THE EQUATIONS FOR THE MODULI SPACE OF n POINTS ON THE LINE

BENJAMIN HOWARD, JOHN MILLSON, ANDREW SNOWDEN AND RAVI VAKIL

ABSTRACT. A central question in invariant theory is that of determining the relations among invariants. Geometric invariant theory quotients come with a natural ample line bundle, and hence often a natural projective embedding. This question translates to determining the equations of the moduli space under this embedding. This article deals with one of the most classical quotients, the space of ordered points on the projective line. We show that under any weighting of the points, this quotient is cut out (scheme-theoretically) by a particularly simple set of quadric relations, with the single exception of the Segre cubic threefold, the space of six points with equal weight. We also show that the ideal of relations is generated in degree at most four, and give an explicit description of the generators. If all the weights are even (e.g. in the case of equal weight for odd n), we show that the ideal of relations is generated by quadrics.

CONTENTS

Date: March 7, 2008.

The first two authors were partially supported by NSF grant DMS-0104006, the second author was also partially supported by NSF grant DMS-0405606 and NSF FRG grant DMS-0554254. The last author was partially supported by NSF Grant DMS-0228011.

2000 Mathematics Subject Classification: Primary 14L24, Secondary 14D22, 14H10.

We consider the space of n ordered points on the projective line, up to automorphisms of the line. In characteristic 0, the best description of this is as a geometric invariant theory quotient $(\mathbb{P}^1)^n//SL_2$, where different choices of linearization yield different compactifications. This is one of the most classical examples of a GIT quotient.

Generators for the coordinate ring in the symmetric case (defined shortly) have been known for a long time: in 1894 Kempe, see [Ke], proved that the lowest degree invariants generate the coordinate ring. We dub these generators the *Kempe generators*. However, the question of the relations has remained open. It was not even known that the relations have bounded degree as n varies. Explicit equations were known classically only for small n, likely only up to $n = 6$.

More generally we consider the space of *weighted* points on $\mathbb{P}^1$. Let the ith point be weighted by w_i and let $\mathbf{w} = (w_1, \ldots, w_n) \in (\mathbb{Z}^+)^n$ (the *weight vector*). The weights can be interpreted as parametrizing the very ample line bundles of $(\mathbb{P}^1)^n$. The most classical case, when all points are treated equally, corresponds to $\mathbf{w} = (1, \ldots, 1) = 1^n$. Let $M_{\mathbf{w}} := \operatorname{Proj} R_{\mathbf{w}}$ be the GIT quotient. We call this the *symmetric* case. At risk of confusion, we denote this important case M_{1^n} by M_n for simplicity.

We say that n points of $\mathbb{P}^1$ are $\mathbf{w}$-*stable* (respectively $\mathbf{w}$-*semistable*) in the sense of geometric invariant theory if the sum of the weights of any set of points that coincide is less than (respectively no more than) half the total weight. The dependence on $\mathbf{w}$ will be clear from the context, so the prefix $\mathbf{w}$- will usually be omitted. The n points are *strictly semistable* if they are semistable but not stable. Then $M_{\mathbf{w}}$ is a projective variety, and Kempe's theorem 2.3 suggests a natural projective embedding. The stable locus of $M_{\mathbf{w}}$ is a fine moduli space for the stable points of $(\mathbb{P}^1)^n$. The strictly semistable locus of $M_{\mathbf{w}}$ is a finite set of points in each characteristic, which are the only singular points of $M_{\mathbf{w}}$ (see for example [KM]). This is an abuse of notation; these points of $M_{\mathbf{w}}$ should be called the *images of the strictly semistable points*, but we will call them *strictly semistable points* in order to avoid too many verbal contortions.

The question we wish to address is: *what are the equations of $M_{\mathbf{w}}$?* There are three possible meanings to this question.

(a) What are "good" equations for $M_{\mathbf{w}}$ set-theoretically?
(b) What are "good" equations cutting it out scheme-theoretically?
(c) What are "good" generators for the *ideal* of all equations for $M_{\mathbf{w}}$?

Each question subsumes the previous one. The third question is the most fundamental. We give a good answer to (b) (First Main Theorem 1.1) and a satisfactory answer to (c) (Second Main Theorem 1.3), and speculate on a good answer to (c) (§1.5).

We prefer to work as generally as possible, over the integers, so we now define the moduli problem of stable n-tuples of points in $\mathbb{P}^1 = \mathbb{P}^1_{\mathbb{Z}} = \operatorname{Proj} \mathbb{Z}[x, y]$. For any scheme B, a *family of stable n-tuples of points in $\mathbb{P}^1$ over B* is a morphism $(\phi_1, \ldots, \phi_n) : B \times \{1, \ldots, n\} \to \mathbb{P}^1$ such that for any $I \subset \{1, \ldots, n\}$ such that $\sum_{i \in I} w_i \geq \sum_{i=1}^n w_i/2$, we have $\cap_{i \in I} \phi_i^{-1}(p) = \emptyset$

for all p. Then there is a fine moduli space for this moduli problem, quasiprojective over $\mathbb{Z}$, which has a natural ample line bundle, the one suggested by GIT. This is well-known, but in any case will fall out of our analysis.

We now state our two main theorems. We will describe a natural equivariant set of "graphical" generators of the algebra of invariants (in §2). The algebraic structure of the invariants is particularly transparent in this language, and as an example we give a short proof of Kempe's Theorem 2.3, and give an easy basis of the $\mathbb{Z}$-module of invariants (by "non-crossing variables", Proposition 2.6). Similar ideas, using certain graphs to describe invariants, appear in the nineteenth century, in the work of Clifford, Sylvester, and Kempe, see [OS]. Another application is the computation in §2.15 of the degrees of all $M_{\mathbf{w}}$. We then describe some geometrically or combinatorially obvious relations, the linear *sign relations*, the linear *Plücker relations*, the quadratic *simple binomial relations*, and the cubic *generalized Segre cubic relations*.

1.1. *First Main Theorem. — Over $\mathbb{Z}$, the space $M_{\mathbf{w}}$ is cut out scheme-theoretically (as a closed subscheme of projective space) by the sign, Plücker, simple binomial, and generalized Segre cubic relations. With the unique exception of $\mathbf{w} = (1,1,1,1,1,1)$, over $\mathbb{Z}[1/3]$, the space $M_{\mathbf{w}}$ is cut out scheme-theoretically by the sign, Plücker, and simple binomial relations.*

In particular, this answers question (b) above, and the ideal of relations (the answer to question (c)) is the radical of the ideal generated by these three "obvious" families of relations. The exceptional case $\mathbf{w} = (1,1,1,1,1,1)$ is the Segre cubic threefold.

The idea of the proof is as follows. We first reduce the question to the symmetric case, where n is even. We do this by showing a stronger result, which reduces such questions about the *ideal of relations* of invariants to the symmetric case. Yi Hu has pointed out to us that the map of semistable points corresponding to γ was constructed (for configurations on Grassmannians) in [Hu, Prop. 2.11].

1.2. *Theorem (reduction to symmetric case, informal statement). — For any weight $\mathbf{w}$, there is a natural map γ from the graded coordinate ring for $1^{|\mathbf{w}|}$ to those for $\mathbf{w}$. Under this map, each of our generators for $1^{|\mathbf{w}|}$ is sent to either a generator for $\mathbf{w}$, or to zero. Moreover, the relations for $1^{|\mathbf{w}|}$ generate the relations for $\mathbf{w}$.*

In §2.17, we will state this result precisely, and prove it, once we have introduced some terminology.

We then verify Theorem 1.1 by ad hoc means in the cases $n = 2m \leq 8$ (§3). The cases $n = 6$ and $n = 8$ are the base cases for our later argument — which is ironic, as in the $n = 6$ case Theorem 1.1 does not hold! In §5, we show that Theorem 1.1 holds set-theoretically, and that the projective variety is a fine moduli space away from the strictly semistable points. The strictly semistable points are more delicate, as the quotient is not naturally a fine moduli space there; we instead give an explicit description of a neighborhood of a strictly semistable point, as the affine variety corresponding to rank one $(n/2-1) \times (n/2-1)$ matrices with entries distinct from 1, using the Gel'fand-MacPherson correspondence. We prove Theorem 1.1 in this neighborhood in §4.

Our Second Main Theorem is about the full ideal of all relations.

1.3. *Second Main Theorem. — For any weights* $\mathbf{w} = (w_1, \ldots, w_n)$*, the ideal of relations in the coordinate ring (over* $\mathbb{Z}$*) is generated by relations of degree at most four. If all the weights are even, then the ideal of relations is generated in degree two.*

The proof is given in Part 3, and is completed in §8.

In Part 3 we choose a filtration of the ring so that the associated graded ring $\mathrm{gr}(R_{\mathbf{w}})$ is simpler to study. By *toric ring* we mean the coordinate ring of a toric variety; by *toric filtration* we mean a filtration of a ring R so that $\mathrm{gr}(R)$ is a toric ring. Here we show that $\mathrm{gr}(R_{\mathbf{w}})$ is a toric ring, by identifying $\mathrm{gr}(R_{\mathbf{w}})$ with the semigroup algebra of the semigroup ring of lattice points in a certain rational cone. We find that the relations among the generators of $\mathrm{gr}(R_{\mathbf{w}})$ are generated in degrees two, three, and four, and by lifting these relations to the original ideal, we obtain the Second Main Theorem 1.3. This is done by first noting that $\mathrm{gr}(R_{\mathbf{w}})$ is generated in degrees one and two. Then we define a normal form for monomials in the degree one and two generators so that the normal monomials are a basis of $\mathrm{gr}(R_{\mathbf{w}})$ as a $\mathbb{Z}$-module. We then show that any monomial can be brought into normal form by relations of degree four and less. The normal form monomials we define are not what one typically encounters. The normal form monomials are not the set of monomials outside some monomial initial ideal, because a normal monomial m may have a factor m' such that m' is not normal. It might be worthwhile in the future to investigate what term orders are well-suited to the study of the combinatorial properties of these toric varieties.

1.4. *Remark.* A toric filtration of the coordinate ring of the Grassmannian $G(2, n)$ for the Plücker embedding was given by Sturmfels in [St1] and all such filtrations appear in Speyer-Sturmfels [SpSt]; furthermore Lakshmibai-Gonciulea [GL] also defined such a filtration in their study of toric degenerations of general flag varieties. Our method is to restrict a toric filtration of the coordinate ring of $G(2, n)$ to the subring of T-invariants (equal to $R_{\mathbf{w}}$, see §4) where T is the maximal torus in SL_n. This method is described more generally in [FHu].

Thus the understanding of the ring of projective invariants of ordered points on the projective line is now quite satisfactory. This is in contrast with the equally classical, and much more complicated, question of *unordered* points, understood only for $n \leq 6$, and (by Shioda [Sh]) $n = 8$. In the unordered case, a generating set for general n is not known, but it is known that the degrees of the generators grow at least linearly in n, unlike the ordered case. Harm Derksen gave us a short and beautiful proof of this fact, with an argument similar to those of his paper [De]. One might dream that the case $n = 10$ might be tractable by computer, given the explicit relations for M_{10} described here. The case of unordered points with even n essentially corresponds to the ring of hyperelliptic modular forms of genus $(n-2)/2$, and their relations. Our case of ordered points, where n is even and the weights are even, essentially corresponds to the ring of hyperelliptic modular forms of level two, and this paper completely describes generators of the ideal of relations among these forms.

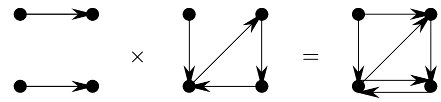

FIGURE 1. Multiplying (directed) graphs

We emphasize that our argument is almost entirely elementary, without reference to heavy machinery.

1.5. Related questions. The simple binomial relations cut out the quotient under the Kempe embedding (see 2.4). *Question:* Do they generate the ideal of relations among invariants if $n \neq 6$? By Theorem 1.2, it suffices to consider the symmetric case. It was classically known that the answer is yes for $n = 5$ (see §2.8), and we have also verified it for $n = 8$ (§2.10) and $n = 10$ (§2.13). The Second Main Theorem 1.3 suggests that one could hope to show that the explicit generators given there lie in the ideal given by these simple binomial quadrics. We have not succeeded in doing so.

Even special cases are simple to state but computationally too complex to verify directly even by computer. For example, we will describe particularly attractive relations for M_n in §2.14; do these lie in the ideal of our simple quadrics? There is a non-simple binomial relation for M_{12}, one for each partition of the 12 points into $6 + 6$. Is this an integral combination of our simple binomial quadrics?

Acknowledgments. We thank Lucia Caporaso, Harm Derksen, Igor Dolgachev, Philip Foth, Brendan Hassett, Roger Howe, Yi Hu, Allen Knutson, Anton Leykin, and Diane Maclagan. We thank Diane Maclagan and the referees for many significant improvements to the manuscript. We thank Allen Knutson for pointing out that the first toric degenerations of these spaces were given in [St1, Prop. 11.10]. We thank Philip Foth for many helpful comments about toric degenerations.

Part 1. A GRAPHICAL DESCRIPTION OF THE RING OF INVARIANTS OF POINTS ON A LINE

2. THE INVARIANTS OF n POINTS ON $\mathbb{P}^1$ AS A GRAPHICAL ALGEBRA

We now give a convenient alternate description of the generators (as a $\mathbb{Z}$-module) of the ring of invariants of n ordered points on $\mathbb{P}^1$. By *graph* we will mean a *directed* graph on n vertices labeled 1 through n. Graphs may have multiple edges, but may not have loops. The *multidegree* of a graph Γ is the n-tuple of valences of the graph, denoted $\mathbf{deg}\,\Gamma$. The bold font is a reminder that this is a vector. We consider each graph as a set of edges. For each edge e of Γ, let $h(e)$ be the head vertex of e and $t(e)$ be the tail. We use multiplicative notation for the "union" of two graphs: if Γ and Δ are two graphs on the same set of vertices, the union is denoted by $\Gamma \cdot \Delta$ (so for example $\mathbf{deg}\,\Gamma + \mathbf{deg}\,\Delta = \mathbf{deg}\,\Gamma \cdot \Delta$), see Figure 1. We will occasionally use additive and subtractive notation when we wish to "subtract" graphs. We apologize for this awkwardness.

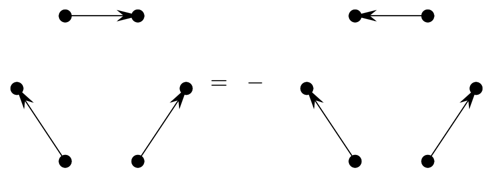

FIGURE 2. An example of the sign relation

We describe the coordinates of the ith point of $(\mathbb{P}^1)^n$ by $[u_i; v_i]$. To simplify formulas, we may write $[u_i; v_i] = [p_i; 1]$ where p_i could take on the value ∞. We leave it to the reader to re-homogenize such formulas.

For each graph Γ, define $X_\Gamma \in H^0((\mathbb{P}^1)^n, \mathcal{O}_{(\mathbb{P}^1)^n}(\mathbf{deg}\,\Gamma))$ by

$$X_\Gamma = \prod_{\text{edge } e \text{ of } \Gamma} (p_{h(e)} - p_{t(e)}) = \prod_{\text{edge } e \text{ of } \Gamma} (u_{h(e)}v_{t(e)} - u_{t(e)}v_{h(e)}). \tag{1}$$

For any non-empty set S of graphs of the same multidegree, the map $(\mathbb{P}^1)^n \dashrightarrow [X_\Gamma]_{\Gamma \in S}$ is easily seen to be invariant under SL_2.

The First Fundamental Theorem of Invariant Theory [Do, Thm. 2.1] implies that, given a weight $\mathbf{w}$, the coordinate ring $R_{\mathbf{w}}$ of $(\mathbb{P}^1)^n//SL_2$ is generated (as a $\mathbb{Z}$-module) by the X_Γ where $\mathbf{deg}\,\Gamma$ is a multiple of $\mathbf{w}$. The translation to the traditional language of tableaux is as follows. Choose any ordering of the edges $e_1, \dots, e_{|\Gamma|}$ of Γ. Then X_Γ corresponds to any $2 \times |\Gamma|$ tableau where the top row of the ith column is $t(e_i)$ and the bottom row is $h(e_i)$. We will soon see advantages of this graphical description as compared to the tableaux description.

We now describe several types of relations among the X_Γ, which will all be straightforward: the sign relations, the Plücker or straightening relations, the simple binomial relations, and the Segre cubic relation.

2.1. *The sign (linear) relations.* The *sign relation* $X_{\Gamma \cdot \vec{xy}} = -X_{\Gamma \cdot \vec{yx}}$ (Figure 2) is immediate from the definition (1). Because of the sign relation, we may omit arrowheads in identities where it is clear how to consistently insert them — see for example Figures 9 and 13, where even the vertices are implicit. We have an equivalence relation on directed graphs, where two are equivalent if their corresponding undirected graphs are the same. Our graphs will have labeled vertices, and when we want to pick a representative of the equivalence class, we will arbitrarily choose the one where if $a < b$, all edges ab are directed $a \to b$. We call such a graph an *upwards* graph. For example, two of the three graphs in Figure 3 are upwards. This choice is completely arbitrary, and breaks symmetry, so we prefer not to do this in general.

2.2. *The Plücker (linear) relations.* The identity of Figure 3 may be verified by direct calculation. If Γ is any graph on n vertices, and Δ_1, Δ_2, Δ_3 are three graphs on the same vertices given by identifying the four vertices of Figure 3 with four of the n vertices of Γ, then

$$X_{\Gamma \cdot \Delta_1} + X_{\Gamma \cdot \Delta_2} + X_{\Gamma \cdot \Delta_3} = 0. \tag{2}$$

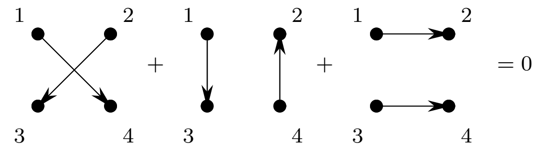

FIGURE 3. The Plücker relation for $n = 4$ (and $\mathbf{w} = (1, 1, 1, 1)$)

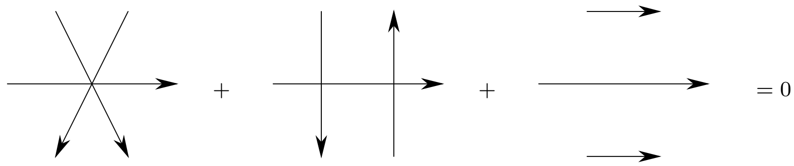

FIGURE 4. An example of a Plücker relation

These relations are called *Plücker relations*, or *straightening rules*. See Figure 4 for an example. We will sometimes refer to this relation as the Plücker relation for $\Gamma \cdot \Delta_1$ with respect to the vertices of Δ_1.

Using the Plücker relations, one can reduce the number of generators to a smaller set, which we will do shortly (Proposition 2.5). However, a central thesis of this article is that this is the wrong thing to do too soon; not only does it obscure the $\mathfrak{S}_n$ symmetry of this generating set, it also makes certain facts opaque. As an example, we give a new proof of Kempe's theorem. The proof will also serve as preparation for the proof of the First Main Theorem 1.1.

2.3. *Kempe's Theorem. — The lowest degree invariants generate the co-ordinate ring $R_{\mathbf{w}}$.*

Note that the lowest-degree invariants are of weight $\epsilon_{\mathbf{w}}\mathbf{w}$, where $\epsilon_{\mathbf{w}} = 1$ if $|\mathbf{w}|$ is even, and $\epsilon_{\mathbf{w}} = 2$ if $|\mathbf{w}|$ is odd.

Proof. We begin in the case when $\mathbf{w} = (1, \ldots, 1)$ where n is even. Recall Hall's Marriage Theorem: given a finite set of men M and women W, and some men and women are compatible (a subset of $M \times W$), and it is desired to compatibly pair each woman with a unique man, then it is necessary and sufficient that for each subset S of women, the number of men compatible with at least one of them is at least $|S|$.

Given a graph Γ of multidegree $(d, \ldots, d)$, we show that we can find an expression $X_\Gamma = \sum \pm X_{\Delta_i} \cdot X_{\Xi_i}$ where $\mathbf{deg}\, \Delta_i = (1, \ldots, 1)$. Arbitrarily divide the vertices into two equal-sized sets, one called the "positive" vertices and one called the "negative" vertices. This creates three types of edges: positive edges (both vertices positive), negative edges (both vertices negative), and neutral edges (one vertex of each sort). When one applies the Plücker relation to a positive edge and a negative edge, all resulting edges are neutral (see Figure 3, and take two of the vertices to be of each type). Also, each regular graph must have the same number of positive and negative edges. Working inductively on the number of positive edges, we can use the Plücker relations so that all resulting graphs have only neutral edges. We thus have an expression $X_\Gamma = \sum \pm X_{\Gamma_i}$ where each Γ_i has

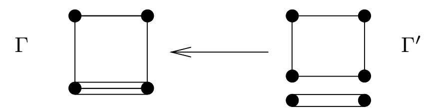

FIGURE 5. Constructing Γ' from Γ (example with $\mathbf{w} = (1, 1, 2, 2)$, $d = 2$)

only neutral edges and is hence a bipartite graph. Each vertex of Γ_i has the same valence d, so any set of p positive vertices must connect to at least p negative edges. By Hall's Marriage Theorem, we can find a matching Δ_i that is a subgraph of Γ_i, with "residual graph" Ξ_i (i.e. $\Gamma_i = \Delta_i \cdot \Xi_i$). (We emphasize that this application of Hall's theorem yields nothing more than König's Theorem.) Thus the result holds in the symmetric case.

We next treat the general case. If $|\mathbf{w}|$ is odd, it suffices to consider the case $2\mathbf{w}$, so by replacing $\mathbf{w}$ by $2\mathbf{w}$ if necessary, we may assume $\epsilon_{\mathbf{w}} = 1$. The key idea is that the quotient $M_{\mathbf{w}}$ is a linear section of $M_{|\mathbf{w}|}$. Suppose $\deg \Gamma = d\mathbf{w}$. Construct an auxiliary graph Γ' on $|\mathbf{w}|$ vertices, and a map of graphs $\pi : \Gamma' \to \Gamma$ such that (i) the preimage of vertex i of Γ consists of w_i vertices of Γ', (ii) π gives a bijection of edges, and (iii) each vertex of Γ' has valence d, i.e. Γ' is d-regular. (See Figure 5 for an illustrative example. There may be choice in defining Γ'.) Then apply the algorithm of the previous paragraph to Γ'. By taking the image under π, we have our desired result for Γ. □

2.4. *The Kempe embedding.* Since the X_Γ for $\deg(\Gamma) = \mathbf{w}$ generate the algebra $R_{\mathbf{w}}$, we shall use the X_Γ to define an embedding of $M_{\mathbf{w}}$ into projective space. We dub this embedding the *Kempe embedding*.

Choosing a planar representation of these graphs, as we shall now describe, makes termination of certain algorithms straightforward as well, as illustrated by the following argument. Consider the vertices of the graph to be the vertices of a regular n-gon, numbered clockwise 1 through n. A graph is said to be *non-crossing* if no two edges cross. Two edges sharing one or two vertices are considered not to cross. A variable X_Γ is said to be non-crossing (resp. upwards, §2.1) if Γ is. In Part 3, we will use regular upwards non-crossing graphs. This is a mouthful, so we dub them *Kempe graphs*.

The following result is well known. We include a proof in this graphical language, because later proofs will follow the same idea.

2.5. *Proposition (graphical version of "straightening algorithm"). — For each* $\mathbf{w}$*, the upwards non-crossing variables of multidegree* $\mathbf{w}$ *generate* $\langle X_\Gamma \rangle_{\deg \Gamma = \mathbf{w}}$ *as a* $\mathbb{Z}$*-module.*

This is essentially the straightening algorithm (e.g. [Do, §2.4] or [St2]) in this situation. This fact first appeared in [Ke] and the proof there is much the same as the one we present. The six non-crossing (undirected) graphs on 5 vertices are given in Figure 6. The fourteen non-crossing graphs on 8 vertices are given in Figure 7.

Proof. We explain how to express X_Γ in terms of upwards non-crossing variables. In this proof, we assume all variables are upwards, using the sign relation (§2.1). If Γ has a

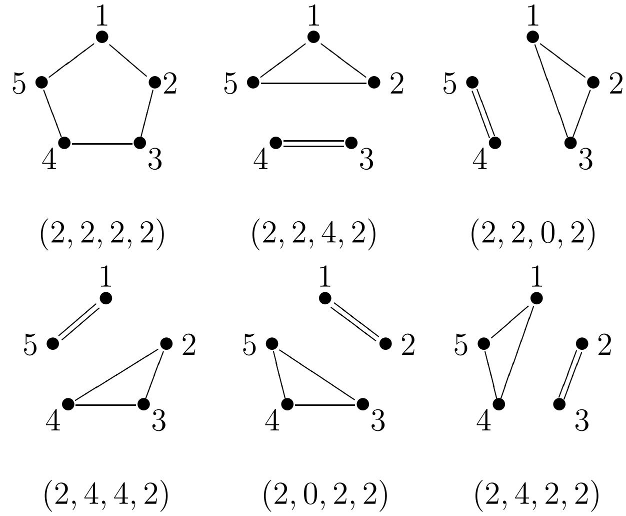

FIGURE 6. The six non-crossing (undirected) graphs on $n = 5$ vertices. The ordered quadruples will be relevant for the toric degeneration of Part 3.

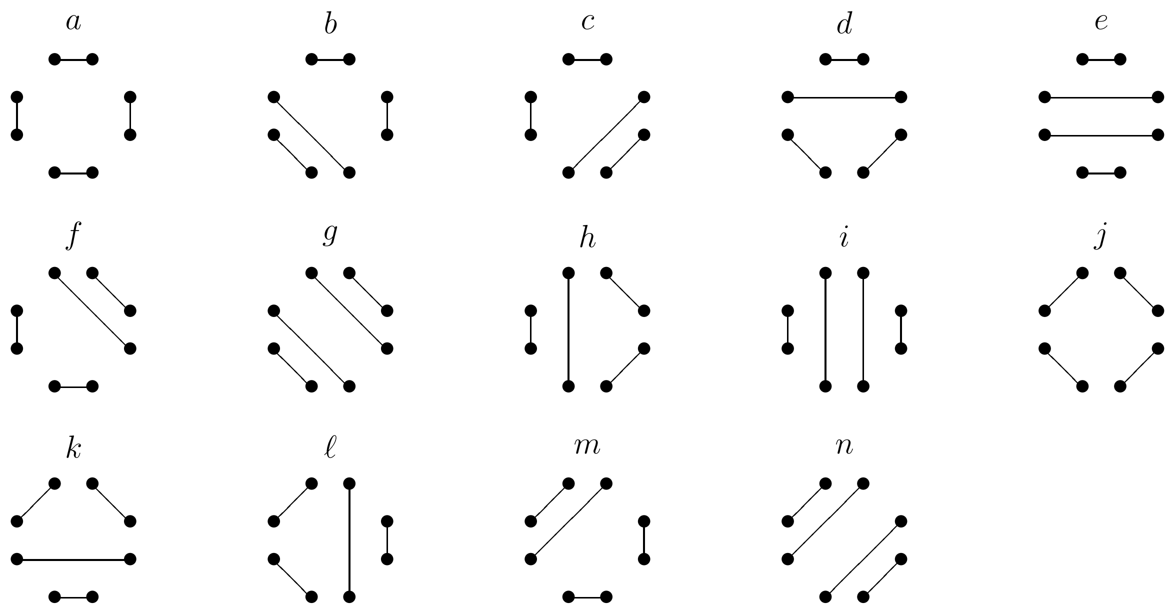

FIGURE 7. The fourteen non-crossing (undirected) graphs on $n = 8$ vertices.

crossing, choose one crossing $wx \cdot yz$ (say $\Gamma = wx \cdot yz \cdot \Gamma'$), and use the Plücker relation (2) involving $wxyz$ to express Γ in terms of two other graphs $wy \cdot xz \cdot \Gamma'$ and $wz \cdot xy \cdot \Gamma'$. Repeat this if possible. We now show that this process terminates, i.e. that this algorithm will express X_Γ in terms of upwards non-crossing variables. Both of these graphs have lower sum of edge-lengths than Γ: see Figure 8, using the triangle inequality on the two triangles with side lengths a, d, f and b, c, e. As there are finite number of graphs of weight w, and hence a finite number of possible sums of edge-lengths, the process must terminate. ☐

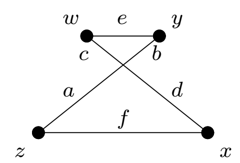

FIGURE 8. The triangle inequality implies termination of straightening: $b + c > e$, $a + d > f$

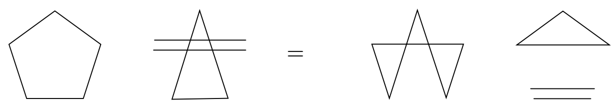

FIGURE 9. A simple binomial relation for $n = 5$

FIGURE 10. The building blocks of Figure 9

The following is well known, so we omit its proof.

2.6. *Proposition (non-crossing basis of invariants). — For each* $\mathbf{w}$*, the upwards non-crossing variables of multidegree* $\mathbf{w}$ *form a basis for* $\langle X_\Gamma \rangle_{\mathbf{deg}\,\Gamma = \mathbf{w}}$.

2.7. *Binomial (quadratic) relations.* We next describe some obvious binomial relations. If $\mathbf{deg}\,\Gamma_1 = \mathbf{deg}\,\Gamma_2$ and $\mathbf{deg}\,\Delta_1 = \mathbf{deg}\,\Delta_2$, then clearly $X_{\Gamma_1 \cdot \Delta_1} X_{\Gamma_2 \cdot \Delta_2} = X_{\Gamma_1 \cdot \Delta_2} X_{\Gamma_2 \cdot \Delta_1}$. We call these the *binomial relations*. A special case are the *simple binomial relations* when $\mathbf{deg}\,\Delta_i = (1,1,1,1,0,\dots,0) = 1^4 0^{n-4}$, or some permutation thereof. Examples are shown in Figures 9 and 13.

2.8. *Example: five points.* As an example, consider the well-known case $n = 5$, with the smallest symmetric linearization $(2,2,2,2,2)$. One of the simple binomial relations is shown in Figure 9. The building blocks Γ_i and Δ_j are shown Figure 10. These quadric relations cut out M_5 in $\mathbb{P}^5$, as can be checked directly, or as follows from Theorem 1.1. It is well-known and easy to verify that they generate the ideal of relations over the integers. The $\mathfrak{S}_5$-representation on the quadrics is visible. In terms of the non-crossing generators, we get a particularly elegant set of equations. If the generators of Figure 6 are x and y_1 through y_5 respectively, the equations are $y_{i-2} y_{i+2} = x y_i + x^2$ as $i = 1, \dots, 5$ and the subscripts are taken modulo 5.

2.9. *The Segre cubic relation (*[DoO, p. 17], [Do, Example 11.6]*).* Other relations are also clear from this graphical perspective. For example, Figure 11 shows an obvious relation for M_6. This space is well-known to be a cubic hypersurface, the Segre cubic hypersurface. As Figure 11 is a nontrivial cubic relation (this can be verified by writing it in terms of a

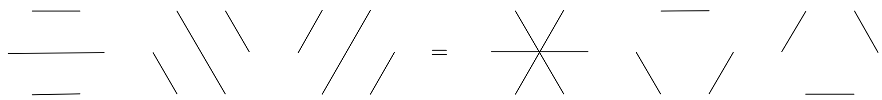

FIGURE 11. The Segre cubic relation (graphical version)

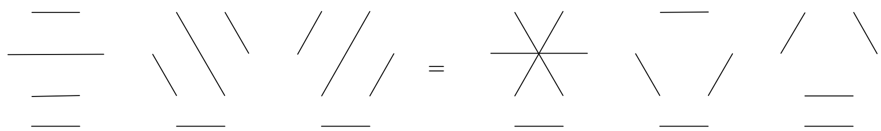

FIGURE 12. The relation s from Proposition 2.10 (a generalized Segre cubic).

non-crossing basis), it must be the Segre cubic relation. Interestingly, although the relation is not $\mathfrak{S}_6$-invariant, it becomes so modulo the Plücker relations (2). Note that there are no nontrivial binomial relations for M_6 so the Segre relation cannot be in the ideal generated by the binomial relations. The "usual" description of the Segre cubic is, in appropriate coordinates,

$$z_1 + z_2 + z_3 + z_4 + z_5 + z_6 = z_1^3 + z_2^3 + z_3^3 + z_4^3 + z_5^3 + z_6^3 = 0$$

which has the advantage over Figure 11 of obvious symmetry. On the other hand, Figure 11 has the advantage of being patently true — while even the definition of the z_i is not so simple, and is best understood in terms of the outer automorphism of $\mathfrak{S}_6$.

2.10. *Proposition. — Assume the Second Main Theorem 1.3. Let* $n = 8$ *and let* $w = (1, \ldots, 1)$ *be the symmetric linearization. Then we have the following:*

- *The simple binomial relations together with a single cubic relation* s *(depicted graphically in Figure 12) generate the ideal of relations over* $\mathbb{Z}$*. In terms of the variables of Figure 7,* s *is given by the formula* $cfi - ah(a + c + f + h + i)$*.*
- *Over* $\mathbb{Z}[1/3]$ *the cubic relation* s *lies in the ideal generated by the simple binomial relations. Therefore, when 3 is inverted, the simple binomial relations generate all relations.*

This proposition will serve as a base case for the First Main Theorem 1.1. We emphasize that the proof of the the Second Main Theorem 1.3 does not rely upon Theorem 1.1.

Proof. We prove this proposition by computation. The basic idea is as follows: we type in the 35 simple binomial relations and the relation s into the computer algebra system Magma [BCP]. We then take the ideal these relations generate, compute the quotient ring and verify that it is a free $\mathbb{Z}$-module of the correct rank in degrees 2, 3 and 4. This shows that these relations generate the full ideal of relations in those degrees and therefore generate the entire ideal by the Second Main Theorem 1.3. Details of the calculation (including the code) are available on the webpage of the third author [HMSV]. □

2.11. *Remark.* As a byproduct of our computer calculations, we obtain a particularly nice basis $\{r_i\}$ for the rank 14 space of quadric relations. We also obtain an explicit formula

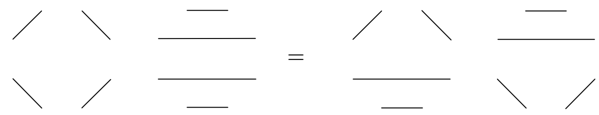

FIGURE 13. One of the simple binomial relations for $n = 8$ points

for s in terms of these relations:

$$\begin{aligned} s \;=\; & \left(-a-\tfrac{2}{3}e\right) r_1 + \left(-f-\tfrac{2}{3}g\right) r_2 + \left(\tfrac{4}{3}i\right) r_3 + \left(-c-\tfrac{2}{3}n\right) r_4 + \left(-f-\tfrac{2}{3}k\right) r_5 \\ & + \left(-a-\tfrac{2}{3}b\right) r_6 + \left(\tfrac{4}{3}h\right) r_7 + \left(\tfrac{1}{3}a-\tfrac{2}{3}m\right) r_8 + \left(-c-\tfrac{2}{3}d-h-\tfrac{2}{3}j\right) r_9 \\ & + \left(\tfrac{4}{3}f\right) r_{10} + \left(-i-\tfrac{2}{3}l\right) r_{11} + \left(\tfrac{4}{3}c\right) r_{12} + \left(\tfrac{4}{3}a\right) r_{13} + \left(-h-\tfrac{2}{3}j\right) r_{14}, \end{aligned} \tag{3}$$

Here the letters $a \dots n$ are the degree one elements from Figure 7. Now, the proposition shows that the natural map $R_1 \otimes I_2 \to I_3$ is surjective over $\mathbb{Q}$, where R_1 denotes the first piece of the ring. Since R_1 is 14-dimensional, I_2 is also 14-dimensional and I_3 is 196-dimensional (to see this count non-crossing degree three graphs to compute $\dim(R_3)$ and subtract this from the dimension of $\mathrm{Sym}^3(R_1)$) it follows that this map is an isomorphism. Thus the expression for s given above is unique. In particular, s does not lie in the ideal generated by the simple binomial relations over $\mathbb{Z}$ and so there is an essential cubic relation when 3 is not inverted.

2.12. *Generalized Segre cubic relations.* One can consider graphical cubic relations analogous to s for any $n \geq 8$, by simply adding more edges to the original Segre cubic relation; see Figure 12. Specifically, the generalized Segre relations are as follows. Choose any six indices a, b, c, d, e, f from $\{1, 2, \dots, n\}$ and any matching Δ of $\{1, \dots, n\} \setminus \{a, b, c, d, e, f\}$. Take the graphical Segre relation depicted in Figure 11 and tack on Δ to each of the six graphs. The result is a cubic relation for n points which we dub a *generalized Segre cubic*. These relations lie in the ideal generated by the simple binomial relations when 3 is inverted. This follows from the case $n = 8$, which was shown in (3). Note that as only 3 appears in the denominator of (3), the second sentence of Theorem 1.1 follows from the first.

2.13. *Example: ten points.* Here there are 42 generators and 300 quadric relations. By a nontrivial computer calculation we find that over the field of rational numbers $\mathbb{Q}$, the quadrics generate all the cubic and quartic relations, and hence by the Second Main Theorem 1.3 they generate the ideal of all relations. We will not be using this fact, so we omit the tedious details of the computer calculation. This gives credence to our conjecture that the ideal is generated by quadrics unless $n = 6$.

2.14. *Other relations.* There are other relations, that we will not discuss further. For example, consider the symmetric case for n even. Then $\mathfrak{S}_n$ acts on the set of graphs. Choose any graph Γ. Then

$$\sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) X^i_{\sigma(\Gamma)} = 0 \tag{4}$$

is a relation for i odd and $1 < i < n-1$. Reason: substituting for X's in terms of p's (or more correctly the u's and v's) using (1) to obtain an expression E, and observing that $\mathfrak{S}_n$ acts oddly on E, we see that we must obtain a multiple of the Vandermonde, which has degree $(n-1, \dots, n-1) > \deg E$. Hence $E = 0$. It is not clear that this is a nontrivial relation, but it is in small cases. In particular, the case $n = 6$, $i = 3$ is the Segre cubic relation. In the introduction, we asked if the relations (4) lie in the ideal generated by the simple binomial quadric relations.

2.15. Degree of the GIT quotient $M_{\mathbf{w}}$. As an application of these coordinates, we compute the degree of all $M_{\mathbf{w}}$, under the Kempe embedding in projective space. We will use this to verify that the degree is 1 when $|\mathbf{w}| = 6$ and $\mathbf{w} \neq (1, \dots, 1)$ (§3), although this can also be done directly.

We would like to intersect the moduli space $M_{\mathbf{w}}$ with $n-3$ coordinate hyperplanes of the form $X_\Gamma = 0$ and count the number of points, but these hyperplanes will essentially never intersect properly. Instead, we note that the intersection of each hyperplane $X_\Gamma = 0$ with $M_{\mathbf{w}}$ is reducible, and consists of a finite number of components of the form $M_{\mathbf{w}'}$ each embedded by Kempe coordinates, where the number of points $\#\mathbf{w}'$ is $n-1$. We can compute the multiplicity with which each of these components appears. The algorithm is then complete, given the base case $n = 4$. Here, more precisely, is the algorithm.

(a) *(trivial case)* If $n = 3$, the moduli space is a point, so the degree is 1.

(b) *(base case)* If $\mathbf{w} = (d, d, d, d)$, then $\deg M_{\mathbf{w}} = d$, as the moduli space is isomorphic to $\mathbb{P}^1$, embedded by the d-uple Veronese: a base-point-free subset of those variables of multidegree (d, d, d, d) are "dth powers" of variables of multidegree $(1, 1, 1, 1)$.

(c) *(main inductive step)* If $n > 4$ and $\mathbf{w}$ satisfies $w_j + w_k \leq \sum w_i/2$ for all j, k, we prove an inductive formula for the degree in terms of degrees for smaller $\mathbf{w}$. Choose any Γ of weight $\mathbf{w}$.

2.16. *Proposition. — There is a bijection between the components of $X_\Gamma = 0$ and those $j < k$ such that $w_j + w_k < \sum w_i/2$, where the component D_{jk} corresponding to (j, k) is isomorphic to $M_{\mathbf{w}'}$, where $\mathbf{w}'$ is the same as $\mathbf{w}$ except w_j and w_k are removed, and $w_j + w_k$ is added. The component D_{jk} appears with multiplicity equal to the number m_{jk} of edges joining j and k in Γ.*

Proof. Consider the morphism $\pi : (\mathbb{P}^1)^n - U_{\mathbf{w}} \to [X_\Gamma]_{\deg \Gamma = \mathbf{w}}$, where $U_{\mathbf{w}}$ is the unstable locus. By the definition (1) of X_Γ,

$$\pi^* X_\Gamma = \prod_{j<k} (u_j v_k - u_k v_j)^{m_{jk}}. \tag{5}$$

For each (j, k), the Weil divisor $D'_{jk} = \{u_j v_k - u_k v_j = 0\}$ on $(\mathbb{P}^1)^n - U_{\mathbf{w}}$ is isomorphic to $(\mathbb{P}^1)^{n-1} - U_{\mathbf{w}'}$, and the linearization $\mathcal{O}(\mathbf{w})$ on $(\mathbb{P}^1)^{n-1}$ restricts to $\mathcal{O}(\mathbf{w}')$ on this locus. This Weil divisor D'_{jk} appears with multiplicity m_{jk} in the Cartier divisor $\pi^* X_\Gamma$ by (5). The map $\pi : D'_{jk} \to D_{jk}$ is precisely the GIT quotient corresponding to $n-1$ points with weight $\mathbf{w}'$. (Indeed, we can even identify the graphical variables. For each Γ' of multidegree $\mathbf{w}'$, we lift $X_{\Gamma'}$ to any X_Γ where Γ is a graph on $\{1, \dots, n\}$ of multidegree $\mathbf{w}$ whose image in

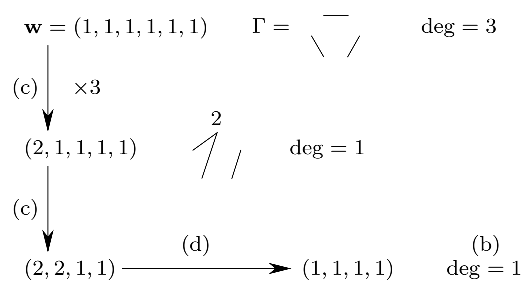

FIGURE 14. Computing $\deg M_6 = 3$ (recall that M_6 is the Segre cubic threefold)

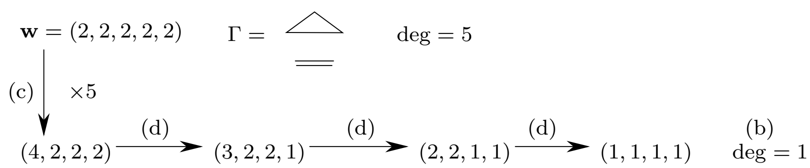

FIGURE 15. Computing $\deg M_{(2,2,2,2,2)} = 5$ using an inconvenient choice of Γ (recall that $M_{(2,2,2,2,2)}$ is a degree 5 del Pezzo surface)

$\{1, \dots, n\} \cup \{0\} \setminus \{j, k\}$ is Γ'. In other words, to w_j of the $w_j + w_k$ edges meeting vertex 0 in Γ', we associate edges meeting vertex j in Γ, and similarly with j and k interchanged. If Γ'' is any other lift, then $X_\Gamma = \pm X_{\Gamma''}$ on D'_{jk}, because using the Plücker relations, $X_\Gamma \pm X_{\Gamma''}$ can be expressed as a combination of variables containing edge jk, which all vanish on D'_{jk}.)

Note that if $w_j + w_k = \sum w_i/2$, then $\pi(D'_{jk}) = M_{\mathbf{w}'} \subset M_{\mathbf{w}}$ is the image of a strictly semistable point, and of dimension $0 < \dim M_{\mathbf{w}} - 1$, and hence is not a component. Our base case is $n = 4$, not 3, for this reason. □

(d) *(reduction of "base locus" case)* If $n \geq 4$ and there are j and k such that $w_j + w_k > \sum w_i/2$, then the rational map $(\mathbb{P}^1)^n \dashrightarrow M_{\mathbf{w}}$ has a base locus. Any graph X_Γ of degree **w** necessarily contains a copy of edge jk, so $(u_j v_k - u_k v_j)$ is a factor of every X_Γ. Hence $M_{\mathbf{w}}$ (and its Kempe embedding) is naturally isomorphic to $M_{\mathbf{w}-e_j-e_k}$ (and its Kempe embedding), so we replace **w** by $\mathbf{w} - e_j - e_k$, and repeat the process. Note that if $n = 4$, then the final resulting quadruple must be of the form (d, d, d, d).

For example, $\deg M_4 = 1$, $\deg M_6 = 3$, $\deg M_8 = 40$, and $\deg M_{10} = 1225$ were computed by hand. (This appears to be sequence A012250 on Sloane's *On-line encyclopedia of integer sequences* [Sl].) The calculations $\deg M_6 = 3$ and $\deg M_{2,2,2,2,2} = 5$ are shown in Figures 14 and 15 respectively. At each stage, **w** is shown, as well as the Γ used to calculate the next stage. In these examples, there is essentially only one such $\mathbf{w}'$ at each stage, but in general there will be many. The vertical arrows correspond to identifying components of X_Γ (step (c)). The first arrow in Figure 14 is labeled $\times 3$ to point out the reader that the next stage can be obtained in three ways. The degrees are obtained inductively from the bottom up. The reader is encouraged to show that $\deg M_8 = 40$, and that this algorithm indeed gives $\deg M_{d\mathbf{w}} = d^{n-3} \deg M_{\mathbf{w}}$.

2.17. Reduction to the symmetric case (proof of Theorem 1.2). We next prove Theorem 1.2, hence reducing questions about relations in general weight to the symmetric case. The argument is similar in spirit to our proof of Kempe's Theorem 2.3.

Suppose that $\mathbf{w} = (w_1, \ldots, w_m)$ and $|\mathbf{w}| = n$ is even. Then $\mathbf{w}$ defines a natural partition of $\{1, \ldots, n\}$ into m parts which we call *clumps*. The first clump is $\{1, \ldots, w_1\}$, the second clump is $\{w_1 + 1, \ldots, w_1 + w_2\}$, etc. For example, if $\mathbf{w} = (2, 2, 1, 1)$ then it partitions $\{1, 2, 3, 4, 5, 6\}$ into four clumps; $\{1, 2\}$, $\{3, 4\}$, $\{5\}$ and $\{6\}$. We shall also be interested in the case where all w_i are even. Then $\mathbf{w}$ naturally partitions $\{1, \ldots, n/2\}$ into m clumps: the first clump is $\{1, \ldots, w_1/2\}$, the second clump is $\{w_1/2 + 1, \ldots, w_1/2 + w_2/2\}$, etc. In what follows, the weight $\widetilde{\mathbf{w}}$ can be either 1^n or $2^{n/2}$, but the latter case is only allowed if each w_i is even.

Consider the commutative diagram:

$$\begin{array}{ccccccccc} 0 & \longrightarrow & I_{\widetilde{\mathbf{w}}} & \longrightarrow & \mathbb{Z}[\widetilde{X}_\Gamma] & \xrightarrow{\phi} & R_{\widetilde{\mathbf{w}}} & \longrightarrow & 0 \\ & & \downarrow{\scriptstyle \alpha} & & \downarrow{\scriptstyle \mathbf{e}} & & \downarrow{\scriptstyle \gamma} & & \\ 0 & \longrightarrow & I_{\mathbf{w}} & \longrightarrow & \mathbb{Z}[\widetilde{X}_\Delta] & \xrightarrow{\psi} & R_{\mathbf{w}} & \longrightarrow & 0, \end{array}$$

where $R_{\widetilde{\mathbf{w}}}$ (resp. $R_{\mathbf{w}}$) is the coordinate ring associated to the weight $\widetilde{\mathbf{w}}$ (resp. $\mathbf{w}$), the Γ's range over all graphs of multidegree $\widetilde{\mathbf{w}}$, and the Δ's range over all graphs of multidegree $\mathbf{w}$. The $\widetilde{X}_\Gamma$'s and $\widetilde{X}_\Delta$' s are formal variables, and the surjective map ϕ (resp. ψ) is given by $\widetilde{X}_\Gamma \mapsto X_\Gamma$ (resp. $\widetilde{X}_\Delta \mapsto X_\Delta$). The map $\mathbf{e}$ takes $\widetilde{X}_\Gamma$ to $\widetilde{X}_\Delta$ where Δ is given by identifying vertices of Γ within the same clump; if Δ has a loop then $\widetilde{X}_\Delta = 0$ by convention.

2.18. *Theorem. — The map $\alpha : I_{\widetilde{\mathbf{w}}} \to I_{\mathbf{w}}$ is surjective.*

In other words, all relations for $M_{\mathbf{w}}$ are "inherited" from $M_{\widetilde{\mathbf{w}}}$.

Proof. It is clear that $\mathbf{e}$ is surjective, and hence γ too. By the Snake Lemma,

$$\ker(\alpha) \to \ker(\mathbf{e}) \to \ker(\gamma) \to \operatorname{coker}(\alpha) \to \operatorname{coker}(\mathbf{e}) \to \operatorname{coker}(\gamma)$$

is exact, where $\ker(\gamma) \to \operatorname{coker}(\alpha)$ is the connecting homomorphism. We know that $\operatorname{coker}(\mathbf{e}) = 0$ so to show α is surjective, it suffices to prove that $\ker \mathbf{e} \to \ker \gamma$ is surjective.

We have that $R_{\widetilde{\mathbf{w}}} = \oplus_G \mathbb{Z} \cdot X_G$, where G ranges over upwards non-crossing graphs of multidegree $N\widetilde{\mathbf{w}}$ for some $N \geq 0$ — such X_G form a $\mathbb{Z}$-basis of $R_{\widetilde{\mathbf{w}}}$. Similarly $R_{\mathbf{w}} = \oplus_H \mathbb{Z} \cdot X_H$ as H ranges over upwards non-crossing graphs of multidegree $N\mathbf{w}$ for some $N \geq 0$. We claim that for each upwards non-crossing H, there is exactly one upwards non-crossing G for which $\gamma(X_G) = X_H$. We'll just prove the claim for the case $\mathbf{w} = 1^n$ and leave the other case to the reader. Recall H has vertices oriented clockwise around the unit circle with all edges drawn inside the circle, where no two edges cross. We number the edges $\epsilon_i(1), \ldots, \epsilon_i(w_i)$ incident with vertex i, so that if $\epsilon_i(s) = ai$, $a < i$, and $\epsilon_i(t) = ib$, $i < b$, then $s < t$. Also, if $\epsilon_i(s) = ai$ and $\epsilon_i(s + 1) = bi$ then $a \geq b$, and if $\epsilon_i(s) = ia$ and $\epsilon_i(s + 1) = ib$ then $a \geq b$. To obtain G from H we split the vertices of H into several vertices. Specifically, we split vertex i of H into w_i vertices $i_1, \ldots, i_{w_i}$ of G. We place the vertices $i_1, \dots, i_{w_i}$ in clockwise orientation. We must attach edge $\epsilon_i(s)$ to vertex i_s in G, for otherwise we would introduce a pair of crossing edges in G. For example if we were to attach edge $\epsilon_i(s)$ to vertex i_t and $s > t$, then there must be some $s' > s$ and $t' < t$, such that $\epsilon_i(s')$ is attached to $i_{t'}$. But then the edges $\epsilon_i(s)$ and $\epsilon_i(s')$ cross each other. This shows uniqueness of G. We also claim that by attaching $\epsilon_i(s)$ to i_s the resulting graph G is noncrossing. Suppose that $i_s j_t$ and $i'_{s'} j'_{t'}$ are edges of G. By way of contradiction suppose that $i_s < i'_{s'}$ and $j_t < j'_{t'}$. If i, i', j, j' are distinct we have a contradiction, since then the edges ij and $i'j'$ are crossing edges in H. By construction of G we know that $i < j$ and $i' < j'$. First suppose that $i = i'$. Hence $s < s'$. But then $j \geq j'$, and since $j_t < j'_{t'}$ we must have $j = j'$ and $t < t'$. However $i_{s'} j$ should precede $i_s j$ in the numbering of edges at vertex j, and so we would have instead edges $i_{s'} j_t$, $i_s j_{t'}$ in G, a contradiction. The case $j = j'$ is similar.

Thus $\ker(\gamma) = \oplus_G \mathbb{Z} \cdot X_G$ where the sum is over those upwards non-crossing G which contain at least one edge which connects two vertices in a single clump. Fix such a G, with an edge ab where a and b are in the same clump.

First we treat the case that $\widetilde{\mathbf{w}} = 1^n$. Partition $\{1, \dots, n\}$ into two equal sized subsets A and B ("positive" and "negative") such that $a \in A$, and $b \in B$. As in the proof of Kempe's theorem 2.3, we can write $X_G = \sum_i X_{\Gamma_i}$, where the Γ_i are bipartite graphs and each Γ_i contains the edge ab. (The process described in the proof of Kempe's Theorem 2.3 involves trading a pair of edges, one positive and one negative, for two neutral edges. No neutral edges such as ab are affected by this process.)

By applying Hall's marriage theorem repeatedly to Γ_i, we can write $X_{\Gamma_i} = \prod_{j=1}^k X_{\Gamma_{i,j}} = \phi(\prod_{j=1}^k \widetilde{X}_{\Gamma_{i,j}})$, where the $\Gamma_{i,j}$ are matchings. There exists some j such that $\Gamma_{i,j}$ contains the edge ab, so $\mathbf{e}(\prod_{j=1}^k \widetilde{X}_{\Gamma_{i,j}}) = 0$. This holds for each i. Therefore $\mathbf{e}\Big(\sum_i \prod_{j=1}^k \widetilde{X}_{\Gamma_{i,j}} \Big) = 0$, and also $\phi\Big(\sum_i \prod_{j=1}^k \widetilde{X}_{\Gamma_{i,j}} \Big) = X_G$. Hence $\ker \mathbf{e}$ surjects onto $\ker \gamma$.

Now suppose that $\widetilde{\mathbf{w}} = 2^{n/2}$ and each vertex of G has valence $2v$. Choose an auxiliary v-regular graph G' with n vertices that maps to G by "clumping" vertices ℓ and $\ell + 1$ for each odd ℓ, $1 \leq \ell \leq n - 1$ (call this map on graphs ρ). Let $a'b'$ be an edge of G' which maps to the edge ab of G. As above, write $X_{G'} = \sum_i X_{\Gamma'_i}$ such that each $X_{\Gamma'_i} = \prod_{j=1}^v X_{\Gamma'_{i,j}}$ and each $\Gamma'_{i,j}$ is an n-matching, where for all i, there exists j such that $\Gamma'_{i,j}$ contains the edge $a'b'$. Let $\Gamma_i = \rho(\Gamma'_i)$ and $\Gamma_{i,j} = \rho(\Gamma'_{i,j})$ for each i, j. Now, $X_G = \sum_i \prod_{j=1}^v X_{\Gamma_{i,j}}$, each $\Gamma_{i,j}$ is a 2-regular graph on $n/2$ vertices, and for each i there is a j such that $\Gamma_{i,j}$ contains the edge ab. As before, $\mathbf{e}\Big(\sum_i \prod_{j=1}^v \widetilde{X}_{\Gamma_{i,j}} \Big) = 0$. Hence $\ker \mathbf{e}$ surjects onto $\ker \gamma$. $\square$

Part 2. THE MODULI SPACE OF n POINTS ON THE LINE IS CUT OUT BY SIMPLE QUADRICS WHEN n IS NOT SIX

In Part 2, we prove the First Main Theorem 1.1. We begin by verifying it in small cases. We then show the result in a neighborhood of a semistable point by explicit calculation. These neighborhoods do not, unfortunately, cover the entire projective space, so in §5 we show the result set-theoretically, and scheme-theoretically *away* from the strictly semistable points. This section is the most difficult part of the proof. (Note: We work over $\mathbb{Z}$ except where specifically noted.)

3. VERIFICATION OF THE FIRST MAIN THEOREM 1.1 IN SMALL CASES

The cases $|\mathbf{w}| = 2$ and $|\mathbf{w}| = 4$ are trivial. If $|\mathbf{w}| = 6$ and $\mathbf{w} \neq (1, \dots, 1)$, then $\mathbf{w} = (3, 2, 1)$, $(2, 2, 2)$, $(2, 2, 1, 1)$, or $(2, 1, 1, 1, 1)$. The first two cases are points, and the next two cases were verified to have degree 1 in §2.15 (see Figure 14). The case $\mathbf{w} = (1, 1, 1, 1, 1, 1, 1, 1)$ was verified in §2.10, so by §2.17, the case $|\mathbf{w}| = 8$ follows. Thus the cases $|\mathbf{w}| \geq 10$ remain.

4. AN ANALYSIS OF A NEIGHBORHOOD OF A STRICTLY SEMISTABLE POINT

We now show the result in a neighborhood of a strictly semistable point, in the symmetric case $\mathbf{w} = 1^{n=2m}$, by explicitly describing an affine neighborhood of such a point. This affine neighborhood has a simple description: it is the space of $(m-1) \times (m-1)$ matrices of rank at most 1, where no entry is 1 (Lemma 4.3). The strictly semistable point corresponds to the zero matrix.

4.1. The Gel'fand-MacPherson correspondence: the moduli space as a quotient of the Grassmannian. We begin by recalling the Gel'fand-MacPherson correspondence, an alternate description of the moduli space. The Plücker embedding of the Grassmannian $G(2, n) \hookrightarrow \mathbb{P}^{\binom{n}{2}-1}$ is via the line bundle $\mathcal{O}(1)$ that is the positive generator for $\operatorname{Pic} G(2, n)$. This generator may be described explicitly as follows. Over $G(2, n)$ we have a tautological exact sequence of vector bundles

$$0 \longrightarrow S \longrightarrow \mathcal{O}^{\oplus n} \longrightarrow Q \longrightarrow 0 \qquad (6)$$

where S is the tautological rank 2 subbundle and Q is the tautological rank $n-2$ quotient bundle. Then $\wedge^2 S = \mathcal{O}(-1)$ is a line bundle, and is the dual to $\mathcal{O}(1)$. Dualizing (6) we get a map $\wedge^2 \mathcal{O}^{\oplus n} \to \wedge^2 S^*$. It can be easily checked that $\wedge^2 S^*$ is generated by the resulting global sections. We call these sections s_{ij}, and note that they satisfy the following relations: the sign relations $s_{ij} = -s_{ji}$ inherited from $\wedge^2 \mathcal{O}^{\oplus n}$ (so $s_{ii} = 0$), and the Plücker relations

$$s_{ij}s_{kl} - s_{ik}s_{jl} - s_{jk}s_{il} = 0.$$

These equations cut out the Grassmannian in $\mathbb{P}^{\binom{n}{2}-1}$.

The connection to n points on $\mathbb{P}^1$ is as follows. Given a general point of the Grassmannian corresponding to the subspace Λ of n-space, we obtain n points on $\mathbb{P}^1$ by considering the intersection of Λ with the n coordinate hyperplanes and projectivizing. This breaks down if Λ is contained in a coordinate hyperplane. (The point $[\Lambda]$ is GIT-stable if the resulting n points in $\mathbb{P}^1$ are GIT-stable, and similarly for semistable. We recover the cross-ratio of four points via $s_{ij}s_{kl}/s_{il}s_{jk}$.)

Let $D(s_{1n})$ be the distinguished open set of the Grassmannian where $s_{1n} \neq 0$. In the correspondence with marked points, this corresponds to the locus where the first point is distinct from the last point. Then $D(s_{1n})$ is isomorphic to $\mathbb{A}^{2(n-2)}$, with good coordinates as follows. Given $\Lambda \notin D(s_{1n})$, choose a basis for Λ, written as a $2 \times n$ matrix. As $\Lambda \notin D(s_{1n})$, the first and last columns are linearly independent, so up to left-multiplication by GL_2 there is a unique way to choose a basis where the first column is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and the last column is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. We choose the "anti-identity" matrix rather than the identity matrix because we will think of the first column as $[0;1] \in \mathbb{P}^1$ and the last column as $[1;0]$. Another interpretation is as follows. If Λ is interpreted as a line in $\mathbb{P}^{n-1}$, and $H_1, \dots, H_n$ are the coordinate hyperplanes, then if Λ does not meet $H_1 \cap H_n$, then it meets H_1 at one point of $H_1 - H_1 \cap H_n$ and H_n at one point of $H_n - H_1 \cap H_n$, and Λ is determined by these two points. The coordinates on the first space are the x's, and the coordinates on the second are the y's.

Thus if the $2 \times n$ matrix is written

$$\begin{bmatrix} 0 & x_2 & x_3 & \cdots & x_{n-1} & 1 \\ 1 & y_2 & y_3 & \cdots & y_{n-1} & 0 \end{bmatrix}$$

then we have coordinates $x_2, \dots, x_{n-1}, y_2, \dots, y_{n-1}$ on our affine chart. For convenience, we define $x_1 = 0$, $y_1 = 1$, $x_n = 1$, $y_n = 0$.

Under the trivialization $(\mathcal{O}(1), s_{1n})|_{D(s_{1n})} \cong (\mathcal{O}, 1)|_{D(s_{1n})}$, in these coordinates the section s_{ij} may be interpreted as

$$s_{ij} = x_j y_i - x_i y_j.$$

We can use this to immediately verify the Plücker relations. We also recover the x_i and y_j from the sections via

$$x_i = s_{1i}/s_{1n} \qquad y_j = s_{jn}/s_{1n}. \tag{7}$$

The Grassmannian has dimension $2(n-2) = 2n-4$. To obtain our moduli space, we take the quotient of $G(2,n)$ by the maximal torus $T \subset SL_n$, which has dimension $n-1$. Thus as expected the quotient has dimension $n-3$. We will write elements of this maximal torus as $\lambda = (\lambda_1, ..., \lambda_n)$. To describe the linearization, we must describe how λ acts on each $s_{k\ell}$: λ_i acts on s_{ij} with weight 1, and on the rest of the $s_{k\ell}$'s by weight 0. This action certainly preserves our relations.

Then we can see how to construct the quotient as a Proj: the terms that have weight $(d, d, \dots, d)$ correspond precisely to d-regular graphs on our n vertices. Hence we conclude that this projective scheme is precisely the GIT quotient of n points on the projective line, as the graded rings are the same. This is the Gel'fand-MacPherson correspondence. The relations we have described on our X_Γ clearly come from the relations on the Grassmannian. That is of course no guarantee that we have them all!

4.2. A neighborhood of a strictly semistable point. We remind the reader that we are currently considering the symmetric case $\mathbf{w} = 1^n = 1^{2m}$. Let $\pi : G(2, 2m)^{ss} \to M_{\mathbf{w}}$ be the quotient map. Let p be the image of a strictly semistable point of the moduli space $M_{\mathbf{w}}$, without loss of generality the image of $(0, \ldots, 0, \infty, \ldots, \infty)$. We say an edge ij on vertices $\{1, \ldots, 2m\}$ is *good* if $i \leq m < j$ (if it "doesn't connect two 0's or two ∞'s"). We say a graph on $\{1, \ldots, 2m\}$ is *good* if all of its edges are good. We say an edge or graph is *bad* if it is not good. Let P be the set of good matchings of $\{1, \ldots, 2m\}$. Let

$$U_P = \{q \in M_{\mathbf{w}} : X_\Gamma(q) \neq 0 \text{ for all } \Gamma \in P\}.$$

In the dictionary to n points on $\mathbb{P}^1$, this corresponds to the set where none of the first m points is allowed to be the same as any of the last m points. Note that $p \in P$, and $\pi^{-1}(U_P) \subset D(s_{1,2m})$.

4.3. *Lemma. — U_P is an affine variety, with coordinate ring generated by W_{ij} and Z_{ij} ($1 < i \leq m < j < 2m$) with relations*

$$W_{ij}W_{kl} = W_{il}W_{kj} \tag{8}$$

(i.e. the matrix $[W_{ij}]$ has rank 1) and

$$Z_{ij}(1 - W_{ij}) = 1 \tag{9}$$

(i.e. the matrix $[W_{ij}]$ has no entry 1).

This has a simple interpretation: U_P is isomorphic to the space of $(m-1) \times (m-1)$ matrices of rank at most 1, where each entry differs from 1, and p is the unique singular point, corresponding to the zero matrix. This is an open subset of the cone over the Segre embedding of $\mathbb{P}^{m-1} \times \mathbb{P}^{m-1}$. Hence we have described a neighborhood of the singular point rather explicitly.

Proof. Let $V_P = \{[\Lambda] \in G(2, 2m) : s_{ij}(\Lambda) \neq 0 \text{ for all } i \leq m < j\}$, so $V_P = \pi^{-1}(U_P)$. Then V_P is an open subset of $D(s_{1,2m})$. In terms of the coordinates on $D(s_{1,2m}) \cong \mathbb{A}^{4m-4}$ described above, V_P is described by

$$x_j y_i - x_i y_j \neq 0, \qquad x_j \neq 0, \qquad y_i \neq 0 \tag{10}$$

for $i \leq m < j$. Let T be the maximal torus of diagonal matrices within the special linear group SL_{2m}. We realize T by $2m$-tuples of variables $(\lambda_1, \ldots, \lambda_{2m})$ such that $\prod_{i=1}^{2m} \lambda_i = 1$. The action of T on the coordinates x_i, y_i, $1 < i < 2m$ is the following:

$$(\lambda_1, \ldots, \lambda_{2m}) \cdot x_i = (\lambda_{2m}/\lambda_i)x_i,$$

$$(\lambda_1, \ldots, \lambda_{2m}) \cdot y_i = (\lambda_1/\lambda_i)y_i.$$

Since T is a torus and the coordinates x_i, y_i are generalized eigenvectors of T, the T invariant subalgebra of the polynomial ring in the variables x_i, y_i is generated by monomials. Now consider a Laurent monomial

$$\mathbf{m} = \prod_{i=2}^{2m-1} {x_i}^{a_i} {y_i}^{b_i}.$$

Here, the exponents $a_2, \ldots, a_m$ and $b_{m+1}, \ldots, b_{2m-1}$ are non-negative but the other exponents may be negative since they correspond to the variables which are units. We have

$$(\lambda_1, \ldots, \lambda_{2m}) \cdot \mathbf{m} = \left(\frac{\lambda_1^{b_2+b_3+\cdots+b_{2m-1}} \lambda_{2m}^{a_2+a_3+\cdots+a_{2m-1}}}{\lambda_2^{a_2+b_2} \lambda_3^{a_3+b_3} \cdots \lambda_{2m-1}^{a_{2m-1}+b_{2m-1}}} \right) \mathbf{m}.$$

Hence $\mathbf{m}$ is T-invariant if and only if the two rows of the matrix

$$\begin{pmatrix} a_2 & \cdots & a_m & a_{m+1} & \cdots & a_{2m-1} \\ b_2 & \cdots & b_m & b_{m+1} & \cdots & b_{2m-1} \end{pmatrix}$$

have a common sum s and all column sums $a_i + b_i$ equal $-s$. However the total sum of all the entries is then $2s = -(2m-2)s$, and so $s = 0$. Therefore the T-invariant $\mathbf{m}$'s are given by integer tuples $(a_2, \ldots, a_{2m-1})$ such that $a_2, \ldots, a_m \geq 0$, $a_{m+1}, \ldots, a_{2m-1} \leq 0$, and $a_2 + \cdots + a_{2m-1} = 0$. Indeed we satisfy the above conditions by setting $b_i = -a_i$ for each i. Such tuples $(a_2, \ldots, a_{2m-1})$ are clearly generated by those for which all $a_k = 0$ except a single $a_i = 1$ for $i \leq m$, and a single $a_j = -1$ for $j \geq m+1$. Thus the T-invariant monomials are generated by the $(m-1)^2$ monomials $W_{ij} = x_i y_i^{-1} x_j^{-1} y_j$, for $2 \leq i \leq m$ and $m+1 \leq j \leq 2m-1$. The condition $s_{ij} \neq 0$ for $2 \leq i \leq m < j \leq 2m-1$ is equivalent to $W_{ij} \neq 1$. Also the $(m-1) \times (m-1)$ matrix of W_{ij}'s is the tensor $[x_2/y_2, \ldots, x_m/y_m] \otimes [y_{m+1}/x_{m+1}, \ldots, y_{2m-1}/x_{2m-1}]$ and so it has rank at most one. □

Now let $I \subset \mathbb{Z}[X_\Gamma]$ be the ideal of relations of the invariants of $M_{\mathbf{w}}$, and let I_V be the ideal generated by the linear Plücker relations, the simple binomial relations, and the generalized Segre cubic relations. We have already shown that $I_V \subset I$. (Note: once 3 is inverted, I_V is generated by the linear Plücker relations and simple binomial relations, using (3).)

Let S be the multiplicative system of monomials in X_Γ generated by those X_Γ where $\Gamma \in P$.

4.4. *Theorem. — If $n = 2m \geq 8$, then $S^{-1}I_V = S^{-1}I$. In other words, the sign, Plücker, simple binomial, and generalized Segre cubic relations cut out the moduli space on this open subset.*

As the First Main Theorem 1.1 is true for $n = 2m = 8$ (§2.10), Theorem 4.4 holds in that "base" case.

Proof. The idea of the proof is to compare the minimal presentation given in Lemma 4.3 for $S^{-1}(R_n)$ with the localization in terms of ratios X_Γ/X_Δ in n-matchings Γ, Δ. This will be done by embedding the polynomial ring in the W_{ij}'s and Z_{ij}'s into $\mathbb{Z}[X_\Gamma/X_\Delta]$, and show that the contractions of $S^{-1}(I)$ and $S^{-1}(I_V)$ to the image of this embedding agree. We will also show that *any* element of $\mathbb{Z}[X_\Gamma/X_\Delta]$ differs from an element of the embedded subring by an element of $S^{-1}I_V$.

By Γ we will mean a general matching, and by Δ, we will mean a matching in P. We have a surjective map

$$\mathbb{Z}[X_\Gamma/X_\Delta]/S^{-1}I_V \to \mathbb{Z}[X_\Gamma/X_\Delta]/S^{-1}I$$

(that we wish to show is an isomorphism), and Lemma 4.3 provides an isomorphism

$$\mathbb{Z}[X_\Gamma/X_\Delta]/S^{-1}I \cong \mathcal{O}(U_P) \cong \mathbb{Z}[W_{ij}, Z_{ij}]/J_{WZ},$$

where $J_{WZ} \subset \mathbb{Z}[W_{ij}, Z_{ij}]$ is the ideal generated by the relations (8) and (9). By comparing the moduli maps, we see that this isomorphism is given by

$$W_{ij} \mapsto \frac{X_{1i\cdot j(2m)\cdot\Gamma}}{X_{1j\cdot i(2m)\cdot\Gamma}}, \qquad Z_{ij} \mapsto \frac{X_{1j\cdot i(2m)\cdot\Gamma}}{X_{1(2m)\cdot ij\cdot\Gamma}} \tag{11}$$

where Γ is any matching on $\{1,\ldots,2m\} - \{1,i,j,2m\}$ such that $1j\cdot i(2m)\cdot\Gamma \in P$. By the simple binomial relations, this is independent of Γ. The description of the isomorphism in the reverse direction is not so pleasant, and we will spend much of the proof avoiding describing it explicitly.

We thus have a surjective map

$$\psi : \mathbb{Z}[X_\Gamma/X_\Delta] \to \mathbb{Z}[W_{ij}, Z_{ij}]/J_{WZ}$$

whose kernel is $S^{-1}I$, which contains $S^{-1}I_V$. We wish to show that the kernel is $S^{-1}I_V$. We do this as follows. For each $1 < i \le m < j < 2m$, fix a matching $\Gamma_{i,j}$ on $\{1,\ldots,2m\} - \{1,i,j,2m\}$ so that $1j\cdot i(2m)\cdot\Gamma_{i,j} \in P$. Consider the subring of $\mathbb{Z}[X_\Gamma/X_\Delta]$ generated by

$$w_{ij} = \frac{X_{1i\cdot j(2m)\cdot\Gamma_{i,j}}}{X_{1j\cdot i(2m)\cdot\Gamma_{i,j}}}, \qquad z_{ij} = \frac{X_{1j\cdot i(2m)\cdot\Gamma_{i,j}}}{X_{1(2m)\cdot ij\cdot\Gamma_{i,j}}}.$$

(Compare this to (11). Note that we want $W_{ij} \mapsto w_{ij}$.) Call this subring $\mathbb{Z}[w_{ij}, z_{ij}]/J_{wz}$.

The proof consists of two steps. *Step 1.* We show that *any* element of $\mathbb{Z}[X_\Gamma/X_\Delta]$ differs from an element of $\mathbb{Z}[w_{ij}, z_{ij}]/J_{wz}$ by an element of $S^{-1}I_V$. We do this in several smaller steps. *Step 1a.* We show that any X_Γ/X_Δ can be written as a linear combination of $X_{\Delta'}/X_\Delta$, where Δ' is also good. *Step 1b.* We show that any such $X_{\Delta'}/X_\Delta$ may be expressed modulo $S^{-1}I_V$ in terms of $X_{ik\cdot jl\cdot\Gamma}/X_{il\cdot jk\cdot\Gamma}$, where $i,j \le m < k,l$, and Γ is good. *Step 1c.* We show that any such expression can be written modulo $S^{-1}I_V$ in terms of w_{ij} and z_{ij}, i.e. modulo $S^{-1}I_V$, any such expression lies in $\mathbb{Z}[w_{ij}, z_{ij}]/J_{wz}$.

Step 2. The kernel of the map $\psi : \mathbb{Z}[w_{ij}, z_{ij}]/J_{wz} \to \mathbb{Z}[W_{ij}, Z_{ij}]/J_{WZ}$ given by $w_{ij} \mapsto W_{ij}$, $z_{ij} \mapsto Z_{ij}$ lies in $S^{-1}I_V$.

We now execute this strategy.

Step 1a. We first claim that X_Γ/X_Δ ($\Delta \in P$) is an integral combination of units $X_{\Delta'}/X_\Delta$ (i.e. $\Delta' \in P$) modulo the Plücker relations (the linear relations, which are in $S^{-1}I_V$). We prove the result by induction on the number of bad edges. The base case — if all edges of Γ are good, i.e. $\Gamma \in P$ — is immediate. Otherwise, Γ has at least two bad edges, say ij and kl, where $i,j \le m < k,l$. Then $X_\Gamma = \pm X_{\Gamma-\{ij,kl\}+\{ik,jl\}} \pm X_{\Gamma-\{ij,kl\}+\{il,jk\}}$ is a Plücker relation, and the latter two terms have two fewer bad edges, completing the induction.

Step 1b. We show that any element $X_{\Delta'}/X_\Delta$ of $\mathbb{Z}[X_\Gamma/X_\Delta]$ (Δ' good) is congruent modulo $S^{-1}I_V$ to an element of the form $X_{ik\cdot jl\cdot\Gamma}/X_{il\cdot jk\cdot\Gamma}$, where $i,j \le m < k,l$, and Γ is good. We prove this by induction on m. If $m = 4$, the result is true (§2.10: the simple binomial relations cut out the quotient scheme-theoretically, and indeed generate the ideal of relations). Assume now that $m > 4$. If Δ' and Δ share an edge e, then let $\overline{\Delta'}$ and $\overline{\Delta}$ be the graphs on $2m-2$ vertices obtained by removing this edge e. Then by the inductive hypothesis, the result holds for $X_{\overline{\Delta'}}/X_{\overline{\Delta}}$. By taking the resulting expression, and "adding edge e to the subscript of each term," we get an expression for $X_{\Delta'}/X_{\Delta}$. Finally, if Δ' and Δ share no edge, suppose in Δ', 1 is connected to $(m+1)$; in Δ, 1 is connected to $(m+2)$; and in Δ', $(m+2)$ is connected to 2. This is true after suitable reordering. Say $\Delta' = 1(m+1)\cdot 2(m+2)\cdot \Gamma'$ and $\Delta = 1(m+2)\cdot\Gamma$. Then

$$\frac{X_{\Delta'}}{X_{\Delta}} = \frac{X_{1(m+2)\cdot 2(m+1)\cdot\Gamma'}}{X_{1(m+2)\cdot\Gamma}} \cdot \frac{X_{1(m+1)\cdot 2(m+2)\cdot\Gamma'}}{X_{1(m+2)\cdot 2(m+1)\cdot\Gamma'}}.$$

For each factor of the right side, the numerator and the denominator "share an edge", so we are done.

Step 1c. We next show that any such ratio $X_{ik\cdot jl\cdot\Gamma}/X_{ij\cdot kl\cdot\Gamma}$ as in Step 1b can be written modulo $S^{-1}I_V$ in terms of w_{ij} and z_{ij}, i.e. modulo $S^{-1}I_V$ lies in $\mathbb{Z}[w_{ij}, z_{ij}]/J_{wz}$. If $2m=8$, the result again holds (§2.10). Assume now that $2m>8$. Given any $X_{ik\cdot jl\cdot\Gamma}/X_{ij\cdot kl\cdot\Gamma}$ as in Step 1b, we will express it modulo $S^{-1}I_V$ in terms of w_{ij} and z_{ij}. By the simple binomial relation (i.e. modulo $S^{-1}I_V$), we may assume that Γ is any good matching on $\{1,\dots,2m\} - \{i,j,k,l\}$, and in particular that there is an edge ab in Γ such that $\{1, 2m\} \subset \{a,b,i,j,k,l\}$. If $m \geq 3$ the relations of I_V are sufficient to write $X_{ik\cdot jl\cdot ab}/X_{ij\cdot kl\cdot ab}$ in terms of w_{ij} and z_{ij} in terms of the "$m=3$ variables". By taking this expression, and "adding in the remaining edges of Γ," we get the desired result for our case.

Step 2. We will show that the kernel of the map $\psi : \mathbb{Z}[w_{ij}, z_{ij}]/J_{wz} \to \mathbb{Z}[W_{ij}, Z_{ij}]/J_{WZ}$ given by $w_{ij} \mapsto W_{ij}$, $z_{ij} \mapsto Z_{ij}$ lies in $S^{-1}I_V$.

In order to do this, we need only verify that the relations (8) and (9) are consequences of the relations in $S^{-1}I_V$.

We first verify (8). By the simple binomial relation, we may write

$$W_{ij} = \frac{X_{1i\cdot j(2m)\cdot kl\cdot\Gamma}}{X_{1j\cdot i(2m)\cdot kl\cdot\Gamma}}, \qquad W_{kl} = \frac{X_{1k\cdot l(2m)\cdot ij\cdot\Gamma}}{X_{1l\cdot k(2m)\cdot ij\cdot\Gamma}}, \tag{12}$$

$$W_{il} = \frac{X_{1i\cdot l(2m)\cdot jk\cdot\Gamma}}{X_{1l\cdot i(2m)\cdot jk\cdot\Gamma}}, \qquad W_{kj} = \frac{X_{1k\cdot j(2m)\cdot il\cdot\Gamma}}{X_{1j\cdot k(2m)\cdot il\cdot\Gamma}}. \tag{13}$$

We wish to show that modulo $S^{-1}I_V$, the product of the terms in (12) equals the product of the terms in (13). The above relation is really just a relation for the six points $1, i, j, k, l, 2m$ by "removing Γ from the subscripts". If we "add Γ back in to the subscripts" this relation will follow as a consequence of the relations in $S^{-1}I_V$.

We next verify (9):

$$1 - W_{i,j} = \frac{X_{1j\cdot i(2m)\cdot\Gamma_{i,j}}}{X_{1j\cdot i(2m)\cdot\Gamma_{i,j}}} - \frac{X_{1i\cdot j(2m)\cdot\Gamma_{i,j}}}{X_{1j\cdot i(2m)\cdot\Gamma_{i,j}}} \equiv \frac{X_{1(2m)\cdot ij\cdot\Gamma_{i,j}}}{X_{1j\cdot i(2m)\cdot\Gamma_{i,j}}} = 1/Z_{i,j} \pmod{S^{-1}I_V}$$

where the equivalence uses a linear Plücker relation. □

We have reduced to the symmetric case $\mathbf{w} = 1^n$, $n = 2m$, where $n \geq 10$. Define V_n to be the scheme cut out by the Plücker and simple binomial relations. We wish to show that $V_n \equiv M_n$.

The reader will notice that we will use the simple binomial and generalized Segre cubic relations very little. In fact we just use the inductive structure of the moduli space: given a matching Δ on $n-k$ of n vertices ($4 \leq k < n$), and a point $[X_\Gamma]_\Gamma$ of V_n, then either these X_Γ with $\Delta \subset \Gamma$ are all zero, or $[X_\Gamma]_{\Delta \subset \Gamma}$ satisfies the Plücker and simple binomial relations for k, and hence is a point of V_k if $k \neq 6$. The reader should think of this rational map $[X_\Gamma] \dashrightarrow [X_\Gamma]_{\Delta \subset \Gamma}$ as a forgetful map, remembering only the moduli of the k points. In fact, even if $k = 6$ and $n \geq 8$, the point must lie in M_6, as the simple binomial relations for $n > 6$ induce the Segre cubic relation (§2.10). The central idea of our proof is, ironically, to use the case $n = 6$, where Theorem 1.1 does not apply.

We will call such Δ, where the X_Γ with $\Delta \subset \Gamma$ are not all zero and the corresponding point of M_6 is stable, a *stable* $(n-6)$*-matching*. One motivation for this definition is that given a stable configuration of n points on $\mathbb{P}^1$ there always exists a stable $(n-6)$-matching. (Hint: Construct Δ inductively as follows. We say two of the n points are in the same *clump* if they have the same image on $\mathbb{P}^1$. Choose any y in the largest clump, and any z in the second-largest clump; yz is our first edge of Δ. Then repeat this with the remaining vertices, stopping when there are six vertices left.) Caution: This is false with 6 replaced by 4 — consider the point of M_6 where $p_1 = p_2$, $p_3 = p_4$, $p_5 = p_6$.

The first sentence of the First Main Theorem 1.1 is a consequence of the following two statements and Theorem 4.4. (As remarked in §2.12, the second sentence of the First Main Theorem follows from the first.) Indeed, **(I)** and **(II)** show Theorem 1.1 set-theoretically, and scheme-theoretically away from the strictly semistable points, and Theorem 4.4 deals with (a neighborhood of) the strictly semistable points.

(I) There is a natural bijection between points of V_n with no stable $(n-6)$-matching, and strictly semistable points of M_n.

(II) If B is any scheme, there is a bijection between morphisms $B \to V_n$ missing the "no stable $(n-6)$-matching" locus (i.e. missing the strictly semistable points of M_n, by (I)) and stable families of n points $B \times \{1, \dots, n\} \to \mathbb{P}^1$. In other words, we are exhibiting an isomorphism of functors.

One direction of the bijection of (I) is immediate. The next result shows the other direction.

5.1. *Claim.* — *If $[X_\Gamma]_\Gamma$ is a point of V_n ($n \geq 10$) having no stable $(n-6)$-matching, then $[X_\Gamma]_\Gamma$ is a strictly semistable stable point of M_n.*

Several of the steps will be used in the proof of (II). We give them names so they can be referred to later.

Proof. We work by induction. We will use the fact that the result is also true for $n = 6$ (tautologically) and $n = 8$, as $V_8 = M_8$ (§2.10).

Our goal is to produce a partition of n into two subsets of size $n/2$, such that the point of M_n given by this partition via the construction of the next paragraph is our point of V_n. Throughout this proof, partitions will be assumed to mean into two equal-sized subsets.

Fix a matching Δ (*not* an $(n-6)$-matching) such that $X_\Delta \neq 0$. By the inductive hypothesis, each edge xy yields a strictly semistable point of M_{n-2}, and hence a partition of $\{1, \dots, n\} - \{x, y\}$, by considering all matchings containing xy. Thus for each $xy \in \Delta$, we get a partition of $\{1, \dots, n\} - \{x, y\}$. If wx, yz are two edges of Δ, then we get the same induced partition of $\{1, \dots, n\} - \{w, x, y, z\}$ (from the inductive hypothesis for $n - 4$), so all of these partitions arise from a single partition $\{1, \dots, n\} = S_0 \coprod S_1$.

5.2. Δ *two-overlap argument.* As this partition is determined using any two edges of Δ, we would get the same partition if we began with any Δ' sharing two edges with Δ, such that $X_{\Delta'} \neq 0$.

Defining the map to $\mathbb{P}^1$. Define $\phi : S_0 \coprod S_1 = \{1, \dots, n\} \to \mathbb{P}^1$ by $S_0 \to 0$ and $S_1 \to 1$. For each matching Γ, define X'_Γ using these points of $\mathbb{P}^1$ and (1), i.e. $X'_\Gamma = \prod_{\text{edge } e \text{ of } \Gamma} (\phi(h(e)) - \phi(t(e)))$. Rescale or normalize all the X_Γ so $X'_\Delta = X_\Delta$. We will show that $X'_\Gamma = X_\Gamma$ for all Γ, which will prove Claim 5.1. The reader should keep in mind that $X_\Gamma \neq 0$ precisely when Γ is a bipartite graph with parts S_0 and S_1.

5.3. *One-overlap argument.* For any Γ sharing an edge xy with Δ we have $X'_\Gamma = X_\Gamma$ for the following reason: $[X_\Xi]_{xy \in \Xi}$ lies in M_{n-2} by the inductive hypothesis, and this point of M_{n-2} corresponds to the map ϕ (as the partition $S_0 \coprod S_1$ was determined using this point of M_{n-2}), so $[X_\Xi]_{xy \in \Xi} = [X'_\Xi]_{xy \in \Xi}$, and the normalization $X'_\Delta = X_\Delta \neq 0$ ensures that $X'_\Xi = X_\Xi$ for all Ξ containing xy.

5.4. *$pqrs$ argument, first version.* First, assume that $X'_\Gamma \neq 0$ and that Γ shares no edge with Δ. See Figure 16. Let qr be an edge of Γ (so $\phi(q) \neq \phi(r)$), and let pq and rs be edges of Δ containing q and r respectively (so $\phi(p) \neq \phi(q)$ and $\phi(r) \neq \phi(s)$). Then $\phi(p) \neq \phi(s)$, as ϕ takes on only two values. Let $\Delta' = \Delta - pq - rs + qr + ps$, so $X'_{\Delta'} \neq 0$ as $\phi(q) \neq \phi(r)$ and $\phi(p) \neq \phi(s)$. Then $X_{\Delta'} = X'_{\Delta'}$ by the one-overlap argument 5.3, as Δ' shares an edge with Δ (indeed all but two edges), so $X_{\Delta'} \neq 0$. Hence by the Δ two-overlap argument 5.2, Δ' defines the same partition $S_0 \coprod S_1$, and hence the same map $\phi : \{1, \dots, n\} \to \mathbb{P}^1$. Finally, Γ shares an edge with Δ', so $X'_\Gamma = X_\Gamma$ by the one-overlap argument 5.3.

5.5. *Reduction to Γ with $X'_\Gamma \neq 0$.* The next idea has already appeared in the proof of Kempe's Theorem 2.3. We now reduce the general case to the case considered in §5.4. It suffices to prove the result for those graphs Γ, all of whose edges connect S_0 and S_1 (i.e. no edge is contained in S_0 or S_1; equivalently, $X'_\Gamma \neq 0$). We show this by showing that *any* X_Γ is an integral combination of such graphs, by induction on the number i of edges of Γ contained in S_0 (= the number contained in S_1). The base case $i = 0$ is tautological. For the inductive step, choose an edge $wx \in \Gamma$ contained in S_0 and an edge yz contained in

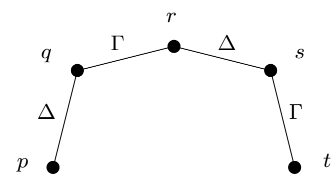

FIGURE 16. The $pqrs$ argument (vertex t is used in §5.11)

S_1. Then the Plücker relation using Γ and $wxyz$ (with appropriate signs depending on the directions of edges) is

$$\pm X_\Gamma \pm X_{\Gamma - wx - yz + wy + xz} \pm X_{\Gamma - wx - yz + wz + xy} = 0,$$

and both $\Gamma - wx - yz + wy + xz$ and $\Gamma - wx - yz + wz + xy$ have $i - 1$ edges contained in S_0, and the result follows.

We have thus completed the proof of Claim 5.1. □

5.6. Proof of (II). This proof will take us to the end of Section 5.

Given any $(n-6)$-matching Δ on some $\{1, \ldots, n\} - \{a, b, c, d, e, f\}$, we will give a bijection between

(a) morphisms $\pi : B \to V_n$ contained in the open subset where Δ is a stable $(n-6)$-matching, and
(b) stable families of points $\phi : B \times \{1, \ldots, n\} \to \mathbb{P}^1$ where $\phi|_{B \times \{a, \ldots, f\}}$ is also a stable family, and for any edge xy of Δ, $\phi|_{B \times \{x\}}$ does not intersect $\phi|_{B \times \{y\}}$.

Thus by Yoneda's lemma, we will have shown that V_n is the fine moduli space of the moduli problem.

We have already described the map (b) ⇒ (a): given a stable family of points ϕ, we get a map $\pi : B \to V_n$ given by the projective variables of §2 equation (1). We now describe the map (a) ⇒ (b), and verify that (a) ⇒ (b) ⇒ (a) is the identity. (It will then be clear that (b) ⇒ (a) ⇒ (b) is the identity: given a stable family of points parameterized by B, we get a map from B to an open subset of M_n, which is a fine moduli space, hence (b) ⇒ (a) is an injection. The result then follows from the fact that (a) ⇒ (b) ⇒ (a) is the identity.) For each morphism π of (a), the stability of the constructed family ϕ follows from the fact that the composition (a) ⇒ (b) ⇒ (a) yields π again, whose image is contained in the open subset where Δ is a stable $(n-6)$-matching.

We work by induction on n. The case $n = 8$ was checked earlier (§2.10): the simple binomial relations generate the full ideal of relations and hence cut out the quotient scheme-theoretically.

The map to $\mathbb{P}^1$. Given an element of (a), define a family of n points of $\mathbb{P}^1$ (an element of (b)) as follows. (i) $\phi : B \times \{a, \ldots, f\} \to \mathbb{P}^1$ is given by the corresponding map $B \to M_6$. (ii) If yz is an edge of Δ, we define $B \times (\{1, \ldots, n\} - \{y, z\}) \to \mathbb{P}^1$ extending (i) by considering the matchings containing yz, which by the inductive hypothesis give a point of M_{n-2}. (iii) The morphisms of (ii) agree "on the overlap," as given two edges wx and yz of Δ, we get $B \times (\{1,\dots,n\} - \{w,x,y,z\}) \to \mathbb{P}^1$ by considering the matchings containing $wx \cdot yz$, which by the inductive hypothesis give a map to M_{n-4}. Here we are using that $n \geq 10$; and if $n = 10$, we need the fact that the Segre cubic relation cutting out M_6 is induced by the quadrics cutting out M_n for $n \geq 8$ (Proposition 2.10). Thus we get a well-defined morphism $\phi : B \times \{1,\dots,n\} \to \mathbb{P}^1$.

5.7. Δ *two-overlap argument, cf.* §*5.2.* If Δ' is another matching on $\{1,\dots,n\} - \{a,\dots,f\}$ sharing at least 2 edges with Δ, with $X_{\Delta'\cdot\Xi} \neq 0$ for some matching Ξ of $\{a,\dots,f\}$, we obtain the same ϕ, as ϕ can be recovered by considering only two edges of Δ when using (ii).

Defining X'. Define X'_Γ for all matchings Γ using ϕ and the moduli morphism of equation (1). The coordinates X_Γ are projective (i.e. the set of X_Γ is defined only up to scalars); scale them so that $X_{\Delta\cdot\Xi} = X'_{\Delta\cdot\Xi}$ for all matchings Ξ of $\{a,\dots,f\}$. Note that if xy is an edge of Δ, then $\phi(x) \neq \phi(y)$, as there exists a matching Ξ of $\{a,\dots,f\}$ such that $X'_{\Delta\cdot\Xi} \neq 0$.

The following result will confirm that (a) ⇒ (b) ⇒ (a) is the identity, concluding the proof of (II).

5.8. *Claim. — We have the equality* $X_\Gamma = X'_\Gamma$ *for all* Γ*.*

Proof. This proof will occupy us until the end of §5.15.

5.9. *One-overlap argument.* As in §5.3, the result holds for those Γ sharing an edge yz with Δ: by considering only those variables $X_{\Gamma'}$ containing the edge yz (including X_Γ), we obtain a point of M_{n-2}. This point of M_{n-2} is the one given by ϕ (this was part of how ϕ was defined), so $[X_{\Gamma'}]_{yz\in\Gamma'} = [X'_{\Gamma'}]_{yz\in\Gamma'}$. By choosing a matching Ξ on $\{a,\dots,f\}$ so that $X_{\Delta\cdot\Xi} \neq 0$, we have that $X_\Gamma X'_{\Delta\cdot\Xi} = X_{\Delta\cdot\Xi} X'_\Gamma$. Using $X_{\Delta\cdot\Xi} = X'_{\Delta\cdot\Xi} \neq 0$, we have $X_\Gamma = X'_\Gamma$, as desired.

We now deal with the remaining case, where Γ and Δ share no edge.

5.10. *Reduction to* Γ *with* $X'_\Gamma \neq 0$ *(cf. §5.5).* It suffices to prove the result for those graphs such that $X'_\Gamma \neq 0$, or equivalently that for each edge xy of Γ, $\phi(x) \neq \phi(y)$. We show this by showing that any X_Γ is an integral combination of such graphs, by induction on the number of edges xy of Γ with $\phi(x) = \phi(y)$. For the purposes of this paragraph, call these *bad edges*. The base case $i = 0$ is tautological. For the inductive step, choose a bad edge $wx \in \Gamma$ (with $\phi(w) = \phi(x)$) and another edge yz such that $\phi(y), \phi(z) \neq \phi(w)$. (Such an edge exists, as by stability, less than $n/2$ elements of $\{1,\dots,n\}$ take the same value in $\mathbb{P}^1$.) Then the Plücker relation using Γ with respect to $wxyz$ is

$$\pm X_\Gamma \pm X_{\Gamma-wx-yz+wy+xz} \pm X_{\Gamma-wx-yz+wz+xy} = 0,$$

and both $\Gamma - wx - yz + wy + xz$ and $\Gamma - wx - yz + wz + xy$ have at most $i - 1$ bad edges, and the result follows.

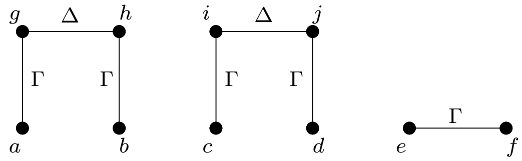

FIGURE 17. The problematic graphs for $n = 10$

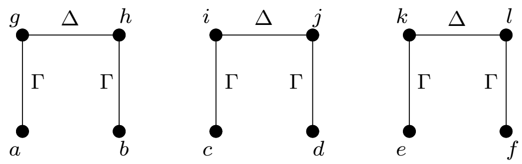

FIGURE 18. The problematic graphs for $n = 12$

Recall that we are proceeding by induction. We first deal with the case $n \geq 14$, assuming the cases $n = 10$ and $n = 12$. We will then deal with these two stray cases. This is logically backward, but the $n \geq 14$ case is cleaner, and the two other cases are similar but more ad hoc.

5.11. The case $n \geq 14$. *$pqrs$ argument, second version.* As $n \geq 14$, there is an edge qr of Γ not meeting $abcdef$. See Figure 16. By §5.10, we may assume $\phi(q) \neq \phi(r)$. Let pq and rs be the edges of Δ meeting q and r respectively (so $\phi(p) \neq \phi(q)$ and $\phi(r) \neq \phi(s)$. (i) (cf. the similar argument of §5.4). If $\phi(p) \neq \phi(s)$, then let $\Delta' = \Delta - pq - rs + qr + ps$; then Δ' defines the same family of n points as Δ by the two-overlap argument §5.7, and Γ and Δ' share an edge, so we are done by the one-overlap argument §5.9. (More precisely, this argument applies on the open subset of B where $\phi(p) \neq \phi(s)$.) (ii) If $\phi(p) = \phi(s)$, then $\phi(p) \neq \phi(r)$. (More precisely, this argument applies on the open set where $\phi(p) \neq \phi(r)$ and $\phi(q) \neq \phi(s)$.) Let st be the edge of Γ containing s. (It is possible that $t = p$.) Let $\Gamma' = \Gamma - qr - st + rs + qt$ and $\Gamma'' = \Gamma - qr - st + qs + rt$ be the other two terms in the Plücker relation for Γ for $qrst$. Then Γ' shares edge rs with Δ, so $X'_{\Gamma'} = X_{\Gamma'}$ by the one-overlap argument §5.9, and by applying (i) to Γ'' (swapping the names of r and s), $X'_{\Gamma''} = X_{\Gamma''}$, so by the Plücker relation, $X'_\Gamma = X_\Gamma$ as desired.

5.12. The cases $n = 10$ **and** $n = 12$. We are assuming that Γ and Δ share no edges. If there is an edge of Γ not meeting $\{a, \dots, f\}$ the $pqrs$-argument §5.11 applies, so assume otherwise. Divide $\{1, \dots, n\}$ into two subsets $abcdef$ and $ghij$ (respectively $ghijkl$) if $n = 10$ (respectively $n = 12$), where the edges of Δ are gh, ij, and (if $n = 12$) kl. By renaming $abcdef$, we may assume the edges of Γ are ag, bh, ci, dj, and either ef (if $n = 10$, see Figure 17) or ek and fl (if $n = 12$, see Figure 18).

5.13. Suppose that $\phi(a) \neq \phi(b)$. Note that we will only use that $ag, bh \in \Gamma$, $gh \in \Delta$, and $\phi(a) \neq \phi(b)$ — we will use this argument again below. There is a matching Ξ of $cdef$ so that if $xy \in \Xi$, then $\phi(x) \neq \phi(y)$. (This is a statement about stable configurations of 6 points on $\mathbb{P}^1$: if we have a stable set of 6 points on $\mathbb{P}^1$, then no three of them are the same point. Hence for any four of them $cdef$, we can find a matching of this sort.) Let

$\Delta' = \Xi \cdot ab \cdot \Delta$. Then by the simple binomial relations (our first invocation!) $X_{\Delta'} X_\Gamma = X_{\Delta'-ab-gh+ag+bh} X_{\Gamma+ab+gh-ag-bh}$ and $X'_{\Delta'} X'_\Gamma = X'_{\Delta'-ab-gh+ag+bh} X'_{\Gamma+ab+gh-ag-bh}$. However, by the one-overlap argument §5.9, $X_{\Delta'} = X'_{\Delta'} \neq 0$ (Δ' and Δ share edge ij), $X_{\Delta'-ab-gh+ag+bh} = X'_{\Delta'-ab-gh+ag+bh}$ ($\Delta' - ab - gh + ag + bh$ and Δ share edge ij), and $X_{\Gamma+ab+gh-ag-bh} = X'_{\Gamma+ab+gh-ag-bh}$ ($\Gamma + ab + gh - ag - bh$ and Δ share edge gh), so we are done.

We are left with the case $\phi(a) = \phi(b)$.

5.14. Suppose now that $n = 10$. As $\phi(a) = \phi(b)$, $\phi(b)$ is distinct from $\phi(e)$ and $\phi(f)$ (as $\phi(a)$, ..., $\phi(f)$ are a stable set of six points on $\mathbb{P}^1$). By the Plücker relations for Γ (using $agef$),

$$\pm X_\Gamma \pm X_{\Gamma-ag-ef+ae+gf} \pm X_{\Gamma-ag-ef+af+eg} = 0,$$

and similarly for the X' variables. By applying the argument of §5.13 with e and a swapped, we have $X'_{\Gamma-ag-ef+af+eg} = X_{\Gamma-ag-ef+af+eg}$, and by applying the argument of §5.13 with f and a swapped, we have $X'_{\Gamma-ag-ef+ae+gf} = X_{\Gamma-ag-ef+ae+gf}$, from which $X'_\Gamma = X_\Gamma$, concluding the $n = 10$ case.

5.15. Suppose finally that $n = 12$. If $\phi(c) \neq \phi(d)$, we are done (by the same argument as §5.13, with ab replaced by cd), and similarly if $\phi(e) \neq \phi(f)$. Hence the only case left is if $\phi(a) = \phi(b)$, $\phi(c) = \phi(d)$, and $\phi(e) = \phi(f)$, and (by stability of the 6 points $\phi(a), \dots, \phi(f)$) these are three distinct points of $\mathbb{P}^1$. Consider the Plücker relation for Γ with respect to $bhci$. One of the other two terms is $\Gamma - bh - ci + bi + ch$, and $X'_{\Gamma-bh-ci+bi+ch} = X_{\Gamma-bh-ci+bi+ch}$ (by the same argument as in §5.13, as $\phi(a) \neq \phi(c)$). We thus have to prove that $X_{\Gamma'} = X'_{\Gamma'}$ for the third term in the Plücker relation, where

$$\Gamma' = ag \cdot bc \cdot hi \cdot dj \cdot ek \cdot fl.$$

For this, apply the argument of §5.14 with $abghef$ replaced by $felkbc$ respectively. $\square$

Part 3. THE IDEAL OF RELATIONS IS GENERATED IN DEGREE AT MOST FOUR

In Part 3, *all graphs will be assumed to be directed upwards* (§2.1): if $a < b$, all edges ab are directed $a \to b$. Recall that *Kempe graphs* are regular upwards non-crossing graphs (§2).

We shall assume here that each point has weight 1, but we will not assume that n is even. (If n is odd then the ring will be zero in odd degrees.) We denote the coordinate ring by R. We will show that the relations of R are generated in degree at most four, by showing the same fact holds for the associated graded ring $\mathrm{gr}(R)$ of some filtration F of R. Indeed, a presentation of $\mathrm{gr}(R)$ may always be lifted to a presentation of R. This is the main method of Part III. All of this works over the integers.

We quickly review how a presentation of $\mathrm{gr}(R)$ lifts to a presentation of R. An (exhaustive) increasing $\mathbb{N}$-filtration of the $\mathbb{Z}$-algebra R is a set of $\mathbb{Z}$-submodules $F_i(R)$ of R, $i \in \mathbb{N}$, such that $1 \in F_0(R)$, $\cup_{i\in\mathbb{N}} F_i(R) = R$, and for all $i, j \in \mathbb{N}$, $F_i(R) \subset F_{i+1}(R)$ and

$F_i(R)F_j(R) \subset F_{i+j}(R)$. Similarly, if M is an R-module and R is filtered, then an increasing $\mathbb{N}$-filtration of M compatible with R is given by an increasing chain of submodules $F_m(M)$, for $m \in \mathbb{N}$, such that for all $i, j \in \mathbb{N}$, $F_i(R)F_j(M) \subset F_{i+j}(M)$.

The associated graded algebra is $\mathrm{gr}(R) = \bigoplus_{m=0}^{\infty} F_m(R)/F_{m-1}(R)$ (taking $F_{-1}(R) = 0$). Likewise the associated graded module $\mathrm{gr}(M) = \bigoplus_{m=0}^{\infty} F_m(M)/F_{m-1}(M)$. If $x \in M$, by the *leading term* of x we mean the image of x in $F_m(M)/F_{m-1}(M)$ where $m = \min\{m' \mid x \in F_{m'}(M)\}$; we sometimes call the number m the *filtration level* of x.

Now suppose that $F_m(R)$, $m \in \mathbb{N}$, is an increasing filtration of R. Suppose that $\overline{x}_1, \ldots, \overline{x}_\ell$ are homogeneous generators for $\mathrm{gr}(R)$ as an algebra. Let J be the kernel of the surjection $\mathbb{Z}[X_1, \ldots, X_\ell] \to \mathrm{gr}(R)$, where $X_i \mapsto \overline{x}_i$. Assign $\deg(X_i) := \deg(\overline{x}_i)$. This yields a filtration of the polynomial ring $\mathbb{Z}[X_1, \ldots, X_\ell]$. The kernel J is given the induced filtration as a $\mathbb{Z}[X_1, \ldots, X_\ell]$-module, and it is homogeneous with respect to the degree of the X_i's.

Now choose any lifts $x_i \in R$, such that $\overline{x}_i$ is the leading term of x_i. Then the x_i generate R. Let I be the kernel of the surjection $\mathbb{Z}[X_1, \ldots, X_\ell] \to R$, where $X_i \mapsto x_i$. Now R has the quotient filtration, that is $F_m(R)$ is the image of $F_m(\mathbb{Z}[X_1, \ldots, X_\ell])$. We give the kernel I the induced filtration $F_m(I) = I \cap F_m(\mathbb{Z}[X_1, \ldots, X_\ell])$, considered as a $\mathbb{Z}[X_1, \ldots, X_\ell]$-module. By [B, p. 169, Prop. 2] we have that $J = \mathrm{gr}(I)$. Let $\overline{r}_j$, $1 \leq j \leq \ell$ be homogeneous generators of J. Let r_j be lifts in I, that is the leading term of r_i is $\overline{r}_i$. Then the r_j generate I. Hence a presentation of $\mathrm{gr}(R)$ may be lifted to a presentation of R.

6. The toric filtration on R

6.1. The toric filtration of R. We shall introduce an $\mathbb{N}$-filtration on R such that the associated graded ring $\mathrm{gr}(R)$ is toric. By "toric ring" we mean a ring which is isomorphic to the quotient of a polynomial ring by a prime ideal which is generated by binomials. The filtration is given by taking the same type of filtration on the standard coordinate ring $\widetilde{R}$ of the Grassmannian $G(2, n)$ (we studied this ring in §4), and restricting to the T invariants $R = \widetilde{R}^T$. For all pairs (i, j) such that $1 \leq i < j \leq n$, let x_{ij} be a formal variable. The polynomial ring $\mathbb{Z}[\{x_{ij}\}_{1 \leq i < j \leq n}]$ surjects onto $\widetilde{R}$, by $x_{ij} \mapsto s_{ij}$. The kernel $\widetilde{I}$ of this map is generated by the quadric Plücker relations $x_{ac}x_{bd} - x_{ad}x_{bc} - x_{ab}x_{cd}$ for $1 \leq a < b < c < d \leq n$. The variables x_{ij} may be given weights $w(x_{ij})$ such that the associated initial ideal $\widetilde{J} = \mathrm{in}_w(\widetilde{I})$ of this weighting is generated by the binomials $x_{ac}x_{bd} - x_{ad}x_{bc}$, for $1 \leq a < b < c < d \leq n$. Such a weighting (which gives a *toric* Gröbner degeneration) first appeared in [St1]. A complete description of all such weights appeared in [SpSt]. The weighting we will use is $w(x_{ij}) := i + 2j$. This weighting also coincides with one of the toric filtrations of $\widetilde{R}$ given by Lakshmibai-Gonciulea in [GL], who studied toric degenerations of flag varieties.

6.2. Kempe graphs. Recall that Kempe graphs are regular upwards non-crossing graphs. Let $\mathcal{K}_{(N)}$ denote the set of N-regular Kempe graphs (on n vertices). Let $\mathcal{K} = \bigcup_{N=0}^{\infty} \mathcal{K}_{(N)}$ be the set of all Kempe graphs. We will use Roman letters (for example G) rather than Greek letters (such as Γ) to denote Kempe graphs.

For a regular graph Γ, recall that X_Γ denotes the associated element of R. We have already seen (Proposition 2.6) that the set $\{X_G\}_{G\in\mathcal{K}_{(N)}}$ is a basis for the N-th graded piece $R_{(N)}$ of R.

For each Kempe graph G let

$$w(G) = \sum_{ij \text{ an edge of } G} (i+2j)\,.$$

Let

$$F_m(R) = \langle X_G \rangle_{w(G)\leq m}.$$

Since the filtration levels $F_m(R) = \langle X_G\rangle_{G\in\mathcal{K}, w(G)\leq m}$ are an increasing chain of free summands of R, and each $F_m(R)$ is a free summand of $F_{m+1}(R)$, it follows that $\mathrm{gr}(R) = \bigoplus_{m=0}^{\infty} F_m(R)/F_{m-1}(R)$ is again a free $\mathbb{Z}$-module. For each Kempe graph G, let Y_G be the leading term of X_G; that is, Y_G is the image of X_G under the surjective map $F_{w(G)}(R) \to F_{w(G)}(R)/F_{w(G)-1}(R)$. Note that $\{Y_G \mid w(G) = m, G \in \mathcal{K}\}$ is a basis for $F_m(R)/F_{m-1}(R)$ as a $\mathbb{Z}$-module. Let the *standard* grading of $\mathrm{gr}(R)$ be given by

$$\mathrm{gr}(R)_{(N)} = \langle Y_G\rangle_{G\in\mathcal{K}_{(N)}}.$$

Hence $\mathrm{gr}(R)$ is bigraded; however we shall not be concerned with the grading of $\mathrm{gr}(R)$ that comes about from the filtration. We are only interested in lifting a presentation for $\mathrm{gr}(R)$. Henceforth when we say that an element of $\mathrm{gr}(R)$ has degree N, we mean the *standard* degree of the element (as defined above) is equal to N.

The following theorem implies that the $F_m(R)$ form an increasing $\mathbb{N}$-filtration of R as an *algebra*, and also that the associated graded ring $\mathrm{gr}(R) = \bigoplus_{m\geq 0} F_m(R)/F_{m-1}(R)$ is toric.

6.3. *Theorem. — Suppose $G_1 \in \mathcal{K}_{(N)}$ and $G_2 \in \mathcal{K}_{(M)}$. Let the integers c_G be the unique coefficients in the expansion of the product,*

$$X_{G_1}X_{G_2} = \sum_{G\in\mathcal{K}_{(N+M)}} c_G X_G.$$

Then there exists an X_G occurring on the right hand side with $c_G = 1$ and $w(G) = w(G_1)+w(G_2)$; furthermore, if $G' \neq G$ and $c_{G'} \neq 0$ then $w(G') < w(G)$.

Proof. Let Δ_1 be the graph consisting of the two crossing edges ik and jl, where $i < j < k < l$. Let Δ_2 be the graph with non-crossing edges il and jk, and let Δ_3 be the graph with non-crossing edges ij and kl. Recall from the proof of Proposition 2.5 that the Plücker relations $X_{\Gamma'\cdot\Delta_1} = X_{\Gamma'\cdot\Delta_2} + X_{\Gamma'\cdot\Delta_3}$ applied two edges at a time are sufficient to enable one to re-express any Γ with upwards oriented edges as a sum of Kempe graphs. However, $w(\Delta_1) = w(\Delta_2) > w(\Delta_3)$ since $i+j+2(k+l) > i+k+2(j+l)$. Hence with each application of said Plücker relations $X_{\Gamma'\cdot\Delta_1} = X_{\Gamma'\cdot\Delta_2} + X_{\Gamma'\cdot\Delta_3}$, we have $w(\Gamma'\cdot\Delta_1) = w(\Gamma'\cdot\Delta_2) > w(\Gamma'\cdot\Delta_3)$. If we always write the higher weight term to the left, then after enough Plücker relations as above have been applied so that all terms are non-crossing, the leftmost term X_G of the expansion will satisfy $w(G) = w(G_1) + w(G_2)$, and if $X_{G'}$ is any term other than the leftmost term X_G then $w(G') < w(G_1) + w(G_2)$. $\square$

It will be useful to have a notation for the unique G above as a function of G_1, G_2. Let this G be denoted $G_1 * G_2$. It is obtained from $G_1 \cdot G_2$ by replacing crossing pairs of edges ik and jl ($i < j < k < l$), with the non-crossing pair il and jk, until no crossing edges remain.

6.4. *Corollary to Theorem 6.3. — The sequence $(F_m(R))_{m=0}^{\infty}$ is an increasing filtration of R as a $\mathbb{Z}$-algebra.*

6.5. *Corollary to Theorem 6.3. — If G_1 and G_2 are Kempe graphs then*

$$Y_{G_1} Y_{G_2} = Y_{G_1 * G_2}.$$

The set of all Y_G for $G \in \mathcal{K}$ forms a graded semigroup. The ring $\mathrm{gr}(R)$ is the $\mathbb{Z}$-algebra generated by the semigroup $\{Y_G \mid G \in \mathcal{K}\}$.

We will now identify the semigroup $\{Y_G \mid G \in \mathcal{K}\}$ explicitly as the set of lattice points in a rational cone.

Let $D \subset \mathbb{R}^{n-1}$ be given by $(d_1, d_2, d_3, \ldots, d_{n-1}) \in D$ if and only if $d_1 = d_{n-1} \geq 0$ and for each i, $1 \leq i \leq n-2$, we have

$$d_i \leq d_{i+1} + d_1, \qquad d_{i+1} \leq d_i + d_1, \qquad d_1 \leq d_i + d_{i+1}.$$

It is easy to see that D is a rational cone of dimension $n-2$. Let $D_{(N)}$ be intersection of the affine hyperplane $d_1 = N$ with D.

6.6. *Remark.* Note that the inequalities defining $D_{(N)}$ hold for the diagonal lengths $d_i = |v_i - v_0|$ of an n-gon with vertices $v_0 = v_n, v_1, \ldots, v_{n-1}$ with all side lengths equal to N. Indeed the triple of inequalities holds for a triangle having side lengths d_1, d_i, d_{i+1}. That *each* d_i is nonnegative follows from $d_1 \geq 0$ and the triangle inequalities. The polytope D is isomorphic to the polytope $GT(\frac{n}{2}\varpi_2, (1, 1, \ldots, 1))$ of Gel′fand-Tsetlin patterns (Gel′fand Tsetlin pattern polytopes are studied for example in [DLMc]).

Let Λ be the lattice in $\mathbb{R}^{n-1}$ given by the conditions $(d_1, d_2, \ldots, d_{n-1}) \in \Lambda$ if and only if $d_1 = d_{n-1} \in \mathbb{Z}$, and

$$d_i \equiv i d_1 \pmod 2,$$

for each i, $1 \leq i \leq n-1$. This is equivalent to the condition that each triple (d_i, d_1, d_{i+1}) sums to an even integer.

Let $S = D \cap \Lambda$ be the semigroup of lattice points in D. We have that S is also graded, where $S_{(N)}$ consists of those elements $(d_1, d_2, \ldots, d_{n-1})$ in S such that $d_1 = N$. Hence $S_{(N)} = D_{(N)} \cap \Lambda$. Let $\mathbb{Z}[S]$ be the graded semigroup algebra.

For each Kempe graph $G \in \mathcal{K}$, let

$$\phi(G) = (d_1, d_2, \ldots, d_{n-1}) \in \mathbb{Z}^{n-1},$$

where d_i is the number of edges kl in G such that $k \leq i$ and $l \geq i+1$.

For example, for $n = 5$, there are six 2-regular Kempe graphs, and Figure 6 illustrates their images under ϕ.

6.7. *Lemma. — For each $N \geq 0$ the map ϕ is a bijection between $\mathcal{K}_{(N)}$ and $S_{(N)}$.*

Proof. First we will show that the image of ϕ is contained within $S_{(N)}$. The valence of a vertex i in G is equal to N by assumption. On the other hand, d_1 is the number of edges ij in G where $i = 1$, and d_{n-1} is the number of edges ij in G where $j = n$. Hence d_1 is the valence of vertex 1 and d_{n-1} is the valence of vertex n. But G is N-regular so $d_1 = d_{n-1} = N$. For each i, $1 \leq i \leq n-1$, let A_i be the multi-set of edges kl in G such that $k \leq i$ and $l \geq i+1$. Hence $|A_i| = d_i$ for each $i \in \{1, \ldots, n-1\}$. Let W_{i+1} be the multi-set of edges kl of G such that $k = i+1$ or $l = i+1$. Hence $|W_{i+1}|$ is the valence of the vertex $i+1$, which is equal to $N = d_1$. It is clear that any edge $kl \in A_i \cup A_{i+1} \cup W_{i+1}$ belongs to exactly two of these three sets. From this the triangle inequalities for the triple d_i, d_{i+1}, d_1 follow easily, and it is also easy to see that $d_i + d_{i+1} + d_1$ must be an even integer, since each edge is counted twice in the sum.

Now we need to show that for each $\mathbf{d} \in S_{(N)}$ there is exactly one Kempe graph $G \in \mathcal{K}_{(N)}$ such that $\phi(G) = \mathbf{d}$. Suppose that we have two planar multigraphs H and H', with respective vertex sets $\{v_1, \ldots, v_s\}$ and $\{v'_1, \ldots, v'_t\}$, such that the valences of v_s and v'_1 are equal. Further we assume that H and H' are drawn as planar graphs in such a way that the vertices $v_1, \ldots, v_s$ are drawn clockwise in a circle where v_{i+1} comes just after v_i. Likewise we assume the same of H'. We define a graph $H \diamond H'$ with vertices $v_1, \ldots, v_{s-1}, v'_2, \ldots, v'_t$ arranged in clockwise order around a circle, with no two edges crossing, by removing vertices v_s from H and v'_1 from H', and joining together their respective edges, without introducing any crossing edges. We now show how to glue the edges of v_s and v'_1 together without introducing crossings. Suppose that $d = \deg(v_s) = \deg(v'_1)$ is the common valence of v_s and v'_1. Label the edges adjacent to v_s as $\epsilon_1, \ldots, \epsilon_d$ where if $1 \leq i < j \leq d$, $\epsilon_i = v_a v_s$, and $\epsilon_j = v_b v_s$, then $a \geq b$. Label the edges adjacent to v'_1 as $\epsilon'_1, \ldots, \epsilon'_d$, where if $1 \leq i < j \leq d$, $\epsilon'_i = v'_1 v'_{a'}$, and $\epsilon'_j = v'_1 v'_{b'}$, then $a' \leq b'$. Now join ϵ_i to ϵ'_i for $1 \leq i \leq d$, to obtain d edges of the graph $H \diamond H'$. Suppose that $v_a v'_{a'}$ and $v_b v'_{b'}$ are (joined) edges of $H \diamond H'$. Suppose $v_a v'_{a'}$ is the join of edges $\epsilon_i = v_a v_s$ and $\epsilon'_i = v'_1 v'_{a'}$, and $v_b v'_{b'}$ is the join of edges $\epsilon_j = v_b v_s$ and $\epsilon'_j = v'_1 v'_{b'}$. If $i < j$ then $b \leq a$ and $a' \leq b'$. If $i > j$ then $a \leq b$ and $b' \leq a'$. In either case $v_a v'_{a'}$ and $v_b v'_{b'}$ do not cross each other. We also claim that there is only one way to join the edges of v_s and v'_1 so that no crossings are introduced. Suppose $\{\delta'_1, \ldots, \delta'_d\} = \{\epsilon'_1, \ldots, \epsilon'_d\}$ and ϵ_i is joined to δ'_i, $1 \leq i \leq d$, so that no two joined edges cross each other. We may assume that if $i < j$, and $\epsilon_i = v_a v_s$, $\epsilon_j = v_a v_s$, $\delta'_i = v'_1 v'_{a'}$, and $\delta'_j = v'_1 v'_{b'}$, then $a' \leq b'$. Let i be the minimal index for which $\delta'_i \neq \epsilon'_i$. Suppose that $\delta'_i = v'_1 v'_{a'}$ and $\epsilon'_i = v'_1 v'_{b'}$. Then $b' < a'$. There exists $j > i$ such that $\delta'_j = v'_1 v'_{b'}$. Suppose that $\epsilon_j = v_b v_s$. Since $i < j$ we have that $a \geq b$. Since $v_a v'_{a'}$ does not cross $v_b v'_{b'}$, $a \geq b$, and $b' < a'$, we have that $a = b$. But now we have a contradiction with our choice of labelling of $\delta'_1, \ldots, \delta'_d$, because $i < j$, $\epsilon_i = v_a v_s$, $\epsilon_j = v_a v_s$, $\delta'_i = v'_1 v'_{a'}$, and $\delta'_j = v'_1 v'_{b'}$, but $b' < a'$.

Now suppose we are given three nonnegative numbers a, b, c which satisfy the triangle inequalities and $a + b + c$ is even. Let $G(a, b, c)$ be the multigraph with three vertices v_1, v_2, v_3, where the number of edges ϵ_{ij} between v_i and v_j is given by $\epsilon_{12} = (a + b - c)/2$, $\epsilon_{13} = (a + c - b)/2$, $\epsilon_{23} = (b + c - a)/2$. Equivalently the valence of vertex v_1 is a, the valence of vertex v_2 is b, and the valence of vertex v_3 is c. We also require $G(a,b,c)$ to be drawn in the plane with no crossing edges, where the vertices v_1, v_2, v_3 are oriented clockwise around a circle. We now describe the desired Kempe graph G as a $\diamond$ product of $n-2$ tripod graphs $G(a,b,c)$. Let $a_1 = c_{n-2} = N$. Let $c_i = a_{i+1} = d_{i+1}$, for $1 \le i \le n-3$. Let $b_i = N$ for all i. Then each triple (a_i, b_i, c_i) satisfies the triangle inequalities, and $a_i + b_i + c_i$ is even. Let

$$G := G(a_1, b_1, c_1) \diamond G(a_2, b_2, c_2) \diamond \cdots \diamond G(a_{n-2}, b_{n-2}, c_{n-2}).$$

Then $\phi(G) = \mathbf{d}$, because the number of edges xy in G where $x \le i < y$ is equal to $a_i = c_{i-1}$, for $2 \le i \le n-2$, and G is regular with all vertices having valence N. The valence of vertex 1 is equal to $a_1 = N$, the valence of vertex n is equal to $a_{n-2} = N$, and the valence of vertex i is equal to $b_{i-1} = N$, for $2 \le i \le n-1$. This shows existence, but we still need to show uniqueness.

We claim that any Kempe graph G decomposes as a $\diamond$-product of tripod graphs. Choose any index $i \ge 2$. Then $G = H \diamond H'$ where H and H' are defined as follows. The graph H will have $i+1$ vertices $1, 2, \ldots, i, *$ where $*$ is an auxiliary vertex, and H' will have vertices $*', i+1, \ldots, n$, where $*'$ is an auxiliary vertex. We split each edge ab of G where $a \le i$ and $b \ge i+1$ into edge $a*$ of H and $*'b$ of H'. The remaining edges ab where $a < b \le i$ resp. $i+1 \le a < b$ are the remaining edges of H resp. H'. Then clearly $H \diamond H' = G$ since we have split each edge ab with $a \le i$ and $i+1 \le b$ into an edge of H and an edge of H', and these same two edges will be joined back together again in $H \diamond H'$, since there is only one way to join edges so that the resulting graph is noncrossing. Now continue to split H, H' into $\diamond$ products until all resulting components are tripods. Now suppose that

$$H = G(a'_1, b'_1, c'_1) \diamond G(a'_2, b'_2, c'_2) \diamond \cdots \diamond G(a'_{n-2}, b'_{n-2}, c'_{n-2}),$$

and each vertex of H has valence N, and $\phi(H) = \mathbf{d}$. Then we have that $b'_i = N = b_i$ for each i, and $a'_1 = c'_{n-2} = N$ since the valence of each vertex must be equal to N. Also we have that $c'_i = a'_{i+1} = d_{i+1}$ for $1 \le i \le n-3$, since the number of edges in H of the form ab with $a \le i+1 < b$ is equal to $c'_i = a'_{i+1} = d_{i+1}$. Hence each tripod graph $G(a'_i, b'_i, c'_i) = G(a_i, b_i, c_i)$, and so $H = G$. We have thus shown uniqueness of G. □

6.8. *Lemma.* — *If G_1 and G_2 are Kempe graphs then*

$$\phi(G_1 * G_2) = \phi(G_1) + \phi(G_2).$$

Hence ϕ is an isomorphism of semigroups, and induces an isomorphism (also denoted ϕ) on the semigroup algebras,

$$\phi : \mathrm{gr}(R) \cong \mathbb{Z}[S].$$

Proof. First extend the domain of ϕ to general multi-graphs Γ (edges may cross) by the same rule:

$$\phi(\Gamma) = (d_1, d_2, d_3, \ldots, d_{n-1}),$$

where d_i is the number of edges kl of Γ such that $k \le i$ and $l \ge i+1$. With this extension of the definition it is clear that $\phi(G_1 \cdot G_2) = \phi(G_1) + \phi(G_2)$.

Now we show that $\phi(G_1 \cdot G_2) = \phi(G_1 * G_2)$. Let Δ_1 be the graph with the two crossing edges ik and jl, where $i < j < k < l$. Let Δ_2 be the graph with non-crossing edges il

and jk, and let Δ_3 be the graph with non-crossing edges ij and kl. Recall from the proof of Proposition 2.5 that the Plücker relations $X_{\Gamma'\cdot\Delta_1} = X_{\Gamma'\cdot\Delta_2} + X_{\Gamma'\cdot\Delta_3}$ applied two edges at a time are sufficient to enable one to re-express any X_Γ with upwards oriented edges as a sum of Kempe graphs. We have that $\phi(\Delta_1) = \phi(\Delta_2)$, hence with each application of a Plücker relation $X_{\Gamma'\cdot\Delta_1} = X_{\Gamma'\cdot\Delta_2} + X_{\Gamma'\cdot\Delta_3}$, we have $\phi(\Gamma'\cdot\Delta_1) = \phi(\Gamma'\cdot\Delta_2)$. Finally once enough Plücker relations have been applied, starting from the initial $X_{G_1}X_{G_2}$, the final leading term $X_{G_1 * G_2}$ of the expansion will satisfy $\phi(G_1 * G_2) = \phi(G_1 \cdot G_2) = \phi(G_1) + \phi(G_2)$. □

6.9. *Corollary. — The rings* $\mathrm{gr}(R)$ *and* $\mathbb{Z}[S]$ *are isomorphic as graded rings. That is,* $\phi : \mathrm{gr}(R)_{(N)} \to \mathbb{Z}\langle S_{(N)}\rangle$ *for each* $N \geq 0$*, and* ϕ *is a ring isomorphism.*

7. The projective coordinate ring $\mathrm{gr}(R)$

We shall show in this section that the degree one and degree two elements of the semigroup algebra $\mathbb{Z}[S]$ generate $\mathbb{Z}[S]$. Furthermore we will show that the ideal of relations of $\mathbb{Z}[S]$ is generated by relations of degrees two, three, and four.

We shall also be interested in the semigroup S^{even} which is defined by $S^{even} = \cup_{m=0}^{\infty} S_{(2m)}$, and we give S^{even} the grading $S^{even}_{(m)} := S_{(2m)}$. Then $\mathbb{Z}[S^{even}]$ is a Veronese subring of $\mathbb{Z}[S]$. Indeed if we set $R^{even} = \bigoplus_{m=0}^{\infty} R_{(2m)}$, and restrict the filtration given in the previous section to R^{even}, then clearly $\mathrm{gr}(R^{even}) \cong \mathbb{Z}[S^{even}]$. It will follow from the proofs given for S that $\mathbb{Z}[S^{even}]$ has a presentation by degree one generators and quadratic relations.

7.1. Generators for $\mathbb{Z}[S]$.

If A, B are subsets of a vector space, let $A + B := \{a + b \mid a \in A,\, b \in B\}$ (the Minkowski sum of A and B).

7.2. *Lemma. — For each positive integer* m*,* $S_{(2m+1)} = S_{(1)} + S_{(2m)}$*.*

Proof. Given $\mathbf{d} \in S_{(2m+1)}$ we shall construct $\mathbf{d}' \in S_{(1)}$ in the proximity of $\mathbf{d}/(2m+1)$ such that $\mathbf{d} - \mathbf{d}' = \mathbf{d}'' \in S_{(2m)}$. Recall that our lattice is Λ. An element $\mathbf{d}' = (d'_1, \ldots, d'_{n-1})$ in $D_{(1)}$ is a lattice point if and only if each $d'_i \equiv i \pmod 2$. On the other hand $\mathbf{d}'' = (d''_1, \ldots, d''_{n-1}) \in D_{(2m)}$ is a lattice point if and only if each d''_i is an even integer. Define

$$d'_i = k, \text{ such that } k \in \mathbb{Z},\ k \equiv i \pmod 2, \text{ and } \left| k - \frac{d_i}{2m+1} \right| < 1.$$

The point $\mathbf{d}'$ is the closest lattice point to the rational point $\mathbf{d}/(2m+1)$. To check that $\mathbf{d}'$ is well-defined we need to show that k exists and is unique. Uniqueness follows immediately since there can be only one integer of a given parity in an open interval of length 2. For existence we must check that $d_i/(2m+1)$ does not have opposite parity to i, since in this case there is no integer of the correct parity less than one unit from $d_i/(2m+1)$. But $d_i \equiv i \pmod 2$ since $\mathbf{d} \in S_{(2m+1)}$ is a lattice point and $2m+1$ is odd. Hence if $d_i/(2m+1)$

is an integer then $d_i/(2m+1) \equiv i \pmod 2$ as well since $2m+1$ is odd. Therefore the parity condition for $\mathbf{d}'$ is satisfied.

Let $\mathbf{d}'' = \mathbf{d} - \mathbf{d}'$. Note that each d_i'' is even since d_i and d_i' have the same parity. We have that in fact d_i'' is the nearest even integer to $2md_i/(2m+1)$. Thus we have checked that $\mathbf{d}', \mathbf{d}'' \in \Lambda$. Now we only need to show that $\mathbf{d}' \in D_{(1)}$ and $\mathbf{d}'' \in D_{(2m)}$.

Since $d_1 = 2m+1$ and $d_{n-1} = 2m+1$ we have that $d_1' = 1$, $d_1'' = 2m$, $d_{n-1}' = 1$, and $d_{n-1}'' = 2m$. It remains to show the triangle inequalities,

(1) $d_{i-1}' \leq d_i' + 1$, $d_{i-1}'' \leq d_i'' + 2m$,
(2) $d_i' \leq d_{i-1}' + 1$, $d_i'' \leq d_{i-1}'' + 2m$,
(3) $1 \leq d_{i-1}' + d_i'$, $2m \leq d_{i-1}'' + d_i''$.

Let $\ell = 2m+1$. We have that $|d_{i-1} - d_i| \leq \ell$, hence $|d_{i-1}/\ell - d_i/\ell| \leq 1$. Recall that d_{i-1}' is the nearest integer to d_{i-1}/ℓ congruent to $i-1 \pmod 2$ and d_i' is the nearest integer to d_i/ℓ with parity $i \pmod 2$. The distance between d_{i-1}/ℓ and d_i/ℓ is at most 1 and also $d_{i-1}' - d_i' \equiv i \pmod 2$. We also have that $|d_{i-1}' - d_{i-1}/\ell| < 1$ and $|d_i' - d_i/\ell| < 1$. Therefore $|d_{i-1}' - d_i'| < 3$ and consequently $|d_{i-1}' - d_i'| \leq 2$. But $d_{i-1}' - d_i'$ is odd so $|d_{i-1}' - d_i'| = 1$. Also $|d_{i-1}'' - \frac{\ell-1}{\ell} d_{i-1}| < 1$ and $|d_i'' - \frac{\ell-1}{\ell} d_i| < 1$. Therefore $|d_{i-1}'' - d_i''| < (\ell-1) + 2 = \ell + 1$, so $|d_{i-1}'' - d_i''| \leq \ell$. But $d_{i-1}'' - d_i''$ is even and ℓ is odd, so $|d_{i-1}'' - d_i''| \leq (\ell - 1) = 2m$. Therefore both (1) and (2) hold.

We have $d_{i-1}' + d_i' > d_{i-1}/\ell + d_i/\ell - 2 \geq 1 - 2 = -1$. Thus $d_{i-1}' + d_i'$ is odd, so $d_{i-1}' + d_i' \geq 1$. We have $d_{i-1}'' + d_i'' > \frac{\ell-1}{\ell}(d_{i-1} + d_i) - 2 \geq (\ell - 1) - 2 = \ell - 3$. But $d_{i-1}'' + d_i''$ is even and so $d_{i-1}'' + d_i'' \geq \ell - 1 = 2m$. Therefore (3) holds. ☐

7.3. *Lemma. — For each positive integer m,*

$$S_{(2m)} = \underbrace{S_{(2)} + \cdots + S_{(2)}}_{m}.$$

Proof. We show the lemma by induction on $m \geq 1$. The case $m = 1$ is a tautology. Suppose that $m \geq 2$. We show that $S_{(2m)} = S_{(2)} + S_{(2m-2)}$. Suppose $\mathbf{d} = (d_1, \ldots, d_{n-1}) \in S_{(2m)}$. We construct $\mathbf{d}' \in S_{(2)}$ and $\mathbf{d}'' \in S_{(2m-2)}$ where $\mathbf{d}' + \mathbf{d}'' = \mathbf{d}$, by placing $\mathbf{d}'$ in the proximity of $\mathbf{d}/m$. Recall that the integrality condition is that the components of $\mathbf{d}$, $\mathbf{d}'$, and $\mathbf{d}''$ are even integers.

Let $e^- : \mathbb{R} \to 2\mathbb{Z}$ be the function which assigns the nearest even integer, where odd integers $2t+1$ are mapped to $2t$. To be concise,

$$e^-(x) = \min\{k \in 2\mathbb{Z} : k \geq x - 1\} = 2\lceil (x-1)/2 \rceil.$$

Similarly let $e^+ : \mathbb{R} \to 2\mathbb{Z}$ assign the nearest even integer where odd integers $2t+1$ are mapped to $2t+2$,

$$e^+(x) = \max\{k \in 2\mathbb{Z} : k \leq x + 1\} = 2\lfloor (x+1)/2 \rfloor.$$

We will often use the following properties of e^- and e^+:

- if x is not an odd integer, then $e^-(x) = e^+(x)$.
- each of e^- and e^+ is weakly increasing.
- if $k \in 2\mathbb{Z}$, then $e^\pm(x+k) = e^\pm(x) + k$.
- $e^+(-x) = -e^-(x)$.
- if $x + y \in 2\mathbb{Z}$, then $e^+(x) + e^-(y) = x + y$.
- if $x + y \geq k \in 2\mathbb{Z}$, then $e^+(x) + e^-(y) \geq k$.

Let $\mathbf{d} = (d_1, \ldots, d_{n-1}) \in S_{(2m)}$. Let

$$\mathcal{J}^0_{\mathbf{d}} = \{i \in \{2, \ldots, n-1\} \mid d_{i-1}/m \text{ and } d_i/m \text{ are odd integers}, d_{i-1} + d_i = 2m\},$$
$$\mathcal{J}^1_{\mathbf{d}} = \{i \in \{2, \ldots, n-1\} \mid d_{i-1} \leq 2m,\ d_i \leq 2m\}.$$

Clearly $\mathcal{J}^0_{\mathbf{d}} \subset \mathcal{J}^1_{\mathbf{d}}$. Let $\mathcal{J}_{\mathbf{d}}$ be such that $\mathcal{J}^0_{\mathbf{d}} \subset \mathcal{J}_{\mathbf{d}} \subset \mathcal{J}^1_{\mathbf{d}}$. Let $\{i_1, \ldots, i_s\} = \mathcal{J}_{\mathbf{d}}$ such that $i_t < i_{t+1}$ for all t, and set $i_0 = 1$ and $i_{s+1} = n$.

Let $\mathbf{d}' = (d'_1, \ldots, d'_{n-1}) \in (2\mathbb{Z})^{n-1}$ be

$$d'_i = \begin{cases} e^-(d_i/m) & \text{for } i_{2t} \leq i < i_{2t+1}, 2t \leq s, \\ e^+(d_i/m) & \text{for } i_{2t+1} \leq i < i_{2t+2}, 2t+1 \leq s. \end{cases}$$

Let $\mathbf{d}'' = (d''_1, \ldots, d''_{n-1}) \in (2\mathbb{Z})^{n-1}$ be

$$d''_i = \begin{cases} e^+((m-1)d_i/m) & \text{for } i_{2t} \leq i < i_{2t+1}, 2t \leq s, \\ e^-((m-1)d_i/m) & \text{for } i_{2t+1} \leq i < i_{2t+2}, 2t+1 \leq s. \end{cases}$$

We will show that $\mathbf{d}' + \mathbf{d}'' = \mathbf{d}$, $\mathbf{d}' \in S_{(2)}$, and $\mathbf{d}'' \in S_{(2m-2)}$. Note that $e^\pm(x) + e^\mp(y) = k$ whenever $x + y = k$ and $k \in 2\mathbb{Z}$. We have that $d_i/m + (m-1)d_i/m = d_i \in 2\mathbb{Z}$ for all i, so

$$d'_i + d''_i = e^\pm(d_i/m) + e^\mp((m-1)d_i/m) = d_i.$$

Thus $\mathbf{d}' + \mathbf{d}'' = \mathbf{d}$.

We show that $\mathbf{d}' \in D_{(2)}$. The proof that $\mathbf{d}'' \in D_{(2m-2)}$ is similar. Since $d_1 = 2m$ and $d_{n-1} = 2m$, we have $d'_1 = 2$ and $d'_{n-1} = 2$. Now suppose $2 \leq i \leq n-1$. We must show the three triangle inequalities that define $D_{(2)}$:

(1) $d'_{i-1} \leq d'_i + 2$,
(2) $d'_i \leq d'_{i-1} + 2$,
(3) $2 \leq d'_{i-1} + d'_i$.

Suppose that $i \notin \mathcal{J}_{\mathbf{d}}$. We have $d'_{i-1} = e^\pm(d_{i-1}/m)$ and $d'_{i-1} = e^\pm(d_i/m)$ (the same function is applied to each). The functions e^- and e^+ are weakly increasing, and since $d_i/m \leq d_{i-1}/m + 2$ we have that

$$d'_i = e^\pm(d_i/m) \leq e^\pm(d_{i-1}/m + 2) = e^\pm(d_{i-1}/m) + 2 = d'_{i-1} + 2$$

so (2) holds. Similarly inequality (1) holds. Since $i \notin \mathcal{J}^0_{\mathbf{d}}$ we know that either $d_{i-1} + d_i > 2m$ or one of d_{i-1}/m or d_i/m is not an odd integer. Suppose $d_{i-1} + d_i > 2m$. We have that $d'_{i-1} + d'_i \geq d_{i-1}/m + d_i/m - 2 > 0$. But $d'_{i-1} + d'_i$ is even so $d'_{i-1} + d'_i \geq 2$ and (3) holds. Suppose that one of d_{i-1}/m or d_i/m is not an odd integer and $d_{i-1} + d_i = 2m$. Without loss of generality suppose that d_{i-1}/m is not odd. Then $e^+(d_{i-1}/m) = e^-(d_{i-1}/m)$. Since the sum $d_{i-1}/m + d_i/m = 2$ is even, we have that $e^+(d_{i-1}/m) + e^-(d_i/m) = 2$. Now

$$d'_{i-1} + d'_i \geq e^-(d_{i-1}/m) + e^-(d_i/m) = e^+(d_{i-1}/m) + e^-(d_i/m) = 2,$$

and again (3) holds.

Suppose that $i \in \mathcal{J}_{\mathbf{d}}$. Hence $i \in \mathcal{J}^1_{\mathbf{d}}$ and so $d_{i-1}/m \leq 2$ and $d_i/m \leq 2$. Therefore each of $d'_{i-1} \leq 2$ and $d'_i \leq 2$. We have $d'_{i-1} = e^{\pm}(d_{i-1}/m)$ and $d'_i = e^{\mp}(d_i/m)$. Whenever two numbers x, y satisfy $x + y \geq k \in 2\mathbb{Z}$, then $e^{\pm}(x) + e^{\mp}(y) \geq k$, thus (3) holds since $d_{i-1}/m + d_i/m \geq 2$. Suppose that $d'_{i-1} = e^{+}(d_{i-1}/m)$ and so $d'_i = e^{-}(d_i/m)$. We show that (1) holds. We have that

$$2 \geq d_{i-1}/m - d_i/m \geq d'_{i-1} - d'_i - 2,$$

so if (1) fails then $d'_{i-1} - d'_i = 4$. But $d'_{i-1} + d'_i \geq 2$ since we have shown (3) already, and hence we get that $d'_{i-1} \geq 3$, a contradiction with $d'_{i-1} \leq 2$. The other cases are similar. □

7.4. *Theorem. — The toric ring* $\mathbb{Z}[S]$ *is generated by elements of degrees one and two; furthermore,* $\mathbb{Z}[S^{even}]$ *is generated by elements of degree one.*

Proof. The first statement is a direct consequence of Lemmas 7.2 and 7.3. The second statement follows from Lemma 7.3. □

7.5. The word problem for S and the relations for $\mathbb{Z}[S]$.

We will solve the presentation problem for $\mathbb{Z}[S]$ by solving the seemingly more difficult *word problem* for the graded semigroup $S = \cup_{N \geq 0} S_{(N)}$. Our technique is to define a normal form for words in S expressed in terms of degree one and degree two elements, then show that any word can be brought into normal form by a sequence of relations of degrees two, three, and four.

Let $\xi_{2m+1} : S_{(2m+1)} \to S_{(1)}$ be given by $\xi_{2m+1}(\mathbf{d}) = \mathbf{d}'$ where $\mathbf{d}'$ is as in the proof of Lemma 7.2.

If A is an integer matrix let the jth column of A be denoted $c_j(A)$. If each column of A belongs to either $S_{(1)}$ or $S_{(2)}$ then we say that A is an S-matrix. (S-matrices represent monomials in $\mathbb{Z}[S]$ which are products of degree one and degree two generators.) We define $\deg(A)$ to be the sum of the degrees of the columns of A whenever A is an S-matrix. Note that $\deg(A)$ is a nonnegative integer, and recall that A has $n-1$ rows.

7.6. *Definition: normal form.* Suppose that A is an S-matrix. Let

$$\mathbf{a} = (a_1, \ldots, a_{n-1}) = \sum_j c_j(A) \in S_{(\deg(A))}.$$

Suppose that $\deg(A) = 2m$ is even. Let $\mathcal{J}_{\mathbf{a}} = \{i_1, \ldots, i_k\}$ be the set of all i, $2 \leq i \leq n-1$, such that $a_{i-1} \leq 2m$ and $a_i \leq 2m$, where $i_t < i_{t+1}$ for all t, $1 \leq t < k$. Let $i_0 = 1$ and let $i_{k+1} = n$. (Note that $\mathcal{J}_{\mathbf{a}} = \mathcal{J}^1_{\mathbf{a}}$ as in the proof of Lemma 7.3.) Suppose A has the following properties:

(N0) Each column of A has degree two.
(N1) For each i the row entries $a_{i,j}$ satisfy $|a_{i,j} - a_i/m| < 2$.

(N2) For $i_{2t} \le i < i_{2t+1}$, row i is weakly increasing. For $i_{2t+1} \le i < i_{2t+2}$, row i is weakly decreasing.

Then we say that A is in normal form.

Now suppose that $\deg(A) = 2m+1$ is odd. Then we say that A is in normal form if the first column of A is equal to $\xi_{2m+1}(\mathbf{a})$ and if the matrix A' obtained from A by removing the first column is in normal form.

7.7. *Lemma (uniqueness). — For any* $\mathbf{a} = (a_1, \dots, a_{n-1}) \in S_{(\ell)}$ *there is at most one matrix* A *in normal form such that the columns of* A *sum to* $\mathbf{a}$*.*

Proof. Suppose $\ell = 2m = \deg(A)$ is even and A is in normal form. Then each column of A is degree two so all the matrix entries are even integers and there are m columns. For each i let k_i be an even integer such that $k_i \le a_i/m \le k_i + 2$. By condition (N1) we know that each $a_{i,j}$ is either k_i or $k_i + 2$. Let t_i be the number of $a_{i,j}$ equal to k_i. Then, $t_i k_i + (m - t_i)(k_i + 2) = a_i$, so $2t_i = m(k_i + 2) - a_i$, and thus t_i is determined by the value of a_i. Finally the monotonicity condition (N2) determines each $a_{i,j}$.

Suppose $\ell = 2m+1 = \deg(A)$ is odd and A is in normal form. The first column of A must be equal to $\xi_{2m+1}(\mathbf{a})$ so it is determined. Now the matrix A' which is A with the first column removed is degree $2m$ and is in normal form, so its entries are determined by the argument given above for matrices of even degree. □

We say that two S-matrices A and B are equivalent if the sum $\mathbf{a}$ of the columns of A is equal to the sum $\mathbf{b}$ of the columns of B. (Note that each equivalence class contains at most one representative in normal form by the above lemma — later we will see that a normal form representative always exists.)

7.8. *Example.* Here is an example of two S-matrices A and B where B is the normal form representative of A. Here $n = 8$.

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 2 & 0 \\ 1 & 0 & 2 \\ 2 & 2 & 4 \\ 3 & 0 & 4 \\ 2 & 2 & 2 \\ 1 & 2 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 2 & 2 \\ 1 & 0 & 2 \\ 2 & 2 & 4 \\ 1 & 2 & 4 \\ 2 & 2 & 2 \\ 1 & 2 & 2 \end{pmatrix}$$

Let A be an S-matrix and let $\mathbf{d}_1$ and $\mathbf{d}_2$ be two different columns of A. We define operations of types (F2), (F3), and (F4) as follows. Here (Fj) corresponds to a degree j relation, for $j = 2, 3, 4$.

(F2) If $\deg(\mathbf{d}_1) = \deg(\mathbf{d}_2) = 1$ then remove columns $\mathbf{d}_1$ and $\mathbf{d}_2$ and place $\mathbf{d}_1 + \mathbf{d}_2$ as the last column.

(F3) If $\deg(\mathbf{d}_1) = 1$ and $\deg(\mathbf{d}_2) = 2$, let $\mathbf{d} = \mathbf{d}_1 + \mathbf{d}_2$. Replace $\mathbf{d}_1$ and $\mathbf{d}_2$ with $\xi_3(\mathbf{d})$ and $\mathbf{d} - \xi_3(\mathbf{d})$, placing $\xi_3(\mathbf{d})$ left of $\mathbf{d} - \xi_3(\mathbf{d})$.

(F4) Suppose $\deg(\mathbf{d}_1) = \deg(\mathbf{d}_2) = \deg(\mathbf{d}'_1) = \deg(\mathbf{d}'_2) = 2$ and $\mathbf{d}'_1 + \mathbf{d}'_2 = \mathbf{d}_1 + \mathbf{d}_2$. Replace $\mathbf{d}_1$ and $\mathbf{d}_2$ with $\mathbf{d}'_1$ and $\mathbf{d}'_2$. (Here we could have been more restrictive; we could have restricted to operations of type f^-, f^+ and g^-, g^+ on pairs of degree two columns as in the proof below of 7.9.)

7.9. *Lemma. — Suppose that A is an S-matrix. Then there is a finite sequence $(A_0, A_1, \ldots, A_p)$ of equivalent S-matrices where $A_0 = A$, the final matrix A_p is in normal form, and for each i the matrix A_{i+1} is obtained from A_i by a single operation of type (F2), (F3), or (F4). (In the special case that A is even degree and all columns are degree two then all these operations are of type (F4).)*

Proof. First note that (F2) operations can be applied to any pair of degree one columns until either every column is degree two (when $\deg(A)$ is even) or there is only one column of degree one (when $\deg(A)$ is odd.) Assume now that A has at most one column of degree one.

Case I: $\deg(A)$ is even. Suppose that $\deg(A) = 2m$ is even, and so each column of A is degree two and there are m columns. We will show that special operations (called f^-, f^+ operations) can be applied enough times so that finally the resulting matrix A will satisfy condition (N1) for normality, but possibly fail (N2). After A satisfies (N1), we may switch to a different types of operation (called g^-, g^+ operations) which will not disrupt the (N1) condition, and will eventually lead to a matrix which also satisfies (N2).

For $\mathbf{d} = (d_1, \ldots, d_{n-1}) \in S_{(4)}$, let $\mathcal{J}_{\mathbf{d}} = \mathcal{J}^0_{\mathbf{d}}$ where $\mathcal{J}^0_{\mathbf{d}}$ is as in the proof of Lemma 7.3, and let $f^-, f^+ : S_{(4)} \to S_{(2)}$ be given by $f^-(\mathbf{d}) = \mathbf{d}'$ and $f^+(\mathbf{d}) = \mathbf{d}''$ where $\mathbf{d}'$ and $\mathbf{d}''$ are again as in the proof of Lemma 7.3. Let an f^-, f^+ operation be the following. Choose j, j' such that $1 \leq j < j' \leq m$. Replace columns $c_j(A), c_{j'}(A)$ with $f^-(c_j(A) + c_{j'}(A)), f^+(c_j(A) + c_{j'}(A))$. Then,

$$c_j(A) + c_{j'}(A) = f^-(c_j(A) + c_{j'}(A)) + f^+(c_j(A) + c_{j'}(A)),$$

so an f^-, f^+ operation is of type (F4). Let $\mathbf{a} = (a_1, \ldots, a_{n-1}) = \sum_j c_j(A)$. We claim that after a finite number of f^-, f^+ operations, each entry $a_{i,j}$ of row i satisfies $|a_{i,j} - a_i/m| < 2$. The ith row (which has even integer entries that sum to a_i) is of minimal distance (using the standard Euclidean metric) from the constant vector $(a_i/m, \ldots, a_i/m)$ if and only if $|a_{i,j} - a_i/m| < 2$ for each j. Suppose x and y are even integers. It is easy to check that

$$(x - a_i/m)^2 + (y - a_i/m)^2 \geq \Big(e^+\Big(\frac{x+y}{2}\Big) - a_i/m\Big)^2 + \Big(e^-\Big(\frac{x+y}{2}\Big) - a_i/m\Big)^2$$

and the inequality is strict if and only if $|x - y| \geq 4$. Hence f^-, f^+ operations cannot take the ith row further from the constant vector $(a_i/m, \ldots, a_i/m)$. Now suppose that the ith row is as close as possible to $(a_i/m, \ldots, a_i/m)$ by applying f^-, f^+ operations. Suppose there is some $a_{i,j}$ such that $|a_{i,j} - a_i/m| \geq 2$. Then there is some j' such that $|a_{i,j'} - a_{i,j}| > 2$ since $\sum_j a_{i,j} = a_i$. Since $a_{i,j'} - a_{i,j}$ is even we have $|a_{i,j'} - a_{i,j}| \geq 4$. But now an f^-, f^+ operation on columns j, j' places row i strictly closer to $(a_i/m, \ldots, a_i/m)$, a contradiction. Therefore after sufficiently many f^-, f^+ operations the resulting matrix satisfies (N1).

Assume now that A satisfies (N1). We shall now switch to a different kind of (F4) operation (which does not disrupt (N1)) which will eventually give us a matrix that also satisfies (N2). These new operations are similar to the f^-, f^+ operations, except they depend globally on the entire matrix A, whereas the f^-, f^+ operations depend only on a pair of columns. Recall the definition of $\mathcal{J}_{\mathbf{a}}$. We have $\mathcal{J}_{\mathbf{a}} = \{i_1, \ldots, i_k\}$ is the set of all i, $2 \le i \le n-1$, such that $a_{i-1} \le 2m$ and $a_i \le 2m$, where $i_t < i_{t+1}$ for all t, $1 \le t < k$. Let $i_0 = 1$ and let $i_{k+1} = n$. Let

$$S_{(4)}(\mathbf{a}) = \{\mathbf{d} \in S_{(4)} \mid \mathcal{J}^0_{\mathbf{d}} \subset \mathcal{J}_{\mathbf{a}} \subset \mathcal{J}^1_{\mathbf{d}}\}.$$

Let

$$g^- : S_{(4)}(\mathbf{a}) \to S_{(2)},$$

$$g^-(\mathbf{d})_i = \begin{cases} e^-(d_i/2) & \text{for } i_{2t} \le i < i_{2t+1}, 2t \le k, \\ e^+(d_i/2) & \text{for } i_{2t+1} \le i < i_{2t+2}, 2t+1 \le k. \end{cases}$$

Let

$$g^+ : S_{(4)}(\mathbf{a}) \to S_{(2)},$$

$$g^+(\mathbf{d})_i = \begin{cases} e^+(d_i/2) & \text{for } i_{2t} \le i < i_{2t+1}, 2t \le k, \\ e^-(d_i/2) & \text{for } i_{2t+1} \le i < i_{2t+2}, 2t+1. \le k \end{cases}$$

We claim that for any two columns $c_j(A)$, $c_{j'}(A)$ the sum $\mathbf{d} = (d_1, \ldots, d_{n-1}) = c_j(A) + c_{j'}(A)$ is a member of $S_{(4)}(\mathbf{a})$. First we show that $\mathcal{J}_{\mathbf{a}} \subset \mathcal{J}^1_{\mathbf{d}}$. Suppose that $i \in \mathcal{J}_{\mathbf{a}}$. Then $a_{i-1}/m \le 2$ and $a_i/m \le 2$, so the entries in rows $i-1$ and i are at most 2 since $|a_{i-1,j} - a_{i-1}/m| < 2$ and $|a_{i,j} - a_i/m| < 2$ for all j. Hence the sum of any two entries in row $i-1$ is at most 4 and the sum of any two entries in row i is at most 4. Therefore each of d_{i-1} and d_i is at most 4 so $i \in \mathcal{J}^1_{\mathbf{d}}$. Next we show that $\mathcal{J}^0_{\mathbf{d}} \subset \mathcal{J}_{\mathbf{a}}$. Suppose $i \in \mathcal{J}^0_{\mathbf{d}}$. This means that $d_{i-1} + d_i = 4$ and each of $d_{i-1}/2$ and $d_i/2$ is an odd integer. Since $a_{i-1,j} + a_{i,j} \ge 2$ and $a_{i-1,j'} + a_{i,j'} \ge 2$ and

$$a_{i-1,j} + a_{i-1,j'} + a_{i,j} + a_{i,j'} = d_{i-1} + d_i = 4,$$

we have that $a_{i-1,j} + a_{i,j} = 2$ and $a_{i-1,j'} + a_{i,j'} = 2$. Suppose by way of contradiction that $a_{i-1} > 2m$. Then each entry of row $i-1$ is at least 2. Then $a_{i-1,j} = a_{i-1,j'} = 2$ and $a_{i,j} = a_{i,j'} = 0$. But now $d_i = a_{i,j} + a_{i,j'} = 0$ which contradicts that $d_i/2$ is an odd integer. Therefore $a_{i-1} \le 2m$. Similarly we can show $a_i \le 2m$. Hence $i \in \mathcal{J}_{\mathbf{a}}$.

We define a g^-, g^+ operation to be the following. Let $j < j'$ and replace columns $c_j(A)$, $c_{j'}(A)$ with $g^-(c_j(A) + c_{j'}(A))$, $g^+(c_j(A) + c_{j'}(A))$ in that order. Clearly any such g^-, g^+ operation is of type (F4) and it preserves the inequalities $|a_{i,j} - a_i/m| < 2$. We claim that a finite number of such operations results in a matrix in normal form. First notice that g^-, g^+ operations don't change the multi-set of entries in any given row since they preserve the (N1) condition. We determine how g^-, g^+ operations affect the order of the row entries. The output of g^- and g^+ is determined by the type of interval i belongs to; either $i_{2t} \le i < i_{2t+1}$ for some t or $i_{2t+1} \le i < i_{2t+2}$ for some t. Let us examine the case $i_{2t} \le i < i_{2t+1}$. Here g^- applies the e^- rule and g^+ applies the e^+ rule. Hence the result of a g^-, g^+ operation to columns j, j' with $j < j'$ is to put entries $a_{i,j}, a_{i,j'}$ into (weakly) increasing order. After applying these operations to all pairs j, j', the resulting ith row is weakly increasing. The case $i_{2t+1} \le i < i_{2t+2}$ is similar; this row will be weakly decreasing after g^-, g^+ operations are performed on all pairs of columns.

Case II: $\deg(A)$ *is odd.* Suppose $\deg(A) = 2m+1$ is odd, so there is one column of degree one and m columns of degree two. Apply a single (F3) operation so that the first column is the degree one column and columns 2 through $m+1$ are degree two. Always let A' denote A without the first column. We will show after enough operations of types (F3) and (F4) that the first column is $\xi_{2m+1}(\mathbf{a})$ and that A' satisfies conditions (N0) and (N1) for normality. Then g^-, g^+ operations can be performed on A' so that A' will eventually satisfy (N2).

The ith row must satisfy that $a_{i,1} \equiv i \pmod 2$, each $a_{i,j}$ is even for $j \geq 2$, and the sum $\sum_j a_{i,j} = a_i$. Clearly row i is closest to the vector

$$v_i = \left(\frac{a_i}{2m+1}, \frac{2a_i}{2m+1}, \frac{2a_i}{2m+1}, \ldots, \frac{2a_i}{2m+1} \right) \in \mathbb{Q}^{m+1}$$

if and only if

$$(*) \quad |a_{i,1} - a_i/(2m+1)| < 1 \text{ and } |a_{i,j} - 2a_i/(2m+1)| < 2 \text{ for all } j \geq 2.$$

These inequalities are necessary for the first column $c_1(A)$ to be $\xi_{2m+1}(\mathbf{a})$ and for A' to satisfy (N1). If each row satisfies $(*)$ then in fact $c_1(A) = \xi_{2m+1}(\mathbf{a})$ and A' satisfies (N1).

Suppose that $(*)$ holds for row i. We claim that operations of types (F3) and (F4) preserve $(*)$. We have that $|2a_{i,1} - a_{i,j}| < 4$ for each $j \geq 2$. But $2a_{i,1} - a_{i,j}$ is even so in fact $|2a_{i,1} - a_{i,j}| \leq 2 < 3$. Therefore, $|(a_{i,1} + a_{i,j})/3 - a_{i,1}| < 1$. But this implies that $\xi_3(c_1(A) + c_j(A))_i = a_{i,1}$ since $a_{i,1}$ has parity $i \pmod 2$ and is less than one unit from $(a_{i,1} + a_{i,j})/3$. Therefore row i is fixed by any (F3) operation. On the other hand if an (F4) operation is applied to columns j and j' then it either fixes $a_{i,j}$ and $a_{i,j'}$ or swaps their order since $|a_{i,j} - a_{i,j'}| \leq 2$.

Suppose that row i is as close as possible to v_i by applying (F3) and (F4) operations. Suppose by way of contradiction that $|a_{i,1} - a_i/(2m+1)| \geq 1$. Then there is some $j_0 \geq 2$ such that $a_i/(2m+1)$ is strictly between $a_{i,j_0}/2$ and $a_{i,1}$ since $a_i/(2m+1)$ is the weighted average of the entries in row i, where $a_{i,1}$ is weighted by 1 and $a_{i,j}$ is weighted by 2 for each $j \geq 2$. Therefore $|2a_{i,1} - a_{i,j_0}| > 2$. But $2a_{i,1} - a_{i,j_0}$ is even so in fact $|2a_{i,1} - a_{i,j_0}| \geq 4$. So $|(a_{i,1} + a_{i,j_0})/3 - a_{i,1}| \geq 4/3$. Without loss of generality suppose that $a_{i,1} < a_{i,j_0}/2$. Then we have

$$a_{i,1} < (a_{i,1} + a_{i,j_0})/3 < a_{i,j_0}/2.$$

Let k be the nearest integer of parity $i \pmod 2$ to $(a_{i,1} - a_{i,j_0})/3$. Then we have $a_{i,1} < k \leq a_{i,j_0}/2$. Let δ_i be the change in the distance between row i and v_i after applying an (F3) operation to columns 1 and j_0. Let $a = a_i/(2m+1)$ and let $t = k - a_{i,1}$. Then

$$\delta_i = (a_{i,1} + t - a)^2 + (a_{i,j_0} - t - 2a)^2 - (a_{i,1} - a)^2 - (a_{i,j_0} - 2a)^2 = 2t(t - (a_{i,j_0} - a_{i,1} - a)).$$

But we know that $0 < t \leq a_{i,j_0}/2 - a_{i,1} < a_{i,j_0} - a_{i,1} - a$. The first inequality follows from the fact that $a_{i,1} < k$ and the last inequality follows from that fact that $a_{i,j_0} > 2a = 2a_i/(2m+1)$. Hence δ_i is negative which means an (F3) operation takes row i strictly closer to v_i, a contradiction. Hence $|a_{i,1} - a_i/(2m+1)| < 1$. Now by our argument above for even degree matrices, we must have that the remaining entries $a_{i,j}$ differ by at most 2 from one another ((F4) operations can accomplish this) and consequently we also have that $|a_{i,j} - 2a_i/(2m+1)| < 2$ for each $j \geq 2$. Therefore, working row by row, we end up with a matrix A such that $c_1(A) = \xi_{2m+1}(\mathbf{a})$ and A' satisfies (N1). Now apply g^-, g^+ operations to A' so that finally A' satisfies (N2) as well. □

7.10. *Corollary. — For any S-matrix A, there is a unique S-matrix $\mathcal{N}(A)$ in normal form which is equivalent to A.*

7.11. *Theorem. — The ideal of relations of $\mathbb{Z}[S] \cong \mathrm{gr}(R)$ is generated by relations of degrees two, three, and four. Furthermore, the ideal of relations of $\mathbb{Z}[S^{even}]$ is generated by quadrics.*

Proof. One only needs to determine if two monomials in degree one and two variables are equal. This corresponds to deciding if two S-matrices are equivalent, which is true if and only if they have the same normal form. Operations of types (F2),(F3), and (F4) correspond to degree two, degree three, and degree four relations in the ideal of $\mathbb{Z}[S]$. By Lemma 7.9 these operations are enough to place any S-matrix A into its normal form $\mathcal{N}(A)$. Hence relations up to degree four must generate the ideal of $\mathbb{Z}[S]$. For the case of $\mathbb{Z}[S^{even}]$ an S-matrix is of even degree and we only need type (F4) operations to place it into normal form. In this case an (F4) operation is a quadric. □

8. Lifting the presentation from $\mathrm{gr}(R)$ to R

We now set up a presentation for R using both degree one and degree two Kempe graphs. Later we will remove the degree two generators (in case n is even) as they are redundant. Let $\widetilde{X}_G$ be a formal variable for each $G \in \mathcal{K}_{(1)} \cup \mathcal{K}_{(2)}$.

Let $\pi : \mathbb{Z}[\widetilde{X}_G]_{G \in \mathcal{K}_{(1)} \cup \mathcal{K}_{(2)}} \to R$ be the surjection sending $\widetilde{X}_G$ to $X_G \in R$. We know this is a surjection for two reasons. One is that X_G has leading term $Y_G \in \mathrm{gr}(R)$, and the Y_G generate $\mathrm{gr}(R)$ for $G \in \mathcal{K}_{(1)} \cup \mathcal{K}_{(2)}$. But we also know by Kempe's theorem that the X_G for $G \in \mathcal{K}_{(1)}$ generate R. Let I'' be the kernel of this map. Also, let $\overline{\pi} : \mathbb{Z}[\widetilde{X}_G]_{G \in \mathcal{K}_{(1)} \cup \mathcal{K}_{(2)}} \to \mathrm{gr}(R)$ be the surjection given by $\widetilde{X}_G \to Y_G$. Let $J = \ker(\overline{\pi})$.

Recall we have the explicit isomorphism between Kempe graphs and S, by $\phi : \mathcal{K} \to S$, where $\mathcal{K}$ has the semigroup structure $(G_1, G_2) \to G_1 * G_2$. If A is an S-matrix with m columns, let $m_A = \prod_{i=1}^m \widetilde{X}_{\phi^{-1}(c_i(A))}$. If A is in normal form, we shall say that m_A is in normal form. Recall that $\mathcal{N}(A)$ denotes the unique S-matrix in normal form such that $\sum_i c_i(\mathcal{N}(A)) = \sum_i c_i(A)$. We also let $\mathcal{N}(\mathbf{a})$ denote the unique S-matrix in normal form whose columns sum to $\mathbf{a}$, for any $\mathbf{a} \in S$.

8.1. *Proposition. — The relations in the X_G for G a Kempe graph of degree one or two are generated in degrees two, three, and four. The ring R^{even} is generated in degree one (by X_G where $G \in \mathcal{K}_{(2)}$) and the relations are generated by quadrics.*

Proof. By the previous section, we have that the ideal J is generated by elements of the form $m_A - m_{\mathcal{N}(A)}$ as A ranges over non-normal form S-matrices of degrees $2, 3, 4$. We lift these relations to get generators of $I'' = \ker(\pi)$. Take any S-matrix A of degree k, where $k \in \{2, 3, 4\}$.

Suppose that A has m columns, which correspond to Kempe graphs $G_1, \ldots, G_m$ via ϕ^{-1}. By Theorem 6.3, we know that

$$\pi(m_A) = X_{G_1 * G_2 * \cdots * G_m} + \sum_G c_G X_G,$$

where the Kempe graphs G appearing in the sum all have degree k, and have strictly lower weight than $G_1 * \cdots * G_m$. Let us assume now that the Kempe graphs of degree k are listed, $K_1, \ldots, K_m$, such that $w(K_i) \leq w(K_{i+1})$. If $B_i = \mathcal{N}(\phi^{-1}(K_i))$, then the change of basis matrix from $\pi(m_{B_1}), \ldots, \pi(m_{B_m})$ to $X_{K_1}, \ldots, X_{K_m}$ is therefore upper triangular, with identity matrices appearing in the diagonal blocks pertaining to the graphs of equal weight. Thus if we write $\pi(m_A) = \sum_{i=1}^m c_i \pi(m_{B_i})$, then the right hand side has the form

$$\pi(m_{\mathcal{N}(A)}) + \text{(terms of strictly lower weight)},$$

The above relation is the lift we need, since now the leading term of

$$m_A - (m_{\mathcal{N}(A)} + \text{(terms of strictly lower weight)}) \in I''$$

is equal to $m_A - m_{\mathcal{N}(A)}$. For the case of R^{even} we need to lift quadric relations only. □

Proof of Second Main Theorem 1.3. Suppose that n is even. Express each X_G of degree two in terms of degree one elements (by Theorem 2.3), and re-write the lifted relations of the above Proposition as relations in the degree one generators. These lifted relations must generate the ideal (and they have degree at most four). Suppose more generally that we have m points with weight $\mathbf{w}$, where $|\mathbf{w}| = n$ is even. Recall by Theorem 2.18 there is a surjection $I_n \to I_{\mathbf{w}}$. So $I_{\mathbf{w}}$ is also generated by relations of degree at most four.

Now suppose that there are m points with weight $\mathbf{w}$ and each weight w_i is even. Let $n = |\mathbf{w}|/2$ (n might be odd). Let $\widetilde{\mathbf{w}} = (2, 2, \ldots, 2) \in \mathbb{Z}^n$ (earlier we abbreviated this as 2^n). Now $R_{\widetilde{\mathbf{w}}} = R^{even}$, and by Proposition 8.1, the ideal of relations $I_{\widetilde{\mathbf{w}}}$ is generated by quadrics. Now by Theorem 2.18 we have that $I_{\widetilde{\mathbf{w}}}$ surjects onto $I_{\mathbf{w}}$, so $I_{\mathbf{w}}$ is generated by quadrics as well. □

References

[BCP] W. Bosma, J. Cannon, and C. Playoust, *The Magma algebra system. I. The user language*, J. Symbolic Comput. **24** (1997), 3-4, 235–265.

[BP] V. Batyrev and O. Popov, *The Cox ring of a del Pezzo surface*, in *Arithmetic of higher-dimensional algebraic varieties (Palo Alto, CA, 2002)*, 85–103, Progr. Math., 226, Birkhäuser Boston, Boston, MA, 2004.

[B] N. Bourbaki, *Commutative Algebra, Chapters 1-7*, Elements of Math. (Berlin), Springer-Verlag, Berlin, 1998.

[DLMc] J. De Loera and T. McAllister, *Vertices of Gelfand-Tsetlin polytopes*, Discrete Comput. Geom. **32** (2004), no. 4, 459–470.

[De] H. Derksen, *Universal denominators of Hilbert series*, J. Algebra **285** (2005), no. 2, 586–607.

[Do] I. Dolgachev, *Lectures on Invariant Theory*, LMS Lecture Note Series, 296, Cambridge U.P., Cambridge, 2003.

[DoO] I. Dolgachev and D. Ortland, *Point sets in projective spaces and theta functions*, Astérisque No. 165, (1988), 210 pp. (1989).

[FHu] P. Foth and Y. Hu, *Toric degenerations of weight varieties and applications*, in *Travaux mathématiques*, Fasc. XVI, 87–105, Univ. Luxemb., Luxembourg, 2005.

[GL] N. Gonciulea and V. Lakshmibai, *Flag Varieties*, Hermann, 2001.

[HK] J.-C. Hausmann and A. Knutson, *Polygon spaces and Grassmannians*, Enseign. Math. **43** (1997), 173-198.

[HMM] B. Howard, C. Manon and J. Millson, *The toric geometry of polygons in Euclidean space*, in preparation.

[HMSV] B. Howard, J. Millson, A. Snowden, and R. Vakil, `http://www.math.princeton.edu/~asnowden/HMSV/`.

[Hu] Y. Hu, *Stable configurations of linear subspaces and quotient coherent sheaves*, Pure and Applied Mathematics Quarterly **1** 2005 127–164

[KY] Y. Kamiyama, T. Yoshida, *Symplectic Toric Space Associated to Triangle Inequalities*, Geometriae Dedicata **93** (2002), 25-36.

[KM] M. Kapovich and J. Millson, *The symplectic geometry of polygons in Euclidean space*, J. Diff. Geom. **44** (1996), no. 3, 479-513.

[KT] S. Keel and J. Tevelev, *Equations for* $\overline{M}_{0,n}$, math.AG/0507093.

[Ke] A. Kempe, *On regular difference terms*, Proc. London Math. Soc. **25** (1894), 343-350.

[MFK] D. Mumford, J. Fogarty, and F. Kirwan, *Geometric Invariant Theory*, 3rd ed., Springer-Verlag, Berlin, 1994.

[OS] P. Olver and C. Shakiban, *Graph Theory and Classical Invariant Theory*, Adv. Math. **75**, no. 2 (1989), 212-245.

[Sh] T. Shioda, *On the graded ring of invariants of binary octavics*, Amer. J. Math. **89** 1967 1022–1046.

[Sl] N.J.A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, available electronically at `http://www.research.att.com/~njas/sequences/`, 2005.

[SpSt] D. Speyer and B. Sturmfels, *The tropical Grassmannian*, Adv. Geom. **4**, no. 3, (2004), 389-411.

[St1] B. Sturmfels, *Gröbner Bases and Convex Polytopes* AMS University Lecture Series **8** (1991).

[St2] B. Sturmfels, *Algorithms in Invariant Theory* Texts and Monographs in Symbolic Computation Series, Springer-Verlag (1994).

B. Howard: Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA, howardbj@umich.edu

J. Millson: Department of Mathematics, University of Maryland, College Park, MD 20742, USA, jjm@math.umd.edu

A. Snowden: Department of Mathematics, Princeton University, Princeton, NJ 08544, USA, asnowden@math.princeton.edu

R. Vakil: Department of Mathematics, Stanford University, Stanford, CA 94305, USA, vakil@math.stanford.edu