Sharpening complexity results in quantified probability logic

Stanislav O. Speranski

`This is a pre-print version of the article published in` *`Logic Journal of the IGPL`*

`DOI: 10.1093/jigpal/jzae114.`

Abstract

We shall be concerned with two natural expansions of the quantifier-free 'polynomial' probability logic of [3]. One of these, denoted by $\mathsf{QPL}^{\mathrm{e}}$, is obtained by adding quantifiers over arbitrary events, and the other, denoted by $\underline{\mathsf{QPL}}^{\mathrm{e}}$, uses quantifiers over propositional formulas — or equivalently, over events expressible by such formulas. The earlier proofs of the complexity lower bound results for $\mathsf{QPL}^{\mathrm{e}}$ and $\underline{\mathsf{QPL}}^{\mathrm{e}}$ relied heavily on multiplication, and therefore on the polynomiality of the basic parts. We shall obtain the same lower bounds for natural fragments of $\mathsf{QPL}^{\mathrm{e}}$ and $\underline{\mathsf{QPL}}^{\mathrm{e}}$ in which only linear combinations of a very special form are allowed. Also, it will be studied what happens if we add quantifiers over reals.

Keywords: probability logic, quantification, undecidability, complexity, second-order arithmetic

Contents

1 Introduction 3

2 Quantifying over events 5

3 Passing to quotients 7

4 Quantifying over propositional formulas 9

5 Concerning second-order arithmetic 10

6 Concerning elementary analysis 14

7 The case of quantifiers over events 15

8 The case of quantifiers over propositional formulas 18

9 Some general upper bounds 21

References 24

1 Introduction

In their famous article [3], Fagin, Halpern and Megiddo considered two quantifier-free languages for reasoning about probabilities. The larger of these, which we shall call 'polynomial' and denote by $\mathcal{L}_{\mathrm{poly}}$, deals with Boolean combinations of expressions of the form

$$f\,(\mu\,(\phi_1), \ldots, \mu\,(\phi_m)) \;\leqslant\; g\,(\mu\,(\psi_1), \ldots, \mu\,(\psi_n)).$$

where f, g are polynomials with integer coefficients, and ϕ_1, …, ϕ_m, ψ_1, …, ψ_n are propositional formulas. The structures for $\mathcal{L}_{\mathrm{poly}}$ — or *measurable probability structures*, in the terminology of [3] — are the tuples

$$\langle\langle \mathscr{A}, \mathsf{P}\rangle, \pi\rangle$$

where $\langle \mathscr{A}, \mathsf{P}\rangle$ is a probability space, and π is a function from the set of propositional variables to $\mathscr{A}$. Briefly, given a $\mathcal{L}_{\mathrm{poly}}$-structure, we first need to extend π to handle all propositional formulas in the natural way, and then interpret each $\mu\,(\phi)$ as $\mathsf{P}\,(\pi\,(\phi))$. The smaller language, which we call 'linear' and denote by $\mathcal{L}_{\mathrm{lin}}$, is obtained from $\mathcal{L}_{\mathrm{poly}}$ by excluding multiplication; so $\mathcal{L}_{\mathrm{lin}}$ is the linear fragment of $\mathcal{L}_{\mathrm{poly}}$. Moreover, as was suggested in [3], one may expand $\mathcal{L}_{\mathrm{poly}}$ by adding quantifiers over reals. We write $\mathcal{L}^*_{\mathrm{poly}}$ for this richer language.

The validity problem for $\mathcal{L}^*_{\mathrm{poly}}$ can be easily reduced to the membership problem for the first-order theory of the ordered field of reals, which was shown to be decidable by Tarski [22]. But if we want to analyze the algorithmic complexity of the validity problems for $\mathcal{L}^*_{\mathrm{poly}}$, $\mathcal{L}_{\mathrm{poly}}$ and $\mathcal{L}_{\mathrm{lin}}$ in terms of polynomial-time reducibility, this does not give us much. Some nice complexity results — as well as certain axiomatizability results — were derived in [3]; see also [11]. In particular, the satisfiability problem for $\mathcal{L}_{\mathrm{lin}}$ is NP-complete, while those for $\mathcal{L}_{\mathrm{poly}}$ and $\mathcal{L}^*_{\mathrm{poly}}$ belong to PSpace and $\mathsf{ExpSpace}$ respectively. [1] Compare [6], [8] and [10, Chapter 3], where similar but somewhat weaker probability logics are examined.

The above three languages can be enriched in different ways, and may often be viewed as fragments of other probability logics. For instance, following [1], one can modify $\mathcal{L}^*_{\mathrm{poly}}$ by taking the underlying language (over the formulas of which measures are distributed) to be first-order rather than propositional, and adding quantifiers over elements of a given domain; see [10] for a related probability logic having no quantifiers over reals.[2] Alternatively, we can keep the underlying language unchanged and add new quantifiers directly to $\mathcal{L}_{\mathrm{poly}}$ or $\mathcal{L}^*_{\mathrm{poly}}$, as was done in [15] and [19]. The present article is concerned with the latter approach. Notice that probability logics containing quantifiers other than those over reals tend to be far from decidable — so their complexity should

[1] In fact, it would be more accurate to use the complexity class $\exists\mathbb{R}$ instead of PSpace; cf. [6]. Evidently, in all three languages, a formula is valid iff its negation is not satisfiable.

[2] The reader might be wondering why quantifiers over reals are avoided in the systems developed by Ognjanović and his colleagues, who are often concerned with axiomatizations (somewhat similar to those provided in [5] and [7]). Technically, the reason is that probability logics with both quantifiers over reals and some other sort of quantifiers are usually at least as complex as complete second-order arithmetic; cf. [1]. This implies that such logics are not axiomatizable by means of reasonable infinitary calculi, which can only handle Π^1_1-sets; see [9] for details.

be analyzed using degrees of undecidability, in terms of reducibility by means of total computable functions.

Following [15], consider the expansion $\underline{\mathsf{QPL}}^{\mathrm{e}}$ of $\mathcal{L}_{\mathrm{poly}}$ obtained by adding quantifiers over propositional formulas — these may be thought of as countable conjunctions of a special kind, but we need to keep the syntax finitary. Of course, the structures for $\underline{\mathsf{QPL}}^{\mathrm{e}}$ are the same as those for $\mathcal{L}_{\mathrm{poly}}$. As was shown in [16, Section 4], for any sufficiently large class of such structures, its $\underline{\mathsf{QPL}}^{\mathrm{e}}$-theory is Π^1_1-hard. The corresponding proof relied heavily on multiplication, and therefore on the polynomiality of the basic part. There, the natural numbers were interpreted as a distinguished sequence E_0, E_1, … of events, whose measures form a geometric progression, e.g.

$$\mathsf{P}\,(E_n) \;=\; \frac{1}{2^{n+1}} \quad \text{for each} \;\; n \in \mathbb{N};$$

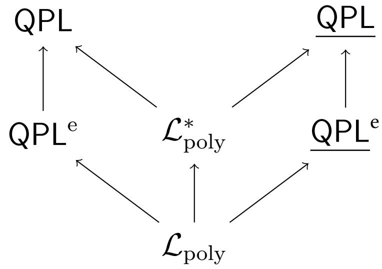
cf. the proofs in [1]. Obviously, for all $i, j, n \in \mathbb{N}$,

$$i + j \;=\; n \quad \iff \quad 2 \cdot \frac{1}{2^{i+1}} \cdot \frac{1}{2^{j+1}} \;=\; \frac{1}{2^{n+1}}.$$

This allows us to interpret addition on $\mathbb{N}$ by using multiplication on probabilities; then we aim to apply Halpern's result from [4]. Such reasoning fails if we restrict ourselves to linear terms. In this article, by providing a more complicated and advanced argument (not related to [4]) we are going to obtain Π^1_1-hardness for a natural fragment of $\underline{\mathsf{QPL}}^{\mathrm{e}}$ in which only linear combinations of a very special form are admitted. Also, we analyze the complexity of the language $\underline{\mathsf{QPL}}$ expanding $\underline{\mathsf{QPL}}^{\mathrm{e}}$ by adding quantifiers over reals, which has not been studied before — this leads to the complexity of complete second-order arithmetic, instead of its Π^1_1-part. Here the upper bound is obtained by a rather general argument, which can be applied to different classes of $\underline{\mathsf{QPL}}$-structures.

Next, a variant $\mathcal{L}'_{\mathrm{poly}}$ of $\mathcal{L}_{\mathrm{poly}}$ can be obtained by replacing propositional variables by 'Boolean variables', intended to range over $\mathscr{A}$. Intuitively, the difference between $\mathcal{L}_{\mathrm{poly}}$ and $\mathcal{L}'_{\mathrm{poly}}$ is that in $\mathcal{L}_{\mathrm{poly}}$, propositional variables are treated as Boolean constants, rather than Boolean variables. This is not essential to the description of $\mathcal{L}_{\mathrm{poly}}$, but it suggests another way of expanding $\mathcal{L}_{\mathrm{poly}}$, namely by adding quantifiers over arbitrary events, as in [19]. We write $\mathsf{QPL}^{\mathrm{e}}$ for the resulting expansion, and QPL for the even richer language with quantifiers over reals; see also [21].[3] As was shown in [19, Section 2], for every sufficiently large class of probability spaces, its $\mathsf{QPL}^{\mathrm{e}}$-theory is at least as complex as complete second-order arithmetic; a general upper bound argument for both $\mathsf{QPL}^{\mathrm{e}}$ and QPL has recently been given in [21]. But the lower bound argument again relied heavily on multiplication, using a suitable generalization of Halpern's result from [17, Section 3]. We are going to show that the corresponding hardness result remains true if we restrict ourselves to special linear combinations, like in $\underline{\mathsf{QPL}}^{\mathrm{e}}$. At this point it might be helpful to represent the relationship between the probabilistic languages mentioned above:

[3] These two can be viewed as elementary languages for reasoning about probability spaces; cf. [14] and other work on elementary theories, e.g. [2], [22] and [23].



Here an arrow from one language to another indicates that the latter expands (conservatively) the former. Notice that QPL and $\underline{\mathsf{QPL}}$ are incomparable, according to the diagram: though the syntax of $\underline{\mathsf{QPL}}$ extends that of QPL, the semantics of quantifiers in $\underline{\mathsf{QPL}}$ differs from that in QPL.

The present work may be compared — or contrasted — with [3] and [6]. For instance, in terms of polynomial-time reducibility, there should be a big difference between $\mathcal{L}_{\text{lin}}$, $\mathcal{L}_{\text{poly}}$ and $\mathcal{L}^*_{\text{poly}}$, but not between $\mathcal{L}_{\text{lin}}$ and certain more restrictive languages. As for the quantified probability logics under consideration, it will be proved that that there is no difference between the polynomial, linear and certain sublinear versions of a given logic; adding quantifiers over reals leads to higher degrees of undecidability in the case of quantifiers over propositional formulas, but not in that of quantifiers over events.

The rest of the article is organized as follows. In Section 2, we define the expansions QPL and $\mathsf{QPL}^{\mathrm{e}}$; Section 3 explains why we can safely pass from each probability space to the corresponding quotient space modulo events of measure zero. In Section 4, we define $\underline{\mathsf{QPL}}$ and $\underline{\mathsf{QPL}^{\mathrm{e}}}$. Sections 5 and 6 contain some technical material on the analytical hierarchy. In Section 7 we show how the earlier hardness results for $\mathsf{QPL}^{\mathrm{e}}$ and QPL can be strengthened, without any use of multiplication. Section 7 does a similar job for quantifiers over propositional formulas. Finally, Section 9 provides general upper bound arguments for $\underline{\mathsf{QPL}}$ and $\underline{\mathsf{QPL}^{\mathrm{e}}}$.

2 Quantifying over events

By a *probability space*, or simply a *space*, we mean a pair $\langle \mathscr{A}, \mathsf{P} \rangle$ where:

- $\mathscr{A}$ is a σ-algebra, i.e. a Boolean algebra in which every countable set of elements has a supremum (and hence an infimum);
- P is a *probability measure* on $\mathscr{A}$, i.e. a function from $\mathscr{A}$ to $[0, 1]$ such that for every sequence $A_0, A_1, \dots$ of pairwise disjoint elements of $\mathscr{A}$,

  $$\mathsf{P}\left(\bigvee_{n\in\mathbb{N}} A_n\right) \;=\; \sum_{n\in\mathbb{N}} \mathsf{P}\,(A_n),$$

  and $\mathsf{P}\,(\top) = 1$ where $\top$ denotes the greatest element of $\mathscr{A}$.

Elements of $\mathscr{A}$ are called *events*, which are measurable with respect to P. Note, in passing, that the countable additivity of P readily implies $\mathsf{P}\,(\bot) = 0$ where $\bot$ is the least element of $\mathscr{A}$.

Since the very definition of a space involves two different sorts of object, our formal language QPL includes two disjoint countable sets of variables:

$$\mathrm{Var} \; := \; \{X, Y, Z, \dots\} \quad \text{and} \quad \mathrm{var} \; := \; \{x, y, z, \dots\}.$$

Elements of Var are intended to range over events, and called *Boolean variables*, while those of var are intended to range over reals, and called *field variables*. This in turn suggests considering two sets of function symbols:

$$\{\bot, \top, \wedge, \vee, \neg\} \quad \text{and} \quad \{0, 1, +, \cdot, -\},$$

viz. the symbols of the language of Boolean algebras and those of the language of fields. In addition to these, we need a special symbol μ to denote a probability measure.

The *Boolean terms* are build up from $\bot$, $\top$ and the Boolean variables by use of $\wedge$, $\vee$ and $\neg$:

- if ϕ_1 and ϕ_2 are Boolean terms, so are $\phi_1 \wedge \phi_1$ and $\phi_1 \vee \phi_2$;
- if ϕ is a Boolean term, so is $\neg\phi$.

Naturally, they represent Boolean combinations of events. By a *μ-term* we mean an expression of the form $\mu(\phi)$ where ϕ is a Boolean term. The *field terms* are built up from 0, 1, the field variables and the μ-terms by use of $+$, $\cdot$ and $-$. Briefly, each field term can be represented as

$$f\,(x_1, \dots, x_m, \mu(\phi_1), \dots, \mu(\phi_n))$$

where f is a polynomial with integer coefficients, x_1, …, x_m are field variables, and ϕ_1, …, ϕ_n are Boolean terms. Now by a *basic* QPL*-formula* we mean an expression of the form $t_1 \leqslant t_2$ where t_1 and t_2 are field terms.

We shall use $\wedge$, $\vee$ and $\neg$ to denote the Boolean operations as well as the ordinary logical connectives. Since their Boolean versions will not occur outside the scope of μ, the interpretations of $\wedge$, $\vee$ and $\neg$ will always be clear from the context. Taking the quantifier symbols to be $\forall$ and $\exists$, the QPL*-formulas* are built up from the basic QPL-formulas in the customary way. We shall adopt the following abbreviations:

$$\begin{aligned} t_1 < t_2 \; &:= \; t_1 \leqslant t_2 \wedge \neg t_2 \leqslant t_1; \\ t_1 = t_2 \; &:= \; t_1 \leqslant t_2 \wedge t_2 \leqslant t_1; \\ t_1 \neq t_2 \; &:= \; \neg t_1 = t_2. \end{aligned}$$

Further, if Φ and Ψ are QPL-formulas, we shall often write $\Phi \to \Psi$ and $\Phi \leftrightarrow \Psi$ instead of $\neg\Phi \vee \Psi$ and $(\Phi \to \Psi) \wedge (\Psi \to \Phi)$ respectively. Given a QPL-formula Φ, define $\mathrm{Free}\,(\Phi)$ to be the set of all variables (of any of the two sorts) that occur free in Φ. Call Φ a QPL*-sentence* if $\mathrm{Free}\,(\Phi) = \varnothing$.

The satisfiability relation $\Vdash$ for QPL can be defined in the obvious way, and it behaves like one would expect. In more detail, let $\mathscr{P} = \langle \mathscr{A}, \mathsf{P} \rangle$ be a space. Then for any QPL-formula Φ, function ζ from Var to $\mathfrak{A}$ and function ι from var to $\mathbb{R}$, we define

$$\mathscr{P} \Vdash \Phi\,[\zeta, \iota]$$

by induction on Φ, as in two-sorted first-order logic. Clearly, it does not matter what values $\langle \zeta, \iota \rangle$ assigns to $(\mathrm{Var} \cup \mathrm{var}) \setminus \mathrm{Free}(\Phi)$. In particular, if Φ is a sentence, we may write $\mathscr{P} \Vdash \Phi$ instead of $\mathscr{P} \Vdash \Phi[\zeta, \iota]$. For example, consider

$$\Theta \;:=\; \forall x\,(0 \leqslant x \leqslant 1 \to \exists X\; x = \mu(X)).$$

Then $\mathscr{P} \Vdash \Theta$ iff for every $r \in [0, 1]$ there exists $A \in \mathscr{A}$ such that $\mathsf{P}(A) = r$.

Let $\mathsf{QPL}^{\mathrm{e}}$ be the sublanguage of QPL obtained by excluding field variables, and hence quantifiers over reals. We shall write Sent and $\mathrm{Sent}^{\mathrm{e}}$ for the collections of all sentences in QPL and $\mathsf{QPL}^{\mathrm{e}}$ respectively. Then, given a class $\mathcal{K}$ of spaces, define the QPL*-theory of* $\mathcal{K}$ to be

$$\mathrm{Th}(\mathcal{K}) \;:=\; \{\Phi \in \mathrm{Sent} \mid \mathfrak{A} \Vdash \Phi \text{ for all } \mathfrak{A} \in \mathcal{K}\}.$$

The $\mathsf{QPL}^{\mathrm{e}}$*-theory of* $\mathcal{K}$, written $\mathrm{Th}^{\mathrm{e}}(\mathcal{K})$, is defined similarly. In general, QPL-theories tend to have very high degrees of undecidability, even without quantifiers over reals.

Say that a class of probability spaces is *rich* iff it contains all infinite discrete spaces. Using an alternative description of the analytical hierarchy given in [17, Section 3], we can derive:

Theorem 2.1 (see [19, Section 2])

Let $\mathcal{K}$ *be a rich class of spaces. Then complete second-order arithmetic — i.e. the second-order theory of the standard model of arithmetic — is reducible to the* $\mathsf{QPL}^{\mathrm{e}}$*-theory of* $\mathcal{K}$*.*[4]

On the other hand, as has been shown in [21, Section 6], with each probability space, one can associate a suitable subset of $\mathbb{N}$. Call a class of spaces *analytical* iff the corresponding collection of subsets of $\mathbb{N}$ is second-order definable in the standard model of arithmetic.

Theorem 2.2 (see [21])

Let $\mathcal{K}$ *be an analytical class of spaces. Then the* QPL*-theory of* $\mathcal{K}$ *is reducible to complete second-order arithmetic.*

This covers a wide range of classes of spaces. Interestingly, the corresponding reduction can be adapted to derive some decidability results or to obtain smaller upper bounds for certain non-rich classes. For instance, the QPL-theory of the class of all atomless spaces turns out to be decidable. Those who would like to know more should consult [21].

Remark 2.3. The fragment of QPL in which quantifiers over events are not allowed — but those over reals are legal — might be viewed as the 'polynomial' logic described earlier in [3, Section 6], viz. $\mathcal{L}^*_{\mathrm{poly}}$. Similarly for the language $\underline{\mathsf{QPL}}$, which will be discussed in Section 4.

3 Passing to quotients

Two spaces $\mathscr{P}_1 = \langle \mathscr{A}_1, \mathsf{P}_1 \rangle$ and $\mathscr{P}_2 = \langle \mathscr{A}_2, \mathsf{P}_2 \rangle$ are called *isomorphic* if there exists $f : \mathscr{A}_1 \to \mathscr{A}_2$ such that:

[4]For more on second-order arithmetic, see Section 5.

- f is an isomorphism of $\mathscr{A}_1$ onto $\mathscr{A}_2$, thought of as Boolean algebras;
- $\mathsf{P}_1(A) = \mathsf{P}_2(f(A))$ for all $A \in \mathscr{A}_1$.

The notion of isomorphism may be relaxed in the following way. Consider the QPL-formula

$$X \approx Y \ := \ \mu((X \wedge \neg Y) \vee (Y \wedge \neg X)) = 0.$$

Let $\mathscr{P} = \langle \mathscr{A}, \mathsf{P} \rangle$ be a space. Evidently, $X \approx Y$ defines (in $\mathscr{P}$) a very natural equivalence relation on $\mathscr{A}$, namely

$$\mathscr{E} \ := \ \{(A_1, A_2) \in \mathscr{A} \times \mathscr{A} \mid \mathscr{P} \Vdash A_1 \approx A_2\}.$$

In particular, $\mathscr{E}$ turns out to be a congruence relation on $\mathscr{A}$. For each $A \in \mathscr{A}$, denote by $[A]_\approx$ the equivalence class of A under $\mathscr{E}$, i.e.

$$[A]_\approx \ := \ \{B \in \mathscr{A} \mid (A, B) \in \mathscr{E}\}.$$

Now take $\mathscr{A}_\approx$ to be the collection of all such classes — or rather, the quotient Boolean algebra of $\mathscr{A}$ modulo $\mathscr{E}$ — and define $\mathsf{P}_\approx : \mathscr{A}_\approx \to [0,1]$ by

$$\mathsf{P}_\approx([A]_\approx) \ := \ \mathsf{P}(A)$$

(note that $\mathsf{P}(A) = \mathsf{P}(B)$ for all $B \in [A]_\approx$). It is easy to check that $\mathscr{P}_\approx = \langle \mathscr{A}_\approx, \mathsf{P}_\approx \rangle$ is a probability space, called the *quotient space of $\mathscr{P}$ modulo events of measure zero.* Moreover, one can show that for any QPL-formula Φ, function ζ from Var to $\mathscr{A}$ and function ι from var to $\mathbb{R}$,

$$\mathscr{P} \Vdash \Phi[\zeta, \iota] \quad \Longleftrightarrow \quad \mathscr{P}_\approx \Vdash \Phi[\zeta_\approx, \iota]$$

where $\zeta_\approx$ is the mapping from Var to $\mathscr{A}_\approx$ given by

$$\zeta_\approx(X) \ := \ [\zeta(X)]_\approx.$$

Finally, we call $\mathscr{P}_1$ and $\mathscr{P}_2$ *weakly isomorphic* if $(\mathscr{P}_1)_\approx$ and $(\mathscr{P}_2)_\approx$ are isomorphic. Consequently, if two spaces are weakly isomorphic, then their QPL-theories coincide.

Let $\mathscr{P} = \langle \mathscr{A}, \mathsf{P} \rangle$ be a space. Now if $\mathscr{P} \Vdash A_1 \approx A_2$, i.e. the symmetric difference of A_1 and A_2 has measure zero, then A_1 and A_2 are indistinguishable in $\mathscr{P}$ by QPL-formulas (with parameters). So definability in $\mathscr{P}$ reduces to definability in $\mathscr{P}_\approx$. For instance, consider

$$X \preccurlyeq Y \ := \ X \wedge Y \approx X.$$

Obviously, $\mathscr{P} \Vdash A_1 \preccurlyeq A_2$ iff $[A_1]_\approx$ is less than or equal to $[A_2]_\approx$ in the Boolean algebra $\mathscr{A}_\approx$. Next, take

$$\mathrm{At}(X) \ := \ \mu(X) \neq 0 \wedge \forall Y (\mu(Y) \neq 0 \wedge Y \preccurlyeq X \to Y \approx X).$$

Clearly, $\mathscr{P} \Vdash \mathrm{At}(A)$ iff $[A]_\approx$ is an atom of $\mathscr{A}_\approx$ — in other words, $[A]_\approx$ is minimal in $\mathscr{A}_\approx \setminus \{[\bot]_\approx\}$. The formula $\mathrm{At}(X)$ will play an important role in our lower bound arguments.

4 Quantifying over propositional formulas

Next, we describe a variation $\underline{\mathsf{QPL}}$ on QPL in which quantifiers over events are replaced by those over ‘propositional formulas’. The syntax of $\underline{\mathsf{QPL}}$ extends that of QPL by adding countably many Boolean constants, called *propositional variables*:

$$p, \quad q, \quad r, \quad \ldots$$

Denote by Prop the set of all these constants. Then the notion of Boolean term is modified in the obvious way. Of course, closed Boolean terms may be called *propositional formulas* in this context. Take Term° to be the set of all these terms. Roughly, in $\underline{\mathsf{QPL}}$ we have

$$\forall X\, \Phi \;\leftrightarrow\; \bigwedge_{\phi \in \mathrm{Term}^{\circ}} \Phi\,(X/\phi) \quad \text{and} \quad \exists X\, \Phi \;\leftrightarrow\; \bigvee_{\phi \in \mathrm{Term}^{\circ}} \Phi\,(X/\phi)$$

where $\Phi\,(X/\phi)$ is the result of replacing all free occurrences of X in Φ by ϕ. To make this precise, one can modify the notion of a space.

By an *expanded space* we mean a tuple $\langle\langle \mathscr{A}, \mathsf{P}\rangle, \pi\rangle$ where $\langle \mathscr{A}, \mathsf{P}\rangle$ is a space and π is a function from Prop to $\mathscr{A}$.[5] Such spaces have been widely employed in probability logic — in particular, in [3] they are called *measurable probability structures.* In $\underline{\mathsf{QPL}}$ the Boolean variables are intended to range over

$$\mathscr{A}^{\pi} \;:=\; \{\pi\,(\phi) \mid \phi \in \mathrm{Term}^{\circ}\},$$

rather than the whole of $\mathscr{A}$. Hence if we want to adapt the material of Section 3 to $\underline{\mathsf{QPL}}$, then $\mathscr{A}$ has to be replaced by $\mathscr{A}^{\pi}$ throughout. The symbol $\Vdash$ will be used for satisfiability in both QPL and $\underline{\mathsf{QPL}}$ — but this should cause no confusion.[6] Naturally, since Term° is countable, one might expect $\underline{\mathsf{QPL}}$ to be less expressive than QPL, at least without quantifiers over reals: quantifiers in $\underline{\mathsf{QPL}}$ are somewhat similar to those over $\mathbb{N}$, not over the power set of $\mathbb{N}$.

Let $\underline{\mathsf{QPL}}^{\mathrm{e}}$ be the sublanguage of $\underline{\mathsf{QPL}}$ obtained by excluding field variables. We shall write $\underline{\mathrm{Sent}}$ and $\underline{\mathrm{Sent}}^{\mathrm{e}}$ for the collections of all sentences in $\underline{\mathsf{QPL}}$ and $\underline{\mathsf{QPL}}^{\mathrm{e}}$ respectively. Then, given a class $\mathcal{K}$ of expanded spaces, define the $\underline{\mathsf{QPL}}$*-theory of* $\mathcal{K}$ to be

$$\underline{\mathrm{Th}}\,(\mathcal{K}) \;:=\; \{\Phi \in \underline{\mathrm{Sent}} \mid \mathfrak{A} \Vdash \Phi \text{ for all } \mathfrak{A} \in \mathcal{K}\}.$$

The $\underline{\mathsf{QPL}}^{\mathrm{e}}$*-theory of* $\mathcal{K}$, written $\underline{\mathrm{Th}}^{\mathrm{e}}\,(\mathcal{K})$, is defined similarly.

Say that a class of expanded probability spaces is *rich* iff it contains all expanded spaces of the form $\langle \mathscr{P}, \pi\rangle$ where $\mathscr{P}$ is an infinite discrete space. Utilizing the result of [4], we can derive:

Theorem 4.1 (see [16, Section 4])
Let $\mathcal{K}$ *be a rich class of expanded spaces. Then the* $\underline{\mathsf{QPL}}^{\mathrm{e}}$*-theory of* $\mathcal{K}$ *is* Π^1_1*-hard.*

[5] By analogy with Boolean valuations, π can be extended to deal with all propositional formulas.

[6] However, we need to keep in mind that $\underline{\mathsf{QPL}}$ is not a (conservative) expansion of QPL. For instance, the sentence $\forall x\,(0 \leqslant x \leqslant 1 \to \exists X\, x = \mu\,(X))$ is satisfiable in QPL but not in $\underline{\mathsf{QPL}}$.

This bound is known to be precise in the case of all expanded spaces and some other cases. A general upper bound argument will be presented in Section 9.

Clearly, each expanded space $\mathfrak{A} = \langle\langle \mathscr{A}, \mathsf{P}\rangle, \pi\rangle$ induces the pair $\langle \mathscr{A}^{\pi}, \mathsf{P}^{\pi}\rangle$ where P^{π} denotes the restriction of P to $\mathscr{A}^{\pi}$. Further, we can pass from P^{π} to the function $\mathsf{p}^{\mathfrak{A}}$ from Term° to $[0, 1]$ given by

$$\mathsf{p}^{\mathfrak{A}}(\phi) \ := \ \mathsf{P}(\pi(\phi)).$$

Let us call $f : \mathrm{Term}^{\circ} \to [0, 1]$ *acceptable* iff $f(\top) = 1$ and for any $\phi, \psi \in \mathrm{Term}^{\circ}$:

- if ϕ and ψ are logically equivalent, then $f(\phi) = f(\psi)$;
- $f(\phi) = f(\phi \wedge \psi) + f(\phi \wedge \neg\psi)$

(cf. the axioms for reasoning about probabilities provided in [3, Section 2.2]). Clearly, $\mathsf{p}^{\mathfrak{A}}$ is always acceptable. The converse also holds, as is easily verified:

Proposition 4.2

For every acceptable $f : \mathrm{Term}^{\circ} \to [0, 1]$ *there exists an expanded space* $\mathfrak{A}$ *such that* $\mathsf{p}^{\mathfrak{A}} = f$.

Thus, instead of expanded spaces, we can work with acceptable functions.

Remark 4.3. Originally, $\underline{\mathsf{QPL}}^{\mathrm{e}}$ was introduced in [15], as a natural expansion of the quantifier-free 'polynomial' logic described in [3, Section 5], viz. $\mathcal{L}_{\mathrm{poly}}$. Unlike the languages studied in [1], $\underline{\mathsf{QPL}}^{\mathrm{e}}$ — as well as $\underline{\mathsf{QPL}}$ — keeps the underlying logic (inside the scope of μ) propositional and does not depend on the choice of external signature.

5 Concerning second-order arithmetic

Remember, in second-order arithmetic we have:

- *individual variables* x, y, z, ..., intended to range over $\mathbb{N}$;
- for each $k \in \mathbb{N}_+$, k*-ary set variables* X^k, Y^k, Z^k, ..., intended to range over the subsets of $\mathbb{N}^k$ — i.e. over the k-ary relations on $\mathbb{N}$.[7]

Let $\mathfrak{N}$ be the standard model of Peano arithmetic and σ be its signature (i.e. the corresponding list of function and predicate symbols). To make things precise, assume that

$$\sigma \ := \ \langle 0, \mathsf{s}, +, \cdot\, ; =\rangle.$$[8]

The *atomic second-order* σ*-formulas* are the identities between σ-terms plus the expressions of the form

$$X^k(t_1, \ldots, t_k)$$

[7]Here $\mathbb{N}_+$ denotes $\mathbb{N} \setminus \{0\}$, i.e. $\{1, 2, \ldots\}$.

where $k \in \mathbb{N}_+$, X^k is a k-ary set variable and $t_1, \ldots, t_k$ are σ-terms. More complicated *second-order σ-formulas* are built up from these in the usual way. In what follows by a *formula* we shall mean a second-order formula, unless otherwise specified.

Before proceeding, let us bring in some concepts from computability theory. Given $A, B \subseteq \mathbb{N}$, we say that A is *reducible* to B, written $A \leqslant B$, iff there exists a computable $f : \mathbb{N} \to \mathbb{N}$ such that for every $m \in \mathbb{N}$,

$$m \in A \iff f(m) \in B.$$

Further, A and B are *equivalent*, written $A \equiv B$, iff they are reducible to each other.

Let $n \in \mathbb{N}_+$. Recall that a σ-formula is in Π^1_n iff it has the form

$$\underbrace{\forall \vec{X}_1 \, \exists \vec{X}_2 \, \ldots \, \vec{X}_n}_{n-1 \text{ alternations}} \Psi$$

where $\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_n$ are tuples of set variables and Ψ contains no set quantifiers. Then $A \subseteq \mathbb{N}$ is:

- Π^1_n-*bounded* iff A can be defined in $\mathfrak{N}$ by a Π^1_n-formula;
- Π^1_n-*hard* iff every Π^1_n-bounded subset of $\mathbb{N}$ is reducible to A;
- Π^1_n-*complete* iff it is both Π^1_n-bounded and Π^1_n-hard.

Traditionally, Π^1_n-bounded sets are called Π^1_n-*sets*.

For each $n \in \mathbb{N}_+$, denote by $\Pi^1_n\text{-Th}(\mathfrak{N})$ the collection of all Π^1_n-sentences true in $\mathfrak{N}$. We write $\text{Th}(\mathfrak{N})$ for the full second-order theory of $\mathfrak{N}$, or *complete second-order arithmetic*. Naturally, given a reasonable coding of σ-formulas, we may identify σ-sentences with their codes.

Folklore 5.1 (see [13, Chapter 16])

Let $n \in \mathbb{N}_+$. Then for every $A \subseteq \mathbb{N}$:

$$\begin{aligned} A \leqslant \Pi^1_n\text{-Th}(\mathfrak{N}) &\iff A \text{ is } \Pi^1_n\text{-bounded;} \\ \Pi^1_n\text{-Th}(\mathfrak{N}) \leqslant A &\iff A \text{ is } \Pi^1_n\text{-hard;} \\ \Pi^1_n\text{-Th}(\mathfrak{N}) \equiv A &\iff A \text{ is } \Pi^1_n\text{-complete.} \end{aligned}$$

Inspired by these equivalences, we shall say that $A \subseteq \mathbb{N}$ is Π^1_∞-*bounded*, Π^1_∞-*hard* or Π^1_∞-*complete* iff $A \leqslant \text{Th}(\mathfrak{N})$, $\text{Th}(\mathfrak{N}) \leqslant A$ or $A \equiv \text{Th}(\mathfrak{N})$ respectively.

Folklore 5.2 (see [13, Chapter 16].)

Every σ-formula can be effectively converted to an equivalent Π^1_n-formula, for a suitable $n \in \mathbb{N}_+$. So the union of the sets $\Pi^1_1\text{-Th}(\mathfrak{N})$, $\Pi^1_2\text{-Th}(\mathfrak{N})$, … is Π^1_∞-complete, i.e.

$$\text{Th}(\mathfrak{N}) \equiv \bigcup_{n \in \mathbb{N}_+} \Pi^1_n\text{-Th}(\mathfrak{N}).$$

[8] Alternatively, we may assume that σ contains a symbol for any primitive recursive function or relation.

Finally, $A \subseteq \mathbb{N}$ is *analytical* iff A is definable in $\mathfrak{N}$ by some σ-formula. By Folklore 5.2, for each $n \in \mathbb{N}_+$, the Π^1_n-sets may be thought of as forming the n-th level of the *analytical hierarchy.*

The lower bound proof in [19, Section 2] utilized the following rather simple result.

Theorem 5.3 (see [17, Section 3])

Let σ_+ be $\langle 0, \mathsf{s}, +; = \rangle$. Then every Π^1_n-subset of $\mathbb{N}$ can be defined in $\mathfrak{N}$ — or rather in its σ_+-reduct — by a σ_+-formula of the form

$$\underbrace{\forall X^1_1 \, \exists X^1_2 \, \ldots \, X^1_n}_{n-1 \text{ alternations}} \, \Psi \tag{$\mathtt{P}_n$}$$

where X^1_1, X^1_2, …, X^1_n are unary set variables and Ψ contains no set quantifiers.

Corollary 5.4 (see [17, Section 3])

The collection of all σ_+-sentences of the form $\mathtt{P}_n$ that are true in $\mathfrak{N}$ is Π^1_n-complete.

Notice that Corollary 5.4 generalizes the result of [4], which was used in [1, Section 5] and [16, Section 4].[9] On the other hand, the fact that the monadic second-order theory of the σ_+-reduct of $\mathfrak{N}$ is Π^1_∞-complete should be viewed as folklore.

However, for our present purposes we shall utilize a more specific characterization, which was briefly discussed at the end of [17]:

Theorem 5.5

Let σ_s be $\langle 0, \mathsf{s}; = \rangle$. Then every Π^1_n-subset of $\mathbb{N}$ can be defined in $\mathfrak{N}$ — or rather in its σ_s-reduct — by a σ_s-formula of the form

$$\underbrace{\forall X^2 \, \exists Y^1_2 \, \ldots \, Y^1_n}_{n-1 \text{ alternations}} \, \Psi \tag{$\mathtt{S}_n$}$$

where X^2 is a binary set variable (intended to range over the binary relations on the natural numbers), Y^1_2, …, Y^1_n are unary set variables and Ψ contains no set quantifiers.

Proof. For convenience, denote by $\mathfrak{N}_\mathsf{s}$ the σ_s-reduct of $\mathfrak{N}$. Take

$$\sigma^\sharp_\mathsf{s} \;:=\; \langle 0, \mathsf{s}; =, X^2 \rangle$$

where X^2 is treated as a binary predicate symbol. If $R \subseteq \mathbb{N}^2$, we write $\langle \mathfrak{N}_\mathsf{s}, R \rangle$ for the $\sigma^\sharp_\mathsf{s}$-structure with domain $\mathbb{N}$ in which the symbols of σ_s are interpreted as in $\mathfrak{N}_\mathsf{s}$, and X^2 is interpreted as R.

Fix some $S \subseteq \mathbb{N}^2$ such that ordinary addition and multiplication are both first-order definable in $\langle \mathfrak{N}_\mathsf{s}, S \rangle$.[10] Let $\Psi_+(x, y, z)$ and $\Psi_\cdot(x, y, z)$ be first-order $\sigma^\sharp_\mathsf{s}$-formulas defining the corresponding functions in $\langle \mathfrak{N}_\mathsf{s}, R \rangle$. Now take Δ to be the conjunction of the following $\sigma^\sharp_\mathsf{s}$-sentences:

- $\forall x, y, z_1, z_2 \left(\Psi_+(x, y, z_1) \land \Psi_+(x, y, z_2) \to z_1 = z_2 \right)$;

[9]For similar results concerning other natural fragments of monadic second-order arithmetic, see [18] and [20].

[10]For instance, as was famously proved in [12], S can be taken to be the divisibility relation. In general, however, we do not even need 0, s and $=$ at this point.

- $\forall x\, \Psi_+(x,0,x) \wedge \forall x,y,z\, (\Psi_+(x,y,z) \to \Psi_+(x,\mathsf{s}(y),\mathsf{s}(z)))$;
- $\forall x,y,z_1,z_2\, (\Psi_\cdot(x,y,z_1) \wedge \Psi_\cdot(x,y,z_2) \to z_1 = z_2)$;
- $\forall x\, \Psi_\cdot(x,0,0) \wedge \forall x,y,z\, (\Psi_\cdot(x,y,z) \to \exists u\, (\Psi_+(z,x,u) \wedge \Psi_\cdot(x,\mathsf{s}(y),u)))$.

It is easy to see that for every $R \subseteq \mathbb{N}^2$,

$$\langle \mathfrak{N}_\mathsf{s}, R\rangle \vDash \Delta \quad \Longleftrightarrow \quad \begin{array}{c}\Psi_+(x,y,z) \text{ and } \Psi_\cdot(x,y,z) \text{ define} \\ \text{addition and multiplication respectively in } \langle \mathfrak{N}_\mathsf{s}, R\rangle.\end{array}$$

Evidently, since $\langle \mathfrak{N}_\mathsf{s}, S\rangle \vDash \Delta$, we get $\mathfrak{N}_\mathsf{s} \vDash \exists X^2\, \Delta$ where Δ is thought of as a σ_s-formula in which X^2 occurs free. So in particular, the σ_s-formulas

$$\forall X^2\, (\Delta \to \Psi_+(x,y,z)) \quad \text{and} \quad \exists X^2\, (\Delta \wedge \Psi_+(x,y,z))$$

both define addition in $\mathfrak{N}_\mathsf{s}$; similarly for multiplication. Moreover, by using $X^2(\mathsf{s}(x),y)$ instead of $X^2(x,y)$ we can turn $X^2(0,y)$ into a free unary predicate. To make this idea precise, given a $\sigma_\mathsf{s}^\sharp$-formula Φ, let

$$\Phi^* \;:=\; \begin{array}{c}\text{the result of replacing each subformula} \\ X^2(t_1,t_2) \text{ in } \Phi \text{ by } X^2(\mathsf{s}(t_1),t_2).\end{array}$$

Then for every $R \subseteq \mathbb{N}^2$,

$$\langle \mathfrak{N}_\mathsf{s}, R\rangle \vDash \Delta^* \quad \Longleftrightarrow \quad \begin{array}{c}\Psi_+^*(x,y,z) \text{ and } \Psi_\cdot^*(x,y,z) \text{ define} \\ \text{addition and multiplication respectively in } \langle \mathfrak{N}_\mathsf{s}, R\rangle.\end{array}$$

As before, we have $\mathfrak{N}_\mathsf{s} \vDash \exists X^2\, \Delta^*$. But the advantage is that $X^2(0,\,\cdot\,)$ can now be treated as a free unary set variable, which does not alter the interpretations of $\Phi_+^*(x,y,z)$ and $\Phi_\cdot^*(x,y,z)$.

Finally, let $A \subseteq \mathbb{N}$ be Π_n^1-bounded — so A is defined in $\mathfrak{N}$ by some Π_n^1-formula $\Phi(x)$. Clearly, since the Cantor pairing function is first-order definable in $\mathfrak{N}$, we can put Φ into the form

$$\underbrace{\forall Y_1^1\, \exists Y_2^1\, \ldots\, Y_n^1}_{n-1 \text{ alternations}}\, \Psi$$

where Y_1^1, Y_2^1, …, Y_n^1 are unary set variables and Ψ contains no set quantifiers. Furthermore, we may assume that every atomic subformula of Ψ has the form

$$x = y \quad \text{or} \quad \mathsf{s}(x) = y \quad \text{or} \quad x + y = z \quad \text{or} \quad x \cdot y = z \quad \text{or} \quad Y_i^1(x)$$

where $i \in \{1,\ldots,n\}$.[11] Take $\Psi^\sharp$ to be the result of replacing:

- each $x + y = z$ in Ψ by $\Phi_+(x,y,z)$;

[11] For example, the atomic σ-formula $x \cdot y + z = u \cdot v$ is equivalent to

$$\exists w_1,w_2,w_3\, (x \cdot y = w_1 \wedge w_1 + z = w_2 \wedge u \cdot v = w_3 \wedge w_2 = w_3)$$

— this illustrates a general method.

- each $x \cdot y = z$ in Ψ by $\Phi_\cdot(x, y, z)$;
- each $Y_1^1(x)$ in Ψ by $X^2(0, x)$.

It is not hard to show that for all $m \in \mathbb{N}$,

$$\mathfrak{N} \vDash \Phi(m) \quad \Longleftrightarrow \quad \mathfrak{N}_\mathsf{s} \vDash \forall X^2\, \exists Y_2^1 \,\ldots\, Y_n^1 \left(\Delta^* \to \Psi^\sharp(m)\right).$$

Thus the σ_s-formula $\forall X^2\, \exists Y_2^1 \,\ldots\, Y_n^1 \left(\Delta^* \to \Psi^\sharp(m)\right)$ does the job. $\square$

Corollary 5.6

The collection of all σ_s-sentences of the form $\mathtt{S}_n$ that are true in $\mathfrak{N}$ is Π_n^1-complete.

Proof. Denote the corresponding collection by T_n. Obviously, we have $\mathrm{T}_n \leqslant \Pi_n^1\text{-}\mathrm{Th}(\mathfrak{N})$, so T_n is Π_n^1-bounded.

Let A be a Π_n^1-complete subset of $\mathbb{N}$. By Theorem 5.5, A can be defined in $\mathfrak{N}_\mathsf{s}$ by a σ_s-formula $\Phi(x)$ of the form $\mathtt{S}_n$. So for every $m \in \mathbb{N}$,

$$m \;\in\; A \quad \Longleftrightarrow \quad \mathfrak{N}_\mathsf{s} \vDash \Phi(m) \quad \Longleftrightarrow \quad \Phi(\underline{m}) \;\in\; \mathrm{T}_n$$

where $\underline{m}$ is the numeral for m. Thus $A \leqslant \mathrm{T}_n$, so T_n is Π_n^1-hard. $\square$

This characterisation will play a key role in strengthening earlier complexity results.

6 Concerning elementary analysis

There is a well-known alternative description of the analytical hierarchy in which set variables are replaced by field variables (also called *real variables*), intended to range over $\mathbb{R}$; see [13, Chapter 16] for details. This version, known as *elementary analysis*, will be convenient for proving complexity upper bound results.

Let ς be the signature of elementary analysis. To make things precise, assume that

$$\varsigma \;:=\; \langle 0, 1, +, \,\cdot\,, -; \mathrm{Nat}, \leqslant\rangle$$

where Nat is a unary predicate symbol, intended to mean 'is a natural number'. We shall write $\mathfrak{R}^2$ for the corresponding ς-structure with domain $\mathbb{R}$. Hence for every $r \in \mathbb{R}$,

$$\mathfrak{R}^2 \vDash \mathrm{Nat}(r) \quad \Longleftrightarrow \quad r \;\in\; \mathbb{N}.$$

Unlike the case of second-order arithmetic, we shall restrict ourselves to first-order ς-formulas; so the adjective 'first-order' may be omitted, as far as $\mathfrak{R}^2$ is concerned.

Call a ς-formula Φ *arithmetical* iff every quantifier occurring in Φ is relativized by $\mathrm{Nat}(x)$, i.e. each subformula of Φ beginning with $\forall$ or $\exists$ has the form

$$\forall u\,(\mathrm{Nat}(u) \to \Theta) \quad \text{or} \quad \exists u\,(\mathrm{Nat}(u) \wedge \Theta).$$

Given $n \in \mathbb{N}_+$, say that a ς-formula is in Π^1_n iff it has the form

$$\underbrace{\forall \vec{x}_1 \, \exists \vec{x}_2 \, \ldots \, \vec{x}_n}_{n-1 \text{ alternations}} \, \Psi$$

where $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ are tuples of (field) variables and Ψ is arithmetical.

Folklore 6.1 (see [13, Chapter 16])

A subset of $\mathbb{N}$ is Π^1_n-bounded iff it can be defined in $\mathfrak{N}^2$ by a Π^1_n-formula.

Moreover, it is easy to obtain the analogues of Folklore 5.1 and 5.2, with $\mathfrak{N}$ replaced by $\mathfrak{N}^2$. In particular, for every $A \subseteq \mathbb{N}$:

$$\begin{aligned} A \ &\leqslant \ \mathrm{Th}\left(\mathfrak{N}^2\right) \quad &\Longleftrightarrow \quad &A \ \text{is} \ \Pi^1_\infty\text{-bounded;} \\ A \ &\leqslant \ \Pi^1_1\text{-}\mathrm{Th}\left(\mathfrak{N}^2\right) \quad &\Longleftrightarrow \quad &A \ \text{is} \ \Pi^1_1\text{-bounded.} \end{aligned}$$

Here $\mathrm{Th}\left(\mathfrak{N}^2\right)$ denotes the collection of all (first-order) ς-sentences true in $\mathfrak{N}^2$, and Π^1_1-$\mathrm{Th}\left(\mathfrak{N}^2\right)$ is its Π^1_1-fragment. We shall utilize these equivalences in Section 9.

7 The case of quantifiers over events

For any $m \in \mathbb{N}$ and Boolean term ϕ, take

$$m\mu(\phi) \ := \ \underbrace{\mu(\phi) + \ldots + \mu(\phi)}_{m \text{ times}}.$$

Here the empty sum is identified with the constant symbol 0, of course. Call a QPL-formula *flat* iff each of its atomic subformulas has the form

$$m\mu(\phi) \ \leqslant \ n\mu(\psi) \quad \text{or} \quad m\mu(\phi) \ \leqslant \ x$$

where m and n are in $\mathbb{N}$, ϕ and ψ are Boolean terms, x is a field variable. Naturally, we shall write $m\mu(\phi) \leqslant n$ instead of $m\mu(\phi) \leqslant n\mu(\top)$.

Recall that a class of spaces is *rich* iff it contains all infinite discrete spaces. Now Theorem 5.5 can be utilized to sharpen Theorem 2.1:

Theorem 7.1

Let $\mathcal{K}$ be a rich class of spaces. Then the flat fragment of the $\mathsf{QPL}^{\mathrm{e}}$-theory of $\mathcal{K}$ is Π^1_∞-hard.

Proof. For each $n \in \{2, 3, \ldots\}$, let

$$\mathrm{Ind}_n(X) \ := \ \forall U \left(\mathrm{At}(U) \wedge U \preccurlyeq X \rightarrow \exists V \left(\mathrm{At}(V) \wedge V \preccurlyeq X \wedge n\mu(V) = \mu(U)\right)\right).$$

In words, modulo events of measure zero, $\mathrm{Ind}_n(X)$ says 'for every atom U below X there exists an atom V below X whose measure is n times smaller than that of U'. So

$$\mathrm{Seq}_n(U, X) \ := \ \mathrm{At}(U) \wedge U \preccurlyeq X \wedge \mathrm{Ind}_n(X) \wedge (n-1)\,\mu(X) = n\mu(U)$$

says ‘U is an atom, and X is the smallest event above U satisfying $\mathrm{Ind}_n(X)$’. [12] These flat formulas will be used to mimic the natural numbers.

Next, consider the sentence

$$\Delta \ := \ \exists U\,\exists X\,(3\mu(U) = 1 \wedge \mathrm{Seq}_2(U,X) \wedge \forall V\,(\mathrm{At}(V) \wedge V \preccurlyeq X \to \exists Y\,\mathrm{Seq}_3(V,Y))).$$

It is instructive to examine the spaces in which Δ is true. To this end, take Ω to be $\{\omega^i_j \mid i,j \in \mathbb{N}\}$, and define $\mathsf{p}_\star : \Omega \to [0,1]$ by

$$\mathsf{p}_\star\left(\omega^i_j\right) \ := \ \frac{1}{2^i \cdot 3^{j+1}}.$$

Observe that

$$\sum_{i,j\in\mathbb{N}} \mathsf{p}_\star\left(\omega^i_j\right) \ = \ \sum_{i,j\in\mathbb{N}} \frac{1}{2^i \cdot 3^{j+1}} \ = \ \sum_{i\in\mathbb{N}} \frac{1}{2^i} \cdot \sum_{j\in\mathbb{N}} \frac{1}{3^{j+1}} \ = \ 2 \cdot \frac{1}{2} \ = \ 1.$$

Denote by $\mathscr{P}_\star = \langle \mathscr{A}_\star, \mathsf{P}_\star \rangle$ the corresponding discrete probability space — so $\mathscr{A}_\star$ is the power set of Ω, and $\mathsf{P}_\star : \mathscr{A}_\star \to [0,1]$ is given by

$$\mathsf{P}_\star(A) \ := \ \sum_{\omega\in A} \mathsf{p}_\star(\omega).$$

Evidently, we have $\mathscr{P}_\star \Vdash \Delta$. Moreover, one easily verifies that for every space $\mathscr{P}$, if $\mathscr{P} \Vdash \Delta$, then the quotient-space of $\mathscr{P}$ modulo events of measure zero is isomorphic to $\mathscr{P}_\star$. Thus, in a sense, the only model of Δ is $\mathscr{P}_\star$.

We are going to interpret arithmetic within $\mathscr{P}_\star$. For this purpose, it is convenient to view Ω as an infinite matrix: for any $i,j \in \mathbb{N}$,

$$\Omega^i \ := \ \left\{\omega^i_j \mid j \in \mathbb{N}\right\} \quad \text{and} \quad \Omega_j \ := \ \left\{\omega^i_j \mid i \in \mathbb{N}\right\}$$

correspond to the ith row and jth column respectively. Now consider the following formulas:

$$\begin{aligned}
\mathrm{Row}^0(X) \ &:= \ \exists U\,(3\mu(U) = 1 \wedge \mathrm{Seq}_3(U,X));\\
\mathrm{Col}^0(X) \ &:= \ \exists U\,(3\mu(U) = 1 \wedge \mathrm{Seq}_2(U,X));\\
\mathrm{Row}(X) \ &:= \ \exists Y\,\exists U\,\left(\mathrm{Col}^0(Y) \wedge \mathrm{At}(U) \wedge U \preccurlyeq Y \wedge \mathrm{Seq}_2(U,X)\right);\\
\mathrm{Col}(X) \ &:= \ \exists Y\,\exists U\,\left(\mathrm{Row}^0(Y) \wedge \mathrm{At}(U) \wedge U \preccurlyeq Y \wedge \mathrm{Seq}_3(U,X)\right);\\
\mathrm{Diag}(X) \ &:= \ \exists U\,(3\mu(U) = 1 \wedge \mathrm{Seq}_6(U,X));\\
\mathrm{Match}(X,Y) \ &:= \ \exists Z\,(\mathrm{Diag}(Z) \wedge \mu(X \wedge Y \wedge Z) \neq 0).
\end{aligned}$$

[12] To express the same condition more directly, one might want to replace the conjunct $(n-1)\mu(X) = n\mu(U)$ by $\neg\exists Y\,(U \preccurlyeq Y \wedge \mathrm{Ind}_n(Y) \wedge Y \prec X)$. However, the original formula is simpler and it works nicely in $\underline{\mathsf{QPL}}^{\mathrm{e}}$, where we deal with Boolean algebras which are not necessarily closed under countable joins.

Their meanings are straightforward:

$$\begin{aligned}
\mathrm{Row}^0(X) &: \text{ `X is the 0th row'};\\
\mathrm{Col}^0(X) &: \text{ `X is the 0th column'};\\
\mathrm{Row}(X) &: \text{ `X is a row''};\\
\mathrm{Col}(X) &: \text{ `X is a column'};\\
\mathrm{Diag}(X) &: \text{ `X is the diagonal'};\\
\mathrm{Match}(X,Y) &: \text{ `X and Y have a diagonal element in common'}
\end{aligned}$$

where by the 'diagonal' we mean

$$\underline{\overline{\Omega}} := \{\omega_i^i \mid i \in \mathbb{N}\}.$$

In addition, $\mathrm{Match}(X,Y)$ allows us to switch from rows to columns, and vice versa, since for any $i, j \in \mathbb{N}$,

$$\mathscr{P}_\star \Vdash \mathrm{Match}\left(\Omega^i, \Omega_j\right) \quad \Longleftrightarrow \quad i = j.$$

Obviously, if we think of natural numbers as rows, then the successor function can be defined by $\mu(X) = 2\mu(Y)$. To interpret a binary set variable, we shall use the formula

$$\Gamma(X,Y,Z) := \exists Y^* (\mathrm{Col}(Y^*) \wedge \mathrm{Match}(Y,Y^*) \wedge \mu(X \wedge Y^* \wedge Z) \neq 0).$$

To see how it works, observe that for every $A \subseteq \mathbb{N}^2$,

$$A = \left\{(i,j) \in \mathbb{N} \mid \mathscr{P}_\star \Vdash \Gamma\left(\Omega^i, \Omega^j, A'\right)\right\}$$

where A' denotes $\left\{\omega_j^i \mid (i,j) \in A\right\}$. Thus elements of $\mathscr{A}_\star$ may be treated as binary relations on $\mathbb{N}$. Unary set variables are even easier to handle.

Finally, we are ready to reduce the union of the sets $\Pi_1^1\text{-}\mathrm{Th}(\mathfrak{N})$, $\Pi_2^1\text{-}\mathrm{Th}(\mathfrak{N})$, … to the flat fragment of $\mathrm{Th}^{\mathrm{e}}(\mathcal{K})$ — this will imply the desired result, by Folklore 5.2. Let Φ be a Π_n^1-sentence, for some $n \in \mathbb{N}_+$. As the proof of Theorem 5.5 shows, Φ can be effectively converted to an equivalent σ-sentence $\Phi^\natural$ of the form S_n. Without loss of generality, we may assume that:

- each atomic subformula of $\Phi^\natural$ has the form

  $$x = y \quad \text{or} \quad \mathsf{s}(x) = y \quad \text{or} \quad X^2(x,y) \quad \text{or} \quad Y_i^1(x)$$

  where $i \in \{2, 3, \ldots, n\}$;
- $\rightarrow$ and $\vee$ do not occur in $\Phi^\natural$, although $\wedge$ and $\neg$ may occur in it.

For convenience, the set variables X^2, Y_2^1, Y_3^1, …, Y_n^1 will also be treated as distinguished Boolean variables in QPL. Now define $\tau\left(\Phi^{\natural}\right)$ recursively:

$$\begin{aligned}
\tau(x=y) &:= \mu(X)=\mu(Y); \\
\tau(\mathsf{s}(x)=y) &:= \mu(X)=2\mu(Y); \\
\tau\left(X^2(x,y)\right) &:= \Gamma\left(X,Y,X^2\right); \\
\tau\left(Y_i^1(x)\right) &:= X \preccurlyeq Y_i^1; \\
\tau(\Psi\wedge\Theta) &:= \tau(\Psi)\wedge\tau(\Theta); \\
\tau(\neg\Psi) &:= \neg\tau(\Psi); \\
\tau(\forall x\,\Psi) &:= \forall X\,(\mathrm{Row}(X)\to\tau(\Psi)); \\
\tau(\exists x\,\Psi) &:= \exists X\,(\mathrm{Row}(X)\wedge\tau(\Psi)); \\
\tau\left(\forall X^2\,\Psi\right) &:= \forall X^2\,\tau(\Phi); \\
\tau\left(\forall Y_i^1\,\Psi\right) &:= \forall Y_i^1\,\tau(\Psi); \\
\tau\left(\exists Y_i^1\,\Psi\right) &:= \exists Y_i^1\,\tau(\Psi).
\end{aligned}$$

By construction, $\tau\left(\Phi^{\natural}\right)$ is always flat. And it is straightforward to verify that

$$\mathfrak{N}\vDash\Phi \quad\Longleftrightarrow\quad \Delta\to\tau\left(\Phi^{\natural}\right)\ \in\ \mathrm{Th}^{\mathrm{e}}(\mathcal{K}).$$

This gives us the desired reduction. □

Moreover, by Theorem 2.2, whenever $\mathcal{K}$ is analytical, then we may replace 'Π^1_∞-hard' by 'Π^1_∞-complete' in the formulation of the last result.

8 The case of quantifiers over propositional formulas

The property of being *flat* in $\underline{\mathsf{QPL}}$ is defined as in QPL, except that we allow elements of Prop to occur in Boolean terms.

Now the proof of Theorem 7.1 can be modified to sharpen Theorem 4.1:

Theorem 8.1

Let $\mathcal{K}$ be a rich class of expanded probability spaces. Then the flat fragment of the $\underline{\mathsf{QPL}}^{\mathrm{e}}$-theory of $\mathcal{K}$ is Π^1_1-hard.

Proof. We shall employ the notation of the proof of Theorem 7.1. Take

$$\mathscr{S}_\star := \left\{\Omega^i \mid i\in\mathbb{N}\right\}\cup\{\Omega_j \mid j\in\mathbb{N}\}\cup\left\{\underline{\Omega}\right\}.$$

Call a function π from Prop to $\mathscr{A}_\star$ *admissible* iff $\mathscr{S}_\star\subseteq\{\pi(\phi)\mid\phi\in\mathrm{Term}^\circ\}$. Here we may replace $\mathscr{S}_\star$ by the Boolean algebra generated by $\mathscr{S}_\star$, of course. Notice that for every $\pi:\mathrm{Prop}\to\mathscr{A}_\star$,

$$\langle\mathscr{P}_\star,\pi\rangle\Vdash\Delta \quad\Longleftrightarrow\quad \pi\ \text{ is admissible.}$$

Since we have '$\subseteq$', not '$=$', there will still be a way of interpreting a binary set variable.

Let Φ be a Π^1_1-sentence. Again, Φ can be effectively converted to an equivalent σ-sentence $\Phi^\natural$ of the form S_1. Without loss of generality, we may assume that:

- each atomic subformula of $\Phi^\natural$ has the form

$$x = y \quad \text{or} \quad \mathsf{s}(x) = y \quad \text{or} \quad X^2(x, y);$$

- $\to$ and $\vee$ do not occur in $\Phi^\natural$.

As before, the set variable X^2 will also be treated as a distinguished Boolean variable. Obviously, for every $A \subseteq \Omega$ there exists $\pi : \mathrm{Prop} \to \mathscr{A}_\star$ such that

$$A \in \{\pi(\phi) \mid \phi \in \mathrm{Term}^\circ\}$$

— even though one needs uncountably many π's to cover all subsets of A. It follows that

$$\mathfrak{N} \models \Phi \quad \Longleftrightarrow \quad \Delta \to \tau\left(\Phi^\natural\right) \in \underline{\mathrm{Th}}^{\mathrm{e}}(\mathcal{K}).$$

This gives us the desired reduction. $\square$

By adding quantifiers over reals we get:

Theorem 8.2

Let $\mathcal{K}$ be a rich class of expanded probability spaces. Then the flat fragment of the $\underline{\mathsf{QPL}}$-theory of $\mathcal{K}$ is Π^1_∞-hard.

Proof. Again, we shall employ the notation of the proof of Theorem 7.1. In particular, the formula $\Gamma(X, Y, Z)$ will be used to interpret a free binary set variable, which is intuitively bounded by the outermost universal quantifier; cf. the proof of Theorem 8.1. However, unary set variables — each of which may be bounded by $\forall$ or $\exists$ — will be handled using field variables.

As is well known, every $\varepsilon \in [0, 1)$ can be uniquely represented as

$$\varepsilon = \sum_{i=0}^{\infty} \frac{\varepsilon_i}{2^{i+1}}$$

where each ε_i is either 0 or 1, and the sequence $\varepsilon_0, \varepsilon_1, \ldots$ contains infinitely many 0's. It is easy to verify that

$$\varepsilon_0 = 1 \quad \Longleftrightarrow \quad \varepsilon \geqslant \frac{1}{2},$$

and moreover, for all $k \in \mathbb{N}_+$,

$$\varepsilon_k = 1 \quad \Longleftrightarrow \quad \varepsilon \geqslant \sum_{i=0}^{k-1} \frac{\varepsilon_i}{2^{i+1}} + \frac{1}{2^{k+1}}.$$

Intuitively, we shall think of ε as the set $\{i \in \mathbb{N} \mid \varepsilon_i = 1\}$. This will give us all subsets of $\mathbb{N}$ whose complements are not finite. Also note that for every $S \subseteq \mathbb{N}$ there exist $\varepsilon^+, \varepsilon^- \in [0,1)$ such that

$$S \;=\; \big\{i \in \mathbb{N} \mid \varepsilon^+_i = 1\big\} \cup \big\{i \in \mathbb{N} \mid \varepsilon^-_i = 0\big\}.$$

Naturally, we want to interpret arbitrary subsets of $\mathbb{N}$ as elements of $[0,1)^2$.

To make the above idea work, we introduce some additional formulas:

$$\begin{aligned} \mathrm{Rows}\,(X) \;&:=\; \forall U\,(\mathrm{Row}\,(U) \wedge \mu\,(X \wedge U) \neq 0 \to U \preccurlyeq X); \\ \mathrm{Upper}\,(U, X) \;&:=\; \mathrm{Rows}\,(X) \wedge \forall V\,(\mathrm{Row}\,(V) \to (V \preccurlyeq X \leftrightarrow \mu\,(U) < \mu\,(V))). \end{aligned}$$

With $\mathscr{P}_\star$ in mind, their meanings are straightforward:

$$\begin{aligned} \mathrm{Rows}\,(X) \;&:\; \text{‘}X \text{ is a union of rows’}; \\ \mathrm{Upper}\,(U, X) \;&:\; \text{‘}X \text{ is the union of all rows that are more probable than } U\text{’}. \end{aligned}$$

Furthermore, if $\mu\,(U) > 0$, then $\mathrm{Upper}\,(U, X)$ implies that X is a Boolean combination of rows — because it includes only finitely many rows. Now take

$$\begin{aligned} \mathrm{Approx}\,(X, x, U) \;:=\;& \forall V\,(\mathrm{Row}\,(V) \wedge V \preccurlyeq X \to \mu\,(U) \leqslant \mu\,(V)) \wedge \\ & \exists V\,(\mathrm{Row}\,(V) \wedge 2\mu\,(V) = 1 \wedge (V \preccurlyeq X \leftrightarrow \mu\,(V) \leqslant x)) \wedge \\ & \forall V\,(\mathrm{Row}\,(V) \wedge \mu\,(U) \leqslant \mu\,(V) \to \\ & \qquad (V \preccurlyeq X \leftrightarrow \exists Y\,(\mathrm{Upper}\,(V, Y) \wedge \mu\,((X \wedge Y) \vee V) \leqslant x))). \end{aligned}$$

It is not hard to verify that for any $k \in \mathbb{N}$, $S \subseteq \mathbb{N}$ and $\varepsilon \in [0,1)$,

$$\mathscr{P}_\star \Vdash \mathrm{Approx}\left(\bigcup\big\{\Omega^i \mid i \in S\big\}, \varepsilon, \Omega^k\right) \quad\iff\quad S \;=\; \{i \in \mathbb{N} \mid i \leqslant k \text{ and } \varepsilon_i = 1\}.$$

Consequently, for the formula

$$\Sigma\,(U, x) \;:=\; \exists X\,(\mathrm{Approx}\,(X, x, U) \wedge U \preccurlyeq X)$$

we have $\mathscr{P}_\star \Vdash \Sigma\left(\Omega^k, \varepsilon\right)$ iff $\varepsilon_k = 1$. The reader should be warned: even though ε plays the role of $\bigcup\{\Omega^i \mid \varepsilon_i = 1\}$, the latter is not necessarily expressible by means of a propositional formula.

Let Φ be a Π^1_n-sentence, for some $n \in \mathbb{N}_+$. Convert Φ to $\Phi^\natural$, as in the proof of Theorem 7.1. For convenience, with each unary set variable Y^1_i, we associate a pair y^+_i, y^-_i of distinguished field variables. Now define $\rho\left(\Phi^{\natural}\right)$ recursively:

$$\begin{aligned}
\rho\left(x=y\right) &:= \mu\left(X\right)=\mu\left(Y\right);\\
\rho\left(\mathsf{s}\left(x\right)=y\right) &:= \mu\left(X\right)=2\mu\left(Y\right);\\
\rho\left(X^{2}\left(x,y\right)\right) &:= \Gamma\left(X,Y,X^{2}\right);\\
\rho\left(Y_i^{1}\left(x\right)\right) &:= \Sigma\left(X,y_i^{+}\right)\vee\neg\Sigma\left(X,y_i^{-}\right);\\
\rho\left(\Psi\wedge\Theta\right) &:= \rho\left(\Psi\right)\wedge\rho\left(\Theta\right);\\
\rho\left(\neg\Psi\right) &:= \neg\rho\left(\Psi\right);\\
\rho\left(\forall x\,\Psi\right) &:= \forall X\left(\mathrm{Row}\left(X\right)\to\rho\left(\Psi\right)\right);\\
\rho\left(\exists x\,\Psi\right) &:= \exists X\left(\mathrm{Row}\left(X\right)\wedge\rho\left(\Psi\right)\right);\\
\rho\left(\forall X^{2}\,\Psi\right) &:= \forall X^{2}\,\rho\left(\Phi\right);\\
\rho\left(\forall Y_i^{1}\,\Psi\right) &:= \forall y_i^{+},y_i^{-}\left(0\leqslant y_i^{+},y_i^{-}<1\to\rho\left(\Psi\right)\right);\\
\rho\left(\exists Y_i^{1}\,\Psi\right) &:= \exists y_i^{+},y_i^{-}\left(0\leqslant y_i^{+},y_i^{-}<1\wedge\rho\left(\Psi\right)\right).
\end{aligned}$$

By construction, $\rho\left(\Phi^{\natural}\right)$ is always flat. And it is straightforward to check that

$$\mathfrak{N}\vDash\Phi \quad\Longleftrightarrow\quad \Delta\to\rho\left(\Phi^{\natural}\right)\ \in\ \underline{\mathrm{Th}}\left(\mathcal{K}\right).$$

This gives us the desired reduction. □

As will be shown shortly, the lower bounds provided by Theorems 8.1 and 8.2 often turn out to be precise.

9 Some general upper bounds

A general upper bound argument for QPL, as well as $\mathsf{QPL}^{\mathrm{e}}$, is given by Theorem 2.2. The cases of $\underline{\mathsf{QPL}}$ and $\underline{\mathsf{QPL}}^{\mathrm{e}}$ are, in fact, much simpler and quite straightforward. Still, to keep the presentation self-contained, we are going to provide the corresponding arguments below.

It is well known that $\mathbb{R}^{\mathbb{N}}$ — i.e. the collection of all functions from $\mathbb{N}$ to $\mathbb{R}$ — can be represented in elementary analysis using only quantifiers over natural numbers. To be precise, one can find an arithmetical ς-formula $\mathrm{Split}\left(x,u,v\right)$ having the following properties:

- $\mathrm{Split}\left(x,u,v\right)$ defines in $\mathfrak{R}^2$ the graph of a function from $\mathbb{R}\times\mathbb{N}$ to $\mathbb{R}$;
- for every $f:\mathbb{N}\to\mathbb{R}$ there exists $c\in\mathbb{R}$ such that

  $$f \;=\; \left\{(n,r)\in\mathbb{N}\times\mathbb{R}\mid\mathfrak{R}^2\vDash\mathrm{Split}\left(c,n,r\right)\right\}.$$

  — intuitively, f is *coded* by c.

Since $\mathbb{R}^{\mathbb{N}}$ has the same cardinality as $\mathbb{R}$, this should not come as a surprise.[13] For convenience, fix a special variable $\mathfrak{m}$, intended to range over $\mathbb{R}^{\mathbb{N}}$. Formally, it means that $\mathfrak{m}$ is, in fact, a field variable, and all expressions of the form $\mathfrak{m}(t)$ in a ς-formula Φ have to be systematically eliminated:

$$\_\_\_\_ \, \mathfrak{m}(t) \, \_\_\_\_ \quad \longmapsto \quad \forall u \left(\mathrm{Split}(\mathfrak{m}, t, u) \to \_\_\_\_ \, u \, \_\_\_\_\right)$$

where u is a fresh field variable. To avoid confusion, we shall always make sure that t (in $\mathfrak{m}(t)$) is interpreted as a natural number. Notice that we may also use

$$\exists u \left(\mathrm{Split}(\mathfrak{m}, t, u) \wedge \_\_\_\_ \, u \, \_\_\_\_\right)$$

instead of $\forall u \left(\mathrm{Split}(\mathfrak{m}, t, u) \to \_\_\_\_ \, u \, \_\_\_\_\right)$. Thus the process of elimination preserves the prefix classification described in Section 6.

Let $\#$ be an effective one-one mapping from Term° onto $\mathbb{N}$. Given $f : \mathrm{Term}^{\circ} \to \mathbb{R}$, denote by $\#f$ the function from $\mathbb{N}$ to $\mathbb{R}$ given by

$$\#f(\#\phi) \; := \; f(\phi).$$

Next, with each class $\mathcal{K}$ of expanded spaces, we associate the set

$$\mathcal{K}^{\#} \; := \; \left\{c \in \mathbb{R} \mid \text{there exists } \mathfrak{A} \in \mathcal{K} \text{ such that } \#\mathsf{p}^{\mathfrak{A}} \text{ is coded by } c\right\}.$$

Naturally, we call $\mathcal{K}$ *analytical* iff $\mathcal{K}^{\#}$ is definable in $\mathfrak{R}^2$.

Theorem 9.1

Let $\mathcal{K}$ be an analytical class of expanded spaces. Then the $\underline{\mathsf{QPL}}$-theory of $\mathcal{K}$ is Π^1_∞-bounded.

Proof. We are going to reduce $\underline{\mathrm{Th}}(\mathcal{K})$ to $\mathrm{Th}(\mathfrak{R}^2)$.

Let $conj, disj : \mathbb{N}^2 \to \mathbb{N}$ and $neg : \mathbb{N} \to \mathbb{N}$ be given by:

$$\begin{aligned} conj(\#\phi, \#\psi) \; &:= \; \#(\phi \wedge \psi); \\ disj(\#\phi, \#\psi) \; &:= \; \#(\phi \vee \psi); \\ neg(\#\phi) \; &:= \; \#(\neg\phi). \end{aligned}$$ [14]

Obviously, these functions are computable, and hence definable in $\mathfrak{R}^2$ by means of arithmetical ς-formulas. To simplify the exposition, we shall pretend that ς contains the corresponding function symbols conj, disj and neg — they may be easily eliminated if needed.

With each Boolean variable X, we associate a distinguished field variable $\underline{x}$, intended to range over $\mathbb{N}$. Then, given a Boolean term ϕ, we define the ς-term $\mathrm{code}(\phi)$ recursively:

$$\begin{aligned} \mathrm{code}(\delta) \; &:= \; \#\delta \quad \text{for all } \; \delta \in \{\bot, \top\} \cup \mathrm{Prop}; \\ \mathrm{code}(X) \; &:= \; \underline{x}; \\ \mathrm{code}(\psi \wedge \theta) \; &:= \; \mathrm{conj}(\mathrm{code}(\psi), \mathrm{code}(\theta)); \\ \mathrm{code}(\psi \vee \theta) \; &:= \; \mathrm{disj}(\mathrm{code}(\psi), \mathrm{code}(\theta)); \\ \mathrm{code}(\neg\psi) \; &:= \; \mathrm{neg}(\mathrm{code}(\psi)). \end{aligned}$$

[13] On the other hand, it is impossible to represent $\mathbb{N}^{\mathbb{R}}$ in elementary analysis.

[14] Here ϕ and ψ range over Term°.

Also, given a field term t, take

$$\gamma(t) \;:=\; \text{the result of replacing each } \mu(\phi) \text{ in } t \text{ by } \mathfrak{m}(\mathrm{code}(\phi))$$

where $\mathfrak{m}$ is intended to range over $\mathbb{R}^{\mathbb{N}}$, as described earlier.

Let Φ be a $\underline{\mathsf{QPL}}$-sentence. Define $\eta(\Phi)$ as follows:

$$\begin{aligned}
\eta(t_1 \leqslant t_2) &:= \gamma(t_1) \leqslant \gamma(t_2);\\
\eta(\Psi \wedge \Theta) &:= \eta(\Psi) \wedge \eta(\Theta);\\
\eta(\Psi \vee \Theta) &:= \eta(\Psi) \vee \eta(\Theta);\\
\eta(\neg\Psi) &:= \neg\eta(\Psi);\\
\eta(\forall X\, \Psi) &:= \forall \underline{x}\,(\mathrm{Nat}(\underline{x}) \to \eta(\Psi));\\
\eta(\exists X\, \Psi) &:= \exists \underline{x}\,(\mathrm{Nat}(\underline{x}) \wedge \eta(\Psi));\\
\eta(\forall x\, \Psi) &:= \forall x\, \eta(\Psi);\\
\eta(\exists x\, \Psi) &:= \exists x\, \eta(\Psi).
\end{aligned}$$

Finally, since $\mathcal{K}$ is analytical, there exists a ς-formula $\Delta(x)$ that defines $\mathcal{K}^{\#}$ in $\mathfrak{R}^2$. It is straightforward to verify that

$$\Phi \in \underline{\mathrm{Th}}(\mathcal{K}) \quad\Longleftrightarrow\quad \mathfrak{R}^2 \vDash \forall \mathfrak{m}\,(\Delta(\mathfrak{m}) \to \eta(\Phi)),$$

which gives us the desired reduction. □

Hence, whenever $\mathcal{K}$ is analytical, we may replace 'Π^1_∞-hard' by 'Π^1_∞-complete' in the formulation of Theorem 8.2. Clearly, Theorem 9.1 covers a wide range of classes of expanded spaces.

Further, we say that a ς-formula is in Σ^1_1 iff it has the form $\exists \vec{x}\, \Psi$ where $\vec{x}$ is a tuple of variables and Ψ is arithmetical. Thus Σ^1_1-formulas may be viewed as negated Π^1_1-formulas. Call a class $\mathcal{K}$ of expanded spaces *existential* iff $\mathcal{K}^{\#}$ is definable in $\mathfrak{R}^2$ by means of a Σ^1_1-formula.

Theorem 9.2

Let $\mathcal{K}$ be an existential class of expanded spaces. Then the $\underline{\mathsf{QPL}}^{\mathrm{e}}$-theory of $\mathcal{K}$ is Π^1_1-bounded.

Proof. Pick a Σ^1_1-formula $\Delta(x)$ that defines $\mathcal{K}^{\#}$ in $\mathfrak{R}^2$. Observe that for each $\underline{\mathsf{QPL}}^{\mathrm{e}}$-sentence Φ, we can effectively convert $\forall \mathfrak{m}\,(\Delta(\mathfrak{m}) \to \eta(\Phi))$ to an equivalent Π^1_1-sentence. □

Therefore, whenever $\mathcal{K}$ is existential, we may replace 'Π^1_1-hard' by 'Π^1_1-complete' in the formulation of Theorem 8.1.

For instance, take $\mathcal{K}$ to be the class of all expanded spaces. It is instructive to see how $\mathcal{K}^{\#}$ can be defined in $\mathfrak{R}^2$. Notice that since propositional logic is decidable, there exists an arithmetical ς-formula $\mathrm{Eq}(x, y)$ such that for any $\phi, \psi \in \mathrm{Term}^{\circ}$,

$$\mathfrak{R}^2 \vDash \mathrm{Eq}(\#\phi, \#\psi) \quad\Longleftrightarrow\quad \phi \text{ and } \psi \text{ are logically equivalent.}$$

Hence — letting conj and neg be as in the proof of Theorem 9.1 — the ς-formula

$$\begin{aligned}\Gamma(\mathfrak{m}) \;:=\; &\forall x\,(0 \leqslant \mathfrak{m}(x) \leqslant 1) \wedge \\ &\forall x\,\forall y\,(\mathrm{Nat}(x) \wedge \mathrm{Nat}(y) \wedge \mathrm{Eq}(x,y) \to \mathfrak{m}(x) = \mathfrak{m}(y)) \wedge \\ &\forall x\,\forall y\,(\mathrm{Nat}(x) \wedge \mathrm{Nat}(y) \to \mathfrak{m}(x) = \mathfrak{m}(\mathrm{conj}(x,y)) + \mathfrak{m}(\mathrm{conj}(x,\mathrm{neg}(y))))\end{aligned}$$

says that $\mathfrak{m}$ codes some acceptable function from Term° to $[0,1]$. Thus $\Gamma(\mathfrak{m})$ defines $\mathcal{K}^{\#}$ in $\mathfrak{R}^2$, by Proposition 4.2. Further, to deal with expanded spaces of a special kind, we may add suitable formulas to $\Gamma(\mathfrak{m})$. As an example, the ς-formula

$$\forall x\,(\mathrm{Nat}(x) \wedge \mathfrak{m}(x) > 0 \to \exists u\,(\mathrm{Nat}(u) \wedge 0 < \mathfrak{m}(u) < \mathfrak{m}(x)))$$

says that the acceptable function f coded by $\mathfrak{m}$ has infinite range — which is equivalent to the condition that for every expanded space $\mathfrak{A} = \langle\langle \mathscr{A}, \mathsf{P}\rangle, \pi\rangle$,

$$\mathsf{p}^{\mathfrak{A}} = f \quad\Longrightarrow\quad \mathscr{A}_{\pi} \text{ is infinite, modulo events of measure zero.}$$

Other reasonable properties can be handled similarly.

Acknowledgments. This work was performed at the Steklov International Mathematical Center and supported by the Ministry of Science and Higher Education of the Russian Federation (agreement no. 075-15-2022-265).

References

[1] M. Abadi, J. Y. Halpern. Decidability and expressiveness for first-order logics of probability. *Information and Computation* 112(1), 1–36, 1994.

[2] Yu. L. Ershov, I. A. Lavrov, A. D. Taimanov, M. A. Taitslin. Elementary theories, *Russian Mathematical Surveys* 20(4), 35–105, 1965.

[3] R. Fagin, J. Y. Halpern, N. Megiddo. A logic for reasoning about probabilities. *Information and Computation* 87(1–2), 78–128, 1990.

[4] J. Y. Halpern. Presburger arithmetic with unary predicates is Π^1_1 complete. *Journal of Symbolic Logic* 56(2), 637–642, 1991.

[5] D. N. Hoover. Probability logic. *Annals of Mathematical Logic* 14(3), 287–313, 1978.

[6] D. Ibeling, T. Icard, K. Mierzewski, M. Mossé. Probing the quantitative–qualitative divide in probabilistic reasoning. *Annals of Pure and Applied Logic* 175(9), 103339, 2024.

[7] H. J. Keisler. Probability quantifiers. In J. Barwise, S. Feferman (eds.), *Model-Theoretic Logics*, pp. 509–556. Springer, 1985.

[8] Z. Marković, M. Rašković. Some new probability operators. In Z. Ognjanović (ed.), *Probabilistic Extensions of Various Logical Systems*, pp. 143–163. Springer, 2020.

[9] Y. N. Moschovakis. *Elementary Induction on Abstract Structures.* North-Holland Publishing Company, 1974.

[10] Z. Ognjanović, M. Rašković, Z. Marković. *Probability Logics.* Springer, 2016.

[11] A. Perović, Z. Ognjanović, M. Rašković, Z. Marković. A probabilistic logic with polynomial weight formulas. In S. Hartmann, G. Kern-Isberner (eds.), *Foundations of Information and Knowledge Systems 2008*, LNCS 4932, pp. 239–252. Springer, 2008.

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

[13] H. Rogers, Jr. *Theory of Recursive Functions and Effective Computability.* McGraw-Hill Book Company, 1967.

[14] R. M. Solovay, R. D. Arthan, J. Harrison. Some new results on decidability for elementary algebra and geometry. *Annals of Pure and Applied Logic* 163(12), 1765–1802, 2012.

[15] S. O. Speranski. Quantification over propositional formulas in probability logic: Decidability issues. *Algebra and Logic* 50(4), 365–374, 2011.

[16] S. O. Speranski. Complexity for probability logic with quantifiers over propositions. *Journal of Logic and Computation* 23(5), 1035–1055, 2013.

[17] S. O. Speranski. A note on definability in fragments of arithmetic with free unary predicates. *Archive for Mathematical Logic* 52 (5–6), 507–516, 2013.

[18] S. O. Speranski. Some new results in monadic second-order arithmetic. *Computability* 4(2), 159–174, 2015.

[19] S. O. Speranski. Quantifying over events in probability logic: An introduction. *Mathematical Structures in Computer Science* 27(8), 1581–1600, 2017.

[20] S. O. Speranski, F. N. Pakhomov. On the coprimeness relation from the viewpoint of monadic second-order logic. *Izvestiya: Mathematics* 86(6), 1225–1239, 2022.

[21] S. O. Speranski. An 'elementary' perspective on reasoning about probability spaces. *Logic Journal of the IGPL*, jzae042, 2024. Published online.

[22] A. Tarski. *A Decision Method for Elementary Algebra and Geometry*. University of California Press, 1951.

[23] A. Tarski, A. Mostowski, R. Robinson. *Undecidable Theories*. North-Holland, 1953.

Stanislav O. Speranski
Steklov Mathematical Institute of Russian Academy of Sciences
8 Gubkina St., 119991 Moscow, Russia
katze.tail@gmail.com