FIRST ORDER DEFINABILITY OF THE INTEGERS IN THE FIELD OF RATIONALS

YAN KOLEZHITSKIY

Supervised by:
André Nies

380 Project in
Computer Science, S1 2016

Department Of Computer Science
University of Auckland
New Zealand

Abstract. We review and present the proof, as originally due to Julia Robinson 1949, that the integers are definable within the field of rationals using first order logic. We discuss consequences of this result, for instance showing that the theory of the rationals is undecidable. We mention Hilbert's tenth problem, and briefly discuss its relevance to first order definability. Further, we look at Poonen's recent work on definability of the integers in the rationals.

Contents

1. Introduction

> Then called the High Prophets: What seest thou, Imbaun? And Imbaun said: I see naught. Then called the High Prophets: What knowest thou, Imbaun? And Imbaun said: I know naught. Then spake the High Prophet of Eld of All the gods save One, who is first on Earth of prophets: O Imbaun! we have all looked upwards in the Hall of Night towards the secret of Things, and ever it was dark, and the Secret faint and in an unknown tongue. And now thou knowest what all High Prophets know.
>
> Lord Dunsany, The Gods of Pegana

In order to use first order definability, we must first explain what it is. The term 'first order' refers to the type of formal logic used, often simply referred to as 'predicate logic'. Primarily, we say 'first order' to differentiate between logics of the 'first order' (quantifying over variables), and 'second order' (quantifying over relations). We say that a concept is *first order definable*, if there exists some formula $A(x)$ that characterizes exactly the properties thereof. Thus we give a definition of some concept in a first order language. Consider the following example:

$$N(k) = \exists x(x^2 = k) \tag{1}$$

Clearly, if we take the field $\mathbb{R}$, (1) would define the nonnegative reals.

In her PhD thesis 'Definability and Decision Problems in Arithmetic', Section 3 [10], Julia Robinson presented a first order definition of $\mathbb{Z}$ in the field of rationals, $(\mathbb{Q}, +, \times, 0, 1)$. She does so by taking the *second order* definition of $\mathbb{N}$ in $\mathbb{Q}$ as the intersection of all inductive sets, and shows that one can convert this into a first order definition by presenting a first order definition with parameters of a sufficiently large subclass of the inductive set.

This has very powerful implications. For starters, since it is a fact that the natural numbers are first order definable in the integers, and since the theory of natural numbers is undecidable, one can conclude that the theory of rationals itself is undecidable.

Most of this essay will be dedicated to the proof of Julia Robinson's theorem, as the author has found that her paper assumes many results without giving sufficient proofs that can be understood by those outside of the discipline of number theory. Near the end we will look at Hilbert's tenth problem and its relevance to Robinson's work. We will also briefly mention Bjorn Poonen's method of defining the integers within rationals, which is in many ways more efficient, as judged by the number of quantifier alternations. Also Jochen Koenigsmann's results a briefly mentioned, who provides similar results, but only using one universal qantifier.

2. Defining $\mathbb{Z}$ in $\mathbb{Q}$

2.1. **Overview.**

Theorem 2.1. *The set of integers $\mathbb{Z}$ is first order definable in the field $\mathbb{Q}$.*

We first consider what it means for some subset S of $\mathbb{Q}$ to be *inductive*. S is inductive if $0 \in S$, and for all y, if $y \in S$, then $y + 1 \in S$. This is enough to obtain a second order definition of $\mathbb{N}$ in $\mathbb{Q}$, even without using multiplication in $\mathbb{Q}$:

$$k \in \mathbb{N} \Leftrightarrow \forall S(S \text{ is inductive } \rightarrow k \in S) \tag{2}$$

In other words, if we take all the inductive sets S over $\mathbb{Q}$, then k is a natural number iff it appears in $\bigcap S$.

The crux of Robinson's definition comes in the idea that we can construct a smaller collection of inductive sets S for the same purpose. This is done via the use of certain constructs of number theory (ie quadratic forms) parameterizable by only two rationals a and b. We consider sets $S_{a,b} = \{k \mid \mathbb{Q} \vDash \Phi(a,b,k)\}$ where $\Phi(a,b,k)$ is:

$$\Phi(a,b,k) \equiv \exists x \exists y \exists z(2 + abk^2 + bz^2 = x^2 + ay^2) \tag{3}$$

Using this way of talking about S, we can take the second order definition (2) and replace the quantification over sets S by the quantification over the parameters a and b, thereby turning it into a first order definition. We define

$$A(k) \equiv \forall a \forall b[(\Phi(a,b,0) \wedge \forall m \langle \Phi(a,b,m) \rightarrow \Phi(a,b,m+1)\rangle) \rightarrow \Phi(a,b,k)] \tag{4}$$

In other words, we have it that k is an integer iff for all inductive sets $S_{a,b}$, $k \in S_{a,b}$. Further, we say that:

$$k \in \mathbb{Z} \Leftrightarrow \mathbb{Q} \vDash A(k) \tag{5}$$

The reason why we replace $\mathbb{N}$ with $\mathbb{Z}$ is because the smaller number of inductive sets we now use cannot be ridded of the negative integers; we note that in (3), k appears in the form of k^2, and thus we have $k \in S_{a,b}$ iff $-k \in S_{a,b}$.

We now set out to prove (5).

First, for convenience we take the hypothesis of (5), and call it $B(a,b)$, so:

$$B(a,b) \equiv (\Phi(a,b,0) \wedge \forall m \langle \Phi(a,b,m) \rightarrow \Phi(a,b,m+1)\rangle) \tag{6}$$

First we must take a closer look at the sets $S_{a,b}$ and prove that they are indeed inductive. Any $k \in \mathbb{Q}$ can be written as $\frac{n}{d}$, where $n \in \mathbb{Z}$, $d \in \mathbb{N}$, and n and d are co-prime; we say that $k = \frac{n}{d}$ *in its lowest terms*, where d is the *denominator of* k *in its lowest terms.*

Fact 2.2. *Let S be a set of rationals. Say S is defined by some condition that holds for 0 and only depends on the denominator of the rational in lowest terms. Then S is inductive.*

Proof. We use the fact that for all $q \in \mathbb{Q}$ the denominator in lowest terms of q is the same as that of $q+1$. In more detail, say $q \in \mathbb{Q}$, and $q = \frac{n}{d}$ in its lowest terms. Then $q+1 = \frac{n}{d}+1 = \frac{n}{d}+\frac{d}{d} = \frac{n+d}{d}$. Since d and n are co-prime, so are $n+d$ and d. □

Now all that needs to be done is to show that (3) satisfies the conditions of (2.2). Robinson does this by showing that there are two specific choices of a and b such that the condition $\Phi(a,b,k)$ for $S_{a,b}$ only depends on the denominator of k in its lowest terms.

The first such choice is of a being a prime p equivalent to $3 \mod 4$, and b being 1.

Lemma 2.3. *If p is a prime and $p \equiv 3 \mod 4$, then the equation $2 + pM^2 + pZ^2 = X^2 + Y^2$ has a solution for X, Y, and Z iff the denominator of M in its lowest terms is odd, and is co-prime to p.*

For the second choice, b is taken to be some prime p equivalent to $1 \mod 4$, and a is determined by b to be some prime q such that the Legendre symbol (q/p) evaluates to -1 (see the glossary at the end for number theoretic terms).

Lemma 2.4. *If p and q are odd primes, where $p \equiv 1 \mod 4$ and $(q/p) = -1$, then $2 + pqM^2 + pZ^2 = X^2 + qY^2$ has a solution for X, Y, and Z iff the denominator of M in its lowest terms is co-prime to both q and p.*

In light of the above, we also need a number theoretic claim that shows that for any such p as discussed above, we can find such a q:

Claim 2.5. *If p is a prime, and $p \equiv 1 \mod 4$, there exists an odd prime q such that $(q/p) = -1$.*

Proof. Take some s that is a quadratic non-residue of p. Either s is odd, or $s+p$ is odd, and clearly both still are non-residues of p. But an odd non-residue of p must have an odd prime factor which is also a non-residue of p, as the Legendre symbol is multiplicative . Call this factor q. □

2.2. **Proof of (5) assuming Lemmas 2.3 and 2.4.** We first consider the left to right direction in (5). We take some $k \in \mathbb{Q}$ such that k is also in $\mathbb{Z}$. Then without loss of generality we take some a and b that satisfy (6). If this were not the case, then trivially $\mathbb{Q} \vDash A(k)$.

There are two cases, either k is positive, or it is not. Presuming that k is positive, we can use the hypothesis to initiate a proof by induction. Since we already know that $\mathbb{Q} \vDash \Phi(a,b,0)$, and we know that for all m, if $\mathbb{Q} \vDash \Phi(a,b,m)$, then $\mathbb{Q} \vDash \Phi(a,b,m+1)$, trivially $\mathbb{Q} \vDash \Phi(a,b,k)$.

If, on the other hand, k is negative, then we have the following; first we recall that $k \in S_{a,b}$ if and only if $-k \in S_{a,b}$, which is the same as $\mathbb{Q} \vDash \Phi(a,b,k)$ if and only if $\mathbb{Q} \vDash \Phi(a,b,-k)$. Then, we note, that if k is negative, $-k$ must be positive, and the proof in the case above can be used to show that $\mathbb{Q} \vDash \Phi(a,b,-k)$. Thus $\mathbb{Q} \vDash \Phi(a,b,k)$.

Thus, in both cases, $\mathbb{Q} \vDash A(k)$.

Now we consider the right to left direction in (5). Suppose $k \in \mathbb{Q}$ such that $\mathbb{Q} \vDash A(k)$. We write k in its lowest terms, $\frac{n}{d}$. By Lemma 2.3, d is odd and not divisible by all primes p equivalent to $3 \mod 4$. But also by Lemma 2.4 in conjunction with Claim 2.5, d is not divisible by any prime p that is equivalent to $1 \mod 4$. But that just means that the d must be 1.
Thus $k \in \mathbb{Z}$.

By the two lemmas the formula $A(k)$ provides a definition of $\mathbb{Z}$ in $\mathbb{Q}$. A variant of the proof in the style of natural deduction can be found in the Appendix, Section 7.

3. Proofs of Lemmas 2.3 and 2.4

We first consider two lemmas that are derived from the Hasse-Minkowski theorem, originally presented in [6]. These are non-trivial results, and to prove them would require a discussion of the p-adic integers to an extent that would fall outside of the scope of this project. An introduction to p-adics and the Hasse-Minkowski theorem can be found, for instance, in [2]. The proofs for the below lemmas can be found in Fraïsse[8], Appendix 1.

Lemma 3.1. *If some prime p has the property $p \equiv 3 \mod 4$, then $X^2 + Y^2 - pZ^2$ represents $M \in \mathbb{Q}$, $M \neq 0$, iff it is not the case that either:*
a) $M = pkS^2$, where $(k/p) = 1$.
b) $M = kS^2$ with $k \equiv p \mod 8$
Where S is some rational.

Lemma 3.2. *If p and q are odd primes with $p \equiv 1 \mod 4$ and $(q/p) = -1$, then there is some non-zero M in the rationals such that $X^2 + qY^2 - pZ^2$ represents M, iff it is not the case that:*
a) $M = pkS^2$ and $(k/p) = -1$
b) $M = qkS^2$ and $(k/q) = -1$
Where S is some rational.

Below are the proofs for Lemmas 2.3, 2.4, and 2.5.

Proof of Lemma 2.3. We note that the statement "$2 + pM^2 + pZ^2 = X^2 + Y^2$ has a solution for X, Y, and Z" is equivalent to:

$$X^2 + Y^2 - pZ^2 \text{ represents } 2 + pM^2 \tag{7}$$

We consider $M = n/d$, where d is the denominator in lowest terms. Then, we have (7) if and only if

$$X^2 + Y^2 - pZ^2 \text{ represents } pn^2 + 2d^2 \tag{8}$$

The proof for this is simple: suppose we have one of the two equations, we divide/multiply through by d^2, and then, because X, Y, and Z are variables ranging over the rationals, we get the other. Hence (7) is equivalent to (8)

By considering three cases, we show that the only scenario where (8) holds, is when d is odd, and co-prime to p.

Case 1: (d is even)
Suppose $d = 2a$. We note that u has to be odd, otherwise d would not be the denominator in lowest terms. We write n as $2b+1$. Thus, $pn^2 + 2d^2 = 8a^2 + p(2b+1)^2$. We now show that this integer is equivalent to $p \mod 8$. Clearly this is the case if $p(2b+1)^2 \equiv p \mod 8$, if $8 \mid p(2b+1)^2 - p$, if $8 \mid p(4b^2+4b)$ if $8 \mid 4b(b+1)$, in which case, either b is even or odd, but in both cases we have it that $b(b+1) = 2c(2d+1)$ for some c and d. Thus $4b(b+1) = 8c(2d+1)$ implies that indeed $8 \mid 4b(b+1)$.

Suppose $S = 1$, then $pn^2 + 2d^2$ can be written as k, and consequently S^2k. Since we have already shown that $k \equiv p \mod 8$, by Lemma 3.1 b, we have it that $X^2 + Y^2 - pZ^2$ does not represent $pn^2 + 2d^2$.

Case 2.1: (d is odd and $p \mid d$)
We write d as pa, then $pn^2 + 2d^2 = p(n^2 + 2pa^2)$. We note that $p \nmid n$, as otherwise d would not be the denominator in lowest terms of M. Thus $p \nmid (n^2 + 2pa^2)$. We then note that $((n^2+2pa^2)/p) = (n^2/p) = 1$. We then take $S = 1$, and write $k = (n^2+2pa^2)$. But this means that $pn^2 + 2d^2pkS^2$, where $(k/p) = 1$. But then, by Lemma 3.1(a), we conclude that $X^2 + Y^2 - pZ^2$ does not represent $pn^2 + 2d^2$.

Case 2.2: (d is odd and $p \nmid d$)
Here, we show that given our assumptions, $pn^2 + 2d^2$ does not satisfy conditions a and b of Lemma 3.1. From the case-given assumption, we get $p \nmid 2d^2$, and thus $p \nmid (pn^2 + 2d^2)$. In other words, p is not a factor of $pn^2 + 2d^2$. Hence the latter cannot be written as pkS^2 and thus $pn^2 + 2d^2$ doesn't satisfy condition a of Lemma 3.1.

As d is odd, $d \equiv 1$ or $3 \mod 4$. In both cases, we have it that $d^2 \equiv 1 \mod 4$. We now consider n. Without loss of generality, we know that $n \equiv 0,\ 1,\ 2,$ or $3 \mod 4$, and so n^2 must be equivalent modulo 4 to either $0,\ 1,\ 4,$ or 6. Thus $n^2 \equiv 0$ or $1 \mod 4$. But this means that $pn^2 + 2d^2 \equiv 3*0 + 2*1 \equiv 2 \mod 4$ or $pn^2 + 2d^2 \equiv 3*1 + 2*1 \equiv 1 \mod 4$.

We can now show that $pn^2 + 2d^2$ cannot be written as kS^2, where $k \equiv p \equiv 3 \mod 4$. We begin by pointing out that if $S \in \mathbb{Q} \setminus \mathbb{Z}$, then $kS^2 \in \mathbb{Q} \setminus \mathbb{Z}$. But since p, n, and d are all integers, $S \in \mathbb{Q} \setminus \mathbb{Z}$ would imply that $pn^2 + 2d^2 \neq kS^2$. Ergo we treat S as an element of $\mathbb{Q} \cap \mathbb{Z}$. Without loss of generality, we note that $S \equiv 0,\ 1,\ 2,$ or $3 \mod 4$, and ergo $S^2 \equiv 0$ or $1 \mod 4$. But this means that, since $k \equiv p \equiv 3 \mod 4$, that $kS^2 \equiv 0$ or $3 \mod 4$. Clearly, we have $pn^2 + 2d^2 \neq kS^2$. Ergo in all cases we have shown that $pn^2 + 2d^2$ cannot be written as kS^2, and consequently doesn't satisfy condition b of Lemma 3.1.

Which means that, by aforementioned lemma, we have it that $X^2 + Y^2 - pZ^2$ represents $pn^2 + 2d^2$. This is enough to prove Lemma 2.3.

□

Proof of Lemma 2.4. Similarly to the above proof, we note that the statement '$2 + pqM^2 + pZ^2 = X^2 + qY^2$ has a solution for X, Y, and Z' is equivalent to:

$$X^2 + qY^2 - pZ^2 \text{ represents } 2 + pqM^2 \tag{9}$$

And, if we rewrite M as n/d, where d is the denominator in lowest terms, then (9) is equivalent to :

$$X^2 + qY^2 - pZ^2 \text{ represents } pqn^2 + 2d^2 \tag{10}$$

The rationale here is identical to the one that justifies teh equivalence of (7) and (8).

We consider 2 cases, and in doing that we show that $X^2 + qY^2 - pZ^2$ represents $pqn^2 + 2d^2$ if and only if d is co-prime to both q and p.

Case 1: ($p \nmid d$ and $q \nmid d$)
We note that both p and q don't divide $2d^2$, as both of them are odd primes. Hence, they don't divide $pqn^2 + 2d^2$. But that means that $pqn^2 + 2d^2$ can't be written as pkS^2 or qkS^2, and thus does not satisfy conditions (a) and (b) of Lemma 3.2. Thus by the same lemma, we conclude that $X^2 + qY^2 - pZ^2$ represents $pqn^2 + 2d^2$ in the rationals.

Case 2: ($p \mid d$ or $q \mid d$)
Without loosing generality, we can consider 2 sub-cases here, treating each of the disjuncts of the assumption, however the cases are identical, with q and p being interchangeable, and so we only treat $p \mid d$.

Suppose $p \mid d$. We write d as pa, and consequently $pqn^2 + 2d^2 = p(qn^2 + 2pa^2)$. We note that $p \nmid n$, as otherwise d would not be the smallest denominator of M. Thus $p \nmid qn^2$ and ergo $p \nmid (qn^2 + 2pa^2)$. We now show that $((qn^2 + 2pa^2)/p) = -1$. First, we recall that by assumption, $(q/p) = -1$[1]. Then we observe that $((qn^2 + 2pa^2)/p) = (qn^2/p) = (q/p) = -1$. We call $(qn^2 + 2pa^2)$ k, and set $S = 1$. Then $pqn^2 + 2d^2 = pkS^2$ where $(k/p) = -1$, but this means that by Lemma 3.2, $X^2 + qY^2 - pZ^2$ does not represent $pqn^2 + 2d^2$.

This suffices to prove Lemma 2.4.

□

4. Undecidability of $\mathbb{Q}$

One of the consequences of first order definability of $\mathbb{Z}$ in $\mathbb{Q}$, is that the theory of rationals, referred to here as $Th(\mathbb{Q})$ can now be proven to be *undecidable*. To do this we need the concept of *mapping reducibility*. We say that some set $A \subseteq \mathbb{N}$ is mapping reducible to set $B \subseteq \mathbb{N}$ if there is a computable function $f : \mathbb{N} \to \mathbb{N}$ such that $\forall x, x \in A \Leftrightarrow f(x) \in B$. We denote this by $A \leq_m B$. Sometimes the domain and codomain of f may be things other than numbers, such as wff in predicate logic. This is fine because these object can be effectively encoded by numbers.

[1] this can also be proven by the Law of Quadratic Reciprocity from the other properties that p and q are assumed to have

We recall that $Th(\mathbb{N})$ is undecidable in accord with Gödel's results shown in [4]. We prove that $Th(\mathbb{N}) \leq_m Th(\mathbb{Z})$ and $Th(\mathbb{Z}) \leq_m Th(\mathbb{Q})$. We then use the properties of $\leq_m$ to show that $Th(\mathbb{Q})$ must be undecidable. An introductory discussion of mapping reducibility and its properties can be found in [11].

4.1. **Defining $\mathbb{N}$ in $\mathbb{Z}$.**

We first consider the following formula:

$$C(n) \equiv \exists x_1 \exists x_2 \exists x_3 \exists x_4 (x_1^2 + x_2^2 + x_3^2 + x_4^2 = n) \tag{11}$$

We claim[2] that $\mathbb{Z} \vDash C(n)$ iff $n \in \mathbb{N}$, (Lagrange's four-square theorem) and consequently we get:

Theorem 4.1. *The natural numbers are first order definable in the integers by $C(n)$.*

4.2. $Th(\mathbb{N}) \leq_m Th(\mathbb{Z})$.

We show this by inductively defining a function $f^{N \to Z}$. The induction is done on the complexity of $\phi \in Th(\mathbb{N})$.

If ϕ is a predicate form with free variables $x_1 ... x_n$, then $f^{N \to Z}(\phi) = C(x_1) \wedge ... \wedge C(x_n) \wedge \phi$.

If ϕ is of the form $\forall x \; \psi$, then $f^{N \to Z}(\phi) = \forall x (C(x) \to (f^{N \to Z}(\psi)))$

If ϕ is of the form $\psi \wedge \chi$, then $f^{N \to Z}(\phi) = f^{N \to Z}(\psi) \wedge f^{N \to Z}(\chi)$

If ϕ is of the form $\neg\psi$, then $f^{N \to Z}(\phi) = \neg f^{N \to Z}(\psi)$

All the other cases can be reduced to these. Clearly f is computable.

Claim 4.2. *For all wff ϕ we have $\phi \in Th(\mathbb{N}) \Leftrightarrow f^{N \to Z}(\phi) \in Th(\mathbb{Z})$.*

Thus, by the definition of mapping reducibility, $Th(\mathbb{N}) \leq_m Th(\mathbb{Z})$.

4.3. $Th(\mathbb{Z}) \leq_m Th(\mathbb{Q})$.

Similarly to what we did in 4.2, we inductively define $f^{Z \to Q}$ by considering the first order definition of $\mathbb{Z}$ in $\mathbb{Q}$, (4) instead of $C(n)$ as above. Thus we have the following claim:

Claim 4.3. *For all wff ϕ we have $\phi \in Th(\mathbb{Z}) \Leftrightarrow f^{Z \to Q}(\phi) \in Th(\mathbb{Q})$.*

And again, by the definition of mapping reducibility, we get $Th(\mathbb{Z}) \leq_m Th(\mathbb{Q})$.

4.4. **Undecidability of $\mathbb{Q}$.**

By the results obtained in 4.2 and 4.3, paired with the transitive nature of $\leq_m$, it is easily shown that $Th(\mathbb{N}) \leq_m Th(\mathbb{Q})$. This is equivalent to taking the composition of the two functions $f^{N \to Z}$ and $f^{Z \to Q}$. It has already been established that $Th(\mathbb{N})$ is undecidable. So all we need now is the following theorem taken from [11], (Corollary 5.23) :

Theorem 4.4. *If $A \leq_m B$ and A is undecidable, then B must also be undecidable.*

Ergo $Th(\mathbb{Q})$ is undecidable.

[2] The proof for this can be found in [5], section 20.5

5. First Order Definability and Hilbert's Tenth

5.1. **H10 and History.** Hilbert's Tenth, referred to as H10, was the 10th problem laid out by Hilbert in his talk at the International Congress of Mathematics in Paris, at the start of the 20^{th} century. He hypothesized that these problems will determine the course of mathematics over the next century.

The problem asks as to whether or not there exists an algorithm that can determine if a solution exists in $\mathbb{Z}$ for some Diophantine equation with integer coefficients.

This can be rephrased using first order logic into the following question: does there exist an algorithm that can determine whether or not

$$\mathbb{Z} \vDash \exists x_1 ... \exists x_n P(x_1, \ldots, x_n) = 0,$$

where P is a polynomial in variables $x_1, \ldots, x_n$ with integer coefficients. Clearly, we can see that if $Th(\mathbb{Z})$ was decidable, then the answer to H10 will be a positive one. However as seen in Section 4, $Th(\mathbb{Z})$ is not decidable. This in itself is not enough to determine an answer to H10, as there still exist subsets of $Th(\mathbb{Z})$ that are decidable.

5.2. **The solution.** Robinson, among other mathematicians, dedicated a lot of time in the pursuit of an answer to H10. In fact, in 1961 Putnam, Davis, and Robinson proved that there is no algorithm that can decide an exponential Diophantine equation over $\mathbb{N}$, ie an equation such as $5^x + z^y = 0$. In 1970, a Russian mathematician called Matiasevic was able to use this to prove a negative result for H10. He did this by using various number theoretic notions and Pell's equation to get rid of the variables in the exponents.

5.3. **From the Perspective of Computability Theory.** Another way of talking about H10 is through Computability Theory, and the recognizability of elements as the members of a language [11]. Under this view, we would ask if there is a Turing Machine (formalized algorithm) that can recognize a given Diophantine equation as a member of the set of all Diophantine equations with a solution in $\mathbb{Z}$. The result of Matiasevic's proof was that Diophantine sets are equivalent to *recursively enumerable* sets. If S is such a set, there is a polynomial with integer coefficients Q such that for each natural number n,

$$n \in S \Leftrightarrow \mathbb{Z} \vDash \exists x_1 ... \exists x_n Q(n, x_1, \ldots, x_n) = 0.$$

Recursively enumerable sets themselves are only recognizable and not necessarily decidable. That is to say there is an algorithm that will ultimately return 'yes' if a given member is in a set, but isn't guaranteed to return 'no' otherwise. In the case where a set is undecidable and recognizable, the algorithm is blind to the difference between the possibility that it simply hasn't searched hard enough for the answer, and the possibility that there is no answer. We know that there are recursively enumerable sets that are not decidable [3]. Thus, by this result, it is concluded that Diophantine sets are not decidable.

Since Diophantine sets are definable in this particularly simple way, Matiasevic's result provides another example of first-order definability: all recursively enumerable sets of natural numbers are definable in the ring $\mathbb{Z}$ using only existential quantifiers, and no connectives.

[3] The halting problem is proven by Alan Turing to define one such set

6. Poonen's Definition of $\mathbb{Z}$ in $\mathbb{Q}$

As demonstrated in the previous sections, Robinson's formula uses quadratic forms to pick out the integers in the rationals. Although this is a proven approach, it is not the only way to define the integers in the rationals. Nor is it necessarily the best.

In [9], Poonen presents yet another first order definition of integers in the field of rationals. Before we present and discuss this definition, it is noteworthy to look at the motivation behind this.

6.1. Positive Arithmetic Hierarchy.

Poonen discusses a notion talked about by Zahidi and Cornelissen in [3], called 'Positive Arithmetic Hierarchy'. Informally, it is the number of quantifier changes that occur in a formula in Prenex Normal Form, with a sensitivity to the main quantifier operator. More formally, we call the set of atomic formulas, ie those without any quantifiers as of the form Σ_0^+, or equally of the form Π_0^+. Then formulas of the form Σ_n^+ are inductively defined to be those with the main operator being an existential quantifier, followed by any number of existential quantifiers, and then by a subformula of the form Π_{n-1}^+. Formulas of the form Π_n^+ are defined similarly, except for with the main operator being a universal quantifier. For example, we recall (11). This clearly is of the form Σ_1^+.

6.2. Further on Undecidability.

We recall that in the previous section it was demonstrated that Robinson essentially proved that any first-order theory of rationals is undecidable. In the light of this new notion of Positive Arithmetic Hierarchy, we can gauge and talk about the type of undecidability entailed by Robinson's formula.

(3) is of the form Π_4^+ with one free variable. What this means is that the Σ_5^+-theory of $\mathbb{Q}$ is undecidable.

Poonen proposes an alternative formula of the form Π_2^+ with one free variable, which means that by the same argument as encountered in Section (4), it can be proven that the Σ_3^+-theory of $\mathbb{Q}$ is undecidable.

H10 also has a version in the rationals: is there an algorithm that determines whether a given polynomial in several variables with integer coefficients have a zero in the rationals? This is one of the big open problems at the interface of logic and number theory. Everyone expects a negative solution, no one has been able to give a proof. This would of course show that the set of Σ_1^+ sentences that hold true in $\mathbb{Q}$ is undecidable.

6.3. Poonen's Formula.

Robinson's formula relies on modifying the second order definition of $\mathbb{Z}$ in $\mathbb{Q}$ by using quadratic forms to define inductive sets. Poonen does away with this, and instead works with quaternion algebras.

He first defines a Π_2^+ formula with one free variable, that begins with two universal quantifiers followed by seventeen existential quantifiers:

$$\begin{gathered}\Psi_1(t) = \forall a \forall b \exists a_1 \exists a_2 \exists a_3 \exists a_4 \exists b_1 \exists b_2 \exists b_3 \exists b_4 \exists x_1 \exists x_2 \exists x_3 \exists x_4 \exists y_1 \exists y_2 \exists y_3 \exists y_4 \exists n \\ [(a + a_1^2 + a_2^2 + a_3^2 + a_4^2)(b + b_1^2 + b_2^2 + b_3^2 + b_4^2) \\ [(x_1^2 - ax_2^2 - bx_3^2 + abx_4^2 - 1)^2 + (y_1^2 - ay_2^2 - by_3^2 + aby_4^2 - 1)^2 \\ +n^2(n-1)^2 \cdot \ldots \cdot (n-2309)^2 + (2x_1 + 2y_1 + n - t)^2] = 0]\end{gathered}$$

Poonen points out that even though he can't see how to get rid of the universal quantifiers, the number of existential quantifiers can be reduced to seven, obtaining:

$$\Psi_2(t) = \forall a \forall b \exists x_1 \exists x_2 \exists x_3 \exists x_4 \exists y_2 \exists y_3 \exists y_4$$
$$(a + x_1^2 + x_2^2 + x_3^2 + x_4^2)(b + x_1^2 + x_2^2 + x_3^2 + x_4^2)$$
$$\cdot[(x_1^2 - ax_2^2 - bx_3^2 + abx_4^2 - 1)^2 + \prod_{n=0}^{2309}((n - t - 2x_1)^2 - 4ay_2^2 - 4by_3^2 + 4aby_4^2 - 4)^2]$$

As discussed by Carol Wood in her talk [4], the top two lines are designed to allow one to ignore negative values of a and b. The two universal quantifiers come from some application of the Local-Global principle. Wood also agrees with Poonen that it is hard to see how one could do away with the universal quantifiers. Nonetheless this is a first reduction in the alterations of quantifiers since Robinson's definition.

6.4. **Kenigsmann's Results.** In his work [7], Kenigsmann presents another Π_2^+ formula for defining $\mathbb{Z}$ in $\mathbb{Q}$ using only one universal quantifier, as opposed to two.

7. Appendix

Below is presented a variant of the proof of 5 in the style of natural deduction. We only prove the left to right direction, as arguably it is the only interesting direction. Natural deduction is a formal proof system that revolves around identifying and formalizing various intuitive leaps in logic often used in mathematical proofs. It is primarily design to prove various sentences in the language of first order logic, however in principle it can be used when proving concepts found beyond the object language. A brief guide to natural deduction can be found in [1]. Note, in some places we skip on detail, for example in the construction of $\Phi(1, p, 0)$ from Lemma 2.3, as constructing sentences such as this can be quite tedious in the style of natural deduction, despite the fact that they follow trivially from the Lemma. We also take for granted the use of induction, which would normally require Peano axioms when done in the object language. Further, as mentioned this isn't, strictly speaking, natural deduction as the statement '$k \in \mathbb{Z}$' is not written in first order logic. But nonetheless the principles and mechanics of natural deduction are applicable.

Proof. Our local assumption is that $A(n)$ holds. The primary and secondary operators are universal quantifiers, and the ternary operator is an implication. We use double universal instantiation (Universal Elimination) on this premise, instantiating $a = 1$ and $b = p$, for some p with the property $p \equiv 3 \mod 4$. Thus we get as resource, the statement $B(1, p) \to \Phi(1, p, n)$. By using Lemma 2.3, we show that we can construct the statement $\Phi(1, p, 0)$. Further, since it is provable that M and $M + 1$ have the same denominator, we can prove by induction that $\forall m \langle \Phi(a, b, m) \to \Phi(a, b, m + 1) \rangle$. Alternatively this can be proven just with universal instantiation. By conjunction introduction, we then get the proposition $B(1, p)$. Then we use modus ponens to show that indeed we have $\Phi(1, p, n)$. From this, given the denominator d of n in its lowest terms, it is not the case that $2 \nmid d$, and $p \nmid d$.

[4] "Defining the integers in the rationals", NYWIMN Conference, May 2, 2008. Slides can be found at http://websupport1.citytech.cuny.edu/faculty/vgitman/nywimn/nywimn2/files/carolslides.pdf

Then, going back to the original premise, $A(n)$, we again use universal instantiation twice such that $a = q$ and $b = p$, where both are primes, $p \equiv 1 \mod 4$, and q is determined based on p by Claim 2.5, so $(q/p) = -1$. We get a premise of the form $B(q,p) \to \Phi(q,p,n)$. By Lemma 2.4 we show that q and p are of the form that $\Phi(q,p,0)$ holds. Again, like in the case above, by induction (or differently, with Universal Instantiation) we prove that $\forall m \langle \Phi(q,p,m) \to \Phi(q,p,m+1) \rangle$. So again, by conjunction introduction we get $B(q,p)$. Thus by modus ponens, we get $\Phi(q,b,n)$. But this implies that the denominator of n, d is such that $p \nmid d$.

Combining the last premise of each of the above derivations, we get it that whatever $n \in \mathbb{Q}$ is, it has to be of the form where its denominator d, in its lowest terms, is not divisible by 2, p, or q. But p and q are arbitrary primes of said properties. It is easy to prove that these 3 properties are necessary and sufficient to conclude that for all primes p $p \nmid n$. We prove this by contradiction, assuming that n is not an integer, from which we derive falsum. Thus, we get $n \in \mathbb{Z}$. □

8. Glossary

Undecidability A set is undecidable if there is no algorithm that, upon input, determine whether or not it is a member of the given set.

Decidability Consequently, we say that a set is decidable iff it is not the case that it is undecidable.

Recursively Enumerable Set A set who's members can be determined to be members of the set by some algorithm, but not necessarily the other way around.

Prenex Normal Form a first order formula is in PNF, when all the quantifiers are prefix to a quantifier-free formula. Every formula has a PNF equivalent.

Law of Quadratic Reciprocity A theorem to do with modular arithmetic that gives conditions for the solvability of quadratic equations modulo prime numbers.

Legendre Symbol (K/P) is a function that returns 1 if k is a quadratic residue $\mod p$, 0 if k is 0, and -1 otherwise.

Quadratic Residue we say that k is a quadratic residue $\mod p$ iff there exists some $i \in mathbbZ_p$ s.t. $k = i^2$.

Residue Class given a ring of integers $\mod p$, $\mathbb{Z}_p$, each element thereof is a residue class.

Quadratic Form a homogenous polynomial over the field $\mathbb{K}$, with degree 2 , and coefficients in $\mathbb{K}$.

A quadratic form is said to represent 0 in $\mathbb{K}$ iff there exist values $a_1....a_n \in \mathbb{K}$ s.t. $f(a_1.....a_n) = 0$ and some of these values are not 0.

A quadratic form is said to represent γ in $\mathbb{K}$ iff there exist values $a_1....a_n \in K$ s.t. $f(a_1.....a_n) = \gamma)$.

Diophantine Equation An equation with solutions in $\mathbb{Z}$ or $\mathbb{Q}$.

Hasse-Minkowski Theorem This states that two quadratic forms over some field $\mathbb{F}$ are equivalent iff they are equivalent over every completion of $\mathbb{F}$.

Local-Global Principle Also known as the Hasse principle, it is related to the Hasse-Minkowski theorem, and states that one can find an integer solution to an equation by using various concepts of number theory, such as the Chinese remainder theorem. A typical way to handle this is via the examination of the rational and p-adic completions of the rationals.

P-adic Numbers The p-adic numbers are a different completion of the rationals to the reals. Instead of using cauchy sequences based on the standard valuation of absolute value, p-adics are based on the p-adic absolute value.

P-adic Absolute Value Given $x \in \mathbb{Q}$, $x = p_1^{a_1}.....p_n^{a_n}$, we can rewrite it as $x = p_i^{a_i}\frac{n}{m}$, where $p_i \nmid n$ and $p_i \nmid m$. We note that if p_i, doesn't appear in the firstly presented representation of x, then $a_i = 0$. A p-adic absolute value is given as $|x|_{p_i} = p_i^{-a_i}$. This is a valuation.

References

[1] Merrie Bergmann, James Moor, and Jack Nelson. The logic book, 1998.

[2] JWS Cassels. Lectures on elliptic curves, volume 24 of London Mathematical Society student texts. *Cambridge University Press, Cambridge*, 19(20):334, 1991.

[3] Gunther Cornelissen and Karim Zahidi. Elliptic divisibility sequences and undecidable problems about rational points. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2007(613):1–33, 2007.

[4] Kurt Gödel. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme i. *Monatshefte für Mathematik und Physik*, 38(1):173–198, 1931.

[5] Godfrey Harold Hardy and Edward Maitland Wright. *An introduction to the theory of numbers*. Oxford University Press, 1979.

[6] Helmut Hasse. Über die Darstellbarkeit von Zahlen durch quadratische Formen im Körper der rationalen Zahlen. *Journal fur die reine und angewandte Mathematik*, 152:129–148, 1923.

[7] Jochen Koenigsmann. Defining Z in Q. *Annals of Mathematics*, 183(1):73–93, 2016.

[8] D. Louvish and R. Fraïssé. *Course of Mathematical Logic: Volume 2 Model Theory*. Synthese Library. Springer Netherlands, 1974.

[9] Bjorn Poonen. Characterizing integers among rational numbers with a universal-existential formula. *American Journal of Mathematics*, pages 675–682, 2009.

[10] Julia Robinson. Definability and decision problems in arithmetic. *The Journal of Symbolic Logic*, 14(2):98–114, 1949.

[11] Michael Sipser. *Introduction to the Theory of Computation*. Cengage Learning, 2012.