

On Presentation Complexity of Number Fields

Victor Selivanov¹

A.P. Ershov IIS SB RAS
and
Kazan Federal University

Chebotarev-Arslanov Conference, Kazan, June 24–28, 2019

1. Introduction
2. Algorithmic problems in field theory.
3. Computably presentable fields of computable reals.
4. Applications to computability of PDE-solving.
5. Polynomial presentations of the field of algebraic reals.
6. Applications to complexity of PDE-solving.
7. Concluding remarks.

Introduction

The algorithms used in mathematics-oriented software can be divided into two big classes: symbolic algorithms which aim to find precise solutions, and approximate algorithms which aim to find “good enough” approximations to precise solutions. The symbolic algorithms are implemented e.g. in computer algebra systems or SMT-solvers while the approximate algorithms - in numerical analysis packages.

The both classes of algorithms are widely used in applications and in mathematical research. The symbolic algorithms correspond well to computations on discrete structures (with mathematical foundations in the classical computability and complexity theory) while the approximate algorithms - to computations on continuous structures (with mathematical foundations in the field of computability and complexity in analysis evolving under the slogan “Exact real computation”).

An important idea relating the both classes of algorithms is to look for approximate solutions to a numerical problem with guaranteed precision. Finding such a solution is of crucial importance for safety-critical applications but it often requires much additional work because it needs a sophisticated algorithm and careful estimations of approximations made during the computation. In many cases the statement of a guaranteed-precision version of some problem on a continuous structure (which requires numerical mathematics and/or computable analysis) reduces it to a problem on a discrete structure which enables to apply the classical computability and complexity theory (sometimes called bit complexity).

Introduction

In this talk we discuss three topics. First, we establish close relations of computably presentable fields of reals to the ordered field of computable reals.

Second, we partially fill the gap between the above-mentioned theories by applying the notions of computability theory to the investigation of some presentations of \mathbb{R}_{alg} and \mathbb{C}_{alg} . In particular, we show that the notion of p -time presentable structure is not applicable to some presentations of \mathbb{R}_{alg} and \mathbb{C}_{alg} in the literature. We introduce a more general notion of p -computable quotient-structure and show that several natural presentations of \mathbb{R}_{alg} and \mathbb{C}_{alg} are p -time equivalent to each other, and are p -time computable. Thus, we clear up the conceptual basis for the complexity theory of structure presentations.

Third, we discuss the complexity of the problem of rational polynomial evaluation in \mathbb{R}_{alg} and \mathbb{C}_{alg} , and of the problem of root-finding for polynomials in $\mathbb{C}_{\text{alg}}[x]$.

Algorithmic problems in field theory

Based on the notion of a computable structure, the computability issues in algebra and model theory were thoroughly investigated. In particular, a rich and useful theory of computable fields was developed.

For instance, M. Rabin in 1960 has shown that the algebraic closure of a computable field is computably presentable, and Yu.L. Ershov in 1977 has shown that the real algebraic closure of a computable ordered field is computably presentable.

Since the ordered field \mathbb{Q} of rationals is computably presentable, the field $\mathbb{C}_{\text{alg}} = (C_{\text{alg}}; +, \times, 0, 1)$ of complex algebraic numbers and the ordered field $\mathbb{R}_{\text{alg}} = (R_{\text{alg}}; \leq, +, \times, 0, 1)$ of algebraic reals are computably presentable.

Presentation complexity of structures

In applications one of course has to pay attention to the complexity of implemented algorithms and of structure presentations. The complexity of structure presentations was first studied by Nerode, Cenzer and Remmel. In particular, the notion of a polynomial-time (p-time) structure was introduced. To our knowledge, the complexity issues for presentations of fields were not studied in computability theory so far.

At the same time, there exists a well-developed theory of symbolic computations (closely related to computer algebra) which investigates the complexity of algorithms in fields, of concrete presentations of fields and rings, and aims to implement these in computer systems. In particular, there is a vast literature around the Tarski theorems on decidability of the theories of algebraically closed fields and of real closed fields. Although the mentioned theories are obviously intimately related, they developed apparently independently and there are essentially no references between them.

Computable reals

The field \mathbb{R}_c of all computable reals is countable, real closed, and not computably presentable. But, in some sense, it is “partially computably presentable”.

Let \varkappa — be a constructivisation of \mathbb{Q} and $\{\varphi_n\}$ be the standard computable numbering of all computable partial functions on \mathbb{N} . Define a partial function ρ from \mathbb{N} onto \mathbb{R}_c : $\rho(n) = x$ iff φ_n is total and $\{\varkappa\varphi_n(i)\}_i$ is a fast Cauchy sequence converging to x .

A numbering μ is *reducible* to a (partial) numbering ν ($\mu \leq \nu$), if $\mu = \nu \circ f$ for some computable function f on \mathbb{N} .

Computable reals

How to measure the complexity of presenting the ordered field \mathbb{R}_c ?
The following result independently obtained by Korovina-Kudinov and by Downey-Greenberg-Miller, characterizes the so called degree spectrum of \mathbb{R}_c .

T h e o r e m. The structure \mathbb{R}_c admits a computable presentation relative to a Turing degree \mathbf{a} if and only if $\mathbf{0}'' \leq \mathbf{a}'$.

Since \mathbb{R}_c does not have a computable presentation, it is natural to ask whether it has a positive or negative presentation (the notions are due to A.I. Maltsev). Recall that a structure has a positive (resp. negative) presentation if there is its numbering modulo which its operations are computable and the relations (including the equality relation) are c.e. (resp. co-c.e.).

P r o p o s i t i o n. The structure \mathbb{R}_c admits neither positive nor negative presentations.

Proposition 1. Let \mathbb{B} be a computable ordered subfield of \mathbb{R} , and β be a constructivisation of \mathbb{B} . Then $\beta \leq \rho$, in particular $\mathbb{B} \subseteq \mathbb{R}_c$.

Proposition 2. Let \mathbb{B} be a subfield of $(\mathbb{R}; +, \cdot, 0, 1)$ and β be a constructivisation of \mathbb{B} such that $\beta \leq \rho$. Then β is a constructivisation of the ordered field $(\mathbb{B}; <)$.

Proposition 3. Let \mathbb{B} be a real closed subfield of $(\mathbb{R}; +, \cdot, 0, 1)$ β be a constructivisation of \mathbb{B} . Then β is a strong constructivisation of the ordered field $(\mathbb{B}; <)$.

Adjoining computable reals

We add the following theorem to the results of the previous slide. The theorem relates constructive fields of reals to the field \mathbb{R}_c of computable reals.

T h e o r e m. For any finite set $F \subseteq \mathbb{R}_c$ there is a strongly constructive real closed subfield (\mathbb{B}, β) of the ordered field \mathbb{R}_c such that $F \subseteq B$.

The results of the last two slides, as well as the results below concerning PDEs, are joint with Svetlana Selivanova.

We consider the initial-value problem

$$\begin{cases} A \frac{\partial \mathbf{u}}{\partial t} + \sum_{i=1}^m B_i \frac{\partial \mathbf{u}}{\partial x_i} = f(t, \mathbf{x}), & t \geq 0, \\ \mathbf{u}|_{t=0} = \varphi(x_1, \dots, x_m). \end{cases} \quad (1)$$

Here $A = A^* > 0$ and $B_i = B_i^*$ are constant symmetric $n \times n$ -matrices, $t \geq 0$, $\mathbf{x} = (x_1, \dots, x_m) \in Q = [0, 1]^m$, $\varphi : Q \rightarrow \mathbb{R}^n$ and $\mathbf{u} : [0, +\infty) \times Q \rightarrow \mathbb{R}^n$ is a partial function acting on the domain H of existence and uniqueness of the Cauchy problem (1). The solution \mathbf{u} depends continuously on $\varphi, A, B_1, \dots, B_m$.

Computability of PDEs

We also consider the boundary-value problem:

$$\begin{cases} A \frac{\partial \mathbf{u}}{\partial t} + \sum_{i=1}^m B_i \frac{\partial \mathbf{u}}{\partial x_i} = f(t, \mathbf{x}), \\ \mathbf{u}|_{t=0} = \varphi(x_1, \dots, x_m), \\ \Phi_i^{(1)} \mathbf{u}(t, x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_m) = 0, \\ \Phi_i^{(2)} \mathbf{u}(t, x_1, \dots, x_{i-1}, 1, x_{i+1}, \dots, x_m) = 0, \\ i = 1, 2, \dots, m, \end{cases} \quad (2)$$

where the boundary coefficients $\Phi_i^{(1)}$, $\Phi_i^{(2)}$ are constant rectangular matrices meeting the following conditions:

- 1) (provides existence) The number of rows of $\Phi_i^{(1)}$ (respectively, $\Phi_i^{(2)}$) is equal to the number of positive (respectively, negative) eigenvalues of the matrices $A^{-1}B_i$;
- 2) (provides uniqueness) Dissipativity:

$$(B_i \mathbf{u}, \mathbf{u}) \leq 0 \text{ for } x_i = 0, \quad (B_i \mathbf{u}, \mathbf{u}) \geq 0 \text{ for } x_i = 1, \quad i = 1, 2, \dots, m.$$

Computability of PDEs

Symmetric hyperbolic systems can be used to describe a wide variety of physical processes like those considered in the theories of elasticity, acoustics, electromagnetism etc., see e.g. [Friedrichs 1954, Godunov 1971,76, Landau, Lifschitz 1986 etc.]. They were first considered in 1954 by K.O. Friedrichs. He proved the existence theorem based on **finite difference approximations**, in contrast with the Schauder-Cauchy-Kovalevskaya method based on approximations by analytic functions and a careful study of infinite series. The notion of a hyperbolic system (applicable also to broader classes of systems) is due to I.G. Petrovski.

Questions: Is the solution \mathbf{u} computable

I. from given initial conditions φ and right-hand part f (with fixed computable coefficients),

II. from φ , f **and** coefficients A, B_i

and in which sense?

III. If yes, what is the complexity of computations?

Most of (few) papers on computability of PDEs rely on explicit solution formulas. As is well-known, **explicit solution formulas exist rarely**. Even for the simplest example of the wave equation the computability of the solution operator for boundary-value problem was formulated in [Weirauch, Zhong 2002] as an open question, and we have not seen any paper where this question would be answered. Results of this paper provide, in particular, a positive answer to this question for dissipative boundary conditions. Moreover, we hope that our methods can be applied to study computability of many other PDEs in the framework of computable analysis going back to A. Turing (1937) and A. Grzegorzczuk (1957), recently developed by M. Pour El, J. Richards, Ker-I Ko, K. Weihrauch and others.

Results on computability in PDEs

I. For fixed computable matrices, the solution operator $(\varphi, f) \mapsto \mathbf{u}$ of (1), (2) is computable provided that the first and second partial derivatives of φ, f are uniformly bounded.

II. 1) The operator $(A, B_1, \dots, B_m) \mapsto H$ is computable;

2) The solution operator $(\varphi, f, A, B_1, \dots, B_m, n_A, n_1, \dots, n_m) \mapsto \mathbf{u}$ of (1), (2) is computable under some additional spectral conditions on A, B_j .

Here n_A is the cardinality of spectrum of A (i.e. the number of different eigenvalues);

n_j are the cardinalities of spectra of the matrix pencils $\lambda A - B_j$.

Eigenvectors are in general not computable!

3) The solution operator $(\varphi, f, A, B_1, \dots, B_m) \mapsto \mathbf{u}$ of (1), (2) is computable when the coefficients of A, B_j run through an arbitrary computable real closed subfield of \mathbb{R} .

Polynomial presentations of structures

All results on polynomial-time presentations below are joint with Pavel Alaev.

A structure $(A; \sigma)$ of a finite signature σ is *p-computably presentable* if it is isomorphic to a polynomial-time computable (p-computable) structure $(B; \sigma)$ where $B \subseteq \Sigma^*$ is p-computable (for a finite alphabet Σ), as well as all the signature functions and predicates.

We illustrate the introduced notion for the structures \mathbb{R}_{alg} and \mathbb{C}_{alg} , using some standard notions and facts. With any $\alpha \in R_{\text{alg}}$ we associate the unique pair (p_α, k) where $p_\alpha \in \mathbb{Q}[x]$ is the minimal (hence, irreducible) unitary polynomial of degree ≥ 1 with $p_\alpha(\alpha) = 0$, and k satisfies $\alpha = \alpha_k$ where $\alpha_1 < \dots < \alpha_m$ is the increasing sequence of all real roots of p_α .

Polynomial presentations of structures

The standard binary encoding $b : \mathbb{Q} \rightarrow \{0, 1\}^*$ induces an encoding $b : \mathbb{Q}[x] \rightarrow \{0, 1, *\}^*$, which associates with a polynomial $p(x) = a_n x^n + \dots + a_1 x + a_0$, $a_n \neq 0$ if $n \neq 0$, the code $b(a_n) * \dots * b(a_0)$. Now we associate with any $\alpha \in R_{\text{alg}}$ the word $b(p_\alpha) * b(k)$ where (p_α, k) is the pair from the previous slide, which yields an injection $b : R_{\text{alg}} \rightarrow \{0, 1, *\}^*$.

Let now $\mathbb{R}_1 = (R_1; \leq, +, \times, 0, 1)$, where $R_1 = b(R_{\text{alg}})$, be the isomorphic copy of \mathbb{R}_{alg} induced by b ; we call it the *order presentation of \mathbb{R}_{alg}* .

Complexity of order presentations

The bijection $b : R_{\text{alg}} \rightarrow R_1$ and the Gauss representation $z = x + iy$ of complex numbers induce a bijection between $R_1 \times R_1$ and C_{alg} . By encoding again the elements of $R_1 \times R_1$ by words in a finite alphabet in a standard way, we obtain a bijection $b : C_{\text{alg}} \rightarrow C_1$ which induces an isomorphism $g : C_{\text{alg}} \rightarrow C_1 = (C_1; +, \times, 0, 1)$. Informally, C_1 is the product $\mathbb{R}_1 \times \mathbb{R}_1$.

T h e o r e m. The structures \mathbb{R}_1 and C_1 are p-computable, and the operations $-x$ and $1/x$ in these fields are also p-computable. As a corollary, \mathbb{R}_{alg} and C_{alg} are p-computably presentable.

Now we define other natural presentations of \mathbb{R}_{alg} and \mathbb{C}_{alg} known in the literature. For any polynomial $p \in \mathbb{Q}[x]$ of degree ≥ 1 without multiple roots, let $p'(x), p''(x), \dots, p^{(n-1)}(x)$ be the sequence of its derivative polynomials. For any $x \in \mathbb{R}$, let $\bar{\varepsilon}_p(x) = (\varepsilon_1(x), \dots, \varepsilon_{n-1}(x))$ where $\varepsilon_i(x) = 1, 0, -1$ iff $p^{(i)}(x)$ is resp. positive, zero, or negative. It is known that $\bar{\varepsilon}_p(\alpha) \neq \bar{\varepsilon}_p(\beta)$ whenever α and β are distinct roots of $p(x)$. Associate with any $\alpha \in R_{\text{alg}}$ the unique pair $(p_\alpha, \bar{\varepsilon}_{p_\alpha}(\alpha))$, and let R_2 be the set of codes of such pairs in a natural word encoding based on the above-mentioned encoding of rational polynomials and a natural encoding of sequences of $1, 0, -1$. Let $\mathbb{R}_2 = (R_2; <, +, \times, 0, 1)$ be the isomorphic copy of \mathbb{R}_{alg} induced by the bijection $\alpha \mapsto (p_\alpha, \bar{\varepsilon}_{p_\alpha}(\alpha))$. We call the presentation \mathbb{R}_2 of \mathbb{R}_{alg} *sign presentation*.

Interval presentations

We can also code a real $\alpha \in \mathbb{R}_{\text{alg}}$ by a pair $(p(x), I)$, where $p(x) \in \mathbb{Q}[x] \setminus \{0\}$, $p(\alpha) = 0$, and $I = (a, b]$ is an isolating rational interval for α including α and no other roots of $p(x)$. Call two pairs *equivalent*, $(p_1(x), I_1) \sim (p_2(x), I_2)$, if they encode the same real. Let

$A_3 = \{b(p(x)) * b(a) * b(b) \mid p(x) \in \mathbb{Q}[x], a, b \in \mathbb{Q} \text{ and } (p(x), I = (a, b]) \text{ encodes some } \alpha \in \mathbb{R}\}$,

let $E_3 \subseteq A_3 \times A_3$ be the relation corresponding to the equivalence of pairs, and let $R_3 = A_3/E_3$ be the corresponding quotient-set.

Let $\mathbb{R}_3 = (R_3; <, +, \times, 0, 1)$ be the corresponding isomorphic copy of \mathbb{R}_{alg} .

We call this presentation of \mathbb{R}_{alg} the *interval presentation*.

Equivalence of presentations

A similar interval presentation of \mathbb{C}_{alg} is also known in the literature. We say that a triple (p, I, K) , where p is a polynomial and I, K are rational intervals as above, *defines the number* $z \in \mathbb{C}$ if z is the unique root of p in the rectangle $I + iK$. Let C be the set of codes of such triples (p, I, K) in a natural encoding, $\gamma : C \rightarrow \mathbb{C}_{\text{alg}}$ be the surjection defined similarly to the previous paragraph (of course, γ is not a bijection), and E be the corresponding equivalence relation on C . Then we have a presentation of \mathbb{C}_{alg} as a quotient-structure $\mathbb{C}_2 = (C/E; +, \times, 0, 1)$.

Theorem. The quotient-structures $\mathbb{R}_2, \mathbb{R}_3$ are p-computably isomorphic to \mathbb{R}_1 and are therefore p-computable. The quotient-structure \mathbb{C}_2 is p-computably isomorphic to \mathbb{C}_1 and is therefore p-computable.

Rational polynomial evaluation

Theorem. There exists an algorithm which, given $k \geq 1$, $\alpha_1, \dots, \alpha_k \in \mathcal{C}_{\text{alg}}$ and $t(x_1, \dots, x_k) \in \mathbb{Q}[x_1, \dots, x_k]$, finds $\beta = t(\alpha_1, \dots, \alpha_k) \in \mathcal{C}_{\text{alg}}$.

Let $n_i = \deg[\alpha_i]$ for each $i \leq k$, and let $n = \max_{i \leq k} \{n_i\}$. Then the working time of the algorithm is bounded by $(n_1 n_2 \cdots n_k)^c L^d$, or $n^{ck} L^d$, where c, d are some constants and L is the input length. In particular, for a fixed k we get a p-computable function that evaluates polynomials from $\mathbb{Q}[x_1, \dots, x_k]$. Also, $\deg[\beta] \leq \prod_{i \leq k} \deg[\alpha_i]$.

It might be shown that the algorithm of this theorem cannot work in polynomial time uniformly on k even when evaluating the polynomials $x_1 + \cdots + x_k$.

Computing the roots of algebraic polynomials

We consider equations of the form

$$t_e(\alpha_1, \dots, \alpha_k)x^e + \dots + t_1(\alpha_1, \dots, \alpha_k)x + t_0(\alpha_1, \dots, \alpha_k) = 0, \quad (3)$$

where $\alpha_1, \dots, \alpha_k \in \mathcal{C}_{\text{alg}}$ and $t_j(\bar{x}) \in \mathbb{Q}[x_1, \dots, x_k]$. The problem is to find a list of (codes of) all roots from given $b(\alpha_1) \& \dots \& b(\alpha_k)$ and $b(t_0(\bar{x})) \& \dots \& b(t_e(\bar{x}))$. The form (1) is convenient since our algorithm remains polynomial for fixed k even if e grows.

T h e o r e m. There exists an algorithm which, given $k \geq 1$, $\alpha_1, \dots, \alpha_k \in \mathcal{C}_{\text{alg}}$ and polynomials $t_0(\bar{x}), \dots, t_e(\bar{x}) \in \mathbb{Q}[x_1, \dots, x_k]$, finds a list $\beta_1, \dots, \beta_g \in \mathcal{C}_{\text{alg}}$ of all complex roots of (1).

The working time of the algorithm is estimated as $(n_1 n_2 \dots n_k)^c L^d$, or $n^{ck} L^d$, where c, d are constant. In particular, if k is fixed or $n = 1$, we get a p-time algorithm for root-finding.

Furthermore, $\deg[\beta_j] \leq e \prod_{i \leq k} \deg[\alpha_i]$ for $j \leq g$.

In a recent joint work with Svetlana Selivanova we investigated **complexity bounds** for computing the solution operator of the Cauchy problem (1).

In this investigation we rely not on computable analysis, but rather on the ideas of **guaranteed precision**. From several approaches to measure the complexity of computation, we choose the classical computational complexity often referred to as **bit complexity**.

Problem Statement

- Informally:** 1) from given $A, B_1, \dots, B_m, \varphi, f$ and precision $\varepsilon = \frac{1}{a}$ (where a is a positive integer), **find** approximation to H (domain of existence and uniqueness) and \mathbf{u} (the precise solution of the Cauchy problem (1)).
- 2) Estimate the **computation time** needed to achieve the prescribed precision.

From various possible specifications of input data and parameters we stick to the following particular case:

$$f = 0;$$

A, B_i ($i = 1, 2, \dots, m$) are rational (or real algebraic) $n \times n$ matrices such that $A = A^* > 0$, $B_i = B_i^*$,

φ_j ($j = 1, 2, \dots, n$) are rational polynomials.

The matrices and polynomials are encoded in a standard way by binary words.

Now we state a guaranteed-precision version of a restricted Cauchy problem.

Let $m, n \geq 2$ be fixed positive integers. We search for an algorithm (and its complexity estimation) which, for any given matrices

$A, B_1 \dots, B_m \in M_n(\mathbb{Q})$, polynomials $\varphi_1 \dots, \varphi_n \in \mathbb{Q}[x_1 \dots, x_m]$, and precision $a \geq 1$ computes a rational $T > 0$ s.th.

$H \subseteq [0, T] \times \mathbb{Q}$, a spatial rational grid step h dividing 1, a time grid step τ dividing T and a rational h, τ -grid function $v : G \rightarrow \mathbb{Q}^n$ such that $\|\mathbf{u} - \tilde{v}|_H\|_{sL_2} < \varepsilon$ where $\varepsilon = \frac{1}{a}$ and \tilde{v} is the multilinear interpolation of v . We abbreviate this problem as $\text{CP}(m, n, \mathbb{Q}, \mathbb{Q})$.

Main results: computation of H

The set H is known to be a nonempty intersection of

$$t \geq 0, \quad x_i - \lambda_{\max}^{(i)} t \geq 0, \quad x_i - 1 - \lambda_{\min}^{(i)} t \leq 0, \quad (i = 1, \dots, m)$$

of \mathbb{R}^{m+1} , where $\{\lambda_k^{(i)}\}_{k=1}^n$ are the eigenvalues of $A^{-1}B_i$. Assume $\lambda_{\min}^{(i)} < 0 < \lambda_{\max}^{(i)}$ for all $i = 1, \dots, m$.

T h e o r e m. Let $m, n \geq 2$ be any fixed integers. There is a polynomial time algorithm which for any given $A, B_1, \dots, B_m \in M_n(\mathbb{A})$ finds the vector $(\lambda_{\max}^{(1)}, \dots, \lambda_{\max}^{(m)}, \lambda_{\min}^{(1)}, \dots, \lambda_{\min}^{(m)})$ and checks the condition $\lambda_{\min}^{(i)} < 0 < \lambda_{\max}^{(i)}$ for all $i = 1, \dots, m$. Thus, the algorithm finds the domain H satisfying the condition above, or reports on the absence of such a domain.

Main results: computation of \mathbf{u}

T h e o r e m. For any fixed $m, n \geq 2$ the problem $\text{CP}(m, n, \mathbb{A}, \mathbb{Q})$ is solvable in EXPTIME.

Remark: increasing any of the parameters m, n makes the computation time double exponential.

Hints to the proof

- The proof heavily relies on: - the Godunov's difference scheme and deep results on its convergence (due to Godunov and his co-authors);
 - proofs of the existence and uniqueness theorems for (1);
 - properties of multilinear interpolations.
- The proof heavily relies on deep results of computer algebra for polynomial arithmetic and computations in the fields of algebraic reals due to Loos, Collins, Grigoriev etc. and those recently considered by Alaev and Selivanov.
- The proof essentially uses polynomial-time computability (in some fields of algebraic reals) of finding eigenvectors of matrix pencils (recall that this problem is not computable in the field of reals). In particular, this is crucial for finding in polynomial time steps h, τ guaranteeing the stability of the Godunov's scheme.
- Our proof is a mix of methods typical for symbolic and numerical computations.

THANK YOU FOR YOUR ATTENTION!!