Noncommutative analogues of a cancellation theorem of Abhyankar, Eakin, and Heinzer

Jason Bell (Joint with Hongdi Huang, Maryam Hamidizadeh, and Helbert Venegas)

May 30, 2021

What is Zariski Cancellation?

A longstanding problem in affine algebraic geometry is the Zariski cancellation problem, which asks whether an affine variety X over an algebraically closed field k having the property that $X \times \mathbb{A}^1 \cong \mathbb{A}^{n+1}$ is necessarily isomorphic to $\mathbb{A}^n$.

At the level of commutative k-algebras, this is just asking:

$$R[x] \cong k[x_1, \ldots, x_n] \implies R \cong k[x_1, \ldots, x_{n-1}]?$$

Some history

- It was solved for $n = 1$ by Abhyankar, Eakin, and Heinzer (1972), who proved cancellation for all affine irreducible curves.
- In $n = 2$, the characteristic zero case being done by Fujita (1979) and Miyanishi-Sugie (1980).
- Russell (1981) did the $n = 2$ case in positive characteristic.
- Gupta (2014) gave counterexamples to the Zariski cancellation problem in dimension at least three in positive characteristic.
- For characteristic zero, the problem is still open for $n \geq 3$.

Of course, we can ask about in general whether an affine varieties X is *cancellative* in the sense that if

$$X \times \mathbb{A}^1 \cong Y \times \mathbb{A}^1$$

for some affine variety Y, then we must have $X \cong Y$.

Here there is some connection with the idea that stably free projective modules need not be free. Danielewski, Hochster, Jelonek have given examples of affine varieties that are not cancellative, and Danielewski's examples are surfaces, so the result of AEH does not hold in higher dimensions.

In fact, Abhyankar, Eakin, and Heinzer prove a very strong version of cancellativity for affine irreducible curves: if

$$X \times \mathbb{A}^n \cong Y \times \mathbb{A}^n$$

then

$$X \cong Y.$$

Working with James

In 2013 at Bedlewo (Poland), James and I started discussing the idea of a noncommutative analogue of the Zariski Cancellation Problem (NCZCP).

In this context, one has a field k, an associative k-algebra A and one asks the question whether $A[x] \cong B[x]$ implies $A \cong B$. Of course, we know it is not true in general, but we can ask about whether it holds for certain classes of algebras (e.g., AS regular in characteristic zero).

In the summer of 2019, Hongdi Huang, Maryam Hamidizadeh, Helbert Venegas, and I started working on one of the basic open questions (Lezama, Wang, and Zhang, also Tang, Venegas, Zhang) about the NCZCP: does the analogue of the Abhyankar, Eakin, Heinzer problem hold in the noncommutative setting?

Of course, we need to do a bit of a translation. Recall AEH says that if R is a commutative affine domain of Krull dimension one over a field k, then $R[x] \cong S[x]$ implies $R \cong S$. Since Krull dimension = Gelfand-Kirillov dimension for affine commutative domains, we can replace "Krull dimension one" with "Gelfand-Kirillov dimension one" and ask whether the same result holds in the noncommutative setting.

Usually, extending about commutative rings to noncommutative rings is a bit hopeless, but here it is reasonable to expect that this should hold.

Theorem

(Small-Warfield) Let k be a field and let R be a finitely generated prime k-algebra of Gelfand-Kirillov dimension one. Then R is a finite module over its centre.

With Stafford, the prime hypothesis was later removed with the conclusion changing to R satisfying a polynomial identity. Van den Bergh noted that if k is algebraically closed and R is a domain of GK dimension one then S-W implies that R is commutative. Why? Look at $Q(R)$: it is a division ring that is finite-dimensional over $Q(Z(R))$, and since $Q(Z(R))$ has transcendence degree one over k, Tsen's theorem gives that $Q(R)$ is a field.

So it felt like it shouldn't be too much of an extension to take the AEH theorem and extend it to noncommutative domains. But the ultimate resolution was a bit surprising to me.

Theorem

(B-Huang-Hamidizadeh-Venegas) We have the following results for affine domains of Gelfand-Kirillov dimension one.

1. *Let k be a field of characteristic zero and let A be an affine domain over k of Gelfand-Kirillov dimension one. Then A is cancellative.*
2. *Let p be prime. Then there exists a field k of characteristic p and an affine domain A of Gelfand-Kirillov dimension one that is not cancellative.*

Lezama, Wang, and Zhang proved that for algebraically closed base fields k, affine prime k-algebras of Gelfand-Kirillov dimension one are cancellative. The algebraically closed property is needed, because the authors invoke Tsen's theorem at one point in their proof.

- In light of Van den Bergh's result, we see that our result is somewhat orthogonal to the LWZ result, since the intersection is just AEH.
- Our example in positive characteristic shows that the LWZ application of Tsen's theorem is in some sense necessary to get their result in positive characteristic.

The Makar-Limanov invariant

Back in the day, Koras-Russell were trying to classify all contractible threefolds admitting a hyperbolic $\mathbb{C}^*$-action. They ended up completely solving the problem, modulo one class of examples they couldn't distinguish.

Is

$$\mathbb{C}[x, y, z, w]/(x + x^2y + z^2 + w^3) \cong \mathbb{C}[u, v, t]?$$

Makar-Limanov showed how to show this isomorphism doesn't hold with an invariant he called AK, which is now denoted ML for Makar-Limanov.

Given a k-algebra R we let

$$\mathrm{ML}(R) = \bigcap \ker(\delta),$$

where δ runs over all k-linear locally nilpotent derivations (LNDs) of R. (Recall that a derivation δ is *locally nilpotent* if for every $a \in R$ there is some $N = N(a) > 0$ such that $\delta^n(a) = 0$ for all $n > N$.)

Notice that $\mathrm{ML}(k[x_1, \dots, x_n] = k$, since d/dx_i is a LND and the kernel of d/dx_i is $k[x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n]$.

Exercise: the Makar-Limanov invariant of

$$\mathbb{C}[x, y, z, w]/(x + x^2y + z^2 + w^3)$$

is the image of $\mathbb{C}[x]$.

Corollary:

$$\mathbb{C}[x, y, z, w]/(x + x^2y + z^2 + w^3) \ncong \mathbb{C}[u, v, t].$$

Derivations lose some of their power in positive characteristic. To get around this, one can work with *Hasse-Schmidt derivations*. A H-S derivation of A is a sequence of k-linear maps $\partial := \{\partial_i\}_{i \geq 0}$ such that:

$$\partial_0 = \mathrm{id}_A, \text{ and } \partial_n(ab) = \sum_{i=0}^{n} \partial_i(a)\partial_{n-i}(b)$$

for $a, b \in A$ and $n \geq 0$.

A Hasse-Schmidt derivation $\partial = (\partial_n)$ is called *locally nilpotent* if for each $a \in A$ there exists an integer $N = N(a) \geq 0$ such that $\partial_n(a) = 0$ for all $n \geq N$ and the k-algebra homomorphism $A[t] \to A[t]$ given by $t \mapsto t$ and $a \mapsto \sum_{i \geq 0} \partial_i(a) t^i$ is a k-algebra isomorphism. If only the first condition holds then the map $A[t] \to A[t]$ is still an injective endomorphism but need not be onto; we will call Hasse-Schmidt derivations for which only the first condition holds (i.e., there exists an integer $N = N(a) \geq 0$ such that $\partial_n(a) = 0$ for all $n \geq N$) a *weakly locally nilpotent Hasse-Schmidt derivation*.

The collection of locally nilpotent Hasse-Schmidt derivations of A (resp. weakly locally nilpotent Hasse-Schmidt derivations) of A is denoted $\mathrm{LND}^H(A)$ (resp. $\mathrm{LND}^{H'}(A)$).

Then we can now define general Makar-Limanov invariants. The *Makar-Limanov* invariant* of A is defined to be

$$\mathrm{ML}^*(A) = \bigcap_{\delta \in \mathrm{LND}^*(A)} \ker(\delta),$$

where $*$ can be blank (the classical ML invariant, H or H').

Why is ML useful for NCZCP?

We say that a domain is *ML-rigid* if $\mathrm{ML}(A) = A$.

Theorem

(ML-B-Zhang) If A has finite GK dimension and is ML-rigid, then:

- $\mathrm{ML}(A[t]) = A$;
- *A is cancellative.*

Sketch of proof of second part

Assume $\mathrm{ML}(A[t]) = A$.

If $A[t] \cong B[t]$, then $A = \mathrm{ML}(A[t]) \cong \mathrm{ML}(B[t])$.

Notice that $\mathrm{ML}(B[t]) \subseteq B$ since d/dt is a locally nilpotent derivation of $B[t]$ with kernel of B. Thus we get an embedding $A \subseteq B$. Now A and B have the same GK dimension so intuitively one "expects" derivations that vanish on A to vanish on B, although this is a bit of work to prove in the noncommutative setting.

Characteristic zero results

In characteristic zero, the main tool we use is a lemma.

Lemma

(Noncommutative slice lemma) Let k be a field and let A be a k-algebra. Then the following statements hold.

1. *Suppose that the characteristic of k is zero and $\delta \in \mathrm{LND}(A)$. If there exists $x \in Z(A)$ such that $\delta(x) = 1$, and if A_0 is the kernel of δ, then the sum $\sum_{i \geq 0} A_0 x^i$ is direct and $A = A_0[x]$.*
2. *Suppose that $\partial := \{\partial_n\}_{n \geq 0} \in \mathrm{LND}^H(A)$ that is iterative. If there exists $x \in Z(A)$ such that $\partial_1(x) = 1$ and $\partial_i(x) = 0$ for $i \geq 2$, and if A_0 is the kernel of ∂, then the sum $\sum_{i \geq 0} A_0 x^i$ is direct and $A = A_0[x]$.*

Proof

Let's just look at the first part and let's work in characteristic zero.

$$A_0 = \{a \in A \mid \delta(a) = 0\}.$$

Goal: $A = A_0[x]$. Notice $\sum A_0 x^i$ is direct. Why?

Thus A_0 and x generate a polynomial ring and $A \supseteq A_0[x]$.

We next claim that $A \subseteq A_0[x]$.

- To see this, suppose that there exists some $a \in A \setminus A_0[x]$.
- Then there is some largest $m \geq 1$ such that $\delta^m(a) \neq 0$.
- Among all $a \in A \setminus A_0[x]$, we choose one with this m minimal.
- Since $\delta^{m+1}(a) = 0$, $\delta^m(a) \in \ker(\delta) = A_0$.
- Let $c = \delta^m(a) \in A_0$ and consider $a' = a - cx^m/m!$.
- Observe that $\delta^m(a') = c - c = 0$.
- Thus by minimality of m, $a' \in A_0[x]$ and hence so is a, a contradiction.

Using slice

There are really two cases: when $\mathsf{ML}(A) = A$ (rigidity) and when $\mathsf{ML}(A) \neq A$.

If $\mathsf{ML}(A) = A$ then $\mathsf{ML}(A[x]) = \mathsf{ML}(A)$ and A is cancellative.

If $\mathsf{ML}(A) \neq A$, then there is a nonzero locally nilpotent derivation δ of A.

Let A_0 denote the kernel of δ. Then $\mathrm{GKdim}(A_0) \leq \mathrm{GKdim}(A) - 1 = 0$, so A_0 is locally finite-dimensional.

So now pick $z \in Z(A)$ that is not in A_0. We may assume without loss of generality that $\delta(z) = c \in Z(A_0)$ and $Z(A_0)$ is an algebraic field extension of k. Then look at $x = zc^{-1}$. Then $\delta(x) = 1$.

Now use slice! So we know by the HHV slice theorem that $A \cong A_0[x]$. Now if $A[t] \cong B[t]$, then we can run the same argument on B and we get $B \cong B_0[x]$ with B_0 GK dimension 0.

We also have $A[t] \cong B[t]$ and so

$$A_0[x, t] \cong B_0[x, t].$$

This implies

$$A_0 \cong B_0$$

and so

$$A \cong B.$$

Characteristic $p > 0$

I thought proving the result in positive characteristic should be a matter of using Hasse-Schmidt derivations and the HS analogue of noncommutative slice. I was wrong....

What is Hénonification?

General principle: Y = some class of examples; X = some nice subclass of Y.

Then understanding Y in dimension d is often like understanding X in dimension $d+1$ for suitable X and Y, only a bit easier.

Examples:

- endomorphisms of d-dimensional varieties vs. automorphisms of $d+1$-dimensional varieties;
- wild automorphisms of d-dimensional varieties vs. tame automorphisms in higher dimensions;
- algebras in GK dimension d vs. graded algebras in GK dimension $d+1$.

When one tries to extend the cancellation results in positive characteristic, one encounters one class of algebras where the proof fails.

Namely, k a field of characteristic $p > 0$, A an affine k-algebra, $Z(A) = k[t]$, $\mathrm{ML}(A) = k$. To construct a counterexample, we combine two ideas: the idea of "Hénonification" along with an idea of Small and Resco.

Let p be a prime, and let $K = \mathbb{F}_p(x_1, \ldots, x_{p^2-1})$.

We let $k = \mathbb{F}_p(x_1^p, \ldots, x_{p^2-1}^p)$ and we let δ be the k-linear derivation of K given by $\delta(x_i) = x_{i+1}$ for $i = 1, \ldots, p^2 - 1$, where we take $x_{p^2} = x_1$.

Since k has characteristic $p > 0$, we have δ^{p^i} is a k-linear derivation for every $i \geq 0$, and since $\delta^{p^2}(x_i) = \delta(x_i) = x_{i+1}$ for $i = 1, \ldots, p^2 - 1$, $\delta^{p^{j+2}} = \delta^{p^j}$ for every $j \geq 0$. We let $\delta' := \delta^p$, which as we have just remarked is a k-linear derivation of K.

Let $A = K[x; \delta]$ and we let $B = K[x'; \delta']$. Then

- $A \ncong B$.
- $A[t] \cong B[t]$.
- A and B are k-algebras of GK dimension one.

Since $\mathrm{ad}_u^p = \mathrm{ad}_{u^p}$ for u in a ring of characteristic p, we have $z := x^{p^2} - x$ and $z' := (x')^{p^2} - x'$ are central by the above remarks. We claim that A and B have Gelfand-Kirillov dimension one, $A \ncong B$, and $A[t] \cong B[t']$.

Recall that $A = K[x;\delta]$ and we let $B = K[x';\delta']$. Since $\mathrm{ad}_u^p = \mathrm{ad}_{u^p}$ for u in a ring of characteristic p, we have $z := x^{p^2} - x$ and $z' := (x')^{p^2} - x'$ are central by the above remarks. We claim that A and B have Gelfand-Kirillov dimension one, $A \ncong B$, and $A[t] \cong B[t']$.

Isomorphism: We define $\Phi : A[t] \to B[t]$ by $\Phi(\alpha) = \alpha$ for $\alpha \in K$, $\Phi(x) = (x')^p + t'$ and $\Phi(t) = (x')^{p^2} - x' + (t')^p$. (Why is this Hénonification?)

One can check that this is one-to-one and onto and is a homomorphism.

Why do we have $A \ncong B$ as k-algebras?

Suppose that $\Psi : A \to B$ is a k-algebra isomorphism. Then since the units group of A and B are both K^*, Ψ induces a k-algebra automorphism of K; furthermore, every $\alpha \in K$ satisfies $\alpha^p \in k$ and for $\beta \in k$ there is a unique $\alpha \in K$ such that $\alpha^p = \beta$.

Since Ψ is the identity on k, Ψ is the identity on K. Thus $\Psi(x) = p(x')$ for some $p(x') \in K[x'; \delta'] \setminus K$.

Let $d \geq 1$ denote the degree of $p(x')$ as a polynomial in x'. If $d > 1$, Ψ cannot be onto.

So $\Psi(x) = \alpha x' + \beta$ with $\alpha \in K^*$ and $\beta \in K$.

Since Ψ is an isomorphism, for $\zeta \in K$ we have

$$\delta(\zeta) = \Psi(\delta(\zeta)) = \Psi([x, \zeta]) = [\Psi(x), \Psi(\zeta)] = [\alpha x' + \beta, \zeta].$$

The right side is just

$$\alpha[x', \zeta] = \alpha \delta'(\zeta).$$

Thus there is some fixed $\alpha \in K^*$ such that

$$\delta(\zeta) = \alpha \delta'(\zeta) \ \forall \zeta \in K$$

.

But by construction $\delta(x_1) = x_2$ and $\delta'(x_1) = x_{p+1}$ and so $\alpha = x_2/x_{p+1}$. We also have $\delta(x_2) = x_3$ and $\delta'(x_2) = x_{p+2}$, and so $\alpha = x_3/x_{p+2}$, which gives $x_2 x_{p+2} = x_3 x_{p+1}$, where we take $x_{p+2} = x_1$ when $p = 2$. This is a contradiction. Thus $A \ncong B$.

What did we learn?

In analogy with terminology from algebraic geometry, given an algebraically closed field k and a finitely generated extension F of k, we will say that F is *uniruled* over k if there is a finitely generated field extension E of k with

$$\mathrm{trdeg}_k(E) = \mathrm{trdeg}_k(F) - 1$$

and an injective k-algebra homomorphism

$$F \to E(t).$$

The idea here is that F is the function field of a normal projective scheme X of finite type over k. Then the condition $F \subseteq E(t)$ says that there is a dominant rational map

$$Y \times \mathbb{P}^1 \dashrightarrow X$$

for some variety Y with $\dim(Y) = \dim(X) - 1$.

Similarly, if F is the function field of a normal projective scheme X of finite type over k, we define the *Kodaira dimension* of F to be the Kodaira dimension of X.

Since Kodaira dimension is a birational invariant, this is well-defined. If k has characteristic zero, a uniruled variety has Kodaira dimension $-\infty$ and the converse holds in dimensions one, two, and three; the main conjectures of the minimal model program imply that the converse should hold in higher dimensions, too.

Over uncountable fields, affine uniruled varieties have a pleasant characterization in terms of being covered by affine lines

Theorem

(Jelonek) Let k be an uncountable algebraically closed field and let X be an irreducible affine variety over k of dimension at least one. Then following conditions are equivalent:

1. *for every $x \in X$ there is a polynomial affine curve Y_x in X that passes through x;*
2. *there is a Zariski-dense open subset U of X, such that for every $x \in U$ there is a polynomial affine curve Y_x in X that passes through x;*
3. *X is uniruled; that is, there exists an affine variety Y with $\dim(Y) = \dim(X) - 1$ and a dominant morphism $Y \times \mathbb{A}^1 \to X$.*

Corollary

Let k be an uncountable algebraically closed field, let A be a finitely generated prime left Goldie k-algebra of finite Gelfand-Kirillov dimension, and suppose that $Z(A)$ is affine. If A is not cancellative then $\mathrm{Frac}(Z(A))$ is uniruled. In particular, if k has characteristic zero and $\mathrm{Frac}(A)$ has nonnegative Kodaira dimension then A is strongly cancellative.

Centre controls cancellation

What does this mean? Not being uniruled is saying that $Z(A)$ is in some sense "rigid". So the above result shows that if the centre of an algebra is sufficiently "rigid" and the base field is sufficiently "nice" then the algebra is cancellative.

Conjecture. Let k be an uncountable algebraically closed field of characteristic zero and let A be an affine noetherian domain over k. Suppose that $Z(A)$ is affine and cancellative. Then A is cancellative.

Skew cancellativity?

There is another natural way to extend the work of Abhyankar, Eakin, and Heinzer.
One can now take R and S affine commutative domains of Krull dimension one and consider an isomorphism of skew polynomial extensions $R[x;\sigma,\delta] \cong S[x;\sigma';\delta']$, and ask whether this implies that R and S are isomorphic.

When σ, σ' are the identity maps and δ, δ' are zero, the question reduces to the classical cancellation problem for affine curves, answered by Abhyankar, Eakin, and Heinzer.

Theorem

Let k be a field, let A and B be affine commutative integral domains of Krull dimension one, and let σ, σ' be k-algebra automorphisms of A and B respectively and let δ, δ' be k-linear derivations of A and B respectively. If $A[x;\sigma] \cong B[x';\sigma']$ then $A \cong B$. If, in addition, k has characteristic zero and if $A[x;\delta] \cong B[x';\delta']$ then $A \cong B$.

A special case of this Theorem was proved by Bergen in the derivation case. Specifically, he proved that if k is a field of characteristic zero and $R[t;\delta]$ is isomorphic to $k[x][y;\delta']$, with $\delta'(x) \in k[x]$ having degree at least one, then $R \cong k[x]$. It would be interesting to give a "unification" of the two results occurring in our and prove that skew cancellation holds for general skew polynomial extensions, although this appears to be considerably more subtle than the cases we consider.

Thanks!