Geometric Complexity Theory I: An approach to the P vs. NP and related problems

Ketan Dattatraya Mulmuley*
The University of Chicago and I.I.T., Bombay

Milind Sohoni
I.I.T., Bombay, and the University of Chicago

April 23, 2001

Abstract

We suggest an approach based on geometric invariant theory to the fundamental lower bound problems in complexity theory concerning formula and circuit size. Specifically, we introduce the notion of a *partially stable* point in a reductive-group-representation, which generalizes the notion of stability in geometric invariant theory due to Mumford [37]. Then we reduce fundamental lower bound problems in complexity theory to problems concerning infinitesimal neighbourhoods of the orbits of partially stable points. We also suggest an approach to tackle the latter class of problems via construction of *explicit obstructions*.

1 Introduction

It is generally agreed by now [41] that the problems in complexity theory, such as P vs. NP [10, 25, 20], and the related lower bound problems for formula or arithmetic-circuit size are of fundamental importance and may have connections with deep issues in other areas of mathematics. In this article we suggest an approach based on geometric invariant theory [37, 38] to several such lower bound problems concerning formula and circuit size. We shall focus on formulae and circuits over integers or algebraically closed fields of arbitrary characteristic, before addressing the ones over finite fields. Our work is based on the classical geometric invariant theory developed by Hilbert [19], Weyl [46] and others, as well as, on the modern theory developed by Mumford [37], Nagata [31], Seshadri [40], Haboush [17], Luna [27], Kempf [23]

*The work of the first author was supported by the Guggenheim Fellowship, and NSF grant CCR 9800042

and others. Especially central in this paper is the powerful notion of stability in geometric invariant theory [19, 37, 23, 27, 40].

For a survey of earlier algebraic and geometric techniques in the lower bound problems see [5]; for a survey of earlier combinatorial techniques, see [3]. A starting point for our investigation was a result by one of the authors [33] that provided a concrete support for the $P \neq NC$ conjecture by an unconditional proof of its weaker implication in a restricted, but natural and realistic *PRAM model without bit operations*, which contains virtually all known fast parallel algorithms for algebraic and weighted-combinatorial-optimization problems. In [33] it is shown that well-known network flow problems such as min-cost flow and max-flow, which can be solved efficiently without bit operations sequentially–i.e., they have the so-called strongly polynomial time algorithms [16]–cannot be solved efficiently in this model. Specifically, consider the max-flow problem. Since it is P-complete, the $P \neq NC$ conjecture implies that it cannot be solved in the usual PRAM model in $O(\text{polylog}(N))$ time using $\text{poly}(N)$ processors, where N is the total bitlength of the input. Since the PRAM model without bit operations is just a restricted version of the usual PRAM model, obtained by not allowing the bit operations in the usual model, it follows from the $P \neq NC$ conjecture that the max-flow problem cannot be solved in the PRAM model without bit operations, as well, in $O(\text{polylog}(N))$ time using $\text{poly}(N)$ processors. In [33], this weaker implication of the $P \neq NC$ conjecture is proved unconditionally, using some algebraic geometry (in fact, a stronger lower bound is proved).

This was a turning point, which provided a compelling argument for believing that algebraic geometry will play a crucial role in the P vs. NC and other separation problems in complexity theory. The approaches attempted earlier (cf. e.g. [3]) were combinatorial, but they had a serious natural-proof limitation as pointed out in [39]. The present approach does not have this limitation, since it is based on *explicit constructions* (cf. Section 5 and part II [34]) and the results, such as the Hilbert-Mumford-Kempf criterion [37, 23] for stability (cf. Section 3), which only work for points with specific types of stabilizers–such points are very rare, because almost all points in the representations under consideration have trivial stabilizers[1]. (Loosely speaking, according to the terminology in [39], the probabilistic proof for existence of expanders would be natural, because it ends up proving that a significant fraction of graphs are expanders, whereas the proof based on explicit construction [26, 29] would be unnatural, because it works for only a specific graph.) A closer analysis of the limitation of the techniques in [33] in the context of the usual P vs. NC problem suggested that geometric invariant theory may help. This led us to explore the role of geometric invariant theory in the fundamental lower bound problems of complexity theory.

Geometric invariant theory is the study of the action of an algebraic, reductive

[1] For example, the proof of Theorem 4.6 works for only points that are stabilized by the stabilizer R of the determinant. But, in the space $P(V)$ under consideration, the determinant is the only point that is stabilized by R.

group on an algebraic variety. One important problem here is to understand the structure of this action in the vicinity of the orbit of any fixed point. If the variety is affine and smooth, the characteristic is zero, and the point is stable in the sense of Mumford [37] and Kempf [23], i.e., its orbit is closed, then Luna's Étale Slice Theorem (cf. [27], and also the appendix of [37]), based on the earlier work on stability of Mumford [37] and others, says that the orbit has a neighbourhood which looks like a certain homogeneous fibre bundle. This gives a complete understanding of the structure of the action in the vicinity of the orbit of a stable point. The action in the vicinity of the null cone, i.e., the set of points in V on which all nonconstant homogeneous G-invariant polynomial functions on V vanish, is, however, not well understood–though there are some results concerning the structure of the null cone; e.g. the Hilbert-Mumford criterion [37], and Luna's stratification [38]. The reason is that the action in the vicinity of the orbit of an arbitrary point of the null cone may be too complex. But the situation may be better if the point under consideration has some weaker form of stability. We define one such notion, called partial stability, that is also applicable to the points in the null cone. Then we show that an in-depth understanding of the structure of a reductive-group action in the vicinity of the orbit of a partially stable point would lead to resolution of the lower bound problems mentioned above. We also suggest an approach–which will be developed in depth in Part II [34]–for studying the group action in the vicinity of the orbit of a partially stable point in the null cone. Such study is of independent mathematical interest in geometric invariant theory.

We mentioned above that the max-flow problem does not have an efficient algorithm in the PRAM model without bit operations. Its dual is the problem of computing a minimum-weight s-t-mincut in a flow network. This is the weighted optimization problem associated with s-t-mincuts. The corresponding weighted counting problem is the problem of computing total weight of all s-t-mincuts in a flow network. Up to fast parallel reductions, this problem is equivalent in the PRAM model without bit operations to the problem of computing the permanent of a matrix, which is $\#P$-complete [44]. Hence it is expected to be much harder than the problem of computing a mimimum weight s-t-mincut. And yet, paradoxically, the task of proving a lower bound for this harder problem in the PRAM model without bit operations seems harder. This is so even if the model is restricted even further by allowing the running time or the number of processors to depend only on the number of input parameters and not their bitlength, and furthermore, by not allowing any branching or indirect references. The lower bound problem in this substantially restricted model is then equivalent to the one concerning formula size over integers. Valiant [45] conjectured that the permanent has no formula of polynomial size; this was posed as an algebraic analogue of the P vs. NP problem with the hope that techniques used in resolving it may shed light on the P vs. NP problem. Indeed, it is the first problem that we address in this paper.

The rest of this paper is organized as follows. In Section 2 we outline our unified approach, which reduces the fundamental lower bound problems under consideration to instances of a certain orbit-closure problem in geometric invariant theory. In Section 3 we review the main results in geometric invariant theory that we need. In Section 4, we illustrate our approach on the lower bound problem for the formula size of the permanent. In Section 5, we formulate an approach to the orbit closure problem via *explicit obstructions*–these will be studied in depth in Part II. In Sections 6, 7 and 8 we consider the P vs. NP problem, with emphasis on its arithmetic implication.

We wish to thank Prof. Madhav Nori, Prof. C. S. Seshadri, and Prof. Burt Totaro for several illuminating discussions and help throughout the course of this research.

2 A unified approach

Let F be an algebraically closed field. Let V be the vector space of homogeneous forms of degree m in l variables over F. Let Y denote the l-vector whose entries are these variables. Consider the reductive group $G = SL_l(F)$. It has a natural action on V. The action of any $\sigma \in SL_l(F)$ on a form $h \in V$ is given by: $(\sigma h)(Y) = h(\sigma^{-1}Y)$. Let $P(V)$ denote the projective space associated with V. Then G acts on $P(V)$ as well. We shall think of a form as belonging to either V or $P(V)$ depending upon the context. Following Mumford [37] and Kempf [23], we shall say that the form $f \in P(V)$ is *stable* if the G-orbit of f in the affine space V is (Zariski) closed.

We shall reduce all lower bound problems under consideration for formulae or circuits over integers or algebraically closed fields of arbitrary characteristic to instances of the following general problem.

The orbit closure problem: Given fixed forms $g, f \in P(V)$, does f belong to the (projective) Zariski-closure $\Delta[g]$ of the G-orbit of g in $P(V)$?

The explicitly given forms g and f will depend on the lower bound problem under consideration. They will be parametrized by l, the number of variables, which asymptotically goes to infinity. The orbit closure $\Delta[g]$ is a compactification of the homogeneous space G/H, where H is the stabilizer of g. The algebraic geometry of compactifications of homogeneous spaces is well understood in some special cases, such as toric varieties [13], spherical embeddings [28], and compactifications of symmetric spaces [9]. But in general, it seems very complex: a rough measure of this complexity, proposed by Luna and Vust [28], is the difference between the dimension of the compactification and that of the orbit of a Borel subgroup. This is zero for spherical embeddings, but quite high–roughly half of $\dim(G)$–in our case. This is one reason why the orbit closure problem may well be intractable if f and g were arbitrary. The crucial result in this paper, which gives structure to this otherwise wild problem and connects complexity theory and geometric invariant theory, is that f which arises in the context of lower bound problems in complexity theory studied here has certain stability property, in the sense of geometric invariant theory [37], and moreover, both f and g have large, nontrivial stabilizers that in a sense characterize them. We then formulate an approach to the orbit closure problem, based on so-called *explicit obstructions*, that exploits this stability and presence of large, characterizing stabilizers (Section 5). This approach is developed in depth in Part II [34].

Specifically f that arises in our context turns out to be a partially stable point in the nullcone of G (Theorems 4.8, 7.4). Partial stability (Section 3) is a version of stability for points in the null cone obtained by relaxing the notion of stability due to Mumford [37], which is for points outside the nullcone. Roughly this means that (1) f is "almost" stable with respect to a reductive subgroup $L \subseteq G$ that is a Levi-subgroup of a maximal parabolic subgroup $P \subseteq G$, (2) the orbit Gf is a fibre bundle over the base space G/P (whose algebraic geometry is very well understood [24]) such that each fibre is isomorphic to the L-orbit of f, (3) and the dimension of the fibre is sufficiently large in comparison to that of the orbit Gf. The precise relationship between these two dimensions will depend on the lower bound problem under consideration. Generally, the dimension of the orbit Gf will be bounded by a polynomial in the dimension of the orbit Lf, as the parameter l goes to infinity. If one could show that the answer to such an instance of the orbit closure problem in a particular lower bound problem is negative for all large enough (or, more generally, infinitely many) l, then it would imply the desired lower bound.

Since g that arises in our context has a nontrivial, reductive stabilizer that in a sense characterizes its structure, it is possible to reduce the orbit closure problem further to a certain stabilizer problem as follows. Let D_g denote the stabilizer of g in $SL_l(F)$. First, we observe that if f lies in the closure of the $SL_l(F)$-orbit of g in $P(V)$, then, thinking of f and g as elements of the affine space V, f will lie in the closure of the $GL_l(F)$-orbit of g in V. But the stabilizer within $SL_l(F)$ of any point in the $GL_l(F)$-orbit of g is a conjugate within $SL_l(F)$ of D_g. Hence, every (Zariski) open neighbourhood of f within V will contain a point whose stabilizer in $SL_l(F)$ is a conjugate of D_g, and the same will hold for any point in the $SL_l(F)$-orbit of f in V. When this happens we will say that the group $D_g \subseteq SL_l(F)$ is present in the infinitesimal neighbourhood of the $SL_l(F)$-orbit of f. The orbit closure problem can now be reduced to the following:

The stabilizer problem Let $D \subseteq G = SL_l(F)$ be an explicitly given reductive subgroup and $f \in V$ an explicitly given form. Can D occur as a stabilizer in the infinitesimal neighbourhood of the G-orbit of f in V?

If $f \in V$ were a stable form, then Luna's Étale Slice Theorem provides a satisfactory answer to this problem. Luna's result in our setting can be roughly stated as follows: Let H be the stabilizer of f. If f is stable then there exists an affine variety N with an H-action such that some G-invariant affine neighbourhood of the orbit of f "looks like" the homogeneous fibre bundle $G \times_H N$–this is a fibre bundle over the base space G/H with fibre N, and it is associated with the primary bundle $G \to G/H$. It easily follows from this result that if D were to occur in the infinitesimal neighbourhood of the $SL_l(F)$-orbit of f then some conjugate of D must be contained in H. This implies, in turn, that if there exists a G-representation W that contains a trivial H-representation but not a trivial D-representation, then D cannot lie in the infinitesimal neighbourhood of $SL_l(F)$-orbit of f, and hence f cannot lie in the (projective) closure of the $SL_l(F)$-orbit of a form $g \in P(V)$, with $D_g = D$. We say that such a W is an *obstruction* for the pair (f, g); i.e., its existence acts as a guarantee that f cannot lie in the projective closure of the orbit of $g \in P(V)$.

In our problem, f would only be partially stable. So what we need is an extension of the preceding approach based on Luna's result from the stable to the partially stable situation. We suggest one such approach in this paper (Section 5) based on explicit construction of representation-based obstructions, as above, in the partially stable situation; this serves as a starting point for our work in Part II [34]. In the context of lower bound problems in complexity theory, one can argue that obstructions ought to exist in plenty, assuming that the conjectured lower bound really holds. However, the problem is such that one has to prove their existence by constructing them more or less explicitly. Explicit construction has already been useful in complexity theory in a different context. For example, Margulis, Lubotzky, Phillips and Sarnak [29, 26] give an explicit construction of expanders that are useful in pseudorandom generation. The essential difference is that proving existence of expanders is easy, while proving existence of obstructions is itself a main problem.

The lower bound problem that arises in the context of the actual P vs. NP question (nonuniform, boolean version) is for circuits defined over finite fields[2]. We shall reduce it (Section 7) to a family of orbit-closure and stabilizer problems where f ranges over forms that are so-called F_p-equivalent to a certain base form f_b. We show that f_b is partially stable, and conjecture the same for all forms F_p-equivalent to it. In any case, the lower bound problem for circuits (or formulae) over integers is a weaker version of the one over finite fields, in the sense that if g is a polynomial function with integer coefficients (e.g. the permanent) and F_p a fixed finite field, then a lower bound on the size of a circuit (or formula) over F_p for computing g, modulo F_p, implies the same lower bound on the size of a circuit (or formula) over integers for computing g. As such, it seems preferable to focus at first only on circuits or formulae over integers. In turn, the lower bound problem for circuits or formulae over integers for computing g is a weaker version of the problem over complex numbers. Hence, we shall concentrate mainly on circuits or formulae over complex numbers, or more generally, over algebraically closed fields of arbitrary characteristic.

[2]This finite field is usually taken to be F_2, but F_p, for any small enough prime p, will also do.

3 Preliminaries

Geometric invariant theory is the study of actions of algebraic groups on algebraic varieties. Here reductive algebraic groups, such as the special linear group $SL_n(F)$, F algebraically closed, play a special role because they are geometrically reductive: this means if V is a representation of a reductive group G then there are enough polynomial functions on V which are invariant under G to distinguish disjoint closed G-invariant subsets of V. This was proved by Mumford [36] for linearly reductive groups, which include the classical simple groups over the complex field (cf. Weyl [46]), and was conjectured in general. It was eventually established by Seshadri [40] for $SL_2(F)$ and Haboush [17] in general. What we shall really need is geometric reductivity of the special linear group $SL_n(F)$, where F is an algebraically closed field of arbitrary characteristic; for general n this was proved by Formanek and Procesi [12] independently of Haboush's general result.

3.1 Stability and partial stability

Consider the action of a reductive group G on an algebraic variety $\mathcal{X}$. It turns out that one cannot always put the structure of an algebraic variety on the set of all orbits of G in $\mathcal{X}$, but only on a subset chosen judiciously. For this purpose Mumford [36] introduced the notion of stability, building on the classical work of Hilbert [19], and exploited it systematically in his construction of quotients and moduli spaces. The definition of stability in the literature, though, has been anything but stable. Here we shall give a simplified definition following Kempf [23], which is somewhat different from the one in [36, 37].

Let V be a representation of a reductive group G. Then G acts on the projective space $P(V)$ of lines in V through the origin. Let y be a point in $P(V)$ and $x \in V$ any nonzero point on y. We say that y is *G-stable*–or simply *stable* if G is clear from the context–iff the orbit $G.x$ is (Zariski) closed, and *properly stable* iff, in addition, the stabilizer of y in G is finite. By abuse of notation, we shall sometimes say that x is stable. We say that y is *semistable* iff the closure of the orbit $G.x$ does not contain zero. These definitions clearly do not depend on the choice of x on y. We will say that $x \in V$ is *unstable* or nilpotent if the the closure of its orbit $G.x$ contains zero. The set of all nilpotent points in V is called the null-cone. It can also be characterized as the set of common zeroes of all homogeneous invariants of positive degree. The concept of the null cone, important to us, was introduced by Hilbert [19].

Next, we define a weaker form of stability called partial stability, which will play a crucial role in this paper, and will be studied in depth in Part II. First assume that the characteristic is zero. Roughly, a point is partially stable if it is stable with respect to a large, nicely embedded, regular[3] reductive subgroup of G. Formally,

[3]this means the root system of the subgroup is a subsystem of the root system of G.

we say that y is partial stable, if there exists a maximal parabolic subgroup $P \subseteq G$ containing the stabilizer H of x in G such that:

1. H contains the unipotent radical R of P–in other words, the unipotent radical fixes x,

2. y is "almost" stable with respect to a Levi-subgroup $L \subseteq P$: This means the stabilizer L' of x in L is reductive (as it would be if y were stable with respect to L [32]), and y is stable with respect to a regular reductive subgroup $K \subseteq L$, whose rank is less than that of L by a small defect δ (in our case δ will be one).

3. The dimension of the orbit $Lx = L/L'$ is large enough in comparison with the dimension of the orbit Gx. The exact ratio Δ between these two dimensions will depend on the lower bound problem under consideration. Generally, the dimension of Gx will be bounded by a polynomial in the dimension of Lx.

When $P = G$, $\delta = 0$ and $\Delta = 1$, this coincides with the usual notion of stability. So δ and Δ measure deviation from (complete) stability. It follows from the definition that H has a Levi decomposition of the form $L'R$, where L' is a reductive subgroup of L. It also follows that the orbit Gx is a fibre bundle over G/P such that each fibre is an affine subvariety of V isomorphic to L/L', the L-orbit of x. Moreover the dimension of the fibre is sufficiently large in comparison with the dimension of Gx. That Gx is a fibre bundle over G/P with affine fibre has an important consequence for us: computation of the cohomology of a quasicoherent sheaf over Gx can be reduced to a similar problem over G/P (cf. Theorem 5.5), where it is well understood [24]. Such explicit computation of cohomology will turn out to be important in our approach to the orbit closure problem (Section 5).

If the characteristic is positive, we assume in the preceding definition of partial stability that G is a connected split reductive F-group (i.e., defined over a finite field F), and that P and H are also F-groups [42]. This is required to ensure that the Levi decomposition holds [4, 42].

3.2 Example

Let Y be a vector space of dimension $m+1$ with coordinates $y_0, \ldots, y_m$, X its subspace with coordinates $x_0, \ldots, x_n$, $n < m$, where $x_i = y_i$, $i < n$, $\hat{X} \subseteq X$ the subspace with coordinates $x_1, \ldots, x_n$. We shall also think of y_i's as variables and, abusing the notation, denote by Y the variable vector with entries $y_0, \ldots, y_n$, and so on. Let $V = \mathrm{Sym}^r(Y^*)$ be the space of forms of degree r in $y_0, \ldots, y_m$, and $W = \mathrm{Sym}^s(\hat{X}^*)$, $s < r$, the space of forms of degree s in $x_1, \ldots, x_n$. We can embed W in V in a natural way using x_0 as the homogenizing variable. Formally, let ϕ be an embedding of W in V given by: $(\phi h)(Y) = x_0^{r-s} h(\hat{X})$. Suppose $h \in W$ is stable with respect to the natural action of $SL_n = SL(\hat{X})$ on W, with (reductive) stabilizer $K \subseteq SL(\hat{X})$.

Claim 3.1 *The embedded form $\phi h \in V$ is partially stable with respect to the action of $SL_{m+1} = SL(Y)$ on V. In other words, any stable form in the $SL(\hat{X})$-representation W becomes partially stable when embedded in the $SL(Y)$-representation V.*

Proof: First notice that the stabilizer of $\phi(h)$ in $SL(X) \subseteq SL(Y)$ is just K, assuming $r-s$ is larger than n. Though ϕh is not stable with respect to $SL(X)$–in fact, it lies in its null cone–it is stable with respect $SL(\hat{X})$. Let $P \subseteq SL(Y)$ be the maximal parabolic subgroup that preserves the space X, i.e., consisting of matrices, acting on columns, of the form

$$\begin{bmatrix} * & * \\ 0 & * \end{bmatrix} \tag{1}$$

where the diagonal stars denote matrices of dimension $n+1$ and $m-n-1$ respectively. This contains the stabilizer H of ϕh, as the reader can easily show. The unipotent radical R of P (i.e., the top right $*$) acts trivially on ϕh, and so belongs to H. Let L be the Levi subgroup of P consisting of block diagonal matrices. The lower block of matrices in L also acts trivially on ϕh. So only the action of the upper block matters. Thus the stabilizer H has a Levi decomposition $L'R$, where $L' \subseteq L$ consists of matrices of the form:

$$\begin{bmatrix} \alpha & 0 & 0 \\ 0 & C & 0 \\ 0 & 0 & D \end{bmatrix} \tag{2}$$

where (1) α is a nonzero constant, (2) a multiple of C belongs to K, and (3) $D \in SL(\bar{X})$, $\bar{X}$ being the complement of X in Y. Since ϕh is stable with respect $SL(\hat{X})$, whose rank is one less than that of $SL(X)$, it is clear that ϕh is almost stable with respect to L with defect $\delta = 1$. Q.E.D.

Thus each $SL(\hat{X})$-stable point in $P(W)$ gives rise to a partially stable point in $P(V)$ with respect to $SL(Y)$-action with defect one. The partially stable points in all our applications will arise from stable points in this simple fashion. Notice that the $SL(Y)$-orbit of ϕy is a fibre bundle over G/P, which is just a Grassmanian in this case. In our applications, m will be typically polynomial in n, and so too the ratio Δ between the dimension of $SL(Y)$-orbit and the dimension of the fibre. The dimension of the fibre will be sufficiently large, in this sense, compared to that of the $SL(Y)$-orbit; the importance of this condition will become clear in Sections 4 and Section 6.

Also notice that explicit decomposition of V as an $SL(\hat{X})$-module can be determined by successive applications of the Pieri formula for the branching rule. The point ϕh belongs to the unique irreducible $SL(\hat{X})$-module isomorphic to $\mathrm{Sym}^s(\hat{X}^*)$ in this decomposition, and it is (completely) stable in this module with respect to the $SL(\hat{X})$-action. This is another way of looking at partially stable points that would arise in our context.

3.3 Criterion for stability

Given a point $x \in V$, how can one effectively decide if it belongs to the null cone? This is rather important, because though the ring of invariants for a reductive group action on affine varieties is finitely generated, only in rare cases can one determine an explicit set of generators for this ring. For $SL_n(C)$ Hilbert [19] proved a remarkable criterion that quite often enables one to prove unstability or semistability without explicitly writing down a single invariant. This criterion was generalized to arbitrary reductive groups over any characteristic by Mumford [36]. The resulting Hilbert-Mumford criterion is a powerful tool for proving semistability. Actually Mumford proved the following more general result, which is what we shall need.

For any closed G-invariant subset S of V, define $x \in V$ to be S-unstable if the closure of the orbit $G.x$ meets S. In the preceding classical case, i.e., when x belongs to the null cone, S contains just the zero point of V. In general, if the orbit of x is not closed, the boundary of this orbit contains a unique closed orbit of minimum dimension–we can let S be this closed orbit, and then x is S-unstable. Mumford [36] proved that if x is S-unstable then there is a one-parameter subgroup λ of G which drives x to S; i.e., the closure of the orbit $\lambda.x$ also meets S. Kempf [23] generalized this result significantly. His result will provide us the main technical tool for proving stability of various forms that arise in the context of lower bound problems.

Kempf's criterion for stability: Assume that G has no non-trivial central one-parameter subgroup. If the stabilizer of $y \in P(V)$ in G is not contained in any proper parabolic subgroup of G then y is stable.

Kempf's other results relevant to us may be summarized as follows. Suppose $x \in V$ is S-unstable. Then

1. One can identify a non-empty subset $\Delta_{S,x}$ of one-parameter subgroups λ which drive x to S "most rapidly".

2. With every one-parameter subgroup λ, there is associated a parabolic subgroup $P(\lambda)$ of G consisting of those elements $g \in G$ such that $\lim_{t\to 0} \lambda(t)g\lambda^{-1}(t)$ exists in G. Then there is a parabolic subgroup $P_{S,x}$ of G such that $P(\lambda) = P_{S,x}$ for all $\lambda \in \Delta_{S,x}$.

3. Every maximal torus of $P_{S,x}$ contains a unique element of $\Delta_{S,x}$.

4. If G' is the stabilizer of x in G, then $P_{S,x}$ contains G'.

5. If, in addition, the characteristic is zero, and H is a reductive subgroup of G fixing x, then there is an element of $\Delta_{S,x}$ contained in the centralizer $Z_G(H)$ of H in G. The result also holds in positive characteristic, provided G is an F-group, F a finite field, and the Levi decomposition holds for parabolic F-subgroups of G, as it will in our case.

3.4 Slice theorem

A powerful tool in the analysis of actions of reductive groups around orbits of stable points is Luna's étale slice theorem. Let x be a stable point in a linear representation V of a reductive group. Let H be the stabilizer of x, which is reductive [32]. Then the orbit Gx of x is isomorphic to G/H. Let T denote the tangent space of the orbit at x; it has a natural H-action. Assume that the characteristic is zero. Since H is reductive, V has an H-module decomposition $T \oplus M$, where M can be thought of as a "normal" space at x. Luna's slice theorem says that the orbit $Gx = G/H$ has a neighbourhood in V that looks like (or more precisely, has a finite sheeted covering by) the induced fibre bundle $G \times_H N$, for some algebraic H-subvariety $N \subseteq M$. Formally

Theorem 3.2 *[27, 37, 38] There is an excellent map ϕ from $G \times_H N$ to a G-invariant Zariski-open neighbourhood U of Gx.*

The map ϕ is called excellent (or strongly étale) if the induced map ϕ/G : $G \times_H N/G \to U/G$ is étale and ϕ is obtained from ϕ/G by base extension. Here $G \times_H N$ is the quotient of $G \times N$ with respect to the action of H given by:

$$h(g,n) = (gh^{-1}, hn).$$

It has a natural G-action induced by the G-action (left multiplication) on $G \times N$ via the translation of the first factor.

When the base field is the field of complex numbers, the slice theorem implies that the orbit Gx has an analytic neighbourhood that is actually isomorphic as a G-variety to $G \times_H N$ for some analytic (but not necessarily algebraic) H-variety $N \subseteq M$; N is called the slice at x. The slice theorem also implies that if Q occurs as a stabilizer in the infinitesimal neighbourhood of the orbit Gx, then some conjugate of Q is a subgroup of H. However, determining if some conjugate of a given $Q \subseteq G$ is a subgroup of another given $H \subseteq G$ is, in general, nontrivial. This problem can be handled using explicit obstructions as in Section 5.

4 Formula size

In this section we illustrate our approach by addressing a concrete lower bound problem for the formula size over an algebraically closed field. Suppose we are given a homogeneous form $g(X)$ of degree d in the n-dimensional vector variable $X = (x_1, \ldots, x_n)$ over an algebraically closed field F of arbitrary characteristic. We want to know if $g(X)$ has a formula of poly(n) (polynomial in n) size, i.e., of size $\leq n^a$, for some fixed constant a. By the formula size we mean the number of arithmetic operators in the formula. The formula size of a function is a good indicator of its parallel complexity [11, 45].

For example, suppose $n = k^2$, for some k. Think of X as a $k \times k$ matrix; so we shall denote the entries of X by x_{ij} instead of x_i. Let $g(X)$ be the permanent function:

$$g(X) = \mathrm{perm}(X) = \sum_{\sigma} \prod_{i} X_{i\sigma(i)},$$

where σ runs over all permutations of size k. This formula for the permanent has size $k \cdot k!$. But there is a non-obvious substantially smaller formula of size $O(2^{O(k)})$ [30]. Valiant [44] conjectured that the permanent does not have a formula of size polynomial in k. No nonlinear lower bound is known so far; cf. [15] for a linear lower bound.

We shall now reduce the general lower bound problem for the formula size to an instance of the orbit closure problem in Section 2. Let Y be an $m \times m$ variable matrix, where $m > n$. Identify X with any $n \times n$-submatrix $\phi(X)$ of Y. Use some entry y of Y not in $\phi(X)$ as a *homogenizing* variable. Then any homogeneous polynomial g of degree d on X can be trivially identified with a homogeneous polynomial g^ϕ of degree m in the entries of the variable Y, where $g^\phi(Y) = y^{m-d} g(\phi(X))$.

Let V be the vector space of homogeneous forms of degree m in Y. Let $P(V)$ be the associated projective space. Let W be the vector space of homogeneous forms of degree d in X, and $P(W)$ the associated projective space. Then the function ϕ, which maps g to g^ϕ, is an embedding of W ($P(W)$) in V (resp. $P(V)$).

Let $G = SL_{m^2}(F)$. It has a natural action on V. Specifically, any $\sigma \in SL_{m^2}(F)$ acts on an $h \in V$ as: $(\sigma h)(Y) = h(\sigma^{-1} Y)$, where we think of Y as an m^2-vector, by linearly ordering its entries in any way, so that the $m^2 \times m^2$-matrix σ^{-1} acts on Y in the usual way. This induces a natural action on $P(V)$ as well. The space V (and $P(V)$) contains the determinant form $\det(Y)$. Let $G(\det(Y)) \subseteq P(V)$ denote the orbit of $\det(Y)$ under this action (thinking of $\det(Y)$ as an element of $P(V)$.) Let $\Delta[\det(Y)]$ denote its Zariski closure in $P(V)$.

Proposition 4.1 *If $f(X)$ has a formula of size l then $f^\phi(Y)$ lies in the projective orbit closure $\Delta[\det(Y)]$ of the determinant, where $m = \dim(Y) = 2l$.*

Proof: We use the well known fact that the determinant is a complete function for the class of polynomials having small formulae [11, 45]. Formally [45], if $f(X)$ had a formula of size l, there would exist a matrix M of size at most $2l$, whose each entry is either a constant or some entry x_{ij} of X, such that

$$f(X) = \det(M).$$

If we identify the entries x_{ij} with the entries of the submatrix $\phi(X)$ of Y, and multiply each constant entry in M with the variable y of Y not in $\phi(X)$, we get a possibly singular matrix whose each entry is a linear combination of the entries of

Y. In other words, we get a possibly singular linear transformation $A : F^{m^2} \to F^{m^2}$, where $m \leq 2l$, such that

$$f^\phi(Y) = \det(AY), \tag{3}$$

where Y is treated as an m^2-vector as above. But $GL_{m^2}(F)$ is dense in the space of all linear transformations from F^{m^2} to itself. So $f^\phi(Y)$ is in the closure of the $GL_{m^2}(F)$-orbit of $\det(Y)$ in V. The proposition follows from this. Q.E.D.

Thus we are led to the following instance of the orbit closure problem in Section 2:

The orbit closure problem Can $f(Y) = \mathrm{perm}^\phi(Y)$ lie in the projective orbit closure $\Delta[\det(Y)]$ of $g(Y) = \det(Y)$ in $P(V)$?

The orbit closure problem can in turn be reduced to an instance of the stabilizer problem in Section 2. Towards that end, let us compute the stabilizer R of $\det(Y)$ in $G = SL_{m^2}(F)$. It is known to be a reductive subgroup which consists of linear transformations in G of the form (thinking of Y as an $m \times m$ matrix):

$$Y \to AY^*B^{-1}, \tag{4}$$

where Y^* is either Y or Y^T, $A, B \in GL_m(F)$. The determinant is completely characterized by its stabilizer in the following sense.

Proposition 4.2 *Every form $h(Y) \in V$ stabilized by R is a multiple of* $\det(Y)$.

Proof: easy.

Now the preceding orbit-closure problem for the determinant can be reduced as in Section 2 to the following:

The stabilizer problem Can R occur as a stabilizer in the infinitesimal neighbourhood of the $SL_l(F)$-orbit of $\mathrm{perm}^\phi(Y)$ in V?

We conjecture that this cannot happen if $m = \mathrm{poly}(n)$. Formally:

Conjecture 4.3 *If $m = n^a$, where a is any fixed constant, then $\mathrm{perm}^\phi(Y)$ cannot lie in the projective orbit closure $\Delta[\det(Y)]$. More specifically, R cannot occur as a stabilizer in the infinitesimal neighbourhood of the $SL_l(F)$-orbit of $\mathrm{perm}^\phi(Y)$ in V.*

By Proposition 4.1, this would imply that the permanent does not have a formula of size polynomial in n. In Section 5 we shall suggest an approach to this conjecture based on the construction of so-called explicit obstructions. The conjecture is not equivalent to the the original lower bound question (cf. Section 4.2). But it is "almost" equivalent in the following sense.

Proposition 4.4 *Suppose F is the field of complex numbers. If Conjecture 4.3 were false then perm(X) can be approximated infinitesimally closely by a formula of quasi-polynomial $n^{O(\log n)}$ size (in the sense that the coefficients of the polynomial computed by this formula would be infinitesimally close to those of the permanent).*

Proof: If the conjecture were false, $\text{perm}^\phi(Y)$ lies in the closure of the SL_{m^2}-orbit of $\det(Y)$ in $P(V)$, with $m = \text{poly}(n)$. This means, as a point in V, it lies in the closure of the GL_{m^2}-orbit of $\det(Y)$. So it can be approximately infinitesimally closely by a function of the form $\det(\sigma^{-1}Y)$, for some $\sigma \in GL_{m^2}$. Let us ignore the terms in this approximation that involve variables other than y or entries of $X = \phi(X)$, which cannot occur in $\text{perm}^\phi(Y)$; i.e., think of setting these variables of Y to zero. What remains is still an approximation to $\text{perm}^\phi(Y)$. Setting y to one, we get an infinitesimally close approximation to $\text{perm}(X)$ by a function of the form $\det(W)$, where W is an $m^2 \times m^2$-matrix whose each entry is a (possibly nonhomogeneous) linear combination of the entries of X. But such a function $\det(W)$ has a formula of size $m^{O(\log m)} = n^{O(\log n)}$ [2]. Q.E.D.

But such an approximation is not expected in view of Valiant's completeness result for the permanent [44]. This provides a good support for the conjecture.

As we remarked in Section 2, the general orbit-closure or the stabilizer problem seems intractable if f and g were arbitrary forms. The situation is better here because they have the following nice properties: (1) g is stable and f is partially stable (Section 4.1), and (2) both have large, nontrivial stabilizers that capture their structure (Propositions 4.2, and 4.5).

We have already seen that $\det(Y)$ has a large characteristic stabilizer (Proposition 4.2). Now let us turn to $\text{perm}(X)$, where the dimension of X is k. Assume that the characteristic is not equal to two (otherwise the permanent is equal to the determinant). X can also be thought of as an n-vector, where $n = k^2$. Consider the natural action of $SL_n(F)$ on the projective space $P(W)$ of forms in X of degree k. Then the stabilizer K of $\text{perm}(X)$ in $SL_n(F)$ is generated [30] by linear transformations of the form (thinking of X as a $k \times k$ matrix):

$$X \to \lambda X \mu^{-1}, \tag{5}$$

where λ and μ are either diagonal or permutation matrices, and $k \geq 3$. The connected component K^0 of K is a $(2k-1)$-dimensional torus.

The following is an analogue of Proposition 4.2 that says that the permanent is completely characterized by its stabilizer.

Proposition 4.5 *Every form $h(X) \in W$ stabilized by K is a multiple of perm(X).*

The proof is easy.

4.1 Stability

Now we turn to the issue of stability.

Theorem 4.6 *The point* $\det(Y) \in P(V)$ *is stable with respect to the action of* $G = SL_{m^2}(F)$ *on* $P(V)$*, where* V *is the space of homogeneous forms of degree* m *in* Y*.*

Proof: The stabilizer R of $\det(Y)$ is not contained in any parabolic subgroup of G, because its representation over the space of $m \times m$ matrices given by eq.(4) is irreducible. So stability follows from the Kempf's criterion [23] (cf. Section 3). Q.E.D.

Remark: Stability of $\det(Y)$ is not needed in our approach to Conjecture 4.3 in Section 5; only the stability of $\mathrm{perm}(X)$ is needed. However, stability of $\det(Y)$ is needed in other problems where the determinant plays the role of f in the orbit closure problem; e.g. in the NC^1 vs. NC^2 problem [11].

Theorem 4.7 *The point perm*$(X) \in P(W)$*, where* $\dim(X) = k$*, is stable with respect to the action of* $SL_n(F)$ *on* $P(W)$*, where* $n = k^2$*.*

Proof: The stabilizer of $\mathrm{perm}(X)$ in $SL_n(F)$ is, again, not contained in a proper parabolic subgroup, because the space of $n \times n$ matrices is its irreducible representation. So the assertion again follows from Kempf's criterion (Section 3).

We shall also give another proof based on the property (5) of Kempf's one-parameter subgroups (cf. Section 3.3), because it illustrates in a simple setting an idea that will be used later (Sections 8) in proving the stability of another form. This second proof ignores the discrete part of the stabilizer.

Let $\mathcal{X}$ be the vector space whose coordinates are the entries x_{ij} of X. It is a representation of K^0 as given by eq.(5). Let E_{ij} be the one-dimensional vector space of all matrices with the only non-zero entry in the (i,j)-th spot. It is easy to check that $\mathcal{X}$ decomposes as a K^0-module into $\oplus_{i,j} E_{i,j}$ and that this representation is multiplicity-free; i.e., all characters are distinct. Now assume that $\mathrm{perm}(X)$ is not stable. Let S be an $SL_n(F)$-orbit of minimum dimension in the closure of the orbit of $\mathrm{perm}(X)$ in V. It must be closed; cf. Borel [4]. In fact, since $SL_n(F)$ is reductive, and hence geometrically reductive (cf. Haboush [17]), there is a unique closed orbit S in the closure of the orbit of $\mathrm{perm}(X)$ in V.

By Kempf's result, there would be a one-parameter subgroup $\lambda : F^* \to SL_n(F)$ which drives $\mathrm{perm}(X)$ into S, and which commutes with K^0. Schur's lemma forces the action of $\lambda(t)$ on $\mathcal{X}$ to be of the form:

$$e_{i,j} \to t^{\alpha_{i,j}} e_{i,j},$$

where $e_{ij} \in E_{ij}$ and α_{ij} is an integer. We also have the constraint that $\sum_{i,j} \alpha_{i,j} = 0$ because $\lambda(t) \in SL_n(F)$. The effect of $\lambda(t)$ on any monomial $m_\sigma = \prod_{i=1}^m x_{i,\sigma(i)}$ in the expansion of $\mathrm{perm}(X)$ is as follows:

$$m_\sigma \to t^{\sum \alpha_{i,\sigma(i)}} m_\sigma$$

Thus for $\lim_{t \to 0} \lambda(t)(m_\sigma)$ to exist, $\sum_{i=1}^k \alpha_{i,\sigma(i)}$ must be nonnegative for every permutation on $[k]$ (since there is no cancellation across monomials). This is possible only when each sum is actually zero, as can be seen using Hall's theorem that every regular bipartite graph is a disjoint union of perfect matchings. Thus if $\lim_{t\to 0}\lambda(t)(\mathrm{perm}(X))$ exists then it is $\mathrm{perm}(X)$ itself. In other words, $\mathrm{perm}(X)$ cannot be S-unstable for any G-invariant closed S. Q.E.D.

In contrast, the point $\mathrm{perm}^{\phi}(Y) \in P(V)$ is not stable with respect to the action of $G = SL_{m^2}(F)$. In fact, it belongs to the null cone; this can be shown using the Hilbert-Mumford criterion [37]. However:

Theorem 4.8 *The point* $\mathrm{perm}^{\phi}(Y) \in P(V)$ *is partially stable with respect to the action of* $G = SL_{m^2}(F)$*, with defect* $\delta = 1$.

Proof: If $h \in P(W)$ is stable then $h^{\phi} \in P(V)$ is partially stable with defect $\delta = 1$, arguing exactly as in Example 3.2: Specifically, X and Y are now matrices instead of vectors, and the homogenizing variable is y here instead of x_0. The parabolic group P consists of the linear transformation that transform the variables in Y to their linear combinations. The Levi subgroup $L \subseteq P$ consists of the linear transformations that transform the variables in Y and $X \setminus Y$ to the linear combinations of the variables in Y and $X \setminus Y$ respectively. The subgroup $L' \subseteq L$ consists of the linear transformation that transform y to its constant multiple, and the variables in Y and $X \setminus (Y \cup \{y\})$ to the linear combinations of the variables in Y and $X \setminus (Y \cup \{y\})$ respectively.

The partial stability of $\mathrm{perm}^{\phi}(Y)$ then follows from Theorem 4.7. Q.E.D.

The SL_{m^2}-orbit Z of $\mathrm{perm}^{\phi}(Y)$ in V is a fibre bundle over G/P, where P is similar to the one in eq.(1), and so G/P is a Grassmanian variety. The fibre is isomorphic to the SL_{n^2}-orbit W of $\mathrm{perm}(X)$ in W. The stabilizer H of $\mathrm{perm}^{\phi}(Y)$ is related to the stabilizer K of $\mathrm{perm}(X)$ just as in eq.(2). The dimension of Z is equal to $\dim(SL_{m^2}) - \dim(H)$, and that of the fibre W is $\dim(SL_{n^2}) - \dim(K)$. In our problem, m is not too large compared to n; in fact it is polynomial in n. Hence the dimension of the global orbit Z is also polynomial in the dimension of the fibre. The parameters δ and Δ in the definition of partial stability (Section 3) that measure deviation from complete stability are not too large: we have $\delta = 1$ and $\Delta = \mathrm{poly}(n)$. Roughly, this says that though $\mathrm{perm}^{\phi}(Y)$ belongs to the null cone, it is not "too unstable".

Remark: The role of the determinant in this section can also be played by the trace: specifically, the analogues of Proposition 4.1 and Theorem 4.6 hold for $\mathrm{trace}(Y^m)$ as well.

4.2 Exterior limit points

Here we shall indicate why Conjecture 4.3 is expected to be somewhat stronger than the original lower bound question for the formula size of the permanent–because $\Delta[\det(Y)]$ may contain points that do not have polynomial size formulae. Let $g = \det(Y)$. We call a point in $\Delta[g]$ an *interior limit point*, if it is of the form $\det(\sigma Y)$, for some *possibly singular* linear transformation σ. In other words, the set of interior points is the image of the map $\alpha : M_{m^2}(F) \to \Delta[g]$, where $M_{m^2}(F)$ is the set of $m^2 \times m^2$-matrices over F, and α maps a possibly singular matrix $\sigma \in M_{m^2}(F)$ to $\det(\sigma Y)$. This image is a constructible subset [18] of $\Delta[g]$–i.e., it is a finite disjoint union of locally closed subsets. But it is not expected to be equal to $\Delta[g]$. The points of $\Delta[g]$ not in this image will be called *exterior limit points*. Interior limit points have formulae of quasi-polynomial size (see the proof of Proposition 4.4), whereas exterior limit points need not.

We now describe some classes of conjectured exterior limit points in $\Delta[g]$. Formally proving that a given point is an exterior limit point amounts to showing that it does not have a formula of $\Omega(\text{poly}(n))$ size, because the determinant is complete for the class of polynomials with small formulae; cf. the proof of Proposition 4.1. This is a lower bound problem in itself, beyond our reach at present.

If f is an interior limit point of $\Delta[g]$, then $f = \det(\sigma Y)$, where σ is possibly singular. Since diagonalizable matrices are dense in $M_{m^2}(F)$, it follows that some conjugate f' of f is of the form $\beta(g')$, β is diagonal, and g' is some conjugate of g. Let $\beta(t)$ be the one-parameter subgroup obtained by replacing the zero diagonal entries of β by t. Then $\beta = \lim_{t\to 0} \beta(t)$. We shall now give an example of a point f in $\Delta[g]$, which is of the form $\lim_{t\to 0} \det(\alpha(t)Y)$ for some one parameter group $\alpha(t)$, but it is not expected to be of the form $\beta(g)$, where $\beta = \lim_{t\to 0} \beta(t)$ for some one parameter group $\beta(t)$.

Let G be a nonbipartite graph on m vertices whose edges have nonnegative integer weights. Consider its skew-symmetric Tutte matrix M whose (i, j)-th entry M_{ij}, $i < j$, is the variable y_{ij} if G has an edge joining vertices i and j, and zero otherwise. We have

$$\det(M) = \sum_{\pi} \text{sign}(\pi)\text{wt}(\pi),$$

where π ranges over all permutations of $[1, \dots, m]$, and $\text{wt}(\pi)$ is the product of the entries $M_{i\pi(i)}$. Define the cycle graph of π with nonzero weight to be the subgraph of G formed by the edges $(i, \pi(i))$. Then it is known and easy to see that

$$\det(M) = \sum_{\pi'} \text{sign}(\pi')\text{wt}(\pi'), \tag{6}$$

where π' ranges over only those permutations with nonzero weight whose cycle graphs contain only cycles of even length–the contributions of other permutations cancel out.

Construct a modified Tutte matrix $\tilde{M}$ by replacing every variable y_{ij}, if it indeed occurs in M, by $t^{w_{ij}} y_{ij}$, where t is an indeterminate parameter and w_{ij} is the weight of the edge joining the nodes i and j. In other words, $\tilde{M} = \beta(t)(M)$, where $\beta(t)$ is the one parameter subgroup that maps such y_{ij} to $t^{w_{ij}} y_{ij}$. Let Y be the vector of the variables y_{ij}. Then eq.(6) implies that

$$D(Y) = \det(\tilde{M}) = t^{2W} h(Y) + \text{higher order terms in } t,$$

where W is the weight of a minimum-weight perfect matching in G and

$$h(Y) = \sum_{\pi} \mathrm{sign}(\pi)\mathrm{wt}(\pi),$$

where π ranges over all permutations with nonzero weight whose cycle graphs can be decomposed as the union of two minimum-weight perfect matchings, not necessarily distinct. Clearly $D(Y) = \det(\tilde{M})$ belongs to the projective orbit closure $\Delta[\det(Y)]$ for every t. Projectively, $D(Y)$ is equal to $h(Y)$ plus higher order terms in t. Since,

$$h(Y) = \lim_{t \to 0} t^{-2W} D(Y),$$

in the projective space $P(V)$, we have

$$h(Y) = \lim_{t \to 0} D(Y) = \lim_{t \to 0} \det(\beta(t)M).$$

In other words, $h(Y)$ belongs to the projective closure $\Delta[\det(Y)]$; but notice that $\det(\beta M)$, $\beta = \lim_{t \to 0} \beta(t)$, is not well defined in the projective space $P(V)$. We conjecture that for every m there are weighted graphs G such that $h(Y)$ does not have a small polynomial size formula. This would imply that $h(Y)$ is an exterior point of $\Delta[\det(Y)]$, since the determinant is complete for the class of polynomials with small formulae.

For most weight assignments on the edges of G, there is a unique minimum-weight perfect matching–this follows from the so-called isolation lemma [35]–which means $h(Y)$ contains just one monomial, and hence surely has a small formula. What the conjecture says is that there exist pathological weighted graphs. Nonbipartiteness of G is essential here. Because if G is bipartite, then for any weight-assignment on G consider the subgraph G' formed by the union of all minimum-weight perfect matchings in G. It easily follows from Hall's theorem that the set of perfect matchings in G' is equal to the set of minimum-weight perfect matchings in G. So $h(Y)$ for G is equal to the determinant of the Tutte matrix for G', and thus has a small formula [2].

Though $h(Y)$ is an exterior limit point, it can be approximated infinitesimally closely by a formula of quasi-polynomial size (Proposition 4.4)–this is not expected for the permanent because of its completeness.

5 Explicit obstructions

Let V be a linear representation of a reductive group G, and $P(V)$ the associated projective space. A nonzero point $v \in V$ also corresponds to a point in $P(V)$. By abuse of notation, we denote the corresponding point in $P(V)$ by v as well. Suppose we are given two explicit points $f, g \in V$, and the goal is to show that that f is not in the closure of the G-orbit of g in the projective space $P(V)$. Conjecture 4.3 is a special case of this problem where V is the space of homogeneous forms of degree m in the entries of an $m \times m$ variable matrix Y, $f = \text{perm}^{\phi}(Y)$, $g = \det(Y)$, and $G = SL(Y)$. Here f lies in the closure of the G-orbit of g in $P(V)$ iff it lies, considered as a point in V, the closure of the $GL(Y)$-orbit of g in V.

We call a mathematical structure an *obstruction* for the pair (f, g) if its existence acts as a guarantee that f does not lie in the closure $\Delta[g]$ of the G-orbit of g in the projective space $P(V)$.

5.1 Polynomials as obstructions

Let Z denote the closure of the G-orbit of g in $P(V)$. Let $I(Z)$ denote the ideal of polynomial functions vanishing on Z. Any polynomial in $I(Z)$ that does not vanish at f is an obstruction for the pair (f, g), because it vanishes on the orbit of g but not on the orbit of f. We say that it *separates* the two orbits. If f does not lie in Z then "almost every" polynomial in $I(Z)$ is an obstruction. In the setting of Conjecture 4.3, where $g = \det(Y)$ and $f = \text{perm}^{\phi}(Y)$, it is thus expected that almost every polynomial in $I(Z)$ will be an obstruction for the pair (f, g) when m is polynomial in n. To prove Conjecture 4.3 it suffices to construct an obstruction for every such m and n. Even though we know–assuming Conjecture 4.3–that such obstructions exist in plenty, the problem is such that one has prove their existence by constructing them more or less *explicitly*.

One approach to constructing an explicit polynomial obstruction is to construct an explicit set of generators of $I(Z)$, and then isolate a generator that does not vanish at f. This problem is addressed in Part II [34], where we give a sufficiently explicit set Σ_g of generators for the ideal of the G-orbit of g in the affine space V, assuming that g is stable and has a "nice" stabilizer, and conjecture that this set actually generates $I(Z)$ for g's that arise in the context of the lower bound problems in this paper. The generators in Σ_g are quite complex, and verifying that one of them does not vanish at f is a difficult task. But they have nice representation-theoretic properties. In part II we develop in detail an approach to the orbit closure problem via explicit construction of obstructions based on these these properties.

Here we only wish to indicate how representation theory comes into play by briefly considering explicit construction of obstructions in the context of the related stabilizer problem. The scheme for explicit construction suggested here depends on the representation-theoretic properties of the stabilizers of f and g, and serves as a starting point for the work in Part II.

Specifically, the scheme will exploit the fact that, in our context, f is partially stable (cf. Theorem 4.8) and both f and g have nontrivial reductive stabilizers that capture their structure (Propositions 4.5 and 4.2); g may or may not be stable. The obstructions that one hopes to construct this way are not polynomials, but rather representations having certain properties (cf. Theorems 5.1 and 5.2).

5.2 Representations as obstructions

In this section, we shall present our scheme when f is (completely) stable with respect to the G-action. In Section 5.5 we shall suggest an extension of this scheme to the case when f is only partially stable as in Conjecture 4.3. There we shall also point out the main problems that are encountered in such an extension.

When f is stable, obstructions based on representations can be constructed using the following immediate consequence of Luna's slice theorem.

Theorem 5.1 *(characteristic zero) Let $H \subseteq G$ be the stabilizer of $f \in V$, and $Q \subseteq G$ the stabilizer of $g \in V$. Suppose f is stable with respect to the action of G. Then a (nonzero) representation W of G is an obstruction for the pair (f, g) if W contains a trivial H-submodule but not a trivial Q-submodule. In other words, if such a W exists then f cannot lie in the closure of the G-orbit of g in the projective space $P(V)$. More generally, W is an obstruction for (f, g) if the multiplicity of the trivial H-representation within W exceeds that of the trivial Q-representation.*

Proof: Since f is stable, by Luna's slice theorem (Section 3), there is a slice N at f, G-invariant neighbourhood U of the orbit Gf, and a map $\tau : G \times_H N \to U$ that is excellent. This implies that (cf. [38]) that the stabilizer of any point in U is a conjugate of a subgroup of H.

Suppose, to the contrary, that f, considered as a point in $P(V)$, lies in the closure of the G-orbit of g in $P(V)$. Then the neighbourhood U must contain a point p whose some multiple p' lies in the G-orbit of g in the affine space V. The stabilizer of p is equal to the stabilizer of p', which is a conjugate of the stabilizer Q of g. Since some conjugate of the stabilizer of p is contained in H, it follows that some conjugate of Q is contained in H. Hence the multiplicity of the trivial Q-representation in any G-module must exceed that of the trivial H-representation.

So if there exists a G-module W such that the multiplicity of the trivial H-representation within W exceeds that of the trivial Q-representation then f cannot lie in the closure of the orbit of g in $P(V)$. Q.E.D.

Second proof: We now prove the first criterion of Theorem 5.1 (though not the second) without using Luna's slice theorem, but using Peter-Weyl theorem instead; the idea in this proof will be useful later in Section 5.5. We shall denote the algebraically closed base field of characteristic zero by k. Because Gf is a closed affine subvariety of V and the characteristic is zero,

$$k[V] = k[Gf] \oplus I, \tag{7}$$

as a G-module, where I is the G-invariant ideal of Gf. Thus $k[Gf] = k[G/H]$ can always be embedded equivariantly within $k[V]$. Now we shall determine an explicit G-module decomposition of the ring $k[G/H]$. First, consider $k[G]$, the ring of regular functions on G, as a $G \times G$ module by letting one copy of G act on it by left translations and the other by right translations. Then $k[G]$ has the following decomposition as a $G \times G$ module [6]:

$$k[G] = \sum_P P \otimes \bar{P}, \tag{8}$$

where P runs over all finite dimensional G-modules, and $\bar{P}$ is its dual. This follows by applying the Peter-Weyl theorem to the Hilbert space $L_2(K)$ of L_2-functions on a maximal compact subgroup K of G, and using the fact that the K-finite vectors in $L_2(K)$ correspond to the regular polynomial functions on G; cf. [6, 22]. Since f is stable, H is reductive [32]. Then $k[G/H]$ is simply the ring of $k[G]^H$ of (right) H-invariants within $k[G]$ [37]. So the decomposition in eq.(8) implies that

$$k[Gf] = k[G/H] = \sum_P P \otimes \bar{P}^H, \tag{9}$$

where P runs over all finite dimensional G-modules, $\bar{P}^H$ is the sum of all trivial H-submodules of $\bar{P}$, and G acts on the first factor of $P \otimes \bar{P}^H$.

Let W be an irreducible representation containing a trivial H-module but not a trivial Q-module. It can be equivariantly embedded within $k[G/H] = k[Gf]$, by eq.(9), and hence within $k[V]$ by eq.(7). Thus each element in W corresponds to a polynomial function on V. Now suppose to the contrary that the G-orbit of f in $P(V)$ lies in the projective closure $\Delta[g]$ of the G-orbit of g in $P(V)$. Not all functions in W can vanish on the affine cone C in V corresponding to the G-orbit of g in $P(V)$, because then they will also vanish on the affine cone in V corresponding to $\Delta[g]$, and hence on the G-orbit of f in V, since $f \in \Delta[g]$. But this is impossible since W is a nonzero submodule of $k[Gf]$. So W contains at least one function ϕ which does not vanish at some point of the cone C, say p. The stabilizer $Q_p \subseteq G$ of p is conjugate to the stabilizer Q of g. Let $W_\phi \subseteq W$ be the Q_p-submodule generated by ϕ. Consider the evaluation map from W_ϕ to k, which maps any function in W_ϕ to its value at p. Since ϕ is nonzero at p, this evaluation map is nonzero. It is also equivariant with respect to the Q_p-action. Since k is the trivial Q_p-representation, and the characteristic is zero, W_ϕ, as a Q_p-module, contains a trivial Q_p-submodule, and so too W. Since, W is a G-module and Q_p is a conjugate of Q in G, it follows that W must contain a trivial Q-module; a contradiction. Q.E.D.

To use the obstruction criterion in this theorem effectively, one needs an explicit formula for the dimension of the trivial H (or Q) representation within a G-module W. This is an instance of the subgroup restriction problem, which will also arise later (Section 5.5).

5.2.1 Example 1

Let X be the variable matrix of size d, $V = \mathrm{Sym}^d(X)$, the space of forms in X of degree d, and $G = SL(X) = SL_{d^2}(\mathbb{C})$. Let $f = \mathrm{perm}(X)$ and $g = \det(X)$, with stabilizers H and Q respectively. We want to show that f is not in the closure of the G-orbit of g in $P(V)$.

In the present case, this can be shown by just dimension considerations. Indeed if f were in the closure of the G-orbit of g in $P(V)$, then by Luna's slice theorem some conjugate of Q will be contained within H. But since the dimension of Q exceeds that of H this is not possible.

However the dimension-based criterion seems too crude to be generalized to the partially stable case. Because when $(f, g) = (\mathrm{perm}^{\phi}(Y), \det(Y))$ as in Conjecture 4.3 the dimension of the stabilizer of f is far greater than that of g when m is only mildly larger than n. On the other hand, representations as obstructions, as in Theorem 5.1, seem suited even for the partially stable case (Section 5.5). With that in mind, we construct an obstruction for the pair (f, g) using Theorem 5.1, which is applicable since $f = \mathrm{perm}(X)$ is stable with respect to the action of $G = SL(X)$ (Theorem 4.7).

Let $W = Sym^r(X)$ be the G-module of homogeneous polynomials of degree r in the entries of X where r is a multiple ad of d for some integer $a > 1$. The multiplicity of the trivial H-representation in W exceeds one, because $\mathrm{perm}(X)^a \in W$ is fixed by H, and so too the permanent of the matrix obtained by replacing each entry of X by its ath power. As a Q-module [7]

$$W = \sum_{\lambda} S_{\lambda} \otimes \bar{S}_{\lambda},$$

where λ ranges over all Young diagrams of size r with at most d rows or columns, S_{λ} denotes the corresponding Weyl module, and $\bar{S}_{\lambda}$ its dual. Here the elements in Q of the form (A, B), $A, B \in GL_d$, $\det(AB^{-1}) = 1$, act on $S_{\lambda} \otimes \bar{S}_{\lambda}$ in the obvious manner, and the transpositions in Q interchange the two factors. When $r = ad$, $S_{\lambda} \otimes \bar{S}_{\lambda}$ is trivial as a Q-module precisely when λ is rectangular with d rows and a columns. Hence the multiplicity of the trivial Q-module in W is precisely one–this trivial module is generated by $\det^a(X)$. By Theorem 5.1, any $\mathrm{Sym}^r(X)$, $r = ad$, $a > 1$, is an obstruction for the pair (f, g).

The preceding obstructions were constructed using the second multiplicity-based criterion of Theorem 5.1. It would be interesting to construct obstructions using the first criterion based on absence or presence of trivial representations. Specifically, it would be interesting to construct a representation W, for each d, such that W contains a trivial H-representation but not a trivial Q-representation. For this one needs to know how a given G-representation W splits when considered as a Q-module or as an H-module. Q contains a copy of $SL_d \otimes SL_d$, and H contains a copy of $S_d \times S_d$, where S_d is the symmetric group. Hence this problem is closely linked to the so-called plethysm problem of representation theory. This asks for explicit decomposition of $S_{\lambda}(V \otimes W)$ as a $GL(V) \times GL(W)$-module and that of $S_{\lambda}(S_{\mu}V)$ as a $GL(V)$-module, where S_{λ} is the Schur-functor [14].

5.3 Importance of stability

The stability of f is crucial for Theorem 5.1 to hold. Otherwise, the first proof does not apply because Luna's slice theorem may fail when f is not stable even if the stabilizer H of f is reductive (cf. Example 2 below). The second proof also fails when f is not stable because it may not be possible to extend the functions in the module $W \subseteq k[Gf] = k[G/H]$ there equivariantly to the whole of V, since Gf is not closed. It may happen that $k[V]$ contains a G-submodule W that contains a trivial H-representation but not a trivial Q-representation. That still does not guarantee that W is an obstruction, because all functions in W may vanish identically on the the affine cones in V corresponding to the G-orbits of g and f in $P(V)$. In other words, an abstract representation W need not be an obstruction even if it contains a trivial H-module but no trivial Q-module. But such a W becomes an obstruction if it has a concrete realization as a G-submodule of $k[V]$ that does not vanish when restricted to the G-orbit of f in V. When this happens (as it does when f is stable), a slight modification of the argument in the second proof of Theorem 5.1 shows that it contains a function that separates the orbits of f and g–i.e., vanishes on the orbit of g but not on the orbit of f– though such a separating function is not explicitly constructed. Thus representation-based obstructions provide an indirect way of constructing–or more precisely, guaranteeing existence of–separating functional obstructions as in Section 5.1.

5.3.1 Example 2

We give a simple example that demonstrates why stability of f is crucial for Theorem 5.1 to hold. Let V be the space of forms of total degree 3 in variables x and y, and $G = SL_2$. Let $f = x^2y$, and $g = x^3 + y^3$ [38]. Then it is easy to show that f lies in the projective closure $\Delta[g]$ of the G-orbit of g in $P(V)$ (the G-orbit of g in $P(V)$ is three dimensional, and that of f is two dimensional, since the G-stabilizer of the line in V containing f and the origin is one dimensional). The stabilizer $H \subseteq G$ of f for the G-action on V is trivial, hence reductive, but f is not stable for the G-action. The stabilizer Q of g is equal to $\{\mathrm{diag}(\epsilon, \epsilon^{-1}) | \epsilon^3 = 1\}$. Since H is trivial, any G-representation contains a trivial H-representation. Now there are G-representations which do not contain a trivial Q-representation; eg. the standard representation. Yet none of these is an obstruction for the pair (f, g), since we already know that f lies in $\Delta[g]$.

5.4 Positive characteristic

Theorem 5.2 *With the same terminology as in Theorem 5.1, suppose, in addition, that the orbit map is separable, and that the ambient space V has an H-module decomposition in which the tangent space of the orbit Gf at f splits off as a direct summand. Then the obstruction criterion given there holds.*

Here, by multiplicity of the trivial H or Q-representation in W, we now mean the total dimension of the sum of such trivial submodules of W. This need not be equal to the number trivial quotients in a Jordan-series filtration of W, since, in positive characteristic, W need not decompose into irreducible Q or H-submodules.

Proof: Luna's slice theorem holds in arbitrary characteristic if the orbit map is separable and V has an H-module decomposition as above [1]. So the proof of Theorem 5.1 holds. Q.E.D.

This points out the issues that arise in positive characteristic because of the lack of linear reductivity.

5.5 Partially stable case

In the preceding section we gave a method (Theorem 5.1 and 5.2) for constructing representation theoretic obstructions for the pair (f, g), when f is stable. We now suggest an extension of this approach to produce obstructions when f is only partially stable. The setting in Conjecture 4.3 falls in this category. Let H denote the stabilizer of f. Let Q denote the stabilizer of g; we assume that it is reductive.

Let α denote the embedding of the orbit $Z = Gf$ in the ambient space V. Let $\mathcal{O}_V$ denote the sheaf of *analytic* functions on V. Let $\mathcal{O}_V^Z = \alpha^{-1}\mathcal{O}_V$ denote the inverse image sheaf on Z (cf. Hartshorne [18]). The stalk of $\mathcal{O}_V^Z$ at any point $z \in Z$ coincides with the stalk of $\mathcal{O}_V$ at z. The global sections, i.e., the elements of $H^0(Z, \mathcal{O}_V^Z)$, of this sheaf correspond to germs of analytic functions that are defined in some analytic open neighbourhood of Z in V. The space $H^0(Z, \mathcal{O}_V^Z)$ is a G-module in a natural way.

We denote the algebraically closed base field by k in this section.

Proposition 5.3 *Assume that the characteristic is zero. Let W be any finite dimensional G-module in $H^0(Z, \mathcal{O}_V^Z)$ that does not vanish identically under the (natural) restriction to Z. Then W is an obstruction for the pair (f, g) if it does not contain a trivial Q-representation.*

Proof: Fix such a G-module W that does not vanish when restricted to Z. Since the G-action on Z is transitive, there exists a global section $\phi \in W$ that does not vanish at f when restricted to Z. Let U_ϕ be an analytic neighbourhood of Z in V on which ϕ is well defined. Suppose to the contrary that the closure of the G-orbit of g in $P(V)$ contains f. Then U_ϕ contains a point p arbitrarily close to f whose multiple $\hat{p}$ lies in the G-orbit of g in the affine space V. Choose p close enough to f so that ϕ does not vanish at p. The stabilizer of p is the same as that of $\hat{p}$, which in turn is equal to the conjugate Q^α of Q for some $\alpha \in G$. Consider any $\beta \in Q^\alpha$. The function $\beta(\phi)$ is defined in the neighbourhood $\beta(U_\phi)$ of Z which contains p, since it is fixed by β. Thus all functions in the Q^α-submodule $W_\phi \subseteq W$ generated by ϕ are well defined at p. Consider the evaluation map from W_ϕ to k, which maps any function in W_ϕ to its value at p. Since ϕ is nonzero at p, this evaluation map is nonzero. It is also equivariant with respect to the Q^α-action. Since k is a trivial Q-representation, and the characteristic is zero, W_ϕ contains a trivial Q^α-module in its decomposition into irreducible Q^α-modules, and so too W. Since, W is a G-module and $\alpha \in G$, its decomposition into irreducible Q-modules also contains a trivial Q-module. Q.E.D.

To construct an obstruction using this proposition, one needs to calculate explicitly an interesting family of finite dimensional G-submodules within $H^0(Z, \mathcal{O}_V^Z)$. When f is stable, one gets such a family as follows.

Theorem 5.4 *Suppose $k = \mathbb{C}$, and f is stable. Then every irreducible G-representation W that contains a trivial representation as an H-module occurs in $H^0(Z, \mathcal{O}_V^Z)$ (with multiplicity at least equal to the multiplicity of the trivial H-representation in W).*

Proof: By Luna's slice theorem (Section 3), the orbit $Gf \cong G/H$ has an analytic G-invariant neighbourhood of the form $G \times_H N$, where N is an analytic slice at f. This implies that any function $\phi \in k[Gf]$ can be extended canonically to a function $\bar{\phi}$ on this neighbourhood so that it is constant on the slice at any point of the orbit. This extension is G-equivariant. So we get a canonical G-equivariant embedding of $k[Gf] = k[G/H]$ in $H^0(Z, \mathcal{O}_V^Z)$. The Peter-Weyl theorem (cf. eq.(9)) gives the following explicit G-module decomposition of the ring $k[G/H]$:

$$k[Gf] = k[G/H] = \sum_P P \otimes \bar{P}^H, \tag{10}$$

where P runs over all finite dimensional G-modules, $\bar{P}^H$ is the sum of all trivial H-submodules of $\bar{P}$, and G acts on the first factor of $P \otimes \bar{P}^H$. This means every G-representation that contains a trivial representation as an H-module occurs in $k[Gf]$, and hence in $H^0(Z, \mathcal{O}_V^Z)$. Q.E.D.

In conjunction with Proposition 5.3, Theorem 5.4 again implies the first criterion of Theorem 5.1. The proof of Theorem 5.4 fails when f is not stable, because then it may not be possible to embed $W \subseteq k[Gf] = k[G/H]$ equivariantly within $H^0(Z, \mathcal{O}_V^Z)$ since Luna's slice theorem can fail (cf. Example 2), and so Proposition 5.3 may not be applied to such a W even if it contains a trivial H-module (also see Section 5.3).

To construct an obstruction using Proposition 5.3 when f is only partially stable, one needs to calculate explicitly an interesting class of finite dimensional G-modules that occur within $H^0(Z, \mathcal{O}_V^Z)$, as we did in the stable case (Theorem 5.4). Just as in the completely stable case, we expect that partial stability of f should play a crucial role in this calculation (cf. Theorem 5.5).

When f is stable, the orbit Z is affine, and the stabilizer of f is reductive [32]. In this case, it can be shown that the higher cohomology groups $H^i(Z, \mathcal{O}_V^Z)$, $i > 0$, all vanish; we omit the proof. This statement is analogous to the Cartan-Serre-Grothendick vanishing theorem for the vanishing of higher cohomology of a (quasi-)coherent algebraic sheaf on an affine variety (or more generally on a formal scheme) [18], and the Cartan-Oka vanishing theorem for coherent analytic sheaves on Stein spaces [43]. However it does not formally follow from the Cartan-Oka vanishing theorem, because, in general, $\mathcal{O}_V^Z$ is neither coherent in the sense of Cartan and Oka, nor an inverse limit of coherent sheaves.

When f is only partially stable, the higher cohomology groups need not vanish, and they can be expected to play a role in the construction of representation-theoretic obstructions. When f is stable, Theorem 5.4 and Proposition 5.3 together say that $H^0(Z, \mathcal{O}_V^Z)$ contains information in the form of finite dimensional G-submodules that puts fairly strong constraints on the stabilizers that can occur in the infinitesimal neighbourhood of the orbit of f in V. Analogously, we expect $H^i(Z, \mathcal{O}_V^Z)$, $i \geq 0$, to contain similar information, when f is partially stable, that will constrain the (reductive) stabilizers that can occur in the infinitesimal neighbourhood of the orbit of f in V. With this in mind, our approach for constructing representation-theoretic obstructions, when f is partially stable, can be formulated as follows.

1. Prove a generalization of the obstruction-based criterion in Proposition 5.3 to higher cohomology groups.
2. Explicitly calculate an interesting class of finite dimensional G-submodules that occur within $H^0(Z, \mathcal{O}_V^Z)$, and more generally, in $H^i(Z, \mathcal{O}_V^Z)$, generalizing the method in the proof of Theorem 5.4 for the stable case.
3. Explicitly isolate a G-submodule in this class that satisfies the obstruction criterion for the pair (f, g). Thus, in the setting of Conjecture 4.3, one hopes to get an explicit obstruction this way when m is polynomial in n, thereby proving that the permanent has no polynomial size formula.

Now we elaborate on the main problems that arise in these three steps.

Generalization of the obstruction criterion

The proof of Proposition 5.3 uses evaluation of functions in $H^0(Z, \mathcal{O}_V^Z)$. As it is, this does not work for higher cohomology, since we do not have an appropriate notion of evaluation for elements of higher cohomology representations.

Explicit computation of $H^i(Z, \mathcal{O}_V^Z)$

Here partial stability is expected to be crucial–just as stability was crucial in the computation of $H^0(Z, \mathcal{O}_V^Z)$ based on Luna's slice theorem and Perter-Weyl theorem (Theorem 5.4) when f was stable. Indeed one important feature of our definition of partial stability is the following:

Theorem 5.5 *Let $\mathcal{F}$ be a coherent (algebraic or analytic) sheaf over the orbit Z. Let $\mathcal{F}^*$ be its direct image over G/P with respect to the projection of the fibre bundle $Z = Gf \to G/P$. Then $H^i(Z, \mathcal{F}) = H^i(G/P, \mathcal{F}^*)$.*

Proof: By the definition of partial stability, the fibre of the bundle $Z \to G/P$ is affine, since it is isomorphic to L/L', where L is a Levi subgroup of P and L' is a stabilizer of f within L that is reductive. Since higher cohomology of a coherent algebraic (or analytic) sheaf over an affine variety vanishes, by the Cartan-Serre theorem [18] (resp. the Cartan-Oka theorem [43]), the result follows by the Leray spectral sequence. Q.E.D.

This indicates why the group action in the vicinity of the orbit of a partially stable point in the null cone ought to have more structure than in the vicinity of an arbitrary point in the null cone. The sheaf $\mathcal{O}_V^Z$ is not coherent. However, the idea in the preceding proof can be pushed towards computation of some interesting families of finite dimensional G-submodules in $H^i(Z, \mathcal{O}_V^Z)$. This will be reported elsewhere.

Construction of an explicit G-module satisfying an obstruction criterion

This step gives rise to the *subgroup restriction problem* of calculating an explicit decomposition of a G-module W when the action is restricted to a reductive subgroup of G. We have already seen this happen, when f is completely stable (cf. Theorem 5.1 and the remark following its proof), the same is expected in the partially stable case. As pointed out in Example 5.2.1, the specific version of the subgroup restriction problem that arises in the context of the determinant vs. permanent problem is the plethysm problem of representation theory, which has no satisfactory solution so far. An explicit combinatorial formula for multiplicities of irreducible Weyl modules within $S_\lambda(V \otimes W)$ and $S_\lambda(S_\mu V)$, based on Weyl's character formula, is known [21]. But this involves alternating signs, and so it is difficult ascertain from the formula whether the specific multiplicity is zero or nonzero. What is desirable is a formula for decomposition that does not involve signs, analogous to the Littlewood-Richardson rule for the decomposition of the tensor product of two Weyl modules. Such a decomposition for the plethysm problem is so far not known.

It would be interesting to use the scheme in this section (using, say, the obstruction criterion in Proposition 5.3) on computer to come up with explicit representation-based obstructions for the pair (f, g) in Conjecture 4.3 for small values of m and n; say $m = 3$ and $n = 2$. Unfortunately, even for very small values, this computational problem seems quite hard (the plethysm problem being one of the computational bottlenecks–for $m = 3$, one requires explicit decomposition of $S_\lambda(V \otimes V)$, where $\dim(V) = 9$).

6 Circuit size

In Section 5 we formulated an approach to a well-behaved instance of the orbit-closure (and the stabilizer) problem for a pair of forms (f, g) with nontrivial characteristic stabilizers, wherein f is partially stable. In Section 4 we reduced the lower bound problem for the formula-size of the permanent to such an instance. Next we shall do the same for certain lower bound questions concerning circuit-size, which will include (Section 7) the usual P vs. NP question (the stronger nonuniform[4] version) and its arithmetic analogue over an algebraically closed field.

Let $f(X)$ be again any fixed form in the variable vector X of dimension n over an algebraically closed field F. Now we wish to know if it has an arithmetic circuit over F of size l. An arithmetic circuit is a directed acyclic graph. Input variables and constants represent nodes with indegree zero. The remaining nodes are labeled as addition or multiplication nodes. All nodes function as per their labels and the result is output at the unique node with outdegree zero. The size of the circuit is the size of the graph.

The role of the determinant function in Section 4 would now be played by a certain function $H(Y)$, which would be complete for the class of polynomials having small circuit size. To define it, consider a generic arithmetic circuit of depth k and width m. It consists of $k+1$ levels of nodes, numbered 0 to k, each level containing m nodes, except the root level zero, which contains a single output node. Each node in the level $i < k$ is connected to every node in level $i+1$. Each node u in the input level k is labeled with the variable y^u; the function computed by this node is defined to be y^u. The function $h(u)$ computed by a node u in level $i < k$ is defined to be $\sum_{v,w} y^u_{v,w} h(v)h(w)$, where v and w range over nodes in level $i+1$ and each $y^u_{v,w}$ is an indeterminate. Let Y be the vector of the variables y^us at the input level k and the variables $y^u_{v,w}$s. Let $H(Y)$ be the function computed at the root level zero. It is a homogeneous form in Y with total degree exponential in k.

Any arithmetic circuit of size r can be obtained from the generic circuit of depth $k \leq r$ and width $m \leq r$ by specializing the indeterminates $y^u_{v,w}$ to some constants and the indeterminates y^us at the input level to the input parameters or constants. Hence $H(Y)$ is complete for the class of polynomials with small circuits. This can be used to prove the following analogue (Proposition 6.1) of Proposition 4.1. Let V be the space of homogeneous forms in Y with degree equal to that of $H(Y)$. Let $l = O(r^2)$ be the size of Y that corresponds to the generic circuit with depth and width r. We can define, as in Section 4, the function ϕ for embedding in V the forms of lower degree in the smaller variable vector X of dimension $n < m$: specifically identify X with some of the input variables in Y and use some remaining variable as a homogenizing parameter. Let $G = SL_l(F)$. Then:

Proposition 6.1 *If $f(X)$ has an arithmetic circuit of size r then $f^\phi(Y)$ lies in the (projective) Zariski-closure $\Delta[H(Y)]$ of the G-orbit of $H(Y)$ in $P(V)$, where $l = O(r^2)$.*

Proof: The proof is similar to that of Proposition 4.1. The role of determinant in that proof is now played by $H(Y)$, which is complete for the class of functions with

[4] In the nonuniform version, the computational circuits for various input sizes n need not be correlated, whereas in the uniform version, correlation is enforced by requiring a $\log(n)$-space machine that can recognize the circuits for all n.

small circuits. Q.E.D.

The form $H(Y)$ turns out to be good like the determinant function because it has a large nontrivial stabilizer that in a sense characterizes it. Indeed, any automorphism of the generic circuit that fixes the nodes in the input level k gives rise to an element of $SL_l(F)$ that stabilizes $H(Y)$. Quite likely, the stabilizer of $H(Y)$ is precisely the discrete group Q of such automorphisms. It contains k copies of the symmetric group S_m, where k is the depth of the generic circuit and m is its width. The total degree of $H(Y)$ in the variables $y^u_{v,w}$ that occur at level i is $d_i = 2^i$, and its total degree in the input variables y_us is $d_{k+1} = 2^{k+1}$. So $H(Y)$ has multidegree $(d_1, \ldots, d_{k+1})$ in these groups of variables. Let $M(Y)$ be the set of monomials over Y with total degree equal to that of $H(Y)$, i.e., with degree $d_1 + \cdots + d_{k+1}$. It has a natural Q-action induced by that on the generic circuit. Order the monomials in $M(Y)$ lexicographically. We call a monomial maximal if it is the largest monomial in its Q-orbit. Let $M_m(Y) \subseteq M(Y)$ be the subset of maximal monomials. Given any monomial α, let $\hat{\alpha}$ denote it symmetrization $\sum_{\sigma \in Q} \sigma(\alpha)$. Let V be the space of homogeneous forms in Y of total degree equal to that of $H(Y)$. The following is easy to prove.

Proposition 6.2 *The set of forms $\hat{\alpha}$, $\alpha \in M_m(Y)$, is a basis of the space V^Q of Q-invariants in V.*

Unlike the determinant (cf.Proposition 4.2), $H(Y)$ is not a unique form stabilized by Q. But all forms stabilized by Q are strongly linked to the generic circuit as in Proposition 6.2. In that sense, Q captures the structure of $H(Y)$.

Via Proposition 6.1 we can reduce our lower bound problem for the circuit size of $f(X)$ to an instance of the orbit closure problem (Section 2) with $H(Y)$ and $f^\phi(Y)$ playing the role of $g(Y)$ and $f(Y)$ there. Since $H(Y)$ has a nice stabilizer that captures its structure (Proposition 6.2), this problem in turn can be reduced to an instance of the stabilizer problem (Section 2), where we let D be equal to the stabilizer of $H(Y)$. For these problems to be well behaved both $H(Y)$ and $f(X)$ should be nice functions like the determinant and permanent. We have already seen that $H(Y)$ has a nontrivial, characteristic stabilizer. The form $f(X)$ should also be nice. This will indeed be the case for the form that arises in the context of the P vs. NP question, as we shall see next.

7 P vs. NP

Before turning to the $P \neq NP$ conjecture, we shall treat its arithmetic version first. Let X be a variable vector (or matrix) of dimension n. Let $f(X)$ be a polynomial–which is defined in some uniform way for each n–such that, when reduced modulo 2, it gives an NP-complete function in the usual sense. If $P \neq NP$ (the nonuniform version), then $f(X)$, modulo 2, cannot be computed by a boolean circuit of poly(n)

size. This implies that $f(X)$ cannot be computed by an integral (arithmetic) circuit of poly(n) size either; otherwise this circuit can be reduced modulo 2 to obtain a boolean circuit of poly(n) size. By arithmetic (integral) version (implication) of the $P \neq NP$ conjecture (nonuniform version), we mean the problem of showing that such $f(X)$ does not have an integral circuit of poly(n) size. Since it is an implication of the usual, nonuniform $P \neq NP$ conjecture, and at the same time does not involve problems of positive characteristic, it is natural to address this arithmetic implication first. In turn, it suffices to consider the complex version of $P \neq NP$ conjecture, where the problem is to show that $f(X)$ does not have a circuit of poly(n) size defined over $\mathbb{C}$. The complex version trivially implies the arithmetic one, since every integral circuit is also a complex circuit. Hence, the version of the $P \neq NP$ conjecture over $\mathbb{C}$, or more generally, over an arbitrary algebraically closed field F should be addressed first.

The above considerations also apply even if $f(X)$ is not NP-complete, but is such that its reduction, modulo a small prime, is in NP, but not expected to be in P. The function $E(X)$, defined below, is one such example. The arithmetic version of the lower bound problem under consideration is again to show that $f(X)$ cannot be computed by a complex, and hence an integral, circuit of polynomial size.

We begin by instantiating the form $f(X)$ in the preceding section to a certain form $E(X)$ that corresponds to a hard function in NP. Consider a system of homogeneous integral polynomial equations of some constant degree $k \geq 3$ in n variables:

$$g_i(y_1, \dots, y_n) = 0, \quad 1 \leq i \leq m, \tag{11}$$

where we assume that each g_i is a product of k homogeneous linear forms. The problem of finding if this system has a nonzero solution over F_p, where $p \leq 2^{\mathrm{poly}(n)}$, is in NP. It is also NP-complete, even for $p = 2$ and $k = 3$ (reduction from 3-CNF[5]).

We shall consider the special case of this problem when $m = n$. It is expected to be hard; i.e., not expected to be in P for large enough p's. The resultant of the sytem (11), when $m = n$, has a very simple form. Take m-dimensional vector variables $\{X_i^j | 1 \leq j \leq k,\ 1 \leq i \leq m\}$, where, for any i, X_i^j, $1 \leq j \leq k$, are the coefficient vectors of the k forms whose product is g_i. For simplicity, say, $k = 3$. Let X be the matrix whose columns consist of these km vectors. For any function $\sigma : \{1, \dots, m\} \to \{1, \dots, k\}$, let $\det_\sigma(X)$ denote the determinant of the matrix whose i-th column is $X_i^{\sigma(i)}$. Define $E(X) = \prod_\sigma \det_\sigma(X)$ where σ ranges over all such functions. This is well defined over any field. The polynomial system (11), when $m = n$, has a nonzero solution over F_p iff $E(X)$ is zero modulo F_p. Let F be either $\mathbb{C}$ or the algebraic closure of F_p. The arithmetic problem then is to show that $E(X)$ does not have an arithmetic circuit of poly(n) size over F. Let $n = km^2$ be

[5]Consider a 3-CNF in l variables $u_1, \dots, u_l$. Let z be a distinct homogenizing variable. For each variable u_i, we include in our system the equation $u_i(z - u_i) = 0$. For each 3-clause ϕ, we include the equation $g_\phi = 0$, where g_ϕ is obtained from ϕ by replacing $\wedge$, i.e., "and" by multiplication, and a literal u_i ($\bar{u}_i$) by u_i (resp. $z - u_i$). The resulting system has a nonzero solution iff the 3-CNF is satisfiable.

the total number entries in X. The stabilizer $K = \mathrm{stab}(E(X))$ of $E(X)$ in $SL_n(F)$ is nice:

Proposition 7.1 *The connected component K_0 of the identity in K is of the form ST, where $S \simeq SL_m(F)$, and $T \simeq (F^*)^{km-1}$ is the $(km-1)$-dimensional algebraic torus. Each $A \in S \simeq SL_m(F)$ corresponds to a linear transformation of the form:*

$$X = [\dots, X_i^j, \dots] \to [\dots, AX_i^j, \dots]. \tag{12}$$

Each element of the torus T corresponds to a linear transformation of the form

$$X = [\dots, X_i^j, \dots] \to [\dots, \lambda_i^j X_i^j, \dots]. \tag{13}$$

The discrete group K/K_0 contains the wreath product of the alternating group $\mathcal{A}_m$ and $\mathcal{S}_k$ (it acts on X by permuting its columns in the obvious way).

Proof: That K_0 contains linear transformation of the form eq.(12) or eq.(13) is easy. Since $E(X) = \prod_\sigma \det_\sigma(X)$, where each $\det_\sigma(X)$ is a distinct and irreducible polynomial, each element of K must permute factors, modulo multiplications of factors by constants as in eq.(13). Since the stabilizer of each determinantal factor is known (cf. eq.(4)), the elements in K that fix all factors $\det_\sigma(X)$ are seen to be precisely those of the form in eq.(12).

The elements of K/K_0 that permute the factors $\det_\sigma(X)$ are those that permute the columns of X appropriately. They form a group isomorphic to the wreath product of $\mathcal{A}_m$ and $\mathcal{S}_k$. Q.E.D.

Moreover, the stabilizer K characterizes $E(X)$ in the following sense. Let d denote the total degree of $E(X)$. Let $[E(X)]$ be the set of forms of deg d in the entries of X that are stabilized by K.

Proposition 7.2 *Any form $h(X) \in [E(X)]$ can be expressed as a homogeneous polynomials in the maximal minors of X (which correspond to Plücker coordinates). Among such forms $E(X)$ is the simplest one, in the sense that its expansion in terms of Plücker coordinates contains only one monomial.*

Proof: Since $h(X)$ is stabilized by SL_n, which acts on X by left multiplication, the first assertion follows from classical invariant theory. In arbitrary characteristic, it follows from the result of De Concini and Procesi [8]. The second assertion then follows immediately. Q.E.D.

Let W be the space of forms in X with degree equal to that of $E(X)$.

Theorem 7.3 *The point $E(X) \in P(W)$ is stable with respect to the action of $G = SL_n(F)$ on $P(W)$. We assume that the characteristic does not divide k, $k-1$, or m.*

This will be proved in Section 8. Stability of $E(X)$ is a key fact that connects the P vs. NP question to geometric invariant theory.

Let V be the space of forms in Y as in Section 6, where the size of Y is $l > n$. Take a suitable embedding ϕ from W $(P(W))$ to V $(P(V))$ as usual. Then Theorem 7.3 immediately implies the following analogue of Theorem 4.8.

Theorem 7.4 *$E^\phi(Y) \in P(V)$ is partially stable with respect to the action of $SL_l = SL(Y)$, with defect $\delta = 1$. (We assume that the characteristic does not divide k, $k-1$, or m.)*

Now we make the following analogue of the first assertion in Conjecture 4.3.

Conjecture 7.5 *If $l = n^a$, where a is any fixed constant, then $E^\phi(Y)$ cannot lie in the (projective) Zariski closure $\Delta[H(Y)]$ of the $SL_l(F)$-orbit of $H(Y)$ in $P(V)$.*

More generally, let $[H(Y)]$ be the set of forms in the entries of Y, with the same degree as $H(Y)$, that are stabilized by the stabilizer of $H(Y)$. Let $\tilde{\Delta}[H(Y)]$ be the projective closure of $SL_l(F) \cdot [H(Y)]$. A stronger conjecture is that:

Conjecture 7.6 *If $l = n^a$, where a is any fixed constant, then $E^\phi(Y)$ cannot lie in $\tilde{\Delta}[H(Y)]$.*

By Proposition 6.1, Conjecture 7.5 would imply that $E(X)$ does not have an arithmetic circuit over F of size polynomial in n. The conjecture is not expected to be equivalent to the original lower bound question over F. But it is "almost" equivalent to it, when F is the field of complex numbers, in the following sense:

Proposition 7.7 *If Conjecture 7.5 were false, then $E(X)$ can be approximated infinitesimally closely by an arithmetic circuit of $poly(n)$ size.*

Proof: Similar to that of Proposition 4.4 Q.E.D.

But such an approximation is not expected in view of the expected computational hardness of $E(X)$.

Since $H(Y)$ has a nontrivial reductive stabilizer that captures its structure (Propositions 6.2), the orbit closure problem in Conjecture 7.5 can be reduced to the stabilizer problem. Specifically, we make the following stronger conjecture, which implies Conjecture 7.5 or 7.6:

Conjecture 7.8 *The stabilizer of $H(Y)$ cannot occur in the infinitesimal neighbourhood of the $SL_l(F)$-orbit of $E^\phi(Y)$ in the affine space V.*

Since $E(X)$ is stable, and both $H(Y)$ and $E(X)$ have nontrivial stabilizers that capture their structure, the approach based on explicit obstructions (Section 5) seems appropriate for Conjecture 7.8. In the context of constructing explicit obstructions, the following problems then arise (cf. Section 5.5):

1. Explicitly compute finite dimensional G-submodules of $H^i(Z, \mathcal{O}_V^Z)$, with $E^\phi(Y)$ playing the role of f in Section 5.5, and $Z = Gf$. As we mentioned there, the fact that the orbit Gf is a fibre bundle over G/P, in conjunction with the the Leray spectral sequence, is important in this context.

2. The instances of the subgroup restriction problem that arise in the present context are the following. How does a Weyl module representation $S_\lambda(Y)$ of $SL(Y)$ decompose as a Q-module, where Q is the stabilizer of $H(Y)$? How does a Weyl module representation $S_\mu(X)$ of $SL(X)$ decompose as a K-module, where K is the stabilizer of $E(X)$?

Now, we turn to the actual P vs. NP question (the stronger nonuniform version). Here the field under consideration is no longer algebraically closed. The base field that is actually used for computation in real computers is finite, in fact, F_2, though, without of loss of generality, we can take it to be F_p for any small enough prime. In this section we shall assume that $p = \text{poly}(n)$, where n is the number of variables in the problem. We will reduce the P vs. NP question to a certain parametrized family of the general orbit-closure or stabilizer problems over F, where F is the algebraic closure of F_p.

The starting point is the following analogue of Proposition 6.1. We say that a form is an F_p-form if its coefficients are in F_p. The two forms $h_1(Y), h_2(Y) \in V$ are said to be F_p-equivalent if they are equal for all F_p-vectors Y; i.e., when we let each entry of Y range over F_p.

Proposition 7.9 *If $E(X)$, considered as a function of the F_p-vector X, has an arithmetic circuit over F_p of size of r, then some F_p-form h that is F_p-equivalent to $E^\phi(Y)$ lies in the projective orbit closure $\Delta[H(Y)] \subseteq P(V)$, where the size l of Y is $O(r^2)$.*

Proof: Analogous to that of Proposition 4.1. Q.E.D.

This leads us to the following stronger version of Conjecture 7.8, which would imply that $P \neq NP$.

Conjecture 7.10 *If $l = n^a$, where a is any fixed constant, then no F_p-form $h(Y)$ F_p-equivalent to $E^\phi(Y)$ can lie in the (projective) Zariski closure $\Delta[H(Y)]$ of the $SL_l(F)$-orbit of $H(Y)$ in $P(V)$. More strongly, the stabilizer of $H(Y)$ cannot occur in the infinitesimal neighbourhood of the $SL_l(F)$-orbit of any such $h(Y)$ in the affine space V.*

For an approach based on explicit obstructions (Section 5.5) to be applicable to this conjecture, we would want $h(Y)$ to be a good form like $E^\phi(Y)$. The latter is partially stable (Theorem 7.4), and we conjecture that, in fact, a stronger statement holds:

Conjecture 7.11 *Any F_p-form $h(Y)$ F_p-equivalent to $E^\phi(Y)$ is partially stable with respect to the action of $SL_l(F)$ on $P(V)$, with defect $\delta = 1$.*

One can approach other fundamental questions of complexity theory, such as NC^1 vs. NC^2, $P^{\#P}$ vs. NP in a similar fashion by reducing them to various instances of the orbit closure and the stabilizer problem.

8 Stability of $E(X)$

We prove Theorem 7.3 in this section. Let us assume that the characteristic is zero; but essentially the same proof also works when the characteristic does not divide m, k or $k-1$. For simplicity, we also assume that $k = 3$.

Let V be a vector space of dimension m. Let I denote the set $\{(i,j) | 1 \le i \le 3, 1 \le j \le m\}$. Let $X = \{X[i,j] | (i,j) \in I\}$ be a collection of (indeterminate) vectors of V. Let 3^m denote the collection of all functions $\sigma : \{1, \ldots, m\} \to \{1,2,3\}$. For a $\sigma \in 3^m$, let $det_\sigma(X)$ denote the determinant of the matrix $[X[\sigma(1),1], \ldots, X[\sigma(m),m]]$. With this notation

$$E(X) = \prod_{\sigma \in 3^m} det_\sigma(X)$$

Clearly, the action of $SL(V)$ on V stabilizes $E(X)$. In other words, if $A \in SL(V)$ and AX denotes the collection $\{AX[i,j] | (i,j) \in I\}$ then $E(AX) = E(X)$. Let D denote the wreath-product of $\mathcal{S}_m$ with $\mathcal{S}_3$. An element of D is given by a tuple $\beta = (\mu; \alpha_1, \ldots, \alpha_m)$ where $\mu \in \mathcal{S}_m$ and each $\alpha_i \in \mathcal{S}_3$. D acts on I naturally, *viz.*, $\beta = (\mu, (\alpha_r))(i,j) = (\alpha_j(i), \mu(j))$. This action on I extends to X naturally, *viz.*, $\beta(X[i,j]) = X[\beta(i,j)]$, and this action stabilizes $E(X)$ up to sign.

Let $\mathcal{Y}$ be the vector space consisting of formal linear combinations of the symbols $\{y_{ij} | (i,j) \in I\}$. D acts on $\mathcal{Y}$ naturally, *viz.*, $\beta(y_{ij}) = y_{\beta(i,j)}$. We shall examine this representation of D.

Let $e' = \sum_{(i,j) \in I} y_{ij}$ and $f'_j = y_{1j} + y_{2j} + y_{3j}$. Let $\mathcal{E}$ be the span of e' and $\mathcal{F}$ be the span of all vectors of the form $\{f'_r - f'_s | 1 \le r, s \le m\}$. It is clear that $\mathcal{E}, \mathcal{F}$ are invariant under the action of D. We claim that $\mathcal{E}$ and $\mathcal{F}$ are, in fact, irreducible representations of D. Irreducibility of $\mathcal{E}$ is clear. For $\mathcal{F}$ note that $\mathcal{F} = \{\sum_{r=1}^m c_r f'_r | \sum_{r=1}^m c_r = 0\}$. Note that the action of D on $\mathcal{F}$ reduces to an action of $\mathcal{S}_m$ on $\mathcal{F}$ which is isomorphic to the irreducible representation of $\mathcal{S}_m$ corresponding to the partition $(m-1,1)$. Let $\mathcal{H}$ be the following subspace of $\mathcal{Y}$:

$$\mathcal{H} = \{\sum c_{ij} y_{ij} | c_{1j} + c_{2j} + c_{3j} = 0 \, \forall j\}$$

It is easy to show that $\mathcal{H}$ is also D-irreducible. Thus we see that the D-module $\mathcal{Y}$ splits into irreducible components $\mathcal{Y} = \mathcal{E} \oplus \mathcal{F} \oplus \mathcal{H}$ as D-modules and thus $\mathcal{Y}$ is a multiplicity-free representation of D.

Let D' be the wreath-product of A_m (the alternating group) with $\mathcal{S}_3$. We note that (i) $D' \subset D$ actually stabilizes $E(X)$ and (ii) for $m > 4$, the irreducible D-modules $\mathcal{E}, \mathcal{F}$ and $\mathcal{H}$ are irreducible as D'-modules as well.

Let $h'_{1j} = 2y_{1j} - y_{2j} - y_{3j}$. Define h'_{2j} and h'_{3j} similarly and note that $h'_{ij} \in \mathcal{H}$. We may check that

$$3m.y_{ij} = e' + \sum_{r:r\neq i} (f'_i - f'_r) + m.h'_{ij} \tag{14}$$

Next, we analyze the stability of the form $E(X)$. If $E(X)$ were unstable, then by Kempf's criteria, there is a one-parameter subgroup $\lambda : k^* \to SL(X)$ "witnessing" its instability and which commutes with the stabilizer of the form $E(X)$. In other words, there would be an $SL(X)$-invariant closed subset of S of forms on X, and a one-parameter subgroup λ commuting with D' above, driving $E(X)$ into S. Since $SL(V)$ acts on X as a diagonal action on $3m$ copies of V, by Schur's lemma, we see that the image of λ must lie in the commutator of this representation of $SL(V)$. This reduces λ to be of the form $\lambda = \gamma \otimes I(V)$, where $I(V)$ is the identity map on V and $\gamma : k^* \to SL(\mathcal{Y})$. Within $SL(\mathcal{Y})$ it must further commute with the D'-representation $\mathcal{Y}$. Since $\mathcal{Y} = \mathcal{E} \oplus \mathcal{F} \oplus \mathcal{H}$ splits as a D'-module, we conclude that D' is actually contained in a suitable parabolic subgroup of $SL(X)$, whence Kempf's first criteria, based on noncontainment of the stabilizer in a proper parabolic group, fails to apply. However, since λ commutes with D', we see that λ can be expressed as

$$\gamma(t)(3m.y_{ij}) = t^a e' + t^b \sum_{r:r\neq i} (f'_i - f'_r) + t^c.m.h'_{ij} \tag{15}$$

Since $dim(\mathcal{E}) = 1, dim(\mathcal{F}) = m - 1$ and $dim(\mathcal{H}) = 2m$, and $\lambda(t) \in SL(X)$ for all t, we have the important relation:

$$a + (m-1)b + 2mc = 0 \tag{16}$$

We define the "unprimed" vectors $e(X), f_j(X), h_{ij}$ (X) as the analogues of e', f'_j and h'_{ij} by replacing the symbol y_{rs} by the vector variable $X[r,s]$. Thus $e(X) = \sum_{(i,j)\in I} X[i,j]$ and $f_j(X) = X[1,j] + X[2,j] + X[3,j]$. The vector $h_{ij}(X)$ is similarly defined. For notational convenience, we drop the argument (X) from $e(X), f_i(X)$ and $h_{ij}(X)$. In this new notation, we have the analogue of equation (14):

$$3m.X[i,j] = e + \sum_{r:r\neq i} (f_i - f_r) + m.h_{ij} \tag{17}$$

Let $\chi_i = \chi_i(X)$ denote the vector $\sum_{r:r\neq i}(f_i - f_r)$. In the form $E(X)$, we may substitute the above expression for each $X[i,j]$ and using the multi-linearity of the determinant, expand $E(X)$ into its "constituents". These constituents may be aggregated by their weights with respect to the one-parameter subgroup λ as follows.

Each constituent is a product of determinants with columns from the collection $\{e, (\chi_i), (h_{ij})\}$. If a particular determinant takes for its arguments k_1 copies of e, k_2 elements of (χ_i) and k_3 elements from the set (h_{ij}), (with $k_1 + k_2 + k_3 = m$), then the degree of that determinant may be taken as $k_1 a + k_2 b + k_3 c$. The degree of a product of determinant factors is clearly the sum of the individual degrees. Then we have

$$\lambda(t)(E(X)) = \sum_w t^w E_w(X) \tag{18}$$

where for an integer w the term $E_w(X)$ is the form of degree w. We say that the integer w is a weight in the above expression, if $E_w(X)$ is non-zero. If $E(X)$ were unstable then there are integers a, b, c satisfying (16) above such that

$$E_w(X) = 0 \ \ for \ \ all \ \ w < 0 \tag{19}$$

This is precisely the condition for $\lim_{t\to 0} \lambda(t)(E(X))$ to exist. We shall show that this condition is not possible for any choice of a, b and c as above.

In the expression $E(X) = \sum_w E_w(X)$, there are many terms which vanish. First note that the vector e cannot occur twice inside any non-trivial determinant. Similarly, we have the relation $\sum_{i=1}^m \chi_i = 0$, thus not all the arguments of a non-trivial determinant can be from the set (χ_i). We also note that for any collections $\{v, (w_i)_{i=1}^m, (y_{ij})_{(i,j)\in I}\}$ of vectors in V such that $\sum_i w_i = 0$ and $y_{1j} + y_{2j} + y_{3j} = 0$ for all j, there are vectors $\overline{x} = \{x[r,s] \in V | (r,s) \in I\}$, such that $e(\overline{x}) = v, \chi_i(\overline{x}) = w_i$ and $h_{ij}(\overline{x}) = y_{ij}$.

So suppose that $E(X)$ were unstable and a, b, c were as above.

Case 1: $a, b \geq c$. In this case, equation (16) tells us that $c \leq 0$. Thus the minimum possible weight is $w_{min} = 3^m.m.c$ coming from terms such as $\prod_{\sigma\in 3^m} det(h_{\sigma(1),1}, \ldots, h_{\sigma(m),m})$, with a few more terms if $b = c$ or $a = c$.

It is easy to show that $E_{w_{min}} \neq 0$: Select $\overline{x} = (x[r,s])$ such that $e(\overline{x}) = \chi_i(\overline{x}) = 0$ (for all i) and $h_{1j}(\overline{x}) = h_{2j}(\overline{x}) = e_j$ and $h_{3j}(\overline{x}) = -2e_j$, where e_j is the j-th "unit" basis vector. For this choice of $\overline{x}$, we see that

$$E_{w_{min}}(\overline{x}) = \prod_{\sigma\in 3^m} det(h_{\sigma(1),1}, \ldots, h_{\sigma(m),m}) \neq 0$$

Thus, unless $c = 0$, which in turn, implies $a, b, c = 0$, there is a negative weight term, contradicting (19).

Case 2: $c \geq a, b$. We choose $\overline{x}$ such that $h_{ij}(\overline{x}) = 0$ for all $(i,j) \in I$. Then the minimum weight term occurs in the expansion of

$$[det(e(\overline{x}) + \chi_1(\overline{x}), \ldots, e(\overline{x}) + \chi_m(\overline{x}))]^{3^m}$$

(with the relation $\chi_1(\overline{x}) + \ldots + \chi_m(\overline{x}) = 0$).

Since e may occur only once and the χ_i's only $m-1$ times, the only weight possible is $w_{min} = (m-1)b + a$. We may check that $E_{w_{min}}$ is actually non-zero, by choosing $\overline{x}$ such that in addition to the requirements above, we have $\chi_i(\overline{x}) = e_i$ for $i = 1, \ldots, m-1$ and $\chi_m(\overline{x}) = -(e_1 + \ldots + e_{m-1})$ and $e(\overline{x}) = e_m$. For this choice of $\overline{x}$, we see that $E_{w_{min}}(\overline{x}) \neq 0$. Since $a + (m-1)b + 2mc = 0$ (equation (16)), we have that $a + (m-1)b < 0$. Thus, again, unless $a, b, c = 0$, there is a negative weight term, again contradicting (19).

Case 3: $a > c > b$. In this case consider the assignment $\overline{x}$ such that (i) $e = 0$, (ii) $h_{1j} = h_{2j} = e_j$, $h_{3j} = -2e_j$ for all j, and (iii) $\chi_i = e_i$ for $i = 1, \ldots, m-1$ with $\chi_m = -(e_1 + \ldots + e_{m-1})$. We see that for any $\sigma \in 3^m$, in the expression

$$det(\chi_1 + m.h_{\sigma(1),1}, \ldots, \chi_m + m.h_{\sigma(m),m})$$

the minimum weight term is

$$\sum_{i=1}^{m} m.det(\chi_1, \ldots, \chi_{i-1}, s.e_i, \chi_{i+1}, \ldots, \chi_m)$$

Here the constant $s = 1$ if $\sigma(i) = 1$ or 2 and -2 otherwise. This expands exactly to either of the numbers m or $-2m$ and is thus non-zero. The weight of this term is minimum possible, which is $(m-1)b + c$. If indeed λ witnesses the instability of $E(X)$ (and therefore, satisfies (19)), then $(m-1)b + c \geq 0$. If $c \geq 0$ then this implies that $(m-1)b + 2mc \geq 0$. Since $a > 0$, we have $a + (m-1)b + 2m.c > 0$ which is a contradiction to equation (16). On the other hand, $c \leq 0$ and $c > b$ together imply $(m-1)b + c < 0$. Thus there is a negative weight non-zero term, contradicting (19).

Case 4: $b > c > a$. In this case consider the assignment $\overline{x}$ such that (i) $e = e_m$, (ii) $h_{1j} = h_{2j} = e_j$, $h_{3j} = -2e_j$ for all j, and (iii) $\chi_i = 0$ for all i. With $b > c > a$, the minimum weight possible is $(m-1)c + a$. We show that this weight is indeed achieved: We analyze the expression:

$$det(e + m.h_{\sigma(1),1}, \ldots, e + m.h_{\sigma(m),m})$$

Since $e = e_m$, the only term which survives is:

$$det(m.h_{\sigma(1),1}, \ldots, m.h_{\sigma(m-1),m-1}, e_m)$$

This is clearly non-zero. Thus there is a non-zero term of weight $(m-1)c + a$. Since (a, b, c) is a witness to the instability of $E(X)$, we must have $(m-1)c + a \geq 0$. The only way this can hold is with $c \geq 0$. But this implies that $2mc + a = (m-1)c + a + (m+1)c \geq 0$. But $b > 0$, and consequently $a + (m-1)b + 2mc > 0$, which is untenable, because of (16).

Thus we must have $(a, b, c) = (0, 0, 0)$ which proves the stability of $E(X)$.

Q.E.D.

References

[1] P. Bardsley, R. Richardson: Étale slices for algebraic transformation groups in characteristic p, Proc. London Math. Soc. (3), 51 (1985), 295-317.

[2] S. Berkowitz: On computing the determinant in small parallel time using a small number of processors. Inform. Process. Lett. 18 (1984), 147-150.

[3] R. Boppana, M. Sipser: The complexity of finite functions, Handbook of Theoretical Computer Science, vol. A, Edited by J. van Leeuwen, North Holland, Amsterdam, 1990, 757–804.

[4] A. Borel: Linear algebraic groups, second edition, Springer 1991.

[5] P. Bürgisser, M. Clausen, M. Shokhrollahi, Algebraic complexity theory, Springer-Verlag, 1997. A.

[6] R. Carter, G. Segal, I. Macdonald: Lectures on Lie groups and Lie algebras, London Mathematical Society, 1995.

[7] C. De Concini, D. Eisendud, C. Procesi: Young diagrams and determinantal varieties, Inventiones mathematicae, 56, 129-165 (1980).

[8] C. De Concini, C. Procesi: A characteristic-free approach to invariant theory, Adv. in Math. 21, 330 (1976).

[9] C. De Concini, C. Procesi: Complete symmetric varieties, Invariant theory (Montecatini, 1982) Lecture Notes in Math. 996, Springer, 1983.

[10] S. Cook: The complexity of theorem-proving procedures. Proceedings of the third annual ACM Symposium on Theory of Computing. 151-158. (1971).

[11] S. Cook: The classification of problems which have fast parallel algorithms. Fundamentals of Computation Theory 4 (1983).

[12] E. Formanek, C. Procesi: Mumford's conjecture for the general linear group. Adv. in math. 19, 292 (1976).

[13] W. Fulton: Introduction to toric varieties, Princeton University Press, 1993.

[14] W. Fulton, J. Harris: Representation theory, Springer-Verlag, 1991.

[15] J. von zur Gathen, Permanent and determinant, Linear algebra and its applications, 96: 87–100 (1987).

[16] M. Grotschel, L. Lovász, A. Schriver: Geometric algorithms and combinatorial optimization. Springer-Verlag, 1988.

[17] W. Haboush: Reductive groups are geometrically reductive. Ann. Math. II Ser. 102, 67-83 (1975).

[18] R. Hartshorne: Algebraic geometry, Springer-Verlag, 1977.

[19] D. Hilbert: Über die vollen Invariantensysteme. Math. Ann. 42, 313-373 (1893).

[20] R. Karp: Reducibility among combinatorial problems. R. E. Miller and J. W. Thatcher (eds.) Complexity of computer computations, Plenum Press, New York, 1972, 85-103.

[21] A. Klymyk: Multiplicities of weights of representations and multiplicities of representations of semisimple Lie algebras, Sov. Math. Dokl. 8 (1967), 1531-1534.

[22] H. Kraft, P. Slodowy, P. Springer: Algebraische Transformationsgruppen aund Invarianttentheorie, DMV Seminar, Bd. 15, Birkhäuser, 1993.

[23] G. Kempf: Instability in invariant theory, Annals of Mathematics, 108 (1978), 299-316.

[24] V. Lakshmibai, C. Musili, C. S. Seshadri: Geometry of G/P, Bull. Amer. Math. Soc. 1 (1979), 432-435.

[25] A. Levin: Universal sequential search problems. Problems of information transmission (translated from Problemy Peredachi Informatsii (Russian)) 9 (1973).

[26] A. Lubotzky, R. Phillips, P. Sarnak: Ramanujan graphs, Combinatorica 8 (1988), 261-277.

[27] D. Luna: Slices Etales. Bull. Soc. Math. France, M emoire 33, 81 (1973).

[28] D. Luna, Th. Vust: Plongements d'espaces homogenes. Comment. Math. Helv. 58, 186 (1983).

[29] G. Margulis: Explicit constructions of concentrators, Problemy Inf. Trans. 9 (1973), 325-332.

[30] H. Minc: Permanents. Addison-Wesley, 1978.

[31] M. Nagata, Polynomial Rings and Affine Spaces. CBMS Regional Conference no. 37, American Mathematical Society, 1978.

[32] Y. Matsushima: Espaces homogènes de Stein des groupes de Lie complexes. Nagoya Math. J. 16, 205-218 (1960).

[33] K. Mulmuley: Lower bounds in a parallel model without bit operations, The SIAM Journal On Computing, vol. 28, no. 4, 1999.

[34] K. Mulmuley, M. Sohoni: Geometric Complexity Theory II: Explicit Obstructions, manuscript under preparation.

[35] k. Mulmuley, U. Vazirani, V. Vazirani: Matching is as easy as matrix inversion, Combinatorica, vol 7, no. 1, 1987, pp 105-113.

[36] D. Mumford: Geometric invariant theory. Ergebnisse der Mathematik 34, Springer-Verlag, 1965.

[37] D. Mumford, J. Fogarty, F. Kirwan: Geometric invariant theory. Springer-Verlag, 1994.

[38] V. Popov, E. Vinberg, Invariant theory, in Encyclopaedia of mathematical sciences, vol. 55, Springer Verlag, 1991.

[39] A. Razborov, S. Rudich, Natural proofs, JCSS 55 (1997), 24–35.

[40] C. Seshadri: Theory of moduli, Algebraic Geometry, Proc. Arcata Symposia in Pure Math. 29, A.M.S., 1975, 263-306.

[41] S. Smale, Mathematical Problems for the next century, in Mathematics: frontiers and perspectives, pg. 271-294. A.M.S., Providence, R.I., 2000.

[42] T. Springer, Linear algebraic groups, in Encyclopaedia of mathematical sciences, vol. 55, Springer Verlag, 1991.

[43] H. Grauert, R. Remmert, Theory of Stein spaces, Springer-Verlag, 1979.

[44] L. Valiant: The complexity of computing the permanent. Theoretical Computer Science 8 , 189-201 (1979).

[45] L. Valiant: Completeness classes in algebra, Proceedings of the ACM Symposium on Theory Of Computing (1979).

[46] H. Weyl: The classical groups. Their invariants and representations. Princeton University Press. 302 (1939).