matrix representing an integral positive semi-definite quadratic form, or — equivalently — of the $\mathbb{Z}$-lattices $\langle\mathcal{S}\rangle_{\mathbb{Z}}$ that might be worth to be studied.

5.2. *Pseudo-affine split systems*

It also seems to be of some interest to note that there is a natural generalization of the class of affine split systems, *viz* the class of *pseudo-affine* split systems. These can be defined simply as the class of split systems that arises as the set of (splits encoded by the) *topes* of an affine regular oriented matroid of rank 3 (i.e. a regular oriented matroid of rank 3 for which the "all-to-one" map is a tope) and which, therefore, clearly encompasses the class of affine split systems.

5.3. *Determinant preserving operations on split systems*

Next note that, while it can be shown that any split system of cardinality of at most 6 is $\mathcal{D}$-independent, there are split systems of cardinality 7 that are not $\mathcal{D}$-independent (provided that

$n \geq 4$ holds). Indeed, given any partition of X into four disjoint non-empty subsets A, B, C, D, we may consider the set $\mathcal{S}(A \,|\, B \,|\, C \,|\, D)$ consisting of altogether the seven splits

$$S_1 := \{A, X - A\}, \quad S_2 := \{B, X - B\}, \quad S_3 := \{C, X - C\}, \quad S_4 := \{D, X - D\},$$
$$S_5 := \{A \cup B, C \cup D\}, \quad S_6 := \{A \cup C, B \cup D\}, \quad S_7 := \{A \cup D, B \cup C\}.$$

Note that the identity

$$d_{S_1} + d_{S_2} + d_{S_3} + d_{S_4} = d_{S_5} + d_{S_6} + d_{S_7} \tag{9}$$

always holds. So, $\mathcal{S}(A \,|\, B \,|\, C \,|\, D)$ is never $\mathcal{D}$-independent. However, these sets provide a way of constructing new maximal split systems from old.

Theorem 11. *Suppose that $\mathcal{S}$ and $\mathcal{S}'$ are proper split systems of cardinality $\binom{n}{2}$, that*

$$\mathcal{S} - \mathcal{S}(A \,|\, B \,|\, C \,|\, D) = \mathcal{S}' - \mathcal{S}(A \,|\, B \,|\, C \,|\, D)$$

and that

$$\#(\mathcal{S} \cap \mathcal{S}(A \,|\, B \,|\, C \,|\, D)) = \#(\mathcal{S}' \cap \mathcal{S}(A \,|\, B \,|\, C \,|\, D)) = 6.$$

Then $\det(\mathcal{S}) = \det(\mathcal{S}')$.

Proof. Applying (9), we see immediately that

$$\langle \mathcal{S} \rangle_{\mathbb{Z}} = \langle \mathcal{S}' \rangle_{\mathbb{Z}}$$

from which the result follows. □

Hence, one can construct split systems $\mathcal{S}'$ of cardinality $\binom{n}{2}$ with $\det(\mathcal{S}') = 2^{(n-1)(n-2)}$ from other such split systems $\mathcal{S}$ by repeatedly exchanging any one split in $\mathcal{S}$ contained in a collection of splits of the form $\mathcal{S}(A \,|\, B \,|\, C \,|\, D)$ with $\#(\mathcal{S}(A \,|\, B \,|\, C \,|\, D) - \mathcal{S}) = 1$ by the (then unique!) split contained in $\mathcal{S}(A \,|\, B \,|\, C \,|\, D) - \mathcal{S}$.

It seems worth investigating whether every split system $\mathcal{S}$ of cardinality $\binom{n}{2}$ with $\det(\mathcal{S}) = 2^{(n-1)(n-2)}$ can be reached by a sequence of these and further similar operations from a split system $\mathcal{S}_0$ of that cardinality that is derived from an affine regular oriented matroid of rank 3.

5.4. Non-affine split system with $\det(\mathcal{S}) = 2^{(n-1)(n-2)}$

Here is, at least, one interesting example of a split system $\mathcal{S}$ of cardinality $\binom{n}{2}$ with $\det(\mathcal{S}) = 2^{(n-1)(n-2)}$ that does arise in this way, but it is neither affine nor pseudo-affine (as it contains all 1-splits and is not circular) and does not even satisfy the pairwise separation property. Consider the set $X := \{1, 2, 3, 4, 5\}$ and the embedding φ of X into $\mathbb{E}^2$ that maps the elements 1, 2, 3, 4 in clockwise fashion onto the vertices of a square in $\mathbb{E}^2$ while it maps the element 5 onto a point that is somewhere in the interior of the triangle whose vertices are the elements $\varphi(1)$, $\varphi(4)$, and the point where the two diagonals of the square $\varphi(1), \varphi(2), \varphi(3), \varphi(4)$ meet. Then $\mathcal{S}_\varphi$ consists of the four "1-splits"

$$\{\{1\}, \{2, 3, 4, 5\}\}, \quad \{\{2\}, \{3, 4, 5, 1\}\}, \quad \{\{3\}, \{4, 5, 1, 2\}\}, \quad \{\{4\}, \{5, 1, 2, 3, \}\},$$

the four 2-splits

$$\{\{1, 2\}, \{3, 4, 5\}\}, \quad \{\{2, 3\}, \{4, 5, 1\}\}, \{\{3, 4\}, \{5, 1, 2\}\}, \quad \{\{4, 1\}, \{2, 3, 5\}\},$$

and the two 2-splits

$$\{\{1,5\},\{2,3,4\}\} \quad \text{and} \quad \{\{4,5\},\{1,2,3\}\}.$$

So, we can apply our procedure relative to the partition of X into the four disjoint subsets $\{2,3\}, \{1\}, \{4\}, \{5\}$ and the split $\{\{2,3\},\{4,5,1\}\} \in \mathcal{S}_\varphi$, yielding a split system $\mathcal{S}$ with $\langle\mathcal{S}\rangle_\mathbb{Z} = \mathcal{D}_{\text{even}}(X \mid \mathbb{Z})$ that consists of all 1-splits and the five 2-splits

$$\{\{1,2\},\{3,4,5\}\}, \quad \{\{3,4\},\{5,1,2\}\}, \quad \{\{4,1\},\{2,3,5\}\},$$
$$\{\{1,5\},\{2,3,4\}\}, \quad \text{and} \quad \{\{4,5\},\{1,2,3\}\}$$

and is easily seen to be non-circular (as the graph with vertex set X and edge set the collection $\{\{1,2\},\{1,4\},\{1,5\},\{3,4\},\{4,5\}\}$ of 2-subsets of X contained in the 2-splits of $\mathcal{S}'$ is non-circular) and therefore, as it contains all 1-splits, neither affine nor pseudo-affine — a fact that, of course, also follows from the observation that $\mathcal{S}$ does not even satisfy the pairwise separation property, thus also providing the counter-example referred to above, which shows that the pairwise separation property is not necessary for a split system $\mathcal{S}$ of cardinality $\binom{n}{2}$ to have determinant $\det(\mathcal{S}) = 2^{(n-1)(n-2)}$. Indeed, consider the two points 1, 3 and note that there is no partition of $\{2,4,5\}$ into two disjoint subsets A, B such that all the four splits $\{\{1,3\} \cup A, B\}$, $\{\{1\} \cup A, \{3\} \cup B\}$, $\{\{3\} \cup A, \{1\} \cup B\}$, and $\{A, \{1,3\} \cup B\}$ are contained in $\mathcal{S}' \cup \{\{X, \emptyset\}\}$ as $\{A, B\} = \{\{2,4,5\}, \emptyset\}$ does not work in view of $\{\{1,3\},\{2,4,5\}\} \notin \mathcal{S}'$, and no choice $\{A, B\}$ with, say, $\#A = 1$ and $\#B = 2$ can work, as this would imply that

$$B \in \binom{\{2,4,5\}}{2} \cap \{\{1,2\},\{1,4\},\{1,5\},\{3,4\},\{4,5\}\}$$

and, therefore, $B = \{4,5\}$ must hold, which is impossible in view of

$$\{\{2,3\},\{1,4,5\}\} \notin \mathcal{S}.$$

5.5. Enumerating split systems on five points

More generally, there are — up to relabelling — 53 split systems with a cardinality of ten on a set of five points. They are not difficult to enumerate, since they correspond to the isomorphism classes of subgraphs with exactly ten edges of the "very complete graph" with five vertices, i.e. the graph with five vertices that contains exactly one edge or loop connecting any given vertex v with any given vertex w (whether distinct from v or not) or, equivalently, the isomorphism classes for simple graphs with six vertices, one of which is distinguished. We can further subdivide these 53 split systems by their determinant and number of 1-splits (see Table 1).

We see that affine split systems provided only a few of the systems with determinant $2^{(n-1)(n-2)} = 4096$. There are three affine split systems on five points (up to relabelling), with three, four, or five trivial splits (1-splits). However there are *thirty* non-equivalent split systems with ten splits on five points and determinant $2^{(n-1)(n-2)} = 4096$.

5.6. Split systems and least squares

Finally, recall that, given any $\mathcal{D}$-independent split system $\mathcal{S}$ of cardinality, say, m and any map $D \in \mathcal{D}(X \mid \mathbb{R})$, there exists a unique map $D_\mathcal{S} \in \langle\mathcal{S}\rangle_\mathbb{R}$ such that the Euclidian distance

Table 1
The numbers of distinct split systems (up to re-labelling) with ten splits on five points, distinguished first by the number of trivial splits (1-splits), then by the absolute value of the determinants of the matrices $\mathbf{B}_z$

# 1-splits	$\lvert\det(\mathbf{B}_z)\rvert$	Number of classes
0	3	1
1	1	1
	2	1
2	0	3
	1	6
	2	1
3	0	6
	1	10
	2	1
	3	1
4	0	6
	1	9
	2	1
5	0	1
	1	4

$$\|D, D_{\mathcal{S}}\|_2 := \sqrt{\sum_{\{x,y\}\in\binom{X}{2}} (D(x,y) - D_{\mathcal{S}}(x,y))^2}$$

between D and $D_{\mathcal{S}}$ minimizes the euclidian distance $\|D, D'\|_2$ of D to all the maps D' in the linear subspace $\langle\mathcal{S}\rangle_{\mathbb{R}}$ and that

$$D_{\mathcal{S}} = \sum_{S\in\mathcal{S}} \lambda_{(D,\mathcal{S})}(S)\, d_S$$

holds for the unique map $\lambda_{(D,\mathcal{S})} \in \mathbb{R}^{\mathcal{S}}$ for which

$$\sum_{S'\in\mathcal{S}} \mu(S,S')\lambda_{(D,\mathcal{S})}(S') = D(S) := \sum_{\{x,y\}\in\binom{X}{2}} D(x,y)d_S(x,y) = \sum_{x\in A, y\in B} D(x,y)$$

holds for every slit $S = \{A, B\} \in \mathcal{S}$, implying that $D_{\mathcal{S}}$ can be computed quite easily from D and the inverse $\mathbf{A}(\mathcal{S})^{-1}$ of the (non-singular!) matrix $\mathbf{A}(\mathcal{S})$ according to the formula

$$\lambda_{(D,\mathcal{S})}(S) = \sum_{S'\in\mathcal{S}} \mathbf{A}(\mathcal{S})^{-1}_{S,S'}\, D(S').$$

This simple — and very “classical” — fact has been used (though, perhaps, employing slightly different *façon de parler* in least squares methods for trees [3] and networks [2,11]). It was actually this specific application of the matrix $\mathbf{A}(\mathcal{S})$ that prompted the investigations communicated in the present note.

References

[1] H.-J. Bandelt, A.W.M. Dress, A canonical decomposition theory for metrics on a finite set, Adv. Math. 92 (1992) 47–105.

[2] D. Bryant, V. Moulton, NeighborNet: An agglomerative algorithm for the construction of planar phylogenetic networks, Mol. Biol. Evol. 21 (2004) 255–265.
[3] L.L. Cavalli-Sforza, A.W.F. Edwards, Phylogenetic analysis: Models and estimation procedures, Evolution 21 (3) (1967) 550–570.
[4] M. Deza, M. Laurent, Geometry of Cuts and Metrics, in: Algorithms and Combinatorics, vol. 15, Springer-Verlag, Berlin, 1997.
[5] A. Dress, B. Holland, K.T. Huber, J.H. Koolen, V. Moulton, J. Weyer-Menkhoff, Δ additive and Δ ultra-additive maps, Gromov's trees, and the Farris transform, Discrete Appl. Math. 146 (1) (2005) 51–73.
[6] H. Edelsbrunner, Algorithms in Combinatorial Geometry, in: EATCS Monographs on Theoretical Computer Science, vol. 10, Springer-Verlag, Berlin, 1987.
[7] J.S. Farris, A.G. Kluge, M.J. Eckart, A numerical approach to phylogenetic systematics, Systematic Zool. 19 (1970) 172–189.
[8] F. Lübeck, On the computation of elementary divisors of interger matrices, J. Symbolic Computation 33 (2002) 57–65.
[9] J.E. Goodman, R. Pollack, Semispaces of configurations, cell complexes of arrangements, J. Combin. Theory Ser. A 37 (3) (1984) 257–293.
[10] J. Steiner, Gesetze über die Theilung der Ebene und des Raumes, J. Reine Angew. Math. 1 (1826).
[11] R. Winkworth, D. Bryant, P. Lockhart, D. Havell, V. Moulton, Biogeographic interpretation of splits graphs: Least squares optimization of branch lengths, Syst. Biol. 54 (1) (2005) 56–65.