Hilbert's Tenth Problem: Refinements and Variants

William Gasarch^{*}

Abstract

Hilbert's 10th problem, stated in modern terms, is

Find an algorithm that will, given $p \in \mathbb{Z}[x_1, \ldots, x_n]$, determine if there exists $a_1, \ldots, a_n \in \mathbb{Z}$ such that $p(a_1, \ldots, a_n) = 0$.

Davis, Putnam, Robinson, and Matijasevič showed that there is no such algorithm. We look at what happens (1) for fixed degree and number of variables, (2) for particular equations, and (3) for variants which reduce the number of variables needed for undecidability results.

1 This Column's Origin

This column is a short version of a long version of an article based on a blog. What? I give the complete history.

- 1. On May 4, 2020 I wrote a blog about Hilbert's 10th problem: This blog caught the attention of Thomas Erlebach who invited me to write a full article for *The Bulletin of the European Association for Theoretical Computer Science (BEATCS)* on this topic.
- 2. The article: Hilbert's Tenth Problem for Fixed d and n appeared in the Bulletin of the European Association for Theoretical Computer Science (BEATCS), Vol 133, February 2021.
- 3. After it appeared I made a few updates to my copy of the article, added a few whole new sections, and posted this longer version on arxiv.

https://arxiv.org/abs/2104.07220

- 4. The article is (1) full of open problems so quite appropriate for this column, but (2) over 20 pages long so not appropriate for this column. Hence I shortened it.
- 5. It is my intent that you read this and then, if you want further enlightenment, read the long version on arxiv.

^{*}The University of Maryland at College Park, gasarch@umd.edu

2 Hilbert's Tenth Problem

In 1900 Hilbert proposed 23 problems for mathematicians to work on over the next 100 years (or longer). The 10th problem, stated in modern terms, is

Find an algorithm that will, given $p \in \mathbb{Z}[x_1, \ldots, x_n]$, determine if there exists $a_1, \ldots, a_n \in \mathbb{Z}$ such that $p(a_1, \ldots, a_n) = 0$.

Hilbert probably thought this would inspire much deep number theory, and it did inspire some. But the work on this problem took a very different direction. Davis, Putnam, and Robinson [DPR61] showed that determining if an exponential diophantine equation has a solution in \mathbb{Z} is undecidable. Their proof coded Turing machines into such equations. Matijasevič [Mat70] extended their work by showing how to replace the exponentials with polynomials. Hence the algorithm that Hilbert wanted is not possible. For a self contained proof from soup to nuts see Davis' exposition [Dav73]. For more about both the proof and the implications of the result see the book of Matijasevič [Mat93].

In Section 3 we will relate the problem of seeking solutions in \mathbb{Z} with the problem of seeking solutions in \mathbb{N} . In Section 4 we will look at what is known for fixed d, n both for solutions over \mathbb{N} and solutions in \mathbb{Z} . In Section 5 we will discuss classes of polynomials with other conditions added. In Section 6 we discuss variants of Hilbert's 10th problem that lead to getting undecidability results with polynomials in fewer variables. In Section 7 we will briefly present Matijasevič's discussion of what Hilbert really wanted in contrast to what happened.

3 Definitions and Reconciling \mathbb{N} with \mathbb{Z}

The undecidability results are usually proven about solutions in \mathbb{N} . The decidability results are usually proven about solutions in \mathbb{Z} . Hence we use the following notation to keep this straight.

Notation 3.1

- 1. $H\mathbb{Z}(d, n)$ is the problem where the degree is $\leq d$, the number of variables is $\leq n$, and we seek a solution in \mathbb{Z} .
- 2. $H\mathbb{N}(d, n)$ is the problem where the degree is $\leq d$, the number of variables is $\leq n$, and we seek a solution in \mathbb{N} .
- 3. $H\mathbb{Z}(d, n) = D$ means that there is an algorithm to decide $H\mathbb{Z}(d, n)$.
- 4. $H\mathbb{Z}(d, n) = U$ means that there is no algorithm to decide $H\mathbb{Z}(d, n)$.
- 5. Similarly for $H\mathbb{N}(d, n)$ equal to D or U.

In the next section we summarize what is known about $H\mathbb{N}(d, n)$.

4 What Happens for Fixed d, n?

The original motivation for my study of this issue was to obtain what is known for fixed d, n. When I blogged about this I noted that there was no website on the question of particular d, n.

Timothy Chow offered this speculation in an email to me: One reason there isn't already a website of the type you envision is that from a number-theoretic (or decidability) point of view, parameterization by degree and number of variables is not as natural as it might seem at first glance. The most fruitful lines of research have been geometric, and so geometric concepts such as smoothness, dimension, and genus are more natural than, say, degree. A nice survey by a number theorist is the book Rational Points on Varieties by Bjorn Poonen [Poo17]. Much of it is highly technical; however, reading the preface is very enlightening. Roughly speaking, the current state of the art is that there is really only one known way to prove that a system of Diophantine equations has no rational solution.

Even so, I am still curious. Hence we proceed.

4.1 When is $H\mathbb{N}(d, n) = U$? $H\mathbb{Z}(d, n) = U$?

In 1980 Jones [Jon80] announced 16 pairs (d, n) for which $H\mathbb{N}(d, n) = U$. In 1982 Jones [Jon82] provided proofs for 13 of these pairs (12 in Theorem 4 and 1 in Section 3). I emailed Jones about the other three and he emailed back the following:

- Those with d < 2668 have proofs similar to the (4,58) case. This was carried out by Dr. Hideo Wada. (No reference is given.)
- The pair with a very large value of d can be obtained using many relation-combining theorems, like the one at the end of the 1982 paper, which allow one to define two squares with one unknown.

In the theorem below we present all 16 statements from the Jones-1980 paper along with a result by Sun [Sun20] from 2020. We note (1) which three do not have proofs in Jones-1982 (though based on Jones's email we are sure the results are true), and (2) the result of Sun. We state the results of the form $H\mathbb{N}(d, n) = U$ and then apply known theorems (which can be found in the arxiv version of this column) to obtain results of the form $H\mathbb{Z}(d', n') = U$ (except for Sun's result which is already about $H\mathbb{Z}$).

The proofs involve very clever use of elementary number theory to get the degrees and number-of-variables reduced.

In some of the results there are absurdly large numbers like 4.6×10^{44} . These are probably upper bounds that might be able to be lowered with a careful examination of the proofs. These large numbers only occur as d since the main concern was to get the number of variables down.

Theorem 4.1

- 1. $H\mathbb{N}(4, 58) = U$ hence $H\mathbb{Z}(8, 174) = U$.
- 2. $H\mathbb{N}(8, 38) = U$ hence $H\mathbb{Z}(16, 114) = U$.

- 3. HN(12, 32) = U hence HZ(24, 96) = U.
- 4. HN(16, 29) = U hence HZ(32, 87) = U. (Not proven in Jones-1982.)
- 5. $H\mathbb{N}(20, 28) = U$ hence $H\mathbb{Z}(40, 84) = U$.
- 6. $H\mathbb{N}(24, 26) = U$ hence $H\mathbb{Z}(48, 78) = U$.
- 7. $H\mathbb{N}(28, 25) = U$ hence $H\mathbb{Z}(56, 75) = U$.
- 8. HN(36, 24) = U hence HZ(72, 72) = U. (Not proven in Jones-1982.)
- 9. $H\mathbb{N}(96, 21) = U$ hence $H\mathbb{Z}(192, 63) = U$.
- 10. $H\mathbb{N}(2668, 19) = U$ hence $H\mathbb{Z}(5336, 57) = U$.
- 11. $H\mathbb{N}(200000, 14) = U$ hence $H\mathbb{Z}(400000, 42) = U$ and $H\mathbb{Z}(31 \times 2^{14}, 30) = U$.
- 12. $H\mathbb{N}(6.6 \times 10^{43}, 13) = U$ hence $H\mathbb{Z}(13.2 \times 10^{43}, 28) = U$. (Not proven in Jones-1982.)
- 13. $H\mathbb{N}(1.3 \times 10^{44}, 12) = U$ hence $H\mathbb{Z}(2.6 \times 10^{44}, 36) = U$.
- 14. $H\mathbb{N}(4.6 \times 10^{44}, 11) = U$ hence $H\mathbb{Z}(9.2 \times 10^{44}, 24) = U$.
- 15. $H\mathbb{N}(8.6 \times 10^{44}, 10) = U$ hence $H\mathbb{Z}(17.2 \times 10^{44}, 22) = U$.
- 16. $H\mathbb{N}(1.6 \times 10^{45}, 9) = U$ hence $H\mathbb{Z}(3.2 \times 10^{45}, 20) = U$. (Jones' 1982 paper presents the proof of this result and credits it to Matijasevič.)
- 17. $H\mathbb{Z}(d, 11) = U$ for some d. The number d is not stated. (This is due to Sun [Sun20].)

4.2 When is $H\mathbb{Z}(d, n) = D$? $H\mathbb{N}(d, n) = D$?

We present statements of the few cases where we know $H\mathbb{Z} = D$ or $H\mathbb{N} = D$. See the arxiv version for details and references.

Theorem 4.2

- 1. For all d, $H\mathbb{Z}(d, 1) = D$ and $H\mathbb{N}(d, 1) = D$. There is an algorithm that finds all of the integer roots (which may be the empty set).
- 2. For all n, $H\mathbb{Z}(1, n) = D$.
- 3. For all n, $H\mathbb{N}(1, n) = D$.
- 4. $H\mathbb{Z}(2,2) = D$.
- 5. $H\mathbb{N}(2,2) = D$.
- 6. For all n, $H\mathbb{Z}(2, n) = D$ and $H\mathbb{N}(2, n) = D$.

4.3 The Curious Case of $H\mathbb{Z}(3,2)$

We give evidence that $H\mathbb{Z}(3,2) = D$; however, this is still open.

Def 4.3 An element of $\mathbb{Q}[x_1, \ldots, x_n]$ is absolutely irreducible if it is irreducible over \mathbb{C} . For example,

 $x^2 + y^2 - 1$ is absolutely irreducible, but $x^2 + y^2 = (x + iy)(x - iy)$ is not.

A combination of results by Baker and Cohen [BC70], Poulakis [Pou93], and Poulakis [Pou02] imply the following theorem:

Theorem 4.4 There is an algorithm which, given any absolutely irreducible polynomial $P(x, y) \in \mathbb{Z}[x, y]$ of degree 3, determines all integer solutions of the equation P(x, y) = 0. (See Poulakis [Pou02] for a more precise definition of "determines all integer solutions" in the case that there are an infinite number of them.)

The original algorithm (from Baker and Coates) is not practical; however, Pethő et al. [PZGH70] and Stroker-Tzankis [ST03] have practical algorithms. There is also an algorithm for solving a large class of cubic equations implemented in SageMath.

So why isn't $H\mathbb{Z}(3,2) = D$? Because the case where P(x,y) has degree 3 but is not absolutely irreducible is still open.

5 Particular Equations

5.1 If the Variables Are Separated...

Ibarra and Dang [ID06] proved the following.

Def 5.1 $P(z_1, \ldots, z_n)$ is a *Presburger Relation* if it can be expressed with \mathbb{Z} , =, +, <, and the usual logical symbols. For example

 $(z_1 + z_2 < z_3 + 12) \land (z_1 + z_4 = 17)$ is a Presburger formula, but $z_1 z_2 = 13$ is not.

Theorem 5.2 *The following is decidable:* **Instance**

(1) For $1 \le i \le k$, polynomials $p_i(y) \in \mathbb{Z}[y]$, and linear functions $F_i(\vec{x}), G_i(\vec{x}) \in \mathbb{Z}[x_1, \ldots, x_n]$, and (2) a Presburger relation $R(z_1, \ldots, z_k)$. Question Does there exist y, \vec{x} such that

$$R(p_1(y)F_1(\vec{x}) + G_1(\vec{x}), \dots, p_k(y)F_k(\vec{x}) + G_k(\vec{x}))$$

holds?

5.2 The Curious Case of $x^3 + y^3 + z^3 = k$

Rather than looking at $H\mathbb{Z}(d, n)$ let's focus on one equation that has gotten a lot of attention:

$$x^3 + y^3 + z^3 = k.$$

It is easy to show that, for $k \equiv 4, 5 \pmod{9}$, there is no solution in \mathbb{Z} . What about for $k \not\equiv 4, 5 \pmod{9}$?

- 1. Heath-Brown [HB92] conjectured that there are an infinite number of $k \not\equiv 4, 5 \pmod{9}$ for which there is a solution in \mathbb{Z} . Others think that, for all $k \not\equiv 4, 5 \pmod{9}$, $x^3 + y^3 + z^k = k$ has a solution in \mathbb{Z} .
- 2. Elkies [Elk00] devised an efficient algorithm to find solutions to

 $x^3 + y^3 + z^3 = k$ if there is a bound on x, y, z.

3. Elsehans and Jahnel [EJ09] modified and implemented Elkies algorithm and determined the following: The only $k \leq 1000, k \not\equiv 4,5 \mod 9$, where they did not find a solution were

33, 42, 74, 114, 165, 390, 579, 627, 633, 732, 795, 906, 921, and 975.

Their work, and the work of all the items below, required hard mathematics, clever computer science, and massive computer time.

- 4. Huisman [Hui16] found a solution for k = 74. For many other values of k where there were solutions, Huisman found additional solutions.
- 5. Booker [Boo19] found a solution for k = 33.
- 6. Booker found solutions for k = 42 and k = 795. These have not been formally published yet; however, the x, y, z can be found on the Wikipedia site:

```
https://en.wikipedia.org/wiki/Sums_of_three_cubes
```

7. As of April 2021 (when this article was written) the only $k \leq 1000, k \not\equiv 4, 5 \mod 9$, where no solution is known are:

114, 165, 390, 579, 627, 633, 732, 906, 921, and 975.

Consider the function that, on input k, determines if $x^3 + y^3 + z^3 = k$ has a solution in \mathbb{Z} . Is this function computable?

- 1. I suspect the function is computable. Why? What would a proof that this function is not computable look like? It would have to code a Turing machine computation into a very restricted equation. This seems unlikely to me. Note also that it may be the case the equation has a solution for every $k \not\equiv 4, 5 \pmod{9}$, in which case the decision problem is not just decidable—it's regular!
- 2. Daniel Varga has suggested there may be a proof that does not go through Turing machines. Perhaps some other undecidable problem? Also, there may be new techniques we just have not thought of yet.

6 Variants that Use Fewer Variables

Hilbert's 10th problem, and the restrictions on it in this article, are about the solvability of the following problem: Given $p(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$, is the following true over \mathbb{Z} :

$$(\exists x_1)\cdots(\exists x_n)[p(x_1,\ldots,x_n)=0].$$

(Undecidability results were usually about truth over \mathbb{N} .)

There has been much work in getting the number of variables needed for an undecidability result to be small. As we saw in Theorem 4.1, $HN(1.6 \times 10^{45}, 9) = U$. As of April 2021 (when this was written) 9 is the lowest *n* such that there is known to be a *d* with HN(d, n) = U. The result of 9 was proven by Matijasevič in the early 1980's (it appears in Jones [Jon82] and credited to Matijasevič). Hence the 9 has not been improved in 29 years. I doubt it will be improved between writing this paper and the appearance of this paper. As we saw in Theorem 4.1, there is a *d* such that $H\mathbb{Z}(d, 11) = U$. This was proven in 2020 so it is plausible to be improved in the near future.

We explore some variants of H10 where the number of variables needed is smaller than 9 (for \mathbb{N}) and 11 (for \mathbb{Z}).

6.1 Different Quantifier Prefixes

Let $Q_1 \cdots Q_n$ be a string of quantifiers. Consider the following problem. Given $p(x_1, \ldots, x_n)$ is

$$(Q_1x_1)\cdots(Q_nx_n)[p(x_1,\ldots,x_n)=0]$$

true over \mathbb{Z} ? Over \mathbb{N} ?

Notation 6.1

- 1. Let $Q_1 \cdots Q_n$ be a string of quantifiers. $Q_1 \cdots Q_n$ is undecidable over \mathbb{N} if the above problem is undecidable over \mathbb{N} . Similar for \mathbb{Z} .
- 2. A quantifier is *bound* if there is an explicit upper and lower bound on it which is a polynomial in the prior variables.

Recall from Theorem 4.1 that ∃⁹ is undecidable over N and that this is the best known. Matijasevič [Mat72] showed that ∃∀∃², with ∀ bounded, over N, is undecidable. From this result one can obtain undecidability with polynomials of four variables. This is much better than nine. See Sun [Sun21] for more of history, references, and results about quantifier prefixes and undecidability over N.

Recall from Theorem 4.1 that \exists^{11} is undecidable over \mathbb{Z} and that this is the best known. Sun [Sun21] proved the following.

These are undecidable over Z: ∀∃⁷, ∀²∃⁴, ∃∀∃⁴, ∃∀∃³, ∃²∀∃³, ∀∃∀∃³, ∀∃∀∃³, ∀∃∀²∃², ∀²∃∀²∃², ∀²∃∀²∃², ∃∀∃∀²∃², ∃∀∃∀²∃², ∃∀⁶∃². Note that the shortest prefixes only use 6 variables which is much better than 11.

These are undecidable if the ∀ are bounded: ∃∀∃³, ∃∀²∃³, ∃²∀²∃³, ∃²∀²∃², ∃²∀∃∀∃², ∃∀⁵∃². Note that the shortest prefixes only use 5 variables which is much better than 11.

6.2 Sets of Polynomials

Matijasevič and Robinson [MR96] (see also Matijasevič [Mat72]) prove the following (All quantifiers are over \mathbb{N}). Let A be an r.e. set.

1. There exist 3n polynomials

 ${P_i(x_1, x_2, x_3)}_{i=1}^n, {Q_i(x_1, x_2, x_3)}_{i=1}^n, {R_i(x_1, x_2, x_3)}_{i=1}^n$ such that

 $a \in A$ iff

$$(\exists b, c) \bigwedge_{i=1}^{n} (\exists d) [P_i(a, b, c) < Q_i(a, b, c) \times d < R_i(a, b, c)].$$

From this result one can obtain a problem with polynomials in 3 variables that is undecidable.

2. There exist polynomials

 $P(x_1, x_2, x_3)$ and $Q(x_1, x_2, x_3, x_4)$ such that

 $a \in A$ iff

$$(\exists b, c)(\forall f)[(f \le P(a, b, c)) \implies (Q(a, b, c, f) > 0)]$$

From this result one can obtain a problem with polynomials in 3 variables that is undecidable.

7 What Would Hilbert Do?

Def 7.1 HQ(d, n) is the problem where the degree is $\leq d$, the number of variables is $\leq n$, and we seek a solution in Q.

Matijasevič [Mat] (Page 18) gives good reasons why Hilbert might have actually wanted to solve H \mathbb{Q} . Hilbert stated the tenth problem as H \mathbb{Z} ; however, if H \mathbb{Z} is solvable then H \mathbb{Q} is solvable. He might have thought that the best way to solve H \mathbb{Q} is to solve H \mathbb{Z} .

What is the status of $H\mathbb{Q}$ now? It is an open question to determine if $H\mathbb{Q}$ is decidable. Hence the problem Hilbert plausibly intended to ask is still open and may yet lead to number theory of interest, which was his intent.

8 Acknowledgement

We thank Blogger vzn, Timothy Chow, Thomas Erlebach, Lance Fortnow, Brogdan Grechuk, Nathan Hayes, James Jones, Emily Kaplitz, Chris Lastowski, David Marcus, Yuri Matijasevič, Andras Salamon, Yuang Shen, Joshua Twitty, Larry Washington, Daniel Varga, Zan Xu, for helpful discussions.

We are particularly grateful to the following people.

- 1. Timothy Chow for his comments in Section 4 and help with the discussion in Section 5.2 of $x^3 + y^3 + z^3 = k$.
- 2. Brogdan Grechuk for telling us about the material that is now in Section 4.3.
- 3. James Jones for discussion of Theorem 4.1.
- 4. Yuri Matijasevič for pointing us to many results of which we were unaware.

References

[BC70]	Alan Baker and John Coates. Integer points on curves of genus 1. Mathematical Proceedings of the Cambridge Philosophical Society, 67:595–602, 1970.
[Boo19]	Andrew Booker. Cracking the problem with 33, 2019. https://arxiv.org/abs/1903.04284.
[Dav73]	Martin Davis. Hilbert's tenth problem is unsolvable. American Mathematical Monthly, pages 233-2695, 1973. https://www.math.umd.edu/~laskow/Pubs/713/Diophantine.pdf.
[DPR61]	Martin Davis, Hillary Putnam, and Julia Robinson. The decision problem for exponential diophantine equations. <i>Annals of Mathematics</i> , 74:425–436, 1961.
[EJ09]	Andreas-Stephan Elsehans and Jorg Jahnel. New sums of three cubes. <i>Mathematics of Computation</i> , 78(266), 2009. https://www.ams.org/journals/mcom/2009-78-266/ S0025-5718-08-02168-6/S0025-5718-08-02168-6.pdf.
[Elk00]	Noam Elkies. Rational points near curves and small nonzero $ x^3 - y^3 $ via lattice reductions. In Algorithmic Number Theory, volume 1838 of Lecture Notes in Computer Science, pages 33–63, 2000. https://arxiv.org/abs/math/0005139.
[HB92]	Roger Heath-Brown. The density of zeros of forms for which weak approximation fails. <i>Mathematics of Computation</i> , 59(200):612–623, 1992.
[Hui16]	Sander Huisman. Newer sums of three cubes, 2016. https://arxiv.org/abs/1604.07746.

- [ID06] Oscar Ibarra and Zhe Dang. On the solvability of a class of diophantine equations and applications. *Theoretical Computer Science*, 352:342–346, 2006.
- [Jon80] James Jones. Undecidable diophantine equations. Bulletin of the American Mathematical Society, 3(2):859-862, 1980. https://www.ams.org/journals/bull/1980-03-02/ S0273-0979-1980-14832-6/S0273-0979-1980-14832-6.pdf.
- [Jon82] James Jones. Universal diophantine equations. Journal of Symbolic Logic, 47(3):549-571, 1982. http://www.jstor.org/stable/2273588?seq=1#metadata_info_tab_ contents.
- [Mat] Yuri Matijasevič. Hilbert's tenth problem: What can we do with diophantine equations? https://logic.pdmi.ras.ru/~yumat/personaljournal/H10history/ H10histe.pdf.
- [Mat70] Yuri Matijasevič. Enumerable sets are diophantine (Russian). Doklady Academy Nauk, SSSR, 191:279–282, 1970. Translation in Soviet Math Doklady, Vol 11, 1970.
- [Mat72] Yuri Matijasevič. Arithmetical representations of enumerable sets with a small number of quantifiers. Zapiski Nauchnykh Seminarov Leningradskogo Otdeleniya Matematicheskogo, 32, 1972. Google the name and you can probably find a translation into English.
- [Mat93] Yuri Matijasevič. *Hilbert's Tenth Problem*. MIT Press, Cambridge, 1993.
- [MR96] Yuri Matijasevič and Julia Robinson. Two universal 3-quantifier representations of recursively enumerable sets. In *The collected works of Julia Robinson*, 1996. Earlier version appeared in Russian in a collection of papers dedicated to A. A. Markov, in 1974. https://arxiv.org/abs/0802.1052.
- [Poo17] Bjorn Poonen. Rational points on varieties, volume 186 of Graduate studies in mathematics. American Mathematical Society, 2017.
- [Pou93] Dimitrios Poulakis. Points entiers sur les courbes de genre 0. Colloquim Mathematicae, 66:1–7, 1993.
- [Pou02] Dimitrios Poulakis. Solving genus zero diophantine equations with at most two infinite valuations. *Journal of Symbolic Computation*, 33:479–491, 2002.
- [PZGH70] Attila Petho, Horst G Zimer, Josef Gebel, and Emanuel Herrmann. Computing all s-integral points on elliptic curves. *Mathematical Proceedings of the Cambridge Philosophical Society*, 127:383–402, 1970.

- [ST03] Roel Stoeker and Nikoluas Tzanakis. Computing all integer solutions of a genus 1 equation. *Mathematics of Computation*, 72(1917–1933), 2003.
- [Sun20] Zhi-Wei Sun. Further results on Hilbert's tenth problem. *Science China Mathematics*, This Journal Does not have Volumes:1–26, 2020.
- [Sun21] Zhi-Wei Sun. Mixed quantifier prefixes over diophantine equations with integer variables, 2021. https://arxiv.org/pdf/2103.08302.pdf.