Pairs of commuting matrices

Victor Wang
(joint work with Tim Browning and Will Sawin)

IST Austria and IoM Academia Sinica

100 Years of the Circle Method, IML, July 2025

This project has received funding from the European Union's Horizon 2020 research and innovation program

under the Marie Skłodowska-Curie Grant Agreement No. 101034413

Diophantine equations at the boundary

Some equations in $\mathbb{Z}^s$ lie at the boundary between heuristic solubility and paucity: in practice, equations of degree s, with reasonable geometry (log Calabi–Yau?). Examples:

- Factoring: $xy = k$, for $k \in \mathbb{N}$. This has $O_\epsilon(k^\epsilon)$ solutions. It is open (?) to factor in $(\log k)^{O(1)}$ digital steps.
- Markov–Hurwitz: $x_1^2 + \cdots + x_s^2 - ax_1 \cdots x_s = k$, for $s \geq 3$, $a \in \mathbb{N}$, $k \in \mathbb{Z}$. If $k - s + 2, k - s - 1 \neq \square$, then this either has 0 or $\sim c(\log T)^{\beta(s)}$ solutions of height $\max(|x_\bullet|) \leq T$ as $T \to \infty$,[1] where $c = c(s, a, k) > 0$, $\beta(3) = 2$, and $\beta(s)$ may well be transcendental for $s \geq 4$.

[1] [Gurwood 1976, ..., Baragar 1998, Gamburd–Magee–Ronan 2019]

Diophantine equations at the boundary

Some equations in $\mathbb{Z}^s$ lie at the boundary between heuristic solubility and paucity: in practice, equations of degree s, with reasonable geometry (log Calabi–Yau?). Examples:

- Factoring: $xy = k$, for $k \in \mathbb{N}$. This has $O_\epsilon(k^\epsilon)$ solutions. It is open (?) to factor in $(\log k)^{O(1)}$ digital steps.
- Markov–Hurwitz: $x_1^2 + \cdots + x_s^2 - ax_1 \cdots x_s = k$, for $s \geq 3$, $a \in \mathbb{N}$, $k \in \mathbb{Z}$. If $k - s + 2, k - s - 1 \neq \square$, then this either has 0 or $\sim c(\log T)^{\beta(s)}$ solutions of height $\max(|x_\bullet|) \leq T$ as $T \to \infty$,[1] where $c = c(s, a, k) > 0$, $\beta(3) = 2$, and $\beta(s)$ may well be transcendental for $s \geq 4$.
- Sums of three cubes (you): $x^3 + y^3 + z^3 = k$. For $k = 42$ the only known solution [Booker–Sutherland 2019] is

$$42 = (-80538738812075974)^3 + (80435758145817515)^3 + (1260$$

[1] [Gurwood 1976, . . . , Baragar 1998, Gamburd–Magee–Ronan 2019]

Diophantine equations at the boundary (cont'd)

Example:

- Factoring: $xy = k$, for $k \in \mathbb{N}$. This has $O_\epsilon(k^\epsilon)$ solutions.

Let $n \geq 2$. Let $\mathrm{M}_n(\mathbb{Z})$ be the set of $n \times n$ matrices with entries in $\mathbb{Z}$. Let $\mathrm{V}_n(\mathbb{Z}) = \{A \in \mathrm{M}_n(\mathbb{Z}) : \mathrm{tr}(A) = 0\}$. Fix $K \in \mathrm{V}_n(\mathbb{Z})$.

- Trace-zero commutators: $XY - YX = K$, to be solved in $(X, Y) \in \mathrm{V}_n(\mathbb{Z})^2$. This can be viewed as a system of $n^2 - 1$ quadratic equations in $2(n^2 - 1)$ integer variables. This is known to be soluble for $n \geq 3$ [Stasinski 2018], even with principal ideal domains in place of $\mathbb{Z}$, and for $n \geq 2$ if we drop the trace-zero condition [Laffey–Reams 1994].

Diophantine equations at the boundary (cont'd)

Example:

- Factoring: $xy = k$, for $k \in \mathbb{N}$. This has $O_\epsilon(k^\epsilon)$ solutions.

Let $n \geq 2$. Let $\mathrm{M}_n(\mathbb{Z})$ be the set of $n \times n$ matrices with entries in $\mathbb{Z}$. Let $\mathrm{V}_n(\mathbb{Z}) = \{A \in \mathrm{M}_n(\mathbb{Z}) : \mathrm{tr}(A) = 0\}$. Fix $K \in \mathrm{V}_n(\mathbb{Z})$.

- Trace-zero commutators: $XY - YX = K$, to be solved in $(X, Y) \in \mathrm{V}_n(\mathbb{Z})^2$. This can be viewed as a system of $n^2 - 1$ quadratic equations in $2(n^2 - 1)$ integer variables. This is known to be soluble for $n \geq 3$ [Stasinski 2018], even with principal ideal domains in place of $\mathbb{Z}$, and for $n \geq 2$ if we drop the trace-zero condition [Laffey–Reams 1994].

It is reasonable to guess that for most K (perhaps all $K \neq 0$), the number of solutions (X, Y) of height $\leq T$ is $O_{K,\epsilon}(T^\epsilon)$. What about for $K = 0$?

Diophantine equations at the boundary (cont'd^2)

Let $n \geq 2$. Let $\mathrm{M}_n(\mathbb{Z})$ be the set of $n \times n$ matrices with entries in $\mathbb{Z}$. Let $\mathrm{V}_n(\mathbb{Z}) = \{A \in \mathrm{M}_n(\mathbb{Z}) : \mathrm{tr}(A) = 0\}$. Fix $K \in \mathrm{V}_n(\mathbb{Z})$.

- Trace-zero commutators: $XY - YX = K$, to be solved in $(X, Y) \in \mathrm{V}_n(\mathbb{Z})^2$. This can be viewed as a system of $n^2 - 1$ quadratic equations in $2(n^2 - 1)$ integer variables.

It is reasonable to guess that for most K (perhaps all $K \neq 0$), the number of solutions (X, Y) of height $\leq T$ is $O_{K,\epsilon}(T^\epsilon)$. However:

- Even after removing special solutions to $XY = YX$ like $Y = X^r$, there may well remain $\geq T^\delta - O_n(1)$ solutions of height $\leq T$. This is because the variety $XY = YX$ has excess dimension (or morally, "redundant equations").

Diophantine equations at the boundary (cont'd^2)

Let $n \geq 2$. Let $\mathrm{M}_n(\mathbb{Z})$ be the set of $n \times n$ matrices with entries in $\mathbb{Z}$. Let $\mathrm{V}_n(\mathbb{Z}) = \{A \in \mathrm{M}_n(\mathbb{Z}) : \mathrm{tr}(A) = 0\}$. Fix $K \in \mathrm{V}_n(\mathbb{Z})$.

- Trace-zero commutators: $XY - YX = K$, to be solved in $(X, Y) \in \mathrm{V}_n(\mathbb{Z})^2$. This can be viewed as a system of $n^2 - 1$ quadratic equations in $2(n^2 - 1)$ integer variables.

It is reasonable to guess that for most K (perhaps all $K \neq 0$), the number of solutions (X, Y) of height $\leq T$ is $O_{K,\epsilon}(T^\epsilon)$. However:

- Even after removing special solutions to $XY = YX$ like $Y = X^r$, there may well remain $\geq T^\delta - O_n(1)$ solutions of height $\leq T$. This is because the variety $XY = YX$ has excess dimension (or morally, "redundant equations").

Other non-examples that resemble examples:

- The cubic surfaces $x^2 + y^3 + z^3 = k$ (Vaughan, Brüdern) and $100x + 10y + z = (x + y + z)^3$ (PU/IAS NTS Zoom password), due to being "insufficiently cubic".

Dimensions of commutator fibers

In what sense does $XY = YX$ have excess dimension? For simplicity, forget the trace-zero condition. Let $n \geq 2$. Let $\mathrm{M}_n(\mathbb{C})$ be the set of $n \times n$ matrices with entries in $\mathbb{C}$.

Theorem (Motzkin–Taussky 1955)

The variety $XY = YX$ in $\mathrm{M}_n(\mathbb{C})^2$ has dimension $n^2 + n$.

[Feit–Fine 1960][2] gave a formula for the number of $\mathbb{F}_q$-points on the commuting variety, which by the Lang–Weil estimate gives another proof of the Motzkin–Taussky theorem.

[2] see also [Fulman–Guralnick 2018]

Dimensions of commutator fibers

In what sense does $XY = YX$ have excess dimension? For simplicity, forget the trace-zero condition. Let $n \geq 2$. Let $\mathrm{M}_n(\mathbb{C})$ be the set of $n \times n$ matrices with entries in $\mathbb{C}$.

Theorem (Motzkin–Taussky 1955)

The variety $XY = YX$ in $\mathrm{M}_n(\mathbb{C})^2$ has dimension $n^2 + n$.

[Feit–Fine 1960][2] gave a formula for the number of $\mathbb{F}_q$-points on the commuting variety, which by the Lang–Weil estimate gives another proof of the Motzkin–Taussky theorem.

Theorem (Browning–Sawin–W. 2024)

$\dim\{(X, Y) \in \mathrm{M}_n(\mathbb{C})^2 : XY - YX = M\} \leq n^2 + 1$ *if* $M \neq 0$.

[Larsen–Lu 2021] proved a similar result for $(g, h) \in \mathrm{SL}_n(\mathbb{C})^2$.

[2] see also [Fulman–Guralnick 2018]

Proof strategy

Theorem (Browning–Sawin–W. 2024)

$\dim\{(X, Y) \in \mathrm{M}_n(\mathbb{C})^2 : XY - YX = M\} \leq n^2 + 1$ *if* $M \neq 0$.

[Larsen–Lu 2021] proved a similar result for $(g, h) \in \mathrm{SL}_n(\mathbb{C})^2$. One might wonder if the SL_n result can imply the M_n result by a Lie-theoretic limiting process; however, Larsen told us that he expected the M_n case to be harder than the SL_n case.

Proof strategy

Theorem (Browning–Sawin–W. 2024)

$\dim\{(X,Y) \in \mathrm{M}_n(\mathbb{C})^2 : XY - YX = M\} \leq n^2 + 1$ *if* $M \neq 0$.

[Larsen–Lu 2021] proved a similar result for $(g,h) \in \mathrm{SL}_n(\mathbb{C})^2$.
One might wonder if the SL_n result can imply the M_n result by a Lie-theoretic limiting process; however, Larsen told us that he expected the M_n case to be harder than the SL_n case.
Nonetheless, both results can be proven by $\mathbb{F}_q$-point counting.
The key difference, as it turns out, is that symmetry is built-in for SL_n, but must be created by hand for M_n. Let

$$\Sigma(M) := (1 - q^{-1}) \sum_{U,V \in \mathrm{M}_n(\mathbb{F}_q)} \mathbf{1}_{UV - VU = M}.$$

By the Lang–Weil estimate, for suitable $q \to \infty$,

$$\Sigma(M) \geq (1 + O_n(q^{-1/2}))\, q^{\dim\{(U,V) \in \mathrm{M}_n(\overline{\mathbb{F}}_p)^2 : UV - VU = M\}}.$$

The polygon method (finite-field circle method)

Let

$$\Sigma(M) := (1 - q^{-1}) \sum_{U,V \in \mathrm{M}_n(\mathbb{F}_q)} \mathbf{1}_{UV - VU = M}.$$

Let $\psi(\cdot) := e_p(\mathrm{tr}_{\mathbb{F}_q/\mathbb{F}_p}(\cdot))$ on $\mathbb{F}_q$. By Fourier orthogonality,

$$\begin{aligned}\Sigma(M) &= \frac{1 - q^{-1}}{\#\mathrm{M}_n(\mathbb{F}_q)} \sum_{U,V,Z \in \mathrm{M}_n(\mathbb{F}_q)} \psi(\mathrm{tr}(Z(UV - VU - M))) \\ &= (1 - q^{-1}) \sum_{\substack{V,Z \in \mathrm{M}_n(\mathbb{F}_q) \\ VZ - ZV = 0}} \psi(\mathrm{tr}(-ZM)),\end{aligned}$$

where we average over U after writing $\mathrm{tr}(ZUV) = \mathrm{tr}(VZU)$.
Since $VZ - ZV = 0$ is homogeneous,

$$\Sigma(M) = \sum_{\substack{V,Z \in \mathrm{M}_n(\mathbb{F}_q) \\ VZ - ZV = 0}} (\mathbf{1}_{\mathrm{tr}(ZM) = 0} - q^{-1}).$$

Blooper reel

From the previous slide, we have

$$\Sigma(M) = \sum_{\substack{V,Z \in \mathrm{M}_n(\mathbb{F}_q) \\ VZ - ZV = 0}} (\mathbf{1}_{\mathrm{tr}(ZM)=0} - q^{-1}).$$

Summing over $V \in C(Z) := \{V \in \mathrm{M}_n(\overline{\mathbb{F}}_q) : VZ - ZV = 0\}$,

$$\Sigma(M) = \sum_{Z \in \mathrm{M}_n(\mathbb{F}_q)} q^{\dim C(Z)} (\mathbf{1}_{\mathrm{tr}(ZM)=0} - q^{-1}).$$

We find on averaging over conjugates of Z that

$$\Sigma(M) = \sum_{Z \in \mathrm{M}_n(\mathbb{F}_q)} q^{\dim C(Z)} E(Z, M),$$

where $E(Z, M) := \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)} (\mathbf{1}_{\mathrm{tr}(gZg^{-1}M)=0} - q^{-1})$.

Blooper reel (cont'd)

If $C(Z) := \{V \in \mathrm{M}_n(\overline{\mathbb{F}}_q) : VZ - ZV = 0\}$, we have

$$\Sigma(M) = \sum_{Z \in \mathrm{M}_n(\mathbb{F}_q)} q^{\dim C(Z)} E(Z, M),$$

where $E(Z, M) := \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{\mathrm{tr}(gZg^{-1}M)=0} - q^{-1}) \ll 1$ is an example of a *Lie-algebra orbital integral*.

- Generically, $\dim C(Z) = n$ and we can bound the tail $\Pr(\dim C(Z) > P)$ for any parameter $P \geq n$.
- For $\dim C(Z) \leq P$, we use Cauchy–Schwarz and a variance bound $\sum_{Z \in \mathrm{M}_n(\mathbb{F}_q)} E(Z, M)^2 \ll_n q^{\dim C(M)} \leq q^{n^2 - \delta_n}$, noting that M is non-scalar (since $\mathrm{tr}(M) = 0 \neq M$).

Blooper reel (cont'd)

If $C(Z) := \{V \in \mathrm{M}_n(\overline{\mathbb{F}}_q) : VZ - ZV = 0\}$, we have

$$\Sigma(M) = \sum_{Z \in \mathrm{M}_n(\mathbb{F}_q)} q^{\dim C(Z)} E(Z, M),$$

where $E(Z, M) := \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{\mathrm{tr}(gZg^{-1}M)=0} - q^{-1}) \ll 1$ is an example of a *Lie-algebra orbital integral*.

- Generically, $\dim C(Z) = n$ and we can bound the tail $\Pr(\dim C(Z) > P)$ for any parameter $P \geq n$.
- For $\dim C(Z) \leq P$, we use Cauchy–Schwarz and a variance bound $\sum_{Z \in \mathrm{M}_n(\mathbb{F}_q)} E(Z, M)^2 \ll_n q^{\dim C(M)} \leq q^{n^2 - \delta_n}$, noting that M is non-scalar (since $\mathrm{tr}(M) = 0 \neq M$).

Optimizing P, we get $\Sigma(M) \ll_n q^{n^2+n-\sqrt{n}+O(1)}$, whence
$\dim\{(U, V) \in \mathrm{M}_n(\overline{\mathbb{F}}_p)^2 : UV - VU = M\} \leq n^2 + n - \sqrt{n} + O(1)$,
$\#\{(X, Y) \in \mathrm{M}_n([-T, T])^2 : XY = YX\} \ll_n T^{n^2+n-\sqrt{n}+O(1)}$.
(More on the latter later. It improves on the *dimension growth bound* $O_n(T^{n^2+n-1})$ for large n. This was our original goal.)

The right button

We pressed L and we should have pressed R instead. We had

$$\Sigma(M) = \sum_{\substack{V,Z \in \mathrm{M}_n(\mathbb{F}_q) \\ VZ - ZV = 0}} (\mathbf{1}_{\mathrm{tr}(ZM)=0} - q^{-1}).$$

Summing over $Z \in C(V) := \{Z \in \mathrm{M}_n(\overline{\mathbb{F}}_q) : VZ - ZV = 0\}$,

$$\begin{aligned} \Sigma(M) &= \sum_{V \in \mathrm{M}_n(\mathbb{F}_q)} q^{\dim C(V)} (q^{-1} \mathbf{1}_{C(V) \not\subseteq M^\perp} + \mathbf{1}_{C(V) \subseteq M^\perp} - q^{-1}) \\ &= \sum_{V \in \mathrm{M}_n(\mathbb{F}_q)} q^{\dim C(V)} (1 - q^{-1}) \mathbf{1}_{C(V) \subseteq M^\perp}, \end{aligned}$$

where $M^\perp := \{A \in \mathrm{M}_n(\overline{\mathbb{F}}_q) : \mathrm{tr}(AM) = 0\}$. By conjugation,

$$\Sigma(M) = \sum_{V \in \mathrm{M}_n(\mathbb{F}_q)} q^{\dim C(V)} (1 - q^{-1}) L(V, M),$$

where $L(V, M) := \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)} (\mathbf{1}_{C(gVg^{-1}) \subseteq M^\perp})$.

If $C(V) := \{Z \in \mathrm{M}_n(\overline{\mathbb{F}}_q) : VZ - ZV = 0\}$, we have

$$\Sigma(M) = \sum_{V \in \mathrm{M}_n(\mathbb{F}_q)} q^{\dim C(V)}(1 - q^{-1})L(V, M),$$

where $L(V, M) := \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{C(gVg^{-1}) \subseteq M^{\perp}}) \leq 1$. If $K_n(\mathbb{F}_q) \subseteq \mathrm{M}_n(\mathbb{F}_q)$ denotes a complete set of representatives for conjugation by $\mathrm{GL}_n(\mathbb{F}_q)$, then breaking $\mathrm{M}_n(\mathbb{F}_q)$ into orbits gives

$$\Sigma(M) \ll_n q^{n^2} \sum_{V \in K_n(\mathbb{F}_q)} L(V, M).$$

If $C(V) := \{Z \in \mathrm{M}_n(\overline{\mathbb{F}}_q) : VZ - ZV = 0\}$, we have

$$\Sigma(M) = \sum_{V \in \mathrm{M}_n(\mathbb{F}_q)} q^{\dim C(V)}(1 - q^{-1})L(V, M),$$

where $L(V, M) := \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{C(gVg^{-1}) \subseteq M^\perp}) \leq 1$. If $K_n(\mathbb{F}_q) \subseteq \mathrm{M}_n(\mathbb{F}_q)$ denotes a complete set of representatives for conjugation by $\mathrm{GL}_n(\mathbb{F}_q)$, then breaking $\mathrm{M}_n(\mathbb{F}_q)$ into orbits gives

$$\Sigma(M) \ll_n q^{n^2} \sum_{V \in K_n(\mathbb{F}_q)} L(V, M).$$

Let n_V be the number of distinct eigenvalues of V in $\overline{\mathbb{F}}_q$. Since $\#\{V \in K_n(\mathbb{F}_q) : n_V = d\} \ll_n q^d$, the following lemma implies $\Sigma(M) \ll_n q^{n^2+1}\mathbf{1}_{\mathrm{tr}(M)=0} + q^{n^2+n}\mathbf{1}_{M=0}$, as desired.

Lemma (Orbital integral bound)

Let $V, M \in \mathrm{M}_n(\mathbb{F}_q)$, not necessarily with trace zero. Then

$$L(V, M) \ll_n q^{1-n_V}\mathbf{1}_{\mathrm{tr}(M)=0} + \mathbf{1}_{M=0}.$$

Let $C(V) := \{Z \in \mathrm{M}_n(\overline{\mathbb{F}}_q) : VZ - ZV = 0\}$. Let n_V be the number of distinct eigenvalues of V in $\overline{\mathbb{F}}_q$.

Lemma (Orbital integral bound)

Let $V, M \in \mathrm{M}_n(\mathbb{F}_q)$, not necessarily with trace zero. Then

$$L(V, M) := \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)}(\mathbf{1}_{\mathrm{tr}(gC(V)g^{-1}M)=0}) \ll_n q^{1-n_V}\mathbf{1}_{\mathrm{tr}(M)=0} + \mathbf{1}_{M=0}.$$

Proof sketch.

Interesting case: $\mathrm{tr}(M) = 0$ and $n_V \geq 2$. WLOG $V = \left[\begin{smallmatrix} V_1 & 0 \\ 0 & V_2 \end{smallmatrix}\right]$, where $V_1 \in \mathrm{M}_k(\mathbb{F}_q)$ and $V_2 \in \mathrm{M}_{n-k}(\mathbb{F}_q)$, with $n_{V_1}, n_{V_2} \geq 1$ and $n_V = n_{V_1} + n_{V_2}$. Then $C(V) = \begin{bmatrix} C(V_1) & 0 \\ 0 & C(V_2) \end{bmatrix}$. Since $\mathrm{tr}(\left[\begin{smallmatrix} A & 0 \\ 0 & B \end{smallmatrix}\right] M) = \mathrm{tr}(Ap_1(M)) + \mathrm{tr}(Bp_2(M))$, it follows that

$$L(V, M) = \mathbb{E}_{g \in \mathrm{GL}_n(\mathbb{F}_q)} L(V_1, p_1(g^{-1}Mg)) L(V_2, p_2(g^{-1}Mg)),$$

by folding the RHS $\mathrm{GL}_k \times \mathrm{GL}_{n-k}$ into GL_n. Induct on n. Average over a large abelian unipotent group of g's. □

Let $V = \begin{bmatrix} V_1 & 0 \\ 0 & V_2 \end{bmatrix}$, where $V_1 \in \mathrm{M}_k(\overline{\mathbb{F}}_q)$ and $V_2 \in \mathrm{M}_{n-k}(\overline{\mathbb{F}}_q)$. After Fourier-manipulating the equation $UV - VU = M$, we used the following fact to induct on n:

Lemma (Well known)

If V_1 and V_2 share no eigenvalues, then $C(V) = \begin{bmatrix} C(V_1) & 0 \\ 0 & C(V_2) \end{bmatrix}$.

The identity $\begin{bmatrix} A & B \\ C & D \end{bmatrix} V - V \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} AV_1 - V_1 A & BV_2 - V_1 B \\ CV_1 - V_2 C & DV_2 - V_2 D \end{bmatrix}$ and the following lemma give an alternative approach to induction, directly at the level of the equation $UV - VU = M$:

Lemma (C. Stephanos 1900)

The eigenvalues of the linear maps $B \mapsto V_1 B - BV_2$ *and* $C \mapsto CV_1 - V_2 C$ *are* $\lambda_1 - \lambda_2$, *where* λ_i *are eigenvalues of* V_i.

Let $V = \left[\begin{smallmatrix} V_1 & 0 \\ 0 & V_2 \end{smallmatrix}\right]$, where $V_1 \in \mathrm{M}_k(\overline{\mathbb{F}}_q)$ and $V_2 \in \mathrm{M}_{n-k}(\overline{\mathbb{F}}_q)$. After Fourier-manipulating the equation $UV - VU = M$, we used the following fact to induct on n:

Lemma (Well known)

If V_1 and V_2 share no eigenvalues, then $C(V) = \begin{bmatrix} C(V_1) & 0 \\ 0 & C(V_2) \end{bmatrix}$.

The identity $\left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] V - V \left[\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right] = \left[\begin{smallmatrix} AV_1 - V_1A & BV_2 - V_1B \\ CV_1 - V_2C & DV_2 - V_2D \end{smallmatrix}\right]$ and the following lemma give an alternative approach to induction, directly at the level of the equation $UV - VU = M$:

Lemma (C. Stephanos 1900)

The eigenvalues of the linear maps $B \mapsto V_1B - BV_2$ and $C \mapsto CV_1 - V_2C$ are $\lambda_1 - \lambda_2$, where λ_i are eigenvalues of V_i.

In fact, [Neubauer 1989] used the latter lemma to show that the variety $\{(X, Y) \in \mathrm{M}_n(\mathbb{C})^2 : \mathrm{rank}(XY - YX) \leq 1\}$ consists of $n - 1$ irreducible components of dimension $n^2 + 2n - 1$.

Sizes of commutator fibers

Theorem (Browning–Sawin–W. 2024)

$\dim\{(X, Y) \in \mathrm{M}_n(\mathbb{C})^2 : XY - YX = M\} \leq n^2 + 1$ *if* $M \neq 0$.

Let $N_M(T) := \#\{(X, Y) \in \mathrm{M}_n([-T, T])^2 : XY - YX = M\}$.

Corollary (Browning–Sawin–W. 2024)

If $T \geq 1$, *then* $T^{n^2+1} \leq N_0(T) \ll_n T^{n^2+2-2/(n+1)}$.

[Chapman–Mudgal 2025] give sharp bounds on $N_0(T)$ for $n \leq 3$.

Sizes of commutator fibers

Theorem (Browning–Sawin–W. 2024)

$\dim\{(X,Y) \in \mathrm{M}_n(\mathbb{C})^2 : XY - YX = M\} \leq n^2 + 1$ if $M \neq 0$.

Let $N_M(T) := \#\{(X,Y) \in \mathrm{M}_n([-T,T])^2 : XY - YX = M\}$.

Corollary (Browning–Sawin–W. 2024)

If $T \geq 1$, then $T^{n^2+1} \leq N_0(T) \ll_n T^{n^2+2-2/(n+1)}$.

[Chapman–Mudgal 2025] give sharp bounds on $N_0(T)$ for $n \leq 3$. Another corollary of our theorem is the following:

Corollary (Browning–Sawin–W. 2024)

$N_M(T) \ll_n T^{n^2+1}$ for $T \geq 1$, uniformly over $M \neq 0$.

(Proof: Embed $[-T,T] \subseteq \mathbb{F}_p$. Improvement to $T^{n^2+\epsilon}$ may be possible by affine dimension growth [Vermeulen 2023].)

It remains to explain this:

Theorem (Browning–Sawin–W. 2024)

$N_0(T) := \#\{(X, Y) \in \mathrm{M}_n([-T, T])^2 : XY = YX\} \ll_n T^{n^2+2}$.

Proof sketch.

Assume $T \ggg 1$. Let $p \sim T^{(n^2+n-2)/(n^2-1)} = T^{(n+2)/(n+1)}$. Relax $XY = YX$ to $p \mid XY - YX$. By Poisson summation,

$$N_0(T) \ll_n (T/p)^{2n^2} \sum_{(A,B)\in \mathrm{M}_n(\mathbb{Z})^2:\, |A|,|B|\le p/T} |S(A, B; p)|,$$

where $S(A, B; p) := \sum_{(U,V)\in \mathrm{M}_n(\mathbb{F}_p)^2:\, UV-VU=0} e_p(\mathrm{tr}(AU + BV))$. But $|S(A, B; p)| \le S(A, 0; p) \ll_n p^{n^2+1} + p^{n^2+n}\mathbf{1}_{p|A}$ by Fourier orthogonality. This "PNT" gives $N_0(T) \ll_n T^{n^2+n/2+O(1)}$. Remove $n/2$ by [Fouvry–Katz 2001] algebraic "large sieve" stratification. □

Upper bounds via harmonic analysis to an auxiliary modulus p go back at least to [Fujiwara 1984]. Also, [Heath-Brown 1994] gave stronger results in many cases via a composite modulus pq. The simplest version of Fujiwara's method, plus [Fouvry–Katz 2001], gives the following:

Theorem (General axiomatization)

Let $V \subset \mathbb{A}^N_{\mathbb{Z}}$ be a subscheme with $\dim(V \otimes \mathbb{C}) = D$. Assume that for all primes p and for all $0 \neq c \in \mathbb{F}_p^N$, we have

$$\sum_{x \in V(\mathbb{F}_p)} e_p(c_1 x_1 + \cdots + c_N x_N) \ll_V p^{D-L},$$

where $2L \in \mathbb{Z}$. Then $\#V([-T,T]) \ll_V T^{D-L+\frac{L^2}{N-D+L}}$.

Upper bounds via harmonic analysis to an auxiliary modulus p go back at least to [Fujiwara 1984]. Also, [Heath-Brown 1994] gave stronger results in many cases via a composite modulus pq. The simplest version of Fujiwara's method, plus [Fouvry–Katz 2001], gives the following:

Theorem (General axiomatization)

Let $V \subset \mathbb{A}^N_{\mathbb{Z}}$ be a subscheme with $\dim(V \otimes \mathbb{C}) = D$. Assume that for all primes p and for all $0 \neq c \in \mathbb{F}_p^N$, we have

$$\sum_{x \in V(\mathbb{F}_p)} e_p(c_1 x_1 + \cdots + c_N x_N) \ll_V p^{D-L},$$

where $2L \in \mathbb{Z}$. Then $\#V([-T,T]) \ll_V T^{D-L+\frac{L^2}{N-D+L}}$.

If say $L \geq \frac{3}{2}$ and $N - D \geq 4$, then this would improve on the *dimension growth bound* $O_{\epsilon,V}(T^{D-1+\epsilon})$, when that applies. For $XY = YX$, take $N = 2n^2$, $D = n^2 + n$, and $L = n - 1$.