Algebras and formal languages

Dmitri Piontkovski

National Research University Higher School of Economics,
Moscow, Russia

Lens, 2019

Graded algebras

We call an associative (or non-associative, or multi-operator) algebra A graded if

$$A = A_0 \oplus A_1 \oplus A_2 \oplus \dots,$$

where $A_0 = k$ is a basic field (or $A_0 = 0$, in non-associtive case), $\dim_k A_i < \infty$. All our algebras are graded.
Hilbert function: $h_A(n) = \dim A_n$
Hilbert series: $H_A(z) = \sum_{n \geq 0} z^n \dim A_n = \sum_{n \geq 0} z^n h_A(n)$.
Let $X = \{x_1, \dots, x_n\}$ with $\deg x_i = 1$, and let $F = k\langle X \rangle$ be the free associative algebra. Then the algebra $A = F/I$ (where I is a two-sided homogeneous ideal) is *standard graded*. This means that A is generated by A_1.

Rationality

An algebra A is called *finitely presented* if it is defined by a finite number of generators and relations.

Theorem (Govorov, 1972)

If the relations of a finitely presented graded algebra A are monomials in generators then $H_A(z)$ is a rational function.

Corollary. If the ideal of relations of A has finite (noncommutative) Groebner basis, then $H_A(z)$ is a rational function.

Irrationality

Conjecture (Govorov)

For each finitely presented algebra A the Hilbert series $H_A(z)$ is a rational function.

First counterexamples: Ufnarovski, 1978 (transcendental) and Shearer, 1980 (algebraic).

Open questions: Govorov conjecture for Noetherian algebras and for Koszul algebras.

Automaton algebras

Let X be a finite generating set of an algebra A. Consider a multiplicative ordering '$<$' of the set of all words in X. A word on X is called normal in A if it is not a linear combination of less words. The set N of all normal words is a linear basis of A.

Definition (Ufnarovski)

An algebra A is called automaton if N is a regular language.

Recall that a language is regular iff it is recognized by a finite automaton.

Theorem (Kleene) A language L is regular if and only if it can be obtained from finite languages by applying a finite number of regular operations, that is, Kleene star, union, concatenation, intersection, and complement.

Suppose $A = F/I$. Let G be a minimal (noncommutative) Groebner basis of I, so that $N = X^* \setminus X^*(\operatorname{lm} G)X^*$. We have

$$A \text{ is automaton} \iff \operatorname{lm} G \text{ is regular}.$$

In particular, finitely presented monomial algebras are automaton.

Automaton algebras

Theorem (Ufnarovski). *If A is graded and automaton, then $H_A(z)$ is a rational function.*

Example. Let $A = \langle x, y | x^2 - xy \rangle$. For the deglex ordering with $x > y$, we the Groebner basis of the relations is $g = \{xy^n x - xy^{n+1} | n \geq 0\}$. Then
$N = \{1, y^a x y^b | a, b \geq 0\} = \{1\} \cup y^* x y^*$ is regular,
$H_A(z) = (1-z)^{-2}$.

Problem 1. How to generalize this theorem to algebras with irrational Hilbert series?

Problem 2. How to generalize Govorov theorem to non-associative and non-binary algebras?

Formal language theory: Chomski's hierarchy

Cf. [Naom Chomski, 1956].

Grammar	Languages	Automaton	Example
Type-0	Recuresively enumerable	Turing machine	{ All terminating computer programs
Type-1	Context-sensitive	Linear bounded	$\{x^n y^n z^n \| n \geq 0\}$
Type-2	Context-free	Pushdown	$\{x^n y^n \| n \geq\}$
Type-3	Regular	Finite	$\{ac^n b \| n \geq 0\}$

A part of the hierarchy for cf languages

Languages	Automaton	Generating functions
cf	Pushdown	(arbitrary)
Unambiguous cf		Algebraic
Deterministic cf	Deterministic pushdown	
Regular	Finite	Rational
Slender regular	Of special kind	$p(x)/(1-x^N)$

Rationality in the linear growth case

An algebra has linear growth, if GK–dim $A \le 1$, that is, for some $c > 0$ we have $h_A(n) = \dim A_n < c$.

Example

Let $A = \langle x, y | x^2, yxy, xy^{2^t}x \text{ for all } t \ge 0 \rangle$. Then
$A_n = k\{y^n, xy^{n-1}, y^{n-1}x, xy^{n-2}x\}$ for $n \ne 2^t + 2$ or
$A_n = k\{y^n, xy^{n-1}, y^{n-1}x\}$ otherwise.
We have $H_A(z) = 1 + 2z + 4z^2/(1-z) - z^2 \sum_{t \ge 0} z^{2^t}$.

Problem (Govorov conjecture for algebras of linear growth, GALG)

Suppose that an algebra A of linear growth is finitely presented. Is $H_A(z)$ a rational function?

For such algebras, $H_A(z)$ is rational iff $h_A(n)$ is eventually periodic, that is, $\exists n_0, T > 0$ such that $h_A(n) = h_A(n+T)$ for all $n > n_0$.

Conjecture (Ufnarovski conjecture for graded algebras, UGA)

A graded finitely presented algebra of linear growth is graded automaton.

UGA implies GALG.

The finite characteristic case

Theorem

Suppose that the field k has a finite characteristic. Then both Govorov conjecture for algebra of linear growth and Ufnarovski conjecture for graded algebras hold if and only if k is an algebraic extension of its prime subfield.

'If' part: essentially, the case of finite field.
'Only if' part (counterexamples to GALG): based on the connections with the dynamical Mordell–Lang conjecture and the set of zeroes of linear recurrent sequences.

The case of infinite field

What about the case char $k = 0$?

Example (Fermat algebras)

For $\alpha, \beta \in k^\times$, let $A = A_{\alpha,\beta}$ be generated by a, b, c, x, y, z subject to 26 relations $xc - \alpha cx, yb - \beta cy$ and others. Then $h_A(n+3)$ is 10 or 11 according to whether the Fermat equality $\alpha^n + \beta^n = 1$ holds. So, it has no nonzero solution in $k^\times$ for each $n \geq 3$ if and only if $h_A(i) = 10$ for all $i \geq 6$ and each $A = A_{\alpha,\beta}$.

Theorem

Let $g \geq 5$ be an integer. If the field k is infinite, then there are infinitely many (periodic) sequences h_A for g-generated quadratic k-algebras of linear growth. If If, in addition, k contains all primitive roots of unity, then both the length d of the initial non-periodic segment and the period T of h_A can be arbitrary large.

Algebras and languages of linear growth

Theorem [Justin, 1971; Belov, Borisenko, Latyshev, 1997; Holt, Owens, Thomas, 2008] *Each finitely generated semigroup of linear growth is a finite union of a finite set and sets of the form $a\langle c\rangle b$, where $\langle c\rangle$ is a monogenic semigroup. Equivalently, if a (non-graded) algebra A of linear growth is generated by a finite set S, then there are $U, V, W \subset S^*$ such that each normal word in A has the form*

$$w = ac^n b, \text{ where } a \in U, b \in V, c \in W, n \geq 0.$$

Languages of linear growth are called *slender*.
Theorem [Paun, Salomaa, 1995]. *Each regular slender language is a finite disjoint union of a finite set and sets of the from ac^*b (where $a, b, c \in X^*$).*

Normal words in f.p. algebras of linear growth

F.p. algebras and monoids of linear growth

Let A be an algebra of linear growth.

Corollary

Suppose that the algebra A is graded finitely presented and the basic field is finite. Then there are a generating set $1 \in S \subset A$ and an ordering such that for some $Q \subset S^3$ the set of normal words in A is

$$\{ac^n b | n \geq 0, (a, b, c) \in Q\}.$$

Corollary

Let S be a homogeneous finitely presented monoid. Then S has linear growth if and only if it is the finite disjoint union of a finite set and sets of the form $a\langle c\rangle b$, where $\langle c\rangle$ is a free monogenic semigroup.

Context free languages: main definitions

Recall that a context-free grammar G is quadruple of finite sets V (variables), X (terminals, or letters), $G \subset V \times (V \cup X)^*$ (rules of the form $A \to \alpha$) and an element $S \in V$ (a start variable).

Compact notation: $A \to \alpha_1 | \dots | \alpha_k$ in place of $A \to \alpha_1, \dots, A \to \alpha_k$.
A language $L \subset X^*$ is context-free if it there is G such that $L = \{w | S \xrightarrow{*} w\}$, that is, for each $w \in L$ there is a *derivation* $S \to a_1 \to a_2 \to \dots \to a_k = w$.
The cf grammar G and the language L are called

- unambiguous, if for each $w \in L$ the leftmost derivation $S \xrightarrow{*} w$ is unique;
- deterministic, if it is unambiguous and the source of each step $a_{i-1} \to a_i$ is uniquely defined by the initial segment of a_i;
- regular, if all rules are of the form $A \to 1$ or $A \to x_i B$ (where $A, B \in P$).

Generating series of languages

Let l_i be the number of the words of L having length i, and let $\gamma_L(z) = \sum_{i=0}^{\infty} l_i z^i$.

Theorem (Chomsky–Schützenberger)

Suppose that a cf grammar G as above is unambiguous. Then $\gamma_L(z)$ is an algebraic function. If, moreover, G is regular, then $\gamma_L(z)$ is a rational function.

In both cases, there are effective algorithms to produce a system of algebraic (or linear) equations which defines $\gamma_L(z)$. Then one can apply the standard elimination technique based on Groebner bases.

Homological approach

Homologically unambiguous algebras

(joint work with R. La Scala and S. Tiwary).

Question. Suppose that $A = F/I$ is a monomial algebra, where the ideal I is generated by an (unambiguous) cf subword-free language $L \subset X^*$. How to describe the language N and the Hilbert series $H_A(z) = \gamma(z)$?

Suppose that A has finite global dimension (say, d). Then there exist a *free resolution*

$$0 \to kL_{d-1} \otimes A \to \ldots \to kL_1 \otimes A \to kL_0 \otimes A \to A \to k \to 0.$$

The languages L_k are called *chain languages* (Anick, 1986). Here $L_0 = X, L_1 = L$, and the elements of L_k are (minimal) intersections of k elements of L:

Then $\mathrm{HS}(A) = \left(1 - nt - \sum_{i=1}^{d-1} (-1)^i \gamma(L_i)\right)^{-1}$

Are chains context free?

Proposition. If L is regular, then the chain language L_k is regular for each k.
Mansson (2002) has provided an algorithm to construct recursively the languages L_k by a regular L (in terms of finite automata).
Question. Suppose L is (unambiguous) cf-language. Does this imply that each chain language L_k is (unambiguous) cf?
Example. Let $X = \{x, y, z\}$ and
$L = \{x^n y^n z \mid n \geq 2\} \cup \{x y^n z^n \mid n \geq 2\}$. Then $L_1 = L$ is an unambiguous cf language generated by the grammar with

$$P = \{S \to Az \mid xB, A \to x^2y^2 \mid xAy, B \to y^2z^2 \mid yBz\}.$$

Still, $L_2 = \{x^n y^n z^n \mid n \geq 2\}$ is not context-free. Here gl. dim $A = 3$ and

$$H = (1 - nt + \gamma(L_1) - \gamma(L_2))^{-1}$$

is a rational function.

Unambiguous algebras

Definition

Let A be a monomial algebra with the relations $L \subset (X^+)^2$. We call A a *homologically unambiguous monomial algebra*, briefly an *unambiguous algebra*, if all chain languages $L_k(A)$ $(k \geq 1)$ are unambiguous cf-languages.

Proposition (algebraic). Let A be an unambiguous algebra having finite global dimension. Then the Hilbert series $\mathrm{HS}(A)$ is an algebraic function.

Proposition (algorithmic). Given unambiguous cf grammars for $L_1 = L, L_2, \ldots, L_{d-1}$, there is an algorithm to construct a system of algebraic equations defining $H_A(z)$.

Unambiguous monomial examples

Example 1. Fix $X = \{x, y, z, c\}$ and $Y = \{a, b\}$. We put $Z = X \cup Y$ and $F = k\langle Z\rangle$. Consider the Lukasiewicz cf-grammar $G = (V, Y, P, S)$ where $V = \{S\}$ and $P = \{S \to a \mid bSS\}$. The corresponding cf-language $L = L(G)$ consists of the algebraic expressions in Polish notation (e.g., $a, baa, babaa$). Put $A = F/(L)$, where

$$L = \{x^2y, x^2z, xy^2, xyz, xzy, xz^2\} \cup yz^2Lc.$$

Then gl. dim $A = 4$ with $L_2 = \{x^2y^2, x^2yz, x^2zy, x^2z^2\} \cup \{xyz^2, xy^2z^2, xzyz^2\}Lc$ and $L_3 = \{x^2y^2z^2, x^2zyz^2\}Lc$. Then $H_A(t) =$ $\left(1 - 6t + \frac{13}{2}t^3 - \frac{9}{2}t^4 - t^5 + t^6 - t^3(1-t)(1-2t^2)\frac{\sqrt{1-4t^2}}{2}\right)^{-1}$.

Finitely presented case: toy example

Toy example. Let $A = k\langle x, y | yxy - y^2x\rangle$. Under the lex-deg ordering with $x > y$, the Groebner basis is $G = \{y^n x^n y - y^{n+1} x^n | n \geq 1\}$. Then the associated monomial algebra $B = k\langle x, y | \text{lm}\,(G)\rangle$ is unambiguous with

$$L_1 = L = \{y^n x^n y | n \geq 1\}$$

and

$$L_k = y^{n_1} x^{n_1} \dots y^{n_k} x^{n_k} y.$$

Moreover, $H_A(z) = H_B(z) = (1 - 2z + z^3)^{-1}$ is rational.
Question (Mansson, Nordbeck, 2002). Are all algebras defined by a single homogeneous relation automaton?
Question. Are all algebras defined by a single homogeneous relation unambiguous?

Finitely presented case: examples

Example 4. Fix $X = \{a', b', x, y\}, Y = \{a, b, e\}$ and put $Z = X \cup Y, F = k\langle Z\rangle$. Let $I \subset F$ be generated by

(i) $a'x - xa', b'x - xe$;

(ii) $a'a - aa', a'b - ab', b'a - ba', b'b - bb', a'e - ab, b'e - b^2$;

(iii) $ay - y^2, by - y^2, a'y - y^2, b'y - y^2$;

(iv) xy.

Let G be the minimal Groebner basis of I for deg-lex with $a' \succ b' \succ a \succ b \succ e \succ x \succ y$, and let $L = \mathrm{lm}\,(G)$. Let $M = (De)^*$ where D is the Dick language on a, b. Note that M is unambiguous defined by the grammar $G = (V, Y, P, S)$, where $V = \{S, T\}$ and

$$P = \{S \to 1 \mid TeS, T \to 1 \mid aTbT\}.$$

Then L is the union of the leading terms of (i)–(iii) and the language xMy.

Then the associated monomial algebra $B = F/(L)$ is unambiguous with gl. dim $B = 3$ and

$$L_2(B) = \{a', b'\}\{a, b\}y \cup \{a', b'\}xMy.$$

Then the function $E = H_B(t)^{-1}$ satisfies a system

$$\begin{cases} E &= 1 - 7t + E_1 - E_2, \\ E_1 &= 12t^2 + t^2 S, \\ E_2 &= 4t^3 + 2t^3 S, \\ S &= tST + 1, \\ T &= t^2T^2 + 1. \end{cases}$$

We obtain

$$H_A(t) = H_B(t) = \left(1 - 7t + \frac{25}{2}t^2 - 5t^3 + t^2\frac{\sqrt{1-4t^2}}{2}\right)^{-1}.$$

Multioperator algebras

We fix a field k.
Multioperator algebra is a vector space with a set of multilinear operations on it.

Example

$A = k[x]$ (polynomials on x),
binary operations: $(f, g) \mapsto f \cdot g$, $\{f, g\} = fg' - gf'$,
$f * g(z) = \int_0^z f'(w)g(w)\,dw$,
unary operation: $f \mapsto f'$, etc.

Examples of *identities*:
$(f \cdot g) \cdot h \equiv f \cdot (g \cdot h)$ (associativity),
$\{f, g\} \equiv -\{g, f\}$ (anti-commutativity),
$(a * b) * c \equiv a * (b * c + c * b)$ (Zinbiel identity).

A variety of multioperator algebras is defined by a set of basic operations (*signature*) and a set of identities.

Operads and varieties

Let V be a variety of multioperator algebras.
A corresponding (symmetric) operad $\mathcal{P} = \mathcal{P}^V$ is the set of all composite multilinear operations on algebras in V.
We have $\mathcal{P} = \mathcal{P}_1 \cup \mathcal{P}_2 \cup \dots$,
where $\mathcal{P}_n \subset F^V(x_1, x_2, \dots)$ is the set of n-linear generic polynomials in $x_1, \dots, x_n$ inside the relatively free algebra $F^V = F^V(x_1, x_2, \dots)$.

Operations on $\mathcal{P}$:

- compositions: $\mathcal{P}_m \circ_t \mathcal{P}_n \to \mathcal{P}_{m+n-1}, t = 1, \dots, m$;
- action of the symmetric group S_n on $\mathcal{P}_n$ (for *symmetric* operads)

with obvious compatibility conditions.

One can recover $\mathcal{P}$ and V by each other:

$\mathcal{P} \rightsquigarrow V = V^{\mathcal{P}}$ and $V \rightsquigarrow \mathcal{P} = \mathcal{P}^V$.

Selection from the history of operads

Operads were introduced by in [May, 1970].

Second born in 1990s after works by Getsler, Jones, Kapranov, Ginzburg, Stasheff, Markl, and others, with applications in topology and mathematical physics.

Selected bibliography

1. J.-L. Loday, J. Stasheff, and A. Voronov, *Operads: proceedings of renaissance conferences*, Contemporary mathematics, **202** (1997)
2. M. Markl, S. Shnider, J. Stasheff, *Operads in algebra, topology and physics*, Mathematical Surveys and Monographs, **96**, AMS, Providence, RI, 2002
3. J.-L. Loday, B. Vallette, *Algebraic Operads,* Grundlehren Math. Wiss. 346, Springer, Heidelberg, 2012
4. M. Bremner and V. Dotsenko, *Algebraic operads: an algorithmic companion,* CRC Press, 2016

   See also:
5. A. Giambruno, M. Zaicev, *Polynomial identities and asymptotic methods*, Mathematical Surveys and Monographs, **122**, AMS, Providence, RI, 2005

A list of some common operads

Contents of Zinbiel's *Encyclopedia of types of algebras 2010*

Generating series of some operads

The operad As is a non-symmetric associativity operad. It is generated by $\mu : (x, y) \mapsto x \cdot y$ subject to $\mu(x_1, \mu(x_2, x_3)) \equiv \mu(\mu(x_1, x_2), x_3))$, or $- \cdot (- \cdot -) = (- \cdot -) \cdot -$. We have $\mathsf{As}(n) = k\{x_1 \dots x_n\}$, $G_{\mathsf{As}}(z) = \frac{z}{1-z}$.
Its symmetrization $\mathcal{A}ssoc$ is a symmetric operad generated by $\mu : (x, y) \mapsto x \cdot y$ and $\nu : (x, y) \mapsto y \cdot x$ subject to $\mu(x_1, \nu(x_2, x_3)) \equiv \mu(\mu(x_1, x_3), x_2))$ and others (6 linearly independent identities).
Then

$$\dim \mathcal{A}ssoc(n) = n!, E_{\mathcal{A}ssoc}(z) = \frac{z}{1-z} = G_{\mathsf{As}}(z).$$

For other common operads:

$$E_{\mathcal{C}om} = e^z - 1, E_{\mathcal{L}ie} = -\ln(1-z).$$

Operads with finite Gröbner bases: a question

Non-symmetric operads As of associative algebras, of q-associative algebras, Dend of dendriform algebras, and others have has finite Gröbner bases (see the book by Dotsenko and Bremner). What does this imply about their generating series?

Analogy. [Govorov, 1972] If A is a graded associative algebra with finite Gröbner basis, then its Hilbert series is a rational function, $H_A(z) = p(z)/q(z)$.
Addition. [Ufnarovsky, 1989] Because A is automaton.

Operads with finite Gröbner bases: answers

The elements of a (free) operad and a free algebra over an are spanned by the words in Polish notations (recall the Lukasiewicz langauge), e.g., $\mu(x_1, \mu(x_2, x_3)) \mapsto \mu x_1 \mu x_2 x_3$ and $\mu(x_1, \mu(x_2, x_3)) \mapsto \mu x_1 \mu x_2 x_3$. For non-symmetric operads and f.g. algebras, they are defined over finite alphabets.

Theorem (P.)

Let P be a non-symmetric operad with finite Groebner basis (e.g., an f.p. monomial operad) and let A be an algebras with finite Groebner basis over such an operad. Then both the set of normal words N_P and N_A of P and of A are detreministic cf langauges.

Corollary[Drensky and Holtkamp, 2008] Each finitely presented monomial algebra over a free finitely generated (non-symmetric) operads have algebraic Hilbert series.
Corollary[Khoroshkin and P., 2015] The ordinary generating series $G_P(z)$ of a non-symmetric operad with a finite Gröbner basis is an algebraic function.

Bibliography

- D. P., *Algebras of linear growth and the dynamical Mordell–Lang conjecture*, Advances in Math., 2019
- R. La Scala, D. P., S. Tiwari, *Noncommutative algebras, context-free grammars and algebraic Hilbert series*, to appear in J. of Symbolic Computation
- A. Khoroshkin, D. P., *On generating series of finitely presented operads*, J. of Algebra, 2015
- D. P., *In preparation*

Thank you!

Unambiguous monomial examples

Example 2. Fix $X = \{x, y, z, c, d\}$ and $Y = \{a, b\}$. We put $Z = X \cup Y$ and $F = k\langle Z \rangle$. Consider the Lukasiewicz cf-grammar $G = (V, Y, P, S)$ where $V = \{S\}$ and $P = \{S \to a \mid bSS\}$. The corresponding cf-language $L = L(G)$ consists of the algebraic expressions in Polish notation (e.g., $a, baa, babaa$). Put $A = F/(L)$, where

$$L = cL\{x^2y, xyz, xzx\} \cup \{xy^2, y^2z, z^2y\}Ld.$$

Then $L_2 = cL\{x^2y^2, x^2y^2z, xyz^2y, xzxy^2\}Ld$ and $L_3 = \emptyset$, so that gl. dim $A = 3$. Then $H_A(t)$ is the inverse of the root of

$$E^2 + (-6t^7 - 2t^6 + 3t^5 + t^4 - 6t^3 + 14t - 2)E + 9t^{14} + 6t^{13} + t^{12}.$$

This is confirmed by its correct power series expansion

$$H_A(t) = 1 + 7t + 49t^2 + 343t^3 + 2401t^4 + 16801t^5 + 117565t^6 + \ldots.$$

Monomial examples: infinite global dimension

Example 3. Let $X = \{x\}, Y = \{a, b\}, Z = X \cup Y$ and $F = k\langle Z \rangle$. Consider the Dyck language D on the alphabet Y. Let

$$\gamma = \gamma(D) = \frac{1 - \sqrt{1 - 4t^2}}{2t^2}.$$

Put $L = xDx \subset Z^*$. and $A = F/(L)$. For any $n \geq 1$, the (unambiguous) n-chain language of A is clearly

$$L_n = x(Dx)^n.$$

We conclude that gl. dim $(A) = \infty$ and $\gamma(L_n) = t^{n+1}\gamma^n$. Finally, $\mathrm{HS}(A)^{-1} = 1 - \sum_{i=0}^{\infty}(-1)^i \gamma(L_i)$

$$= 1 - 3t + t^2 \frac{\gamma}{(1 - t\gamma)} = \frac{1 - 6t + 6t^2 - (1 - 4t)\sqrt{1 - 4t^2}}{1 - 2t - \sqrt{1 - 4t^2}}.$$

More general classes

Theorem. Let $M \subset Y^+$ be an unambiguous context-free language and let $R_0 \subset X^*, R_1, R_1', \dots, R_k, R_k' \subset X^+$ be regular languages such that their disjoint union

$$R_0 \cup R_1 \cup R_1' \cup \cdots \cup R_k \cup R_k'$$

is subword free. Then the monomial algebra

$$A = \langle X \cup Y | R_0 \cup R_1 M R_1' \cup \cdots \cup R_k M R_k'$$

is homologically unambiguous.

A phrase-book

variety	—	operad
subvariety	—	quotient operad
signature	—	set of generators
identities	—	relations
free algebra	—	free algebra
(exponential) codimension series	—	(exponential) generating function
T-space	—	right ideal
T-ideal	—	ideal
Specht properties	—	Noether properties