A polynomial that detects the consistency of set theory

Matt Booth

PG Colloquium, University of Edinburgh

19 January 2017

Theorem

One can write down a (multivariate) polynomial p, with integer coefficients, that has a solution in natural numbers if and only if ZFC[a] is inconsistent.

[a] **Z**ermelo-**F**raenkel Set Theory with the Axiom of **C**hoice

Theorem

One can write down a (multivariate) polynomial p, with integer coefficients, that has a solution in natural numbers if and only if ZFC[a] is inconsistent.

[a] **Z**ermelo-**F**raenkel Set Theory with the Axiom of **C**hoice

- In fact, one can take p to have at most 9 variables, or to be a quartic (but not both at once). Moreover, the theorem doesn't use any special properties of ZFC!

Theorem

One can write down a (multivariate) polynomial p, with integer coefficients, that has a solution in natural numbers if and only if ZFC[a] is inconsistent.

[a] **Z**ermelo-**F**raenkel Set Theory with the Axiom of **C**hoice

- In fact, one can take p to have at most 9 variables, or to be a quartic (but not both at once). Moreover, the theorem doesn't use any special properties of ZFC!
- Disclaimer: I'm not a logician. I'll skim over some technicalities and maybe make some nonstandard definitions.

What kind of polynomial is p?

- I don't mean a polynomial like

$$p(x) = \begin{cases} x+1 & \text{if ZFC is consistent} \\ x-1 & \text{if ZFC is inconsistent} \end{cases}$$

  This is cheating!

- Nor do I mean a polynomial like $p(x) = x^2 + 1$, since ZFC proves that this has no (natural) solutions. If this is our p, then ZFC is consistent, and also proves its own consistency, which contradicts Gödel's second incompleteness theorem.
- So p must be fairly complicated - if ZFC is consistent, then p has no solutions, but ZFC doesn't prove this!

Some terminology

Let $S \subseteq \mathbb{N}$.

- Say that S is **recursively enumerable** (r.e. for short) if there is an algorithm A, that takes natural numbers as input, that halts on input n if and only if $n \in S$.
- Equivalently, there's an algorithm B, that takes infinite time to run, that prints out precisely the elements of S. Intuitively, given A, run $A[n]$ at time n, and whenever $A[n]$ finishes print n. Given B and n, just run B and check whether n shows up.
- For example the set of prime numbers is r.e.
- Note that non-r.e. sets must exist by a cardinality argument!

The MRDP theorem

- Say that S is **diophantine**[1] if there is a polynomial $q(x_0, \ldots, x_m)$ with coefficients in $\mathbb{Z}$ such that $n \in S$ if and only if there exist $a_1, \ldots, a_m \in \mathbb{N}$ with $q(n, a_1, \ldots, a_m) = 0$.

[1] Named after **Diophantus**, 3rd century Greek mathematician who studied these kinds of equations.

The MRDP theorem

- Say that S is **diophantine**[1] if there is a polynomial $q(x_0, \ldots, x_m)$ with coefficients in $\mathbb{Z}$ such that $n \in S$ if and only if there exist $a_1, \ldots, a_m \in \mathbb{N}$ with $q(n, a_1, \ldots, a_m) = 0$.
- Clearly a diophantine set is r.e. - just enumerate all tuples of natural numbers $(n, a_1, \ldots, a_m)$, plug them into q, and if the answer is zero then add n to the list. Is the converse true?

[1]Named after **Diophantus**, 3rd century Greek mathematician who studied these kinds of equations.

The MRDP theorem

- Say that S is **diophantine**[1] if there is a polynomial $q(x_0, \ldots, x_m)$ with coefficients in $\mathbb{Z}$ such that $n \in S$ if and only if there exist $a_1, \ldots, a_m \in \mathbb{N}$ with $q(n, a_1, \ldots, a_m) = 0$.
- Clearly a diophantine set is r.e. - just enumerate all tuples of natural numbers $(n, a_1, \ldots, a_m)$, plug them into q, and if the answer is zero then add n to the list. Is the converse true?

Theorem (Matiyasevich–Robinson–Davis–Putnam, 1970)

Any recursively enumerable set S is diophantine. Moreover, given an algorithm that prints out S, we can write down a polynomial q.

---

[1] Named after **Diophantus**, 3rd century Greek mathematician who studied these kinds of equations.

Motivation: Hilbert's Tenth Problem

- Question: Is there an algorithm A, that accepts polynomials over $\mathbb{Z}$ as arguments, that will tell us whether p has a root in $\mathbb{N}$ or not?

Motivation: Hilbert's Tenth Problem

- Question: Is there an algorithm A, that accepts polynomials over $\mathbb{Z}$ as arguments, that will tell us whether p has a root in $\mathbb{N}$ or not?
- MRDP tells us that if the answer is yes, then every r.e. set S is **recursive**, meaning that there's an algorithm that accepts natural numbers n as input and tells us whether or not $n \in S$. To see this, given S, we can write down p, and then apply A to $q := p(n, x_2, \ldots, x_m)$, since $n \in S$ if and only if q has roots.

Motivation: Hilbert's Tenth Problem

- Question: Is there an algorithm A, that accepts polynomials over $\mathbb{Z}$ as arguments, that will tell us whether p has a root in $\mathbb{N}$ or not?
- MRDP tells us that if the answer is yes, then every r.e. set S is **recursive**, meaning that there's an algorithm that accepts natural numbers n as input and tells us whether or not $n \in S$. To see this, given S, we can write down p, and then apply A to $q := p(n, x_2, \ldots, x_m)$, since $n \in S$ if and only if q has roots.
- Since there exist r.e. sets which are not recursive, no such A can exist.

Putnam's trick: generating polynomials

- If S is a r.e. set (with $0 \notin S$), by the MRDP theorem it's diophantine. Take a polynomial $M(x_0, \ldots, x_m)$ such that $n \in S$ if and only if $M(n, x_1, \ldots, x_m)$ has a solution in natural numbers. Replacing M by M^2 if necessary, we may assume that M is nonnegative. Then setting $Q := x_0(1 - M)$, we see that the positive values taken by Q as $x_0, \ldots, x_m$ range across $\mathbb{N}$ are precisely the members of S.
- Jones, Sato, Wada and Wiens wrote down such a polynomial Q when S is the set of prime numbers:

A prime generating polynomial

$$
\begin{aligned}
&(k+2)\{1-[wz+h+j-q]^2-[(gk+2g+k+1)\cdot(h+j)+h-z]^2-[2n+p+q+z-e]^2\\
&-[16(k+1)^3\cdot(k+2)\cdot(n+1)^2+1-f^2]^2-[e^3\cdot(e+2)(a+1)^2+1-o^2]^2-[(a^2-1)y^2+1-x^2]^2\\
&-[16r^2y^4(a^2-1)+1-u^2]^2-[((a+u^2(u^2-a))^2-1)\cdot(n+4dy)^2+1-(x+cu)^2]^2-[n+l+v-y]^2\\
&-[(a^2-1)l^2+1-m^2]^2-[ai+k+1-l-i]^2-[p+l(a-n-1)+b(2an+2a-n^2-2n-2)-m]^2\\
&-[q+y(a-p-1)+s(2ap+2a-p^2-2p-2)-x]^2-[z+pl(a-p)+t(2ap-p^2-1)-pm]^2\}
\end{aligned}
$$

Figure: A polynomial whose positive values (as $a, \ldots, z$ range across $\mathbb{N}$) are precisely the prime numbers. It also takes negative values; e.g. -76.

Idea of today's proof

- Code up proofs in ZFC as natural numbers (Gödel numbering)
- Write an algorithm that looks at a natural number and decides whether it codes a proof of a contradiction in ZFC
- Get a r.e. set S expressing consistency of ZFC
- Apply the MRDP theorem to S to get the polynomial p.

Model theory, 1

- A **theory** T is a pair (σ, A) where σ, the **signature**, is a tuple (F, R, C) of function symbols, relation symbols, and constant symbols, and A is a set of axioms.
- Example: the theory of abelian groups has signature $(+, 0)$ and axioms including $\forall x \forall y \forall z((x+y)+z = x+(y+z))$ and $\forall x \exists y(x+y=0)$.
- A **model** of a theory T is a set M with functions $M \to M$, relations on M, and constants in M, all satisfying the axioms.
- $(\mathbb{Z}/5\mathbb{Z}, +)$ is a model of the theory of abelian groups. $(\mathbb{R}, +)$ is a model. S_3 is not a model. $(\mathbb{N}, +)$ is not a model.

Model theory, 2

- A **proof** of a sentence ϕ from a set of sentences Σ is a list of sentences $\phi_1, \ldots, \phi_n$ with $\phi_n = \phi$ and where ϕ_{i+1} follows from ϕ_i by some deduction rules applied to the previous sentences and Σ.
- A **theorem** of T is a sentence ϕ with a proof. Write $T \vdash \phi$ to mean that T proves ϕ. T proves ϕ if and only if ϕ is true in all models (Gödel's completeness theorem).
- For the theory of abelian groups, $0 + 0 = 0$ is a theorem. $\forall x(x + x + x = 0)$ is not a theorem, since not every abelian group is 3-torsion. But its negation $\exists x(x + x + x \neq 0)$ is not a theorem either, since it's false in $\mathbb{Z}/3\mathbb{Z}$. We might say that $\forall x(x + x + x = 0)$ is **independent**.

Model theory, 3

- A theory T is **effectively axiomatisable** if there's an algorithm that runs in infinite time that prints out precisely all of the theorems of T.
- Some examples of effectively axiomatised theories:
  - 'Algebraic' theories: groups, rings, Lie algebras,...
  - 'Arithmetic' theories: Peano arithmetic, Robinson arithmetic (PA without induction), Presburger arithmetic (PA without multiplication),...
  - Set theories: ZFC, NBG (ZFC with proper classes),...
  - Order theories: Partial orders, total orders, well-orders, real closed fields,...
  - Any theory with a finite list of axioms and a finite signature.
- Non-examples: any complete undecidable theory, e.g. True Arithmetic - all statements true in $\mathbb{N}$.

Gödel numbering

- Suppose T is effectively axiomatisable. Then T has countably many symbols (logical symbols, plus symbols from σ, plus variables) so we can associate to each symbol a positive natural number. If ϕ is a sentence then we can associate a number $\ulcorner \phi \urcorner$ by taking prime powers.
- For example if we say $\ulcorner 0 \urcorner = 1$, $\ulcorner + \urcorner = 2$ and $\ulcorner = \urcorner = 3$ then we get $\ulcorner 0 + 0 = 0 \urcorner = 2^1 \cdot 3^2 \cdot 5^1 \cdot 7^3 \cdot 11^1 = 339,570$. The number $\ulcorner \phi \urcorner$ is called the **Gödel number** of ϕ.
- We can code up proofs in T similarly. Whether or not a number codes a proof can be checked algorithmically.
- If T can also talk about arithmetic (e.g. if T is ZFC or PA), then we can do this encoding *inside* T, and so T can talk about itself.

Digression on the incompleteness theorems 1

Gödel's First Incompleteness Theorem

If T is any consistent effectively axiomatisable theory which contains enough arithmetic, then there is a sentence G_T, the **Gödel sentence**, which is neither provable nor disprovable in T.

Proof idea:

- Define a predicate $NP(n)$ to mean 'T does not prove the sentence with Gödel number n'.
- use a clever diagonalisation argument to show that for any predicate Q, there is a sentence ϕ such that $T \vdash (\phi \leftrightarrow Q(\lceil \phi \rceil)$.
- Apply the above to the predicate NP to obtain a sentence G_T such that $T \vdash (G_T \leftrightarrow NP(\lceil G_T \rceil))$

Digression on the incompleteness theorems 2

So informally, G_T is true if and only if T does not prove G_T. It follows that G_T is neither provable nor disprovable, and in reasonable circumstances[2] must be true.

Gödel's Second Incompleteness Theorem

If T is any consistent effectively axiomatisable theory which contains enough arithmetic, then T does not prove its own consistency.

Proof idea:

- Define the sentence $\mathrm{Con}(T)$ to mean $NP(\ulcorner 0 = 1 \urcorner)$.
- Code up the proof of the first incompleteness theorem *inside* T to see that $T \vdash (\mathrm{Con}(T) \rightarrow G_T)$.
- So if $T \vdash \mathrm{Con}(T)$ then $T \vdash G_T$, which is a contradiction. So T can't prove $\mathrm{Con}(T)$.

---

[2] For example, in models of $T + \mathrm{Con}(T)$.

Back to the magic polynomial

- Consider your favourite contradiction $\perp$ of T. This might be $\forall x(x \neq x)$, or if T contains enough arithmetic it might be $0{=}1$.

Back to the magic polynomial

- Consider your favourite contradiction $\perp$ of T. This might be $\forall x(x \neq x)$, or if T contains enough arithmetic it might be $0{=}1$.
- Since T is effectively axiomatisable, given a natural number n we can algorithmically check whether n encodes a proof of $\perp$.

Back to the magic polynomial

- Consider your favourite contradiction $\perp$ of T. This might be $\forall x(x \neq x)$, or if T contains enough arithmetic it might be 0=1.
- Since T is effectively axiomatisable, given a natural number n we can algorithmically check whether n encodes a proof of $\perp$.
- This gives us a r.e. set S such that $n \in S$ if and only if n codes for a proof of $\perp$ in T. The MRDP theorem tells us that S must be diophantine.

Back to the magic polynomial

- Consider your favourite contradiction $\perp$ of T. This might be $\forall x(x \neq x)$, or if T contains enough arithmetic it might be $0{=}1$.
- Since T is effectively axiomatisable, given a natural number n we can algorithmically check whether n encodes a proof of $\perp$.
- This gives us a r.e. set S such that $n \in S$ if and only if n codes for a proof of $\perp$ in T. The MRDP theorem tells us that S must be diophantine.
- So any effectively axiomatisable theory T has an associated polynomial p_T that has solutions in $\mathbb{N}$ if and only if T is inconsistent. One can write down such a p_T algorithmically from the axioms of T. There are lots of different polynomials, since e.g. they depend on our choice of Gödel numbering.

What does T know about p_T?

- Let's suppose that T is consistent and contains enough arithmetic (so it satisfies the hypotheses of the Incompleteness Theorems). In particular p_T has no roots.
- We can code up the previous proof inside T to see that $T \vdash$ ('p_T has no roots' $\leftrightarrow \mathrm{Con}(T)$). Hence $T \nvdash$ 'p_T has no roots', otherwise T would prove $\mathrm{Con}(T)$.
- But if T doesn't prove a sentence ϕ, there must be models where ϕ is false. In particular there are models of T (necessarily models of $T + \neg\mathrm{Con}(T)$) where p_T has a root!
- In such a model, any root of p_T must necessarily be a nonstandard natural number, corresponding to a proof of $\perp$ of nonstandard length.

References

- Martin Davis, *Hilbert's Tenth Problem is Unsolvable*, The American Mathematical Monthly, March 1973.
- James P. Jones, Daihachiro Sato, Hideo Wada and Douglas Wiens, *Diophantine representation of the set of prime numbers*, The American Mathematical Monthly, June 1976.
- Panu Raatikainen, *Gödel's Incompleteness Theorems*, The Stanford Encyclopedia of Philosophy, 2015, available at `https://plato.stanford.edu/archives/spr2015/entries/goedel-incompleteness/`