# Infinite objects in constructive mathematics

Thierry Coquand

July 31 2009

# Introduction

This talk will be about Hilbert's program and the connections between

*reasoning* and *computation*

in mathematics

The situation is especially interesting in *algebra*

# Some history

The word *algorithm* comes from the name of the author, *Al-Khwarizmi* of the treatise *Al-jabr wa-all-Muqabilah* (around 825)

The word *algebra* comes from the title of the book!

Until 1800 works in algebra consist mostly of computations and clever algebraic manipulations (like in computer algebra)

Example: elimination theory (Bezout, Poisson), Lagrange

# Some history

The situation changes with Gauss, Abel, Galois

Concept of *irreducible* polynomial in Gauss' work, which plays a fundamental role for Abel and Galois

Construction of the splitting field of a polynomial, cf. H. Edwards

*Essays in constructive mathematics*

Rational functions of given quantities which will evoluate later to the notion of field

# Some history

All the arguments that prove the existence of an object can still be interpreted algorithmically

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 without to obliging either myself or anyone else to do so. In other word, the calculations are impracticable.*"

Same for Kronecker (cf. H. Edwards). The connection with computations is however essential: we have to predict some informations about the results of the computation (without actually doing them)

# Some history

The connection between reasoning and algorithms became then less and less clear, typically through the different versions that Dedekind gave to his theory of ideals (cf. J. Avigad *Methodology and Metaphysics in the Development of Dedekind's Theory of Ideals*)

Cauchy, Puiseux, Riemann: introduction of complex analysis in the theory of algebraic functions (treated mainly as pure algebra by Abel, Galois, Kronecker)

The connection with computation, maybe not feasible but which was always possible in theory, is now *lost*, because of the use of the law of excluded-middle

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

# Reasoning and computations

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), even for $X^2 + 1$. (Simply take $\mathbb{Q}[\alpha i]$ where we don't know whether $\alpha = 0$ or $\alpha = 1$.)

# Existence in mathematics

*If we prove in commutative algebra the existence of an object, this proof may not give a way to compute this object. What is the meaning of mathematical existence then?* Hilbert thought deeply about this new situation from the 1890s on, as one can see from his Mathematical Notebooks

New meaning of mathematical existence (1890s; translation, S. Hayashi)

"To exist means that the conditions defining the concept do not contradict to themselves"

Replace "semantics" by "syntax"

# Hilbert's Program

Introduction/elimination of ideal elements

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

Ideal objects (non constructive): prime ideals, maximal ideals, valuation rings, local-global principle, non constructive reasoning, ...

These ideal objects are suggestive means for proofs with no real existence

# Hilbert's Program

Any non trivial ring has a maximal ideal

Let $R$ be a ring, if a linear map $R^2 \to R^3$ is *surjective* then $R$ is trivial (we have $1 = 0$ in $R$)

This can be formulated as the fact that we can derive $1 = 0$ in an equational theory

Proof: if $\mathfrak{m}$ is a maximal ideal of $R$ and $k = R/\mathfrak{m}$ then we have a surjective map $k^2 \to k^3$ contradiction

By Birkhoff's completness theorem for equational logic, there should be a purely equational derivation. Is it contained (hidden) in this non effective argument?

# Hilbert's Program

Serre's problem (Quillen-Suslin's Theorem)

**Theorem:** *Any finitely generated projective module on a polynomial ring is free*

**Theorem:** (concrete formulation) *An idempotent matrix over a polynomial ring is similar to a canonical projection matrix of the form* $I_{r,n} = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$

Given such a matrix $M$ satisfying $M^2 = I_n$ we can find an invertible matrix $P$ such that $PMP^{-1} = I_{r,n}$

The proof by Suslin uses a maximal ideal. Does this proof indicates a way to compute the matrix $P$ given $M$?

# Hilbert's Program

Examples: Dirichlet Theorem proved with analysis, or

**Theorem:** (Krivine, 1964) *If* $P \in \mathbb{Q}[x_1, \ldots, x_k]$ *is* $> 0$ *on* $[0,1]^n$ *then it can be written as a polynomial in* $x_i, 1 - x_i$ *with rational* positive *coefficients*

This is also proved with the Axiom of Choice

It is not true if $P$ is only $\geqslant 0$: take $(2x-1)^2$

(but it works for $(2x-1)^2 + \epsilon$ if $\epsilon > 0$)

Krivine provides two different proofs

# Formal topology

The method we use for representing Hilbert's idea of *introduction* and *elimination* of ideal objects comes from the theory of *locales* also known as *formal* or *point-free topology*

This presents a topological space, not as a set of points, but as a *logical theory* describing its lattice of open sets

Reverse of the traditional conceptual order in topology, where open sets are thought of as primitive symbolic objects (observable) and points are infinite ideal objects, defined as particular filters of neighborhoods

Typically, the existence of points is proved using Zorn's Lemma

# Elimination of points

Hilbert's *ideal objects* are represented by *points* of a formal space

The *introduction* of a point of a formal space corresponds to working in the sheaf model over this space

The *elimination* of this point is achieved by Beth-Kripke-Joyal explanation of the logic of this sheaf model

Some roots of this approach involve Brouwer's notion of choice sequences and an analysis of universal quantification over these objects in constructive mathematics

# Elimination of points

This involves replacing an ideal object by a syntactical theory that describes this object

"Phenomenological" description of infinite objects (only using "observable" properties)

# Point-Free Topology

An early point-free description of a space can be found in the work of Dedekind and Weber

*Theorie des algebraischen Funktionen einer Veründerlichen* , J. de Crelle t. XCII (1882) 181-290

aiming at giving a rigorous (and almost algebraic) presentation of Riemann surfaces

(An early application of this method is due to N. de Bruijn 1967 who could eliminate in this way the use of the Axiom of Choice in the context of Banach algebras analysing the proof of Wiener's Theorem on inverse of Fourier series.)

# Elimination of points

We can apply this method to proofs of purely existential statements in algebra which uses ideal objects (prime ideals, maximal ideals, valuation rings, ...)

We get elementary statements (essentially first-order) and constructive proofs,which are extracted from the non effective proofs

This can be seen as a partial realization of Hilbert's Program in commutative algebra

# Reasoning and computations

These proofs, being constructive can be considered as algorithms that compute a witness of these existential statements

Question: has the use of ideal/non effective arguments in mathematics some computational relevance??

Ideal methods seem to correspond to clever algorithmic ideas, reminiscent of the technique of lazy evaluation in functional programming (we cannot compute completely an infinite object but we can use partial finite amount of information about this object during a computation)

# Reasoning and computations

Some methods to analyse the computational meaning of non effective reasoning

-negative translation (Kolmogorov, Gödel, Gentzen)

-Dialectica interpretation (Gödel, + monotone interpretation, cf. the tutorial of U. Kolhenbach)

-formal topology/elimination mappings (this method is suggested in Troelstra *Choice sequences*, 6.12, with an application to Riemann's permutation theorem)

# Logic

Gödel's *incompleteness* theorem shows that Hilbert's Program does not work for arithmetic (Dirichlet's Theorem on arithmetic progressions of primes), even for purely universal statement

Gödel's *completeness* theorem indicates that Hilbert's Program should work for a large part of algebra, where statements are *first-order*

However Gödel's proof is *non constructive* and does not provide a way to eliminate ideal elements in proofs of first-order statement

(Gödel's *constructible sets* give a general method to eliminate the use of the Axiom of Choice)

# Zariski spectrum

$R$ commutative ring

The prime filters (classically complement of prime ideals) may be very difficult to build (prime factorisation), but it is simple to describe their logical theory of "observable" properties

Joyal's definition: free (bounded) distributive lattice generators $D(a)$, thought of as a pure symbols, and relations

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

We write $D(a_1, \ldots, a_n)$ for $D(a_1) \vee \cdots \vee D(a_n)$

# Zariski spectrum

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

We have $D(a^n) = D(a),\ n \geqslant 1$ and so $D(a) = 0$ if $a$ is nilpotent

We have for instance $D(a, b) = D(a + b, ab)$ and hence $D(a, b) = D(a + b)$ if $D(ab) = 0$

Question: how many elements $l$ are needed to write an arbitrary element $D(a_1, \ldots, a_l)$ of $\mathsf{Zar}\ R$?

# Nullstellensatz

We clearly have $D(a^2) = D(a)$ and so $D(a) \leqslant D(b_1, \ldots, b_m)$ whenever $a$ belongs to the radical of the ideal generated by $b_1, \ldots, b_m$

The *formal Nullstellensatz* states that conversely if $D(a) \leqslant D(b_1, \ldots, b_m)$ then $a$ belongs to the radical of the ideal generated by $b_1, \ldots, b_m$

This can be seen as a cut-elimination theorem

# Nullstellensatz

This follows from the following explicit description of the Zariski lattice: define $D(a_1, \ldots, a_n)$ to be the radical ideal generated by $a_1, \ldots, a_n$

It is *always* a distributive lattice; the product is also the intersection

(In general the lattice of finitely generated ideals of a ring is *not* distributive: take in $k[X, Y]$ the ideals $\langle X \rangle$, $\langle Y \rangle$ and $\langle X + Y \rangle$

A ring is *arithmetical* iff its lattice of ideals is distributive)

# Zariski spectrum, application

$D(a) = 0$ iff $a$ is nilpotent (we have $a^n = 0$ for some $n$)

This corresponds to the fact that the intersection of all prime ideals of a ring is the ideal of nilpotent elements

$D(a_1, \ldots, a_n) = 1$ iff $a_1, \ldots, a_n$ is *unimodular* that is $\langle a_1, \ldots, a_n \rangle = 1$

# Zariski spectrum, application

Gauss-Joyal identity

$$D(a_1, \ldots, a_n) \wedge D(b_1, \ldots, b_m) = D(c_1, \ldots, c_l)$$

if $(\Sigma a_i X^i)(\Sigma b_j X^j) = \Sigma c_k X^k$

Application: the product of primitive polynomials (ideal of coefficient is $1$) if primitive

This can be interpreted as using a generic prime filter (in the sheaf model over the Zariski spectrum)

We *force* the existence of such a prime filter

# Krull dimension

We think of a distributive lattice as a *formal space* (Stone)

The points of the space are the prime filters; the elements of the distributive lattice can be thought of as symbols for the compact open subsets of the space

In general non Haussdorf space: we can have non trivial chains of prime filters $\alpha_0 \subset \cdots \subset \alpha_n$ (chain of length $n$)

Krull dimension of a distributive lattice/commutative ring: maximal length of such chains

# Krull dimension

Here is a purely "phenomenological" approach

Given a bounded distributive lattice $L$, define the *boundary ideal* $B_a$ of an element $a$ of $L$ to be the ideal generated by $a$ and the ideal

$$a^{\perp} = \{b \in L \mid a \wedge b = 0\}$$

In term of points, the quotient $L/B_a$ describes the topological boundary of the compact open set $a$

# Krull dimension

We define inductively Kdim $L < 0$ iff $L$ is trivial and Kdim $L < n+1$ iff Kdim $L/B_a < n$ for all $a$ in $L$

Geometrically, the dimension of the space is $< n+1$ iff the dimension of each boundary of any compact open subspace is $< n$ (cf. Menger-Uryshon dimension)

# Krull dimension

In this way Kdim $L = 0$ iff any element $a$ has a *complement* i.e. there exists $b$ such that $a \vee b = 1$ and $a \wedge b = 0$ iff $L$ is a Boolean algebra

In general Kdim $L < n$ iff any sequence $a_1, \ldots, a_n$ has a "complement" $b_1, \ldots, b_n$ such that

$$\begin{array}{lcl} 1 & = & a_1 \vee b_1 \\ a_1 \wedge b_1 & \leqslant & a_2 \vee b_2 \\ & \ldots & \\ a_{n-1} \wedge b_{n-1} & \leqslant & a_n \vee b_n \\ a_n \wedge b_n & = & 0 \end{array}$$

# Krull dimension

Intermediate logics between classical (Boolean algebra) and intuitionistic logic (Heyting algebra)

dim $0$: $p \vee \neg p = 1$

dim $1$: $p \vee (p \to (q \vee \neg q)) = 1$

$\ldots$

In term of Kripke models: the height of the tree is bounded

These logics are *decidable* (Ono, Smorynski)

Finitely generated algebras are finite (cf. recent work of Katarzyna Słomczyǹska)

# Krull dimension of a ring

We define Kdim $R < n$ to mean that for any $a_1, \ldots, a_n$ in $R$ there exists $b_1, \ldots, b_n$ in $R$ such that

$$\begin{array}{lcl} 1 & = & D(a_1, b_1) \\ D(a_1 b_1) & \leqslant & D(a_2, b_2) \\ & \cdots & \\ D(a_{n-1} b_{n-1}) & \leqslant & D(a_n, b_n) \\ D(a_n b_n) & = & 0 \end{array}$$

**Theorem:** Kdim $R < n$ *iff* Kdim (Zar $R$) $< n$

# Krull dimension

**Theorem:** *We have* Kdim $k[X_1, \ldots, X_{n-1}] < n$

This follows from the fact that $n$ polynomials in $k[X_1, \ldots, X_{n-1}]$ are algebraically dependent

# Kronecker's Theorem

**Theorem:** *If* Kdim $R < n$ *then any element of* Zar $R$ *can be written on the form* $D(a_1, \ldots, a_l)$ *with* $l \leqslant n$

For instance $D(a, c) = D(a + bc)$ if $D(ab) = 0$ and $D(a, b) = 1$

In general

$$D(a_1, \ldots, a_n, c) = D(a_1 + cb_1, \ldots, a_n + cb_n)$$

where $b_1, \ldots, b_n$ is a complement of $a_1, \ldots, a_n$

# Kronecker's Theorem

This gives an algorithm for Kronecker's Theorem: given $m$ polynomials $Q_1, \ldots, Q_m$ in $\mathbb{Q}[X_1, \ldots, X_{n-1}]$ find $n$ polynomials $P_1, \ldots, P_n$ such that the two lists have the same common complex zeros

(More precisely we find $P_1, \ldots, P_n$ in the ideal generated by $Q_1, \ldots, Q_m$ and $Q_1, \ldots, Q_m$ in the radical of the ideal generated by $P_1, \ldots, P_l$)

# Kronecker's Theorem

Geometrically this means that an algebraic variety in $\mathbb{C}^{n-1}$ is the intersection of $n$ hypersurfaces (the bound is not optimal; $n-1$ is enough)

Section 10 of *Grundzüge einer arithmetischen Theorie der algebraischen Grössen*
*J. reine angew. Math.* 92, 1-123 (1882)

This result was generalized from polynomial ring to the case of Noetherian rings and Krull dimension by van der Waerden 1941

It is usually presented only in the case with the additional hypothesis that the ring is Noetherian, for instance in the text book of Kunz, Eisenbud, Ischebeck-Rao, ...

# 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

# Krull dimension

The same method can be used to extract an algorithm for Forster's Theorem (1964), which generalizes Serre's splitting-off Theorem (1957)

**Theorem:** *If* Kdim $R < n$ *and* $M$ *is a rectangular matrix such that* $\Delta_n(M) = 1$ *then we can find an unimodular linear combination of the column of* $M$

Here $\Delta_n(M) = \vee_\nu D(\nu)$ where $\nu$ ranges over $n \times n$ minor of $M$

# Forster(-Swan)'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)

# Serre's Splitting-Off Theorem

If we specialise to the case where $M$ is an idempotent matrix we get Serre's Splitting-Off Theorem

In this case, the geometric intuition is that we have a vector bundle over a space of dimension $< n$ and each fibers having a dimension $\geqslant n$, and the unimodular combination corresponds to a non vanishing section

# Point-free formulation

In this example of Krull dimension and finding generators for modules the search for a point-free formulation *does* simplify the statements and the proofs

The proofs are elementary and constructive

One can in this way find a purely *first-order formulation* of one open question in Heitman's 1984 paper (and then solves this question)

# Valuation rings

We can similarly give a point-free formulation of the notion of valuation ring

If $R$ is a domain of fied of fraction $K$ a *valuation ring* is a subring $V$ of $K$ containing $R$ such that $s$ or $s^{-1}$ is in $V$ whenever $s \neq 0$ in $K$

We have $s/(1-s)$ in $V$ or $(1-s)/s$ in $V$, hence $1/s$ in $V$ or $1/(1-s)$ in $V$

A valuation ring is a *local* ring: if $r$ is in $V$ either $r$ or $1-r$ is invertible in $V$. Let $\mathfrak{m}_V$ be the maximal ideal of $V$

$V \longmapsto \mathfrak{m}_V \cap R$ is the *center map*

# Space of valuations

By analogy with Joyal's definition of the Zariski spectrum of a ring we define the formal space $\mathrm{Val}(R)$ to be the distributive lattice generated by the symbols $V(s)$ for $s$ in $K$ with the relations

$$1 = V(s) \vee V(1/s), \qquad V(s_1) \wedge V(s_2) \leqslant V(s_1 + s_2) \wedge V(s_1 s_2)$$

and, for $r$ in $R$

$$1 = V(r)$$

# Space of valuations

We write $V(s_1, \ldots, s_n)$ for $V(s_1) \vee \cdots \vee V(s_n)$

In general we do not have $V(s_1) \wedge V(s_2) = V(s_1 s_2)$

We have $V(s) \wedge V(1/s) = V(s + 1/s)$

# Zariski spectrum and space of valuations

We have a lattice map

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

This is the *center map*, built here only using the initiatility condition

**Theorem:** *The point-free center map is injective*

# Zariski spectrum and space of valuations

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

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

This can be seen as a *conservativity* result between two logical theories and it corresponds to the following extension Theorem

**Theorem:** (Chevalley) *For any prime ideal* $\mathfrak{p}$ *of* $R$ *there exists a valuation ring* $V$ *such that* $\mathfrak{m}_V \cap R = \mathfrak{p}$

# Elimination of Noetherian hypotheses

In most examples analysed so far in algebra, Noetherian hypotheses can be eliminated and replaced by pure first-order formulation (for instance, Kronecker's or Forster's Theorem).

Regular Element Theorem: *if* $R$ *Noetherian* and if $I = \langle a_1, \ldots, a_n \rangle$ is regular (i.e. $Ix = 0$ implies $x = 0$) then there exists $u \in I$ such that $u$ is regular (i.e. $ux = 0 \to x = 0$)

# Abstract functional analysis

The same method can be used to analyse some non effective results in functional analysis

Hahn-Banach theorem (Mulvey): the *extension* theorem becomes a *conservativity* result between theories

Gelfand representation theorem: we get an alternative constructive approach to the representation theorem

Point-free analysis of Krivine's paper *Anneaux préordonnés* Journal d'Analyse Math., 12, p. 307-326 (1964)

# Minimal and maximal primes

Some arguments in algebra using a generic maximal/minimal prime ideal

-Suslin's proof of Serre's problem: the argument has been analysed by I. Yengui who could extract an algorithm from this proof

-Traverso-Swan's characterisation of seminormal rings

$R$ is seminormal iff $b^2 = c^3 \ \rightarrow \exists a.\ b = a^3 \wedge c = a^2$ iff

the canonical map Pic $R \rightarrow$ Pic $R[X]$ is an isomorphism

(application: if $R$ is seminormal then so is $R[X]$) work of Th. C., H. Lombardi, C. Quitté, S. Baroumi

-Peskine's proof of Zariski's Main Theorem

# What next?

H. Edwards in several works

*Essays in Constructive Mathematics*, Springer, New York, 2005

*Divisor Theory*, Birkhauser, Boston, 1995

*A Normal Form for Elliptic Curves*, Bulletin of the AMS, vol. 44 (2007) 393-422

has shown how to give a constructive treatment of algebraic function theory following Kronecker. I believe that this treatment can be simplified and made closer to the non effective presentation using a point-free presentation

Goal: purely algebraic presentation of Abel's work on algebraic functions

# What next?

Coste M., Lombardi H., Roy M.F.
*Dynamical method in algebra: Effective Nullstellensätze* J.P.A.A. 155 (2001)

contains a constructive presentation of some quantifier elimination results, using a method inspired from model theory and close to point-free presentations

Connections with the "dynamical method" D5 used in computer algebra

This should be extended to the case of differential closed fields and connected to the early work of Drach

*Sur le problème logique de l'intégration des équations différentielles.* Annales de la faculté des sciences de Toulouse, Sér. 2, 10 (1908), p. 393-472

# What next?

Constructive model theory?

Decidability of the theory of algebraically closed fields of a given characteristic by the fact that any two fields of the same uncoutable cardinality are isomorphic. What is the computational meaning of this proof?

Krivine's thesis, part 2 (available at his home page): use of Krein-Milman which does not seem direct to interpret constructively

# Some references

P. Johnstone *The point of pointless topology* Bull. Amer. Math. Soc. (N.S.) Volume 8, Number 1 (1983), 41-53

P. Johnstone *Stone Spaces*, Cambridge University Press, 1984.

M. Fourman and D. Scott *Sheaves and Logic*, in: Applications of Sheaves, Proc. Durham, LNiMath 753, 1979

# Some references

Coste M., Lombardi H., Roy M.F.
"Dynamical method in algebra: Effective Nullstellensätze"
J.P.A.A. 155 (2001)

L. Ducos, H. Lombardi, C. Quitté and M. Salou.
"Théorie algorithmique des anneaux arithmétiques, de Prüfer et de Dedekind."
Journal of Algebra 281, (2004), 604-650.

Edwards, Harold M.
*Essays in constructive mathematics.* New York, NY: Springer (2005)

# Some references

Th.C., H. Lombardi, M.-F. Roy
"An elementary characterisation of Krull dimension"
Selected articles from a workshop in San Servolo. Venice, Italy, May 12-16, 2003.
Oxford Logic Guides, Oxfort University Press.

L. Español
"*Constructive Krull dimension of lattices.*" Rev. Acad. Cienc. Zaragoza (2) 37 (1982), 5–9.

L. Español
"*Dimension of Boolean valued lattices and rings.*" J. Pure Appl. Algebra 42 (1986), no. 3, 223–236.

# Some references

Th.C. "*Sur un théorème de Kronecker concernant les variétés algébriques.*"
C.R.Acad.Sci., Paris, Ser I 338 (2004), Pages 291-294

Th. C., H. Lombardi, C. Quitté
"*Generating non-Noetherian modules constructively.*"
Manuscripta Mathematica, 1115, 513-520 (2004)

L. Ducos "*Vecteurs unimodulaires et système générateurs.*"
Journal of Algebra 297, 566-583 (2005)