Infinite objects in constructive mathematics

Thierry Coquand

Munich, December 14, 2007

Introduction

The general theme will be the connections between

reasoning and *computation*

in mathematics, mainly here abstract commutative algebra.

I will survey some recent works in constructive mathematics which bring a new viewpoint on this question

Some history

The word *algebra* comes from the title of a book *Hibab al-jabr wal-muqubala* (around 825)

The word *algorithm* comes from the name of the author of this book *Al-Khwarizmi*

Until 1800 most works in algebra are mostly computations (like in computer algebra)

Example: elimination theory (Poisson), Lagrange

Some history

The situation changes with Gauss, Abel, Galois

concept of *irreducible* polynomial: Gauss (cyclotomic polynomial), fundamental notion for Abel and Galois

Construction of the splitting field of a polynomial

Rational functions of given quantities (which will become *domain of rationality* for Kronecker, and later our notion of field)

Some history

All the proofs have still a direct algorithmic interpretation

Galois insists on the *ideal* character of these computations

"*If now, you give me an equation that you have in any way you like and you want to know whether it is or not solvable by radicals, I have nothing to do but to indicate to you the way to reply to the question, but whithout to obliging either myself or anyone else to do so. In other word, the calculations are impracticable.*"

Same for Kronecker. The connection with computations is however essential

Some history

The connection between reasoning and algorithms became then less and less clear, typically through the different versions that Dedekind will give to his theory of ideals

Hilbert: all ideals of $K[X_1, \ldots, X_n]$ are of finite type

Noetherian: all ideals are of finite type

This proposition has *no* computational content, and it is logically complex (technically it cannot be expressed in first-order logic)

Some history

Example: any polynomial P of degree $\geqslant 1$ in $K[X]$ has an irreducible factor, given a field K

If P is not irreducible, $P = QR$ with $1 \leqslant d(Q) < d(R)$ and we can find an irreducible factor of Q by induction. This looks like an algorithm but even if K is concretely given and computable it can be shown that there is *no* algorithm for finding an irreducible factor in general

The property: "to be irreducible" is not decidable in general (in some special cases it is)

The connection with computation, maybe unfeasible but which was always possible in theory, even for Dedekind, is *lost*

Some history

Abstract commutative algebra has become one of the "less computational" part of mathematics (if one looks at the proofs). One of the rare part which uses the general form of the Axiom of Choice

Theorem: (Krull) *An element is nilpotent iff it belongs to all prime ideals*

Theorem: *Any field has an algebraic closure, unique up to isomorphism*

If we prove in commutative algebra the existence of an object satisfying a simple "concrete" property, it is not clear if this proof gives a way to compute this object

Example: Kronecker's theorem

Theorem: *An algebraic variety in* $\mathbb{C}^n$ *is the intersection of* $n+1$ *hypersurfaces*

Given (let say) $n = 3$ and 5 given polynomials with 3 variables, can one use an abstract proof of this theorem and compute 4 polynomials that define the same variety?

Hilbert's Program

This was one issue raised during the debate between Hilbert and Brouwer

Hilbert's program: if we prove using ideal methods a *concrete statement*, one can always eliminate the use of these ideal elements and obtain a purely elementary proof

Ideal methods: use of prime ideals, maximal ideals, valuation rings, local-global principle, non constructive reasoning, ...

Hilbert: existence = logical consistency

Hilbert's Program

What is constructive mathematics?

Mathematics developped using intuitionistic logic: Richman, Lombardi, ...

(No need to be explicit about algorithms)

Proof theory: completeness theorem for first-order logic and cut-elimination results

Hilbert's Program

Recent work in constructive mathematics shows that Hilbert's program works for a large part of abstract algebra providing a constructive explanation of some abstract methods used in mathematics

Furthermore this follows Hilbert's idea of replacing an "infinite ideal object" by a syntactical theory that describes it

Zariski spectrum

Fundamental object in abstract algebra, usually defined as a set of prime ideals of a ring R with the basic open

$$D(a) = \{\mathfrak{p} \mid a \notin \mathfrak{p}\}$$

This is a *spectral space*: the compact open form a distributive lattice. They are exactly the finite union $D(a_1) \vee \cdots \vee D(a_n)$

However, even if the ring R is given concretely (discrete) it may be difficult to show effectively the existence of *one* prime ideal

Often, what matters is not *one* particular prime ideals, but the collection of *all* prime ideals.

Zariski spectrum

Zariski spectrum is best seen as a point-free space (cf. Menger, 1940, de Bruijn 1967)

A. Joyal (1972) definition of the Zariski spectrum

We consider the distributive lattice defined by the generators $D(a),\ a \in R$ (seen as formal symbols) and the relations

$$D(0) = 0 \qquad D(1) = 1 \qquad D(ab) = D(a) \wedge D(b) \qquad D(a+b) \leqslant D(a) \vee D(b)$$

Zariski spectrum

We have $D(a^2) = D(a)$ and $D(a^n) = D(a)$ if $n \geqslant 1$

All elements can be written on the form

$$D(a_1, \ldots, a_n) = D(a_1) \vee \cdots \vee D(a_n)$$

We have $D(a, b) = D(a + b, ab)$

If $D(ab) = 0$ then $D(a + b) = D(a, b)$

(Intuitively $D(f)$ is the "open set" over which the function f is $\neq 0$)

$D(a) \leqslant D(b_1, \ldots, b_m)$ if a is in the radical of the ideal generated by $b_1, \ldots, b_m$

Zariski spectrum

Theorem: $D(a_1) \wedge \cdots \wedge D(a_n) \leqslant D(b_1, \ldots, b_m)$ *holds iff the product* $a_1 \ldots a_n$ *is in the radical of the ideal generated by* $b_1, \ldots, b_m$

This is also known as the *formal* version of the Nullstellensatz. This can be seen as a *cut-elimination* result: any proof can be reduced to a direct proof

If R polynomial ring over $\mathbb{Q}$, $D(p)$ can be thought of as the complement of the set of zeros of p (in some algebraic closure). But following Kronecker we see $D(p)$ as a pure symbol.

Zariski spectrum

This definition is purely algebraic: we manipulate only rings and lattices, $R \longmapsto \mathrm{Zar}(R)$ is a functorial construction

Even if R is discrete (we have an algorithm to decide the equality in R), the lattice $\mathrm{Zar}(R)$ does not need to be discrete

Counter-example with Kripke model: $\mathbb{Z} \to \mathbb{Z}[1/2]$ is injective but $\mathrm{Zar}(\mathbb{Z}) \to \mathrm{Zar}(\mathbb{Z}[1/2])$ is not

Krull dimension of a ring

The *Krull dimension* of a ring is defined to be the maximal length of proper chain of prime ideals.

In fact, one can give a purely algebraic definition of the Krull dimension of a ring

Inductive definition of dimension of spectral spaces/distributive lattice: Kdim $X \leqslant n$ iff for any compact open U we have Kdim $Bd(U) < n$ (cf. Menger-Urysohn definition of dimension)

To be zero-dimensional is to be a Boolean lattice

Krull dimension of a lattice

If L is a lattice, we say that $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ are *(n-)complementary* iff

$$u_1 \vee v_1 = 1,\ u_1 \wedge v_1 \leqslant u_2 \vee v_2, \ldots, u_{n-1} \wedge v_{n-1} \leqslant u_n \vee v_n,\ u_n \wedge v_n = 0$$

For $n = 1$: we get that u_1 and v_1 are complementary

Proposition: Kdim $L < n$ *iff any n-sequence of elements has a complementary sequence*

Krull dimension of a ring

Kdim $R < n$ is defined as Kdim $(\mathrm{Zar}(R)) < n$

Proposition: Kdim $R < n$ *iff for any sequence* $a_1, \ldots, a_n$ in R *there exists a sequence* $b_1, \ldots, b_n$ in R *such that, in* Zar(R)*, we have*

$$D(a_1, b_1) = 1,\ D(a_1 b_1) \leqslant D(a_2, b_2), \ldots, D(a_{n-1} b_{n-1}) \leqslant D(a_n, b_n),\ D(a_n b_n) = 0$$

This is a *first-order* condition in the multi-sorted language of rings and lattices

Example: Kronecker's theorem

Kronecker in section 10 of

Grundzüge einer arithmetischen Theorie der algebraischen Grössen.
J. reine angew. Math. 92, 1-123 (1882)

proves a theorem which is now stated in the following way

An algebraic variety in $\mathbb{C}^n$ is the intersection of $n+1$ hypersurfaces

Kronecker's Theorem

Theorem: *If* Kdim $R < n$ *then for any* $b_0, b_1, \ldots, b_n$ *there exist* $a_1, \ldots, a_n$ *such that* $D(b_0, \ldots, b_n) = D(a_1, \ldots, a_n)$

This is a (non Noetherian) generalisation of Kronecker's Theorem

For each fixed n this is a first-order tautology. So, by the completeness Theorem for first-order logic, it has a first-order proof

Kronecker's Theorem

This concrete proof/algorithm, is *extracted* from R. Heitmann "*Generating non-Noetherian modules efficiently*" Michigan Math. J. 31 (1984), 167-180

Though seeemingly unfeasible (use of prime ideals, topological arguments on the Zariski spectrum) this paper contains implicitely a clever and simple algorithm which can be instantiated for polynomial rings

Forster's Theorem

We say that a sequence $s_1, \ldots, s_l$ of elements of a commutative ring R is *unimodular* iff $D(s_1, \ldots, s_l) = 1$ iff $R = <s_1, \ldots, s_l>$

If M is a matrix over R we let $\Delta_n(M)$ be the ideal generated by all the $n \times n$ minors of M

Theorem: *Let M be a matrix over a commutative ring R. If $\Delta_n(M) = 1$ and* Kdim $R < n$ *then there exists an unimodular combination of the column vectors of M*

This is a non Noetherian version of Forster's 1964 Theorem

Forster's Theorem

We get a first-order (constructive) proof.

It can be interpreted as an algorithm which produces the unimodular combination.

The motivation for this Theorem comes from differential geometry

If we have a vector bundle over a space of dimension d and all the fibers are of dimension r then we can find $d + r$ generators for the module of global sections

Forster-Swan's and Serre Splitting-Off Theorem

The same method applies for Forster-Swan's and Serre Splitting-Off Theorem that applies to the *maximal* spectrum

One can represent similarly the maximal spectrum in a first-order way

Space of valuation

Let L be a field, and R a subring of L

Another spectral space important in mathematics is the space $\mathrm{Val}(L, R)$ of *valuation rings* of L containing R

Such a ring is a subring $V \subseteq L$ containing R and such that if s in L and $s \neq 0$ then s is in V *or* $1/s$ is in V

We have always the solution $V = L$

Space of valuation

We define the lattice $\mathsf{Val}(L, R)$ as the universal solution of the problem $V_R : L \to \mathsf{Val}(L, R)$ with the conditions

$$V_R(r) = 1 \qquad (r \in R)$$

$$V_R(s_1) \wedge V_R(s_2) \leqslant V_R(s_1 s_2) \wedge V_R(s_1 + s_2)$$

$$1 = V_R(s) \vee V_R(1/s) \qquad (s \neq 0)$$

Space of valuation

In general we cannot simplify $V_R(s_1) \wedge \cdots \wedge V_R(s_l)$

$V_R(s) \wedge V_R(1/s) = V_R(s + s^{-1})$

$V_R((x + y)^{-1}) \leqslant V_R(1/x) \vee V_R(1/y)$

$1 = V_R(x^{-1}) \vee V_R((1 - x)^{-1})$

Space of valuation

Theorem: $V_R(t_1) \wedge \cdots \wedge V_R(t_n) \leqslant V_R(s_1) \vee \cdots \vee V_R(s_m)$ *holds iff we have an equality of the form* $1 = \Sigma 1/s_i P_i(t_j, 1/s_i)$

This is a *cut-elimination* Theorem, proved by *algebraic* elimination

This is proved by *algebraic* elimination of variables

Space of valuation

Special case: $1 = V_R(s/t_1) \vee \cdots \vee V_R(s/t_n)$ iff s is *integral* over the ideal I generated by $t_1, \ldots, t_n$ in $R[t_1, \ldots, t_n, s]$. This means that we have an equality

$$s^l = a_1 s^{l-1} + \cdots + a_l$$

where a_k is in I^k

Special case: $1 = V_R(s)$ iff $1/s$ is invertible in $R[1/s]$ iff s is integral over R

We get a constructive reading of the fact that the intersection of valuation rings containing R is the integral closure of R

Application: Dedekind Prague's Theorem

Theorem: *If* $(\Sigma a_i X^i)(\Sigma b_j X^j) = \Sigma c_k X^k$ *then each product* $a_i b_j$ *is* integral *over the coefficients* c_k

This generalises a famous result of Gauss: if all a_i, b_j are *rationals* and all c_k are *integers* then all products $a_i b_j$ are *integers*

This "may be considered as one of the most basic result in commutative algebra of the XIXth century ... It ended up as one exercise in Bourbaki, but here it is proved in a non constructive way" (Olaf Neumann)

This appears as an exercise in Algebra, Chapter 7 (Diviseurs)

Application: Dedekind Prague's Theorem

We get a proof-theoretic reading of the non constructive argument. We take $L = \mathbb{Q}(a_0, \ldots, a_n, b_0, \ldots, b_m),\ R = \mathbb{Q}$ and we prove

$$1 = V(a_i b_j / c_0) \vee \cdots \vee V(a_i b_j / c_m)$$

This corresponds to the non constructive argument: prove this for an *arbitrary* valuation

Application: Dedekind Prague's Theorem

For $n = m = 2$ a proof certificate of $1 = V(a_0b_1/c_0) \vee \cdots \vee V(a_0b_1/c_4)$ is

$$(a_0b_1)^6 = p_1(a_0b_1)^5 + p_2(a_0b_1)^4 + p_3(a_0b_1)^3 + p_4(a_0b_1)^2 + p_5(a_0b_1) + p_6$$

where

$$p_1 = 3c_1, \quad p_2 = -3c_1^2 - 2c_0c_2, \quad p_3 = c_1^3 + 4c_0c_1c_2$$
$$p_4 = -c_0^2c_1c_3 - 2c_0c_1^2c_2 - c_0^2c_2^2 + 4c_0^3c_4$$
$$p_5 = c_0^2c_1^2c_3 + c_0^2c_1c_2^2 - 4c_0^3c_1c_4$$
$$p_6 = -c_0^3c_1c_2c_3 + c_0^4c_3^2 + c_0^3c_1^2c_4$$

Application: Dedekind Prague's Theorem

Constructively $L \to \mathrm{Val}_R(L)$ is seen as a (clever) system of notations which records polynomial identities

Classically $\mathrm{Val}_R(L)$ is seen as a set of points

Zariski spectrum and space of valuations

Given any domain R of field of fractions L we have a lattice map

$\mathsf{Zar}(R) \to \mathsf{Val}(L, R), \qquad D(a) \longmapsto V(1/a)\ (a \neq 0)$

This is the *center map*. It is *always* injective.

The (constructive) proof of this fact requires cut-elimination

Intuitively: the function f is $\neq 0$ iff $1/f$ is finite

Prüfer domains

For Dedekind the crucial/fundamental property of Dedekind domains is that any finitely generated (non zero) ideals is invertible

This fundamental property is hard to recover from the (now) standard definition of Dedekind rings: integrally closed domain Noetherian and such that any non zero prime ideals is maximal

There is a first-order notion which captures most of the interesting computations: Prüfer domains

Prüfer domains

R Domain: if $ab = 0$ then $a = 0$ or $b = 0$

R Arithmetic: the lattice of ideals is distributive

$$\exists u\ v\ w.\quad au = bv \quad \wedge \quad b(1-u) = aw$$

This is a point-free description of: any localisation at any prime is a valuation domain

Prüfer domains

Proposition: *If R is a Prüfer domain the center map is a bijection. Conversely, if the center map is injective and R is integrally closed then R is a Prüfer domain*

We use the fact that if $au = bv$ and $b(1-u) = aw$ then we have

$$V(a/b) = V(1/u) \vee V(1/w)$$

The proof of the isomorphism can be done without using cut-elimination

Prüfer domains

Let R be a domain, L its fraction field

Proposition: *R is a Prüfer domain iff R is integrally closed and any s in L is root of a* primitive *polynomial*

Corollary: *If S is a Bezout domain and L is an extension of its field of fraction then the integral closure of R in L is a Prüfer ring*

Bezout domain: any finitely generated ideal is principal

In particular: $S = \mathbb{Z}$ (number theory) and $S = k[X]$ (algebraic curves)

These two results have simple constructive proofs (hence simple corresponding algorithms)

Prüfer domains

We have constructive proofs of the following results.

Proposition: *If R is integrally closed and* Kdim $R[X] \leqslant 2$ *then R is a Prüfer domain*

Proposition: *If R is integrally closed and* Kdim $(\mathrm{Val}(L, R)) \leqslant 1$ *then R is a Prüfer domain*

Structure sheaf

Any element of Zar(R) can be written $D(b_1, \ldots, b_m) = D(b_1) \vee \cdots \vee D(b_m)$

To simplify we assume that R is an integral domain

We define the *structure sheaf* $\mathcal{O}$ on Zar(R) by

$$\mathcal{O}(D(b_1, \ldots, b_m)) = R[1/b_1] \cap \cdots \cap R[1/b_m]$$

Classically, we have a continuous family of local rings $R_{\mathfrak{p}}$, and any element of R defines a global section

Local-global principle

$\mathcal{O}$ is a *sheaf*, called the structure sheaf

If in each $R_{\mathfrak{p}}$ the linear system $AX = B$ has a solution then it has a global solution

Constructively we have a covering $1 = D(s_1, \ldots, s_n)$

The sytem $AX = B$ has a solution in the ring $R[1/s_i]$

We find X_i, k_i such that $AX_i = s_i^{k_i} B$

We have $\Sigma u_i s_i^{k_i} = 1$ and so $X = \Sigma u_i X_i$ satisfies $AX = B$

Exactly like "partition of unity" in analysis

Towards point-free algebraic topology

Let L be an field algebraic extension of a field of $k(X)$

We define a sheaf on the (point-free) space $X = \mathsf{Val}(L, k)$ by

$\mathcal{O}_X(V(s_1) \wedge \cdots \wedge V(s_n)) = E(s_1, \ldots, s_n)$ integral closure of $k[s_1, \ldots, s_n]$ in L

$E(s)$ is a Prüfer domain so $V(s)$ is isomorphic to $\mathsf{Zar}(E(s))$ and this sheaf reduces to the structure sheaf of $V(s)$

We get a canonical example of *scheme* (glueing of two affine schemes)

Towards point-free algebraic topology

If u in L non zero then $V(u)$ and $V(u^{-1})$ covers the space X

One can show that $E(u, u^{-1})/E(u + u^{-1})$ (which is a k-vector space) is independent of u non zero element of L (non algebraic over L)

This defines $H^1(X, \mathcal{O}_X)$ as an *invariant* of the algebraic curve X

Towards point-free algebraic topology

For $L = k(X)$ one finds $H^1(X, \mathcal{O}_X) = 0$: any element integral over $k[X, X^{-1}]$ is a sum of an element integral over $k[X]$ and an element integral over $k[X^{-1}]$

For $L = k(x, y), \;\; 1 = x^2 + y^2 + x^2y^2$ one finds $H^1(X, \mathcal{O}_X) = k$: the element $u = y(1 + x^2)/x$ is integral over $k[x, x^{-1}]$ but cannot be written as the sum of an element integral over $k[x]$ and an element integral over $k[x^{-1}]$.