Constructive real algebra

M.A.P. Meeting - Leiden - January 12th 2007

H. Lombardi, Besançon

Henri.Lombardi@univ-fcomte.fr, http://hlombardi.free.fr

Printable version of these slides:

http://hlombardi.free.fr/publis/AlrecoDocMAP.pdf

A more developped version

http://hlombardi.free.fr/publis/AlrecoProgram.pdf

Investigating the algebraic properties of real numbers

i.e., properties of real numbers

w.r.t. $+$, $-$, $\times$, $>$, $\geq$

Summary:

Many questions, few answers.

Why studying constructive real algebra?

Constructive real algebra is not well understood!
Constructive analysis is much more developped.

From a constructive point of view, real algebra is far away from the theory of discrete real closed fields (which was settled by Artin in order to understand real algebra in the framework of classical logic).

Most algorithms for discrete real closed fields fail for real numbers, because we have no sign test for real numbers.

Within constructive analysis, it should be interesting to drop dependent choice. A study of real agebra without dependent choice could help.

Why studying constructive real algebra?

Understanding real algebra should be a first important step for obtaining a constructive version of O-minimal structures.

Real algebra can be seen instead as the simplest O-minimal structure. Indeed classical O-minimal structures give effectiveness results inside classical mathematics, but they are not completely effective, because the sign test on real numbers is needed for the corresponding "algorithms".

Heyting fields . . . without negation

$$(\mathbf{K}, = 0, \neq 0, +, -, \times, 0, 1)$$

- $x = y$ means $x - y = 0$
- $x \neq 0$ means x invertible
- $x \neq y$ means $x - y \neq 0$

Axioms:

F1 $(\mathbf{K}, = 0, +, -, \times, 0, 1)$ is a commutative ring.
I.e., computational machinery of commutative rings, plus direct axioms:

$$\vdash 0 = 0, \quad x = 0 \vdash xy = 0, \quad x = 0,\, y = 0 \vdash x + y = 0.$$

F2 $x^2 = 0 \vdash x = 0$ (simplification axiom)

F3 $x + y \neq 0 \vdash x \neq 0 \vee y \neq 0$ (dynamical axiom)

F4 $(x \neq 0 \Rightarrow 1 = 0) \vdash x = 0$ (complicated axiom)

Heyting fields

In other words, an Heyting field is a local ring whose Jacobson radical is reduced to 0.
Examples: $\mathbb{R}$, $\mathbb{C}$, $\mathbb{R}(t)$, $\mathbb{C}(t)$, primitive recursive real numbers, ...

Remark : Axiom **F4**

F4 $\quad (x \neq 0 \ \Rightarrow \ 1 = 0) \ \vdash \ x = 0$

means that the local ring (axiom **F3**) has its Jacobson radical equal to 0. It is a very unpleasant axiom. This can be seen as a weakened form of the TEM axiom **DF** for discrete fields.

DF $\quad x = 0 \ \vee \ x \neq 0$

Note that **F4** and **DF** are formulations without negation: the trivial ring is allowed to be a discrete field.

What is an algebraically closed Heyting field?

"An homogenous bivariate polynomial with at least one invertible coefficient splits into linear factors"?

Richman version without dependent choice? How to formalize it in algebra?

Considering only separable polynomials?

Simultaneous collapses for commutative rings and fields

A commutative ring which collapses as a dynamic discrete field collapses.

A dynamic discrete field which collapses as an algebraically closed discrete field collapses.

As a particular case this allows a dynamic version of the algebraic closure of an Heyting field. It should be better if we were able to construct an Heyting field containing $\mathbb{C}(t)$ where separable polynomials split into linear factors.

For simultaneus collapses, see:

http://hlombardi.free.fr/publis/NullstellensatzDynamic.pdf

(*Dynamical method in algebra: Effective Nullstellensätze*. Coste M., L. H., Roy M.-F. Annals of Pure and Applied Logic **111**, (2001) 203-256)

Ordered Heyting fields

$$(\mathbf{K}, =0, \neq 0, >0, \geq 0, +, -, \times, \sup, 0, 1)$$

- $x = y$ means $x - y = 0$
- $x > y$ means $x - y > 0$
- $x \neq y$ means $x - y \neq 0$
- $x \leq y$ means $x - y \leq 0$

Direct rules

1. $(\mathbf{K}, =0, +, -, \times, 0, 1)$ is a commutative ring.
2. $\vdash\ 1 > 0$
3. $x = 0\ \vdash\ x \geq 0$
4. $x > 0\ \vdash\ x \geq 0$
5. $\vdash\ x^2 \geq 0$
6. $(x > 0,\ y \geq 0)\ \vdash\ x + y > 0$
7. $(x > 0,\ y > 0)\ \vdash\ xy > 0$
8. $(x \geq 0,\ y \geq 0)\ \vdash\ x + y \geq 0$
9. $(x \geq 0,\ y \geq 0)\ \vdash\ xy \geq 0$

Collapsus axiom

10. $0 > 0\ \vdash\ 1 = 0$

Ordered Heyting fields

Simplification rules

11. $x^2 \leq 0 \vdash x = 0$
12. $(c \geq 0, cs > 0) \vdash s > 0$
13. $(s > 0, cs \geq 0) \vdash c \geq 0$
14. $(c \geq 0,\ x(x^2 + c) \geq 0) \vdash x \geq 0$

Dynamic rules

15. $x + y > 0 \vdash x > 0 \vee y > 0$
16. $xy > 0 \vdash x > 0 \vee -y > 0$
17. $x^2 > 0 \vdash \exists y\, xy = 1$

Discrete ordered fields

DOF $\vdash x \geq 0 \vee -x > 0$

Heyting ordered fields

HOF $(x > 0 \Rightarrow 1 = 0) \vdash x \leq 0$

Simultaneous collapsus and provable facts

Theorem 1. *Let* $\mathbf{A}$ *be a commutative ring. Let* Z, P, S *be three subsets of* $\mathbf{A}$*. Consider the* "dynamical preordered ring" *defined by these data (i.e., let* $x = 0$ *for* $x \in Z$*,* $x \geq 0$ *for* $x \in P$*,* $x > 0$ *for* $x \in S$*). Then the collapsus occurs simultaneously in the following cases:*

a) Use only direct rules.

b) Use direct rules and simplification rules.

c) Use direct rules, dynamic rules and **DOF** *(simplification rules follow).*

d) Add a real closure rule: a monic polynomial whose sign changes between a *and* b *has a root on* (a, b)

Moreover the dynamical structures b), c) and d) prove the same facts.

Simultaneous collapsus and provable facts, 2.

So adding **DOF** as an axiom in an ordered Heyting field does not change facts, and does not produce a collapsus.

Samething with real closure rules.

Feel free of using **DOF** and real closure axioms in an ordered Heyting field if you have only to prove a fact.

Problem with the function sup

sup is a well defined function on $\mathbb{R}^2$, but the previously given theory of ordered Heyting fields does **not** prove the existence of a sup z for any x, y, i.e., the following statement is not provable

$$\forall x, y\, \exists z \quad (z - x)(z - y) = 0,\ z \geq x,\ z \geq y$$

So the theory has to be improved by adding a symbol for the function sup with the following axioms.

Rules for sup

18. $\vdash \sup(x, y) = \sup(y, x)$

19. $\vdash \sup(x, y) \geq x$

20. $\vdash (\sup(x, y) - x)(\sup(x, y) - y) = 0$

Properties of sup

Define $\inf(a,b) = -\sup(-a,-b)$.

- $\sup(x+z, y+z) = \sup(x,y) + z$
- $\sup(x,y) + \inf(x,y) = x + y$
- $\sup(x,y)\inf(x,y) = xy$
- $\sup(x,y) > 0 \iff (x > 0 \vee y > 0)$
- $x = \sup(x,y) \iff x \geq y$
- $\sup(x,y) < 0 \iff (x < 0 \wedge y < 0)$
- $\sup(x,y) \leq 0 \iff (x \leq 0 \wedge y \leq 0)$

Remark : The two sets $\{a,b\}$ and $\{\inf(a,b), \sup(a,b)\}$ have the same adherence, which is the set of roots of $(T-a)(T-b)$.
Similar things with $(T-a_1)\cdots(T-a_n)$.

Some nonprovable properties in ordered Heyting fields

$$x = 0 \vee x \neq 0$$
$$\forall x \; \exists y \;\; x^2 y = x$$
$$xy = 0 \vdash (x = 0 \vee y = 0)$$
$$x \geq 0 \vee x \leq 0$$
$$\sup(x, y) = x \vee \sup(x, y) = y$$
$$(x \leq 0 \Rightarrow 1 = 0) \vdash x > 0$$

For the (Bishop) real number field,

- the two first assertions are equivalent to **LPO**,
- the three following ones to **LLPO**,
- and the last one to **MP**.

What exactly is available?

Is "real linear algebra"
correctly described by our axioms?
If not, what is missing?

Same questions with the real linear programming.

Other "rational" problems

E.g., $$\frac{(ax+by)xy}{x^2+y^2}$$

The above rational function is the prototype of a family (with parameters a, b) of continuous functions definable on $\mathbb{R}^2$ in a rational way.

Nevertheless it seems that the existential statement

$(*)$ $$\forall a, b, x, y \ \exists z \quad z(x^2+y^2) = (ax+by)xy$$

is not provable with our axiomatisation of Heyting ordered fields.

So we have to add axioms as $(*)$, or, better, symbols of functions, each time we have a continuous function which is definable from an element of $\mathbb{Q}(X_1, \ldots, X_n)$.

Other continuous "rational" maps

Related question

Is it the case that every continuous function defined by an element of $\mathbb{R}(X_1, \ldots, X_n)$ is a real point in a continuous family defined over $\mathbb{Q}(X_1, \ldots, X_n)$?

Semialgebraic sets

A *semipolynomial*, or sup-inf-polinomially-defined (SIPD) function is given by a term in the language $(\mathbf{K}, +, -, \times, \sup, 0, 1)$ (with $\mathbb{Q} \subseteq \mathbf{K}$ if $\neg(1 = 0)$)

A *closed* (resp. *open*) semialgebraic set in $\mathbf{R}^n$ defined over $\mathbf{K}$, where $\mathbf{R}$ is an ordered field containing $\mathbf{K}$ is a set $\{\, x \in \mathbf{R}^n \,|\, h(x) \geq 0 \,\}$ (resp. $\{\, x \in \mathbf{R}^n \,|\, h(x) > 0 \,\}$) where h is an SIPD in n variables over $\mathbf{K}$. "Union" and intersection correspond to sup and inf.
A *locally closed* semialgebraic set in $\mathbf{R}^n$ defined over $\mathbf{K}$ is the intersection of a closed and an open semialgebraic sets in $\mathbf{R}^n$ defined over $\mathbf{K}$.

It seems better to avoid "other" semialgebraic sets such as

$$\Big\{(x, y) \in \mathbf{R}^2 \,|\, x \neq 0 \vee x = y = 0\Big\},$$

where the "$\vee$" leads to many problems.

Real closure properties

Recall the real closure axiom in a discrete setting.

RCF1: A univariate polynomial P such that $P(a) < 0$, $P(b) > 0$, $a < b$ has a zero on (a, b).

RCF1 is not available for real numbers without dependent choice. The following one is constructively valid:

RCF2: A univariate polynomial P such that $P(a) < 0$, $P(b) > 0$, $a < b$ and $P' > 0$ on (a, b) has a zero on (a, b).

But it is not sufficient. We will need virtual roots:

(see: *Virtual roots of real polynomials.* Gonzalez-Vega L., L. H., Mahé L. Journal of Pure and Applied Algebra **124**, (1998) 147–166.

Coste M., Lajous T., L. H., Roy M.-F. *Generalized Budan-Fourier theorem and virtual roots*. Journal of Complexity **21** (2005), 479–486.

http://hlombardi.free.fr/publis/AVirtualRealRoots.html)

Virtual real roots

Lemma 2. *A continuous increasing (resp. decreasing) function f on $[a, b] \subseteq \mathbb{R}$ $(a \leq b)$ attains its (unique) minimum absolute value.*

Corollary 3. *One can define on the set of real univariate polynomials of (well defined) degree d, d* virtual root functions *$\rho_{d,k}$ $(k = 1, \ldots, d)$ with the following characteristic properties,*

$$\begin{aligned} f(\rho_{1,1}(f)) = 0 \qquad & \text{if} \quad d = 1 \\ \rho_{d-1,k-1}(f') \leq \rho_{d,k}(f) \leq \rho_{d-1,k}(f') \qquad & \text{if} \quad d \geq 2 \\ \left|f(\rho_{d,k}(f))\right| \leq |f(x)| \qquad & \text{if} \quad \rho_{d-1,k-1}(f') \leq x \leq \rho_{d-1,k}(f') \end{aligned}$$

(with the convention $f(\rho_{d,0}(f)) = \varepsilon(-1)^d \infty$, $f(\rho_{d,d+1}(f)) = \varepsilon \infty$, where $\varepsilon = \pm 1$ is the sign of the leading coefficient)

Virtual roots, 2.

1. If $f(T) = (T-a)(T-b)$ then
$$\rho_{2,1}(f) = \inf(a,b),\ \rho_{2,2}(f) = \sup(a,b).$$

2. If $\deg(f) = d$ and $f(x) = 0$ then $\prod_{i=1}^{d}(x - \rho_{d,i}(f)) = 0$.

3. Constructive version of **RCF1**:
if $\deg(f) = d$, $a < b$ and $f(a)f(b) < 0$ then
$$\prod_{i=1}^{d} f(\mu_{d,i}(f)) = 0,$$
where $\mu_{d,i}(f) = \inf(b, \sup(a, \rho_{d,i}(f)))$. This implies **RCF2**.

4. Each $\rho_{d,i}(f)$ is a locally uniformly continuous function, and is a zero of the product
$$\prod_{k=0}^{d-1} f^{(k)}(T).$$

A result à la Pierce-Birkhoff

An interesting result concerning virtual roots is the following one:

Theorem 4.
Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous semialgebraic function defined over $\mathbb{Q}$ which is integral over the ring $\mathbb{Q}[X_1, \ldots, X_n]$. Then f is a combination of virtual root functions and polynomials defined over $\mathbb{Q}$.

Remark : In the previous theorem, it is possible to replace $\mathbb{Q}$ by a discrete subfield of $\mathbb{R}$.

Remark : Is it possible to replace $\mathbb{Q}$ by $\mathbb{R}$? (the exact meaning of the hypothesis becomes not so clear). We should need a good definition for: "$f : \mathbb{R}^n \to \mathbb{R}$ is a continuous semialgebraic function."!

A plausible definition

Definition 5. *A* real closed field *is given when you have an (Heyting) ordered field with virtual root functions in each degree satisfying the characteristic properties given in the real number field case.*

(We may use only virtual root functions of monic polynomials.)

Examples of nondiscrete real closed subfields of $\mathbb{R}$ in this meaning.

- Primitive recursive real numbers.
- Polytime computable real numbers.
- Turing computable real numbers.

Related questions

Construction of the real closure of an ordered field

Other closure properties

Projection Theorem

Constructive Positivstellensätze

Construction of the real closure of an (Heyting) ordered field

This seems not problematic.

Add the virtual root functions as (formal) operators. Apply the axioms. From the simultaneous collapsus theorem, no collapsus can occur.

So no catastrophe. But it is not sufficient.

E.g., if an axiom gives a conclusion which is a disjunction, how can we find a good branch (this is stronger than: open two branches, if one branch collapses the other is good).

The solution comes from the fact that the real closure of a discrete ordered field is **strongly unique** (and the virtual roots are uniquely defined by their defining axioms).

Probably this works, but we need a more precise argument, giving clearly an algorithm.

Construction of the real closure, 2.

Does this show the possibility to add a positive infinitesimal ε to $\mathbb{R}$ and to construct the real closure?

No. But the obstacle does not come from the real closure.

The classical object $\mathbb{R}(\varepsilon)$ is *not* an ordered Heyting field.
It is a noncollapsing dynamic ordered discrete field.

Question : giving a structure or ordered Heyting field over $\mathbb{R}(X)$ is impossible in a constructive way?

Fundamental Theorem of Algebra

One can prove constructive versions of the **FTA** (in $\mathbf{R} + i\mathbf{R}$) for a real closed field $\mathbf{R}$ in the above meaning (i.e., with virtual root functions symbols and axioms).

The first one is: every monic separable polynomial splits into linear factors.

Probably this can be deduced from the second "continuous" version.

Remark : It should be interesting to find a good setting for the Richman version, which uses the space of d-multisets of complex numbers.

Fundamental Theorem of Algebra

A second one is a continuous version, giving a version "without dependent choice" for $\mathbb{C}$.

In degree d the real parts of the d complex roots, enumerated in increasing order, are continuous "integral" semialgebraic functions of the coefficients. Same thing for the imaginary parts. So we can define d^2 continuous functions that "cover the complex roots", $\theta_{d,i}(f)$ $(1 \leq i \leq d^2)$, with the following meaning:

- $f(z) = 0 \vdash \prod_{i=1}^{d^2}(z - \theta_{d,i}(f)) = 0$
- for any $J \subseteq \left\{1, \ldots, d^2\right\}$ of cardinality $d^2 - d + 1$, $\prod_{i \in J} f(\theta_{d,i}(f)) = 0$.

Other continuous semialgebraic functions

1. Distance map to a located closed semialgebraic set?

2. Projection map on a located closed semialgebraic convex set?

Other continuous semialgebraic functions, 2.

Distance map to a located closed semialgebraic set?

It seems that a located closed semialgebraic set $S \subseteq \mathbb{R}^n$ appears always as a "real point" $S(\alpha)$ in a family $S(a)$ $(a \in J \subseteq \mathbb{R}^k)$ defined over $\mathbb{Q}$, the distance function $\varphi = d(x, S(a))$ being a continuous semialgebraic function of $(x, a) \subseteq \mathbb{R}^n \times J$.
Here φ is defined over $\mathbb{Q}$, J is locally closed.

Projection map on a located closed semialgebraic convex set?
Same thing?

So we are lead to the following general context.

Other continuous semialgebraic functions, 3.

Let $\mathbf{R}$ be a discrete real closed subfield of $\mathbb{R}$.

A semialgebraic continuous function $S \subseteq \mathbf{R}^n \to \mathbf{R}$ defined over $\mathbb{Q}$, having as domain a $\mathbb{Q}$-semialgebraic locally closed set S, has a natural extension to $\mathbb{R}$, since it is uniformly continuous on each compact, for the natural topology of locally compact metric space of the domain.

Do these extensions can be expressed using only virtual root functions? (we allow taking the inverse of an everywhere positive function).

If it is not the case, we need a better definition for real closed fields.

The projection theorem

Let $\mathbf{K}$ be a subfield of a discrete real closed field $\mathbf{R}$, $S \subseteq \mathbf{R}^n$ a semialgebraic set defined over the subfield $\mathbf{K}$ and $\pi_n = \mathbf{R}^n \to \mathbf{R}^{n-1}$: $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_{n-1})$. The projection theorem says that $\pi_n(S)$ is a semialgebraic set defined over $\mathbf{K}$.

We need good constructive versions when $\mathbf{R}$ is replaced by $\mathbb{R}$.

The following weakened version is likely to be true.

In the nondiscrete case let us call "compact semialgebraic subset of $\mathbf{R}^n$" a located closed bounded semialgebraic set.

Theorem 6. *(we hope)*
If S is a compact semialgebraic subset of $\mathbb{R}^n$ then so is $\pi_n(S)$.

Projection Theorem, 2.

If Theorem 6 is true, we expect that it will be true for "Heyting real closed fields". Perhaps this would force us to add new axioms in the definition of real closed fields.

Constructive Positivstellensätze

Let us recall that in the case of a discrete real closed field, the constructive Positivstellensatz follows directly from the simulatneaous collapsus theorem, and from the fact that the formal theory is complete.

The simultaneous collapsus theorem says us how to transform a simple (i.e., dynamical) proof of impossibility in the real closure into an algebraic identity.

Moreover a "cut elimination theorem" shows how to transform a first order proof into a dynamical one.

Constructive Positivstellensätze, 2.

All this remains true in the nondiscrete context.
If you find a proof of the impossibility of a system of conditions on signs (on polynomials) in $\mathbf{R}^n$ by using our axiomatisation of real closed fields, you will get a corresponding Positivstellensatz.

Moreover, since our theory is weaker than the discrete one, a proof is more informative and has to give a better form of Positivstellensatz, where the dependence of the algebraic identity w.r.t. the coefficients is best controlled (this dependence must have some continuity properties).

On the other side the formal theory is no more complete and there is no more a systematic way of testing the compatibility of a system of signs conditions.

Constructive Positivstellensätze, 3.

Such kind of continuity results have been obtain by C. Delzell and other authors for the 17-th Hilbert problem (and for other variants of Positivstellensätze), in a discrete context.

In the following paper, you find a rather complete bibliography on the subject and a discussion about the consequences of the results for the Bishop real number field.

A Real Nullstellensatz and Positivstellensatz for the Semipolynomials over an Ordered Field. Gonzalez-Vega L., L. H., Journal of Pure and Applied Algebra **90**, (1993) 167–188.

http://hlombardi.free.fr/publis/PstSemiPols.pdf