The Geometry of Matrix Rigidity

Joseph M. Landsberg * Jacob Taylor † Nisheeth K. Vishnoi ‡

November 26, 2003

Abstract

Consider the following problem: *Given an $n \times n$ matrix A and an input x, compute Ax.* This problem has a simple algorithm which runs in time $O(n^2)$. The question thus is: Is this is the best possible ?

Valiant showed ([12]) that if an $n \times n$ matrix A is **rigid,** then the smallest straight line program computing Ax is either super-linear size, or has super logarithmic depth. Roughly a matrix of rank n is said to be **rigid,** if to bring its rank down to $n/2$, one has to change at least $n^{1+\epsilon}$ of its entries, for some $\epsilon > 0$. After considerable effort by many researchers, the problem of finding an explicit rigid matrix (hence proving circuit lower bounds) remains elusive.

This paper casts the problem of matrix rigidity in the language of algebraic geometry. As a first step, we provide the basic setup and prove some elementary results about this problem. This setting facilitates our understanding of the difficulty of determining the rigidity of an explicit matrix (like Hadamard). On the brighter side, we hope that tools from algebraic geometry might eventually help establish rigidity of explicit matrices.

1 Introduction

One of the important problems in Numerical Analysis and Computer Science is:

Given an $n \times n$ matrix A with entries from a field K, and an input x, compute Ax.

This problem has a simple algorithm which runs in time $O(n^2)$. The question thus is: Is this is the best possible ?

Valiant in [12] noticed that if an $n \times n$ matrix A is **rigid,** then the smallest *straight line program* computing Ax is either super-linear size, or has super logarithmic depth. Roughly a matrix of rank n is said to be **rigid** if to bring its rank down to $n/2$, one has to change at least $n^{1+\epsilon}$ of its entries, for some $\epsilon > 0$.

Before proceeding let us define **rigiditiy** formally.

Definition 1.1. *For a field K, the* **rigidity** *of a matrix A is the function $R_A^K(r) : \{1, \ldots, \min(m, n)\} \to \{0, 1, \ldots, mn\}$ defined by*

$$R_A^K(r) := \min\{s | \exists B \in K^{m\times n},\ \mathit{supp}(B) = s \ \mathit{and} \operatorname{rank}(A + B) \leq r\}.$$

Here $\mathit{supp}(B)$ denotes the number of non-zero entries in B.

*School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332. Email: jml@math.gatech.edu.
†School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332. Email: taylor@math.gatech.edu.
‡College of Computing, Georgia Institute of Technology, Atlanta GA 30332. Emailnkv@cc.gatech.edu

Valiant proved that when K is infinite, most matrices $A \in K^{m\times n}$ have $R_A^K(r) = (m-r)(n-r)$. To prove computational lower bounds for computing Ax, one has to come up with an explicit matrix with high rigidity, and existential arguments are not sufficient. He conjectured that the Hadamard matrix is rigid. This matrix arises in computing the Fourier transform over the group $\mathbb{Z}_2^m$ (let $n := 2^m$) and also from many other settings. The best result for the rigidity of a $n \times n$ Hadamard matrix is $\Omega(n^2/r)$ due to Kashin and Razborov [6]. For $r = n/2$, this reduces to $\Omega(n)$.

In other explicit results, a lower bound of $\Omega(n^2/r \log n/r)$ was shown by Friedman [4] showed for certain explicit matrices, while a similar result was obtained by Shokrollahi et al. [10].

Establishing high rigidity of explicit matrices has other applications and the reader is referred to the papers [8, 2, 3, 7, 1, 9].

Valiant's Contribution

Valiant proved the following theorem relating the straight line complexity of computing Ax and the rigidity of A. (For definition of straight line program refer Appendix A.)

Theorem 1.2. *[12] Let $A_1, \ldots, A_n, \ldots$ be an infinite family where A_n is a real $n \times n$ matrix and for some constants $c, \epsilon > 0$, $R_{A_n}(n/2) \geq cn^{1+\epsilon}$. Then there does not exist a family of straight line programs for the corresponding sets of linear forms that for some $c_1, c_2 > 0$ achieve size $c_1 n$ and depth $c_2 \log n$ simultaneously for all n.*

He also proved the following:

Theorem 1.3. *[12] For an infinite field K, for all n there are $n \times n$ A such that $R_A^K(r) = (n-r)^2$.*

Our Contribution

The main contribution of this paper is to cast the problem of Matrix Rigidity in the language of Algebraic Geometry. In the set up, three sets of algebraic objects arise naturally:

- The set of full rank matrices.
- The set of matrices of rank at most r.
- The set of matrices having at least p zeros.

We study the associated *algebraic varieties* and their basic invariants. Using some of the known results and techniques from geometry, we establish that almost all matrices are *maximally rigid*.

It becomes clear that the hardness of establishing rigidity of explicit matrices turns out to be the same as deciding whether a given point lies on a certain implicitly define algebraic varieties. (Its worth mentioning here that Strassen [11] showed that determining the complexity of a related problem- Matrix Multiplication- reduces to determining whether a point lies on a certain variety.)

We also define the concept of **defect** (again borrowed from geometry) for matrices capturing: *How far a matrix is from being rigid?*

We hope that providing this language will eventually help resolve some of the important open problems in this area. Thus our contribution is to be viewed as yet another connecting bridge between Algebraic Geometry and Complexity Theory.

2 Geometric Preliminaries

For this paper $K = \mathbb{C}$. Let the vector space $V := \mathbb{C}^{N+1}$. Then $\mathbb{P}V = \mathbb{P}^N$, the projective space of lines through the origin in V. We define a **projective variety** $X \subset \mathbb{P}V$ to be the zero set of a finite collection of homogeneous polynomials in $N+1$ variables. For $x \in X$, we will denote the line over x as $\hat{x} \subset V$.

Definition 2.1. *For $Y \subset \mathbb{P}V$, the closure of the* **cone** $\widehat{Y} \subset V$ *over Y is the union of $0 \in V$ and inverse image of Y under the projection $\pi : V \backslash 0 \to \mathbb{P}V$.*

Definition 2.2. *A variety X is* **irreducible** *if for any pair of subvarieties $Y, Z \subseteq X$ such that $Y \cup Z = X$, either $Y = X$ or $Z = X$.*

Definition 2.3. *x is a* **smooth** *point of X, if X is a manifold in some neighborhood of x (in the analytic topology). The set of smooth points of X is denoted by X_{sm}. Also denote by X_{sing} the set of* **singular** *points of X ($:= X \backslash X_{\mathrm{sm}}$).*

Definition 2.4. *The* **dimension** *of X, $\dim X$, is the smallest integer n such that a general $(N-n-1)$-plane $\Lambda \subset \mathbb{P}^N$ is disjoint from X.*

Definition 2.5. *Given a submanifold $M \subset V$ and $v \in M$, we let $T_v^{\mathrm{aff}} M \subset V$ denote the affine tangent space to M at v, that is the naive tangent space by taking the union of all embedded tangent lines to curves on M at v. (We distinguish this from the abstract tangent space.) Given $X \subset \mathbb{P}V$ and $x \in X_{\mathrm{sm}}$, we let $\widehat{T}_x X = T_v^{\mathrm{aff}} \widehat{X}$, where $v \in \widehat{x}$ is any nonzero point and we observe that this is well defined as the tangent space is constant along the fibers of π. We let $\widetilde{T}_x X = \pi(\widehat{T}_x X) \subset \mathbb{P}^N$ denoted the tangent projective space to X at x.*

Definition 2.6. *The* **join** *of two varieties $X, Y \subset \mathbb{P}V$ as*

$$J(X,Y) = \overline{\bigcup_{x \in X, y \in Y} \mathbb{P}^1_{xy}}.$$

Here $\mathbb{P}^1_{xy}$ is the projective line joining the points x and y.

Although an algebraic variety is a topological manifold, to better keep track of the algebraic nature of a variety we use a different topology:

Definition 2.7. *The* **Zariski topology** *on a variety X is the topology where closed sets are the zero sets of homogeneous polynomials.*

3 The Geometric Connection

In this section we give a sketch of the geometric setting. The three algebraic varieties related to this question are:

1. $M_{m \times n} := \mathbb{P}(\mathbb{C}^n \otimes \mathbb{C}^m)$. The set of full rank matrices forms a Zariski closed set of this variety and is a *a quasi projective variety.* [1] (Henceforth we refer this variety as M.)
2. For $0 \leq r \leq \min\{m, n\}$, let $M_r := \mathbb{P}(\{A \in M | \operatorname{rank}(A) \leq r\})$.
3. For $p \geq 0$, let $S_p := \mathbb{P}(\{B \in M | B$ has atleast p zeros$\})$.

[1]See Harris [5] for more details.

Note that the definition of S_p depends on choices of coordinate axes in $\mathbb{C}^n$ and $\mathbb{C}^m$ while the first two varieties are well defined independent of coordinate choices.

If there is an r and a $p := p(m,n,r)$ such that every full rank matrix A, can be written as $A = L + Z$, for $L \in M_r$ and $Z \in S_p$, then

$$M = J(M_r, S_p).$$

It is clear that the larger p is the harder it is for the above equality to be satisfied. The problem thus is:

Problem 3.1. *For given nonnegative integers m, n, r, with $r \leq \min\{m,n\}$, what is the largest $p := p(m,n,r)$ such that*

$$M = J(M_r, S_p)?$$

We sketch the technique to determine the optimal $p(m,n,r)$ here. Let $p := p(m,n,r)$ be the largest integer such that

1. $J(M_r, S_p) \subseteq M$, and
2. $\dim M = \dim J(M_r, S_p)$.

The first condition is trivially true. Thus our task reduces to finding the largest $p(m,n,r)$ for which the second condition holds.

The following theorem shows that $p(m,n,r) \geq r(m+n-r)$ for all fields (see appendix B for its proof):

Theorem 3.2. *For a given $m \times n$ matrix A, (say of rank m, hence $m \leq n$) there is a always a way to write it as a sum of a $r(\leq m)$ rank matrix and a matrix with at most $(m-r)(n-r)$ non zero entries.*

Next we show that $\dim J(M_r, S_p) = \dim M + (r(m+n-r) - p)$. This implies that if $p > r(m+n-r)$, then $J(M_r, S_p) < \dim M$, and hence M cannot be contained in $J(M_r, S_p)$. This establishes that $p \leq r(m+n-r)$. We show this using Terracini's Lemma ([13], Chapter 2, Proposition 1.10). The lemma says that the tangent space at a general point of a join is the sum of the tangent spaces at the respective points of the two varieties. This allows us to retrieve the dimension of the join.

Terracini's lemma applied to our situation states

Theorem 3.3. *For given non-negative integers m, n, r, with $r \leq \min\{m,n\}$, and $p := p(m,n,r) \geq r(m+n-r)$,*

$$\dim(\widehat{T}_{[A+y]} J(M_r, S_p)) = \dim(\widehat{T}_{[A]} M_r) + \dim(\widehat{T}_{[y]} S_p),$$

where $[A]$ and $[y]$ are general points of M_r and S_p respectively and $[A+y]$ is a general point of $J(M_r, S_p)$.

A corollary of Theorems 3.2 and 3.3 is the main result:

Corollary 3.4. *The largest $p(m,n,r)$ such that $J(M_r, S_p) = M$ is $r(n+m-r)$.*

Hence to reduce the rank of a *generic* full rank $m \times n$ matrix down to r, one needs to change exactly $(m-r)(n-r)$ entries. The exceptional set is a Zariski closed set in M. Also since rational points are dense in $\mathbb{R}$, there are integer matrices which attain maximum rigidity.

To prepare for the use of Terracini's Lemma, in Section 3.1 we give a characterization of the smooth points of M_r and compute the affine tangent space at its smooth points. In Section 3.2 we do the same for S_p. In Section 3.3 we prove Theorem 3.3 by finding $p(m,n,r)$ such that M_r and S_p intersect transversally.

For $p = d + r(m+n-r)$, we get a *stratification* of the space M into $J(M_r, S_{p+d})$ (indexed by d). Determining the rigidity of explicit matrices (like Hadamard) now reduces to checking which strata does the matrix lie in. Testing whether points belong to varieties is a rather hard question if there is no short set of polynomials describing the variety. In our case we need efficient representation of the variety $J(M_r, S_{p+d})$. A set of polynomials describing $J(M_r, S_{p+d})$ does exist but the problem is that this variety is reducible with many components, so there is a huge number of equations to check.

3.1 The Variety M_r

In this section we take $V = \mathbb{C}^n \otimes \mathbb{C}^m$, and let $M = \mathbb{P}V$. Define

$$M_r := \mathbb{P}(\{A \in V |\, \mathrm{rank}(A) \leq r\}).$$

M_r is the common zero set of the $(r+1)\times(r+1)$ minors, which are homogeneous polynomials of degree $r+1$ on the projective space M. Hence M_r is a projective variety.

The following are well known facts about M_r and the reader is referred to the book by Harris [5].

Proposition 3.5.

1. $\dim M_r = r(m+n-r) - 1$.

2. *The* $(r+1)\times(r+1)$ *minors generate the homogeneous ideal* $I(M_r)$ *of* M_r.

3. M_r *is irreducible and quasi-homogeneous; i.e.* $M_r \setminus M_{r-1}$ *is a group orbit of the action of* $\mathbf{PGL}_m\mathbb{C} \times \mathbf{PGL}_n\mathbb{C}$ *on* M. *(Since this is the case, the smoothness or singularity of a point* $A \in M$ *can only depend on rank.)*

First we characterize the smooth points of M_r. This again can be found in Harris.

Lemma 3.6. *[5]* $(M_r)_{\mathrm{sm}} = M_r \setminus M_{r-1}$

Now we give the description of the affine tangent space of M_r at its smooth points.

Lemma 3.7. *For* $A \in (M_r)_{\mathrm{sm}}$*,*

$$\widehat{T}_A M_r = \{B \in M_{\mathrm{m\times n}} \,|\, (\mathrm{Im}\, B)|_{\mathrm{Ker}\, A} \subset \mathrm{Im}\, A\}$$

Proof. Define A as above. the only minors with nonzero differential at A are those having $x_{ij}, i, j > r$ as the linear terms in their expansions (i.e. those involving the first r rows and columns). So matrices

$$\begin{pmatrix} X & Y \\ Z & 0 \end{pmatrix} \in \mathrm{Ker}\, \{df_{ij}|_A\}$$

where X, Y, Z are arbitrary since minors involving the last $m-r$, $n-r$ rows and columns vanish identically. The bases $\{e_i\}, \{f_j\}$ were chosen so that Ker (A) is exactly span$\{e_{k+1}, \ldots, e_n\}$ and Im (A) is span$\{f_1, \ldots, f_k\}$. so,

$$\mathrm{Ker}\, \{df_{ij}|_A\} = \{B \in M_{\mathrm{m\times n}} \,|\, (\mathrm{Im}\, B)|_{\mathrm{Ker}\, A} \subset \mathrm{Im}\, A\}$$

By definition, $T_A^{\mathrm{aff}} M_r = \widehat{T}_A M_r$, hence the claim.

□

3.2 The Variety S_p

The other projective variety we'll be interested in is

$$S_p = \mathbb{P}(\{B \in V | B \text{ has atleast } p \text{ zeros}\}).$$

S_p is the union of $\binom{mn}{p}$ $(mn-p-1)$-planes in $\mathbb{P}V$. It is a reducible variety of dimension $\dim S_p = mn - p - 1$.

In this section we characterize the smooth points of S_p and compute the affine tangent space at its smooth points.

Lemma 3.8. $(S_p)_{\rm sm} = S_p \setminus S_{p+1}$

Proof. Any time a variety has multiple components, its smooth points are the smooth points on each component that do not intersect any other component. $\square$

Note that since S_p is a union of linear components, the tangent space T_xS_p to S_p at a smooth point x is just the particular linear component of S_p in which x lives.

3.3 Dimension of $J(M_r, S_p)$

Definition 3.9. *For $X, Y \subset \mathbb{P}V$, the* **expected dimension** *of $J(X,Y)$ is*

$$\dim J(X,Y) = \min \begin{cases} \dim X + \dim Y + 1 \\ \dim V \end{cases}$$

Notice that $J(M_r, S_p)$ is reducible

$$J(M_r, S_p) = \bigcup_I J(M_r, L_I)$$

where L_I is a coordinate $(mn-p-1)$-plane, and I is an index representing which entries of L_I are zero. All the L_I's are isomorphic since $\mathbf{PGL}_m\mathbb{C} \times \mathbf{PGL}_n\mathbb{C}$ acts transitively on the coordinate planes, so it suffices to find some L_0 such that $A \in J(M_r, L_0)$. Hence a solution to this problem would be $[L_r] \in M_r, [L_0] \in S_p$ such that $[A] = [L_r + L_0]$.

In general we expect that the dimension of $J(M_r, S_p)$ to be

$$\dim J(M_r, S_p) = \min \begin{cases} \dim M_r + \dim S_p + 1 \\ mn - 1 \end{cases}$$

However, this is not always the case: for example, for $X = M_1 \subset \mathbb{P}(\mathbb{C}^3 \otimes \mathbb{C}^3)$,

$$\dim J(M_2, S_0) = \dim M_2 - 1$$

The following lemma characterizes when the expected dimension is the actual dimension.

Lemma 3.10 (Terracini's Lemma [13]). *For $[x] \in X_{\rm sm}, [y] \in Y_{\rm sm}$,*

$$\widehat{T}_{[x+y]} J(X,Y) = \widehat{T}_{[x]}X + \widehat{T}_{[y]}Y$$

In particular, for $X, Y \subset \mathbb{P}^N$ and $N \geq \dim X + \dim Y + 1$, we get the expected dimension of $J(X,Y)$ if and only if for $x \in X_{\rm sm}, y \in Y_{\rm sm}$

$$\widehat{T}_{[x]}X \cap \widehat{T}_{[y]}Y = (0).$$

In our case, for fixed rank $= r$, if $p \geq r(m+n-r)$ we get the **expected dimension** of the join iff the intersection is (0) (**transversal**).

Let

$$A = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}$$

and recall (by Lemma 3.7) that

$$\widehat{T}_A M_r = \{B \in M_{\rm m\times n} \,|\, (\operatorname{Im} B)|_{\operatorname{Ker} A} \subset \operatorname{Im} A\}$$

Let $y \in S_p \setminus S_{p+1}$. Since S_p is a union of linear spaces, $T_y S_p$ is the particular linear component in which y lives. Consider

$$B = \begin{pmatrix} 0 & 0 \\ 0 & \widetilde{B} \end{pmatrix}$$

where $\widetilde{B}$ is an $(m-r) \times (n-r)$ with only 1 entry equal to zero. For example

$$\widetilde{B} = \begin{pmatrix} 0 & x & \dots & x \\ x & x & & \\ \vdots & & \ddots & \\ x & & & x \end{pmatrix}$$

The number of zeros in B is $mn - [(m-r)(n-r) - 1]$, which is $r(n+m-r)+1$. Therefore $B \in T_y S_p$. Notice that $\widehat{T}_{[A]} M_r \bigcap \widehat{T}_{[y]} S_p \neq (0)$ iff $(b)_{ij} = 0$ for all $i, j : \ r+1 \leq i \leq m, \ r+1 \leq j \leq n$. Hence $\widehat{T}_{[A]} M_r \cap \widehat{T}_{[y]} S_p = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Hence we have proved the following lemma

Lemma 3.11. *For $p \geq r(m+n-r)$, $A \in (M_r)_{\mathrm{sm}}$ and $y \in (S_p)_{\mathrm{sm}}$,*

$$\widehat{T}_{[A]} M_r \cap \widehat{T}_{[y]} S_p = (0).$$

This proves Theorem 3.3.

Corollary 3.12. *For $p \geq r(m+n-r)$*

$$\dim J(M_r, S_p) = \dim M_r + \dim S_p + 1$$

Proof. By Terracini's Lemma,

$$\widehat{T}_{[A+y]} J(M_r, S_p) = \widehat{T}_{[A]} M_r + \widehat{T}_{[y]} S_p.$$

But since $\widehat{T}_{[A]} M_r$ and $\widehat{T}_{[y]} S_p$ intersect only at the origin, the above is a direct sum. Therefore,

$$\dim(\widehat{T}_{[A+y]} J(M_r, S_p)) = \dim(\widehat{T}_{[A]} M_r) + \dim(\widehat{T}_{[y]} S_p).$$

By projectivizing the corollary follows. □

4 Rigidity and Defect

Definition 4.1. *A is r-rigid if $A \notin J(M_r, S_{p(r)+1})$.*

We call A **r-nonrigid** if $A \in J(M_r, S_{p(r)+1})$.

Definition 4.2. *The r-defect of A is $\delta_r(A) = p(r) - p_r(A)$, where $p_r(A)$ is the largest number of zeros such that $A \in J(M_r, S_{p_r(A)}), A \notin J(M_r, S_{p_r(A)+1})$.*

We call A **totally rigid** if $\delta_r(A) = 0$ for all r. We showed earlier that there exist matrices with integer entries that are totally rigid.

4.1 Hadamard Matrix

An $n \times n$ matrix H_n is said to be **Hadamard** if

1. All entries of H_n are from the set $\{-1, 1\}$.
2. $H_n H_n^T = nI_n$.

Note that for each fixed n there are only a finite number of Hadamard matrices. The Hadamard matrix can also be thought of as a point in the **conformal orthogonal group.** This is a matrix group which consists of matrices $A \in \mathbb{C}^{n^2}$ such that $AA^T = \lambda I$ for some $\lambda \in \mathbb{C}\backslash 0$. The conformal orthogonal group, denoted $\mathrm{CO}(n)$, is a smooth $n(n-1)/2+1$ dimensional manifold. Its tangent space at the identity is the Lie algebra of the span of the space of $n \times n$ skew symmetric matrices and the homotheties cId, and is denoted by $\mathfrak{co}(n)$.

The conformal orthogonal group is also an irreducible algebraic variety. It seems easier and more interesting to study the rigidity of a generic matrix in $\mathrm{CO}(n)$ rather than a specific point in it. (We are investigating this problem and will include the details in the final version of this paper.)

4.2 Rigidity and Defect of Hadamard Matrix

Determining $R_{H_n}(n/2)$ is an important open problem. The best known theorem in this regard is due to Kashin and Razborov.

Theorem 4.3. *[6] If less than $\Omega\left(\frac{n^2}{r}\right)$ entries of an $n \times n$ Hadamard matrix H_n are changed (over the reals) then the rank of the resulting matrix remains at least r.*

This result says that to get the rank of H_n down to a constant one has to change $\Omega(n^2)$ entries. Hence establishing a lower bound on the defect. In this section we consider the problem from the other direction, and upper bound the defect of H_n. [2]

Let $H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$. It is easy to see that H_2 is totally rigid.

One way to construct Hadamard matrices of order $n := 2^m$ is via the following $H_{2^m} = H_2 \otimes H_{2^{m-1}}$. This gives

$$H_{2^m} = \begin{pmatrix} H_{2^{m-1}} & H_{2^{m-1}} \\ H_{2^{m-1}} & -H_{2^{m-1}} \end{pmatrix}, \quad for \quad m \geq 2$$

We would like to know the defect of a Hadamard matrix A generated by the above tensoring operation. What is the largest d such that $A \in J(M_r, S_{p(r)+d})$?

Proposition 4.4. *$\delta_1(H_{2^m})$ is at least $2^{2m-1} - 2^{m-1} - 2^m - 1$.*

Proof. It is easy to see that the number of -1's in H_{2^m} (denoted $\nu(m)$) is $2^{2m-1} - 2^{m-1}$. This follows easily from the recursion $\nu(m) = 3\nu(m-1) + (2^{m-1})^2 - \nu(m-1)$, for $n \geq 2$ and $\nu(1) = 1$. Hence if we change all t he -1's of H_{2^m} to $+1$'s, its rank will become one. $\square$

We leave as an interesting open problem to investigate defect of other explicit matrices which have been considered in the rigidity literature.

[2] We just show upper bounds on the defect for the case $r = 1$. Same ideas can be used to obtain upper bounds on $\delta_r(H_n)$. We omit the details.

References

[1] N. Alon. On the rigidity of an Hadamard matrix. unpublished.

[2] B. Codenotti. Matrix rigidity. *Linear Algebra and its Applications*, 304(1–3):181–192, 2000.

[3] B. Codenotti, P. Pudlák, and G. Resta. Some structural properties of low-rank matrices related to computational complexity. *Theoretical Computer Science*, 235(1):89–107, 2000.

[4] J. Friedman. A note on matrix rigidity. *Combinatorica*, 13(2):235–239, 1993.

[5] J. Harris. *Algebraic Geometry: A First Course*. Number 133 in Graduate Texts in Mathematics. Springer-Verlag New York, Inc., 1992.

[6] B. Kashin and A. Razborov. Improved lower bounds on the rigidity of hadamard matrices. *Matematicheskie Zametki*, 63(4):535–540, 1998.

[7] S. V. Lokam. Spectral methods for matrix rigidity with applications to size-depth tradeoffs and communication complexity. In *IEEE Symposium on Foundations of Computer Science*, pages 6–15, 1995.

[8] S. V. Lokam. On the rigidity of Vandermonde matrices. *Theoretical Computer Science*, 237(1–2):477–483, 2000.

[9] A. A. Razborov. On rigid matrices (in russian). unpublished.

[10] M. A. Shokrollahi, D. A. Spielman, and V. Stemann. A remark on matrix rigidity. *Information Processing Letters*, 64(6):283–285, 1997.

[11] V. Strassen. Relative bilinear complexity and matrix multiplication. *J. reine. angew. Math.*, 375/376:406–443, 1987.

[12] L. G. Valiant. Graph-theoretic arguments in low-level complexity. In *Proceedings of the 6th Symposium on Mathematical Foundations of Computer Science*, volume 53, pages 162–176. LNCS, 1977.

[13] F. Zak. Tangents and secants to algebraic varieties. *AMS translations of mathematical monographs*, 1993.

A Straight Line Programs

We restrict ourselves here to the fairly general model of **straight line programs.** A straight line program is a sequence of assignments each of the form $x := f(y, z)$ where f belongs to a set of binary functions. The restriction is that a variable appearing on the left hand side of some assignment cannot appear on the right hand side of another assignment before in the sequence. The variables that never occur on the right left hand side of any assignment are the input variables. The **size** of the program is the number of assignments in the sequence.

The underlying structure of a straight line program is an acyclic directed graph. The **depth** of the program is the length of the longest path in this directed graph.

A **linear program** over a field k is a straight line program with its input set $\{x_1, \ldots, x_n\}$ and the function set $\{f_{\lambda,\mu} | \lambda, \mu \in k\}$.

It is not difficult to see that if a straight line program computes a set of linear forms, then it can be converted to a linear program with at most a constant factor blow up in size.

B Proofs

of Theorem 3.2. Since rank of A is m, there are r linearly independent columns of A. We may assume they are the first r columns. Let U be the subspace of dimension r generated by them. Among the standard basis vectors $e_1, \ldots, e_m$, there are at least $m - r$ vectors which do not lie in U. Let W be the subspace generated by these vectors. Now decompose each column c_j (for $r < j < n$) of A into its projection on to the spaces U and W. Denote these projections by c_j^U and c_j^W respectively. Let B be the $m \times n$ matrix, the first r columns of which are the ones from A and the remaining the vectors c_j^U, for $r < j < n$. Let $C = A - B$. By construction B has rank r and C has at most $(m-r)(n-r)$ non zero entries, hence proving the theorem.

□