Baby Alg. Geom. defeats Landau–Siegel zeroes

Daniel Rui 10/19/26-10/20/26

Many expositions of the recent Landau-Siegel zero result have already been written: Banks, Kerr, Green, Tao, Barreto, etc.

I for some reason find the approach in Kerr more "tangible" than the derivative/multiplicity approach in Green, Tao, etc. and I think this makes a great opportunity to learn/exposit some basic algebraic geometry (of a combinatorial flavor).

The proof is essentially the one found in Kerr, but Kerr's Masser theorem blackbox is replaced by a beautiful "sumset to subspace"/"Marton conjecture"-like tool; we present this tool in a self-contained manner. The article is therefore self-contained, modulo undergraduate level math (especially undergraduate abstract algebra and linear algebra). As Green wrote, "The proof was obtained by a 'vibemathing' session with ChatGPT 6" and "No human originality is claimed" — merely human taste.

I consider this Landau-Siegel breakthrough as another entry in the We Suck At The Polynomial Method Cinematic Universe, cf. Terry Tao's hilarious-in-retrospect https://mathoverflow.net/questions/43317/how-to-recognise-that-the-polynomial-method-might-work

> My feeling is that it may be premature to declare what types of problems look tractable with respect to the polynomial method. For instance, the idea of using the polynomial method to attack the finite field Kakeya problem, while "obvious" in retrospect, was certainly a big shock to many of us working on the Kakeya problem at the time Dvir's argument came out.
>
> In contrast, the question of getting a non-trivial Szemeredi-Trotter or sum-product theorem in finite fields, while closely related to the Kakeya problem (see e.g. my paper with Bourgain and Katz on this topic), has so far resisted all attempts at a polynomial method proof.

(Narrator voice: the capset problem was solved (Ellenberg–Gijswit) by the polynomial method 6 years after Terry wrote this.)

Theorem 0.1: Main result

Let $q \in \mathbb{Z}$ be a fundamental discriminant, and $\chi(p) := (\frac{q}{p})$ be the Legendre symbol. For any sufficiently large $|q|$, we have that $X \geqslant |q|^{13}$ implies

$$\sum_{\substack{p \leqslant X \\ \chi(p)=-1}} \frac{\log p}{p} \leqslant \frac{12}{13} \log X.$$

This is the form of "Landau Siegel" Green uses.

Parameters:

- q fundamental discriminant
- X large real number
- $N := N(X) := \lceil X^{10/13} \rceil$
- $D := D(X) := \lfloor \frac{1}{2}(\frac{N-1}{3})^{4/3} \rfloor$, so then $D \asymp X^{40/39} \gg X$ (equivalently $\frac{X}{D} = O(X^{-1/39}) \ll 1$). For $|q|$ sufficiently large, $D > X$.
- $S := S(X) := \{p : 32 < p \leqslant X, \chi(p) = -1\}$ set of "bad primes". By previous bullet all such $p < D$.
- Later (5) we will have parameters $B := 8N\sqrt{|q|}$, $H := \frac{P}{4M}$. Thm. 0.2 has a $T := \lfloor \frac{N-1}{3} \rfloor + 1$.
- We will show (Theorem 0.3)
$$\sum_{p \in S} \frac{\log p}{p} \leqslant \frac{9}{8} \log(8N\sqrt{|q|}).$$
By previous bullet points $N \leqslant 2X^{10/13}$ and $|q| \leqslant X^{1/13}$, an easy calculation reveals the RHS to be $\leqslant \frac{9}{8}(\frac{10}{13} + \frac{1}{26}) \log X + O(1) = \frac{189}{208} \log X + O(1)$. The contribution from primes $p \leqslant 32$ is absorbed into $O(1)$, and $\frac{189}{208} < \frac{192}{208} =: \frac{12}{13}$ completes the deduction of Theorem 0.1 from Theorem 0.3.

0.1 Grids

- Fundamental discriminants $q \neq 1, 8$ in particular satisfy $1, \sqrt{2}, \sqrt{q}, \sqrt{2q}$ are $\mathbb{Q}$-LI.
- Let $R \coloneqq \mathbb{Z}[\sqrt{q}, \sqrt{2}]$. By the previous bullet, all elements have a unique representation as a $\mathbb{Z}$-linear combination of $1, \sqrt{2}, \sqrt{q}, \sqrt{2q}$.
- Let $\theta \coloneqq \theta_{a,b,c,d} \coloneqq (a, b, c, d) \cdot (1, \sqrt{q}, \sqrt{2}, \sqrt{2q})$ and let $\Theta \coloneqq \{\theta_{a,b,c,d} : a, b, c, d \in [0..N)\}$ (cardinality N^4, by $\mathbb{Q}$-LI).
- We will later motivate the definition of the grid (also cardinality N^4)

$$G \coloneqq \{(\theta, \sigma\theta, \sigma\tau\theta) \in \mathbb{C}^3 : \theta \in \Theta\}.$$

We already see that all grid points $\in G$ have absolute value

$$\leqslant \max\{|\theta| : \theta \in \Theta\} + \max\{|\sigma\theta|\} + \max\{|\sigma\tau\theta|\} \leqslant 3N \max\{1, \sqrt{|q|}, \sqrt{2}, \sqrt{2|q|}\} \leqslant 8N\sqrt{|q|}. \quad (1)$$

Goal 1: Polynomial method

I want to force a (non-zero) polynomial $F(X, Y, Z)$ to vanish at every grid point $\in G$.
The key trick is to show "sparseness" of residue patterns (of values of polynomials f on G) mod P (for some large P), use "pigeonhole and subtraction" to get a polynomial $F = f - g$ that takes values divisible by P on G but has small coefficients, and use essentially the "discreteness phenomenon" that the only multiple of P in $[0, P)$ is 0.

0.2 Arithmetic (mod p) Compression by Frobenius

- For any θ, and odd prime p with $\chi(p) \coloneqq (\frac{q}{p}) = -1$, binomial expansion ("Freshman's dream") yields

$$\theta^p = \sum_{\mathbb{N} \ni i+j+k+l=p} \frac{p!}{i!j!k!l!} \sqrt{q}^j \sqrt{2}^k \sqrt{2q}^l,$$

which mod p (i.e. in the quotient R/pR) simplifies to $1 + \sqrt{q}^p + \sqrt{2}^p + \sqrt{2q}^p$ (all terms with $\max\{i, j, k, l\} < p$ die), which we can further massage to

$$\begin{aligned} \theta^p \pmod p &= 1 + (q^{\frac{p-1}{2}})\sqrt{q} + (2^{\frac{p-1}{2}})\sqrt{2} + ((2q)^{\frac{p-1}{2}})\sqrt{2q} \pmod p \\ &= 1 - (\tfrac{q}{p})\sqrt{q} + (\tfrac{2}{p})\sqrt{2} + (\tfrac{q}{p})(\tfrac{2}{p})\sqrt{2q} \pmod p \\ &= \begin{cases} \sigma\theta & \text{if } (\frac{2}{p}) = 1 \\ \sigma\tau\theta & \text{if } (\frac{2}{p}) = -1 \end{cases} \pmod p \end{aligned},$$

where the 2nd equality comes from Euler's criterion, and where $\sigma : R \to R$ flips the sign of $\sqrt{q}$, and $\tau : R \to R$ flips the sign of $\sqrt{2}$.

- Motivated by the previous "Freshman dream"/"Frobenius" observation, we define the grid (also cardinality N^4)

$$G \coloneqq \{\theta, \sigma\theta, \sigma\tau\theta) \in \mathbb{C}^3 : \theta \in \Theta\},$$

so the previous bullet can be said as: every grid point $\in G$ satisfies $X^p - U = 0 \bmod p$, where we take $U \coloneqq Y$ if p satisfies $(\frac{2}{p}) = 1$, and $U \coloneqq Z$ otherwise.

Definition 0.1: Monomials and their statistics

Fix $D \geqslant 0$, and let $\mathscr{M} \subseteq \mathbb{C}[x,y,z]$ be the set of monomial $M_{i,j,k} \coloneqq x^i y^j z^k$ with "weighted degree" $i + 32(j+k) \leqslant D$. (Picture slicing the corner off the cube $\mathbb{N}^3$ to form the "simplex" $\mathscr{M}$.)
We define $M \coloneqq |\mathscr{M}|$, $W \coloneqq \sum_{(i,j,k)\in\mathscr{M}}(i+j+k)$, and $E_p \coloneqq \sum_{(i,j,k)\in\mathscr{M}} \lfloor \frac{i}{p} \rfloor$.
Define $\mathcal{L}_D \coloneqq \mathrm{span}_R\{\mathscr{M}\} \subseteq \mathbb{C}[x,y,z]$.

- Fix a prime $p > 32$, and define an operation on monomials $M_{i,j,k} \coloneqq X^i Y^j Z^k$ to be: write $i = pe + r$ with $r \in [0..p)$ (i.e. $e \coloneqq \lfloor \frac{i}{p} \rfloor$), and map

$$M_{i,j,k} \coloneqq X^i Y^j Z^k \quad \mapsto \quad M^{(p)}_{i,j,k} \coloneqq X^r (X^p - U)^e Y^j Z^k.$$

  The main benefit of this is that $M^{(p)}_{i,j,k}$ evaluated any grid point $\in G$ is in $p^e R$, i.e. the $\bullet^{(p)}$ operation improves a monomial's p-adic valuation (by $e \geqslant 0$) at *all* grid points $\in G$.

  Observe that the leading monomial of $\breve{M}_{i,j,k}$ is exactly the original $M_{i,j,k}$, and every other term/monomial has strictly smaller X-degree and weighted degree $\leqslant i + 32(j+k)$ because $p > 32$. By "Gaussian elimination" (in X-degree), we see that $\mathcal{L}_D \coloneqq \mathrm{span}_R\{\mathscr{M}\} = \mathrm{span}_R\{\mathscr{M}^{(p)}\}$.

Definition 0.2: Residue pattern

For a polynomial $f \in \mathcal{L}_D$, we define the "residue pattern mod (prime) P" $\mathsf{ev}_P(f)$ to be the function $[\vec{x} \mapsto f(\vec{x}) \bmod P] : G \to R/PR$ (which we can think of as an element of $(R/PR)^{N^4}$). Naively the number of residue patterns is P^{4N^4}, but we will show that the "compression" phenomenon above pushes this number down significiantly.

- Expand $f \in \mathcal{L}_D$ in the R-basis $\mathscr{M}^{(p)}$ of $\mathcal{L}_D$. When we look mod $p^{\lfloor D/p \rfloor} =: p^{k_p}$, the coefficient of $M^{(p)}_{i,j,k}$ only matters mod $p^{\lfloor D/p \rfloor - \lfloor i/p \rfloor}$ because the monomial evaluation already contributes $p^{\lfloor i/p \rfloor}$. Since R is as a $\mathbb{Z}$-module is free rank 4, $|R/p^h R| = p^{4h}$ so we conclude

$$\#\{\text{distinct residue patterns of } f \in \mathcal{L}_D, \bmod p^{k_p}\} \leqslant \prod_{(i,j,k)\in\mathscr{M}} p^{4(\lfloor \frac{D}{p} \rfloor - \lfloor \frac{i}{p} \rfloor)} =: p^{4Mk_p - 4E_p}.$$

- Recalling $S \coloneqq \{p : 32 < p \leqslant X, \chi(p) = -1\}$, we define $P \coloneqq \prod_{p\in S} p^{k_p}$. The Chinese remainder theorem says

$$R/PR \cong \prod_{p\in S} R/p^{k_p} R,$$

  so we see that the residue pattern $\mathsf{ev}_P(f)$ is the same data as the pattern-tuple $(\mathsf{ev}_{p^{k_p}}(f))_{p\in S}$, and hence we conclude

$$\#\{\text{distinct residue pattern of } f \in \mathcal{L}_D, \bmod P\} \leqslant \prod_{p\in S} p^{4Mk_p - 4E_p} = \boxed{P^{4M} \prod_{p\in S} p^{-4E_p}} \tag{2}$$

0.3 "Pigeonhole and Subtract" to Find Polynomial Vanishing on Grid

This section corresponds to Lemma 3 and Lemma 5 from Kerr.

- Consider $f \in \mathcal{L}_D$ written in the $\mathscr{M}$ (R-)basis. For $B, H \geqslant 1$, we say $\|f\|_B \leqslant H$ to mean all R-coefficients c_{ijk} satisfy $|\rho(c_{ijk})| \leqslant HB^{-(i+j+k)}$ for every embedding $\rho : R \hookrightarrow \mathbb{C}$.

The intuition behind this definition is: for $B \coloneqq 8N\sqrt{|q|}$, G grid points have absolute value $\leqslant B$ (see (1)), so $\|f\|_B \leqslant H \implies$ (all conjugates of) all terms $c_{ijk}x^iy^jz^k$ of f are bounded in absolute value by $\leqslant H$, so then (recall $f \in \mathcal{L}_D$)

$$\|f\|_B \leqslant H \implies \|f\|_{\text{conjsup}(G)} \coloneqq \max\{|\rho(f(\vec{x}))| : \vec{x} \in G, \rho : R \hookrightarrow \mathbb{C}\} \leqslant MH. \tag{3}$$

We want a lower bound on the "number of available polynomials"

$$\#\{f \in \mathcal{L}_D : \|f\|_B \leqslant H\} > \dots$$

- Because each coefficient $c_\bullet = (a,b,c,d) \cdot (1, \sqrt{q}, \sqrt{2}, \sqrt{2q})$ has 4 independent integer coefficients, and $|c_\bullet| \leqslant 4\sqrt{2}\sqrt{|q|}\max\{a,b,c,d\} \leqslant 8\sqrt{|q|}\max\{a,b,c,d\}$, one easy source of such f is to take a,b,c,d from 0 to $\lfloor HB^{-(i+j+k)}/8\sqrt{|q|}\rfloor$ to get the lower bound

$$\#\{f \in \mathcal{L}_D : \|f\|_B \leqslant H\} > \prod_{(i,j,k)\in\mathcal{M}} \left(\frac{H}{8\sqrt{|q|}}B^{-(i+j+k)}\right)^4 = \left(\frac{H}{8\sqrt{|q|}}\right)^{4M} B^{-4W}. \tag{4}$$

- If we choose $H \coloneqq \frac{P}{4M}$, then the "number of available polynomials" (RHS(4)) definitely exceeds the "number of distinct residue patterns" (RHS(2)) if the following key inequality holds

$$\sum_{p\in S} E_p \log p > W \log B + M \log(32M\sqrt{|q|}) \tag{5}$$

(to reiterate, we've chosen the parameters $B \coloneqq 8N\sqrt{|q|}$, $H \coloneqq \frac{P}{4M}$).

- Pigeonhole then produces two distinct f, g with identical residue pattern mod P, so subtracting yields a polynomial $F \coloneqq f - g \in \mathcal{L}_D$ satisfying $\|F\|_{\text{conjsup}(G)} \leqslant 2MH \coloneqq \frac{P}{2}$ (see (3)), and taking values in $PR \subseteq R$ on G.

Can replace Kerr's norm argument in Lemma 5 with a superficially more geometric version ("discreteness phenomenon of algebraic integers", like Fundamental Lemma of Transcendental Number Theory: there is no integer between 0 and 1):

- Any non-zero $\alpha \coloneqq a + b\sqrt{q} + c\sqrt{2} + d\sqrt{2q} \in R$ has 4 distinct conjugates (independently toggle sign of $\sqrt{2}, \sqrt{q}$). The average of their squared absolute values is

$$\frac{1}{4}\sum_{\rho:R\to\mathbb{C}} |\rho(\alpha)|^2 = a^2 + b^2|q| + c^2 2 + d^2 2|q| \geqslant 1.$$

This tells us that every non-zero element of R has a conjugate of absolute value $\geqslant 1$. (This is in general true of all algebraic integers.)

- Multiplying by P, we see that every non-zero element of PR has a conjugate with absolute value $\geqslant P$.

- But above we showed every conjugate of our residue pattern $\mathsf{ev}_P(F)$ had absolute value $\leqslant \frac{P}{2}$. This forces $F \in \mathcal{L}_D$ to vanish on all of G! Magic!

We have shown that the technical inequality (5) being true implies the existence of $F \in \mathcal{L}_D$ vanishing on G. Because $1 \leqslant 32$,

$$\deg F \leqslant \text{"weighted degree" of } F \leqslant D. \tag{6}$$

0.4 Degree of F is Large by Baby Algebraic Geometry

We will use linear algebra and rationality to show that $\deg F \geqslant T^{4/3}$ for some large parameter T (roughly $\sim \frac{N}{3}$).

Theorem 0.2: Degree of polynomial vanishing on G is large

Define $T \coloneqq \lfloor \frac{N-1}{3} \rfloor + 1$ (so the parameter D defined above satisfies $D \leqslant \frac{1}{2}T^{4/3}$). Any non-zero polynomial F that vanishes on G must have

$$\deg F \geqslant T^{4/3}.$$

Proof: we can think of G in terms of the following "gamma vectors" in $\mathbb{C}^3$: defining

$$\begin{aligned}\gamma_1 &= (1,1,1),\\ \gamma_2 &= (\sqrt{q},-\sqrt{q},-\sqrt{q}),\\ \gamma_3 &= (\sqrt{2},\sqrt{2},-\sqrt{2}),\\ \gamma_4 &= (\sqrt{2q},-\sqrt{2q},\sqrt{2q}).\end{aligned}$$

we see that for $\theta \coloneqq \theta_{a,b,c,d}$,

$$(\theta, \sigma\theta, \sigma\tau\theta) \coloneqq (a,b,c,d)\cdot(\gamma_1,\dots,\gamma_4),$$

so $G \coloneqq \{(a,b,c,d)\cdot(\gamma_1,\dots,\gamma_4) : a,b,c,d \in [0..N)\}$, i.e. G is exactly the image of $[0..N)^4$ under the $\mathbb{C}$-linear map

$$\phi \coloneqq [(a,b,c,d) \mapsto (a,b,c,d)\cdot(\gamma_1,\dots,\gamma_4)] : \mathbb{C}^4 \to \mathbb{C}^3.$$

Above we constructed a polynomial F that vanishes on the grid $G \subseteq \mathbb{C}^3$. Define $T \coloneqq \lfloor \frac{N-1}{3} \rfloor + 1$ (so the parameter D defined above satisfies $D \leqslant \frac{1}{2}T^{4/3}$) and

$$E \coloneqq \phi([0..T)^4) \coloneqq \{(a,b,c,d)\cdot(\gamma_1,\dots,\gamma_4) : a,b,c,d \in [0..T)\} \subsetneq G \subseteq \mathbb{C}^3,$$

where we defined T that way so that the triple sumset $E+E+E \subseteq G \subseteq \mathbb{C}^3$. Applying the below Theorem 0.4, we get some subspace V of some codimension $r \in \{1,2,3\}$ so that

$$|(E+V)/V| \leqslant (\deg F)^r.$$

On the other hand, we'll show $|(E+V)/V| \geqslant T^{r+1}$ using rationality. "Geometry forces an upper bound (dependence/conspiracy), but arithmetic forces a lower bound (independence)."

Define $\pi : \mathbb{C}^3 \to \mathbb{C}^3/V$ to be the projection $\mathbb{C}$-linear map. It suffices to show that $\pi(\gamma_1),\dots,\pi(\gamma_4) \in \mathbb{C}^3/V$ — let's rename these as $w_1,\dots,w_4 \in W$ — have rational rank $\geqslant r+1$, i.e. $W_{\mathbb{Q}} \coloneqq \operatorname{span}_{\mathbb{Q}}\{w_1,\dots,w_4\}$ has $\dim_{\mathbb{Q}} \geqslant r+1$.

$$\mathbb{C}^4 \xrightarrow[\phi]{e_i \mapsto \gamma_i} \mathbb{C}^3 \xrightarrow[\pi]{\gamma_i \mapsto w_i} W$$

Well, assume f.s.o.c. the $\mathbb{Q}$-span $W_{\mathbb{Q}}$ of the $w_1,\dots,w_4 \in W$ has $\dim_{\mathbb{Q}} \leqslant r$. Their $\mathbb{C}$-span is the full codomain W ($\dim_{\mathbb{C}} = r$), so r of them — denoted say $b_1,\dots,b_r$ — are $\mathbb{C}$-LI hence $\mathbb{Q}$-LI (so in fact we upgrade to $\dim_{\mathbb{Q}} W_{\mathbb{Q}} = r$).

Because $\vec{b}$ $\mathbb{Q}$-spans $W_{\mathbb{Q}}$, we have a rational matrix $Q \in M^{4\times r}(\mathbb{Q})$ so that "$\vec{w} = Q\vec{b}$":

$$\begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} = Q \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix}.$$

Consider $\vec{z} \in \ker \pi\phi \subseteq \mathbb{C}^4$, which means exactly that

$$0 = \pi(\phi(\vec{z})) \coloneqq (z_1, \dots, z_4) \cdot \underbrace{(w_1, \dots, w_4)}_{=Q\vec{b}} = (\vec{z}^\top Q)\vec{b} = (Q^\top \vec{z}) \cdot \vec{b}.$$

But remember $\vec{b}$ forms a $\mathbb{C}$-basis of W, so $Q^\top \vec{z} = \vec{0} \in \mathbb{C}^4$. Because Q is a rational matrix, this shows that every $\vec{z} \in \ker \pi\phi$ is $\mathbb{Q}$-LD!

But there's an obvious element in $\ker \pi\phi$, $c \coloneqq (-\sqrt{2q}, -\sqrt{2}, \sqrt{q}, 1)$, whose coordinates are $\mathbb{Q}$-LI! ※

We've shown that $T^{r+1} \leqslant (\deg F)^r$, and since $r \leqslant 3$ we definitely have $\deg F \geqslant T^{4/3}$, as desired. ■

But recalling our definition of T, D and (6) ($f \in \mathcal{L}_D$!) we see

$$T^{4/3} \leqslant \deg F \leqslant D \leqslant \frac{1}{2}T^{4/3}$$

which is of course impossible. Thus by (6), the key inequality (5) must be false. We show in the last Appendix that

Theorem 0.3: Landau Siegel zero implies Key Inequality true

$$\sum_{p\in S} \frac{\log p}{p} > \frac{9}{8} \log B$$

being true for some large enough X (recall S, B depend on X) implies the key inequality (5). Because we have just proven that the key inequality (5) to be false, we conclude that for all large enough X,

$$\sum_{p\in S} \frac{\log p}{p} \leqslant \frac{9}{8} \log B$$

Key geometric input to the main result

Theorem 0.4: Low sumset (polynomial) complexity forces low subspace (translation) cxty.

Let $0 \in E \subseteq \mathbb{C}^3$ be finite and contain the origin. Suppose a (non-zero) degree D polynomial $F \in \mathbb{C}[x, y, z]$ vanishes on the sumset $E + E + E$. Then there is a (proper) $\mathbb{C}$-VS $V \subsetneq \mathbb{C}^3$ s.t. E is covered by $\leqslant D^{\mathrm{codim}(V)}$ many translates of V.

The conclusion can be written equivalently as: letting $r \coloneqq \mathrm{codim}(V)$,

$$|(E + V)/V| \leqslant D^r \qquad \text{or } \#(E \bmod V) \leqslant D^r.$$

Slogan: "if repeated addition doesn't increase complexity much, then there must be hidden translation symmetry".

Some history of "Sumset to Subspace" deductions in additive combinatorics.

- A main one that comes to mind for me is a key step in Kelley–Meka's 2023 work on 3AP's, namely "almost periodicity" (Key Step 4 in Bloom–Sisask exposition). Theirs is a much stronger "filling" conclusion (finding a small-codimension subspace W that fills up a significant fraction of a sumset $A+A$ of an α-dense A), than the "covering" conclusion we need here, but I think it has a similar spirit.
- Polynomial Frieman-Ruzsa/Marton's conjecture says that if a sumset is small (doesn't increase size by more than a constant factor), then it is covered by a small number (polynomial in the constant factor) translates of a subspace/subgroup. Very similar flavor to what our key tool here.

Proof of Theorem 0.4: we use the basic algebraic geometry we develop in the Appendix (look there for definitions of dimension, irreducible component, Bézout component bound).

Let $Z \coloneqq Z_0 \coloneqq V(F)$ be the zero/vanishing locus of F. We define for $j = 1, 2, 3$

$$Z_j \coloneqq \{\vec{x} \in \mathbb{C}^3 : \vec{x} + \underbrace{E + \ldots + E}_{j \text{ times}} \subseteq Z\} = \{\vec{z} \in \mathbb{C}^3 : \vec{x} + \vec{e}_1 + \ldots + \vec{e}_j \in Z, \forall \vec{e}_1, \ldots, \vec{e}_j \in E\}$$

so $\mathbb{C}^3 \supsetneq Z_0 \supseteq Z_1 \supseteq Z_2 \supseteq Z_3$, where $0 \in Z_3$ (non-empty!) using our assumption about $E + E + E \subseteq Z$!

By Theorem 0.6 (about dimension), $\dim Z_0 \leqslant 2$, so by the pigeonhole principle (4 pigeons and 3 holes!) there is some $j \in \{1, 2, 3\}$ and $d \in \{0, 1, 2\}$ s.t.

$$\dim Z_j = \dim Z_{j-1} = d.$$

Choose an irreducible component of Y of Z_j of (maximal) dimension d (one exists by Thm. 0.6). By construction, $Y + e \subseteq Z_{j-1}$ for all $e \in E$, so we have many maximal-dimension irreducible components $Y + e$ in Z_{j-1}.

But the Bézout component bound (Theorem 0.5) tells us there are $\leqslant D^{3-d}$ such components, so

$$\#\{Y + e : e \in E\} \leqslant D^{3-d}.$$

This is telling us that Y has to have many translation symmetries, in order to have $|E|$-many translates compress into $\leqslant D^{3-d}$ many.

We define the set of translation symmetries $V \coloneqq \{v \in \mathbb{C}^3 : Y + v = Y\}$ of Y; which turns out to be a $\mathbb{C}$-VS: it is easy to see closure under addition and negation; and for scalar multiplication observe that $v \in V$ implies $\forall y \in Y$ the entire discrete line $y + \mathbb{Z}$ is contained $\subseteq Y$, and the only way for an algebraic set to contain a discrete line is for it to contain the whole line $y + \mathbb{C}v \subseteq Y$.

Thus Y contains an affine shift of the $\mathbb{C}$-subspace V, so $\dim_{\mathbb{C}} V = \dim V \leqslant \dim Y =: d \in \{0, 1, 2\}$ (monotonicty, and the equality of two notions of dimension for subspaces is in Thm. 0.6).

The rest is linear algebra. Clearly, two translates $Y + e, Y + e'$ coincide $\iff e - e' \in V$, so we have proven that

$$|(E + V)/V| \leqslant D^{3-d}. \quad \blacksquare$$

Baby Algebraic Geometry Appendix

We state the key result; the definitions (algebraic set, irreducible component, dimension, etc.) and lemmas will come after.

Theorem 0.5: Bézout Component Bound

If an algebraic set $X \subseteq \mathbb{C}^3$ is defined by polynomials with degree $\leqslant D$, and its largest irreducible component has dimension d, then X has

$$\leqslant D^{3-d}$$

many irreducible components.
Special cases that may help communicate the picture: an algebraic set X cut out by degree D surfaces has $\leqslant D^2$ curve components; or if X is finite, $\leqslant D^3$ isolated points.

We follow a more combinatorial route to develop some basic dimension theory for algebraic sets. A set $X \subseteq \mathbb{C}^3$ being algebraic means there exists a set of polynomials $\mathscr{F} \subseteq \mathbb{C}[x, y, z]$ s.t.

$$X = \{(x, y, z) : P(x, y, z) = 0 \quad \forall P \in \mathscr{F}\}.$$

An algebraic set X is irreducible if it can not be written as the union of two proper algebraic subsets $\subsetneq X$. An irreducible component of an algebraic set X is maximal irreducible algebraic subset.

A polynomial function on a subset $X \subseteq \mathbb{C}^3$ is a function $f : X \to \mathbb{C}$ s.t. there exists a polynomial $P \in \mathbb{C}[x, y, z]$ s.t. $f(x) = P(x)$ for all $x \in X \subseteq \mathbb{C}^3$.

Definition 0.3: Hilbert H function/polynomial, to define dimension

Define

$$H_X(m) \coloneqq \dim_{\mathbb{C}\text{-VS}} \mathcal{P}_m(X) \coloneqq \dim_{\mathbb{C}\text{-VS}}\{\text{polynomial functions on } X \text{ of degree} \leqslant m\}.$$

Note that denoting by $\mathcal{P}_m$ the $\mathbb{C}$-VS of polynomials of total degree $\leqslant m$, the restriction map $P \mapsto P\big|_X : \mathcal{P}_m \to \mathcal{P}_m(X)$ has kernel exactly the ideal $I(X)$ of polynomials vanishing on X, so $H_X(m) = \dim \mathcal{P}_m - \dim(I(X) \cap \mathcal{P}_m)$.

Lemma 0.1: Polynomial function space dimension is eventually polynomial

For any algebraic set $X \subseteq \mathbb{C}^3$, the function $m \mapsto H_X(m)$ is eventually a polynomial.

Definition 0.4: Dimension of algebraic set

Let $X \subseteq \mathbb{C}^3$ be an algebraic set. By the previous lemma $m \mapsto H_X(m)$ is eventually polynomial, which has a degree we'll denote $\deg H_X(m)$, so it is well-defined to write

$$\dim X \coloneqq \deg H_X(m).$$

In the case of $\mathbb{C}^3$ it is easy to see (especially after reading proof of Lemma 0.1 below) that the possible dimensions of algebraic sets $X \subseteq \mathbb{C}^3$ are $0, 1, 2, 3$.

Proof of Lemma: the difficulty with understanding $H_X(m)$ is that it is the dimension of a vector space where the vectors are equivalence classes of polynomials

$$[f] \coloneqq \{P \in \mathbb{C}[x,y,z] : P\big|_X \equiv f \text{ as functions } X \to \mathbb{C}\},$$

and equivalence classes are always a little tricky to work with. The standard fix is to take any polynomial representative $P \in [f] \subseteq \mathbb{C}[x,y,z]$, and "reduce its complexity" until we reach some kind of "minimal"/"canonical" form $P(f) \in [f]$, i.e. somehow get a **canonical/minimal representative from each equivalence class** $[f]$ — **this makes up the first half of the proof**.

We take our notion of "complexity" to be given by the largest monomial under the order $\prec$ of comparing total sum (of exponents) and breaking ties lexicographically:

$$x^0y^0z^0 \prec x^0y^0z^1 \prec x^0y^1z^0 \prec x^1y^0z^0 \prec x^0y^0z^2 \prec x^0y^1z^1 \prec \dots$$

The rough idea is to reduce the complexity of P, by subtracting polynomials that vanish on X to cancel away leading monomials one by one.

More rigorously, let $I(X) \subseteq \mathbb{C}[x,y,z]$ be the set of polynomials that vanish on X. Consider the set $\mathscr{M} \subseteq \mathbb{C}[x,y,z]$ of $\prec$-largest monomials $x^ay^bz^c$ of (monic say) polynomials $P_{a,b,c}(x) = x^ay^bz^c +$ (lesser monomials) in $I(X)$:

$$\mathscr{M} \coloneqq \{\text{largest monomial of } P : P \in I(X)\}$$

There is an obvious correspondence between $\mathscr{M}$ and its exponents-set $\tilde{\mathscr{M}} \subseteq \mathbb{N}^3$. To finish the discussion on reducing complexity, given any polynomial function $f : X \to \mathbb{C}$ and polynomial representative $P \in [f]$, we can kill off the largest monomials of P step by step, remaining in the equivalence class $[f]$, until we reach $P(f) \in [f]$ whose largest monomial $\notin \mathscr{M}$.

(It is easy to see there is only one monic representative in $[f]$ with largest monomial $\notin \mathscr{M}$; if there were 2, subtract to get a polynomial in $I(X)$ whose largest monomial $\notin \mathscr{M}$; ※.)

This work to cook up a canonical representative reveals that

$$H_X(m) = \#\{\text{monomials} \notin \mathscr{M} \text{ of total exponent sum} \leqslant m\}.$$

The **second half of the proof uses inclusion-exclusion (and general combinatorics) to argue that the above formula for $H_X(m)$ equals a certain polynomial in m**, for all large enough m.

By multiplying any above such $P_{a,b,c}$ by monomials, we see that $\tilde{\mathscr{M}} \subseteq \mathbb{N}^3$ is an upwards-closed subset of the poset $(\mathbb{N}^3, \leqslant_{\text{dom}})$. It is easy to see that $\tilde{\mathscr{M}}$ is the upwards-closure of its subset of $\leqslant_{\text{dom}}$-minimal elements $\tilde{\mathscr{M}}_{\min}$, of which there are finitely many by Dickson (Lemma 0.7)!

Define $Q_{a,b,c} \subseteq \mathbb{N}^3$ to be the "quadrant" of all things $\geqslant_{\text{dom}} (a,b,c) \in \mathbb{N}^3$, and the "simplex" $\Delta^{\leqslant m} \coloneqq \{(a,b,c) : a+b+c \leqslant m\}$ of monomials of total exponent sum $\leqslant m$. Also observe that

$$Q_{a,b,c} \cap Q_{a',b',c'} = Q_{a\vee a', b\vee b', c\vee c'}$$

where $x \vee y \coloneqq \max\{x,y\}$. (Draw these shapes in $\mathbb{N}^2$!)

We leave it as easy exercises to show

$$|\Delta^{\leqslant m}| = \binom{m+3}{3}$$

and (by "shifting" to move (a,b,c) to $(0,0,0)$)

$$|Q_{a,b,c} \cap \Delta^{\leqslant m}| = \binom{m-(a+b+c)+3}{3},$$

which are polynomials in m for sufficiently large m (that the top number in the binomial coefficient is $\geqslant 1$).

Inclusion-exclusion (draw a picture in $\mathbb{N}^2$! The below monster formula makes this very visual argument more scary than it actually is!) tells us

$$H_X(m) = |\Delta^{\leqslant m}| - \sum_{\alpha:=(a,b,c)\in \tilde{\mathcal{M}}_{\min}} |Q_\alpha \cap \Delta^{\leqslant m}| + \sum_{\alpha<_{\text{lex}}\alpha'\in\tilde{\mathcal{M}}_{\min}} |Q_{\alpha\vee\alpha'} \cap \Delta^{\leqslant m}| - \dots$$

which by the exercises in the previous sentence is a linear combination of polynomials in m, for large enough m. Note how important finiteness of $\tilde{\mathcal{M}}_{\min}$ is! ■

Now that we have a well-defined dimension function, we should prove some basic facts about it.

Theorem 0.6: Dimension function properties

The dimension function on algebraic subsets defined above satisfies:

1. non-empty, proper algebraic subsets $\varnothing \neq X \subsetneq \mathbb{C}^3$ have dimensions $\in \{0,1,2\}$
2. every dimension d algebraic set contains an irreducible component of dimension d.
3. $Y \subseteq X \implies \dim(Y) \leqslant \dim(X)$
4. suppose $Y \subsetneq X$ with X irreducible. Then $\dim Y < \dim X$.
5. if $V \subseteq \mathbb{C}^3$ is an affine $\mathbb{C}$-subspace, then $\dim V = \dim_{\mathbb{C}} V$ (algebraic set dimension equals linear algebra dimension)

I will postpone the proof of Theorem 0.6 until after two lemmas on irreducible components.

Lemma 0.2: Finite irreducible decomposition

Every algebraic set $X \subseteq \mathbb{C}^3$ is a finite union of irreducible algebraic sets. Its maximal irreducible subsets (called irreducible components) are finite in number and uniquely determined.

Proof of Lemma 0.2: The ascending-chain condition on polynomial ideals (implied by Dickson, Lemma 0.7 below) implies the descending-chain condition on algebraic sets.

Suppose some algebraic set cannot be expressed as a finite union of irreducible sets. By the descending-chain condition, choose a minimal such set X. Then X is reducible, so $X = Y \cup Z$ for proper algebraic subsets $Y, Z \subsetneq X$. By minimality, both Y and Z are finite unions of irreducible sets, hence so is X, a contradiction.

Thus write $X = X_1 \cup \dots \cup X_t$ with each X_i irreducible, discarding redundant members.

If an irreducible set W is contained in a finite union of algebraic sets, then W is contained in one of them; this follows by induction from the definition of irreducibility. Hence every irreducible subset of X is contained in some X_i.

It follows that the X_i are precisely the maximal irreducible subsets of X: a larger irreducible set containing X_i would lie in some X_j, contradicting irredundancy unless $i = j$. Thus the decomposition

is unique up to reordering. ■

Lemma 0.3: Irreducibility and prime ideals

A nonempty algebraic set $X \subseteq \mathbb{C}^3$ is irreducible if and only if its vanishing ideal $I(X)$ is prime.

Proof of Lemma 0.3: Suppose X is irreducible. Observe that no matter what X is, $fg \in I(X) \implies$ every $x \in X$ has $f(x)g(x) = 0$, so $x \in V(f)$ or $\in V(g)$; from this it is easy to get to

$$X = (X \cap V(f)) \cup (X \cap V(g)).$$

Both sets on the RHS are algebraic. By irreducibility, one equals X, so $f \in I(X)$ or $g \in I(X)$. Since $X \neq \varnothing$, $I(X) \neq (1)$. Thus $I(X)$ is prime.

Conversely, suppose $I(X)$ is prime and $X = Y \cup Z$ for proper algebraic subsets $Y, Z \subsetneq X$. Choose

$$f \in I(Y) \setminus I(X), \qquad g \in I(Z) \setminus I(X).$$

Such polynomials exist because Y, Z are proper algebraic subsets of X. Then fg vanishes on $Y \cup Z = X$, so $fg \in I(X)$, contradicting primeness. Hence X is irreducible. ■

Proofs of the big AG theorems

Proof of Theorem 0.6: recall

$$H_X(m) \coloneqq \dim_{\mathbb{C}}\{f|_X : \deg f \leqslant m\}, \qquad \dim X \coloneqq \deg H_X,$$

where $\dim \varnothing \coloneqq -1$ and $H_\varnothing \coloneqq 0$. We use the previously proved eventual polynomiality of H_X.

1. For $X \neq \varnothing$, we have $H_X(m) \geqslant 1$, so $\dim X \geqslant 0$. For $X \subsetneq \mathbb{C}^3$, some nonzero polynomial f vanishes on X. The polynomials $fg \in I(X) \cap \mathcal{P}_m$ are $\mathbb{C}$-LI as g ranges over the standard monomial basis of $\mathcal{P}_{m-\deg f}$ because multiplication by (nonzero) f is injective (UFD!) in $\mathbb{C}[x, y, z]$, so (by alternate formula in Def. 0.3)

$$H_X(m) \leqslant \binom{m+3}{3} - \binom{m-\deg f+3}{3} = O(m^2).$$

   Thus $\dim X \leqslant 2$.

2. Write $X = X_1 \cup \cdots \cup X_t$ as its finite irreducible-component decomposition (Lemma 0.2 above). Restriction of polynomial functions gives

$$\max_i H_{X_i}(m) \leqslant H_X(m) \leqslant \sum_i H_{X_i}(m).$$

   Comparing polynomial growth shows $\dim X = \max_i \dim X_i$.

3. If $Y \subseteq X$, restriction of polynomial functions on X to Y is surjective (degree-wise, in fact), or slightly rephrased in notation (Def. 0.3): $P\restriction_X \mapsto P\restriction_{Y \subseteq X} : \mathcal{P}_m(X) \to \mathcal{P}_m(Y)$ is surjective. Hence $H_Y(m) \leqslant H_X(m)$ and $\dim Y \leqslant \dim X$.

4. Suppose $Y \subsetneq X$, with X irreducible. Choose $h \in I(Y) \setminus I(X)$. Since $I(X)$ is prime, and $h \notin I(X)$, multiplication by h is injective on polynomial functions on X ($hf\restriction_X \equiv 0 \implies f\restriction_X \equiv 0$). Consider

the diagram

$$\mathcal{P}_{m-\deg h}(X) \xhookrightarrow{\cdot h} \mathcal{P}_m(X) \xrightarrow{\text{restriction}} \mathcal{P}_m(Y)$$

Observe that the kernel of the second map contains the image of the first, and the dimensions are respectively $H_X(m-\deg h), H_X(m), H_Y(m)$. Rank-nullity gives

$$H_Y(m) \leqslant H_X(m) - H_X(m-\deg h).$$

If the eventual polynomial $H_X(\bullet)$ has degree $d \geqslant 1$, the right side is an eventual polynomial of degree $d-1$. Else, $d=0$, and the RHS is eventually zero, so $Y=\varnothing$. In either case $\dim Y < \dim X$.

5. Write the affine subspace as $V = v + W$, where $\dim_{\mathbb{C}} W = k$. Choose affine linear coordinates $t_1, \dots, t_k$ on V, extended to affine linear functions on $\mathbb{C}^3$. Restriction identifies the degree-$\leqslant m$ polynomial functions on V with polynomials of degree $\leqslant m$ in these k coordinates. Thus

$$H_V(m) = \binom{m+k}{k}, \qquad \dim V = k = \dim_{\mathbb{C}} W. \quad \blacksquare$$

Proof of Theorem 0.5: Write

$$X =: V(P_1, \dots, P_s) \subseteq \mathbb{C}^3, \qquad \deg P_i \leqslant D,$$

and put $d := \dim X$. We prove that X has at most D^{3-d} irreducible components of dimension d.

Preliminary (distilling out a trick we used twice above): multiplication–restriction. Recall that for an algebraic set Z, we write

$$\mathcal{P}_m(Z) := \{P|_Z : \deg P \leqslant m\}, \qquad H_Z(m) := \dim_{\mathbb{C}} \mathcal{P}_m(Z).$$

Suppose $Y = Z \cap V(g)$, where $\deg g = b$, and multiplication by g is injective on polynomial functions on Z. Consider

$$\mathcal{P}_{m-b}(Z) \xhookrightarrow{\times g} \mathcal{P}_m(Z) \xrightarrow{\text{res}} \mathcal{P}_m(Y).$$

The first map is injective, the second is surjective, and their composition is zero. Thus by rank-nullity

$$\boxed{H_Y(m) \leqslant H_Z(m) - H_Z(m-b).} \tag{MR}$$

Case $d=2$: at most D surface components. Choose any nonzero P_i and factor it into irreducibles:

$$P_i = c\, u_1^{e_1} \cdots u_t^{e_t}.$$

Every irreducible surface (=2-dim) component C of X lies inside $V(P_i) = \bigcup_{j=1}^{t} V(u_j)$. By irreducibility, $C \subseteq V(u_j)$ for some j.

Since u_j is irreducible, $V(u_j)$ is irreducible, by the Nullstellensatz and the prime-ideal criterion. Moreover, writing $a = \deg u_j$,

$$H_{V(u_j)}(m) = \binom{m+3}{3} - \binom{m-a+3}{3},$$

so $\dim V(u_j) = 2$. The strict dimension-drop theorem (Thm. 0.6) forces $C = V(u_j)$. Therefore the

number of surface components is at most

$$t \leqslant \deg P_i \leqslant D.$$

Preparation for $d \leqslant 1$: choose two equations. Choose a nonzero P_i and let f be the product of its distinct irreducible factors. Then

$$X \subseteq V(f), \qquad \deg f =: a \leqslant D, \qquad I(V(f)) = (f).$$

We claim that we can choose a linear combination

$$g = \sum_{i=1}^{s} \lambda_i P_i$$

such that $\gcd(f, g) = 1$.

Indeed, let u be any irreducible factor of f. Not every P_i is divisible by u: otherwise $V(u) \subseteq X$, contradicting $\dim X \leqslant 1$. The condition $u \mid g$ is therefore a proper linear condition on $(\lambda_1, \ldots, \lambda_s)$. There are only finitely many factors u, and finitely many proper linear subspaces cannot cover $\mathbb{C}^s$. Choose the coefficients outside their union.

Put $C \coloneqq V(f, g)$. We have $X \subseteq C$, and $b \coloneqq \deg g \leqslant D$.

Because f is squarefree and $\gcd(f, g) = 1$, multiplication by g on polynomial functions P on $V(f)$ is injective: if P is s.t. $gP \in (f)$, Euclid's lemma gives $P \in (f)$.

Applying (MR) with $Z \coloneqq V(f)$ and $Y \coloneqq C$ gives

$$\begin{aligned} H_C(m) &\leqslant H_{V(f)}(m) - H_{V(f)}(m-b) \\ &= \left[\binom{m+3}{3} - \binom{m-a+3}{3}\right] \\ &\quad - \left[\binom{m-b+3}{3} - \binom{m-a-b+3}{3}\right] \\ &= abm + O(1). \end{aligned} \tag{B}$$

This is our basic two-equation bound.

Case $d = 1$: at most D^2 curve components. Let $C_1, \ldots, C_t$ be the one-dimensional irreducible components of X.

For each i, choose a polynomial q_i which vanishes on every C_j with $j \neq i$, but not identically on C_i. Such a polynomial exists: for each $j \neq i$, choose an equation vanishing on C_j but not on C_i, and multiply these equations. Primeness of $I(C_i)$ ensures that the product does not vanish identically on C_i.

Also choose a linear polynomial ℓ_i nonconstant on C_i; some coordinate function must be nonconstant, since $\dim C_i = 1$.

For each i, we claim the functions

$$q_i,\ q_i\ell_i,\ q_i\ell_i^2, \ldots$$

are linearly independent on C_i. Indeed, q_i is nonzero in the integral domain of polynomial functions on C_i, and ℓ_i takes infinitely many values on C_i.

Moreover, the families belonging to distinct components $C_\bullet$ are $\mathbb{C}$-LI on X: restricting a linear relation to C_i kills every family except the ith.

Consequently, for sufficiently large m,

$$H_X(m) \geqslant \sum_{i=1}^{t}(m - \deg q_i + 1) = tm - O(1).$$

But $X \subseteq C$, so (B) gives

$$H_X(m) \leqslant H_C(m) \leqslant abm + O(1).$$

Comparing coefficients of m, we obtain

$$\boxed{t \leqslant ab \leqslant D^2.}$$

Case $d = 0$: at most D^3 points. First, X is finite. Indeed, since $H_X(m)$ is eventually constant, the powers of each coordinate function are linearly dependent on X. Hence each coordinate takes only finitely many values, and so does the triple (x, y, z).

Retain the coprime equations f, g chosen above, and put $C \coloneqq V(f, g)$.

Because f, g are coprime, C has no surface component. By the codimension-two intersection lemma (Lemma 0.4 below), it has no isolated point component. Hence every irreducible component of C is a curve. Write these curve components as $C_1, \ldots, C_t$.

We claim that we can choose another defining combination

$$h = \sum_{i=1}^{s} \mu_i P_i$$

which does not vanish identically on any C_j.

Indeed, no C_j is contained in X, because X is finite. Thus for each C_j, some P_i does not vanish identically on C_j. The condition $h|_{C_j} = 0$ is therefore a proper linear condition on $(\mu_1, \ldots, \mu_s)$. Avoiding these finitely many conditions gives h.

Put $Y \coloneqq C \cap V(h)$ and $c \coloneqq \deg h \leqslant D$. Then $X \subseteq Y$. Moreover, each $C_j \cap V(h)$ is a proper algebraic subset of the irreducible curve C_j, so it has dimension at most zero, by strict dimension drop. Thus Y is finite.

Multiplication by h on polynomial functions on C is injective. Indeed, if hu vanishes on C, then on each irreducible component C_j, primeness of $I(C_j)$ and $h|_{C_j} \not\equiv 0$ force $u|_{C_j} \equiv 0$. Hence $u \equiv 0$ on C.

Applying (MR) again, now with $:== C$, gives

$$H_Y(m) \leqslant H_C(m) - H_C(m - c). \tag{*}$$

Since $\dim C = 1$, its eventual Hilbert polynomial has the form

$$H_C(m) = \delta m + \eta.$$

By (B), $\delta \leqslant ab$. Thus (*) gives, for all sufficiently large m,

$$H_Y(m) \leqslant c\delta \leqslant abc.$$

Finally, since Y is finite, polynomial interpolation gives $H_Y(m) = |Y|$ for sufficiently large m. Therefore

$$\boxed{|X| \leqslant |Y| \leqslant abc \leqslant D^3.}$$

This completes all three cases. ■

Lemma 0.4: Codim 2 intersection

Let $f, g \in \mathbb{C}[x, y, z]$. Every irreducible component of

$$V(f, g) = \{(x, y, z) \in \mathbb{C}^3 : f(x, y, z) = g(x, y, z) = 0\}$$

has dimension at least 1. Equivalently, $V(f, g)$ has no isolated point components.

Proof of Lemma 0.4: (uses UFD/Gauss's lemma, Euclid's lemma, polynomial division, localization, finite irreducible decomposition, and the Nullstellensatz).

The idea is to establish two incompatible conclusions: an isolated common zero would produce a nonzero polynomial class annihilated by every coordinate, whereas two coprime equations admit a linear combination of the coordinates whose multiplication map is injective.

We may assume f, g are nonzero and coprime. Indeed, if $h \coloneqq \gcd(f, g)$, then

$$V(f, g) = V(h) \cup V(f/h, g/h).$$

The first set has only surface components, while the second is defined by coprime equations. Thus it suffices to treat the latter case.

Step 1: An isolated point produces an annihilated element.

Suppose, for contradiction, that $V(f, g)$ has an isolated irreducible component consisting of a point. Translate coordinates so that this point is 0.

By finite irreducible decomposition, we can choose a polynomial s vanishing on all the other components, but with $s(0) \neq 0$. Hence

$$V(f, g) \cap \{s \neq 0\} = \{0\}.$$

Consider the ring

$$A \coloneqq \mathbb{C}[x, y, z, 1/s]/(f, g).$$

Intuitively, A records polynomial functions near the origin, subject to the equations $f = g = 0$. It is nonzero, since evaluation at 0 defines a homomorphism $A \to \mathbb{C}$.

By the Nullstellensatz, some powers of x, y, z vanish in A. More precisely, the equations

$$f = g = ts - 1 = 0$$

in four variables have the unique solution $(0, 0, 0, 1/s(0))$. Thus, for some $M \geqslant 1$,

$$x^M, y^M, z^M \in (f, g, ts - 1).$$

Substituting $t = 1/s$ gives

$$\boxed{x^M = y^M = z^M = 0 \quad \text{in } A.} \tag{1}$$

These coordinates need not themselves be zero: compare $\mathbb{C}[x]/(x^3)$, where $x \neq 0$ but $x^3 = 0$.

Put $\mathfrak{m} \coloneqq (x, y, z)A$. Equation (1) implies that $\mathfrak{m}$ is nilpotent: every monomial of sufficiently large total degree is divisible by x^M, y^M, or z^M.

Choose k maximal such that $\mathfrak{m}^k \neq 0$, and take $0 \neq H \in \mathfrak{m}^k$. Since $\mathfrak{m}^{k+1} = 0$, we obtain

$$\boxed{xH = yH = zH = 0 \quad \text{in } A.} \tag{2}$$

In particular, every linear combination of x, y, z annihilates H.

Step 2: A generic plane preserves coprimality.

Introduce independent indeterminates a, b, and put

$$K \coloneqq \mathbb{C}(a, b), \qquad \ell \coloneqq z - ax - by.$$

We claim that the specialized polynomials

$$\overline{f} \coloneqq f(x, y, ax + by), \qquad \overline{g} \coloneqq g(x, y, ax + by)$$

are coprime in $K[x, y]$.

Indeed, after inverting x, the substitution $z = ax + by$ becomes reversible, since $a = (z - by)/x$. It therefore induces an isomorphism

$$\mathbb{C}[a, b, x^{\pm 1}, y] \cong \mathbb{C}[b, x^{\pm 1}, y, z].$$

Under this isomorphism, $\overline{f}, \overline{g}$ correspond to the original coprime polynomials f, g. They remain coprime after adjoining b and inverting x, by Gauss's lemma and unique factorization.

Consequently, any common irreducible factor of $\overline{f}, \overline{g}$ in $\mathbb{C}[a, b, x, y]$ would have to be associated to x: all other factors survive upon inverting x.

But x cannot divide both. For example, $x \mid \overline{f}$ would imply

$$f(0, y, by) = 0$$

identically in b, y. Since the substitution $z = by$ is injective on $\mathbb{C}[y, z]$, this implies $f(0, y, z) = 0$, hence $x \mid f$. The same argument applies to g, contradicting $\gcd(f, g) = 1$.

By Gauss's lemma, we conclude that

$$\boxed{\gcd(\overline{f}, \overline{g}) = 1 \quad \text{in } K[x, y].} \tag{3}$$

Step 3: Multiplication by ℓ is injective.

We claim that multiplication by ℓ on

$$B \coloneqq K[x, y, z]/(f, g)$$

is injective. Equivalently,

$$\ell H \in (f, g) \implies H \in (f, g). \tag{4}$$

To prove this, suppose

$$\ell H = fU + gW.$$

Reduce modulo ℓ, or equivalently substitute $z = ax + by$. We obtain

$$\overline{f}\,\overline{U} + \overline{g}\,\overline{W} = 0.$$

By (3) and Euclid's lemma, there exists $T \in K[x, y]$ such that

$$\overline{U} = \overline{g}T, \qquad \overline{W} = -\overline{f}T.$$

Thus $U - gT$ and $W + fT$ vanish upon setting $\ell = 0$. By polynomial division by the monic linear polynomial ℓ, there exist U', W' such that

$$U = gT + \ell U', \qquad W = -fT + \ell W'.$$

Substitution gives

$$\begin{aligned}\ell H &= f(gT + \ell U') + g(-fT + \ell W') \\ &= \ell(fU' + gW').\end{aligned}$$

Canceling ℓ yields $H = fU' + gW' \in (f, g)$, proving (4).

Thus the essential argument is:

$$\boxed{\text{reduce modulo } \ell \longrightarrow \text{ apply Euclid } \longrightarrow \text{ lift and cancel.}}$$

Step 4: Contradiction.

Recall from Step 1 that $0 \neq H \in A$ satisfies $xH = yH = zH = 0$.

Extend scalars to $K = \mathbb{C}(a, b)$ and put

$$A_K \coloneqq A \otimes_{\mathbb{C}} K \cong K[x, y, z, 1/s]/(f, g).$$

Since extension of scalars by a field preserves nonzero vectors, H remains nonzero in A_K. But (2) implies

$$\ell H = (z - ax - by)H = 0.$$

On the other hand, Step 3 proves that multiplication by ℓ is injective on B. This remains true after localizing at s, hence on $A_K = B[1/s]$.

Therefore $H = 0$ in A_K, a contradiction.

We conclude that $V(f, g)$ has no isolated point component. Since every zero-dimensional irreducible algebraic set is a point, every irreducible component of $V(f, g)$ has dimension at least 1. ■

Weak Nullstellensatz

Theorem 0.7: (contrapositive of) Hilbert Nullstellensatz over $\mathbb{C}$

IF $g_1, \ldots, g_n$ do NOT generate the unit ideal in $A[\frac{1}{f}]$ — or equiv., $g_1, \ldots, g_n \in \mathfrak{m}$ a maximal ideal of $A[\frac{1}{f}]$;

THEN $\bigcap_{i=1}^n V(g_i \mid \mathbb{C}^d) \not\subseteq V(f \mid \mathbb{C}^d)$ — or equiv., $\exists$ a solution $\overline{a} \in \mathbb{C}^d$ to the system $\left\{ \begin{smallmatrix} \text{all } g_i(\overline{a})=0 \\ f(\overline{a})\neq 0 \end{smallmatrix} \right.$.

(Of course, conversely, such solution to the system implies the impossibility of $g_1, \ldots, g_n$ generating the unit ideal in $A[\frac{1}{f}]$.)

First-half of proof of Nullstellensatz: of course, given $g_1, \ldots, g_n \in \mathfrak{m}$ a maximal ideal of $A[\frac{1}{f}]$, we quotient by $\mathfrak{m}$ to get a field

$$L \coloneqq \mathbb{C}[x_1, \ldots, x_d]\big[\tfrac{1}{f}\big]/_{\mathfrak{m}}.$$

Observe that in this field, we have a solution to the system $\left\{ \begin{smallmatrix} \text{all } g_i(\overline{a})=0 \\ f(\overline{a})\neq 0 \end{smallmatrix} \right.$, namely the cosets $\frac{x_1}{1} + \mathfrak{m}, \ldots, \frac{x_d}{1} + \mathfrak{m} \in L$, because g_i evaluated at those land in $\mathfrak{m}$ (i.e. zero in L), and f evaluated a those has a multiplicative inverse $\frac{1}{f} + \mathfrak{m}$ in L, hence isn't zero in L.

It remains to transfer this L-solution to the system to a $\mathbb{C}$-solution to the system. There are 2 ways of doing this: the model theory ACF transfer principle (transferring the satisfaction of the 1st order ($\mathscr{L}_{\text{ring}}$-)sentence $\varphi \coloneqq \exists \mathsf{x}_1 \ldots \mathsf{x}_d \big(\bigwedge_{i=1}^n g_i(\overline{\mathsf{x}})\mathsf{=}0 \wedge \neg f(\overline{\mathsf{x}})\mathsf{=}0 \big)$ between L and $\mathbb{C}$); or Zariski's [f.g. field algebraicity] lemma (which tells us in fact $L \cong \mathbb{C}$). We prove Zariski's lema below.

Lemma 0.5: Zariski's (finitely generated) field algebraicity lemma

Let L be a field extension of a field k that is f.g. as a k-algebra (i.e. $L = k[\alpha_1, \ldots, \alpha_n]$ for some $\alpha_1, \ldots, \alpha_n \in L$). Then L is algebraic over k (i.e. all α_i are algebraic over k).

Allcock exposits his proof quite well. No fancy technology/concepts, very concrete (a bit computational...); things are first explained intuitively and then rigorously. There is literally no way to improve upon it. Lay your eyes upon its full majesty https://web.ma.utexas.edu/users/allcock/expos/nullstellensatz3.pdf...

MO (+ Keith Conrad comments and reference to Fulton's book *Algebraic Curves* (Chp.1 §10 Prop. 4)) informs me Azarang has also written up a short proof, which I think is **exactly the same** as Allcock's proof, but phrased slightly differently like using induction and using fancier language (integral extension).

It is interesting to compare Allcock and Azarang proofs, and also the more well-known proofs using Noether normalization. Zanachan presents a/the proof via Noether normalization in 5 minutes.

In particular, Noether normalization gives a **direct proof**: we (assuming algebraic dependence, and a variable change trick — D&F pg.699) actually find a monic polynomial with coefficients in $k[\alpha_1, \ldots, \alpha_{n-1}]$ that α_n is a root of; finish the proof by induction and transitivity of integrality (overall, I think can unwind this to get an explicit polynomial out — (transitivity of) integrality intimately tied to determinant trick; kind of gross but still constructive).

Whereas Allcock/Azarang both prove by the **contrapositive** (Allcock)/**contradiction** (Azarang): we assume α_1 is NOT algebraic, so all polynomials in α_1 are not zero and can be inverted in the fraction field. I was actually able to twist Allcock/Azarang's proof to make it constructive, by keeping careful track of the fractions involved.

0.4.1 Azarang's proof of Zariski's [f.g. field algebraicity] lemma

Zanachan and Azarang both use/isolate/distill the following baby lemma (in fact iff, see Thm. 2.7 in KConrad article):

Lemma 0.6: Fields can only integrally extend other fields

Suppose a field F is integral over a subdomain D. Then D must be a field.

Proof (taken from Zanachan slide): given nonzero $0 \neq f \in D$, want to show $f^{-1} \in D$ (where f^{-1} refers to the inverse of f in F). We assumed F is integral over D, so we have a monic polynomial relation ($n \geqslant 1, d_i \in D$):

$$(f^{-1})^m + d_{m-1}(f^{-1})^{m-1} + \ldots + d_0 = 0.$$

Multiplying by f^{m-1} and rearranging yields of course $f^{-1} \in D$. ■

We are now ready to prove Zariski's [f.g. field algebraicity] lemma, by contradiction.

Proof: by induction on the number of generators $\alpha_1, \ldots, \alpha_n$ ($\in$ the larger field L) it takes to get from the base field k to to the larger field L.

For $n = 1$, we assume $L = k[\alpha]$.

Then L is a field implies $\alpha^{-1} \in L$, i.e. $\alpha^{-1} = \text{poly}(\alpha)$. Multiplying by α and rearranging, we get that α is a root of a polynomial $\in k[t]$.

For $n > 1$, we have $L = k[\alpha, \beta_2, \ldots, \beta_n]$, which also equals $= k(\alpha)[\beta_2, \ldots, \beta_n]$ ($\subseteq$ is obvious; and $\supseteq$ because L is a field). Note $\beta_2, \ldots, \beta_n$ are algebraic over $k(\alpha)$.

From here, we consider $e_1, \ldots, e_\ell$ a (finite!) basis of L as a $k(\alpha)$-VS (i.e. e_i is a (culled if necessary) enumeration of products of $\beta_\bullet$; finitely many span L because $\beta_\bullet$ are algebraic over $k(\alpha)$). We write down the multiplication table for these basis elements:

$$e_i e_j = \sum_k \underbrace{\frac{p_{ijk}(\alpha)}{q_{ijk}(\alpha)}}_{\in k(\alpha)} \cdot e_k,$$

where $p_{ijk}, q_{ijk} \in k[t]$ and $q_{ijk}(\alpha) \neq 0$ (the $k(\alpha)$ coefficients in terms of the basis $\{e_\bullet\}$ are well-defined).

Consider arbitrary $x \in L = k[\alpha, \beta_2, \ldots, \beta_n]$; i.e. x is a (finite!) k-linear combination (LC) of products of $\alpha, \beta_2, \ldots, \beta_n$, which are all themselves $k(\alpha)$-LCs of the basis $\{e_\bullet\}$:

$$\beta_j = \sum_k \frac{r_{jk}(\alpha)}{s_{jk}(\alpha)} \cdot e_k,$$

where $r_{jk}, s_{jk} \in k[t]$ and $s_{jk}(\alpha) \neq 0$. (Yes, most of these r, s polynomials are $\equiv 1$ or $\equiv 0$; unless β_j was culled when we extracted the basis $\{e_\bullet\}$ from the list of products of the $\beta_\bullet$.)

Plugging these $\{e_\bullet\}$ representations of β_j into the k-LC representing x, and using the multiplication table, we get

$$x = \sum_k \frac{P_x(\alpha)}{\prod q_\bullet(\alpha) \prod s_\bullet(\alpha)} \cdot e_k,$$

for $P_x \in k[t]$ (may depend on x). Call this big product denominator polynomial $=: Q \in k[t]$ — $Q(\alpha) \neq 0$.

We have shown that every $k(\alpha)$-LC of the basis $\{e_\bullet\}$ equals a $k[\alpha][\frac{1}{Q(\alpha)}]$ -LC, so by uniqueness of coefficients of LCs w.r.t. a basis, we conclude

$$k(\alpha) = k[\alpha]\left[\frac{1}{Q(\alpha)}\right].$$

To finish, let $Z \in k[t]$ relatively prime to $Q \in k[t]$ (can use Euclid infinitude of primes trick). Then we know (if $Z(\alpha) = 0$, immediately done; else:)

$$\frac{1}{Z(\alpha)} = \frac{P(\alpha)}{Q(\alpha)} \implies Q(\alpha) - P(\alpha)Z(\alpha) = 0.$$

The LHS polynomial $Q(t) - P(t)Z(t)$ can not be identically $\equiv 0$, since that would imply $Q(t) = P(t)Z(t)$ in $k[t]$, but we demanded $Z(t)$ to be relatively prime to $Q(t)$. ■

This combines the induction of Azarang's proof, with Allcock's basis techniques and avoidance of fancy technology/terminology (like integral extensions), and my own twist to make it constructive:

- given $k(\alpha)[\beta_2, \ldots, \beta_n]$, where I know the $k(\alpha)$-polynomial with β_j as zero (actually just the degree suffices),

- I can form a finite spanning set of products of $\beta_\bullet$,
- which I can cull to get a basis $\{e_\bullet\}$,
- w.r.t. which I can write out LC representations for β_j, and also a multiplication table,
- whose coefficients' denominator polynomials I can multiply together to get $Q \in k[t]$,
- to which I can find $Z \in k[t]$ relatively prime and $P \in k[t]$ such that $\frac{1}{Z(\alpha)} = \frac{P(\alpha)}{Q(\alpha)}$,
- thus giving me the final explicit non-zero polynomial $Q - ZP \in k[t]$ which has α as a root.

Baby Combinatorics Appendix

Lemma 0.7: Dickson's Lemma

Let $k \in \mathbb{N}^+$. Any subset $S \subseteq \mathbb{N}^k$ (really the poset $(\mathbb{N}^k, \leqslant_{\text{dom}})$ of domination in every coordinate) has finitely many minimal elements.
Equivalently, there are no infinite antichains in $(\mathbb{N}^3, \leqslant_{\text{dom}})$. ichinger
Equivalently, any infinite list $(v_1, v_2, \ldots)$ of integer vectors $v_i \in \mathbb{N}^k$, has $i \neq j$ s.t. $v_i \leqslant_{\text{dom}} v_j$.

We show in fact the superficially stronger statement that any such infinite list $(v_1, v_2, \ldots)$, has $i < j$ s.t. $v_i \leqslant_{\text{dom}} v_j$.

Proof: I was delighted to learn that there is a clever proof (§3 of Aichinger–Aichinger 2020; they claimed originality) via the Infinite Ramsey Theorem (Lemma 0.8 below). Suppose f.s.o.c. that every $i < j$ has v_i strictly exceeding v_j in some coordinate $c_{(i,j)} \in [1..k]$. For $i < j$, we color the edge (i, j) (in the complete graph on $\mathbb{N}$) by $c_{(i,j)} \in [1..k]$.

Well, by Infinite Ramsey, we have some $c^\star \in [1..k]$ s.t. there is an infinite sequence of vertices $i_1 < i_2 < \ldots$ so that v_{i_1} strictly exceeds v_{i_2} in the $c^\star$-th coordinate, and so on, i.e.

$$(v_{i_1})_{c^\star} > (v_{i_2})_{c^\star} > (v_{i_3})_{c^\star} > \ldots;$$

but there are no infinite decreasing chains of natural numbers! ※■

For sake of completeness, we include the following classic proof, mainly because we wish to share our illustration for the proof. (A query "infinite ramsey theorem picture" on Google reveals that the same diagram has appeared independently elsewhere, e.g. in these online notes by Jungic.)

Lemma 0.8: Infinite Ramsey Theorem

Consider the complete graph on vertex set $V := \mathbb{N}$, and a k-coloring of the edges. Then there is an infinite subset $U \subseteq V$ so that all edges "inside U" have the same color.

Proof: in a slogan, "$\omega + 1$ applications of the (infinite) pigeonhole principle". Consider node ①: it has infinitely many edges of some color c_1. Color node ① color c_1, and keep the infinite cloud $C_1 \subseteq V$ of vertices reached from ① by a c_1-edge.

Pick any (say, the minimal index) node n_2 in the previous cloud C_1. It has infinitely many edges of some color c_2. Color node n_2 color c_1, and keep the infinite cloud $C_2 \subseteq C_1$ of vertices reached from n_2 by a c_2-edge.

Continue in this way until we get an infinite sequence of nodes ① =: $n_1 < n_2 < \ldots$, where if n_i is colored color c_i, then all edges (n_i, n_j) for $i < j$ lie in the cloud C_i colored c_i, and hence are colored c_i.

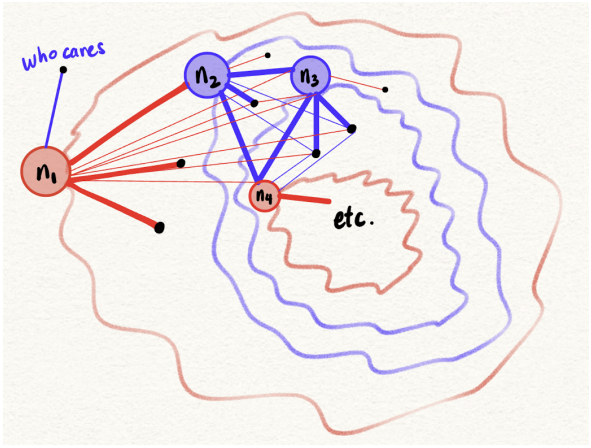


Figure 1: Infnite Ramsey Theorem

Above accounts for ω many applications of infinite pigeonhole, but we promised $\omega + 1$. Well, to finish, there are infinitely many nodes in the sequence $\textcircled{n_1} < \textcircled{n_2} < \ldots$ of the same color $c^\star$, so by one last application of infinite pigeonhole, there is an infinite subsequence, called $U \subseteq \mathbb{N}$ say, that form an infinite monochromatic clique (indeed any edge (u_i, u_j) will be colored $c^\star$ by the argument in the penultimate paragraph).

Appendix for Riemann sums for weighted lattice point counts

This section is about estimating M, W, E_p by Riemann sums, and verifying the key inequality (5). Perhaps this would make for good exercises, but as the details are tedious, I copy paste them below directly from GPT 6.

Our monomials satisfy

$$i + 32(j + k) \leqslant D, \qquad i, j, k \geqslant 0.$$

The simplest way to count them is to group together monomials with the same value of $j + k$.

Put

$$K = \left\lfloor \frac{D}{32} \right\rfloor, \qquad L_t = D - 32t \quad (0 \leqslant t \leqslant K).$$

For a fixed $t = j + k$, there are exactly $t + 1$ choices of (j, k), namely $(0, t), (1, t - 1), \ldots, (t, 0)$. For each choice, i ranges from 0 to L_t.

This gives exact formulas immediately.

C1. Count the monomials: M

For each t, there are $(t + 1)(L_t + 1)$ monomials. Therefore

$$\boxed{M = \sum_{t=0}^{K} (t + 1)(D - 32t + 1).} \tag{C1}$$

We can even evaluate this sum exactly. Using

$$\sum_{t=0}^{K} (t + 1) = \frac{(K + 1)(K + 2)}{2}, \qquad \sum_{t=0}^{K} t(t + 1) = \frac{K(K + 1)(K + 2)}{3},$$

we obtain

$$\boxed{M = \frac{(K + 1)(K + 2)}{6} \big(3D + 3 - 64K\big).}$$

This is an exact integer formula for every $D \geqslant 0$.

C2. Count the total ordinary degree: W

By definition,

$$W = \sum_{i+32(j+k) \leqslant D} (i + j + k).$$

For fixed $t = j + k$, the ordinary degree is $i + t$. There are $t + 1$ pairs (j, k), so the contribution is

$$(t + 1) \sum_{i=0}^{L_t} (i + t).$$

Using $\sum_{i=0}^{L} i = L(L + 1)/2$, we obtain the exact formula

$$\boxed{W = \sum_{t=0}^{K} (t + 1) \left[\frac{L_t(L_t + 1)}{2} + t(L_t + 1) \right].} \tag{C2}$$

For clarity, separate this into two contributions:

$$I = \sum_{t=0}^{K} (t+1)\frac{L_t(L_t+1)}{2}, \qquad J = \sum_{t=0}^{K} t(t+1)(L_t+1).$$

Thus $W = I + J$.

The quantity I is the sum of all the i's, and J is the sum of all the $j+k$'s. The distinction matters because Frobenius only exploits the i-coordinate.

C3. Count the Frobenius savings: E_p

By definition,

$$\boxed{E_p = \sum_{t=0}^{K} (t+1) \sum_{i=0}^{L_t} \left\lfloor \frac{i}{p} \right\rfloor.} \tag{C3}$$

This is an exact finite formula. There is also a closed expression for its inner sum.

Write $L_t = pa_t + b_t$, where $0 \leqslant b_t < p$. The values of $\lfloor i/p \rfloor$, for $0 \leqslant i \leqslant L_t$, consist of p copies of each integer $0, \ldots, a_t - 1$, followed by $b_t + 1$ copies of a_t.

Therefore

$$\boxed{E_p = \sum_{t=0}^{K} (t+1) \left[\frac{p\, a_t(a_t-1)}{2} + a_t(b_t+1) \right].} \tag{C4}$$

This is completely explicit, with no asymptotic estimate.

C4. Numerical check: $D = 64$

Here $K = 2$, so the three layers are

$t = j + k$	L_t	Number of monomials
0	64	65
1	32	66
2	0	3

Thus

$$M = 65 + 66 + 3 = 134.$$

For I,

$$I = \frac{64 \cdot 65}{2} + 2\frac{32 \cdot 33}{2} = 3136.$$

For J,

$$J = 0 + 2 \cdot 33 + 2 \cdot 3 = 72.$$

Consequently, $W = 3208$.

Finally, for $p = 37$, the only integers $i \leqslant 64$ with $\lfloor i/37 \rfloor > 0$ are $37, \ldots, 64$. They all occur in the $t = 0$ layer, so

$$E_{37} = 28.$$

This confirms the numerical example in Appendix A2.

0.4.2 Derive the asymptotics, with errors justified

Now assume $D \to \infty$. All implied constants below are absolute because the weight 32 is fixed.

The standard estimates $\sum_{t \leqslant K} t^r = K^{r+1}/(r+1) + O(K^r)$, for $r = 0, 1, 2, 3$, follow from the elementary power-sum formulas. We may therefore replace our sums by the corresponding integrals with the stated errors.

C5. The number of monomials

From (C1),

$$M = \sum_{t=0}^{K} (t+1)(D - 32t + 1).$$

Replacing the summand by $t(D - 32t)$ costs $O(D)$ per term. There are $O(D)$ terms, so the total error is $O(D^2)$.

The remaining sum differs from its integral by $O(D^2)$. Hence

$$\begin{aligned} M &= \int_0^{D/32} t(D - 32t)\, dt + O(D^2) \\ &= \frac{D^3}{6 \cdot 32^2} + O(D^2). \end{aligned} \tag{7}$$

Thus

$$\boxed{M = \frac{D^3}{6144} + O(D^2).}$$

C6. The sum of the first exponents

Recall

$$I = \frac{1}{2} \sum_{t=0}^{K} (t+1) L_t (L_t + 1).$$

Replacing the summand by $\frac{1}{2} t(D - 32t)^2$ costs $O(D^2)$ per term, hence $O(D^3)$ overall.

The integral calculation is

$$\begin{aligned} I &= \frac{1}{2} \int_0^{D/32} t(D - 32t)^2\, dt + O(D^3) \\ &= \frac{D^4}{2 \cdot 32^2} \int_0^1 s(1-s)^2\, ds + O(D^3). \end{aligned} \tag{8}$$

The elementary integral is

$$\int_0^1 s(1-s)^2\, ds = \frac{1}{2} - \frac{2}{3} + \frac{1}{4} = \frac{1}{12}.$$

Therefore

$$\boxed{I = \frac{D^4}{24 \cdot 32^2} + O(D^3).}$$

Comparing (7) and (8),

$$\boxed{\frac{I}{M} = \frac{D}{4} + O(1).}$$

This is the rigorous version of the statement that the average i-exponent is $D/4$.

C7. The other two exponents

Recall

$$J = \sum_{t=0}^{K} t(t+1)(L_t+1).$$

Replace the summand by $t^2(D-32t)$. As before, this produces an $O(D^3)$ error.

Thus

$$\begin{aligned} J &= \int_0^{D/32} t^2(D-32t)\,dt + O(D^3) \\ &= \frac{D^4}{32^3}\int_0^1 s^2(1-s)\,ds + O(D^3). \end{aligned} \tag{9}$$

Since the integral equals $1/12$,

$$\boxed{J = \frac{D^4}{12\cdot 32^3} + O(D^3).}$$

Combining $W = I + J$, we obtain

$$\begin{aligned} W &= \frac{D^4}{24\cdot 32^2}\left(1+\frac{2}{32}\right) + O(D^3) \\ &= \frac{17D^4}{384\cdot 32^2} + O(D^3). \end{aligned}$$

In particular,

$$\boxed{\frac{W}{I} = \frac{17}{16} + O(D^{-1}).} \tag{10}$$

This ratio, rather than the individual leading coefficients, is what matters for the final numerical inequality.

C8. Uniform estimate for E_p

Here there is a much simpler method than expanding the exact formula (C4).

For every integer $i \geqslant 0$,

$$\frac{i}{p} - 1 < \left\lfloor \frac{i}{p} \right\rfloor \leqslant \frac{i}{p}.$$

Summing over all our monomials gives

$$\boxed{\frac{I}{p} - M < E_p \leqslant \frac{I}{p}.} \tag{11}$$

Therefore, uniformly for every prime p,

$$\boxed{E_p = \frac{D^4}{24\cdot 32^2 p} + O(D^3).} \tag{12}$$

We have now proved all three counting estimates with explicit leading constants and uniform errors.

0.4.3 Finish Appendix C: the prime-sum hypothesis implies (PC)

This is where uniformity in X must be checked.

Write

$$T_S = \sum_{p \in S} \frac{\log p}{p}, \qquad L_S = \sum_{p \in S} \log p.$$

By (11),

$$\sum_{p \in S} E_p \log p \geqslant I T_S - M L_S.$$

Consequently, the exact pigeonhole condition (PC) follows if

$$I T_S - M L_S > W \log B + M \log(32M\sqrt{m}).$$

Since $I > 0$, this is equivalent to the sufficient condition

$$\boxed{T_S > \frac{W}{I} \log B + \frac{M}{I}\big[L_S + \log(32M\sqrt{m})\big].} \tag{13}$$

We now estimate its right-hand side.

From (7)–(10),

$$\frac{W}{I} = \frac{17}{16} + O(D^{-1}), \qquad \frac{M}{I} = \frac{4}{D} + O(D^{-2}).$$

The elementary bound $L_S \leqslant X \log X$ requires no theorem about prime distribution: there are at most X primes, and each has $\log p \leqslant \log X$.

Also, $M = O(D^3)$, and our choices $N = \lceil X^{10/13} \rceil$, $m \leqslant X^{1/13}$ imply

$$\log B = O(\log X), \qquad \log(32M\sqrt{m}) = O(\log X).$$

It follows that the right-hand side of (13) is at most

$$\frac{17}{16} \log B + O\left(\frac{X \log X}{D}\right). \tag{14}$$

But our degree satisfies $D \asymp X^{40/39}$. Thus

$$\frac{X \log X}{D} = O(X^{-1/39} \log X) = o(1).$$

The error tends to zero uniformly as $X \to \infty$, and consequently uniformly for all $X \geqslant m^{13}$ as $m \to \infty$.

Hence (13) — i.e. (5) — certainly holds whenever

$$T_S > \frac{9}{8} \log B,$$

because

$$\frac{9}{8} - \frac{17}{16} = \frac{1}{16} > 0,$$

and $\log B \to \infty$.

References

[AA20] Erhard Aichinger and Florian Aichinger. "Dickson's Lemma, Higman's Theorem and Beyond: A survey of some basic results in order theory". In: *Expositiones Mathematicae* 38.4 (2020), pp. 537–547. DOI: 10.1016/j.exmath.2019.05.003. URL: https://www.sciencedirect.com/science/article/pii/S0723086918302226.

[All05] Daniel Allcock. *Hilbert's Nullstellensatz*. Expository note, 3 pp. Jan. 9, 2005. URL: https://web.ma.utexas.edu/users/allcock/expos/nullstellensatz3.pdf.

[Aza15] Alborz Azarang. *A one-line undergraduate proof of Zariski's lemma and Hilbert's nullstellensatz*. 2015. arXiv: 1506.08376. URL: https://arxiv.org/abs/1506.08376.

[Ban26] William D. Banks. *On the OpenAI proof that Landau–Siegel zeros do not exist*. Oct. 11, 2026. arXiv: 2610.15483 [math.NT]. URL: https://arxiv.org/abs/2610.15483.

[BS23] Thomas F. Bloom and Olof Sisask. "The Kelley–Meka bounds for sets free of three-term arithmetic progressions". In: *Essential Number Theory* 2 (2023). Linked manuscript is version 3 (2025), pp. 15–44. DOI: 10.2140/ent.2023.2.15. arXiv: 2302.07211. URL: https://arxiv.org/pdf/2302.07211v3.

[Con] Keith Conrad. *Maximal ideals in polynomial rings*. Expository notes; see Theorems 2.7 and 2.11. URL: https://kconrad.math.uconn.edu/blurbs/ringtheory/maxideal-polyring.pdf.

[DF04] David S. Dummit and Richard M. Foote. *Abstract Algebra*. 3rd ed. See discussion of Noether normalization near p. 699; manuscript linked to a scanned PDF. John Wiley & Sons, 2004. URL: https://danielrui.com/texts/dummit&foote.pdf#page=713.

[EG17] Jordan S. Ellenberg and Dion Gijswijt. "On large subsets of F_q^n with no three-term arithmetic progression". In: *Annals of Mathematics* 185.1 (2017), pp. 339–343. arXiv: 1605.09223. URL: https://arxiv.org/abs/1605.09223.

[Ful69] William Fulton. *Algebraic Curves: An Introduction to Algebraic Geometry*. Originally published in 1969; freely available revised edition online. New York: W. A. Benjamin, 1969. URL: https://dept.math.lsa.umich.edu/~wfulton/CurveBook.pdf.

[Gre26] Ben Green. *On Landau–Siegel zeros*. Hexagon preprint, version 1, hexagon:2610.00311v1. Oct. 2026. URL: https://hexagonmath.org/pdf/2610.00311v1.

[Jun20] Veselin Jungić. *Introduction to Ramsey Theory: Lecture Notes for Undergraduate Course*. See Section 2.4: Ramsey's Theorem, Infinite Case. 2020. URL: https://opentext.uleth.ca/Ramsey/sec_Infinite.html.

[KM23] Zander Kelley and Raghu Meka. *Strong Bounds for 3-Progressions*. 2023. arXiv: 2302.05537. URL: https://arxiv.org/abs/2302.05537.

[Ker26] Bryce Kerr. *A simpler proof that there are no Landau–Siegel zeros*. Oct. 15, 2026. arXiv: 2610.21099 [math.NT]. URL: https://arxiv.org/abs/2610.21099.

[Mas81] David W. Masser. "On polynomials and exponential polynomials in several complex variables". In: *Inventiones Mathematicae* 63 (1981), pp. 81–95. DOI: 10.1007/BF01389194.

[Mat] Math Curator Zanachan. *Noether normalization and the Nullstellensatz (video)*. Descriptive title taken from the manuscript's context; exact video title not verified. YouTube. URL: https://www.youtube.com/watch?v=MU_nw2Pg0YQ.

[] *MathOverflow answer concerning short proofs of Zariski's lemma and the Nullstellensatz*. Answer link quoted in the source; answer metadata not independently confirmed. MathOverflow. URL: https://mathoverflow.net/a/222638/112504.

[McG24] William McGovern. *Lecture 5-10: Noether normalization and the proof of the Nullstellensatz*. University of Washington course notes. May 10, 2024. URL: https://sites.math.washington.edu/~mcgovern/506%20pdf%20files%202024/506.5-10.pdf.

[Ope26] OpenAI. *Uniform exclusion of Landau–Siegel zeros*. Preprint. OpenAI Math Release. Oct. 1, 2026. URL: https://github.com/openai/math/blob/main/preprints/Uniform-exclusion-of-Landau-Siegel-zeros-October-1-2026/paper.pdf.

[Tao10] Terence Tao. *Answer to "How to recognise that the polynomial method might work"*. MathOverflow. Oct. 25, 2010. URL: https://mathoverflow.net/questions/43317/how-to-recognise-that-the-polynomial-method-might-work.

[Tao24] Terence Tao. *An abridged proof of Marton's conjecture*. The manuscript links to the blog's polynomial Freiman–Ruzsa tag. What's New (blog). June 22, 2024. URL: https://terrytao.wordpress.com/2024/06/22/an-abridged-proof-of-martons-conjecture/.

[Tao26] Terence Tao. *A sketch of the proof of non-existence of Siegel zeroes*. What's New (blog). Oct. 19, 2026. URL: https://terrytao.wordpress.com/2026/10/19/a-sketch-of-the-proof-of-non-existence-of-siegel-zeroes/.

[Wika] Wikipedia contributors. *Discriminant — Fundamental discriminants*. Wikipedia. URL: https://en.wikipedia.org/wiki/Discriminant#Fundamental_discriminants (visited on 10/20/2026).

[Wikb] Wikipedia contributors. *Euler's criterion*. Wikipedia. URL: https://en.wikipedia.org/wiki/Euler%27s_criterion (visited on 10/20/2026).

[Wikc] Wikipedia contributors. *Polynomial interpolation*. Wikipedia. URL: https://en.wikipedia.org/wiki/Polynomial_interpolation (visited on 10/20/2026).

[Wikd] Wikipedia contributors. *Unit (ring theory) — Associatedness*. Wikipedia. URL: https://en.wikipedia.org/wiki/Unit_(ring_theory)#Associatedness (visited on 10/20/2026).