INTERPOLATION: AN ALGEBRAIC POINT OF VIEW

Abstract. In one variable, Lagrange and Hermite interpolation are explicit forms of the Chinese remainder theorem in $k[x]$. In several variables we study the evaluation map of polynomials of bounded degree at a scheme of fat points. Lagrange interpolation has maximal generic rank, and we give three proofs. Hermite interpolation of order two, by contrast, can fail for every configuration of points. The Alexander–Hirschowitz theorem lists exactly when it fails, and Terracini's lemma translates the answer into the generic Waring rank of forms. We end with a discussion of the role of the characteristic.

Contents

1. Introduction 2
2. One variable 2
2.1. Lagrange 2
2.2. Hermite 2
3. Several variables: the evaluation at a scheme of fat points 2
4. Several variables: Lagrange interpolation 4
4.1. The problem 4
4.2. First proof: the monomials of the determinant are distinct 4
4.3. Second proof: adding one point at a time 4
4.4. Third proof: an explicit unisolvent set 4
4.5. Why none of this works for Hermite 5
5. Several variables: Hermite interpolation of order two 5
5.1. The problem and its matrix 5
5.2. The simplest failure 6
5.3. A case with an elementary proof 6
6. The Alexander–Hirschowitz theorem 7
7. Terracini's lemma: secant varieties and Waring rank 7
7.1. Veronese and secant varieties 8
7.2. Terracini's lemma 8
7.3. The conormal space to the s-th secant variety of the Veronese variety 8
7.4. Waring rank 9
8. The role of the characteristic 9
9. Conclusion 10
10. Authorship 10
References 11

Date: November 9, 2026.

1. Introduction

The interpolation problem is to find a polynomial with prescribed values, and possibly prescribed derivatives, at prescribed points. In one variable the answer is classical and always positive.

In several variables two new phenomena appear. First, whether a problem can be solved depends on the position of the points. The natural question becomes whether the determinant of the problem, viewed as a polynomial in the coordinates of the points, is nonzero: the problem then has *maximal generic rank*. Second, for derivative conditions even this can fail. The precise answer for first derivatives is the Alexander–Hirschowitz theorem.

Throughout, k is a field. Section 8 collects what depends on the characteristic.

2. One variable

2.1. **Lagrange.** Let $a_1, \dots, a_s \in k$ be distinct and $y_1, \dots, y_s \in k$.

Theorem 2.1 (Lagrange)**.** *There is a unique $p \in k[x]$ with $\deg p \leqslant s-1$ and $p(a_i) = y_i$ for all i, namely*

$$p = \sum_{i=1}^{s} y_i \, L_i, \qquad L_i(x) = \prod_{j \neq i} \frac{x - a_j}{a_i - a_j}.$$

Proof. Since $L_i(a_j) = \delta_{ij}$, the polynomial p has the required values. If p and p' both work, then $p - p'$ has degree $\leqslant s-1$ and s distinct roots, so $p = p'$. □

Equivalently, the Vandermonde matrix $(a_i^j)_{1 \leqslant i \leqslant s,\ 0 \leqslant j \leqslant s-1}$ is invertible, with determinant $\prod_{i<j}(a_j - a_i) \neq 0$.

2.2. **Hermite.** Let $a_1, \dots, a_s \in k$ be distinct and $m_1, \dots, m_s \geqslant 1$. Now prescribe the class of p modulo $(x - a_i)^{m_i}$ rather than just its value at a_i.

Theorem 2.2 (Hermite)**.** *Let $\lambda = \sum_i m_i$. For any $r_i \in k[x]$ there is a unique p with $\deg p \leqslant \lambda - 1$ and $p \equiv r_i \pmod{(x-a_i)^{m_i}}$ for every i. More generally, for every d, the map*

$$\{p : \deg p \leqslant d\} \longrightarrow \prod_i k[x]/(x - a_i)^{m_i}$$

has rank $\min(\lambda, d+1)$.

Proof. Put $Q = \prod_i (x - a_i)^{m_i}$, of degree λ. By the Chinese remainder theorem, $k[x]/(Q) \simeq \prod_i k[x]/(x-a_i)^{m_i}$, and the polynomials of degree $\leqslant \lambda - 1$ form a system of representatives of $k[x]/(Q)$. For general d, the kernel is the set of multiples of Q of degree $\leqslant d$, and counting dimensions gives the rank. □

3. Several variables: the evaluation at a scheme of fat points

Put

$$V_{n,d} = \{p \in k[x_1, \dots, x_n] :\ \deg p \leqslant d\}, \qquad N = \dim V_{n,d} = \binom{n+d}{n}.$$

For $\alpha \in \mathbb{N}^n$ write $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and $|\alpha| = \alpha_1 + \cdots + \alpha_n$, and let $e_1, \dots, e_n$ be the standard basis of $\mathbb{Z}^n$, so $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ with 1 in position j.

Let $P_1, \dots, P_s$ be distinct points of $\mathbb{A}^n$ and $m_1, \dots, m_s \geqslant 1$ integers. For the sake of simplicity we assume that the points P_i are in $\mathbb{A}^n(k)$, and we set $P_i = (P_{i,1}, \dots, P_{i,n})$ and $P_i^\alpha = x^\alpha(P_i)$.

Let $\mathfrak{m}_{P_i}$ be the maximal ideal of P_i and $I_Z = \bigcap_i \mathfrak{m}_{P_i}^{m_i}$. The subscheme $Z = m_1P_1 + \cdots + m_sP_s$ of $\mathbb{A}^n$ defined by I_Z is a *scheme of fat points*. By the Chinese remainder theorem,

$$\mathcal{O}_Z := k[x_1, \ldots, x_n]/I_Z \xrightarrow{\sim} \prod_{i=1}^{s} k[x_1, \ldots, x_n]/\mathfrak{m}_{P_i}^{m_i}, \qquad \lambda(Z) := \dim_k \mathcal{O}_Z = \sum_{i=1}^{s} \binom{n+m_i-1}{n}.$$

The integer $\lambda(Z)$ is called the *length* of Z. The *evaluation map* of Z in degree d is

$$\rho_{Z,d} \colon V_{n,d} \longrightarrow \mathcal{O}_Z, \qquad p \longmapsto (p \bmod \mathfrak{m}_{P_i}^{m_i})_{1 \leqslant i \leqslant s}.$$

The class of p modulo $\mathfrak{m}_{P_i}^{m_i}$ is its Taylor expansion at P_i up to order $m_i - 1$. So $\rho_{Z,d}$ records the values of p when all $m_i = 1$ (Lagrange), and the values and gradients when all $m_i = 2$ (Hermite of order 2). Its kernel is $I_Z \cap V_{n,d}$, and its rank is the value at d of the Hilbert function of Z.

We say that $\rho_{Z,d}$ has *maximal rank* if $\operatorname{rank} \rho_{Z,d} = \min(\lambda(Z), N)$. Either $\rho_{Z,d}$ is surjective, and all data can be interpolated, or it is injective, so an interpolant is unique when it exists. We do not restrict to the square case $\lambda(Z) = N$.

Fix the multiplicities $m_1, \ldots, m_s$. The configurations of s distinct points form the *configuration space*

$$\mathcal{C} = (\mathbb{A}^n)^s \setminus \Delta, \qquad \Delta = \bigcup_{i<i'} \{P_i = P_{i'}\},$$

an open subvariety of the affine space $\mathbb{A}^{ns}$. It is irreducible, and its function field is $k(P_{i,j})$, where the coordinates $P_{i,j}$ are now indeterminates. In the monomial basis, the matrix of $\rho_{Z,d}$ has entries in $k[P_{i,j}]$; its rank over $k(P_{i,j})$ is the *generic rank* of the problem.

Definition 3.1. The problem has *maximal generic rank* if its generic rank is $\min(\lambda(Z), N)$, that is, if some maximal minor of its matrix is a nonzero polynomial in the $P_{i,j}$.

The maximal minors define a closed subvariety $B \subset \mathcal{C}$, the configurations of non-maximal rank. The problem has maximal generic rank if and only if $B \neq \mathcal{C}$, that is, if and only if $\rho_{Z,d}$ has maximal rank on the nonempty Zariski open subset $U = \mathcal{C} \setminus B$. This is a statement about varieties and needs no hypothesis on k. Since the minors have integer coefficients, it depends only on the characteristic of k. More generally, a minor that is nonzero at one configuration, with coordinates in k or in any extension of k, is a nonzero polynomial. So the rank at any configuration is at most the generic rank, and the generic rank is the rank on a nonempty open subset of $\mathcal{C}$.

Whether U has k-rational points is a different question. Over $\mathbb{F}_q$ it may not: a nonzero polynomial can vanish at every point of $\mathbb{F}_q^r$, as $x^q - x$ does. However U has k-rational points as soon as k is large enough, by the following lemma.

Lemma 3.2. *Let $G \in k[y_1, \ldots, y_r]$ have degree at most h in each variable, and let $A \subset k$ with $|A| > h$. If G vanishes on A^r, then $G = 0$.*

Proof. Induction on r. For $r = 1$, a polynomial of degree $\leqslant h$ with more than h roots is zero, as in the uniqueness part of Theorem 2.1. For $r \geqslant 2$, write $G = \sum_{j \leqslant h} c_j(y_1, \ldots, y_{r-1})\, y_r^j$. For $u \in A^{r-1}$, the polynomial $G(u, y_r)$ has more than h roots, so $c_j(u) = 0$ for all j. By induction, every c_j is zero. □

In other words, the grid A^r is unisolvent for polynomials of degree $\leqslant h$ in each variable: this is tensor-product Lagrange interpolation.

In one variable, Theorem 2.2 shows that $\rho_{Z,d}$ has rank $\min(\lambda(Z), d+1)$ for *every* scheme of fat points Z. In several variables this fails, and the question becomes which schemes of fat points have maximal rank.

4. Several variables: Lagrange interpolation

4.1. **The problem.** For $Z = P_1 + \cdots + P_s$, the map $\rho_{Z,d}$ is $p \mapsto (p(P_i))_i$. Its matrix is the $s \times N$ Vandermonde matrix

$$M(P_1, \dots, P_s) = \left(P_i^\alpha\right)_{1 \leqslant i \leqslant s,\ |\alpha| \leqslant d}.$$

Example 4.1. Unlike in one variable, special configurations fail. If $n = d = 2$ and six points lie on a conic $q = 0$, then q is in the kernel, so $\operatorname{rank} M \leqslant 5$.

Theorem 4.2. *Lagrange interpolation at s points in degree d has maximal generic rank: the generic rank of $M(P_1, \dots, P_s)$ is $\min(s, N)$.*

Reduction. We must show that *one* maximal minor is a nonzero polynomial in the $P_{i,j}$. If $s \leqslant N$, choose s distinct exponents $\alpha_1, \dots, \alpha_s$; if $s \geqslant N$, take the first N points and all exponents. In both cases we will show that

$$D = \det\left(P_i^{\alpha_j}\right)_{1 \leqslant i,j \leqslant t}, \qquad t = \min(s, N), \quad \alpha_1, \dots, \alpha_t \text{ distinct},$$

is not the zero polynomial. We give three arguments. The first one shows that $D \neq 0$ for every choice of the exponents. The other two exhibit a configuration, with coordinates in k or in an extension of k, at which the matrix has maximal rank, so that some maximal minor does not vanish there; this suffices, since a polynomial that does not vanish at one point is nonzero. □

4.2. **First proof: the monomials of the determinant are distinct.**

Proof. By the Leibniz formula,

$$D = \sum_{\sigma \in \mathfrak{S}_t} \operatorname{sgn}(\sigma) \prod_{i=1}^{t} P_i^{\alpha_{\sigma(i)}}.$$

The part of the σ-term in the variables $(P_{i,1}, \dots, P_{i,n})$ of the i-th point is exactly $P_i^{\alpha_{\sigma(i)}}$. So the monomial determines each $\alpha_{\sigma(i)}$, hence σ, because the α_j are distinct. The $t!$ terms are distinct monomials with coefficients ± 1; nothing cancels, and $D \neq 0$. □

4.3. **Second proof: adding one point at a time.**

Proof. Let $W_j = \{p \in V_{n,d} : \ p(P_1) = \cdots = p(P_j) = 0\}$, so $\dim W_j \geqslant N - j$. Suppose $\dim W_j = N - j$ with $j < N$. Then W_j contains some $p \neq 0$. By Lemma 3.2, applied over an infinite extension of k (or over k itself if $|k| > d$), there is P_{j+1} with $p(P_{j+1}) \neq 0$, and then $\dim W_{j+1} = N - j - 1$. By induction there is a configuration of $\min(s, N)$ points imposing independent conditions, so a maximal minor is nonzero. □

4.4. **Third proof: an explicit unisolvent set.**

Proposition 4.3. *Assume* $\operatorname{char} k = 0$ *or* $\operatorname{char} k > d$. *The principal lattice*

$$\Lambda_{n,d} = \{\alpha \in \mathbb{N}^n : \ |\alpha| \leqslant d\} \subset \mathbb{A}^n(k),$$

which has N points, is unisolvent for $V_{n,d}$: the only $p \in V_{n,d}$ vanishing on $\Lambda_{n,d}$ is $p = 0$.

Proof. Induction on $n + d$. If $d = 0$, p is a constant vanishing at the origin. If $n = 1$, p has degree $\leqslant d$ and vanishes at $0, 1, \dots, d$, which are distinct in k by hypothesis. Let $n \geqslant 2$ and $d \geqslant 1$.

(i) The points with $\alpha_n = 0$ form a copy of $\Lambda_{n-1,d}$ in the hyperplane $x_n = 0$. By induction p vanishes on this hyperplane, so $p = x_n\, g$ with $\deg g \leqslant d - 1$.

(ii) The points with $\alpha_n \neq 0$ form $e_n + \Lambda_{n,d-1}$. There $x_n = \alpha_n$ with $1 \leqslant \alpha_n \leqslant d$, which is nonzero in k. So g vanishes on $e_n + \Lambda_{n,d-1}$, and $g(x + e_n)$ vanishes on $\Lambda_{n,d-1}$. By induction $g = 0$. □

Remark 4.4. $\Lambda_{n,d}$ is the classical node set of Lagrange finite elements on a simplex. On a simplex with vertices $v_0, \dots, v_n$, the nodes of the element of degree d are the points $\sum_i (\beta_i/d)\, v_i$ with $\beta_i \in \mathbb{N}$ and $\sum_i \beta_i = d$; for the unit simplex this is $\frac{1}{d}\Lambda_{n,d}$, and in general it is an affine image of $\Lambda_{n,d}$. Since $V_{n,d}$ is stable under affine changes of coordinates, Proposition 4.3 is exactly the unisolvence result that makes this element well defined: this is [6, Theorem 2.2.1]. These nodes are called the *principal lattice* of order d of the simplex [6, Exercise 2.2.2], after Nicolaides [8].

4.5. **Why none of this works for Hermite.**

- *First proof.* In the Hermite matrix of Section 5, the $n+1$ rows attached to one point involve the same variables. Different assignments of columns inside a block give the same monomial, and terms can cancel. They do cancel in Example 5.1.
- *Second proof.* A double point imposes $n+1$ conditions at once. Choosing the new point off the zero set of one polynomial guarantees only one new condition.
- *Third proof.* Restricting to a hyperplane keeps the tangential derivatives but loses the normal one, so the induction does not close. Hirschowitz's *méthode d'Horace* is a refined version of this step that keeps track of the lost information.

5. Several variables: Hermite interpolation of order two

5.1. **The problem and its matrix.** Take $Z = 2P_1 + \cdots + 2P_s$, so $\lambda(Z) = s(n+1)$. Since $p \bmod \mathfrak{m}_P^2$ is the pair $(p(P), \nabla p(P))$, the evaluation map is

$$\rho_{Z,d} \colon V_{n,d} \to k^{s(n+1)}, \qquad p \mapsto \Big(p(P_i),\ \tfrac{\partial p}{\partial x_1}(P_i), \dots, \tfrac{\partial p}{\partial x_n}(P_i)\Big)_{1 \leqslant i \leqslant s}.$$

It has maximal rank if $\operatorname{rank} \rho_{Z,d} = \min\big(s(n+1), N\big)$. Its kernel consists of the polynomials whose zero set is singular at every P_i. Each point P_i contributes a block of $n+1$ rows; in the column indexed by α the block is

$$\big(P_i^{\alpha},\ \alpha_1 P_i^{\alpha - e_1},\ \dots,\ \alpha_n P_i^{\alpha - e_n}\big)^T,$$

where a term with a negative exponent is zero.

For $n = 1$, Theorem 2.2 shows that $\rho_{Z,d}$ has maximal rank for *every* choice of distinct nodes. Several variables are genuinely different.

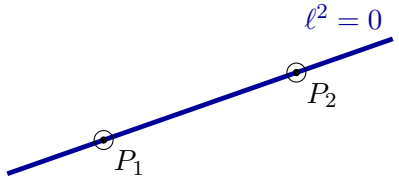
5.2. **The simplest failure.**

Example 5.1. Take $n = d = 2$ and the points $P_1 = (0,0)$ and $P_2 = (1,0)$. With the columns indexed by the monomials and the rows by the conditions at each point, the matrix of $\rho_{Z,2}$ is

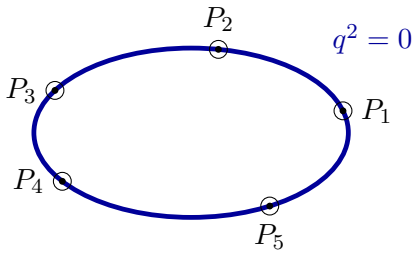
$$\begin{array}{r} \\ p(P_1) \\ \partial_x p(P_1) \\ \partial_y p(P_1) \\ p(P_2) \\ \partial_x p(P_2) \\ \partial_y p(P_2) \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & x & y & x^2 & xy & y^2 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

The count gives $6 = N$ conditions, so we expect an invertible 6×6 matrix. But the last column vanishes, and the rank is 5.

This happens for *every* pair of distinct points. Let ℓ be an affine linear form vanishing on the line through P_1 and P_2. Then ℓ^2 has degree 2, and $\nabla(\ell^2) = 2\ell\,\nabla\ell$ vanishes wherever ℓ does. So $0 \neq \ell^2 \in \ker \rho_{Z,2}$, so $\rho_{Z,2}$ never has maximal rank.



Example 5.2. Let $n = 2$, $d = 4$, $s = 5$. The count gives $5 \cdot 3 = 15 = N$ conditions, so we expect an invertible 15×15 matrix. Let q be the equation of the conic through the five points. Then q^2 has degree 4 and $\nabla(q^2) = 2q\,\nabla q$ vanishes on the conic, so $0 \neq q^2 \in \ker \rho_{Z,4}$, and the rank is at most 14.

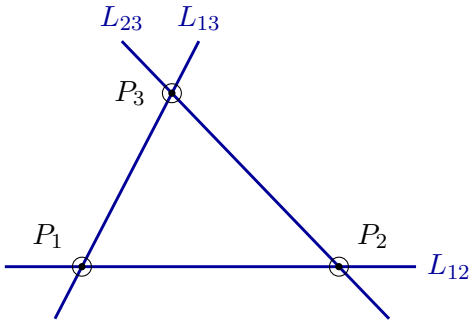


5.3. **A case with an elementary proof.**

Proposition 5.3. *Let $n = 2$, $d = 3$, $s = 3$. For any three non-collinear points the problem has rank 9, which is maximal since $9 < N = 10$. The kernel is spanned by the product of the three lines joining the points.*

Proof. Let $F \in V_{2,3}$ vanish to order 2 at non-collinear points P_1, P_2, P_3. For $i \neq j$, let L_{ij} be the line through P_i and P_j, and ℓ_{ij} an equation of L_{ij}. The restriction of F to L_{ij} has degree at most 3 and vanishes with multiplicity at least 2 at P_i and at P_j, so it is zero. Hence ℓ_{ij} divides F. The three lines are distinct, so $\ell_{12}, \ell_{13}, \ell_{23}$ are pairwise coprime and their product divides F. Comparing degrees, $F = c\,\ell_{12}\ell_{13}\ell_{23}$. Conversely, $\ell_{12}\ell_{13}\ell_{23}$ vanishes to order 2 at each P_i, because

two of its factors vanish there; the corresponding lines are distinct, so they meet transversally at P_i, which is a node of the cubic. So $\dim\ker\rho_{Z,3}=1$ and $\operatorname{rank}\rho_{Z,3}=9$. □



Remark 5.4. The proof uses only restriction to a line and counting roots with multiplicity. These are the basic tools of the degeneration proofs of the Alexander–Hirschowitz theorem.

6. The Alexander–Hirschowitz theorem

Theorem 6.1 (Alexander–Hirschowitz [3])**.** *Let k be any field, and let $Z=2P_1+\cdots+2P_s$ be s double points of $\mathbb{P}^n$. Hermite interpolation of order 2 on Z in degree d has maximal generic rank $\min\bigl(s(n+1),\binom{n+d}{n}\bigr)$, except in the following cases:*

n	d	s	*explanation*
any	2	$2\leqslant s\leqslant n$	*quadrics singular along the span of the points (Ex. 5.1)*
2	4	5	*square of the conic through the* 5 *points (Ex. 5.2)*
3	4	9	*square of the quadric through the* 9 *points*
4	4	14	*square of the quadric through the* 14 *points*
4	3	7	*secant variety of the rational normal quartic through the* 7 *points*

Remark 6.2 (About the proof). Bronowski claimed a proof in 1931, which was incomplete. Alexander and Hirschowitz settled the problem in a series of papers ending with [3]. Their method, the *méthode d'Horace*, specialises some of the points onto a hyperplane and inducts on n and d; its differential version handles the normal derivatives lost in §4.5.

- Chandler [5] simplified the argument; she treats cubics over any infinite field.
- Brambilla and Ottaviani [4] give a self-contained exposition.
- Postinghel [9] gives a proof by degeneration.

Remark 6.3 (The theorem lives over the prime field). The matrix of $\rho_{Z,d}$, with the coordinates of the points as indeterminates, has entries in $\mathbb{Z}[P_{i,j}]$: the rows attached to P_i are P_i^{α} and $\alpha_j\,P_i^{\alpha-e_j}$. Its generic rank over k is its rank over $k_0(P_{i,j})$, where k_0 is the prime field of k, since $k_0[P_{i,j}]\subset k[P_{i,j}]$ and a polynomial is nonzero in one ring if and only if it is in the other. So Theorem 6.1 is a statement about $\mathbb{Q}$ and the fields $\mathbb{F}_p$: it depends only on the characteristic, and the size of k plays no role in it. The size of k only matters for the existence of configurations with coordinates in k (Lemma 3.2).

7. Terracini's lemma: secant varieties and Waring rank

In this section k **is algebraically closed of characteristic 0**. Let

$$S_{n,d}=\Bigl\{\sum_{|\beta|=d} c_\beta\,x_0^{\beta_0}x_1^{\beta_1}\cdots x_n^{\beta_n}\ :\ c_\beta\in k\Bigr\}\subset k[x_0,\dots,x_n]$$

be the space of forms (homogeneous polynomials) of degree d in the $n+1$ variables $x_0, \dots, x_n$, where β runs over $\mathbb{N}^{n+1}$. Setting $x_0 = 1$ identifies $S_{n,d}$ with $V_{n,d}$, so $\dim S_{n,d} = N = \binom{n+d}{n}$. In particular $S_{n,1}$ is the space of linear forms.

7.1. **Veronese and secant varieties.** The *Veronese variety* $\mathcal{V} = \{[\ell^d] : \ \ell \in S_{n,1}\backslash\{0\}\} \subset \mathbb{P}(S_{n,d})$ is the image of $\mathbb{P}^n$ under the d-th Veronese embedding, so $\dim \mathcal{V} = n$. The *s-th secant variety* $\sigma_s(\mathcal{V})$ is the closure of the union of the spans of s points of $\mathcal{V}$. Since k is algebraically closed, scalars can be absorbed into d-th powers, and

$$\sigma_s(\mathcal{V}) = \overline{\{[\ell_1^d + \cdots + \ell_s^d]\}}.$$

Counting parameters gives $\dim \sigma_s(\mathcal{V}) \leqslant \min\big(s(n+1)-1,\ N-1\big)$, the *expected dimension*. When the inequality is strict, $\sigma_s(\mathcal{V})$ is said to be *defective*.

7.2. **Terracini's lemma.** Consider $\varphi\colon S_{n,1}^s \to S_{n,d}$, $(\ell_1, \dots, \ell_s) \mapsto \sum_i \ell_i^d$. The closure of its image is the affine cone $\widehat{\sigma_s(\mathcal{V})}$, and its differential is

$$D\varphi_{(\ell_i)}(\mu_1, \dots, \mu_s) = d\sum_{i=1}^{s} \ell_i^{d-1}\mu_i.$$

Lemma 7.1 (Terracini [10])**.** *For $(\ell_1, \dots, \ell_s)$ in a nonempty open subset of $S_{n,1}^s$, the form $\ell_1^d + \cdots + \ell_s^d$ is a smooth point of $\widehat{\sigma_s(\mathcal{V})}$, and the tangent space there is*

$$T_{\ell_1^d+\cdots+\ell_s^d}\widehat{\sigma_s(\mathcal{V})} = \ell_1^{d-1}S_{n,1} + \cdots + \ell_s^{d-1}S_{n,1}.$$

In particular $\dim \widehat{\sigma_s(\mathcal{V})} = \dim\big(\ell_1^{d-1}S_{n,1} + \cdots + \ell_s^{d-1}S_{n,1}\big)$.

Proof. φ is a dominant morphism from the smooth irreducible variety $S_{n,1}^s$ onto the irreducible variety $\widehat{\sigma_s(\mathcal{V})}$. In characteristic 0, generic smoothness gives a nonempty open subset W of the smooth locus of $\widehat{\sigma_s(\mathcal{V})}$ such that φ is smooth over W. At every point of $\varphi^{-1}(W)$, which is a nonempty open subset of $S_{n,1}^s$, the differential $D\varphi$ is therefore surjective onto the tangent space of $\widehat{\sigma_s(\mathcal{V})}$. Its image is $\ell_1^{d-1}S_{n,1} + \cdots + \ell_s^{d-1}S_{n,1}$, since $d \neq 0$ in k. □

7.3. **The conormal space to the s-th secant variety of the Veronese variety.** Let $T = k[y_0, \dots, y_n]$ act on $S = k[x_0, \dots, x_n]$, with y_j acting as $\partial/\partial x_j$. For $F \in T_{n,d}$ and $G \in S_{n,d}$, $F(\partial)G$ is a scalar. Since $\partial^\alpha x^\beta = \alpha!\,\delta_{\alpha\beta}$ for $|\alpha| = |\beta| = d$, this pairing is perfect in characteristic 0.

Lemma 7.2. *Let $\ell = \sum_j a_j x_j$, $\mu = \sum_j b_j x_j$ and $a = (a_0, \dots, a_n)$. For $F \in T_{n,d}$,*

$$F(\partial)\,\ell^d = d!\,F(a), \qquad F(\partial)(\ell^{d-1}\mu) = (d-1)!\sum_j b_j\,\frac{\partial F}{\partial y_j}(a).$$

Proof. For $F = y^\alpha$ with $|\alpha| = d$ we have $\partial^\alpha \ell^d = d!\,a^\alpha$, which gives the first identity by linearity. Apply it to $(\ell + t\mu)^d$ and differentiate at $t = 0$: the left side gives $d\,F(\partial)(\ell^{d-1}\mu)$, the right side $d!\sum_j b_j\,\partial_j F(a)$. □

Proposition 7.3. *Let $\ell_i = \sum_j P_{i,j}x_j$, with corresponding points $P_i \in \mathbb{P}^n$. Then*

$$\big(\ell_1^{d-1}S_{n,1} + \cdots + \ell_s^{d-1}S_{n,1}\big)^\perp = \{F \in T_{n,d} : \ F \text{ singular at } P_1, \dots, P_s\}.$$

Consequently $\dim \widehat{\sigma_s(\mathcal{V})}$ *equals the generic rank of Hermite interpolation of order 2 at s double points of $\mathbb{P}^n$.*

Proof. By Lemma 7.2, F annihilates $\ell_i^{d-1}\mu$ for all $\mu \in S_{n,1}$ if and only if $\nabla F(P_i) = 0$. Euler's formula $d\,F(P_i) = \sum_j P_{i,j}\,\partial_j F(P_i)$ then gives $F(P_i) = 0$ too. The second statement follows from Lemma 7.1 and the perfectness of the pairing. □

Since the pairing identifies $T_{n,d}$ with the dual of $S_{n,d}$, the left-hand side of Proposition 7.3 is the conormal space of the span of the tangent spaces to the cone over $\mathcal{V}$ at $\ell_1^d, \dots, \ell_s^d$. By Lemma 7.1, for $(\ell_1, \dots, \ell_s)$ in a nonempty open subset of $S_{n,1}^s$, this span is the tangent space to the cone over $\sigma_s(\mathcal{V})$ at $\ell_1^d + \cdots + \ell_s^d$. So the conormal space to $\sigma_s(\mathcal{V})$ at such a point is the space of forms singular at $P_1, \dots, P_s$:

$$N^*_{\sigma_s(\mathcal{V})} = \left(I_{2P_1+\cdots+2P_s}\right)_d.$$

Corollary 7.4. *$\sigma_s(\mathcal{V})$ has the expected dimension if and only if Hermite interpolation of order 2 at s double points has maximal generic rank in degree d. Hence the Alexander–Hirschowitz theorem is equivalent to the statement that all secant varieties of all Veronese varieties have the expected dimension, except in the five families of Theorem 6.1.*

7.4. **Waring rank.** The *generic Waring rank* $r_{\mathrm{gen}}(n,d)$ is the least s with $\sigma_s(\mathcal{V}) = \mathbb{P}(S_{n,d})$: the number of d-th powers of linear forms needed to write every form of degree d in $n+1$ variables outside a proper closed subset. By the corollary,

$$r_{\mathrm{gen}}(n,d) = \left\lceil \frac{1}{n+1}\binom{n+d}{n} \right\rceil$$

except in the following cases, where it is larger: by $\lfloor n/2 \rfloor$ for quadrics, and by one in the four sporadic cases.

forms	expected value	generic rank
quadrics in $n+1 \geqslant 3$ variables	$\lceil (n+2)/2 \rceil$	$n+1$
ternary quartics	5	6
quaternary quartics	9	10
quinary quartics	14	15
quinary cubics	7	8

Remark 7.5. The quadric row is elementary. A sum of s squares of linear forms is a quadratic form of rank $\leqslant s$, so $\sigma_s(\mathcal{V})$ is the variety of quadrics of rank $\leqslant s$, and a quadric of maximal rank needs $n+1$ squares. This is the Waring counterpart of the ℓ^2 phenomenon in Example 5.1.

8. The role of the characteristic

In one variable everything is characteristic free once Hermite conditions are written as congruences modulo $(x-a_i)^{m_i}$. The derivative formulation needs care: in characteristic p, the p-th derivative of every polynomial vanishes, so one uses Hasse derivatives $D^{(r)}x^j = \binom{j}{r}x^{j-r}$ in place of $f^{(r)}/r!$. For first derivatives, as in Sections 5–6, no such issue arises, and the Alexander–Hirschowitz theorem holds in every characteristic (Remark 6.3).

Table 1 summarises how the multivariate results depend on the characteristic.

Example 8.1. Let $\operatorname{char} k = p$ and $d = p$. Then

$$\Big(\sum_j a_j x_j\Big)^p = \sum_j a_j^p\, x_j^p,$$

statement	valid over	comment
Thm. 4.2, first proof	any field	coefficients ± 1, so $D \neq 0$ in every characteristic; k-points exist if k is large enough (Lemma 3.2)
Thm. 4.2, second proof	any field	needs a point off a hypersurface of degree $\leqslant d$, over k if $\lvert k\rvert > d$ (Lemma 3.2)
Prop. 4.3 (lattice)	$\operatorname{char} k = 0$ or $> d$	for $\operatorname{char} k = p \leqslant d$ the integers $0, \dots, d$ are not distinct in k
order 2 = value + gradient	any field	first-order Taylor expansion involves no division
Ex. 5.1, Ex. 5.2, Prop. 5.3	any field	restriction to lines and root counting only
Thm. 6.1	any field	a statement over the prime field (Remark 6.3); proved by the authors in every characteristic [2, §1.5], [1, Th. 0.1]
Lemma 7.1, Prop. 7.3	$\operatorname{char} k = 0$; secant statement also for $\operatorname{char} k > d > 2$ [7]	uses generic smoothness and $d! \neq 0$; $\operatorname{char} k = d$ open according to [4]

TABLE 1. Dependence on the characteristic.

so all p-th powers of linear forms lie in the $(n+1)$-dimensional span of the x_j^p. A form of degree p outside this span is not a sum of p-th powers at all, and the Waring side of Section 7 collapses. The double-point interpolation problem still makes perfect sense. The equivalence between the two is a characteristic-zero phenomenon.

9. Conclusion

In one variable, interpolation is the Chinese remainder theorem: the evaluation map at any scheme of fat points has maximal rank, with explicit formulas. In several variables, the Chinese remainder theorem still describes the target $\mathcal{O}_Z$, but polynomials of bounded degree no longer map onto it as well as possible. For values, the determinant remains a nonzero polynomial in the points: the problem has maximal generic rank. For derivatives even this can fail, and the Alexander–Hirschowitz theorem says exactly when. Through Terracini's lemma, the same theorem computes the generic number of powers of linear forms needed to write a polynomial.

10. Authorship

These notes were written by Claude, an AI model made by Anthropic, under the supervision of Jean-Marc Couveignes, in a single working session from 4 to 6 November 2026.

- *Model.* The session was configured for the model `claude-opus-5-5` (Claude Opus 5.5); the model that served a given turn may occasionally have differed. The effort level selected in the interface, as reported by Jean-Marc Couveignes at the end of the session, was *Medium*.
- *Usage.* From the first logged turn on, the session made about 350 calls to the model, which read about 112 million tokens of context from cache, wrote about 2.1 million tokens to cache, and produced about 0.26 million output tokens. Weighting these by their approximate relative cost gives about 17 million token-equivalents, of which about

43% went to re-reading the model's own instructions, 33% to working with files (writing and compiling LaTeX, checking the rendered pages, reading articles), 21% to writing and reasoning, and 3% to web searches. No other agents were used.

- *Cost.* Six hours of full-time work for a French professor nearing the end of his career, that is about 450 euros. Plus the cost of AI. At Anthropic's published API prices for this model (about \$4 per million input tokens, \$20 per million output tokens, \$0.20 per million tokens read from cache and \$8 per million tokens written to a one-hour cache), these figures amount to about \$45, that is about 40 euros at the rate of late October 2026 ($1\,\mathrm{EUR} \approx 1.12\,\mathrm{USD}$). This is a list-price equivalent: the amount actually paid depends on the subscription under which the session ran.
- *Carbon footprint.* Anthropic publishes no energy or emission figures per token, so only a very rough order of magnitude can be given. Using the API price as a proxy for computation, the session is equivalent to about 2 million output tokens. Published external estimates for large models range from well under 1 Wh to about 20 Wh per thousand output tokens, and a typical grid emits about 0.4 kg of CO_2 per kWh. This gives between about 1 and 50 kWh, that is between about 0.5 and 20 kg of CO_2-equivalent, the lower end being more consistent with recent disclosures by other providers. These numbers are indicative only. These are the costs according to the price list, not the production costs.
- *Division of labour.* Jean-Marc Couveignes chose the subject, the outline and the material kept in the notes, and directed every revision: the definitions (the evaluation map at a scheme of fat points, the configuration space, maximal generic rank), the notation, the terminology, the choice of examples and the layout. Claude drafted the text, the proofs, the figures and the tables, compiled the document, and checked the references it could reach.
- *Sources.* The statements on the characteristic rely on the articles of Hirschowitz [2] and Alexander–Hirschowitz [1], and the finite-element remark on Ciarlet [6], all read in copies supplied by him. Some other bibliographic details were taken from secondary sources and should be checked before citation. The exceptions of Theorem 6.1 were also checked by a computer experiment over finite fields of characteristic 2, 3, 5 and 7.

Jean-Marc Couveignes is responsible for the final content of these notes.

References

[1] J. Alexander and A. Hirschowitz, *La méthode d'Horace éclatée: application à l'interpolation en degré quatre*, Invent. Math. **107** (1992), 585–602.

[2] A. Hirschowitz, *La méthode d'Horace pour l'interpolation à plusieurs variables*, Manuscripta Math. **50** (1985), 337–388.

[3] J. Alexander and A. Hirschowitz, *Polynomial interpolation in several variables*, J. Algebraic Geom. **4** (1995), 201–222.

[4] M. C. Brambilla and G. Ottaviani, *On the Alexander–Hirschowitz theorem*, J. Pure Appl. Algebra **212** (2008), 1229–1251. arXiv:math/0701409.

[5] K. A. Chandler, *Linear systems of cubics singular at general points of projective space*, Compositio Math. **134** (2002), 269–282.

[6] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland, 1978; reprinted as Classics in Applied Mathematics **40**, SIAM, 2002.

[7] A. Iarrobino and V. Kanev, *Power sums, Gorenstein algebras, and determinantal loci*, Lecture Notes in Math. **1721**, Springer, 1999.

[8] R. A. Nicolaides, *On a class of finite elements generated by Lagrange interpolation*, SIAM J. Numer. Anal. **9** (1972), 435–445.

[9] E. Postinghel, *A new proof of the Alexander–Hirschowitz interpolation theorem*, Ann. Mat. Pura Appl. **191** (2012), 77–94. arXiv:1003.0323.

[10] A. Terracini, *Sulle V_k per cui la varietà degli S_h $(h+1)$-seganti ha dimensione minore dell'ordinario*, Rend. Circ. Mat. Palermo **31** (1911), 392–396.