Automata and (generalized) power series: beyond Christol

Kiran S. Kedlaya

Department of Mathematics, University of California, San Diego
kedlaya@ucsd.edu
http://kskedlaya.org/slides/

Bridges between Automatic Sequences, Algebra and Number Theory
Centre de Recherches Mathématiques, Montréal
May 4, 2017

Reference: K.S. Kedlaya, On the algebraicity of generalized power series, arXiv:1508.01836, to appear in *Beiträge zur Algebra und Geometrie*.

Kedlaya was supported by NSF grant DMS-1501214 and UCSD (Warschawski chair).

Contents

Christol's theorem

Throughout this talk, fix a prime number p and let k be a field of characteristic p.

Let $x = \sum_{n=0}^{\infty} x_n T^n \in k[\![T]\!]$ be a power series with coefficients in k. We say that x is *algebraic* if there exists a nonzero polynomial $P \in k[y,z]$ such that $P(T,x) = 0$. That is, as an element of the Laurent series field $k((T))$, x is integral over the subfield $k(T)$ of rational functions.

Suppose now that k is a *finite* field of characteristic p. We say that x is *automatic* if for each $c \in k$, the set $\{n \geq 0 : x_n = c\}$ is p-automatic, that is, the corresponding base-p expansions (read right-to-left) form a regular language on the alphabet $\Sigma_p := \{0, \dots, p-1\}$.

Theorem (Christol)

For k finite, $x \in k[\![T]\!]$ is algebraic if and only if it is automatic.

Christol's theorem

Throughout this talk, fix a prime number p and let k be a field of characteristic p.

Let $x = \sum_{n=0}^{\infty} x_n T^n \in k[\![T]\!]$ be a power series with coefficients in k. We say that x is *algebraic* if there exists a nonzero polynomial $P \in k[y, z]$ such that $P(T, x) = 0$. That is, as an element of the Laurent series field $k((T))$, x is integral over the subfield $k(T)$ of rational functions.

Suppose now that k is a *finite* field of characteristic p. We say that x is *automatic* if for each $c \in k$, the set $\{n \geq 0 : x_n = c\}$ is p-automatic, that is, the corresponding base-p expansions (read right-to-left) form a regular language on the alphabet $\Sigma_p := \{0, \dots, p-1\}$.

Theorem (Christol)

For k finite, $x \in k[\![T]\!]$ is algebraic if and only if it is automatic.

Christol's theorem

Throughout this talk, fix a prime number p and let k be a field of characteristic p.

Let $x = \sum_{n=0}^{\infty} x_n T^n \in k[\![T]\!]$ be a power series with coefficients in k. We say that x is *algebraic* if there exists a nonzero polynomial $P \in k[y, z]$ such that $P(T, x) = 0$. That is, as an element of the Laurent series field $k((T))$, x is integral over the subfield $k(T)$ of rational functions.

Suppose now that k is a *finite* field of characteristic p. We say that x is *automatic* if for each $c \in k$, the set $\{n \geq 0 : x_n = c\}$ is p-automatic, that is, the corresponding base-p expansions (read right-to-left) form a regular language on the alphabet $\Sigma_p := \{0, \dots, p-1\}$.

Theorem (Christol)

For k finite, $x \in k[\![T]\!]$ is algebraic if and only if it is automatic.

A few variations on Christol's theorem

Continue to assume that k is finite, of cardinality q.

- Bridy: for x in a fixed finite extension of $k(T)$, one can give a good estimate of the complexity of the corresponding automaton in terms of geometric invariants (degree, height, genus).
- Kedlaya: "algebraic equals automatic" also for generalized (univariate) power series $\sum_i x_i T^i$, where i can run over any *well-ordered* set of nonnegative rationals.
- Furstenberg: if $x = \sum_{n=0}^{\infty} x_n T^n$, $y = \sum_{n=0}^{\infty} y_n T^n$ are algebraic, then so is the Hadamard product $\sum_{n=0}^{\infty} x_n y_n T^n$.
- Salon: "algebraic equals automatic" also for multivariate power series.

A few variations on Christol's theorem

Continue to assume that k is finite, of cardinality q.

- Bridy: for x in a fixed finite extension of $k(T)$, one can give a good estimate of the complexity of the corresponding automaton in terms of geometric invariants (degree, height, genus).
- Kedlaya: "algebraic equals automatic" also for generalized (univariate) power series $\sum_i x_i T^i$, where i can run over any *well-ordered* set of nonnegative rationals.
- Furstenberg: if $x = \sum_{n=0}^{\infty} x_n T^n$, $y = \sum_{n=0}^{\infty} y_n T^n$ are algebraic, then so is the Hadamard product $\sum_{n=0}^{\infty} x_n y_n T^n$.
- Salon: "algebraic equals automatic" also for multivariate power series.

A few variations on Christol's theorem

Continue to assume that k is finite, of cardinality q.

- Bridy: for x in a fixed finite extension of $k(T)$, one can give a good estimate of the complexity of the corresponding automaton in terms of geometric invariants (degree, height, genus).
- Kedlaya: "algebraic equals automatic" also for generalized (univariate) power series $\sum_i x_i T^i$, where i can run over any *well-ordered* set of nonnegative rationals.
- Furstenberg: if $x = \sum_{n=0}^{\infty} x_n T^n$, $y = \sum_{n=0}^{\infty} y_n T^n$ are algebraic, then so is the Hadamard product $\sum_{n=0}^{\infty} x_n y_n T^n$.
- Salon: "algebraic equals automatic" also for multivariate power series.

A few variations on Christol's theorem

Continue to assume that k is finite, of cardinality q.

- Bridy: for x in a fixed finite extension of $k(T)$, one can give a good estimate of the complexity of the corresponding automaton in terms of geometric invariants (degree, height, genus).
- Kedlaya: "algebraic equals automatic" also for generalized (univariate) power series $\sum_i x_i T^i$, where i can run over any *well-ordered* set of nonnegative rationals.
- Furstenberg: if $x = \sum_{n=0}^{\infty} x_n T^n$, $y = \sum_{n=0}^{\infty} y_n T^n$ are algebraic, then so is the Hadamard product $\sum_{n=0}^{\infty} x_n y_n T^n$.
- Salon: "algebraic equals automatic" also for multivariate power series.

A few variations on Christol's theorem

Continue to assume that k is finite, of cardinality q.

- Bridy: for x in a fixed finite extension of $k(T)$, one can give a good estimate of the complexity of the corresponding automaton in terms of geometric invariants (degree, height, genus).
- Kedlaya: "algebraic equals automatic" also for generalized (univariate) power series $\sum_i x_i T^i$, where i can run over any *well-ordered* set of nonnegative rationals.
- Furstenberg: if $x = \sum_{n=0}^{\infty} x_n T^n$, $y = \sum_{n=0}^{\infty} y_n T^n$ are algebraic, then so is the Hadamard product $\sum_{n=0}^{\infty} x_n y_n T^n$.
- Salon: "algebraic equals automatic" also for multivariate power series.

Beyond finite fields

The primary goal of this talk is to indicate how theorems about automaticity of power series can be extended to *nonfinite* fields of characteristic p. There are several reasons one might want to do this.

- In 1998, I thought I had given an explicit description of an algebraic closure of $k((T))$ using generalized power series, but this turned to be incorrect for k nonfinite! Automaticity provides a way to correct it (and recover many corollaries).
- Derksen showed that if $(x_n)_{n=0}^{\infty}$ is a linear recurrent sequence over k, then $\{n \geq 0 : x_n = 0\}$ is p-automatic. Adamczewski–Bell extended this to algebraic power series. This is very suggestive!
- Furstenberg's theorem on Hadamard products remains true for nonfinite k (Deligne, Sharif–Woodcock). This is also very suggestive!
- The linear-algebraic perspective of Bridy's lecture sets this up nicely...

Beyond finite fields

The primary goal of this talk is to indicate how theorems about automaticity of power series can be extended to *nonfinite* fields of characteristic p. There are several reasons one might want to do this.

- In 1998, I thought I had given an explicit description of an algebraic closure of $k((T))$ using generalized power series, but this turned to be incorrect for k nonfinite! Automaticity provides a way to correct it (and recover many corollaries).
- Derksen showed that if $(x_n)_{n=0}^{\infty}$ is a linear recurrent sequence over k, then $\{n \geq 0 : x_n = 0\}$ is p-automatic. Adamczewski–Bell extended this to algebraic power series. This is very suggestive!
- Furstenberg's theorem on Hadamard products remains true for nonfinite k (Deligne, Sharif–Woodcock). This is also very suggestive!
- The linear-algebraic perspective of Bridy's lecture sets this up nicely...

Beyond finite fields

The primary goal of this talk is to indicate how theorems about automaticity of power series can be extended to *nonfinite* fields of characteristic p. There are several reasons one might want to do this.

- In 1998, I thought I had given an explicit description of an algebraic closure of $k((T))$ using generalized power series, but this turned to be incorrect for k nonfinite! Automaticity provides a way to correct it (and recover many corollaries).
- Derksen showed that if $(x_n)_{n=0}^{\infty}$ is a linear recurrent sequence over k, then $\{n \geq 0 : x_n = 0\}$ is p-automatic. Adamczewski–Bell extended this to algebraic power series. This is very suggestive!
- Furstenberg's theorem on Hadamard products remains true for nonfinite k (Deligne, Sharif–Woodcock). This is also very suggestive!
- The linear-algebraic perspective of Bridy's lecture sets this up nicely...

Beyond finite fields

The primary goal of this talk is to indicate how theorems about automaticity of power series can be extended to *nonfinite* fields of characteristic p. There are several reasons one might want to do this.

- In 1998, I thought I had given an explicit description of an algebraic closure of $k((T))$ using generalized power series, but this turned to be incorrect for k nonfinite! Automaticity provides a way to correct it (and recover many corollaries).
- Derksen showed that if $(x_n)_{n=0}^{\infty}$ is a linear recurrent sequence over k, then $\{n \geq 0 : x_n = 0\}$ is p-automatic. Adamczewski–Bell extended this to algebraic power series. This is very suggestive!
- Furstenberg's theorem on Hadamard products remains true for nonfinite k (Deligne, Sharif–Woodcock). This is also very suggestive!
- The linear-algebraic perspective of Bridy's lecture sets this up nicely...

Beyond finite fields

The primary goal of this talk is to indicate how theorems about automaticity of power series can be extended to *nonfinite* fields of characteristic p. There are several reasons one might want to do this.

- In 1998, I thought I had given an explicit description of an algebraic closure of $k((T))$ using generalized power series, but this turned to be incorrect for k nonfinite! Automaticity provides a way to correct it (and recover many corollaries).
- Derksen showed that if $(x_n)_{n=0}^{\infty}$ is a linear recurrent sequence over k, then $\{n \geq 0 : x_n = 0\}$ is p-automatic. Adamczewski–Bell extended this to algebraic power series. This is very suggestive!
- Furstenberg's theorem on Hadamard products remains true for nonfinite k (Deligne, Sharif–Woodcock). This is also very suggestive!
- The linear-algebraic perspective of Bridy's lecture sets this up nicely...

Contents

Algebraicity revisited

Assume again that k is finite. Let $x = \sum_{n=0}^{\infty} x_n T^n$ be an automatic power series over k. Instead of describing each of the sets $\{n \geq 0 : x_n = c\}$ with a separate automaton, it is already more efficient to combine resources.

Take $\Sigma = \{0, \dots, p-1\}$. Let L_p^0 be the subset of the language Σ^* consisting of strings not starting with 0. For $s \in L_p^0$, let $|s|$ be the nonnegative integer represented by s. Then x is automatic if the function $s \mapsto x_{|s|}$ is the function f_M arising from a *deterministic finite automaton with output*. (Equivalently, the level sets of this function form a *regular partition* of L_p^0.)

For example, the Thue–Morse series $\sum_{n=0}^{\infty} x_n T^n$ over $\mathbb{F}_2$ is computed by a DFAO with 2 states (see previous lecture).

Algebraicity revisited

Assume again that k is finite. Let $x = \sum_{n=0}^{\infty} x_n T^n$ be an automatic power series over k. Instead of describing each of the sets $\{n \geq 0 : x_n = c\}$ with a separate automaton, it is already more efficient to combine resources.

Take $\Sigma = \{0, \dots, p-1\}$. Let L_p^0 be the subset of the language Σ^* consisting of strings not starting with 0. For $s \in L_p^0$, let $|s|$ be the nonnegative integer represented by s. Then x is automatic if the function $s \mapsto x_{|s|}$ is the function f_M arising from a *deterministic finite automaton with output*. (Equivalently, the level sets of this function form a *regular partition* of L_p^0.)

For example, the Thue–Morse series $\sum_{n=0}^{\infty} x_n T^n$ over $\mathbb{F}_2$ is computed by a DFAO with 2 states (see previous lecture).

Algebraicity revisited

Assume again that k is finite. Let $x = \sum_{n=0}^{\infty} x_n T^n$ be an automatic power series over k. Instead of describing each of the sets $\{n \geq 0 : x_n = c\}$ with a separate automaton, it is already more efficient to combine resources.

Take $\Sigma = \{0, \ldots, p-1\}$. Let L_p^0 be the subset of the language Σ^* consisting of strings not starting with 0. For $s \in L_p^0$, let $|s|$ be the nonnegative integer represented by s. Then x is automatic if the function $s \mapsto x_{|s|}$ is the function f_M arising from a *deterministic finite automaton with output*. (Equivalently, the level sets of this function form a *regular partition* of L_p^0.)

For example, the Thue–Morse series $\sum_{n=0}^{\infty} x_n T^n$ over $\mathbb{F}_2$ is computed by a DFAO with 2 states (see previous lecture).

Reminder: finite automata with output

A *deterministic finite automaton with output* (DFAO) is a tuple $M = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ in which:

- Q is a finite set (the *states*);
- Σ is a finite set (the *alphabet*);
- δ is a function from $\Sigma \times Q$ to Q (the *transition function*);
- $q_0 \in Q$ is a state (the *initial state*);
- Δ is a set (the *output alphabet*);
- τ is a function from Q to Δ (the *output function*).

The function δ formally extends to $\delta^* : Q \times \Sigma^* \to Q$ thus:

$$\delta^*(q, \emptyset) = q, \quad \delta^*(q, aw) = \delta(a, \delta^*(q, w)) \qquad (q \in Q, a \in \Sigma, w \in \Sigma^*).$$

We then obtain a function $f_M : \Sigma^* \to \Delta$ by setting $f_M(w) = \tau(\delta^*(q_0, w))$.

Reminder: finite automata with output

A *deterministic finite automaton with output* (DFAO) is a tuple $M = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ in which:

- Q is a finite set (the *states*);
- Σ is a finite set (the *alphabet*);
- δ is a function from $\Sigma \times Q$ to Q (the *transition function*);
- $q_0 \in Q$ is a state (the *initial state*);
- Δ is a set (the *output alphabet*);
- τ is a function from Q to Δ (the *output function*).

The function δ formally extends to $\delta^* : Q \times \Sigma^* \to Q$ thus:

$$\delta^*(q, \emptyset) = q, \quad \delta^*(q, aw) = \delta(a, \delta^*(q, w)) \qquad (q \in Q, a \in \Sigma, w \in \Sigma^*).$$

We then obtain a function $f_M : \Sigma^* \to \Delta$ by setting $f_M(w) = \tau(\delta^*(q_0, w))$.

Composed functions

Let V be a finite-dimensional k-vector space, choose a function $\tau : \Sigma \to \mathrm{End}_k(V)$, and construct the function $f : \Sigma^* \to \mathrm{End}_k(V)$ by

$$f(a_1 \cdots a_n) = \tau(a_1) \circ \cdots \circ \tau(a_n).$$

We will say that any f occurring this way is *composed*.

We say that f is *potentially composed* if there exist a finite-dimensional k-vector space V', a k-linear injection $\iota : V \to V'$, a composed function $f' : \Sigma^* \to \mathrm{End}_k(V')$, and a k-linear surjection $\pi : V' \to V$ such that $f(s) = \pi \circ f'(s) \circ \iota$.

This is basically the same thing as using a *p-representation* in the notation of Bridy's lecture.

Composed functions

Let V be a finite-dimensional k-vector space, choose a function $\tau : \Sigma \to \mathrm{End}_k(V)$, and construct the function $f : \Sigma^* \to \mathrm{End}_k(V)$ by

$$f(a_1 \cdots a_n) = \tau(a_1) \circ \cdots \circ \tau(a_n).$$

We will say that any f occurring this way is *composed*.

We say that f is *potentially composed* if there exist a finite-dimensional k-vector space V', a k-linear injection $\iota : V \to V'$, a composed function $f' : \Sigma^* \to \mathrm{End}_k(V')$, and a k-linear surjection $\pi : V' \to V$ such that $f(s) = \pi \circ f'(s) \circ \iota$.

This is basically the same thing as using a *p-representation* in the notation of Bridy's lecture.

Composed functions

Let V be a finite-dimensional k-vector space, choose a function $\tau : \Sigma \to \mathrm{End}_k(V)$, and construct the function $f : \Sigma^* \to \mathrm{End}_k(V)$ by

$$f(a_1 \cdots a_n) = \tau(a_1) \circ \cdots \circ \tau(a_n).$$

We will say that any f occurring this way is *composed*.

We say that f is *potentially composed* if there exist a finite-dimensional k-vector space V', a k-linear injection $\iota : V \to V'$, a composed function $f' : \Sigma^* \to \mathrm{End}_k(V')$, and a k-linear surjection $\pi : V' \to V$ such that $f(s) = \pi \circ f'(s) \circ \iota$.

This is basically the same thing as using a *p-representation* in the notation of Bridy's lecture.

Automata and composed functions

Suppose that in the formula

$$f(a_1 \cdots a_n) = \tau(a_1) \circ \cdots \circ \tau(a_n),$$

we had made $\tau(a_i)$ dependent also on the class of $a_{i+1} \cdots a_n$ in some regular partition of L_p^0. Then the notion of a potentially composed function would not change: we can encode the regular partition by replacing V' with some finite direct sum of copies thereof.

From this observation, it follows easily that a function $f : \Sigma^* \to k$ is potentially composed if and only if it is automatic. That is, Christol's theorem can be stated (and proved, with good complexity bounds!) in terms of potentially composed functions.

Automata and composed functions

Suppose that in the formula

$$f(a_1 \cdots a_n) = \tau(a_1) \circ \cdots \circ \tau(a_n),$$

we had made $\tau(a_i)$ dependent also on the class of $a_{i+1} \cdots a_n$ in some regular partition of L_p^0. Then the notion of a potentially composed function would not change: we can encode the regular partition by replacing V' with some finite direct sum of copies thereof.

From this observation, it follows easily that a function $f : \Sigma^* \to k$ is potentially composed if and only if it is automatic. That is, Christol's theorem can be stated (and proved, with good complexity bounds!) in terms of potentially composed functions.

Contents

Nonfinite fields

Let k be a perfect, but not finite, field of characteristic p. In order to describe algebraic series over k, we must modify the previous setup in one crucial way: instead of *linear* transformations on V, we consider φ-*semilinear* transformations $T : V \to V$, which satisfy

$$T(r_1 v_1 + r_2 v_2) = \varphi(r_1) T(v_1) + \varphi(r_2) T(v_2) \qquad (r_1, r_2 \in k; v_1, v_2 \in V)$$

where $\varphi(r) = r^p$. (For compatibility with the previous discussion, I could instead take $\varphi(r) = r^q$, but going forward I won't need to.)

Note that the composition of φ-semilinear transformations is *not* itself φ-semilinear, so we need to be careful about definitions. For a composed function, we now want

$$f(a_1 \cdots a_n) = \tau(a_1) \circ \cdots \circ \tau(a_n)$$

where each $\tau(a_i)$ is φ-semilinear (but $f(a_1 \cdots a_n)$ is φ^n-semilinear).

Nonfinite fields

Let k be a perfect, but not finite, field of characteristic p. In order to describe algebraic series over k, we must modify the previous setup in one crucial way: instead of *linear* transformations on V, we consider φ-*semilinear* transformations $T : V \to V$, which satisfy

$$T(r_1 v_1 + r_2 v_2) = \varphi(r_1) T(v_1) + \varphi(r_2) T(v_2) \qquad (r_1, r_2 \in k; v_1, v_2 \in V)$$

where $\varphi(r) = r^p$. (For compatibility with the previous discussion, I could instead take $\varphi(r) = r^q$, but going forward I won't need to.)

Note that the composition of φ-semilinear transformations is *not* itself φ-semilinear, so we need to be careful about definitions. For a composed function, we now want

$$f(a_1 \cdots a_n) = \tau(a_1) \circ \cdots \circ \tau(a_n)$$

where each $\tau(a_i)$ is φ-semilinear (but $f(a_1 \cdots a_n)$ is φ^n-semilinear).

A Christol-type theorem for nonfinite fields

Theorem

Suppose that k is perfect. For $x = \sum_{n=0}^{\infty} x_n T^n \in k[\![T]\!]$, x is algebraic if and only if the function $f : L_p^0 \to k$ given by $f(s) = x_{|s|}$ is potentially composed. (To interpret "potentially composed", treat k as a one-dimensional vector space over itself.)

The proof is an immediate adaptation of the Speyer–Bridy approach to Christol's theorem, using the Cartier operator

$$(u_0^p + u_1^p \, dT + \cdots + u_{p-1}^p \, dT) \mapsto u_{p-1} \, dT.$$

Note that this is not k-linear, but φ^{-1}-semilinear. (Presumably the bound on dimension is the same as in Bridy, but I didn't optimize for this.)

Corollary (Deligne, Sharif–Woodcock)

For any k, the Hadamard product of two algebraic series is algebraic.

A Christol-type theorem for nonfinite fields

Theorem

Suppose that k is perfect. For $x = \sum_{n=0}^{\infty} x_n T^n \in k[\![T]\!]$, x is algebraic if and only if the function $f : L_p^0 \to k$ given by $f(s) = x_{|s|}$ is potentially composed. (To interpret "potentially composed", treat k as a one-dimensional vector space over itself.)

The proof is an immediate adaptation of the Speyer–Bridy approach to Christol's theorem, using the Cartier operator

$$(u_0^p + u_1^p \, dT + \cdots + u_{p-1}^p \, dT) \mapsto u_{p-1} \, dT.$$

Note that this is not k-linear, but φ^{-1}-semilinear. (Presumably the bound on dimension is the same as in Bridy, but I didn't optimize for this.)

Corollary (Deligne, Sharif–Woodcock)

For any k, the Hadamard product of two algebraic series is algebraic.

A Christol-type theorem for nonfinite fields

Theorem

Suppose that k is perfect. For $x = \sum_{n=0}^{\infty} x_n T^n \in k[\![T]\!]$*, x is algebraic if and only if the function* $f : L_p^0 \to k$ *given by* $f(s) = x_{|s|}$ *is potentially composed. (To interpret "potentially composed", treat k as a one-dimensional vector space over itself.)*

The proof is an immediate adaptation of the Speyer–Bridy approach to Christol's theorem, using the Cartier operator

$$(u_0^p + u_1^p\, dT + \cdots + u_{p-1}^p\, dT) \mapsto u_{p-1}\, dT.$$

Note that this is not k-linear, but φ^{-1}-semilinear. (Presumably the bound on dimension is the same as in Bridy, but I didn't optimize for this.)

Corollary (Deligne, Sharif–Woodcock)

For any k, the Hadamard product of two algebraic series is algebraic.

An application to zero sets

Corollary (Derksen, Adamczewski–Bell)

For any k, for $x = \sum_{n=0}^{\infty} x_n T^n \in k[\![T]\!]$ algebraic, the set $\{n \geq 0 : x_n = 0\}$ is p-automatic.

Proof.

Formally reduce to the case where k is finite over $\mathbb{F}_p(t_1, \dots, t_m)$ for some m. By the previous corollary, we may take norms to reduce to the case $k = \mathbb{F}_p(t_1, \dots, t_m)$, then rescale to force $x_n \in \mathbb{F}_p[t_1, \dots, t_m]$ for all n. Now view x in $\mathbb{F}_p[\![t_1, \dots, t_m, T]\!]$ and apply Salon's theorem and the following lemma with $\Sigma_1 = \{0, \dots, p-1\}^{m+1}$, $\Sigma_2 = \{0, \dots, p-1\}$. □

Lemma (exercise)

Let $\Sigma_1 \to \Sigma_2$ be any map of finite sets. Let L be a regular language on Σ_1. Then the image of substitution $L \to \Sigma_2^*$ is again a regular language.

An application to zero sets

Corollary (Derksen, Adamczewski–Bell)

For any k, for $x = \sum_{n=0}^{\infty} x_n T^n \in k[\![T]\!]$ algebraic, the set $\{n \geq 0 : x_n = 0\}$ is p-automatic.

Proof.

Formally reduce to the case where k is finite over $\mathbb{F}_p(t_1, \dots, t_m)$ for some m. By the previous corollary, we may take norms to reduce to the case $k = \mathbb{F}_p(t_1, \dots, t_m)$, then rescale to force $x_n \in \mathbb{F}_p[t_1, \dots, t_m]$ for all n. Now view x in $\mathbb{F}_p[\![t_1, \dots, t_m, T]\!]$ and apply Salon's theorem and the following lemma with $\Sigma_1 = \{0, \dots, p-1\}^{m+1}$, $\Sigma_2 = \{0, \dots, p-1\}$. □

Lemma (exercise)

Let $\Sigma_1 \to \Sigma_2$ be any map of finite sets. Let L be a regular language on Σ_1. Then the image of substitution $L \to \Sigma_2^*$ is again a regular language.

An application to zero sets

Corollary (Derksen, Adamczewski–Bell)

For any k, for $x = \sum_{n=0}^{\infty} x_n T^n \in k[\![T]\!]$ algebraic, the set $\{n \geq 0 : x_n = 0\}$ is p-automatic.

Proof.

Formally reduce to the case where k is finite over $\mathbb{F}_p(t_1, \dots, t_m)$ for some m. By the previous corollary, we may take norms to reduce to the case $k = \mathbb{F}_p(t_1, \dots, t_m)$, then rescale to force $x_n \in \mathbb{F}_p[t_1, \dots, t_m]$ for all n. Now view x in $\mathbb{F}_p[\![t_1, \dots, t_m, T]\!]$ and apply Salon's theorem and the following lemma with $\Sigma_1 = \{0, \dots, p-1\}^{m+1}$, $\Sigma_2 = \{0, \dots, p-1\}$. □

Lemma (exercise)

Let $\Sigma_1 \to \Sigma_2$ be any map of finite sets. Let L be a regular language on Σ_1. Then the image of substitution $L \to \Sigma_2^*$ is again a regular language.

Generalized power series

A *generalized power series* over a field k is a formal sum $\sum_{i \in \mathbb{Q}_{\geq 0}} x_i T^i$ with $x_i \in k$ such that the support $\{i \in \mathbb{Q}_{\geq 0} : x_i \neq 0\}$ is *well-ordered* (contains no infinite decreasing sequence). These were introduced by Hahn (1904); a noncommutative analogue was considered by Malcev–Neumann.

One reason to consider generalized power series is that when k is of characteristic p, the algebraic closure of $k((T))$ cannot be obtained simply by forming the union of $k((T^{1/d}))$ over all $d > 0$ (i.e., the field of *Puiseux series*). For example, the roots of the equation

$$z^p - z = T^{-1}$$

have the form

$$z = c + T^{-1/p} + T^{-1/p^2} + \cdots \qquad (c \in \mathbb{F}_p).$$

Generalized power series

A *generalized power series* over a field k is a formal sum $\sum_{i \in \mathbb{Q}_{\geq 0}} x_i T^i$ with $x_i \in k$ such that the support $\{i \in \mathbb{Q}_{\geq 0} : x_i \neq 0\}$ is *well-ordered* (contains no infinite decreasing sequence). These were introduced by Hahn (1904); a noncommutative analogue was considered by Malcev–Neumann.

One reason to consider generalized power series is that when k is of characteristic p, the algebraic closure of $k((T))$ cannot be obtained simply by forming the union of $k((T^{1/d}))$ over all $d > 0$ (i.e., the field of *Puiseux series*). For example, the roots of the equation

$$z^p - z = T^{-1}$$

have the form

$$z = c + T^{-1/p} + T^{-1/p^2} + \cdots \qquad (c \in \mathbb{F}_p).$$

Automatic generalized power series

Let L_p be the language on $\{0, \dots, p-1, .\}$ consisting of strings with a single "." which do not start or end with 0. We regard these as the base-p expansions of elements of $\mathbb{Z}[p^{-1}]_{\geq 0}$; let $\|s\|$ denote the number represented by the string s.

We again define *composed* and *potentially composed* functions, but with a twist: in the formula

$$f(a_1 \cdots a_m.b_1 \cdots b_m) = \tau(a_1) \cdots \tau(a_m)\tau(.)\tau(b_1) \cdots \tau(b_m),$$

the linear transformations $\tau(a_i)$ are φ-semilinear, but the $\tau(b_j)$ are φ^{-1}-semilinear (and $\tau(.)$ is k-linear).

Warning: most potentially composed functions do not give rise to formal sums with well-ordered support!

Automatic generalized power series

Let L_p be the language on $\{0, \dots, p-1, .\}$ consisting of strings with a single "." which do not start or end with 0. We regard these as the base-p expansions of elements of $\mathbb{Z}[p^{-1}]_{\geq 0}$; let $\|s\|$ denote the number represented by the string s.

We again define *composed* and *potentially composed* functions, but with a twist: in the formula

$$f(a_1 \cdots a_m . b_1 \cdots b_m) = \tau(a_1) \cdots \tau(a_m) \tau(.) \tau(b_1) \cdots \tau(b_m),$$

the linear transformations $\tau(a_i)$ are φ-semilinear, but the $\tau(b_j)$ are φ^{-1}-semilinear (and $\tau(.)$ is k-linear).

Warning: most potentially composed functions do not give rise to formal sums with well-ordered support!

Automatic generalized power series

Let L_p be the language on $\{0, \dots, p-1, .\}$ consisting of strings with a single "." which do not start or end with 0. We regard these as the base-p expansions of elements of $\mathbb{Z}[p^{-1}]_{\geq 0}$; let $\|s\|$ denote the number represented by the string s.

We again define *composed* and *potentially composed* functions, but with a twist: in the formula

$$f(a_1 \cdots a_m.b_1 \cdots b_m) = \tau(a_1) \cdots \tau(a_m)\tau(.)\tau(b_1) \cdots \tau(b_m),$$

the linear transformations $\tau(a_i)$ are φ-semilinear, but the $\tau(b_j)$ are φ^{-1}-semilinear (and $\tau(.)$ is k-linear).

Warning: most potentially composed functions do not give rise to formal sums with well-ordered support!

An "algebraic equals automatic" theorem for generalized power series

Theorem

For k perfect, a generalized power series $\sum_{i \in \mathbb{Z}[p^{-1}]_{\geq 0}} x_i T^i$ is algebraic over $k(T)$ if and only if the function $L_p \to k$ taking s to $x_{\|s\|}$ is potentially composed. (For k finite, one may read "automatic" for "potentially composed.")

After adjoining T^{-1} and $T^{1/d}$ for d coprime to p, this produces the full integral closure of $k(T)$ within the field of generalized power series.

A bizarre corollary: if one reverses all the base-p-expansions, one gets another algebraic generalized power series *provided* that the support is still well-ordered. Some special cases which can be shown explicitly are actually exploited in the proof!

An "algebraic equals automatic" theorem for generalized power series

Theorem

For k perfect, a generalized power series $\sum_{i \in \mathbb{Z}[p^{-1}]_{\geq 0}} x_i T^i$ is algebraic over $k(T)$ if and only if the function $L_p \to k$ taking s to $x_{\|s\|}$ is potentially composed. (For k finite, one may read "automatic" for "potentially composed.")

After adjoining T^{-1} and $T^{1/d}$ for d coprime to p, this produces the full integral closure of $k(T)$ within the field of generalized power series.

A bizarre corollary: if one reverses all the base-p-expansions, one gets another algebraic generalized power series *provided* that the support is still well-ordered. Some special cases which can be shown explicitly are actually exploited in the proof!

An "algebraic equals automatic" theorem for generalized power series

Theorem

For k perfect, a generalized power series $\sum_{i \in \mathbb{Z}[p^{-1}]_{\geq 0}} x_i T^i$ is algebraic over $k(T)$ if and only if the function $L_p \to k$ taking s to $x_{\|s\|}$ is potentially composed. (For k finite, one may read "automatic" for "potentially composed.")

After adjoining T^{-1} and $T^{1/d}$ for d coprime to p, this produces the full integral closure of $k(T)$ within the field of generalized power series.

A bizarre corollary: if one reverses all the base-p-expansions, one gets another algebraic generalized power series *provided* that the support is still well-ordered. Some special cases which can be shown explicitly are actually exploited in the proof!

Corollaries

Here are some corollaries of the previous theorem (for any k).

Corollary

For $\sum_{i \in \mathbb{Z}[p^{-1}]_{\geq 0}} x_i T^i$ algebraic over $k(T)$, any truncation $\sum_{i<j} x_i T^i$ is again algebraic.

Corollary

The Hadamard product of two algebraic generalized power series over k is algebraic.

Corollary

For $\sum_{i \in \mathbb{Z}[p^{-1}]_{\geq 0}} x_i T^i$ algebraic over $k(T)$, the set $\{i \in \mathbb{Z}[p^{-1}]_{\geq 0} : x_i = 0\}$ is p-automatic.

Corollaries

Here are some corollaries of the previous theorem (for any k).

Corollary

For $\sum_{i \in \mathbb{Z}[p^{-1}]_{\geq 0}} x_i T^i$ algebraic over $k(T)$, any truncation $\sum_{i<j} x_i T^i$ is again algebraic.

Corollary

The Hadamard product of two algebraic generalized power series over k is algebraic.

Corollary

For $\sum_{i \in \mathbb{Z}[p^{-1}]_{\geq 0}} x_i T^i$ algebraic over $k(T)$, the set $\{i \in \mathbb{Z}[p^{-1}]_{\geq 0} : x_i = 0\}$ is p-automatic.

Corollaries

Here are some corollaries of the previous theorem (for any k).

Corollary

For $\sum_{i \in \mathbb{Z}[p^{-1}]_{\geq 0}} x_i T^i$ algebraic over $k(T)$, any truncation $\sum_{i<j} x_i T^i$ is again algebraic.

Corollary

The Hadamard product of two algebraic generalized power series over k is algebraic.

Corollary

For $\sum_{i \in \mathbb{Z}[p^{-1}]_{\geq 0}} x_i T^i$ algebraic over $k(T)$, the set $\{i \in \mathbb{Z}[p^{-1}]_{\geq 0} : x_i = 0\}$ is p-automatic.

Corollaries

Here are some corollaries of the previous theorem (for any k).

Corollary

For $\sum_{i \in \mathbb{Z}[p^{-1}]_{\geq 0}} x_i T^i$ algebraic over $k(T)$, any truncation $\sum_{i<j} x_i T^i$ is again algebraic.

Corollary

The Hadamard product of two algebraic generalized power series over k is algebraic.

Corollary

For $\sum_{i \in \mathbb{Z}[p^{-1}]_{\geq 0}} x_i T^i$ algebraic over $k(T)$, the set $\{i \in \mathbb{Z}[p^{-1}]_{\geq 0} : x_i = 0\}$ is p-automatic.

Contents

Mixed characteristic

As far as I know, there is no mixed-characteristic analogue of Christol's theorem to pick out, say, the integral closure of $\mathbb{Q}$ within $\mathbb{Q}_p$.

However, there is a mixed-characteristic analogue of the generalized power series field (considered by Poonen), and in that field one can identify the integral closure of $\mathbb{Q}_p$ using automata. The new result shows that this can also be achieved with $\mathbb{Q}_p$ replaced by $W(k)[p^{-1}]$ for k perfect.

Mixed characteristic

As far as I know, there is no mixed-characteristic analogue of Christol's theorem to pick out, say, the integral closure of $\mathbb{Q}$ within $\mathbb{Q}_p$.

However, there is a mixed-characteristic analogue of the generalized power series field (considered by Poonen), and in that field one can identify the integral closure of $\mathbb{Q}_p$ using automata. The new result shows that this can also be achieved with $\mathbb{Q}_p$ replaced by $W(k)[p^{-1}]$ for k perfect.

More questions

- Bridy's bounds should carry over easily to ordinary power series over a perfect field. What about generalized power series over a finite field? Or a perfect field?
- Is there an "algebraic implies automatic" theorem that simultaneously includes multivariate power series and univariate generalized power series? Probably yes, but the tricky part is to give a formulation that captures a full algebraic closure of $k(T_1, \dots, T_n)$.

More questions

- Bridy's bounds should carry over easily to ordinary power series over a perfect field. What about generalized power series over a finite field? Or a perfect field?
- Is there an "algebraic implies automatic" theorem that simultaneously includes multivariate power series and univariate generalized power series? Probably yes, but the tricky part is to give a formulation that captures a full algebraic closure of $k(T_1, \ldots, T_n)$.

More questions

- Bridy's bounds should carry over easily to ordinary power series over a perfect field. What about generalized power series over a finite field? Or a perfect field?
- Is there an "algebraic implies automatic" theorem that simultaneously includes multivariate power series and univariate generalized power series? Probably yes, but the tricky part is to give a formulation that captures a full algebraic closure of $k(T_1, \ldots, T_n)$.