(Un)decidability results on real vector spaces arising from the formalization of mathematics

John Harrison, Intel Corporation

Joint work with Robert M. Solovay and Rob Arthan

Computability in Europe 2008, “Proofs” special session

University of Athens

Mon 16th June 2008 (15:15 – 15:55)

The state of formalization

Formalization of mathematics in theorem provers is attracting increasing interest, for intellectual and practical reasons.

`http://www.cs.ru.nl/~freek/100/` lists some notable theorems that have been formally proved, e.g.

- Four-Colour Theorem (Gonthier)
- Prime Number Theorem (Avigad, Harrison)
- Jordan Curve Theorem (Hales)

Ambitious projects in progress to formally prove

- Hales's proof of Kepler conjecture (Flyspeck project)
- Feit-Thomson theorem (from classification of finite simple groups)

The interaction-automation spectrum

Theorem provers offer widely different levels of automation:

AUTOMATH (de Bruijn)

Mizar (Trybulec)

. . .

LCF (Milner and others)

. . .

ACL2 (Boyer, Kaufmann, Moore)

Vampire (Voronkov)

Arguably most productive for formalization are those that fall in the middle, e.g. Coq, HOL, Isabelle, Nuprl, PVS.

The user provides guidance but many “routine” steps are automated.

Current automation

Many major proof assistants offer efficient automated proof of facts from linear real, integer or natural number arithmetic.

```
# time ARITH_RULE ‘!y j:num. y < j ==> y + 1 <= (y + 1 + j) DIV 2‘;;
CPU time (user): 0.11
val it : thm = |- !y j. y < j ==> y + 1 <= (y + 1 + j) DIV 2
```

Some also offer automation for nonlinear arithmetic over reals, but this is typically much slower and often impractical

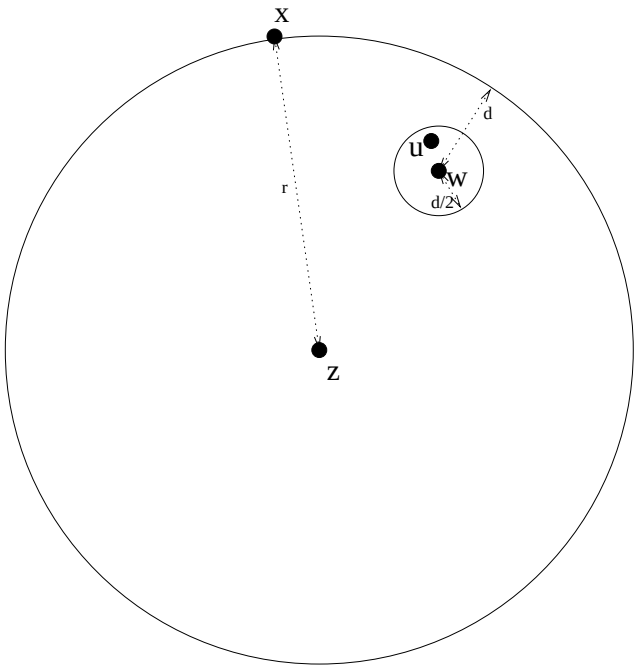
```
# time REAL_SOS
   ‘!x:real. abs(x) <= &1
     ==> abs(&64 * x pow 7 - &112 * x pow 5 + &56 * x pow 3 - &7 * x) <= &1‘;;
CPU time (user): 3.75
...
```

Of course, by Gödel/Tarski/Matiyasevich, nonlinear arithmetic over naturals or integers is in general impossible.

But often useful to prove relaxations over reals or over all rings etc.

Automation gap in formalizing complex analysis

```
|- abs(norm(w - z) - r) = d /\ norm(u - w) < d / &2 /\ norm(x - z) = r
   ==> d / &2 <= norm(x - u)
```



This is not immediately solvable by HOL Light's standard automation, even though the analogous property over $\mathbb{R}$ would be.

Straightforward approach and questions

We could just introduce two real coordinates for each point and reduce everything to reals.

However, the property doesn't depend on the fact that we are working in $\mathbb{C} = \mathbb{R}^2$.

It would work equally well over $\mathbb{R}^n$ for any n, or indeed *any* real inner product space.

Question: is the theory of real inner product spaces decidable?

The theory of real vector spaces

Two-sorted first-order theory with sorts of vectors $\mathcal{V}$ and scalars $\mathcal{S}$.

Language has zero vector $\mathbf{0}$, addition and negation of vectors, and multiplication of vector by scalar, plus the usual constants, addition, negation and multiplication of scalars.

The models of the theory are those structures where $\mathcal{S}$ and its operations are interpreted over $\mathbb{R}$ in the usual way, and the vector space axioms are satisfied.

Vector space axioms

$$\forall \mathbf{u}\ \mathbf{v}\ \mathbf{w}.\ \mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$$

$$\forall \mathbf{v}\ \mathbf{w}.\ \mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$$

$$\forall \mathbf{v}.\ \mathbf{0} + \mathbf{v} = \mathbf{v}$$

$$\forall \mathbf{v}.\ -\mathbf{v} + \mathbf{v} = \mathbf{0}$$

$$\forall a\ \mathbf{v}\ \mathbf{w}.\ a(\mathbf{v} + \mathbf{w}) = a\mathbf{v} + a\mathbf{w}$$

$$\forall a\ b\ \mathbf{v}.\ (a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$$

$$\forall \mathbf{v}.\ 1\mathbf{v} = \mathbf{v}$$

$$\forall a\ b\ \mathbf{v}.\ (ab)\mathbf{v} = a(b\mathbf{v})$$

The theory of real inner product spaces

The language of vector spaces plus an inner product operation $\mathcal{V} \times \mathcal{V} \to \mathcal{S}$ written $\langle -, - \rangle$ and satisfying:

$$\forall \mathbf{v}\ \mathbf{w}.\ \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$$

$$\forall \mathbf{u}\ \mathbf{v}\ \mathbf{w}.\ \langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$$

$$\forall a\ \mathbf{v}, \mathbf{w}.\ \langle a\mathbf{v}, \mathbf{w} \rangle = a\langle \mathbf{v}, \mathbf{w} \rangle$$

$$\forall \mathbf{v}.\ \langle \mathbf{v}, \mathbf{v} \rangle \geq 0$$

$$\forall \mathbf{v}.\ \langle \mathbf{v}, \mathbf{v} \rangle = 0 \Leftrightarrow \mathbf{v} = \mathbf{0}$$

In Euclidean space, the inner product is $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} \mathbf{x}_i \mathbf{y}_i$.

Decidability of inner product spaces

Answer (Solovay): Yes: the theory of real inner product spaces is decidable, and admits quantifier elimination in a language expanded with inequalities on dimension.

In fact (Arthan) a formula with k vector variables holds in all inner product spaces iff it holds in each $\mathbb{R}^n$ for $0 \leq n \leq k$, which is in the decidable Tarski subset.

These results directly give rise to methods for testing if a formula holds in all real inner product spaces, or those satisfying some particular constraints on dimension.

Inner product spaces are a conservative extension of vector spaces (use any basis to define an inner product), so we also have quantifier elimination and decidability for vector spaces.

The problem of nonlinearity

In Euclidean space, the norm is defined by $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$, and we can similarly define a norm this way for *any* inner product space.

Unfortunately, problems that look entirely "linear" but involve the norm, like our example:

$$|\|\mathbf{w} - \mathbf{z}\| - r| = d \wedge \|\mathbf{u} - \mathbf{w}\| < d/2 \wedge \|\mathbf{x} - \mathbf{z}\| = r \Rightarrow d/2 \leq \|\mathbf{x} - \mathbf{u}\|$$

then give rise to *nonlinear* problems over the reals, whether we use the general decision procedure or just a reduction to $\mathbb{R}^2$.

Naive reduction of our example

Just introduce coordinates for each point and use n_i for the norms:

$$
\begin{aligned}
&(w_1 - z_1)^2 + (w_2 - z_2)^2 = n_1^2 \wedge n_1 \geq 0 \wedge \\
&(u_1 - w_1)^2 + (u_2 - w_2)^2 = n_2^2 \wedge n_2 \geq 0 \wedge \\
&(x_1 - z_1)^2 + (x_2 - z_2)^2 = n_3^2 \wedge n_3 \geq 0 \wedge \\
&(x_1 - u_1)^2 + (x_2 - u_2)^2 = n_4^2 \wedge n_4 \geq 0 \wedge \\
&|n_1 - r| = d \wedge n_2 < d/2 \wedge n_3 = r \\
&\Rightarrow d/2 \leq n_4
\end{aligned}
$$

This is within the scope of automation in principle, but it's quite inefficient in practice.

Can we be even more general and prove that our property holds in all *normed* real vector spaces?

The theory of real normed spaces

The language of vector spaces plus a norm operation $\mathcal{V} \to \mathcal{S}$ written $\| - \|$ and satisfying:

$$\forall \mathbf{v}.\ \|\mathbf{v}\| = 0 \Rightarrow \mathbf{v} = \mathbf{0}$$

$$\forall a\ \mathbf{v}.\ \|a\mathbf{v}\| = |a|\|\mathbf{v}\|$$

$$\forall \mathbf{v}\ \mathbf{w}.\ \|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$$

For example, on $\mathbb{R}^n$, can use the 1-norm $\|\mathbf{x}\| = \sum_{i=1}^{n} |\mathbf{x}_i|$ or the ∞-norm $\|\mathbf{x}\| = \max\{|\mathbf{x}_i| \mid 1 \leq i \leq n\}$.

Relation between decision problems

Every inner product space gives rise to a normed space by defining $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$.

Not every norm arises from an inner product in this way, but *if it does*, we can recover the inner product from the norm, e.g. by $\langle \mathbf{x}, \mathbf{y} \rangle = (\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2)/2$.

Write p^* for such a replacement inside a formula p, and let I be the inner product axioms. Then p holds in all inner product spaces iff $I^* \wedge p^*$ holds in all normed spaces.

Thus, on general grounds, the decision problem for normed spaces is at least as hard as that for vector spaces. But is it harder?

Normed spaces: better or worse?

(Solovay) Yes, the full theory of real normed spaces is strongly undecidable. In fact, it is not even arithmetical, and is actually harder than deciding second-order arithmetic!

(Arthan) Even the purely *additive* theory of *2-dimensional* normed spaces is strongly undecidable.

(Harrison, Arthan) However universally quantified and linear problems in the theory of normed spaces can be decided by a generalization of parametrized linear programming, and the full universal theory is still decidable.

Related results: constraints on dimension

There is a striking contrast between the well-behaved decidable theory of inner product spaces and the strongly undecidable theory of normed spaces.

One way of understanding this is to recall the 'finite-dimensional model' property of inner product spaces and see that this fails for normed spaces:

> There exist non-zero vectors, and the unit disc has no extreme points. (An extreme point of a set is one that is not on a line between two other distinct points of the set.)

This has an infinite-dimensional model, e.g. $\mathbb{R}^*$ with the ∞-norm, but by the Krein-Milman theorem, no finite-dimensional model.

Related results: dependence on field

It has been known since Tarski that all real-closed fields are elementarily equivalent to $\mathbb{R}$.

For the theory of *inner product* spaces, we have a similar property: the theory is the same over $\mathbb{R}$ as over any real-closed field.

In fact, the reduction of a vector formula to an equivalid scalar formula depends on very few properties, mainly the existence of square roots.

On the other hand, because the theory of real *normed* spaces is non-arithmetical, it must differ from the theory over real-closed fields in general, since that theory is recursively axiomatizable.

Completeness

We say that a space is *complete* if every Cauchy sequence

$$\forall \epsilon > 0.\, \exists N.\, \forall m, n \geq N.\, \|\mathbf{x}_m - \mathbf{x}_n\| < \epsilon$$

converges

$$\exists \mathbf{l}.\, \forall \epsilon > 0.\, \exists N.\, \forall n \geq N.\, \|\mathbf{x}_n - \mathbf{l}\| < \epsilon$$

The following is standard terminology.

- Hilbert space = complete inner product space
- Banach space = complete normed space

Related results: significance of completeness

Completeness cannot be expressed in the language we consider here.

The theories of Hilbert spaces and inner product spaces are *the same*, because all finite-dimensional inner product spaces are complete.

(Solovay) The theories of Banach spaces and normed spaces are *different*.

(Solovay) The decision problems for Banach spaces and normed spaces *are*, however, mutually many-one reducible, to each other and to a certain fragment of third-order number theory.

Related results: metric spaces

Results for normed spaces are echoed by the simpler theory of metric spaces, where we have no operations on points.

$$\forall \mathbf{x}\ \mathbf{y}.\ d(\mathbf{x},\mathbf{y}) \geq 0$$

$$\forall \mathbf{x}\ \mathbf{y}.\ d(\mathbf{x},\mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$$

$$\forall \mathbf{x}\ \mathbf{y}.\ d(\mathbf{x},\mathbf{y}) = d(\mathbf{y},\mathbf{x})$$

$$\forall \mathbf{x}\ \mathbf{y}\ \mathbf{z}.\ d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z}) \leq d(\mathbf{x},\mathbf{z})$$

The full theory is strongly undecidable, but the universal subset (in fact the AE subset) is decidable.

Completeness cannot be expressed in the metric language.

The theories of metric spaces and complete metric spaces are different.

Interpreting first-order arithmetic

For a formula $N(x)$ with one free scalar variable, we can assert that its interpretation within $\mathbb{R}$ is the natural numbers by this formula Nat:

$$\begin{array}{l}(\forall x.\ N(x) \Rightarrow x \geq 0) \wedge \\ (\forall x.\ x \geq 0 \Rightarrow (N(x) \Leftrightarrow N(x+1))) \wedge \\ (\forall x.\ 0 \leq x \wedge x < 1 \Rightarrow (N(x) \Leftrightarrow x = 0))\end{array}$$

Let $\phi_{\mathbb{N}}$ be the result of relativizing all quantifiers in ϕ, e.g. $(\forall n.\ P[n])_{\mathbb{N}} =_{def} \forall n.\ N(n) \Rightarrow P[n]_{\mathbb{N}}$.

Then provided there is at least one model of the metric space axioms where the formula $N(x)$ does indeed define $\mathbb{N}$, we have:

ϕ holds in $\mathbb{N}$ iff Nat $\Rightarrow \phi_{\mathbb{N}}$ holds in all metric spaces.

Interpreting second-order arithmetic

(Folklore? See similar results in Moschovakis and Kechris) In the theory of reals with an integer or natural number predicate, can even interpret *second-order* arithmetic.

One way is to encode a set $S \subseteq \mathbb{N}$ with characteristic function χ_S as the real $\sharp S = \sum_{n=0}^{\infty} \chi_S(n)/3^n$, replacing quantification over sets with quantification over $\mathbb{R}$.

Thus, provided there is at least one model of the metric space axioms where the formula $N(x)$ does indeed define $\mathbb{N}$, the theory of metric spaces is at least as hard as second-order arithmetic.

And there is indeed such a model, just the integers with the usual metric and the formula $N(x) =_{def} \exists \mathbf{a}\ \mathbf{b}.\ d(\mathbf{a}, \mathbf{b}) = x$.

Interpretation in a linear theory

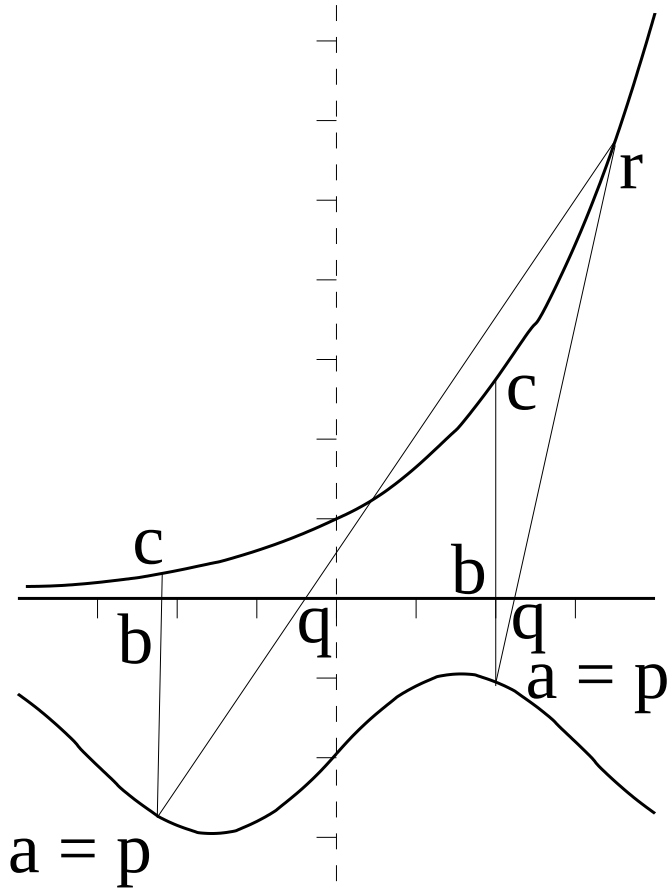
With a bit more work, we can even avoid assuming multiplication in the language by similarly characterizing it for a formula $M(x, y, z)$ and finding a model where some such formula works too:

$$
\begin{aligned}
&(\forall x\ y.\ \exists! z.\ M(x, y, z)) \wedge \\
&(\forall x\ y\ z.\ M(x, y, z) \Rightarrow M(y, x, z)) \wedge \\
&(\forall y\ z.\ M(0, y, z) \Leftrightarrow z = 0) \wedge (\forall y\ z.\ M(1, y, z) \Leftrightarrow z = y) \wedge \\
&(\forall x_1\ x_2\ y\ z_1\ z_2.\ M(x_1, y, z_1) \wedge M(x_2, y, z_2) \Rightarrow M(x_1 + x_2, y, z_1 + z_2)) \wedge \\
&(\forall x\ y\ z\ \epsilon.\ \ M(x, y, z) \wedge \epsilon > 0 \\
&\qquad\qquad \Rightarrow \exists \delta.\ \delta > 0 \wedge \forall x'\ z'.\ \ |x - x'| < \delta \wedge M(x', y, z') \Rightarrow |z - z'| < \epsilon)
\end{aligned}
$$

NB: we need full *real* multiplication to interpret second-order arithmetic.

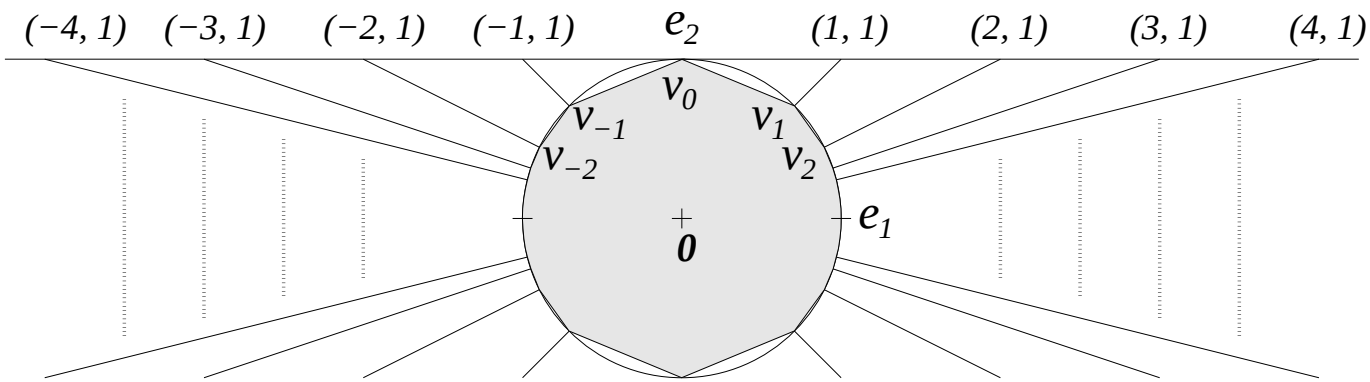
A more exotic metric space

One can indeed come up with an exotic metric space where this works.



The same thing for normed spaces

Constructing a normed space where we can define the integers is harder. One way is using this 'infinigon' in $\mathbb{R}^2$:



This can be used as the unit circle of a norm, and one can characterize the integers using just the language of normed spaces.

Decidability of AE fragment for metric spaces

Validity of an AE formula reduces, after negation and Skolemization, to unsatisfiability of a formula with some vector constants $\mathbf{c}_i$ and real constants, but only universally quantified variables, at least over vectors:

$$\forall \overline{\mathbf{y}} / Q \overline{z}.\ P[\mathbf{c}_1, \ldots, \mathbf{c}_n, \overline{\mathbf{y}}, \overline{z}]$$

This formula is satisfiable in a metric space iff it is satisfiable in a metric space with a point domain of size $\leq n$.

We can test it by instantiating the universally quantified vector variables in this formula and the metric axioms with constants $\mathbf{c}_i$ in all possible ways.

A similar ‘finite model’ property holds for normed spaces, but replacing ‘size’ by ‘dimension’. The universal fragment at least is decidable.

Current status and conclusions

A paper by Solovay, Arthan and Harrison proving all the main results is almost complete.

HOL Light contains implementations of:

- A limited case of the decision procedure for inner product spaces. The practical usefulness is limited but it can, for example, prove the Cauchy-Schwartz inequality automatically.
- The procedure for universal linear problems in normed spaces. This is very useful for solving routine problems like the motivating example at the beginning.

A good example of how problems arising from real formalization problems can lead to interesting theoretical investigation.