5. Grassmannians and the Space of Trees

In this lecture we shall be interested in a very particular ideal. The ambient polynomial ring $\mathbb{C}[p]$ has $\binom{n}{d}$ variables, which are called *Plücker coordinates*:

$$\mathbb{C}[p] \ := \ \mathbb{C}\big[\, p_{i_1 i_2 \cdots i_d} \,:\, 1 \leq i_1 < i_2 < \cdots < i_d \leq n \,\big].$$

The *Plücker ideal* is the homogeneous prime ideal in $\mathbb{C}[p]$ consisting of the algebraic relations among the $d \times d$-minors of an arbitrary $d \times n$-matrix. The ideal $I_{d,n}$ is generated by quadrics, and it has a well-known quadratic Gröbner basis (see e.g. [75, Theorem 3.1.7]. The projective variety of $I_{d,n}$ is the *Grassmannian* $G_{d,n}$ which parametrizes d-dimensional linear subspaces of $\mathbb{C}^n$. The Grassmannian $G_{d,n}$ is a smooth and irreducible variety of dimension $d(n-d)$. Hence $\dim(I_{d,n}) = d(n-d)+1$. The parametrization of d-dimensional subspaces of $\mathbb{C}^n$ by points p in $G_{n,d}$ works as follows: if a subspace is given as the row space of a $d \times n$-matrix then its Pücker coordinate vector p consists of the $d \times d$-minors of that matrix. This is unique up to scaling. Conversely, if p any point in the Grassmannian $G_{d,n}$ then the corresponding subspace equals

$$(1) \quad \Big\{ x \in \mathbb{R}^n : \sum_{r=0}^{d} (-1)^r \cdot p_{i_0 i_1 \cdots \widehat{i_r} \cdots i_d} \cdot x_{i_r} \ = \ 0 \ \text{ for all } 1 \leq i_0 < i_1 < \cdots < i_d \leq n \Big\}.$$

We shall be particularly interested in the case $d = 2$. Here the Grassmannian $G_{2,n}$ has dimension $2n-4$ and it parametrizes planes through the origin in $\mathbb{C}^n$, or, equivalently, $G_{2,n}$ parametrizes lines in the projective space $\mathbb{P}^{n-1}$.

Proposition 5.1. *The ideal $I_{2,n}$ is generated by the special quadrics*

$$(2) \qquad p_{ij}p_{kl} - p_{ik}p_{jl} + p_{il}p_{jk}, \quad 1 \leq i < j < k < l \leq n.$$

For suitable term order, these constitute the reduced Gröbner basis for $I_{2,n}$.

The *three-term Plücker relations* (2) can be generalized to $d > 2$ as follows:

$$(3) \qquad p_{ijS} \cdot p_{klS} \ - \ p_{ikS} \cdot p_{jlS} \ + \ p_{ilS} \cdot p_{jkS}$$

for $S \subset \{1, \ldots, n\}$ with $|S| = d-2$ and $i, j, k, l \in \{1, \ldots, n\} \backslash S$. Here we are employing the standard convention that Plücker coordinates $p_{i_1 i_2 \cdots i_d}$ for non-increasing index strings $i_1 i_2 \cdots i_d$ are defined by permuting that string to be increasing and multiplying with the sign of the permutation. For example,

$$p_{5342} \ = \ -p_{3542} \ = \ p_{3452} \ = \ \cdots \ = \ -p_{2345}.$$

In general, for $n \geq d+3 \geq 6$, the three-term Plücker relations (3) do not generate the ideal $I_{d,n}$. However, they do so when $I_{d,n}$ is extended to the Laurent polynomial ring $\mathbb{C}[p^{\pm}]$, so they will be close enough for our purposes.

We define the *tropical Grassmannian* $\mathcal{G}_{d,n}$ to be the tropical variety $\mathcal{T}(I_{d,n})$ specified by the Plücker ideal $I_{d,n}$. Thus $\mathcal{G}_{d,n}$ is a pure fan in the vector space $\mathbb{R}^{\binom{d}{n}}$ whose coordinates we now denote by $w_{i_1 i_2 \cdots i_d}$ for $1 \leq i_1 < i_2 < \cdots < i_d \leq$

n. The dimension of $\mathcal{G}_{d,n}$ equals $d(n-d)+1 = \dim(I_{d,n})$. Among the tropical hypersurfaces containing the tropical Grassmannian $\mathcal{G}_{d,n}$ are those defined by the three-term Plücker relations. The *tropical three-term Plücker relations* are

$$w_{ijS} \odot w_{klS} \ \oplus\ w_{ikS} \odot w_{jlS} \ \oplus\ w_{ilS} \odot w_{jkS} \tag{4}$$

where $|S| = d-2$ and $i,j,k,l \in \{1,\ldots,n\}\backslash S$. The intersection of the tropical hypersurfaces $\mathcal{T}(w_{ijS} \odot w_{klS} \oplus w_{ikS} \odot w_{jlS} \oplus w_{ilS} \odot w_{jkS})$ is a tropical prevariety, denoted *pre*$\mathcal{G}_{d,n}$, which is a combinatorial approximation to the tropical Grassmannian $\mathcal{G}_{d,n}$. For $n \geq d+4 \geq 7$ the approximation *pre*$\mathcal{G}_{d,n}$ is strictly larger than $\mathcal{G}_{d,n}$, as we shall see in Example ???, but for $d=2$ they are equal:

Theorem 5.2. *The three-term Plücker relations (2) form a tropical basis for the Plücker ideal $I_{2,n}$, and hence the tropical prevariety pre$\mathcal{G}_{2,n}$ equals $\mathcal{G}_{2,n}$.*

The proof of this theorem will be derived as a corollary to our discussion of phylogenetic trees. First, however, we reduce the dimension of the tropical Grassmannian $\mathcal{G}_{d,n}$. Consider the linear map from $\mathbb{R}^n$ to $\mathbb{R}^{\binom{n}{d}}$ which sends an n-vector $(a_1, a_2, \ldots, a_n)$ to the $\binom{n}{d}$-vector whose coordinates are $a_{i_1} + \cdots + a_{i_d}$. The map ϕ is injective and its image L equals the intersection of all cones in $\mathcal{G}_{d,n}$. The image of $\mathcal{G}_{n,d}$ in $\mathbb{R}^{\binom{n}{d}}/L$ is a pointed fan of dimension $d(n-d)+1-n = nd - n - d^2 + 1$. Intersecting the image with a sphere, we obtain a pure polyhedral complex $\mathcal{G}'_{d,n}$ of dimension $nd - n - d^2$. By standard abuse of notation, we refer to this polyhedral complex as the tropical Grassmannian.

Example 5.3. ($d = 2, n = 4$) The smallest non-zero Plücker ideal is the principal ideal $I_{2,4} = \langle p_{12}p_{34} - p_{13}p_{24} + p_{14}p_{23}\rangle$. Its tropical variety $\mathcal{G}_{2,4} = \mathcal{T}(I_{2,4})$ is a fan with three five dimensional cones $L \times \mathbb{R}_{\geq 0}$ glued along $L \simeq \mathbb{R}^4$. The polyhedral complex $\mathcal{G}'_{2,4}$ is zero-dimensional and consists of three points. $\square$

Example 5.4. ($d = 2, n = 5$) The tropical Grassmannian $\mathcal{G}_{2,5}$ is a pure fan of dimension 7 in $\mathbb{R}^{10}$, and its reduced version $\mathcal{G}'_{2,5}$ is a one-dimensional complex. This complex the *Petersen graph*, which has 10 vertices and 15 edges. $\square$

The following theorem generalizes both of these examples. It concerns the $d = 2$, that is, the tropical version of the Grassmannian of lines in $\mathbb{P}^{n-1}$.

Theorem 5.5. *The tropical Grassmannian $\mathcal{G}'_{2,n}$ is a simplical complex known as the space of phylogenetic trees. It has $2^{n-1} - n - 1$ vertices, $1 \cdot 3 \cdot 5 \cdots (2n-5)$ facets, and its homotopy type is a bouquet of $(n-2)!$ spheres of dimension $n-4$.*

We begin our steps towards a proof with a description of the simplicial complex. It is denoted by $\mathbf{T}_n$ and is defined as follows. The vertex set consists of all unordered pairs $\{A, B\}$ where A and B are disjoint subsets of $[n] := \{1, 2, \ldots, n\}$ having cardinality at least two, and $A \cup B = [n]$. Such pairs are called *splits*. The number of splits is $2^{n-1} - n - 1$. Two splits $\{A, B\}$ and $\{A', B'\}$ are connected by an edge in the simplicial complex $\mathbf{T}_n$ if and only if

$$A \subseteq A' \text{ or } A \subseteq B' \text{ or } B \subseteq A' \text{ or } B \subseteq B'. \tag{5}$$

We define $\mathbf{T}_n$ as the largest simplicial complex having this edge graph. Thus, a subset $\sigma \subset \mathrm{Vert}(\mathbf{T}_n)$ is a face of $\mathbf{T}_n$ if any pair $\big\{\{A,B\},\{A',B'\}\big\} \subseteq \sigma$ satisfies (5). In the language of algebraic combinatorics, $\mathbf{T}_n$ is the *flag complex* of the compatibility graph specified by (5) on the set of all $2^{n-1}-n-1$ splits.

The two smallest cases $n=4$ (three points) and $n=5$ (the Petersen graph) are discussed in Examples 5.3 and 5.4. Here is a description of the next case:

Example 5.6. ($n=6$) The two-dimensional simplicial complex $\mathbf{T}_6$ has 25 vertices, 105 edges and 105 triangles, each coming in two symmetry classes:

$$
\begin{gathered}
15 \text{ vertices like } \{12,3456\}\,, \quad 10 \text{ vertices like } \{123,456\},\\
60 \text{ edges like } \{\{12,3456\},\{123,456\}\},\\
45 \text{ edges like } \{\{12,3456\},\{1234,56\}\},\\
90 \text{ triangles like } \{\{12,3456\},\{123,456\},\{1234,56\}\},\\
15 \text{ triangles like } \{\{12,3456\}\},\{34,1256\}\},\{56,1234\}\}.
\end{gathered}
$$

Each edge lies in three triangles, so it looks locally like a tropical in $\mathbb{R}^2$. □

The simplicial complex $\mathbf{T}_n$ is well-known in phylogenetic combinatorics [66]. Its faces σ correspond to semi-labeled trees with leaf labels $1,2,\ldots,n$. Here each internal node is unlabeled and has at least three neighbors. Each internal edge of such a tree defines a partition $\{A,B\}$ of the set of leaves $\{1,2,\ldots,n\}$, and we encode the tree by the set of splits representing its internal edges.

The facets (= maximal faces) of $\mathbf{T}_n$ correspond to *trivalent trees*, that is, semi-labeled trees whose internal nodes all have three neighbors. All facets of $\mathbf{T}_n$ have the same cardinality $n-3$, the number of internal edges of any trivalent tree. Hence $\mathbf{T}_n$ is pure of dimension $n-4$. The number of facets (i.e. trivalent semi-labeled trees on $\{1,2,\ldots,n\}$) is the *Schröder number*

$$
(2n-5)!! \quad = \quad (2n-5)\times(2n-7)\times\cdots\times 5\times 3\times 1. \tag{6}
$$

We now describe an embedding of $\mathbf{T}_n$ as a simplicial fan into the $\frac{1}{2}n(n-3)$-dimensional vector space $\mathbb{R}^{\binom{n}{2}}/\mathrm{image}(\phi)$. For each trivalent tree σ we first define a cone B_σ in $\mathbb{R}^{\binom{n}{2}}$ as follows. By a *realization* of a semi-labeled tree σ we mean a one-dimensional cell complex in some Euclidean space whose underlying graph is a tree isomorphic to σ. Such a realization of σ is a metric space on $\{1,2,\ldots,n\}$. The *distance* between i and j is the length of the unique path between leaf i and leaf j in that realization. Then we set

$$
\begin{aligned}
B_\sigma \quad = \quad & \big\{\, (w_{12},w_{13},\ldots,w_{n-1,n}) \in \mathbb{R}^{\binom{n}{2}} \;:\; -w_{ij} \text{ is the distance from}\\
& \qquad \text{leaf } i \text{ to leaf } j \text{ in some realization of } \sigma \big\} \quad + \quad \mathrm{image}(\phi).
\end{aligned}
$$

Let C_σ denote the image of B_σ in the quotient space $\mathbb{R}^{\binom{n}{2}}/\mathrm{image}(\phi)$. Passing to this quotient has the geometric meaning that two trees are identified if their only difference is in the lengths of the n edges adjacent to the leaves.

Lemma 5.7. *The closure $\overline{C}_\sigma$ is a simplicial cone of dimension $|\sigma|$ with relative interior C_σ. The set of all cones $\overline{C}_\sigma$, as σ runs over $\mathbf{T}_n$, is a simplicial fan. The support of this fan is the space of tree metrics in $\mathbb{R}^{\binom{n}{2}}$ modulo* $\text{image}(\phi)$.

Proof. Metric spaces that can be realized as trees are characterized by the *Four Point Condition* (see e.g. [61, 66]). This condition states that for any quadruple of leaves i, j, k, l there exists a unique relabeling such that

$$w_{ij} + w_{kl} \ = \ w_{ik} + w_{jl} \ \leq \ w_{il} + w_{jk}. \tag{7}$$

Given any tree σ, this gives a system of $\binom{n}{4}$ linear equations and $\binom{n}{4}$ linear inequalities. The solution set of this linear system is precisely the closure $\overline{B}_\sigma$ of the cone B_σ in $\mathbb{R}^{\binom{n}{2}}$. The Neighbor Joining Algorithm [61, Algorithm 2.41] easily reconstructs the combinatorial tree σ from any point w in B_σ.

All of our cones share a common linear subspace, namely,

$$\overline{B}_\sigma \ \cap \ -\overline{B}_\sigma \quad = \quad \text{image}(\phi). \tag{8}$$

This is seen by replacing the inequalities in (7) by equalities. The cone $\overline{B}_\sigma$ is the direct sum (9) of this linear space with a $|\sigma|$-dimensional simplicial cone. Let $\{e_{ij} \,:\, 1 \leq i < j \leq n\}$ denote the standard basis of $\mathbb{R}^{\binom{n}{2}}$. Adopting the convention $e_{ji} = e_{ij}$, for any split $\{A, B\}$ of $\{1, 2, \ldots, n\}$ we define

$$E_{A,B} \quad = \quad \sum_{i \in A} \sum_{j \in B} e_{ij}.$$

These vectors give the generators of our cone as follows:

$$\overline{B}_\sigma \quad = \quad \text{image}(\phi) \quad + \quad \mathbb{R}_{\geq 0} \big\{ \, E_{A,B} \,:\, \{A, B\} \in \sigma \, \big\}. \tag{9}$$

From the two presentations (7) and (9) it follows that

$$\overline{B}_\sigma \ \cap \ \overline{B}_\tau \quad = \quad \overline{B}_{\sigma \cap \tau} \qquad \text{for all } \sigma, \tau \in \mathbf{T}_n. \tag{10}$$

Therefore the cones B_σ form a fan in $\mathbb{R}^{\binom{n}{2}}$, and this fan has face poset $\mathbf{T}_n$. It follows from (9) that the quotient $\overline{C}_\sigma \, = \, \overline{B}_\sigma / \text{image}(\phi)$ is a pointed cone.

We get the desired conclusion for the cones $\overline{C}_\sigma$ by taking quotients modulo the common linear subspace (8). The resulting fan in $\mathbb{R}^{\binom{n}{2}} / \text{image}(\phi)$ is simplicial of pure dimension $n - 3$ and has face poset $\mathbf{T}_n$. The support of this fan is as stated because the support of the fan of $\overline{B}_\sigma$'s is the space of tree metrics with $\text{image}(\phi)$ added to it. We note that this space is isometric to the Billera-Holmes-Vogtmann space in [11] because their metric is flat on each cone $\overline{C_\sigma} \simeq \mathbb{R}^{|\sigma|}_{\geq 0}$ and extended by the gluing relations $\overline{C}_\sigma \ \cap \overline{C}_\tau \ = \ \overline{C}_{\sigma \cap \tau}$. □

We now turn to the tropical Grassmannian and prove the theorem stated earlier. The simplicial complex $\mathbf{T}_n$ is identified with the fan in Lemma 5.7.

Proof of Theorem 5.5: The Plücker ideal $I_{2,n}$ is generated by the $\binom{n}{4}$ quadrics

$$p_{ij} p_{kl} \ - \ p_{ik} p_{jl} \ + \ p_{il} p_{jk} \qquad \text{for } 1 \leq i < j < k < l \leq n.$$

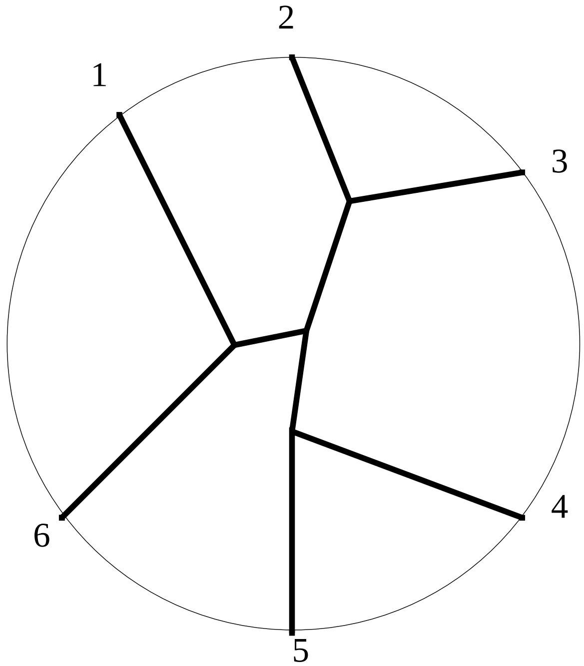

FIGURE 1. A Circular Labeling of a Tree with Six Leaves

The tropicalization of this polynomial is the disjunction of linear systems

$$\begin{array}{rrcccl} & w_{ij}+w_{kl} & = & w_{ik}+w_{jl} & \leq & w_{il}+w_{jk} \\ \text{or} & w_{ij}+w_{kl} & = & w_{il}+w_{jk} & \leq & w_{ik}+w_{jl} \\ \text{or} & w_{ik}+w_{jl} & = & w_{il}+w_{jk} & \leq & w_{ij}+w_{kl}. \end{array}$$

The solution set to these constraints is the tropical prevariety $pre\mathcal{G}_{2,n}$. The Four Point Condition, cited in the previous proof, shows that $|\mathbf{T}_n| = pre\mathcal{G}_{2,n}$.

We need to prove that the fans $\mathbf{T}_n$ and $\mathcal{G}_{2,n}$ are equal. Equivalently, every cone C_σ is actually a cone in the Gröbner fan of $I_{2,n}$ and the corresponding initial ideal contains no monomial. In view of the fan property (10), it suffices to consider maximal cones σ of $\mathbf{T}_n$. Fix a trivalent tree σ and a weight vector $w \in C_\sigma$. Then, for every quadruple i, j, k, l, the inequality in (7) is strict. This means combinatorially that $\big\{\{i,l\},\{j,k\}\big\}$ is a four-leaf subtree of σ.

Let J_σ denote the ideal generated by the quadratic binomials $p_{ij}p_{kl} - p_{ik}p_{jl}$ corresponding to all four-leaf subtrees of σ. Our discussion shows that $J_\sigma \subseteq \mathrm{in}_w(I_{2,n})$. The proof will be complete by showing that the two ideals agree:

$$J_\sigma \;=\; \mathrm{in}_w(I_{2,n}). \tag{11}$$

This identity will be proved by showing that the two ideals have a common initial monomial ideal, generated by square-free quadratic monomials.

We may assume, without loss of generality, that $-w$ is a strictly positive vector, corresponding to a planar realization of the tree σ in which the leaves $1, 2, \ldots, n$ are arranged in circular order to form a convex n-gon (Figure 1).

Let M be the ideal generated by the monomials $p_{ik}p_{jl}$ for $1 \leq i < j < k < l \leq n$. These are the crossing pairs of edges in the n-gon. By a classical construction of invariant theory, known as *Kempe's circular straightening law* (see [75, Theorem 3.7.3]), there exists a term order $\prec_{\mathrm{circ}}$ on $\mathbb{Z}[p]$ such that

$$M \quad = \quad \mathrm{in}_{\prec_{\mathrm{circ}}}(I_{2,n}). \tag{12}$$

Now, by our circular choice w of realization of the tree σ, the crossing monomials $p_{ik}p_{jl}$ appear as terms in the binomial generators of J_σ. Moreover, the term order $\prec_{\mathrm{circ}}$ on $\mathbb{Z}[p]$ refines the weight vector w. This implies

$$\mathrm{in}_{\prec_{\mathrm{circ}}}(\mathrm{in}_w(I_{2,n})) \quad = \quad \mathrm{in}_{\prec_{\mathrm{circ}}}(I_{2,n}) \quad = \quad M \quad \subseteq \quad \mathrm{in}_{\prec_{\mathrm{circ}}}(J_\sigma). \tag{13}$$

Using $J_\sigma \subseteq \mathrm{in}_w(I_{2,n})$ we conclude that equality holds in (13) and in (11).

Vogtman [88] proved that $\mathbf{T}_n$ has the homotopy type of a bouquet of $(n-2)\,!$ spheres of dimension $n-4$. A stronger combinatorial result, stating that $\mathbf{T}_n$ is a shellable complex, was shown by Trappman and Ziegler [86]. An alternative derivation of the homotopy type was given by Ardila and Klivans in [7]. $\square$

The simplicial complex $\Delta(M)$ represented by the squarefree monomial ideal M is an iterated cone over the boundary of the polar dual of the *associahedron*; see [75, page 132]. The facets of $\Delta(M)$ are the triangulations of the n-gon. Their number is the common degree of the ideals $I_{2,n}$, J_σ and M:

$$\text{the } (n-2)^{\mathrm{nd}} \text{ Catalan number} \quad = \quad \frac{1}{n-1}\binom{2n-4}{n-2}.$$

Corollary 5.8. *There exists a maximal cone in the Gröbner fan of the Plücker ideal $I_{2,n}$ which contains, up to symmetry, all cones of the Grassmannian $\mathcal{G}_{2,n}$.*

Proof. The cone corresponding to the initial ideal (12) has this property. $\square$

Corollary 5.9. *Every monomial-free initial ideal of $I_{2,n}$ is a prime ideal.*

Proof. If an initial ideal of a given ideal is prime then that ideal is prime as well. It therefore suffices to consider the binomial ideals $\mathrm{in}_w(I_{2,n})$ where w is in C_σ for some maximal cone σ. Then the vector w satisfies all four point conditions (7) with strict inequalities. Hence $\mathrm{in}_w(I_{2,n}) = J_\sigma$ for some semi-labeled trivalent tree σ. The ideal J_σ is radical and equidimensional because its initial ideal $M = \mathrm{in}_{\prec_{\mathrm{circ}}}(J_\sigma)$ is radical and equidimensional (unmixed).

To show that J_σ is prime, we use the following argument. For each edge e of the tree σ we introduce an indeterminate y_e. Consider the polynomial ring

$$\mathbb{C}[y] \quad = \quad \mathbb{C}\big[\, y_e \,:\, e \text{ edge of } \sigma \,\big].$$

Let ψ denote the homomorphism $\mathbb{C}[p] \to \mathbb{C}[y]$ which sends p_{ij} to the product of all indeterminates y_e corresponding to edges on the unique path between leaf i and leaf j. The kernel of ψ is a prime ideal. We claim that $\mathrm{kernel}(\psi) = J_\sigma$. The convex polytope corresponding to the toric ideal $\mathrm{kernel}(\psi)$ has a canonical triangulation into $\frac{1}{n-1}\binom{2n-4}{n-2}$ unit simplices (namely, $\Delta(M)$). Hence $\mathrm{kernel}(\psi)$ and J_σ are both unmixed of the same dimension and the same degree. Since $\mathrm{kernel}(\psi)$ is contained in J_σ, it follows that the two ideals are equal. $\square$

We next study the case $d = 3$ and $n = 6$ in detail. The Plücker ideal $I_{3,6}$ is minimally generated by 35 quadrics in the polynomial ring in 20 variables,

$$\mathbb{C}[p] \quad = \quad \mathbb{C}[p_{123}, p_{124}, \ldots, p_{456}].$$

We are interested in the 10-dimensional fan $\mathcal{G}_{3,6}$ which consists of all vectors $w \in \mathbb{R}^{20}$ such that $\mathrm{in}_w(I_{3,6})$ is monomial-free. The four-dimensional pointed quotient fan of $\mathcal{G}_{3,6}$ sits in $\mathbb{R}^{20}/\mathrm{image}(\phi) \simeq \mathbb{R}^{14}$ and is a fan over the three-dimensional polyhedral complex $\mathcal{G}'_{3,6}$. Our aim is to prove the following result:

Theorem 5.10. *The tropical Grassmannian $\mathcal{G}'_{3,6}$ is a 3-dimensional simplicial complex with 65 vertices, 550 edges, 1395 triangles and 1035 tetrahedra. The homology of $\mathcal{G}'_{3,6}$ is concentrated in (top) dimension 3 and $H_3(\mathcal{G}'_{3,6}, \mathbb{Z}) = \mathbb{Z}^{126}$.*

We begin by listing the vertices. Let E denote the set of 20 standard basis vectors e_{ijk} in $\mathbb{R}^{\binom{6}{3}}$. For each 4-subset $\{i, j, k, l\}$ of $\{1, 2, \ldots, 6\}$ we set

$$f_{ijkl} \quad = \quad e_{ijk} + e_{ijl} + e_{ikl} + e_{jkl}.$$

Let F denote the set of these 15 vectors. Finally consider any of the 15 *tripartitions* $\{\{i_1, i_2\}, \{i_3, i_4\}, \{i_5, i_6\}\}$ of $\{1, 2, \ldots, 6\}$ and define the vectors

$$\begin{aligned} & g_{i_1 i_2 i_3 i_4 i_5 i_6} \; := \; f_{i_1 i_2 i_3 i_4} + e_{i_3 i_4 i_5} + e_{i_3 i_4 i_6} \\ \text{and} \qquad & g_{i_1 i_2 i_5 i_6 i_3 i_4} \; := \; f_{i_1 i_2 i_5 i_6} + e_{i_3 i_5 i_6} + e_{i_4 i_5 i_6}. \end{aligned}$$

This gives us another set G of 30 vectors. All 65 vectors in $E \cup F \cup G$ are regarded as elements of the quotient space $\mathbb{R}^{\binom{6}{3}}/\mathrm{image}(\phi) \simeq \mathbb{R}^{14}$. Note that

$$g_{i_1 i_2 i_3 i_4 i_5 i_6} \quad = \quad g_{i_3 i_4 i_5 i_6 i_1 i_2} \quad = \quad g_{i_5 i_6 i_1 i_2 i_3 i_4}.$$

The following identity will be used later in the proof of Theorem 5.10:

$$(14) \qquad g_{i_1 i_2 i_3 i_4 i_5 i_6} + g_{i_1 i_2 i_5 i_6 i_3 i_4} \quad = \quad f_{i_1 i_2 i_3 i_4} + f_{i_1 i_2 i_5 i_6} + f_{i_3 i_4 i_5 i_6}.$$

The following lemma was found by computation:

Lemma 5.11. *The set of vertices of $\mathcal{G}'_{3,6}$ equals $E \cup F \cup G$.*

We next describe all the 550 edges of the tropical Grassmannian $\mathcal{G}_{3,6}$.

(EE) There are 90 edges like $\{e_{123}, e_{145}\}$ and 10 edges like $\{e_{123}, e_{456}\}$, for a total of 100 edges connecting pairs of vertices both of which are in E. (By the word “like”, we will always mean “in the S_6 orbit of, where S_6 permutes the indices $\{1, 2, \ldots 6\}$.)

(FF) This class consists of 45 edges like $\{f_{1234}, f_{1256}\}$.
(GG) Each of the 15 tripartitions gives exactly one edge, like $\{g_{123456}, g_{125634}\}$.
(EF) There are 60 edges like $\{e_{123}, f_{1234}\}$ and 60 edges like $\{e_{123}, f_{1456}\}$, for a total of 120 edges connecting a vertex in E to a vertex in F.
(EG) This class consists of 180 edges like $\{e_{123}, g_{123456}\}$. The intersections of the index triple of the e vertex with the three index pairs of the g vertex must have cardinalities $(2, 1, 0)$ *in this cyclic order.*
(FG) This class consists of 90 edges like $\{f_{1234}, g_{123456}\}$.

Lemma 5.12. *The* 1*-skeleton of* $\mathcal{G}'_{3,6}$ *is the graph with the* 550 *edges above.*

Let Δ denote the *flag complex* specified by the graph in the previous lemma. Thus Δ is the simplicial complex on $E \cup F \cup G$ whose faces are subsets σ with the property that each 2-element subset of σ is one of the 550 edges. We will see that $\mathcal{G}_{3,6}$ is a subcomplex homotopy equivalent to Δ.

Lemma 5.13. *The flag complex* Δ *has* $1,410$ *triangles,* $1,065$ *tetrahedra,* 15 *four-dimensional simplices, and it has no faces of dimension five or more.*

The facets of Δ are grouped into seven symmetry classes:

Facet FFFGG: There are 15 four-dimensional simplices, one for each partition of $\{1, \ldots, 6\}$ into three pairs. An example of such a tripartition is $\{\{1, 2\}, \{3, 4\}, \{5, 6\}\}$. It gives the facet $\{f_{1234}, f_{1256}, f_{3456}, g_{123456}, g_{125634}\}$. The 75 tetrahedra contained in these 15 four-simplices are not facets of Δ.

The remaining 990 tetrahedra in Δ are facets and they come in six classes:

Facet EEEE: There are 30 tetrahedra like $\{e_{123}, e_{145}, e_{246}, e_{356}\}$.
Facet EEFF1: There are 90 tetrahedra like $\{e_{123}, e_{456}, f_{1234}, f_{3456}\}$.
Facet EEFF2: There are 90 tetrahedra like $\{e_{125}, e_{345}, f_{3456}, f_{1256}\}$.
Facet EFFG: There are 180 tetrahedra like $\{e_{345}, f_{1256}, f_{3456}, g_{123456}\}$.
Facet EEEG: There are 240 tetrahedra like $\{e_{126}, e_{134}, e_{356}, g_{125634}\}$.
Facet EEFG: There are 360 tetrahedra like $\{e_{234}, e_{125}, f_{1256}, g_{125634}\}$.

While Δ is an abstract simplicial complex on the vertices of $\mathcal{G}'_{3,6}$, it is not embedded on the given vertices because of the relation (14) which says that the five involved vertices form a bipyramid with the F-vertices as the base and the G-vertices as the two cone points. We modify the flag complex Δ to a new simplicial complex Δ' which has pure dimension three. The complex Δ' is obtained from Δ by removing the 15 FFF-triangles $\{f_{1234}, f_{1256}, f_{3456}\}$, along with the 30 tetrahedra FFFG and the 15 four-simplices FFFGG containing the FFF-triangles. In the three-dimensional complex Δ', the bipyramids are each divided into three tetrahedra arranged around the GG-edges.

Proof of Theorem 5.10: It remains to show that the Grassmannian $\mathcal{G}'_{3,6}$ equals the simplicial complex Δ'. This is accomplished an explicit computation. The integral homology groups Δ' were computed independently by Michael Joswig and Volkmar Welker. The assertion $\Delta' = \mathcal{G}_{3,6}$ can be verified by the following method. One first checks that the seven types of cones described above are indeed Gröbner cones of $I_{3,6}$ whose initial ideals are monomial-free. Next one checks that the list is complete. This relies on our result in Section 3 which states that $\mathcal{G}_{3,6}$ is connected in codimension 1. The completeness check is done by computing the link of each of the known classes of triangles. Algebraically, this amounts to computing the (truly zero-dimensional) tropical variety of $\mathrm{in}_w(I_{3,6})$ where w is any point in the relative interior of the triangular cone in question. For all but one class of triangles the link consists of three points, and each neighboring 3-cell is found to be already among our seven classes. The links of the triangles are as follows:

Triangle EEE: The link of $\{e_{146}, e_{256}, e_{345}\}$ consists of e_{123}, g_{163425}, g_{142635}.
Triangle EEF: The link of $\{e_{256}, e_{346}, f_{1346}\}$ consists of f_{1256}, g_{132546}, g_{142536}.
Triangle EEG: The link of $\{e_{156}, e_{236}, g_{142356}\}$ consists of e_{124}, e_{134}, f_{1456}.
Triangle EFF: The link of $\{e_{135}, f_{1345}, f_{2346}\}$ consists of e_{236}, e_{246}, g_{153426}.
Triangle EFG: The link of $\{e_{235}, f_{2356}, g_{143526}\}$ consists of e_{145}, f_{1246}, e_{134}.
Triangle FFG: The link of $\{f_{1236}, f_{1345}, g_{134526}\}$ consists of e_{126}, e_{236}, g_{132645}.
Triangle FGG: The link of $\{f_{1456}, g_{142356}, g_{145623}\}$ consists of f_{2356} and f_{1234}.

The FGG triangle lies in the interior of our bipyramid FFFGG and is incident to two of the three FFGG tetrahedra which make up the triangulation of that bipyramid. It is not contained in any other facet of $\mathcal{G}''_{3,6}$. □

One of the consequences which can be inferred from our detailed study of $\mathcal{G}_{3,6}$ is that the quadrics in $I_{3,6}$ form a tropical basis, and hence $\mathcal{G}_{3,6} = pre\mathcal{G}_{3,6}$. It turns out that the same statement no longer holds for $n \geq d+4 \geq 7$.

Proposition 5.14. *For $n \geq d+4 \geq 7$, the tropical Grassmannian $\mathcal{G}_{d,n}$ is strictly contained in the tropical prevariety $pre\mathcal{G}_{d,n}$ defined by the relations (4).*

Proof. We consider the case $d = 3$ and $n = 7$. An easy lifting argument extends our example to the general case $n \geq d+4 \geq 7$. The Plücker ideal $I_{3,7}$ is minimally generated by 140 quadrics in a polynomial ring $\mathbb{C}[p_{123}, p_{124}, \ldots, p_{567}]$ in 35 unknowns. We fix the following zero-one vector. The appearing triples are gotten by a cyclic shift, and they correspond to the lines in the *Fano plane*:

$$w \quad = \quad e_{124} + e_{235} + e_{346} + e_{457} + e_{156} + e_{267} + e_{137} \quad \in \quad \mathbb{R}^{\binom{6}{3}}.$$

The vector w satisfies all the three-term Plücker relations (4), so it lies in the prevariety $pre\mathcal{G}_{3,7}$. To show that it is not in the Grassmannian $\mathcal{G}_{3,7}$, we compute the initial ideal $\mathrm{in}_w(I_{3,7})$. In a computer algebra system, this is done by computing the reduced Gröbner basis of $I_{3,7}$ over the field of rational numbers with respect to the (reverse lexicographically refined) weight order defined by $-w$. The reduced Gröbner basis is found to have precisely 196 elements, namely, 140 quadrics, 52 cubics, and 4 quartics. The initial ideal $\mathrm{in}_w(I_{3,7})$ is generated by the w-leading forms of the 196 elements in that Gröbner basis.

Among the 52 cubics in the Gröbner basis of $I_{3,7}$, we find the special cubic

$$\begin{aligned} f \quad = \quad & 2 \cdot p_{123}p_{467}p_{567} \; - \; p_{367}p_{567}\underline{p_{124}} - p_{167}p_{467}\underline{p_{235}} - p_{127}p_{567}\underline{p_{346}} \\ & -p_{126}p_{367}\underline{p_{457}} - p_{237}p_{467}\underline{p_{156}} + p_{134}p_{567}\underline{p_{267}} + p_{246}p_{567}\underline{p_{137}} \; + \; p_{136}\underline{p_{267}}\,\underline{p_{457}}. \end{aligned}$$

The weights of the underlined variables are one, while the weights of the non-underlined variables are zero. We conlcude that the leading form of this special cubic polynomial is the monomial

$$\mathrm{in}_w(f) \quad = \quad p_{123}p_{467}p_{567}.$$

This proves that w does not lie in the tropical Grassmannian $\mathcal{G}_{3,7}$. □

The Fano plane in the proof of Proposition 5.14 indicates an intimate connection between the tropical Grassmannian and the theory of matroids. The basis exchange axiom in matroid theory can be interpreted as consistency with the three-term Plücker relations (3). This leads to the following result.

Corollary 5.15. *Let $\mathcal{B}$ be any subset of the collection $\binom{[n]}{d}$ of d-element subsets of $\{1, \ldots, n\}$, and consider the negative of the corresponding incidence vector:*

$$w_{\mathcal{B}} \quad := \quad -\sum_{\sigma \in \mathcal{B}} e_\sigma \quad \in \quad \mathbb{R}^{\binom{n}{d}}.$$

Then $\mathcal{B}$ is the set of bases of a matroid if and only if $w_{\mathcal{B}}$ lies in $pre\mathcal{G}_{d,n}$. The vector $w_{\mathcal{B}}$ also lies in $\mathcal{G}_{d,n}$ if and only if that matroid can be realized over $\mathbb{C}$.

Proof. The first statement is a re-interpretation of the basis exchange axiom for matroids. Consider the second statement. The if-direction is seen as follows: Suppose B is a complex $d \times n$-matrix which realizes the given matroid, i.e., the column bases of B are precisely the d-sets in $\mathcal{B}$. Then take a generic complex $d \times n$-matrix C and consider the matrix $t^{-1}B + C$ over $K = \mathbb{C}\{\{t\}\}$. Consider the vector of maximal minors of this matrix and multiply that vector by t^{d-1}. The resulting vector lies in the Grassmannian $G_{d,n}$ and its image under the valuation map is precisely the vector $w_{\mathcal{B}}$ Thus $w_{\mathcal{B}}$ lies in $\mathcal{G}_{d,n}$. The same argument works in reverse, using the extended Kapranov theorem [45, 82]. □

The combinatorial study of tropical Grassmannians and its relation to matroid theory has its origin in the work of Dress and Wenzel [28]. They introduced the concept of *valuated matroids*, and these are precisely the points in the prevariety $pre\mathcal{G}_{d,n}$. The Dress-Wenzel theory was further developed by Murota [59] who showed that the three-term relations (3) can be replaced by the set of all quadrics in $I_{d,n}$ without changing the prevariety. Dress and Terhalle [27] explained the connection between valuated matroids and affine buildings, which leads to the interpretation of points in $\mathcal{G}_{d,n}$ as higher-dimensional trees. For that reason, Pachter and Sturmfels [61, §3.5] refer to $pre\mathcal{G}_{d,n}$ as the *space of d-trees*. The state of the art on this subject is the work of Speyer [72, 73] on tropical linear spaces which will be discussed in the next section.