Annals of Pure and Applied Logic 163 (2012) 1765–1802

Contents lists available at SciVerse ScienceDirect

Annals of Pure and Applied Logic

www.elsevier.com/locate/apal

Some new results on decidability for elementary algebra and geometry

Robert M. Solovay [a], R.D. Arthan [b,c,*], John Harrison [d]

[a] PO Box 5949, Eugene, OR 97405, USA
[b] Lemma 1. Ltd., 27 Brook St., Twyford, Berkshire, RG10 9NX, UK
[c] School of Electronic Engineering and Computer Science, Queen Mary, University of London, London, E1 4NS, UK
[d] Intel Corporation, RA2-451, 2501 NW 229th Avenue, Hillsboro, OR 97124, USA

ARTICLE INFO

Article history:
Received 15 April 2009
Received in revised form 2 March 2012
Accepted 9 April 2012
Available online 10 May 2012
Communicated by I. Moerdijk

MSC:
03D35
03C10
03C65

Keywords:
Decidability
Undecidability
Banach spaces
Hilbert spaces

ABSTRACT

We carry out a systematic study of decidability for theories (a) of real vector spaces, inner product spaces, and Hilbert spaces and (b) of normed spaces, Banach spaces and metric spaces, all formalized using a 2-sorted first-order language. The theories for list (a) turn out to be decidable while the theories for list (b) are not even arithmetical: the theory of 2-dimensional Banach spaces, for example, has the same many-one degree as the set of truths of second-order arithmetic.

We find that the purely universal and purely existential fragments of the theory of normed spaces are decidable, as is the ∀∃ fragment of the theory of metric spaces. These results are sharp of their type: reductions of Hilbert's 10th problem show that the ∃∀ fragments for metric and normed spaces and the ∀∃ fragment for normed spaces are all undecidable.

1. Introduction

It is natural to formulate the theory of real vector spaces using a 2-sorted first-order language with a sort for the scalars and a sort for the vectors. Introduction of coordinates reduces the theory VS^n of a vector space of a given finite dimension n to the first-order theory of the real numbers, known to be decidable since the pioneering work of Tarski [42] to which our title alludes. The purpose of this paper is to investigate decidability for more general classes of real vector spaces.

We will consider real vector spaces equipped with an inner product or a norm, possibly required to be complete (i.e., to be Hilbert spaces or Banach spaces) and under various restrictions on the dimension and often with multiplication disallowed. So, for example, we will find that the theories IP^∞ and HS^∞ of infinite-dimensional inner product spaces and Hilbert spaces respectively are both decidable, and in fact by the same decision procedure, so that the two theories coincide. By contrast, we will see that the analogous theories NS^∞ and BS^∞ of infinite-dimensional normed spaces and Banach spaces differ, and both are undecidable, as is the purely additive fragment BS^d_+ of the theory of d-dimensional Banach spaces for $d \geqslant 2$.

In fact, all the theories of inner product spaces we consider are decidable, while for normed spaces, only the most trivial example, namely the theory of a 1-dimensional space, is decidable. The undecidable normed space theories are not recursively axiomatizable or even arithmetical, as we will see by constructing primitive recursive reductions of the set of

* Corresponding author at: School of Electronic Engineering and Computer Science, Queen Mary, University of London, London, E1 4NS, UK.
E-mail addresses: solovay@gmail.com (R.M. Solovay), rda@lemma-one.com (R.D. Arthan), johnh@ichips.intel.com (J. Harrison).

0168-0072/$ – see front matter
http://dx.doi.org/10.1016/j.apal.2012.04.003

truths of second-order arithmetic to these theories. In fact, if we restrict to normed spaces of finite dimension, the normed space theories have the same degree of unsolvability (many-one degree) as second-order arithmetic, while for arbitrary dimensions, the normed space and Banach space theories are many-one equivalent to the set of true Π^2_1 sentences in third-order arithmetic.

Normed spaces and inner product spaces are vector spaces with a metric that relates nicely to the algebraic structure. We therefore consider metric spaces as a source of motivating examples and for their own interest. The theory MS of metric spaces is known to be undecidable [9,28]. We give an alternative proof which shows that the theory is not arithmetical.

We obtain positive decidability results for normed spaces by restricting the use of quantifiers: rather trivially, the set of valid purely existential sentences is decidable, but much more interestingly, so is the set of valid purely universal sentences. The decision procedure for the purely universal case is via a computational process which (at least in principle) produces a concrete counter-example in the shape of an explicit norm on $\mathbb{R}^n$ for some n which fails to satisfy a given invalid sentence. This algorithm has been implemented in the special case where multiplication is not allowed. For metric spaces, we do even better: the set of valid ∀∃ sentences is decidable, as we see using an analogue of the Bernays–Schönfinkel decision procedure for valid ∀∃ sentences in a first-order language with no function symbols. However, by reducing satisfiability for quantifier-free formulas of arithmetic to the dual satisfiability problem, we will find that validity for the ∃∀ fragment is undecidable for both metric spaces and normed spaces, as is validity for the ∀∃ fragment for normed spaces. Finding other useful decidable fragments is an interesting challenge.

The structure of the sequel is as follows:

Section 2 introduces notation and terminology and then gives some preliminary observations and results. We assume that the reader is acquainted with the concept of a many-sorted first-order language as described, for example, in the book by Manzano [29]. However, as an *aide-memoire* to make the material more easily accessible to readers without a professional background in pure mathematics, we review many of the ideas from vector algebra and affine geometry that we will use. We then make some initial observations on the possibilities for decision procedures in the theories of interest. This leads on to a number of examples showing the expressive strength of the language of normed spaces compared with the language of inner product spaces. For example, while we will later prove that a first-order property of inner product spaces that holds in all finite dimensional spaces holds in any inner product space, there are very simple first-order properties of normed spaces that only have infinite-dimensional models. The section concludes with a proof that there are first-order properties that hold in all Banach spaces but not in all normed spaces.

Section 3 introduces our basic method for proving undecidability in a language equipped with a sort whose intended interpretation is the real numbers. The method is to exhibit a structure $\mathcal{M}$ for the language and a formula $\nu(x)$ with the indicated free variable of the real sort which holds in the structure iff x is interpreted as a natural number. If such a structure $\mathcal{M}$ exists, the method provides a reduction of the set of truths of second-order arithmetic to the set of sentences that hold in any class of structures containing $\mathcal{M}$. Thus the method shows that a theory for which such a structure $\mathcal{M}$ exists is not even arithmetical.

Section 4 applies the method of the previous section to the case of metric spaces, which gives a new proof that the first-order theory of a metric space is undecidable. Here we also give a decision procedure for (a superset of) the ∀∃ fragment of the theory of metric spaces and show that this is the best possible result of its type by reducing the satisfiability problem for Diophantine equations to ∀∃ satisfiability for metric spaces.

Section 5 gives the main undecidability results for normed spaces and Banach spaces: it turns out that in every dimension $d \geqslant 2$ we can apply the methods of Section 3 and prove undecidability of the corresponding theories of normed spaces and Banach spaces, even for the purely additive fragments where multiplication is disallowed. This section concludes with a more detailed investigation into the degrees of unsolvability of these theories: the theories for spaces of finite dimension $d \geqslant 2$ turn out to have the same many-one degree as the set of truths in second-order arithmetic, while if we allow infinite-dimensional spaces, the theories have the same many-one degree as the set of true Π^2_1 sentences in third-order arithmetic.

In Section 6 we turn to inner product spaces and find that they are quite tractable: the key result implies that a sentence holds in every space of dimension $d \geqslant k$ iff it holds in $\mathbb{R}^k$ where k is the number of distinct vector variables in the sentence. From this we find that the theories of inner product spaces and Hilbert spaces with various dimensional constraints can all be decided via a simple reduction to the first order theory of the real numbers.

Section 7 complements our investigation with some results on decidable fragments of the normed space theory analogous to the decidability results for metric spaces in Section 4. The purely existential fragment admits a very simple reduction to the first-order language of the real numbers. The purely universal fragment is also decidable via a more sophisticated method.

Again these results are the best possible of their type: in Section 8 we give reductions of satisfiability for Diophantine equations to both the ∀∃ and the ∃∀ satisfiability problems for normed spaces; in fact both these reductions are subsumed by our final result which gives the undecidability of the set of $\forall \Rightarrow \forall$ sentences valid for normed spaces.

Some of the results presented here have been foreshadowed by several authors and some have been strengthened since the present paper was first written. We conclude in Section 9 with a brief survey of related work.

The genesis of this paper was a question about decision procedures for vector spaces asked several years ago of Solovay by Harrison, and quickly answered with the first proofs of decidability and quantifier elimination for inner product spaces. Some time later, Harrison became interested in corresponding questions for the theory of normed spaces and implemented a decision procedure for the universal additive theory. Arthan conjectured, however, that the full theory of normed spaces is undecidable. On hearing this conjecture, Solovay rapidly proved it and precisely characterized the theory as many-one equivalent to the fragment of third-order arithmetic discussed below. Arthan refined these results to cover finite-dimensional spaces, purely additive theories and formulas with limited quantifier alternations, while Harrison extended the decidability to the full universal theory and has done further practical work on implementations. All hands have contributed to the numerous improvements leading to the present account.

2. The languages and their interpretation

We will study sublanguages of a 2-sorted first-order language $\mathcal{L}$. $\mathcal{L}$ itself provides a full repertoire of first-order features for work in a Hilbert space. It includes the operations of a vector space equipped with an inner product together with the induced norm and metric. After introducing $\mathcal{L}$, we define sublanguages for other kinds of structure: $\mathcal{L}_V$ for vector spaces, $\mathcal{L}_M$ for metric spaces, $\mathcal{L}_N$ for normed spaces and $\mathcal{L}_I$ for inner product spaces.

2.1. Sorts

The two sorts in $\mathcal{L}$ are as follows:

1. $\mathcal{R}$ — scalars
2. $\mathcal{V}$ — vectors or points

The variables and constants in our many-sorted languages all carry a label indicating their sort. In $\mathcal{L}$, we adopt the familiar convention of bold font ($\mathbf{x}$, $\mathbf{y}$, $\mathbf{0}$, etc.) for vectors or points and regular font (x, y, 0, etc.) for scalars. If we need to write sort labels explicitly, we will use superscripts, e.g., $x^{\mathcal{N}}$ will be a variable of the natural number sort in second-order arithmetic.

2.2. Language

We describe here the constant, function and predicate symbols of $\mathcal{L}$ and then define important sublanguages in later sections. Following mathematical custom, we overload many of the arithmetic operations like '+' for both scalars and vectors, but this should not cause confusion given that our notation distinguishes vector variables and constants from their scalar counterparts. $\mathcal{L}$ has the following constants and function symbols:

1. Scalar constants n for all rational numbers n.
2. Addition $(x + y)$, negation $(-x)$ and multiplication (xy) of scalars.
3. The zero vector or origin, $\mathbf{0}$.
4. Addition $(\mathbf{x} + \mathbf{y})$ and negation $(-\mathbf{x})$ of vectors.
5. Multiplication of a scalar and a vector, with type $\mathcal{R} \times \mathcal{V} \to \mathcal{V}$. We write the product of a scalar c and vector $\mathbf{x}$ as $c\mathbf{x}$.
6. The inner (dot) product of vectors, with type $\mathcal{V} \times \mathcal{V} \to \mathcal{R}$. We write the inner product of vectors $\mathbf{x}$ and $\mathbf{y}$ as $\langle \mathbf{x}, \mathbf{y} \rangle$.
7. The norm operation on vectors, with type $\mathcal{V} \to \mathcal{R}$. We write $\|\mathbf{x}\|$ for the norm of a vector $\mathbf{x}$.
8. The distance function for metric spaces, with type $\mathcal{V} \times \mathcal{V} \to \mathcal{R}$. We write $d(\mathbf{x}, \mathbf{y})$ for the distance between $\mathbf{x}$ and $\mathbf{y}$.

We will also use the usual shorthands such as $x - y$ (for $x + (-y)$), x^2 (for xx) and $\mathbf{v}/2$ (for $(1/2)\mathbf{v}$). Nothing of substance would change if we also added a multiplicative inverse operation. However, it can always be eliminated if necessary. In any case, if a multiplicative inverse is to be included, adapting the results of this paper is much more straightforward and efficient if 0^{-1} has a specific known value.

The predicate symbols are:

1. Equality $\mathbf{v} = \mathbf{w}$ of vectors.
2. All the usual equality and inequality comparisons for scalars: $x = y$, $x < y$, $x \leqslant y$, $x > y$, $x \geqslant y$.

We use $|x|$ as a shorthand for the absolute value of x and $\max\{x, y\}$ as a shorthand for the maximum of x and y: $\phi(\max\{x, y\})$ stands for $x \geqslant y \wedge \phi(x) \vee x < y \wedge \phi(y)$ and $\phi(|x|)$ stands for $\phi(\max\{x, -x\})$. Recursively, we write $\max\{x_1, x_2, \ldots, x_k\}$ for $\max\{x_1, \max\{x_2, \ldots, x_k\}\}$.

2.3. Interpretation

All the languages considered here include a symbol for equality for every sort and this is to be interpreted as actual equality in any structure. For sublanguages of $\mathcal{L}$, unless otherwise stated, we require the sort $\mathcal{R}$ and the symbols for the field operations and the ordering to be interpreted as the ordered field of real numbers. Thus all the first-order properties of $\mathbb{R}$ form part of the theory while we may make free use of higher-order properties such as completeness when we reason about it.

2.3.1. Vector spaces

The language $\mathcal{L}_V$ of vector spaces has the scalar constant, function and predicate symbols together with the constant $\mathbf{0}$ and addition, negation and scalar multiplication for vectors. A vector space is a structure for this language satisfying the vector space axioms listed below. These state that the vectors form an Abelian group on which the field of scalars acts a ring of homomorphisms:

- $\forall \mathbf{u}\,\mathbf{v}\,\mathbf{w} \cdot \mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$
- $\forall \mathbf{v}\,\mathbf{w} \cdot \mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$
- $\forall \mathbf{v} \cdot \mathbf{0} + \mathbf{v} = \mathbf{v}$
- $\forall \mathbf{v} \cdot -\mathbf{v} + \mathbf{v} = \mathbf{0}$
- $\forall a\,\mathbf{v}\,\mathbf{w} \cdot a(\mathbf{v} + \mathbf{w}) = a\mathbf{v} + a\mathbf{w}$
- $\forall a\,b\,\mathbf{v} \cdot (a + b)\mathbf{v} = a\mathbf{v} + b\mathbf{v}$
- $\forall \mathbf{v} \cdot 1\mathbf{v} = \mathbf{v}$
- $\forall a\,b\,\mathbf{v} \cdot (ab)\mathbf{v} = a(b\mathbf{v})$.

The simplest example of a vector space comprises the single element $\mathbf{0}$ and is called 0. One can define a vector space structure component-wise on the set $\mathbb{R}^n$ of n-tuples of real numbers, and 0 can be considered the degenerate case $n = 0$. The space $\mathbb{R}^n$ contains the *standard basis* $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$:

$$\begin{aligned}
\mathbf{e}_1 &= (1, 0, 0, \ldots, 0) \\
\mathbf{e}_2 &= (0, 1, 0, \ldots, 0) \\
&\ldots \\
\mathbf{e}_n &= (0, 0, 0, \ldots, 1).
\end{aligned}$$

A fundamental result is that every vector space V has a *basis*, i.e., a set of vectors B such that (i) any vector $\mathbf{x} \in V$ can be represented as a linear combination $\mathbf{x} = x_1\mathbf{b}_1 + \cdots + x_m\mathbf{b}_m$ for some $m \in \mathbb{N}$, $x_i \in \mathbb{R}$ and $\mathbf{b}_i \in B$ (B *spans* V) and (ii) this representation is unique (B is *linearly independent*). The standard basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is indeed a basis for $\mathbb{R}^n$. Any two bases of a vector space V have the same cardinality called the *dimension* of V and we will write $\mathsf{dim}(V) = n$ if V has a finite basis with n elements, otherwise we write $\mathsf{dim}(V) = \infty$.

A *subspace* of a vector space is a substructure that also interprets $\mathcal{R}$ as the field of all real numbers. A subspace is automatically a vector space, since the vector space axioms are purely universal. An analogous definition applies to all our notions of a "space", a subspace being given by any subset of the vectors or points that is closed under all relevant operations. Two subspaces U and W of a vector space V are said to be *complementary* if every $\mathbf{v} \in V$ can be written uniquely as $\mathbf{v} = \mathbf{u} + \mathbf{w}$ with $\mathbf{u} \in U$ and $\mathbf{w} \in W$, in which case the dimension of W depends only on U and V and is said to be the *codimension* of U in V. Any subspace of a vector space has at least one complementary subspace.

If A is any set, the set $A \to \mathbb{R}$ of all real-valued functions on A becomes a vector space if one defines $(\mathbf{f} + \mathbf{g})(a) = \mathbf{f}(a) + \mathbf{g}(a)$ and $(x\mathbf{f})(a) = x\mathbf{f}(a)$. Taking $A = \mathbb{N}$, the elements of $A \to \mathbb{R}$ are sequences of real numbers and we define $\mathbb{R}^*$ to be the subspace comprising sequences $(\mathbf{x}_0, \mathbf{x}_1, \ldots)$ whose support $\{n \mid \mathbf{x}_n \neq 0\}$ is finite. This space is infinite-dimensional since the unit vectors $(0, \ldots, 0, 1, 0, \ldots)$ are linearly independent. Identifying the n-tuple $(x_1, \ldots, x_n)$ with the sequence $(x_1, \ldots, x_n, 0, \ldots)$, $\mathbb{R}^*$ can be viewed as the union of the spaces $\mathbb{R}^n$ for $n \in \mathbb{N}$.

Many useful geometric notions can be defined just in terms of the vector space operations. If $\mathbf{v}$ and $\mathbf{w}$ are distinct vectors, the *affine line* passing through them comprises the set of points that can be written as linear combinations $a\mathbf{v} + b\mathbf{w}$ where $a + b = 1$. The points of this form with $a, b \geqslant 0$ comprise the *closed line segment* $[\mathbf{v}, \mathbf{w}]$, while those with $a, b > 0$ form the *open line segment* $(\mathbf{v}, \mathbf{w})$. We say the line segment $[\mathbf{v}, \mathbf{w}]$ is *parallel* to a subspace W iff, for some $\mathbf{u}$, $[\mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{w}]$ is contained in W.

A set of vectors A is said to be *convex* if it contains the line segment connecting any two of its points. Following the convention that quantifiers have lower precedence than propositional operators (so the scope of a quantifier extends as far to the right as possible), we express this formally as follows:

$$\forall \mathbf{v}\,\mathbf{w} \cdot \mathbf{v} \in A \wedge \mathbf{w} \in A \quad \Rightarrow \quad \forall a\,b \cdot 0 \leqslant a \wedge 0 \leqslant b \wedge a + b = 1 \quad \Rightarrow \quad a\mathbf{v} + b\mathbf{w} \in A.$$

If A is any set of vectors, its *convex hull*, $\mathsf{conv}(A)$, is the smallest convex set containing A (this is well-defined because the intersection of any family of convex sets is convex). $\mathsf{conv}(A)$ comprises all the *convex combinations* of elements of A, i.e., all finite sums $a_1\mathbf{v}_1 + \cdots + a_m\mathbf{v}_m$ where $a_i \geqslant 0$, $\mathbf{v}_i \in A$, $i = 1, \ldots, m$, $a_1 + \cdots + a_m = 1$ and $m \geqslant 1$. If A is a finite set with n elements and if each element of $\mathsf{conv}(A)$ has a unique representation as a convex combination of elements of A, then $\mathsf{conv}(A)$ is said to be an $(n-1)$-*simplex* and the points of A are its *vertices*. So, for example, a 1-simplex is a closed line segment while a 2-simplex is a triangle.

2.3.2. Metric spaces

The language $\mathcal{L}_M$ of metric spaces has all the scalar constant, function and predicate symbols together with the metric $d(\_,\_)$ as the only function symbol involving the point type $\mathcal{V}$. A metric space is a structure for this language satisfying the metric space axioms listed below: positive definiteness, symmetry and the triangle inequality.

- $\forall \mathbf{x}\,\mathbf{y} \cdot d(\mathbf{x},\mathbf{y}) \geqslant 0 \wedge (d(\mathbf{x},\mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y})$
- $\forall \mathbf{x}\,\mathbf{y} \cdot d(\mathbf{x},\mathbf{y}) = d(\mathbf{y},\mathbf{x})$
- $\forall \mathbf{x}\,\mathbf{y}\,\mathbf{z} \cdot d(\mathbf{x},\mathbf{z}) \leqslant d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{z})$

A metric space is said to be *complete* if every Cauchy sequence converges. Unsurprisingly, it turns out to be impossible to capture this notion by first-order axioms in our language (see Theorem 1), but if we allow quantification over infinite sequences of points, we can express it as follows, where $\mathbf{x}$ ranges over such sequences:

$$\begin{aligned}&\forall \mathbf{x} \cdot \big(\forall \epsilon \cdot \epsilon > 0 \Rightarrow \exists N \cdot \forall m\, n \cdot m \geqslant N \wedge n \geqslant N \Rightarrow d\big(\mathbf{x}(n), \mathbf{x}(m)\big) < \epsilon\big)\\ &\quad \Rightarrow \quad \exists \mathbf{l} \cdot \forall \epsilon \cdot \epsilon > 0 \quad \Rightarrow \quad \exists N \cdot \forall n \cdot n \geqslant N \quad \Rightarrow \quad d\big(\mathbf{x}(n), \mathbf{l}\big) < \epsilon.\end{aligned}$$

2.3.3. Normed spaces

The language $\mathcal{L}_N$ of normed spaces includes all the symbols of the language $\mathcal{L}_V$ of vector spaces together with a norm $\|\_\|$. A metric d may also be used as a notational convenience (see below). A normed space is a structure for this language that satisfies the axioms for a vector space together with the axioms for norms listed below: positive definiteness, scaling and the triangle inequality:

- $\forall \mathbf{v} \cdot \|\mathbf{v}\| \geqslant 0 \wedge (\|\mathbf{v}\| = 0 \Leftrightarrow \mathbf{v} = \mathbf{0})$
- $\forall a\, \mathbf{v} \cdot \|a\mathbf{v}\| = |a|\,\|\mathbf{v}\|$
- $\forall \mathbf{v}\,\mathbf{w} \cdot \|\mathbf{v}+\mathbf{w}\| \leqslant \|\mathbf{v}\| + \|\mathbf{w}\|$

As a function from the space to the real numbers the norm is continuous with respect to a topology defined by the *induced metric*: $d(\mathbf{v},\mathbf{w}) = \|\mathbf{v}-\mathbf{w}\|$. This will be a very useful fact in our later arguments. The continuity of the norm at the point $\mathbf{v}$ can be expressed in our first-order language as follows:

$$\forall \epsilon \cdot \epsilon > 0 \quad \Rightarrow \quad \exists \delta \cdot \delta > 0 \wedge \forall \mathbf{w} \cdot \|\mathbf{w}\| < \delta \quad \Rightarrow \quad \big|\big(\|\mathbf{v}+\mathbf{w}\| - \|\mathbf{v}\|\big)\big| < \epsilon.$$

A *Banach space* is a normed space that is also metrically complete, i.e., with respect to the induced metric, every Cauchy sequence converges. As with metric spaces, we shall prove later that it is impossible to capture this by first-order axioms in our language (see Theorem 1).

The usual euclidean norm on $\mathbb{R}^n$ is defined by $\|\mathbf{x}\| = \sqrt{\sum_{i=1}^n \mathbf{x}_i^2}$, but there are plenty of other possibilities satisfying the axioms, such as the 1-norm ("Manhattan distance") $\|\mathbf{x}\| = \sum_{i=1}^n |\mathbf{x}_i|$ and the ∞-norm $\|\mathbf{x}\| = \max\{|\mathbf{x}_i| \mid 1 \leqslant i \leqslant n\}$. Similar norms can be defined on $\mathbb{R}^*$ by summing or maximizing over the support rather than from 1 to n. Other examples from functional analysis include the norm $\|f\| = \sup\{|f(x)| \mid x \in [a,b]\}$ on the Banach space of continuous functions $f : [a,b] \to \mathbb{R}$.

All norms on a finite-dimensional vector space, V, can be shown to be equivalent in the sense that if $\|\_\|_1$ and $\|\_\|_2$ are norms on V, then there are positive real numbers s and t such that for any $\mathbf{v} \in V$, $\frac{1}{s}\|\mathbf{v}\|_1 \leqslant \|\mathbf{v}\|_2 \leqslant t\|\mathbf{v}\|_1$. Although this implies that many properties of interest, in particular topological ones, are independent of the norm, we shall see that there are very great differences in the general first-order properties satisfied by different norms on the same finite-dimensional vector space.

Each norm defines a corresponding *unit circle* $S = \{\mathbf{x} \mid \|\mathbf{x}\| = 1\}$ and a *unit disc* $D = \{\mathbf{x} \mid \|\mathbf{x}\| \leqslant 1\}$. In spaces of higher dimension we also sometimes refer to S and D as the *unit sphere* and *unit ball* respectively. For the usual euclidean norm $\|(x_1,x_2)\| = \sqrt{x_1^2 + x_2^2}$ on $\mathbb{R}^2$, S and D are indeed a circle and a disc respectively. However many other shapes are possible, e.g. for the ∞-norm on $\mathbb{R}^2$, D is a square. However, D is always a convex set: if $\mathbf{x}$ and $\mathbf{y}$ are in D then $\|a\mathbf{x}+b\mathbf{y}\| \leqslant \|a\mathbf{x}\| + \|b\mathbf{y}\| = |a|\|\mathbf{x}\| + |b|\|\mathbf{y}\| \leqslant |a| + |b|$, and if $a, b \geqslant 0$ with $a + b = 1$ we have $|a| + |b| = 1$. D is also always symmetric about the origin in the sense that $\mathbf{v} \in D$ iff $-\mathbf{v} \in D$. As D is convex and $S \subseteq D$, any convex combination of unit vectors (i.e., members of S) has norm at most 1.

Conversely, it is often convenient to define a norm by nominating a suitable set as its unit disc D and defining the norm by taking $\|\mathbf{x}\|$ to be the smallest non-negative real number λ such that for some $\mathbf{d} \in D$ one has $\lambda\mathbf{d} = \mathbf{x}$. Provided the set D is convex and meets every line through the origin in a closed line segment $[-\mathbf{v},\mathbf{v}]$ where $\mathbf{v} \neq \mathbf{0}$, this is well-defined and satisfies the norm properties. For example, if $\mathbf{x} = \|\mathbf{x}\|\mathbf{d}$ and $\mathbf{y} = \|\mathbf{y}\|\mathbf{e}$ for $\mathbf{d}, \mathbf{e} \in D$ then

$$\mathbf{x}+\mathbf{y} = \big(\|\mathbf{x}\| + \|\mathbf{y}\|\big)\left(\frac{\|\mathbf{x}\|}{\|\mathbf{x}\| + \|\mathbf{y}\|}\mathbf{d} + \frac{\|\mathbf{y}\|}{\|\mathbf{x}\| + \|\mathbf{y}\|}\mathbf{e}\right)$$

and since by convexity $(\frac{\|\mathbf{x}\|}{\|\mathbf{x}\| + \|\mathbf{y}\|}\mathbf{d} + \frac{\|\mathbf{y}\|}{\|\mathbf{x}\| + \|\mathbf{y}\|}\mathbf{e}) \in D$, we have $\|\mathbf{x}+\mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\|$, i.e., the triangle inequality holds.

Under the euclidean norm, the unit circle meets any affine line in at most two points. However, there are many interesting norms for which this is not the case: in the ∞-norm in $\mathbb{R}^2$, for example, the unit circle comprises the union of four line segments. In working with such norms, it is useful to note that if L is an affine line, then $L \cap D$, the set of points on L of norm at most 1, is a bounded convex subset of L whose endpoints are contained in S. So if $L \cap D$ is non-empty, either $L \cap D = L \cap S = \{\mathbf{a}\}$ for some $\mathbf{a}$ with $\|\mathbf{a}\| = 1$, or $L \cap D = [\mathbf{a}, \mathbf{b}]$ for some distinct $\mathbf{a}$, $\mathbf{b}$ with $\|\mathbf{a}\| = \|\mathbf{b}\| = 1$. In the latter case, either $L \cap D \subseteq S$, i.e., $\|\mathbf{x}\| = 1$ for every $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$, or $L \cap S = \{\mathbf{a}, \mathbf{b}\}$ and $\|\mathbf{x}\| < 1$ for every $\mathbf{u} \in (\mathbf{a}, \mathbf{b})$. In particular, the condition $\|\mathbf{v}\| = \|\mathbf{w}\| = \|(\mathbf{v} + \mathbf{w})/2\| = 1$ implies that $[\mathbf{v}, \mathbf{w}] \subseteq S$.

More generally, let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be the vertices of an $(n-1)$-simplex, Δ, and let $\mathbf{u} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$ be any proper convex combination of the $\mathbf{v}_i$, i.e., $a_i > 0$, $i = 1 \ldots n$, and $a_1 + \cdots + a_n = 1$. Then $\{\mathbf{u}, \mathbf{v}_1, \ldots, \mathbf{v}_n\} \subseteq S$ implies that $\Delta \subseteq S$, i.e., $\|\mathbf{u}\| = \|\mathbf{v}_1\| = \cdots = \|\mathbf{v}_n\| = 1$ implies that $\|\mathbf{x}\| = 1$ for every $\mathbf{x} \in \Delta$. To see this, first note that $\Delta \subseteq D$ because a convex combination of unit vectors has norm at most 1. For $1 \leqslant i \leqslant n$, let L_i be the affine line passing through $\mathbf{u}$ and $\mathbf{v}_i$ and let Δ_i be the $(n-2)$-simplex whose vertices are $\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_n$. L_i meets Δ_i at a point $\mathbf{u}_i$, say, that must be a proper convex combination of the vertices of Δ_i. Since $\mathbf{v}_i$ and $\mathbf{u}$ are unit vectors and $\mathbf{u}$ lies on the open line segment $(\mathbf{u}_i, \mathbf{v}_i)$, we have $\|\mathbf{u}_i\| \geqslant 1$. Then as $\mathbf{u}_i \in \Delta \subseteq D$, $\|\mathbf{u}_i\| = 1$ and so, by induction, $\Delta_i \subseteq S$. Let $\mathbf{x}$ be any point of Δ. As $\Delta \subseteq D$, to show that $\|\mathbf{x}\| = 1$, it suffices to show that $\|\mathbf{x}\| \geqslant 1$. If $\mathbf{x} = \mathbf{u}$, then $\|\mathbf{x}\| = 1$ by assumption. So assume $\mathbf{x} \neq \mathbf{u}$. For some i, the half-line starting at $\mathbf{x}$ and passing through $\mathbf{u}$ meets the $(n-2)$-simplex Δ_i at a point $\mathbf{w}$. As $\mathbf{w}$ and $\mathbf{u}$ are both unit vectors and $\mathbf{u}$ lies on the open line segment $(\mathbf{x}, \mathbf{w})$, we find $\|\mathbf{x}\| \geqslant 1$ completing the proof. As a special case we have that if the unit vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are the vertices of an $(n-1)$-simplex Δ, then $\Delta \subseteq S$ iff the barycentre $\frac{1}{n}(\mathbf{v}_1 + \cdots + \mathbf{v}_n)$ is a unit vector. Note that this gives a considerable economy from a logical point of view: we can assert $\Delta \subseteq S$ without using any quantifiers or scalar variables.

2.3.4. Inner product spaces

The language $\mathcal{L}_I$ of inner product spaces includes all the symbols of the language $\mathcal{L}_V$ of vector spaces together with an inner product $\langle \_, \_ \rangle$. A norm may also be used as a notational convenience (see below). An inner product space satisfies the axioms for a vector space together with the axioms asserting that inner product is a positive definite symmetric bilinear form, which means:

- $\forall \mathbf{v}\, \mathbf{w} \cdot \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$
- $\forall \mathbf{u}\, \mathbf{v}\, \mathbf{w} \cdot \langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$
- $\forall a\, \mathbf{v}\, \mathbf{w} \cdot \langle a\mathbf{v}, \mathbf{w} \rangle = a\langle \mathbf{v}, \mathbf{w} \rangle$
- $\forall \mathbf{v} \cdot \langle \mathbf{v}, \mathbf{v} \rangle \geqslant 0 \wedge (\langle \mathbf{v}, \mathbf{v} \rangle = 0 \Leftrightarrow \mathbf{v} = \mathbf{0})$.

For example, n-dimensional euclidean space is $\mathbb{R}^n$ equipped with the inner product $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^n \mathbf{x}_i\mathbf{y}_i$. Note that $\langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2$ for the euclidean norm, and in general given any inner product we define the *induced norm* by $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$.

A *Hilbert space* is an inner product space that is also complete for the induced norm. Any finite-dimensional inner product space is a Hilbert space. The vector space of sequences $\mathbf{x} : \mathbb{N} \to \mathbb{R}$ such that the sum $\sum_{i=0}^\infty \mathbf{x}_i^2$ is convergent is an infinite-dimensional Hilbert space under an inner product defined by $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=0}^\infty \mathbf{x}_i\mathbf{y}_i$. This Hilbert space, called l_2, is one of many Hilbert spaces that occur naturally in functional analysis. The vector space $\mathbb{R}^*$ of finitely-supported sequences viewed as a subspace of l_2 gives an example of an incomplete inner product space.

If $\mathbf{u}$ and $\mathbf{v}$ are elements of an inner product space V, we say $\mathbf{v}$ is orthogonal to $\mathbf{u}$ if $\langle \mathbf{u}, \mathbf{v} \rangle = 0$. If $\mathbf{u}$ is non-zero then the set W of all vectors orthogonal to $\mathbf{u}$ forms a subspace W of V called the *orthogonal complement* of $\mathbf{u}$. Every element $\mathbf{v}$ of V can be written uniquely in the form $\mathbf{v} = a\mathbf{u} + \mathbf{w}$ where $\mathbf{w}$ is a member of W.

2.4. Additive sublanguages

The so-called linear fragment of real arithmetic admits a very simple quantifier elimination procedure [19,14] and enjoys many other pleasant properties. Here "linear" means that the multivariate polynomials that are the terms of the language are restricted to have total degree at most one. To define an analogous notion for $\mathcal{L}$ (or any of its sublanguages or extensions thereof by the addition of extra vector constants), we say a term or formula is *additive* if the left operand of every subterm of the form xy, $x\mathbf{v}$ or $\langle \mathbf{v}, \mathbf{w} \rangle$ is a constant. In $\mathcal{L}$ itself, which has only rational scalar constants and the vector constant $\mathbf{0}$, an additive formula is equivalent to one in which multiplication and inner product do not occur. E.g., one can write $\mathbf{q} + \mathbf{q} = \mathbf{p} + \mathbf{r}$ rather than $\mathbf{q} = (\mathbf{p} + \mathbf{r})/2$ to indicate that $\mathbf{q}$ is the midpoint of the line between $\mathbf{p}$ and $\mathbf{r}$. We write $\mathcal{L}^+$ for the additive sublanguage of $\mathcal{L}$.

Unless otherwise stated, in a structure for one of the additive sublanguages, we will require the sort $\mathcal{R}$, the symbols for the additive group operations and the ordering to be interpreted as the ordered additive group of real numbers with the rational number constants interpreted accordingly.

2.5. Initial observations on decidability

The principal results of this paper are connected with decidability or undecidability for the various 2-sorted languages introduced above. We now make some initial observations about the possibilities for decision procedures, e.g., via quantifier elimination, and about the interrelations among the decision problems for the languages.

2.5.1. Reductions among decision problems

Recall that every vector space V has a basis, i.e., a subset B such that any vector $\mathbf{x} \in V$ can be written uniquely as a sum $\sum_{\mathbf{b}\in B} x_{\mathbf{b}}\mathbf{b}$ (where all but finitely many $x_{\mathbf{b}}$ are zero). Given any basis we can regard the scalar coefficient $x_{\mathbf{b}}$ as the "**b**th coordinate" of **x**, and for $\mathbf{y} = \sum_{\mathbf{b}\in B} y_{\mathbf{b}}\mathbf{b}$, we can define $\langle \mathbf{x}, \mathbf{y}\rangle = \sum_{\mathbf{b}\in B} x_{\mathbf{b}} y_{\mathbf{b}}$ and show that this satisfies the inner product properties. Thus every vector space can be made into an inner product space; in logical parlance, this implies that the theory of inner product spaces is a conservative extension of the theory of vector spaces:

> A formula using neither the inner product nor norm operation holds in all vector spaces [optionally with constraints on the dimension] iff it holds in all inner product spaces [with corresponding constraints].

As noted already, in any inner product space we can define $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x}\rangle}$ and this satisfies the norm properties, so any model of the inner product space axioms immediately gives a model of the normed space axioms. The converse is not true, i.e. not every normed space is an inner product space. (See the remarks at the end of Section 6 for a more quantitative statement on this topic.) However if a normed space *is* derived from an inner product as above, the inner product can be recovered from the norm, e.g. by $\langle \mathbf{x}, \mathbf{y}\rangle = \frac{1}{2}(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}\|^2 - \|\mathbf{y}\|^2)$. It is a classic result of Jordan and von Neumann [24] that a norm is induced by an inner product iff it satisfies the parallelogram identity:

$$\forall \mathbf{x}\,\mathbf{y} \cdot \|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2\big(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2\big).$$

See Section 2.5.3 below for more about characterizations of inner product spaces.

Let ι be a sentence in the language $\mathcal{L}_I$ of inner product spaces asserting that '$\langle \_, \_\rangle$' satisfies the inner product axioms. Given any formula ϕ in $\mathcal{L}_I$, let ϕ^* be the corresponding formula in the language $\mathcal{L}_N$ of normed spaces where each term $\langle \mathbf{a}, \mathbf{b}\rangle$ is replaced by $(\|\mathbf{a}+\mathbf{b}\|^2 - \|\mathbf{a}\|^2 - \|\mathbf{b}\|^2)/2$. If M is an inner product space in which ϕ holds, then $\iota \wedge \phi$ holds in M. In that case $\iota^* \wedge \phi^*$ holds in the normed space N derived from M by defining $\|\mathbf{x}\|_N = \sqrt{\langle \mathbf{x}, \mathbf{x}\rangle_M}$. Conversely, if $\iota^* \wedge \phi^*$ holds in a normed space N, then setting $\langle \mathbf{x}, \mathbf{y}\rangle_M = (\|\mathbf{x}+\mathbf{y}\|_N^2 - \|\mathbf{x}\|_N^2 - \|\mathbf{y}\|_N^2)/2$ makes N into an inner product space, M say, in which ϕ holds. Both these constructions preserve dimensions and completeness and so restating in terms of validity, we have:

> A sentence ϕ in the language of inner product spaces holds in all inner product spaces [with or without constraints on the dimension and with or without the requirement for completeness] iff the sentence $\iota^* \Rightarrow \phi^*$ (as defined above) holds in all normed spaces [with or without corresponding limitations].

This establishes that the decision problem for normed spaces is at least as general as the decision problem for inner product spaces, which in turn is at least as general as the decision problem for vector spaces. It will emerge in what follows that the decision problem for normed spaces is in fact dramatically harder than the other two. Intuitively, one might see this as expressing the fact that one has freedom to describe very "exotic" norms, whereas the freedom to define inner products is more constrained.

2.5.2. Possibility of quantifier elimination

It is not hard to see that we cannot have quantifier elimination in the basic language we are considering, for any of the vector space theories. For if so, any closed formula would be equivalent to a ground formula. Now the vector-valued subterms in a ground formula are formed from **0** using addition and scalar multiplication and so evaluate to **0** in any model. Thus the truth of a ground formula is independent of the space in which it is interpreted. So quantifier elimination would imply that all models are elementarily equivalent. This is certainly not the case however: we can write down formulas expressing non-trivial properties of the dimension and/or the norm. For example, the dimension is finite and $\leqslant n$ iff there is a spanning set of at most n vectors:

$$\exists \mathbf{v}_1 \ldots \mathbf{v}_n \cdot \forall \mathbf{w} \cdot \exists a_1 \ldots a_n \cdot a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{w}.$$

We will see in Section 6 that, if these sentences $\mathsf{D}_{\leqslant n}$ are treated as atomic predicates, there *is* a full quantifier elimination algorithm for vector spaces and for inner product spaces. This allows us to decide validity in all vector spaces or all those with a specific restriction on the dimension. Moreover, it implies the existence, for any formula ϕ in this theory, of a bound k such that ϕ holds in all vector (or inner product) spaces iff it holds in all those of dimension at most k. In other words, if a formula ϕ in the language of inner product spaces is satisfiable, it is satisfiable in an inner product space with a specific finite upper bound on the dimension.

If we turn to normed spaces, however, the situation changes dramatically. We will see in Section 5 that the theory is undecidable, so no algorithmically useful quantifier elimination in an expanded language exists. We will show below that there are satisfiable formulas that are satisfiable only in infinite-dimensional normed spaces. Moreover, quantifier elimination in the unexpanded language must even fail for purely additive formulas (no scalar multiplication or inner products, and scalar–vector multiplication only for integer constants), since we can for example express the fact that the dimension is $\leqslant 1$ by:

$$\exists \mathbf{x} \cdot \forall \mathbf{y} \cdot \|\mathbf{y}\| = 1 \quad\Rightarrow\quad \mathbf{y} = \mathbf{x} \vee \mathbf{y} = -\mathbf{x}$$

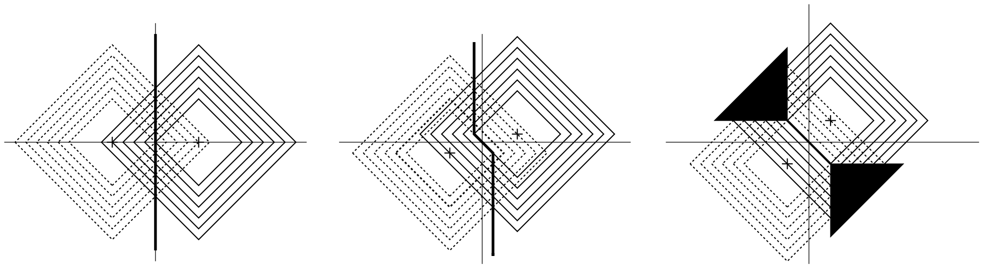

Fig. 1. Examples of $\{\mathbf{x} \mid \mathsf{O}(\mathbf{x}, \mathbf{y})\}$ in the 1-norm on $\mathbb{R}^2$.

and distinguish the 1-norm and 2-norm by:

$$\forall \mathbf{x}\, \mathbf{y} \cdot \|\mathbf{x}\| = \|\mathbf{y}\| \wedge \|\mathbf{x}+\mathbf{y}\| = \|\mathbf{x}\| + \|\mathbf{y}\| \quad \Rightarrow \quad \mathbf{x} = \mathbf{y}.$$

(This holds for the euclidean norm in any number of dimensions, but fails in $\mathbb{R}^2$ with the 1-norm $\|(x_1, x_2)\| = |x_1| + |x_2|$, as can be seen by setting $\mathbf{x} = (1, 0)$, $\mathbf{y} = (0, 1)$.)

2.5.3. Further expressiveness results for normed spaces

There are (purely additive) formulas in the language of normed spaces that are satisfiable yet have only infinite-dimensional models. To see this, define a 1-place predicate $\mathsf{E}(\mathbf{v})$ that holds iff $\mathbf{v}$ is a unit vector that is not the midpoint of the line connecting two distinct vectors in the unit disc, i.e., $\mathbf{v}$ is an extreme point of the unit disc:

$$\mathsf{E}(\mathbf{v}) := \|\mathbf{v}\| = 1 \wedge \forall \mathbf{u}\, \mathbf{w} \cdot \|\mathbf{u}\| \leqslant 1 \wedge \|\mathbf{w}\| \leqslant 1 \wedge \mathbf{v} = (\mathbf{u}+\mathbf{w})/2 \quad \Rightarrow \quad \mathbf{u} = \mathbf{w}.$$

Now consider the sentence Inf asserting that there exist non-zero vectors but that the unit disc has no extreme points:

$$\mathsf{Inf} := (\exists \mathbf{v} \cdot \mathbf{v} \neq \mathbf{0}) \wedge \big(\forall \mathbf{v} \cdot \neg \mathsf{E}(\mathbf{v})\big).$$

In a finite-dimensional normed space, the Krein–Milman theorem implies that the unit disc is the convex hull of its extreme points, so Inf cannot hold in finite dimensions. But when equipped with the ∞-norm, the space $\mathbb{R}^*$ considered above (sequences of real numbers with finite support) has a unit disc with no extreme points: given any unit vector $\mathbf{v}$, pick an n so that $\mathbf{v}_n = 0$ and set $\mathbf{u}_n = -1$, $\mathbf{w}_n = 1$ and $\mathbf{u}_i = \mathbf{w}_i = \mathbf{v}_i$ for $i \neq n$; then $\mathbf{v} = (\mathbf{u}+\mathbf{w})/2$. Hence Inf holds in $\mathbb{R}^*$, so Inf is satisfiable but only has infinite-dimensional models.

It is also interesting to observe that using the norm, we can find purely additive sentences that are satisfiable, but only in certain models with a specific finite dimension. In fact without using multiplication we can even characterize specific norms, e.g., the 1-norm and the euclidean norm on $\mathbb{R}^n$.

To give these characterizations, we will write $\mathsf{O}(\mathbf{x}, \mathbf{y})$ for $\|\mathbf{x}+\mathbf{y}\| = \|\mathbf{x}-\mathbf{y}\|$, i.e. $\mathsf{O}(\mathbf{x}, \mathbf{y})$ holds iff $\mathbf{x}$ is equidistant from $\mathbf{y}$ and $-\mathbf{y}$. Intuitively, this is intended as an approximation to the concept of orthogonality in an inner product space. Indeed, for a norm derived in the usual way from an inner product, this says exactly that $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. In a general normed space, $\mathsf{O}(\mathbf{x}, \mathbf{y})$ will not enjoy all the properties of orthogonality, and, in particular, the "orthogonal complement", $C_{\mathbf{y}} := \{\mathbf{x} \mid \mathsf{O}(\mathbf{x}, \mathbf{y})\}$, need not be a subspace, as illustrated for the 1-norm on $\mathbb{R}^2$ in Fig. 1.

Assume $\|\mathbf{x}+\mathbf{y}\| < \|\mathbf{x}-\mathbf{y}\|$ for some $\mathbf{x}$ and $\mathbf{y}$ in some normed space V, so that certainly $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$. Let $f(s) = \|s\mathbf{x} + (2-s)\mathbf{y}\| - \|s(\mathbf{x}-\mathbf{y})\|$ so that $f(s) = 0$ iff $\mathsf{O}(s\mathbf{x} + (1-s)\mathbf{y}, \mathbf{y})$ holds. $f(s)$ is a continuous function of s with $f(0) = 2\|\mathbf{y}\| > 0$ and $f(1) = \|\mathbf{x}+\mathbf{y}\| - \|\mathbf{x}-\mathbf{y}\| < 0$. By the intermediate value theorem, $f(t) = 0$ for some $t > 0$. So $\mathbf{x} = a\mathbf{y} + b\mathbf{z}$, where $a = \frac{t-1}{t}$, $b = \frac{1}{t}$ and $\mathbf{z} = t\mathbf{x} + (1-t)\mathbf{y}$. As $f(t) = 0$, $\mathsf{O}(\mathbf{z}, \mathbf{y})$ holds. Similarly, if $\|\mathbf{x}+\mathbf{y}\| > \|\mathbf{x}-\mathbf{y}\|$, we can also find $\mathbf{z}$ such that $\mathsf{O}(\mathbf{z}, \mathbf{y})$ holds and $\mathbf{x} = a\mathbf{y} + b\mathbf{z}$ for some a and b. Since $\|\mathbf{x}+\mathbf{y}\| = \|\mathbf{x}-\mathbf{y}\|$ implies $\mathsf{O}(\mathbf{x}, \mathbf{y})$, any $\mathbf{x} \in V$ can be written as a linear combination $\mathbf{x} = a\mathbf{y} + b\mathbf{z}$, where $\mathsf{O}(\mathbf{z}, \mathbf{y})$, i.e., $\mathbf{y}$ and $C_{\mathbf{y}}$ span V.

Now assume that for some $\mathbf{y} \neq \mathbf{0}$, the set $C_{\mathbf{y}}$ is a subspace. Then if $\mathbf{z} \in C_{\mathbf{y}}$, so also is $b\mathbf{z}$. Thus any $\mathbf{x} \in V$ can be written as $a\mathbf{y} + \mathbf{z}$ where $\mathbf{z} \in C_{\mathbf{y}}$, and, as $\mathbf{y} \notin C_{\mathbf{y}}$, this representation is unique. Thus $C_{\mathbf{y}}$ has codimension 1 and $\mathbf{y}$ spans a complementary subspace. If $\mathbf{z} \in C_{\mathbf{y}}$ and $a \neq 0$, then $\mathbf{z}/a \in C_{\mathbf{y}}$ and we have $\|a\mathbf{y} + \mathbf{z}\| = |a| \cdot \|\mathbf{y} + \mathbf{z}/a\| = |a| \cdot \|\mathbf{y} - \mathbf{z}/a\| = \|-a\mathbf{y} + \mathbf{z}\|$. Thus, if $C_{\mathbf{y}}$ is a subspace, there is a (unique) linear isometry from V to itself that fixes $C_{\mathbf{y}}$ and maps $\mathbf{y}$ to $-\mathbf{y}$. For example, for $\mathbf{y} \neq \mathbf{0}$ in $\mathbb{R}^2$ under the 1-norm, $C_{\mathbf{y}}$ is a subspace iff $\mathbf{y}$ lies on one of the coordinate axes, in which case $C_{\mathbf{y}}$ is the other axis and reflection in it gives the linear isometry mapping $\mathbf{y}$ to $-\mathbf{y}$ (see Fig. 1).

For any $n \in \mathbb{N}$, there is a sentence ϕ_n of $\mathcal{L}_N^+$ which holds in a normed space iff there are vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ such that:

- $\|\mathbf{e}_i\| = 1$ for each i
- $\mathsf{O}(\mathbf{e}_i, \mathbf{e}_j)$ for each $i \neq j$
- $\forall \mathbf{v}\, \mathbf{w} \cdot \mathsf{O}(\mathbf{v}, \mathbf{e}_i) \wedge \mathsf{O}(\mathbf{w}, \mathbf{e}_i) \Rightarrow \mathsf{O}(\mathbf{v}+\mathbf{w}, \mathbf{e}_i)$ for each i
- $\forall \mathbf{v} \cdot \mathsf{O}(\mathbf{v}, \mathbf{e}_i) \Rightarrow \mathsf{O}(\frac{1}{2}\mathbf{v}, \mathbf{e}_i)$ for each i
- $\forall \mathbf{v} \cdot \mathsf{O}(\mathbf{v}, \mathbf{e}_1) \wedge \cdots \wedge \mathsf{O}(\mathbf{v}, \mathbf{e}_n) \Rightarrow \mathbf{v} = \mathbf{0}$.

I claim that in any model V of ϕ_n, the set $W_i = C_{\mathbf{e}_i} = \{\mathbf{x} \mid \mathsf{O}(\mathbf{x}, \mathbf{e}_i)\}$ is a subspace, and hence, by the above remarks, a subspace of codimension 1. To see that W_i is indeed a subspace, note that, by induction, if $\mathsf{O}(\mathbf{v}, \mathbf{e}_i)$ holds then so does $\mathsf{O}(\frac{m}{2^k}\mathbf{v}, \mathbf{e}_i)$ for any integers m and k, and so by continuity $\mathsf{O}(a\mathbf{v}, \mathbf{e}_i)$ holds for all real a. Now, setting $V_0 = V$ and $V_{i+1} = W_{i+1} \cap V_i$, we see that each V_{i+1} is a subspace of V_i of codimension 1. By the final hypothesis in our list, we must have $V_{n+1} = 0$, and so V must have dimension n.

Moreover, if we add the additional property $\|\frac{1}{n}(\mathbf{e}_1 + \cdots + \mathbf{e}_n)\| = 1$, then the resulting sentence actually has a *unique* model up to isomorphism, namely $\mathbb{R}^n$ with the 1-norm w.r.t the usual basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$. For, by the remarks at the end of Section 2.3.3, these revised hypotheses imply that the $(n-1)$-simplex with vertex set $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is contained in the unit sphere. Also, there is a linear isometry mapping $\mathbf{e}_i$ to $-\mathbf{e}_i$ and fixing the other basis elements, which means that each of the 2^n $(n-1)$-simplices with vertex sets $\{\pm\mathbf{e}_1, \ldots, \pm\mathbf{e}_n\}$ is contained in the unit sphere. It follows that the unit sphere is the generalized octahedron whose facets are these simplices and this is the unit sphere of the 1-norm with respect to the basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$.

Many characterizations of inner product spaces amongst normed spaces have been discovered and rediscovered over the years, often based on abstractions of orthogonality (our "isosceles orthogonality" $\mathsf{O}(\mathbf{v}, \mathbf{w})$ was proposed by James [23]). Amir [1] gives a systematic presentation of some 350 characterizations involving a wide range of ideas from geometry and analysis. One characterization due to Aronszajn [2] says that a normed space is an inner product space if the norms of two sides and of one diagonal of any parallelogram determine the norm of the other diagonal. Aronszajn's theorem implies that if we add the following additional hypothesis to our original ϕ_n, we obtain a purely additive characterization of euclidean n-space:

$$\forall \mathbf{v}_1\, \mathbf{w}_1\, \mathbf{v}_2\, \mathbf{w}_2 \cdot \|\mathbf{v}_1\| = \|\mathbf{v}_2\| \wedge \|\mathbf{w}_1\| = \|\mathbf{w}_2\| \wedge \|\mathbf{v}_1 - \mathbf{w}_1\| = \|\mathbf{v}_2 - \mathbf{w}_2\| \Rightarrow \|\mathbf{v}_1 + \mathbf{w}_1\| = \|\mathbf{v}_2 + \mathbf{w}_2\|.$$

See Amir [1] or Arthan [3] for a proof of Aronszajn's characterization and see Mok [31] for another interesting purely additive characterization.

2.5.4. Completeness in metric spaces and normed spaces

In Section 6 we will show that the theories of inner product spaces and of Hilbert spaces coincide. In this section we investigate the analogous question for metric spaces compared with complete metric spaces and for normed spaces compared with Banach spaces and find, by contrast, that for these theories the assumption of completeness does make a difference to the first-order theory.

Consider the following properties of a relation R between the real numbers and the points of a metric space X.

- R is a partial function whose domain comprises positive numbers:

$$\forall x\, \mathbf{p}\, \mathbf{q} \cdot R(x, \mathbf{p}) \wedge R(x, \mathbf{q}) \quad \Rightarrow \quad x > 0 \wedge \mathbf{p} = \mathbf{q}.$$

- The domain of R has no positive lower bound:

$$\forall \epsilon > 0 \cdot \exists x\, \mathbf{p} \cdot x < \epsilon \wedge R(x, \mathbf{p}).$$

- R satisfies a form of the Cauchy criterion as its argument tends to 0:

$$\forall \epsilon > 0 \cdot \exists \delta > 0 \cdot \forall x\, y\, \mathbf{p}\, \mathbf{q} \cdot x < \delta \wedge y < \delta \wedge R(x, \mathbf{p}) \wedge R(y, \mathbf{q}) \quad \Rightarrow \quad d(\mathbf{p}, \mathbf{q}) < \epsilon.$$

- R has no limit as its argument tends to 0:

$$\forall \mathbf{q} \cdot \exists \epsilon > 0 \cdot \forall \delta > 0 \cdot \exists x\, \mathbf{p} \cdot x < \delta \wedge R(x, \mathbf{p}) \wedge d(\mathbf{p}, \mathbf{q}) \geqslant \epsilon.$$

Write Q_R for the conjunction of the above properties and say R *represents* a sequence, $\mathbf{s}_n$, of points of X iff there is a strictly decreasing subsequence x_n contained in the domain of R such that x_n tends to 0 as n tends to ∞ and $R(x_n, \mathbf{s}_n)$ for all n. Thus Q_R implies that R represents at least one Cauchy sequence but that no Cauchy sequence represented by R has a limit, so that Q_R cannot hold in a complete metric space. Moreover if Q_R holds and R is definable in some space S by a formula $\mathsf{R}(x, \mathbf{p})$ of the language of metric spaces, then the sentence asserting $\neg\mathsf{Q}_\mathsf{R}$ belongs to the theory of complete metric spaces but not to the theory of metric spaces in general, since it does not hold in S. A similar argument applies to normed spaces and Banach spaces, the construction below being slightly complicated by the need for a parameter in the formula $R(x, \mathbf{v})$.

For the metric space case, consider the subset $\mathbb{M}$ of the real plane comprising points $\mathbf{p}_n = (\frac{1}{2^n}, 0)$ and circles C_n of radius $\frac{1}{2^{n+2}}$ with centre $\mathbf{p}_n$ for $n = 1, 2, \ldots$ (see Fig. 2). Taking $\mathbb{M}$ as a metric space under the euclidean metric, the sequence $\mathbf{p}_n$ is Cauchy but has no limit in $\mathbb{M}$, since its limit in the plane is the origin, which is not in $\mathbb{M}$. Define a predicate $\mathsf{P}(x, \mathbf{p})$ as follows:

$$\mathsf{P}(x, \mathbf{p}) := \big(\exists \mathbf{q} \cdot \mathbf{q} \neq \mathbf{p} \wedge d(\mathbf{p}, \mathbf{q}) = x\big) \wedge \big(\forall \mathbf{q} \cdot \mathbf{q} \neq \mathbf{p} \Rightarrow d(\mathbf{p}, \mathbf{q}) \geqslant x\big).$$

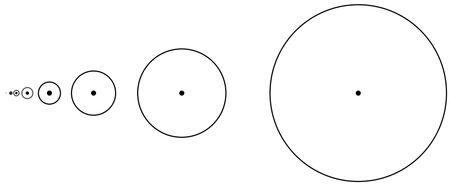

Fig. 2. The incomplete metric space $\mathbb{M}$.

I.e., $\mathsf{P}(x, \mathbf{p})$ holds iff $\mathbf{p}$ is an isolated point such that for some $\mathbf{q}$, $x = d(\mathbf{p}, \mathbf{q})$ is minimal for $\mathbf{q} \neq \mathbf{p}$, i.e., in $\mathbb{M}$, iff $x = \frac{1}{2^{n+2}}$ and $\mathbf{p} = \mathbf{p}_n$ for some n, $\mathbf{q}$ being any point of C_n. Thus $\mathsf{P}(x, \mathbf{v})$ represents the divergent sequence $\mathbf{p}_n$ and Q_P holds in $\mathbb{M}$ so that a first-order sentence asserting $\neg Q_\mathsf{P}$ holds in all complete metric spaces but not in $\mathbb{M}$.

For the normed space case, we start with the vector space $\mathbb{R}^*$ of finitely non-zero sequences of real numbers, which we think of as the union of the finite dimensional spaces $\mathbb{R}^n$. We will construct a normed space $\mathbb{Y}$ by making modifications to the euclidean unit ball to make a certain divergent sequence representable. So until further notice we work with the euclidean metric on $\mathbb{R}^*$ which we write as $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$. Also if X is any non-empty subset of $\mathbb{R}^*$, we write $d(\mathbf{v}, X)$ for the distance between $\mathbf{v}$ and X, i.e., the infimum of the numbers $d(\mathbf{v}, \mathbf{x})$ as $\mathbf{x}$ ranges over X.

If $\mathbf{v}$ and $\mathbf{w}$ are distinct, non-antipodal unit vectors (i.e., $\|\mathbf{v}\| = \|\mathbf{w}\| = 1$ and $\mathbf{v} \neq \pm\mathbf{w}$), the *great circle* through $\mathbf{v}$ and $\mathbf{w}$ is defined to be the intersection of the unit sphere S in $\mathbb{R}^*$ and the plane through the origin spanned by $\mathbf{v}$ and $\mathbf{w}$. Writing $\mathbf{e}_1, \mathbf{e}_2, \ldots$ for the standard basis vectors, define a sequence of unit vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots$ as follows:

$$\mathbf{v}_1 = \mathbf{e}_1$$

$$\mathbf{v}_{n+1} = \text{the unique point on the great circle through } \mathbf{v}_n \text{ and } \mathbf{e}_{n+1} \text{ such that}$$

$$d(\mathbf{v}_{n+1}, \mathbf{v}_n) = \tfrac{1}{4^n} \text{ and } d(\mathbf{v}_{n+1}, \mathbf{e}_{n+1}) < d(\mathbf{v}_n, \mathbf{e}_{n+1}).$$

So each $\mathbf{v}_n$ lies in $\mathbb{R}^n \setminus \mathbb{R}^{n-1}$ and of the two points on the great circle at distance $\frac{1}{4^n}$ from $\mathbf{v}_n$, $\mathbf{v}_{n+1}$ is the one on the same side of $\mathbb{R}^n$ as $\mathbf{e}_{n+1}$ in $\mathbb{R}^{n+1}$.

It is a straightforward exercise in using the triangle inequality to prove the following bounds on the distance between two members of the sequence $\mathbf{v}_n$ (e.g., prove the upper bound first by induction on k and then derive the lower bound using the upper bound for $d(\mathbf{v}_{n+1}, \mathbf{v}_{n+k})$).

$$\frac{2}{3} \cdot \frac{1}{4^n} < d(\mathbf{v}_n, \mathbf{v}_{n+k}) < \frac{4}{3} \cdot \frac{1}{4^n} = \frac{1}{3} \cdot \frac{1}{4^{n-1}}.$$

These upper bounds show that the $\mathbf{v}_n$ form a Cauchy sequence. Also if α is the angle between $\mathbf{v}_n$ and $\mathbf{v}_{n+1}$, one has that $d(\mathbf{v}_{n+1}, \mathbb{R}^n) = \sin(\alpha) \geqslant \sin(\frac{\alpha}{2}) = \frac{1}{2} d(\mathbf{v}_n, \mathbf{v}_{n+1}) = \frac{1}{2}\frac{1}{4^n}$. Whence using the triangle inequality and the above bounds, we have $d(\mathbf{v}_{n+k}, \mathbb{R}^n) \geqslant (\frac{1}{2} - \frac{1}{3})\frac{1}{4^n} = \frac{1}{6}\frac{1}{4^n}$. It follows that a limit of the $\mathbf{v}_n$ could not belong to any $\mathbb{R}^n$, and so the sequence $\mathbf{v}_n$ has no limit in $\mathbb{R}^*$.

We have $d(\mathbf{e}_1, \mathbf{v}_n) = d(\mathbf{v}_1, \mathbf{v}_n) < \frac{1}{3}$ implying the following bound for any $m, n \geqslant 1$.

$$d(\mathbf{v}_m, -\mathbf{v}_n) > d(\mathbf{e}_1, -\mathbf{e}_1) - \frac{2}{3} = \frac{4}{3}.$$

Let O_n be the open disc with centre $\mathbf{v}_n$ and radius $\frac{1}{2}\frac{1}{4^n}$. Our estimates imply that the sets $O_1, -O_1, O_2, -O_2, \ldots$ have pairwise disjoint closures. Let E be the convex hull of the set $A \cup \{\mathbf{v}_1, -\mathbf{v}_1, \mathbf{v}_2, -\mathbf{v}_2, \ldots\}$ where A is the set obtained from the (euclidean) unit disc D in $\mathbb{R}^*$ by removing any points that are within $\frac{1}{2}\frac{1}{4^n}$ of $\pm\mathbf{v}_n$, i.e., $A = D \setminus \bigcup\{O_1, -O_1, O_2, -O_2, \ldots\}$.

E satisfies the conditions for a unit disc in a normed space. Let T be the unit sphere in this normed space, i.e., the boundary of E. Writing S for the unit sphere in $\mathbb{R}^*$, T comprises $S \setminus \bigcup\{O_1, -O_1, O_2, -O_2, \ldots\}$ together with a set of truncated cones made up of line segments $[\pm\mathbf{v}_n, \mathbf{w}]$ joining each $\pm\mathbf{v}_n$ to each (euclidean) unit vector $\mathbf{w}$ such that $d(\pm\mathbf{v}_n, \mathbf{w}) = \frac{1}{2}\frac{1}{4^n}$. Since the closures of the sets $\pm O_n$ are pairwise disjoint, the $\mathbf{v}_n$ are the only isolated extreme points of T and the points on the open line segments $(\pm\mathbf{v}_n, \mathbf{w})$ are the only points of T that are not extreme points. (All these claims are most easily seen by considering the possible ways in which T can intersect a plane through the origin.)

Clearly, $\frac{1}{2}D \subseteq E \subseteq D$. Thus writing $\|\_\|_X$ for the norm with unit disc X (so $\|\_\|_D$ is the euclidean norm), we have that our two norms are equivalent in the sense that each is bounded by a constant multiple of the other:

$$2\|\mathbf{v}\|_D \geqslant \|\mathbf{v}\|_E \geqslant \|\mathbf{v}\|_D \geqslant \frac{1}{2}\|\mathbf{v}\|_E.$$

As a consequence, under $\|\_\|_E$, just as under the euclidean norm, the $\mathbf{v}_n$ form a Cauchy sequence that has no limit in the normed space $\mathbb{Y}$ whose underlying vector space is $\mathbb{R}^*$ and whose unit disc is E. Now let $\mathsf{R}(x, \mathbf{v}, \mathbf{e})$ be a formula in the language of normed spaces expressing the following properties:

(i) $\|\mathbf{e}-\mathbf{v}\| < \frac{2}{3}$;
(ii) $\mathbf{v}$ is an isolated point in the set of extreme points of the unit disc;
(iii) there exists an extreme point $\mathbf{w} \neq \mathbf{v}$ of the unit disc such that the line segment $[\mathbf{v}, \mathbf{w}]$ lies on the unit disc and $x = \|\mathbf{w}-\mathbf{v}\|$.

In $\mathbb{Y}$, take $\mathbf{e} = \mathbf{e}_1 = \mathbf{v}_1$, and let $\mathbf{v} \in \mathbb{Y}$ and $x \in \mathbb{R}$ be given. By the above estimates and remarks, conditions (i) and (ii) are satisfied iff $\mathbf{v}$ is one of the $\mathbf{v}_n$. If $\mathbf{v} = \mathbf{v}_n$, then condition (iii) is satisfied iff $x = \|\mathbf{v}-\mathbf{w}\|_E$ where $\mathbf{w}$ is a (euclidean) unit vector with $\|\mathbf{v}-\mathbf{w}\|_D = \frac{1}{2}\frac{1}{4^n}$, and, for such a $\mathbf{w}$, we have:

$$\frac{1}{4^n} = 2\|\mathbf{v}-\mathbf{w}\|_D \geqslant \|\mathbf{v}-\mathbf{w}\|_E = x \geqslant \|\mathbf{v}-\mathbf{w}\|_D = \frac{1}{2} \cdot \frac{1}{4^n}.$$

We conclude that when the parameter $\mathbf{e}$ is interpreted by $\mathbf{e}_1$, the relation defined in $\mathbb{Y}$ by $\mathsf{R}(x, \mathbf{v}, \mathbf{e})$ represents the divergent sequence $\mathbf{v}_n$. Thus $\exists \mathbf{e} \cdot Q_{\mathsf{R}}$ holds in $\mathbb{Y}$ and a sentence asserting $\forall \mathbf{e} \cdot \neg Q_{\mathsf{R}}$ holds in all Banach spaces but does not hold in the normed space $\mathbb{Y}$.

In Section 6, we shall prove that for every set of sentences A in the language of inner product spaces there is a subset D of $\mathbb{N} \cup \{\infty\}$, such that an inner product space V is a model of A iff $\mathsf{dim}(V) \in D$ (see Corollary 35). So if A is any set of sentences in the language of metric spaces, then the class of metric space models of A cannot coincide with the class of complete metric spaces, since if that were the case then the inner product space models of A would comprise precisely the class of Hilbert spaces, but this is impossible since, if D is the set of dimensions associated with A, either $\infty \notin D$, so that A does not admit any infinite-dimensional Hilbert space as a model, or $\infty \in D$, so that A admits every infinite-dimensional inner product space as a model and hence any incomplete inner product space is a model of A (incomplete spaces being necessarily infinite-dimensional). Essentially the same argument shows that no set of sentences in the language of normed spaces can have the class of Banach spaces as its class of models.

Collecting together the results of this section gives us the following theorem.

Theorem 1. *There are first-order sentences that hold in all complete metric spaces* (*resp. Banach spaces*) *but not in all metric spaces* (*resp. normed spaces*)*. However, the class of complete metric spaces* (*resp. Banach spaces*) *is not an axiomatizable subclass of the class of metric spaces* (*resp. normed spaces*). □

3. On undecidability in languages with a sort for the real numbers

We will demonstrate the undecidability of various theories over languages containing a sort for the real numbers by showing how to interpret second-order arithmetic in them. In this section we describe a general procedure for doing this.

3.1. Interpreting first-order arithmetic

Consider a first-order language L that includes symbols for the field operations and the ordering relation on a sort $\mathcal{R}$ whose intended interpretation is the ordered field $\mathbb{R}$, e.g., our language $\mathcal{L}_N$ for normed spaces. Let $\mathcal{C}$ be some class of structures for L in which the sort $\mathcal{R}$ has its intended interpretation, e.g., the class of all Banach spaces is such a class for $\mathcal{L}_N$.

Let a formula $\nu(x)$ of L with one free variable of sort $\mathcal{R}$ be given. The following sentence Peano holds in a structure $\mathcal{M}$ in the class $\mathcal{C}$ iff in $\mathcal{M}$, $\nu(x)$ defines the set $\mathbb{N} \subseteq \mathbb{R}$.

$$\begin{aligned}\mathsf{Peano} := {} & \nu(0) \wedge \\ & \big(\forall x \cdot \nu(x) \Rightarrow x \geqslant 0 \wedge \nu(x+1)\big) \wedge \\ & \big(\forall x\, y \cdot \nu(x) \wedge \nu(y) \wedge x \neq y \Rightarrow |x-y| \geqslant 1\big).\end{aligned}$$

Now take any sentence ϕ in the language of first-order arithmetic and reinterpret it as a sentence $\phi_{\mathbb{N}}$ of L by labelling all variables and constants in ϕ with sort $\mathcal{R}$ and relativizing all quantifiers using the formula $\nu(x)$, i.e., replacing every subformula of the form $\exists x \cdot \psi$ by $\exists x \cdot \nu(x) \wedge \psi$ and every subformula of the form $\forall x \cdot \psi$ by $\forall x \cdot \nu(x) \Rightarrow \psi$.

I claim that if Peano is satisfiable in $\mathcal{C}$, then $\mathsf{Peano} \Rightarrow \phi_{\mathbb{N}}$ holds in $\mathcal{C}$ iff ϕ holds in $\mathbb{N}$. For, in any structure with the intended interpretation of $\mathcal{R}$ in which Peano holds, $\phi_{\mathbb{N}}$ holds iff ϕ holds in $\mathbb{N}$. So if Peano holds in some structure $\mathcal{M} \in \mathcal{C}$, then $\mathsf{Peano} \Rightarrow \phi_{\mathbb{N}}$ holds in $\mathcal{M}$ iff ϕ is true, iff $\mathsf{Peano} \Rightarrow \phi_{\mathbb{N}}$ holds in $\mathcal{C}$. Thus, if we can find a single model of the sentence Peano in the class $\mathcal{C}$, then the theory of $\mathcal{C}$ must be undecidable, since a decision procedure for it would lead to a decision procedure for the set of truths of first-order arithmetic, contradicting Tarski's theorem on the undefinability of truth.

This method of relativization has often been used to show that extending decidable theories such as Presburger arithmetic or the theory of a real closed field with a new uninterpreted unary function or predicate leads to undecidability [43,13]. Even though our $\nu(x)$ is not just an uninterpreted unary predicate but rather a complex formula in a language with a constrained interpretation, we have to exhibit just *one* model of the characterizing sentence Peano in order to get a reduction of first-order arithmetic to the theory of the class $\mathcal{C}$.

3.2. Interpreting second-order arithmetic

We will obtain still stronger undecidability results by observing that in a first-order theory of the real numbers with a predicate for the natural numbers, one can interpret not only first-order arithmetic as we did above but even *second-order* arithmetic. This is "well-known" but since we know of no reference for it in the literature we will give the proof. The setting is as in the previous section with L a language including a sort $\mathcal{R}$ for the real numbers with the usual operations and $\mathcal{C}$ a class of structures for L in which these things have their intended interpretation.

First we will briefly describe second-order arithmetic; see, e.g., Simpson [39] for more details. The language $\mathcal{L}_A^2$ of second-order arithmetic is a 2-sorted language with a sort $\mathcal{N}$ called "type 0" whose intended interpretation is the set of natural numbers $\mathbb{N}$ and a sort $\mathcal{P}$ called "type 1" whose intended interpretation is the set $\mathbb{P}(\mathbb{N})$ of all sets of natural numbers. The expressions are those of first-order arithmetic which have sort $\mathcal{N}$ together with variables of sort $\mathcal{P}$. Atomic formulas can be built from numeric terms by the usual predicates of first-order arithmetic, and also if t is a numeric term and A a set variable we can form the atomic formula $t \in A$. Quantification is allowed over both numeric and set variables.

We have already seen how to interpret first-order arithmetic by relativizing quantifiers using the natural number predicate $\nu(x)$. In order to interpret type 1 variables in the first-order theory of the real numbers, we use the mapping taking a set A with characteristic function χ_A:

$$\chi_A(n) = \begin{cases} 1 & \text{if } n \in A \\ 0 & \text{otherwise} \end{cases}$$

into the real number whose ternary expansion is determined by the values $\chi_A(n)$:

$$\sharp A = \sum_{n=0}^{\infty} \chi_A(n)/3^n.$$

Note that a binary version of the same method would not give an injective map because of $1.000\cdots = 0.111\cdots$ etc., and so would require workarounds like treating terminating expansions differently or encoding the function in even digits of the binary expansion. Using ternary, we can straightforwardly and unambiguously recover the set A from the number $\sharp A$. Let $h_n(x)$ be the value of the first n ternary digits of x, considered as an integer, where the zeroth 'digit' is simply $\lfloor x \rfloor$:

$$h_n(x) = \lfloor 3^n x \rfloor.$$

Then defining

$$d_n(x) = \begin{cases} h_0(x) & \text{if } n = 0 \\ h_n(x) - 3h_{n-1}(x) & \text{otherwise} \end{cases}$$

we have $d_n(\sharp A) = \chi_A(n)$. We will show below that the function $d_n(x)$ is definable, or more precisely that we can find a formula $\mathsf{D}(n, x)$ of our language with two free variables whose interpretation corresponds to $d_n(x) = 1$ in all standard models (i.e. those interpreting the real sort in the usual way).

Assume that $\mathcal{M}$ is a structure for the language L and that $\nu(x)$ is a formula in L with the indicated free variable of sort $\mathcal{R}$ which defines the natural numbers in $\mathcal{M}$, i.e., $\nu(x)$ holds in $\mathcal{M}$ iff x is interpreted as a natural number. Then the relation $\mathsf{D}(n, x)$ can be defined in terms of $\nu(x)$ using the following relational translations of the definitions given above, first for $h_n(x)$:

$$h_n(x) = l \quad \Leftrightarrow \quad \nu(n) \wedge \nu(l) \wedge \exists k \cdot \nu(k) \wedge 3^n = k \wedge l \leqslant k \cdot x \wedge k \cdot x < l + 1$$

then $d_n(x)$:

$$\begin{aligned} d_n(x) = y \quad \Leftrightarrow \quad & \nu(n) \wedge \nu(y) \wedge \\ & \big((n = 0 \wedge h_0(x) = y\big) \vee \\ & \big(\exists m\, l\, k \cdot \ \nu(m) \wedge \nu(l) \wedge \nu(k) \wedge \\ & \qquad n = m + 1 \wedge h_n(x) = l \wedge h_m(x) = k \wedge \\ & \qquad l = y + 3 \cdot k)\big) \end{aligned}$$

and finally:

$$\mathsf{D}(n, x) \quad \Leftrightarrow \quad d_n(x) = 1.$$

It remains to define the exponential relation $3^n = k$ in L, but this can be done by taking any of the usual definitions in the language of first-order arithmetic, e.g. the one given by [40], and translating into L using the numeric sort $\mathcal{R}$ and its operations and relativizing with respect to the predicate $\nu(x)$.

If we define:

$$\mathsf{S}(x) := x \geqslant 0 \wedge \forall y \cdot y \geqslant 0 \wedge \big(\forall n \cdot \mathsf{D}(n,x) \Leftrightarrow \mathsf{D}(n,y)\big) \quad \Rightarrow \quad x \leqslant y$$

then $\mathsf{S}(x)$ holds iff $x = \sharp\{n \in \mathbb{N} \mid \mathsf{D}(n,x)\}$. Thus, we can interpret second-order arithmetic in L using $\mathsf{D}(n,x)$ to represent sets of natural numbers as real numbers and using $\mathsf{S}(x)$ to pick a canonical representative: given a formula ϕ of second-order arithmetic, we take ϕ^* to be the result of the following sequence of transformations:

1. Replace subformulas of the form $\exists x^{\mathcal{N}} \cdot \psi$ by $\exists x^{\mathcal{R}} \cdot \nu(x^{\mathcal{R}}) \wedge \psi$ and subformulas of the form $\forall x^{\mathcal{N}} \cdot \psi$ by $\forall x^{\mathcal{R}} \cdot \nu(x^{\mathcal{R}}) \Rightarrow \psi$;
2. Replace subformulas of the form $\exists A^{\mathcal{P}} \cdot \psi$ by $\exists A^{\mathcal{R}} \cdot \mathsf{S}(A^{\mathcal{R}}) \wedge \psi$ and subformulas of the form $\forall A^{\mathcal{P}} \cdot \psi$ by $\forall A^{\mathcal{R}} \cdot \mathsf{S}(A^{\mathcal{R}}) \Rightarrow \psi$;
3. Replace remaining occurrences of the sort labels $\mathcal{N}$ and $\mathcal{P}$ by $\mathcal{R}$;
4. Replace subformulas of the form $t \in A$ by $\mathsf{D}(t, A)$.

Here recall that each variable and constant in our many-sorted language comprises a name labelled with a sort, which we write as a superscript, and note that there are no constants of sort $\mathcal{P}$. Now given a *sentence* ϕ of second-order arithmetic, we may assume (up to a logical equivalence) that bound variables have been renamed if necessary so that no variable name appears in ϕ with two different sorts and distinct variables remain distinct even after a relabelling that identifies two sorts. Assuming that $\nu(x)$ does indeed define the natural numbers, we then find by induction on the structure of a formula in which no variable name appears with two different sorts that the sentence ϕ is true iff ϕ^* holds in the structure $\mathcal{M}$. The details of the induction are straightforward: in the inductive step for the type 1 quantifiers, one notes that by the discussion above, $\sharp$ defines a 1–1 correspondence between $\mathbb{P}(\mathbb{N})$ and the set of real numbers s such that $\mathsf{S}(s)$ holds.

Theorem 2. *Let L be a (many-sorted) first-order language including a sort $\mathcal{R}$, constants $0 : \mathcal{R}$ and $1 : \mathcal{R}$ and function symbols $\_ + \_$, $\_ \times \_ : \mathcal{R} \times \mathcal{R} \to \mathcal{R}$ whose intended interpretations form the field of the real numbers. Let $\mathcal{C}$ be some class of structures for L in which $\mathcal{R}$ and these symbols have their intended interpretations and let $\mathcal{T}$ be the theory of $\mathcal{C}$, i.e., the set of all sentences that hold in every member of $\mathcal{C}$. If there is a formula $\nu(x)$ of L with one free variable x of sort $\mathcal{R}$ such that in some structure $\mathcal{M}$ in the class $\mathcal{C}$, $\nu(x)$ defines the set of natural numbers, then there is a primitive recursive reduction of second-order arithmetic to $\mathcal{T}$.*

Proof. The reduction maps a sentence ϕ of second-order arithmetic to the sentence $\phi' := \mathsf{Peano} \Rightarrow \phi^*$ where Peano is defined as above using the $\nu(x)$ that we are given by hypothesis and ϕ^* is the above translation of ϕ into the language L. By the discussion above, ϕ' then holds in every member of $\mathcal{C}$ iff ϕ is true. □

3.3. Interpretation in an additive theory

Since the *linear* theory of integer arithmetic is decidable [33] we need multiplication in our language in order to interpret the full, undecidable theory, even though the characterizing formula Peano itself does not involve multiplication. But we will later want to show the undecidability of additive theories of metric and vector spaces where multiplication is not available. In some interesting cases we can construct a structure in which we can define not only the natural numbers but also the graph of the multiplication function $(x, y) \mapsto xy$. In order to interpret first-order arithmetic we only need to be able to define and characterize the multiplication of natural numbers. But to achieve the full reduction of second-order arithmetic, we require multiplication of arbitrary real numbers, since this is used in the formulas defining $h_n(x) = l$ and $d_n(x) = l$ above.

To make this programme work, we need an analogue Mult of the sentence Peano, asserting that a formula $\mu(x, y, z)$ with three free variables defines the multiplication relation $x \cdot y = z$ on the real numbers. Let us define Mult as follows:

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

The first conjunct asserts that $\mu(x, y, z)$ does indeed define a function $f(x, y) = z$, and the second that $f(x, y) = f(y, x)$. The next three conjuncts ensure that this function coincides with multiplication in the case where x is a natural number because they give $f(0, y) = 0$, $f(1, y) = y$ and $f(x + 1, y) = f(x, y) + y$. They also imply that this holds for $x \in \mathbb{Z}$, because $f(-x, y) + f(x, y) = 0$ and therefore $f(-x, y) = -f(x, y)$. Using the additivity property repeatedly we also see that for any real number x and natural number $q > 0$ we have $f(x, y) = f(x/q + \cdots + x/q, y) = f(x/q, y) + \cdots + f(x/q, y) = q \cdot f(x/q, y)$ and therefore $f(x/q, y) = f(x, y)/q$. Together these imply that $f(x, y) = x \cdot y$ when $x \in \mathbb{Q}$. Now the final conjunct implies that for any $y \in \mathbb{R}$ the function $g(x) = f(x, y) - x \cdot y$ is continuous. Since $\{x \mid g(x) \neq 0\}$ is the preimage of the open set $\mathbb{R} - \{0\}$ under a continuous function, it is open. Since it contains no rational numbers, it must be empty, so $g(x)$ is identically zero as required.

Hence our characterizing formula Mult works as claimed and we can summarize the import of all this in the following result:

Theorem 3. *Let L be a (many-sorted) first-order language including a sort $\mathcal{R}$, together with function symbol $\_+\_ : \mathcal{R} \times \mathcal{R} \to \mathcal{R}$, a binary predicate symbol $\_ < \_$ on the sort $\mathcal{R}$ and a constant $1 : \mathcal{R}$ whose intended interpretations form the ordered group of real numbers under addition with 1 as a distinguished positive element. Let $\mathcal{C}$ be some class of structures for L in which $\mathcal{R}$ and these symbols have their intended interpretations and let $\mathcal{T}$ be the theory of $\mathcal{C}$, i.e., the set of all sentences that hold in every member of $\mathcal{C}$. Let $\nu(x)$ (resp. $\mu(x, y, z)$) be a formula of L with one free variable x of sort $\mathcal{R}$ (resp. free variables x, y and z all of sort $\mathcal{R}$). If in some structure in the class $\mathcal{C}$, $\nu(x)$ defines the set of natural numbers with the intended interpretation of the constant 1 and $\mu(x, y, z)$ defines the multiplication relation on the set of real numbers, then there is a primitive recursive reduction of second-order arithmetic to $\mathcal{T}$.*

Proof. Given a sentence ϕ in second-order arithmetic, let ϕ^* be the translation of ϕ into the language L used in the proof of Theorem 2. There is a primitive recursive function that maps any formula in L to a logically equivalent one in which all instances of multiplication are unnested, i.e., multiplication only appears in atomic predicates of the form $xy = z$ where x, y and z are variables; see, e.g., Hodges [20]. Let ϕ^+ be the result of applying this function to ϕ^* and then replacing each atomic predicate of the form $xy = z$ by $\mu(x, y, z)$. If we then set $\phi'' := \mathsf{Peano} \wedge \mathsf{Mult} \Rightarrow \phi^+$, ϕ'' holds in every member of $\mathcal{C}$ iff ϕ is true. $\square$

In fact, both Theorems 2 and 3 can easily be strengthened to allow the formulas $\nu(x)$ and $\mu(x, y, z)$ to have additional free variables acting as parameters: if for some structure and some choice of values for the parameters, $\nu(x)$ defines the natural numbers, then the conclusion of Theorem 2 will obtain, while if also $\mu(x, y, z)$ defines the graph of multiplication, then the conclusion of Theorem 3 will also obtain. The formulations without parameters are all we need in the sequel.

4. Metric spaces

We begin the main work of this paper by considering metric spaces. The generality of the metric space axioms gives us considerable freedom to construct spaces in which various arithmetic sets and relations are definable as needed to apply the methods of Section 3. Many of the same ideas will appear later for normed spaces but in a more intricate form.

The elementary theory of metric spaces is known to be undecidable. This was first proved by Bondi [9]. Kutz et al. [28] give a very simple proof by encoding an arbitrary reflexive symmetric binary relation R (i.e. an undirected graph) as a metric via:

$$d(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y \wedge R(x, y) \\ 2 & \text{if } \neg R(x, y) \end{cases}$$

This allows the decision problem for the theory of a reflexive symmetric binary relation, known to be hereditarily undecidable [35], to be reduced to the theory of metric spaces. In this proof, few special properties of $\mathbb{R}$ are needed and almost any other set of valuations would work; the set of points takes centre stage and the set of scalars plays only a supporting role.

The theory of a reflexive symmetric binary relation is undecidable, but is recursively (indeed finitely) axiomatizable. The arguments of Bondi [9] and Kutz et al. [28] do not preclude the possibility that the theory of metric spaces might be recursively axiomatizable. By exploiting the methods of Section 3, we obtain a much stronger result:

Theorem 4. *There is a primitive recursive reduction of second-order arithmetic to the theory of metric spaces* MS.

Proof. Let $\mathbb{Z}$ be the set of integers with the usual metric $d(x, y) = |x - y|$. Clearly in this metric space the formula:

$$\mathsf{N}(x) := \exists a\, b \cdot d(a, b) = x$$

defines the natural numbers as a subset of the real numbers. Applying Theorem 2 completes the proof. $\square$

The theory of metric spaces is therefore not arithmetical, i.e., it is not definable by any formula of first-order arithmetic, and hence it is not recursively enumerable and it is not recursively axiomatizable.

If K is an ordered field, define a *metric space over K* to be a structure for the language $\mathcal{L}_M$ of metric spaces in which the scalar sort $\mathcal{R}$ and its operations are interpreted in K and which satisfies the metric space axioms. Let $\mathcal{C}$ be the class of all structures for $\mathcal{L}_M$ that are metric spaces over K, where K ranges over all real closed fields. Then $\mathcal{C}$ is clearly a recursively axiomatizable class of structures and so the set of sentences of $\mathcal{L}_M$ that are valid in $\mathcal{C}$ is recursively enumerable. Given

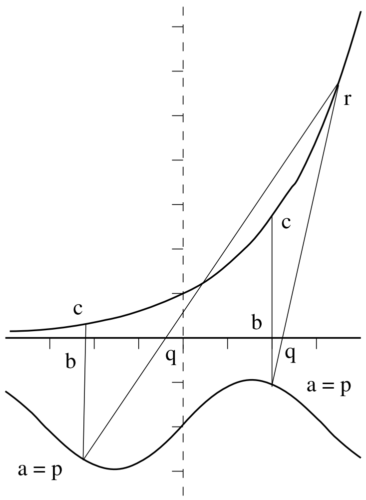

Fig. 3. Defining exp and sin in the metric space $\mathbb{G}$.

Theorem 4, we must conclude that there is a real closed field K and a sentence of $\mathcal{L}_M$ that holds in any metric space over $\mathbb{R}$ but fails in some metric space over K.

The situation is much the same even if we disallow multiplication:

Theorem 5. *There is a primitive recursive reduction of second-order arithmetic to the additive theory of metric spaces* MS_+.

Proof. We will exhibit a metric space $\mathbb{G}$ such that the set of natural numbers and the graph of the real multiplication function are *additively definable* in $\mathbb{G}$, i.e., definable using formulas that do not involve multiplication. $\mathbb{G}$ is the subspace of the euclidean plane comprising the x-axis together with the graphs of two functions e and s where e is the exponential function, $e(x) = \mathsf{exp}(x)$, and s is defined by $s(x) = \mathsf{sin}(x) - 2$. Thus $\mathbb{G}$ has three connected components: the graph of e lying strictly above the x-axis, the x-axis itself and the graph of s lying strictly below the x-axis, as illustrated in Fig. 3 (which actually shows $\mathsf{exp}(x/2)$ rather than $\mathsf{exp}(x)$ for reasons of space).

Our first task is to show that the connected components of $\mathbb{G}$ are additively definable. In the euclidean plane, a point $\mathbf{q}$ lies on the line segment $[\mathbf{p}, \mathbf{r}]$ iff $d(\mathbf{p}, \mathbf{r}) = d(\mathbf{p}, \mathbf{q}) + d(\mathbf{q}, \mathbf{r})$. A point $\mathbf{p}$ of $\mathbb{G}$ lies on the x-axis iff $\mathbb{G}$ contains the entire line segment $[\mathbf{p}, \mathbf{q}]$ for some $\mathbf{q} \neq \mathbf{p}$. So the x-axis is additively definable in $\mathbb{G}$. Now if f is a real-valued function of a real variable and x is any real number, then $(x, 0)$ is the point on the x-axis nearest to the point $(x, f(x))$ on the graph of f. Therefore, if $\mathbf{p}$ is a point of $\mathbb{G}$ and $\mathbf{q}$ is the point on the x-axis nearest to $\mathbf{p}$, then $d(\mathbf{p}, \mathbf{q}) > 3$ iff $\mathbf{p} = (x, e(x))$ for some x with $e(x) > 3$, so the set of such $\mathbf{p}$ is additively definable. But then the graph of s comprises precisely those points $\mathbf{p}$ of $\mathbb{G}$ for which there are a point $\mathbf{r} = (x, e(x))$ with $e(x) > 3$ and a point $\mathbf{q} \neq \mathbf{p}$ on the x-axis such that $\mathbf{q}$ lies on the line segment $[\mathbf{p}, \mathbf{r}]$ (see Fig. 3). Thus the graph of s is additively definable and hence so is the graph of e (which comprises the points of $\mathbb{G}$ that are neither on the x-axis nor on the graph of s).

The point $\mathbf{0} = (0, 0)$ is now additively definable in $\mathbb{G}$ as the point on the x-axis for which there is a point $\mathbf{p}$ on the graph of e with $d(\mathbf{p}, \mathbf{0}) = 1$ and $d(\mathbf{p}, \mathbf{q}) > 1$ for every other point $\mathbf{q}$ on the x-axis. The functions exp and sin are then additively definable: given a real number t, there are collinear points $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ with $\mathbf{a}$ on the graph of s, $\mathbf{c}$ on the graph of e and $\mathbf{b}$ the point on the x-axis closest to $\mathbf{c}$ with $d(\mathbf{0}, \mathbf{b}) = |t|$ and with $d(\mathbf{b}, \mathbf{c}) \geqslant 1$ if $t \geqslant 0$ and $d(\mathbf{b}, \mathbf{c}) < 1$ if $t < 0$ (see Fig. 3). With this unique choice of $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$, $\mathsf{exp}(t) = d(\mathbf{b}, \mathbf{c})$ and $\mathsf{sin}(t) = 2 - d(\mathbf{b}, \mathbf{a})$.

For positive x, we may now define $\mathsf{log}(x)$ by $\mathsf{exp}(\mathsf{log}(x)) = x$, then define multiplication for positive real numbers using $xy = \mathsf{exp}(\mathsf{log}(x) + \mathsf{log}(y))$ and extend the definition to all real numbers using $0y = x0 = 0$, $(-x)y = x(-y) = -xy$ and $(-x)(-y) = xy$. The real number π is additively definable as the smallest $x > 0$ such that $\mathsf{sin}(x) = 0$ and then the natural numbers are additively definable as the set of $n \geqslant 0$ such that $\mathsf{sin}(n\pi) = 0$. Thus multiplication and the natural numbers are additively definable in the metric space $\mathbb{G}$ and we may conclude by Theorem 3 that there is a primitive recursive reduction of second-order arithmetic to the additive theory of any class of metric spaces including $\mathbb{G}$. □

4.1. Decidability of the ∀∃ fragment

A sentence is said to be ∀∃ if it is in prenex normal form with no universal quantifier in the scope of an existential one, i.e. it has the following form for some $n \geqslant 0$, and $m \geqslant 0$ with ϕ quantifier-free:

$$\forall x_1 \ldots x_n \cdot \exists y_1 \ldots y_m \cdot \phi$$

the set of ∃∀ sentences being defined analogously exchanging '∀' with '∃'.

It is a classic theorem of Bernays and Schönfinkel that the set of valid first-order $\forall\exists$ sentences with no function symbols is decidable [7]: in fact, such a sentence with n initial universal quantifiers is valid iff it holds in all interpretations with at most $\max\{n, 1\}$ elements; but then it is a finite problem to enumerate all such interpretations. By working in many-sorted logic, this can be generalized to some important cases where function symbols occur [15]. We will prove the decidability of the set of valid $\forall\exists$ sentences in the language of metric spaces using similar ideas exploiting the fact that if $K \subseteq M$ and d is a metric on M then the restriction of d to $K \times K$ is also a metric on K. In fact our decision procedure will decide validity for a superset of the $\forall\exists$ sentences. We say a sentence is:

- $\forall\exists_p$ if it is prenex and no universal quantifier over points is in the scope of an existential quantifier (of any sort);
- $\exists\forall_p$ if it is prenex and no existential quantifier over points is in the scope of a universal quantifier (of any sort).

We have the following analogue of the theorem of Bernays and Schönfinkel:

Theorem 6. *Let ϕ be an $\exists\forall_p$ sentence in the language of metric spaces, and let n be the number of existential quantifiers of the point sort in ϕ. Then ϕ is satisfiable in a metric space iff it is satisfiable in a finite metric space with no more than* $\max\{n, 1\}$ *points.*

Proof. The right-to-left direction of the theorem is immediate. For the left-to-right direction, assume that the $\exists\forall_p$ sentence ϕ is satisfiable in some metric space M. As existential quantifiers commute up to logical equivalence, we can assume without loss of generality that ϕ consists of a block of $n \geqslant 0$ *existential* quantifiers over points followed by a block comprising *universal* quantifiers over points and scalar quantifiers of either kind. We write this as follows:

$$\phi \equiv \exists \mathbf{x}_1 \ldots \mathbf{x}_n \cdot \forall \bar{\mathbf{y}}/\mathsf{Q}\bar{z} \cdot \psi$$

where ψ is a quantifier-free formula whose free variables are contained in

$$\{\mathbf{x}_1, \ldots, \mathbf{x}_n, \mathbf{y}_1, \ldots, \mathbf{y}_k, z_1, \ldots, z_l\}.$$

If $n = 0$, we may replace ϕ by the logically equivalent formula $\exists \mathbf{x} \cdot \phi$ (hence replacing n by $1 = \max\{n, 1\}$), and so we may assume that $n \geqslant 1$. We have that $\rho := \forall \bar{\mathbf{y}}/\mathsf{Q}\bar{z} \cdot \psi$ holds for some points $\mathbf{x}_1, \ldots, \mathbf{x}_n \in M$. But then *a fortiori*, ρ and hence ϕ hold in the subspace $K = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ of M. But K has at most n points and we are done. $\square$

Corollary 7. *An $\forall\exists_p$ sentence in the language of metric spaces with n universally quantified point variables, which we can write as*

$$\forall \mathbf{x}_1 \ldots \mathbf{x}_n \cdot \exists \bar{\mathbf{y}}/\mathsf{Q}\bar{z}.\phi$$

holds in all metric spaces iff it holds in all finite metric spaces with at most $\max\{n, 1\}$ *points.*

Proof. Apply the theorem to the negation of the sentence. $\square$

These ideas lead to a decision procedure for valid $\forall\exists_p$ sentences:

Theorem 8. *The set of valid $\forall\exists_p$ sentences in the language of metric spaces is decidable.*

Proof. Since ϕ is valid iff $\neg\phi$ is not satisfiable, it suffices to describe a decision procedure for satisfiable $\exists\forall_p$ sentences. If ϕ is an $\exists\forall_p$ sentence, then as in the proof of Theorem 6, we may assume ϕ has the form $\exists \mathbf{x}_1 \ldots \mathbf{x}_n \cdot \forall \bar{\mathbf{y}}/\mathsf{Q}\bar{z} \cdot \psi$ where ψ is quantifier-free and $n \geqslant 1$, and then ϕ is satisfiable iff it is satisfiable in a metric space comprising just the interpretations of $\mathbf{x}_1, \ldots, \mathbf{x}_n$ under a satisfying assignment for $\forall \bar{\mathbf{y}}/\mathsf{Q}\bar{z} \cdot \psi$. So if we replace each subformula of ϕ of the form $\forall \mathbf{y} \cdot \rho$, by the conjunction $\rho[\mathbf{x}_1/\mathbf{y}] \wedge \cdots \wedge \rho[\mathbf{x}_n/\mathbf{y}]$ we obtain a sentence that is equisatisfiable with ϕ and has no point universal quantifiers. So we may assume ϕ has the form $\exists \mathbf{x}_1 \ldots \mathbf{x}_n \cdot \psi$ where ψ contains only scalar quantifiers. Now if M is a finite metric space with n points $\mathbf{p}_1, \ldots, \mathbf{p}_n$, say, define a function $f_M : M \mapsto \mathbb{R}^n$ by $f_M(\mathbf{p}) = (d(\mathbf{p}, \mathbf{p}_1), \ldots, d(\mathbf{p}, \mathbf{p}_n))$. If we equip $\mathbb{R}^n$ with the metric d_∞ induced from the ∞-norm, $d_\infty(\mathbf{v}, \mathbf{w}) = \max\{|\mathbf{v}_i - \mathbf{w}_i| \mid 1 \leqslant i \leqslant n\}$, then it is easy to check that f_M is an isometric embedding of M in $(\mathbb{R}^n, d_\infty)$. It follows that $\exists \mathbf{x}_1 \ldots \mathbf{x}_n \cdot \psi$ is satisfiable in general iff it is satisfiable in $(\mathbb{R}^n, d_\infty)$. Thus if we choose fresh variables x_{ij}, $1 \leqslant i, j \leqslant n$, and let ψ' be the result of replacing each subterm $\mathbf{x}_s = \mathbf{x}_t$ in ψ by $x_{s1} = x_{t1} \wedge \cdots \wedge x_{sn} = x_{tn}$ and each subterm $d(\mathbf{x}_s, \mathbf{x}_t)$ by $\max\{|x_{s1} - x_{t1}|, \ldots, |x_{sn} - x_{tn}|\}$, then $\exists \mathbf{x}_1 \ldots \mathbf{x}_n \cdot \psi$ is satisfiable iff $\phi' := \exists x_{11}\, x_{12} \ldots x_{nn} \cdot \psi'$ is satisfiable. But ϕ' contains no point variables so we may apply a decision procedure for real closed fields to complete the proof. $\square$

4.2. Undecidability of the $\exists\forall$ fragment

The following result shows that Theorem 8 is the best possible decidability result of its type:

Theorem 9. *If $\mathcal{C}$ is any class of metric spaces that includes the metric space $\mathbb{Z}$, then the set of $\exists\forall$ sentences that are valid in $\mathcal{C}$ is undecidable.*

Proof. We will prove the equivalent claim that the set of $\forall\exists$ sentences that are satisfiable in $\mathcal{C}$ is undecidable. Note that the formula $\mathsf{N}(x)$ used in the proof of Theorem 4 is purely existential, and so the corresponding sentence Peano of Section 3 is logically equivalent to an $\forall\exists$ sentence. Let $\phi(x_1, \ldots, x_k)$ be a quantifier-free formula in the language of arithmetic and consider the following sentence in the language of metric spaces:

$$\phi_1 := \mathsf{Peano} \wedge \exists x_1 \ldots x_k \cdot \mathsf{N}(x_1) \wedge \cdots \wedge \mathsf{N}(x_k) \wedge \phi(x_1, \ldots, x_k).$$

ϕ_1 is logically equivalent to an $\forall\exists$ sentence and ϕ_1 is satisfiable in $\mathbb{Z}$ and hence in $\mathcal{C}$ iff $\phi(x_1, \ldots, x_k)$ is satisfiable over the natural numbers. Thus a decision procedure for $\forall\exists$ sentences that are satisfiable in $\mathcal{C}$ would lead to a decision procedure for satisfiability of quantifier-free formulas in arithmetic and, in particular, for systems of Diophantine equations, contradicting the famous resolution of Hilbert's 10th problem by Matiyasevich [30]. □

5. Undecidability of theories of normed spaces

The theory NS^1 of 1-dimensional normed spaces reduces easily to the theory of the real numbers, since every such space is isomorphic to $\mathbb{R}$ with absolute value as the norm. Thus NS^1 is decidable. We will show in this section that this is the strongest possible positive decidability result of its type: even the additive theory NS^2_+ of 2-dimensional normed spaces is undecidable. In fact, NS^2_+ is not even arithmetical.

The main argument giving undecidability is in Section 5.1. We exhibit a 2-dimensional normed space, $\mathbb{X}$, and describe geometric constructions in that space of the set of natural numbers and of the graph of the multiplication function. Formalizing these constructions in the additive language of normed spaces and applying the methods of Section 3 immediately gives a reduction of second-order arithmetic to the (additive) theory of any class of normed spaces including $\mathbb{X}$. Taking a product with a Hilbert space of appropriate dimension, the construction lifts into any desired dimension $\geqslant 2$.

In Section 5.2, we obtain tighter estimates of the degrees of unsolvability of the normed space theories. We prove a kind of Skolem–Löwenheim theorem for normed spaces and use it to give reductions of the normed space theories to fragments of third-order arithmetic. We find that for any integer $d \geqslant 2$, the theory NS^d is many-one equivalent to second-order arithmetic, as is the theory $\mathsf{NS}^{\mathbb{F}}$ of all finite-dimensional normed spaces. We then strengthen the results of Section 5.1: using a variant of the approach of Section 3, we show that the theory NS^∞ of infinite-dimensional normed spaces and the theory NS of all normed spaces are both many-one equivalent to the set of true Π^2_1 sentences in third-order arithmetic. All of this goes through for the purely additive theories with little extra work. The results also hold equally well for Banach spaces: even though, by Theorem 1, the theory NS of all normed spaces is a proper subset of the theory BS of all Banach spaces, the two theories turn out to be many-one equivalent.

5.1. Reducing second-order arithmetic to the theory of a normed space

To apply the results of Section 3, we will exhibit a particular 2-dimensional normed space, $\mathbb{X}$, and give additive predicates that, in $\mathbb{X}$, define the natural numbers as a subset of the scalars and the graph of the scalar multiplication function. We define the norm by describing its unit disc. Let C be the unit circle in $\mathbb{R}^2$ with respect to the standard euclidean norm. For each $i \in \mathbb{Z}$, let l_i be the line passing through $\mathbf{0}$ and the point $(i, 1)$. Then l_i meets C in two points $\mathbf{v}_i$, say, in the upper half-plane and $-\mathbf{v}_i$ in the lower (see Fig. 4). The set E comprising the $\pm\mathbf{v}_i$ together with the two points $\mathbf{e}_1 = (1, 0)$ and $-\mathbf{e}_1$ is a closed and bounded subset of $\mathbb{R}^2$ and is symmetric about the origin. If we write D for the convex hull of E, D satisfies the requirements for a unit disc. Let us define $\mathbb{X}$ to be $\mathbb{R}^2$ with the norm $\|\_\|$ that has D as its unit disc. Note that as D is symmetric with respect to the x-axis and y-axis, $\|\_\|$ is invariant under reflection in these axes.

If we let S be the boundary of D, i.e., S is the set of unit vectors under $\|\_\|$, then clearly S consists of an infinite family of line segments, $\pm[\mathbf{v}_i, \mathbf{v}_{i+1}]$, together with the points $\pm\mathbf{e}_1$. The extreme points of D comprise the set E, i.e., the $\pm\mathbf{v}_i$ and $\pm\mathbf{e}_1$. Any neighbourhood of $\mathbf{e}_1$ or $-\mathbf{e}_1$ contains infinitely many extreme points of D; moreover, no other point of S, or indeed of $\mathbb{X}$, has this property.

We now define formulas in the additive language $\mathcal{L}^+$ that express various topological and geometric properties that will let us define a set of vectors in $\mathbb{X}$ whose norms comprise the natural numbers.

$$\mathsf{EP}(\mathbf{v}) := \forall \mathbf{u}\, \mathbf{w} \cdot \|\mathbf{u}\| = \|\mathbf{v}\| = \|\mathbf{w}\| \wedge \mathbf{v} = \frac{1}{2}(\mathbf{u} + \mathbf{w}) \quad \Rightarrow \quad \mathbf{u} = \mathbf{v} = \mathbf{w}$$

$$\mathsf{O}(\mathbf{v}, \mathbf{w}) := \|\mathbf{v} - \mathbf{w}\| = \|\mathbf{v} + \mathbf{w}\|$$

$$\mathsf{ACC}(\mathbf{v}) := \mathsf{EP}(\mathbf{v}) \wedge \left(\forall \epsilon \cdot \epsilon > 0 \Rightarrow \exists \mathbf{u} \cdot \|\mathbf{u}\| = \|\mathbf{v}\| \wedge \mathsf{EP}(\mathbf{u}) \wedge \mathbf{u} \neq \mathbf{v} \wedge \|\mathbf{u} - \mathbf{v}\| < \epsilon\right)$$

$$\mathsf{B}(\mathbf{p}, \mathbf{q}) := \|\mathbf{p}\| = \|\mathbf{q}\| = 1 \wedge \mathsf{ACC}(\mathbf{p}) \wedge \mathsf{EP}(\mathbf{q}) \wedge \mathsf{O}(\mathbf{q}, \mathbf{p}).$$

So $\mathsf{EP}(\mathbf{v})$ holds iff $\mathbf{v}$ is an extreme point of the disc $D_{\|\mathbf{v}\|}$ centred on the origin and of radius $\|\mathbf{v}\|$ (this is true in $\mathbb{X}$ iff $\mathbf{v}$ lies on the x-axis or on one of the lines l_i); $\mathsf{O}(\mathbf{v}, \mathbf{w})$ holds iff $\mathbf{v}$ is equidistant from the points $\pm\mathbf{w}$; $\mathsf{ACC}(\mathbf{v})$ holds iff $\mathbf{v}$ is a

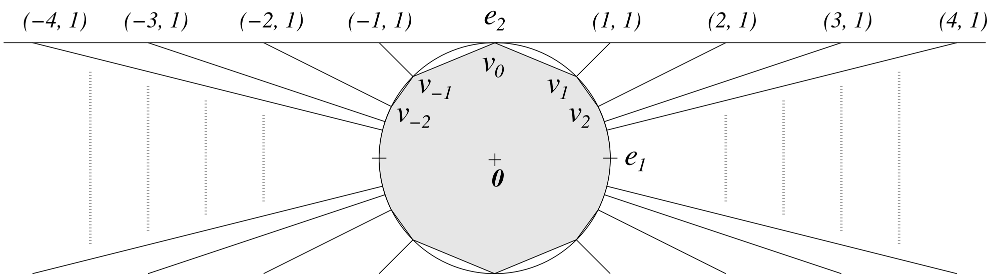

Fig. 4. The unit disc D in the space $\mathbb{X}$.

point of accumulation in the set of extreme points of the disc $D_{\|\mathbf{v}\|}$ (by the remarks above this is true in $\mathbb{X}$ iff $\mathbf{v}$ lies on the x-axis); and $\mathsf{B}(\mathbf{p},\mathbf{q})$ holds iff $\mathbf{p}$ is an accumulation point in the set of extreme points of the unit disc D and $\mathbf{q}$ is an extreme point of the unit disc equidistant from the points $\pm\mathbf{p}$.

If $\mathbf{p}=\pm\mathbf{e}_1$ and $\mathbf{q}=\pm\mathbf{e}_2$, we refer to $\mathbf{p}$ and $\mathbf{q}$ as a *standard basis pair*. Since the norm on $\mathbb{X}$ is invariant under reflection in the y-axis, if $\mathbf{v}$ lies on the y-axis, then $\mathsf{O}(\mathbf{v},\mathbf{e}_1)$ holds in $\mathbb{X}$. The following lemma gives the converse, which means that the predicate $\mathsf{B}(\mathbf{p},\mathbf{q})$ characterizes the standard basis pairs in $\mathbb{X}$.

Lemma 10.

(i) $\mathsf{O}(\mathbf{v},\mathbf{e}_1)$ *holds in* $\mathbb{X}$ *iff* $\mathbf{v}$ *lies on the* y*-axis, whence*
(ii) $\mathsf{B}(\mathbf{p},\mathbf{q})$ *holds in* $\mathbb{X}$ *iff* $\mathbf{p}=\pm\mathbf{e}_1$ *and* $\mathbf{q}=\pm\mathbf{e}_2$.

Proof. We have already observed that the points $\pm\mathbf{e}_1$ are the only accumulation points in the set of extreme points of the unit disc. Thus (ii) follows from (i) since (i) implies that the vectors $\pm\mathbf{e}_2$ are the only unit vectors that are equidistant from $\pm\mathbf{e}_1$. By the remarks above, we have only to prove that if $\mathbf{v}$ is equidistant from $\pm\mathbf{e}_1$, then $\mathbf{v}$ lies on the y-axis. Replacing $\mathbf{v}$ by $-\mathbf{v}$ if necessary, we may assume that $\mathbf{v}$ lies in the upper half plane. So, writing $\mathbf{v}=(a,b)$, we may assume $b\geqslant 0$ and what we have to prove is that if $\mathbf{v}$ is equidistant from $\pm\mathbf{e}_1$ then $a=0$.

So assume that $\mathbf{v}$ is equidistant from the points $\pm\mathbf{e}_1$, which means that $\mathbf{v}$ lies in the intersection of the sets $F=\mathbf{e}_1+\lambda S$ and $G=-\mathbf{e}_1+\lambda S$, where $\lambda=\|\mathbf{v}-\mathbf{e}_1\|=\|\mathbf{v}+\mathbf{e}_1\|$. By the triangle inequality, $2=\|\mathbf{e}_1+\mathbf{e}_1\|\leqslant\|\mathbf{e}_1-\mathbf{v}\|+\|\mathbf{e}_1+\mathbf{v}\|=2\lambda$, so $\lambda\geqslant 1$. The upper half of the set F comprises the graph of a function $f:[1-\lambda,1+\lambda]\to\mathbb{R}$ and the upper half of G comprises the graph of a function $g:[-1-\lambda,-1+\lambda]\to\mathbb{R}$. Since $\mathbf{v}=(a,b)$ is in the upper half-plane by assumption, a must lie in the intersection $[1-\lambda,-1+\lambda]$ of the domains of f and g and we have $b=f(a)=g(a)$. As the norm on $\mathbb{X}$ is invariant under reflection in the y-axis, we have $f(x)=g(-x)$ for $x\in[1-\lambda,-1+\lambda]$, thus $f(0)=g(0)$ and the point $(0,f(0))$ lies in the intersection of the two graphs. Now f is strictly increasing on $[1-\lambda,1]$ and strictly decreasing on $[1,1+\lambda]$ and $g(x)=f(x+2)$. So in the (possibly empty) closed interval where f and g are both defined and g is increasing, we have $g(x)>g(x-2)=f(x)$, while where f and g are both defined and f is decreasing we have $f(x)>f(x+2)=g(x)$. Thus $f(a)=g(a)$ implies that a is in the interval where f is increasing and g is decreasing and there can be at most one such a. Hence we must have $(a,b)=(0,f(0))$ so that $a=0$ as required. $\square$

With a few more definitions, we can give a formula of $\mathcal{L}^+$ that in $\mathbb{X}$ characterizes the natural numbers.

$$\mathsf{XAX}(\mathbf{v},\mathbf{p},\mathbf{q}):=\mathbf{v}=\mathbf{0}\vee\big(\mathsf{ACC}(\mathbf{v})\wedge\|\mathbf{v}+\mathbf{p}\|=\|\mathbf{v}\|+\|\mathbf{p}\|\big)$$

$$\mathsf{YAX}(\mathbf{v},\mathbf{p},\mathbf{q}):=\mathsf{O}(\mathbf{v},\mathbf{p})\wedge\|\mathbf{v}+\mathbf{q}\|=\|\mathbf{v}\|+\|\mathbf{q}\|$$

$$\mathsf{Z}(\mathbf{v},\mathbf{p},\mathbf{q}):=\mathsf{XAX}(\mathbf{v},\mathbf{p},\mathbf{q})\wedge\mathsf{EP}(\mathbf{v}+\mathbf{q})$$

$$\mathsf{Nat}(x):=\exists\mathbf{v}\,\mathbf{p}\,\mathbf{q}\cdot x=\|\mathbf{v}\|\wedge\mathsf{B}(\mathbf{p},\mathbf{q})\wedge\mathsf{Z}(\mathbf{v},\mathbf{p},\mathbf{q}).$$

Thus in $\mathbb{X}$, if $\mathbf{p}$ and $\mathbf{q}$ are a standard basis pair: $\mathsf{XAX}(\mathbf{v},\mathbf{p},\mathbf{q})$ holds iff $\mathbf{v}$ lies on the x-axis on the same side as $\mathbf{p}$; $\mathsf{YAX}(\mathbf{v},\mathbf{p},\mathbf{q})$ holds iff $\mathbf{v}$ lies on the y-axis on the same side as $\mathbf{q}$; and for $\mathsf{Z}(\mathbf{v},\mathbf{p},\mathbf{q})$ and $\mathsf{Nat}(x)$ we have:

Lemma 11.

(i) *If* $\mathbf{p}$ *and* $\mathbf{q}$ *are a standard basis pair in* $\mathbb{X}$, $\mathsf{Z}(\mathbf{v},\mathbf{p},\mathbf{q})$ *holds iff* $\mathbf{v}=x\mathbf{p}$ *for some* $x\in\mathbb{N}$, *whence*
(ii) $\mathsf{Nat}(x)$ *holds in* $\mathbb{X}$ *iff* $x\in\mathbb{N}$.

Proof. The right-to-left direction of the claim about $\mathsf{Z}(\mathbf{v},\mathbf{p},\mathbf{q})$ is easy to check. So assume $\mathsf{Z}(\mathbf{v},\mathbf{p},\mathbf{q})$ holds. By Lemma 10, $\mathbf{p}=\pm\mathbf{e}_1$ and $\mathbf{q}=\pm\mathbf{e}_2$. Also $\mathbf{v}$ lies on the x-axis on the same side as $\mathbf{p}$. Thus as $\mathsf{EP}(\mathbf{v}+\mathbf{q})$ holds, $\mathbf{v}+\mathbf{q}=\mathbf{v}\pm\mathbf{e}_2$ is the point of intersection of the line $y = \pm 1$ and one of the lines l_i (since it cannot lie on the x-axis). Thus $\mathbf{v}$ is indeed a natural number multiple of $\mathbf{p} = \pm \mathbf{e}_1$. The claim about $\mathsf{Nat}(x)$ follows, since $\mathsf{B}(\mathbf{p}, \mathbf{q})$ implies that $\|\mathbf{p}\| = 1$. □

The above lemma will give us the undecidability of the theory of any class of normed spaces that includes the 2-dimensional normed space $\mathbb{X}$. The next lemma lets us transfer information about definability in $\mathbb{X}$ to definability in normed spaces and Banach spaces of higher dimensions.

Lemma 12. *For any $d \in \{2, 3, 4, \ldots\} \cup \{\infty\}$, there is a Banach space $\mathbb{X}^d$ with $\mathsf{dim}(\mathbb{X}^d) = d$ such that for any formula $\rho(x_1, \ldots, x_k)$ of $\mathcal{L}_N$ with the indicated free variables (all scalar), there is a formula $\rho^*(x_1, \ldots, x_k)$ of $\mathcal{L}_N$ with the same free variables such that under any assignment of real numbers to the x_i, $\rho^*(x_1, \ldots, x_k)$ holds in $\mathbb{X}^d$ iff $\rho(x_1, \ldots, x_k)$ holds in $\mathbb{X}$. Moreover, ρ^* is additive if ρ is.*

Proof. If V and W are normed spaces, their *1-sum*, $V + W$, is the product vector space $V \times W$ equipped with the norm defined by $\|(\mathbf{p}, \mathbf{q})\| = \|\mathbf{p}\|_V + \|\mathbf{q}\|_W$. $V + W$ has dimension $\mathsf{dim}(V) + \mathsf{dim}(W)$ and is a Banach space iff V and W are both Banach spaces. The subspaces $V \times 0$ and $0 \times W$ are isomorphic to V and W respectively, and the extreme points of the unit disc in $V + W$ comprise the points $(\mathbf{v}, \mathbf{0})$ and $(\mathbf{0}, \mathbf{w})$ where $\mathbf{v}$ and $\mathbf{w}$ are extreme points of the unit discs in V and W respectively.

Let W be the euclidean space $\mathbb{R}^{d-2}$ if $d \neq \infty$ or any infinite-dimensional Hilbert space, e.g., l_2, if $d = \infty$, and let $\mathbb{X}^d = \mathbb{X} + W$. Now every unit vector in the Hilbert space W is an extreme point of the unit disc (a counter-example would give rise to a counter-example in a 2-dimensional subspace and hence a counter-example in $\mathbb{R}^2$). On the other hand, the unit disc in $\mathbb{X}$ has only countably many extreme points. Moreover a point of $\mathbb{X}^d$ lies in $\mathbb{X} \times 0$ iff it is equidistant from $\pm \mathbf{u}$ for every unit vector $\mathbf{u} \in 0 \times W$ (as may be seen by noting that for any unit vectors $\mathbf{x} \in \mathbb{X}$ and $\mathbf{w} \in W$, there is an isomorphism from $\mathbb{R}^2$ under the 1-norm to the subspace of $\mathbb{X} + W$ spanned by $(\mathbf{x}, \mathbf{0})$ and $(\mathbf{0}, \mathbf{w})$ that maps $\mathbf{e}_1$ to $(\mathbf{x}, \mathbf{0})$ and $\mathbf{e}_2$ to $(\mathbf{0}, \mathbf{w})$). It follows that if we define:

$$\mathsf{U}(\mathbf{u}) := \forall \delta \cdot 1 > \delta \geqslant 0 \quad \Rightarrow \quad \exists \mathbf{b} \cdot \|\mathbf{b} - \mathbf{u}\| = \delta \wedge \|\mathbf{b}\| = 1 \wedge \mathsf{EP}(\mathbf{b})$$

$$\mathsf{X}(\mathbf{v}) := \forall \mathbf{u} \cdot \mathsf{U}(\mathbf{u}) \Rightarrow \mathsf{O}(\mathbf{v}, \mathbf{u})$$

then $\mathsf{U}(\mathbf{u})$ holds iff $\mathbf{u}$ is a unit vector in $0 \times W$ and $\mathsf{X}(\mathbf{v})$ holds iff $\mathbf{v}$ is in $\mathbb{X} \times 0$.

Let ρ^* be the relativization of ρ to $\mathsf{X}(\mathbf{v})$, i.e., let ρ^* be obtained from ρ by replacing every subformula of the form $\exists \mathbf{v} \cdot \phi$ by $\exists \mathbf{v} \cdot \mathsf{X}(\mathbf{v}) \wedge \phi$ and every subformula of the form $\forall \mathbf{v} \cdot \phi$ by $\forall \mathbf{v} \cdot \mathsf{X}(\mathbf{v}) \Rightarrow \phi$. Clearly ρ^* is in $\mathcal{L}_N$ and, as $\mathsf{X}(\mathbf{v})$ is additive, ρ^* is additive if ρ is. Since $\mathsf{X}(\mathbf{v})$ holds iff $\mathbf{v}$ belongs to $\mathbb{X} \times 0$, under any assignment of real numbers to the x_i, $\rho^*(x_1, \ldots, x_k)$ holds in $\mathbb{X}^d$ iff $\rho(x_1, \ldots, x_k)$ holds in $\mathbb{X}$. □

We write NS, NS^n, $\mathsf{NS}^{\mathbb{F}}$ and NS^{∞} for the theories of normed spaces where the dimension is respectively unconstrained, constrained to be n, constrained to be finite and constrained to be infinite. We write BS, BS^n etc. for the theories of Banach spaces with the corresponding constraints on the dimension. As finite-dimensional normed spaces are Banach spaces, $\mathsf{BS}^n = \mathsf{NS}^n$ and $\mathsf{BS}^{\mathbb{F}} = \mathsf{NS}^{\mathbb{F}}$.

Theorem 13. *There is a primitive recursive reduction of second-order arithmetic to any of the theories BS, BS^{∞}, NS, NS^n, $\mathsf{NS}^{\mathbb{F}}$, and NS^{∞} ($n \geqslant 2$).*

Proof. What we need to apply Theorem 2 is provided by part (ii) of Lemma 11 using Lemma 12 in the cases of NS^n for $n > 2$, BS^{∞} and NS^{∞}. □

Theorem 13 is already a satisfyingly sharp result, since as we observed at the beginning of this section, the theory of 1-dimensional normed spaces reduces to the theory of the real numbers. But with a little more work, we can show that scalar multiplication can be defined in our space $\mathbb{X}$ in the additive language $\mathcal{L}^+$ and so get a reduction of second-order arithmetic to purely additive normed space theory. To this end we define some more geometric predicates. "ESD" stands for "extreme points, same direction".

$$\mathsf{ESD}(\mathbf{v}, \mathbf{w}) := \mathsf{EP}(\mathbf{v} + \mathbf{w}) \wedge \|\mathbf{v} + \mathbf{w}\| = \|\mathbf{v}\| + \|\mathbf{w}\|.$$

I.e., $\mathsf{ESD}(\mathbf{v}, \mathbf{w})$ holds iff $\mathbf{v} + \mathbf{w}$ is an extreme point of the disc $D_{\|\mathbf{v}+\mathbf{w}\|}$ and equality holds in the triangle inequality for $\mathbf{v}$ and $\mathbf{w}$. I claim that $\mathsf{ESD}(\mathbf{v}, \mathbf{w})$ holds in any normed space iff either $\mathbf{v} = \mathbf{w} = \mathbf{0}$ or there is an extreme point $\mathbf{u}$ of the unit disc such that $\mathbf{v} = x\mathbf{u}$ and $\mathbf{w} = y\mathbf{u}$ for some non-negative x and y. Thus $\mathsf{ESD}(\mathbf{v}, \mathbf{w})$ holds in $\mathbb{X}$ iff $\mathbf{v}$ and $\mathbf{w}$ lie on the same side of the origin on the x-axis or on one of the lines l_i. My claim follows easily from the following lemma:

Lemma 14. *Let $\mathbf{v}$ and $\mathbf{w}$ be non-zero vectors in a normed space. If $\mathbf{v} + \mathbf{w}$ is an extreme point of the disc $D_{\|\mathbf{v}+\mathbf{w}\|}$ of radius $\|\mathbf{v} + \mathbf{w}\|$ and if $\|\mathbf{v} + \mathbf{w}\| = \|\mathbf{v}\| + \|\mathbf{w}\|$, then $\mathbf{v} = \frac{\|\mathbf{v}\|}{\|\mathbf{w}\|}\mathbf{w} = \frac{\|\mathbf{v}\|}{\|\mathbf{v}+\mathbf{w}\|}(\mathbf{v} + \mathbf{w})$.*

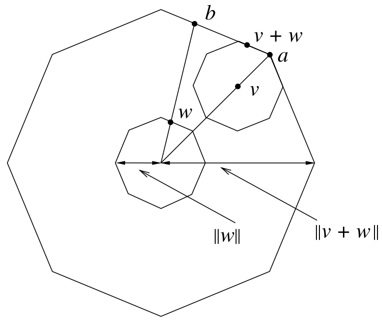

Fig. 5. If $\|\mathbf{v}+\mathbf{w}\| = \|\mathbf{v}\| + \|\mathbf{w}\|$ and $\mathbf{a} \neq \mathbf{b}$, then $\mathbf{v}+\mathbf{w} \in (\mathbf{b}, \mathbf{a})$.

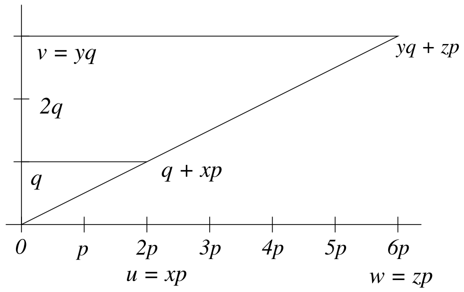

Fig. 6. $z = xy$.

Proof. Under the given hypotheses on $\mathbf{v}$ and $\mathbf{w}$, let $\mathbf{a} = \frac{\|\mathbf{v}+\mathbf{w}\|}{\|\mathbf{v}\|}\mathbf{v}$ and $\mathbf{b} = \frac{\|\mathbf{v}+\mathbf{w}\|}{\|\mathbf{w}\|}\mathbf{w}$ (see Fig. 5). As $\|\mathbf{v}+\mathbf{w}\| = \|\mathbf{v}\| + \|\mathbf{w}\|$, we have:

$$\mathbf{v}+\mathbf{w} = \frac{\|\mathbf{v}\|}{\|\mathbf{v}+\mathbf{w}\|}\mathbf{a} + \frac{\|\mathbf{w}\|}{\|\mathbf{v}+\mathbf{w}\|}\mathbf{b} = \frac{\|\mathbf{v}\|}{\|\mathbf{v}\|+\|\mathbf{w}\|}\mathbf{a} + \left(1 - \frac{\|\mathbf{v}\|}{\|\mathbf{v}\|+\|\mathbf{w}\|}\right)\mathbf{b}.$$

Thus $\mathbf{v}+\mathbf{w}$ is a proper convex combination of $\mathbf{a}$ and $\mathbf{b}$. As $\|\mathbf{a}\| = \|\mathbf{b}\| = \|\mathbf{v}+\mathbf{w}\|$ and $\mathbf{v}+\mathbf{w}$ is an extreme point of the disc $D_{\|\mathbf{v}+\mathbf{w}\|}$, we must have $\mathbf{a} = \mathbf{b}$, i.e., $\frac{\|\mathbf{v}+\mathbf{w}\|}{\|\mathbf{v}\|}\mathbf{v} = \frac{\|\mathbf{v}+\mathbf{w}\|}{\|\mathbf{w}\|}\mathbf{w}$ implying $\mathbf{v} = \frac{\|\mathbf{v}\|}{\|\mathbf{w}\|}\mathbf{w}$ and so also $\mathbf{v} = \frac{\|\mathbf{v}\|}{\|\mathbf{v}+\mathbf{w}\|}(\mathbf{v}+\mathbf{w})$. $\square$

We now give the geometric predicate that will allow us to define multiplication (see Fig. 6).

$$\begin{aligned}\mathsf{NTIMES}(x, y, z) := \exists \mathbf{p}\,\mathbf{q}\,\mathbf{u}\,\mathbf{v}\,\mathbf{w} \cdot x = \|\mathbf{u}\| \wedge y = \|\mathbf{v}\| \wedge z = \|\mathbf{w}\| \wedge \\ \mathsf{B}(\mathbf{p}, \mathbf{q}) \wedge \mathsf{Z}(\mathbf{u}, \mathbf{p}, \mathbf{q}) \wedge \mathsf{YAX}(\mathbf{v}, \mathbf{p}, \mathbf{q}) \wedge \mathsf{XAX}(\mathbf{w}, \mathbf{p}, \mathbf{q}) \wedge \\ \mathsf{ESD}(\mathbf{q}+\mathbf{u}, \mathbf{v}+\mathbf{w}).\end{aligned}$$

Lemma 15. *In $\mathbb{X}$, $\mathsf{NTIMES}(x, y, z)$ holds iff $x \in \mathbb{N}$, $y, z \in \mathbb{R}_{\geqslant 0}$ and $z = xy$.*

Proof. By reference to Fig. 6, it is easy to see that the right-to-left direction of the lemma holds (put $\mathbf{p} = \mathbf{e}_1$, $\mathbf{q} = \mathbf{e}_2$, $\mathbf{u} = x\mathbf{p}$, $\mathbf{v} = y\mathbf{q}$ and $\mathbf{w} = z\mathbf{p}$). Conversely, let $\mathbf{p}$, $\mathbf{q}$, $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ be witnesses to the truth of the existential formula $\mathsf{NTIMES}(x, y, z)$, so that $\|\mathbf{u}\| = x$, $\|\mathbf{v}\| = y$, $\|\mathbf{w}\| = z$. Since $\mathbf{p}$ and $\mathbf{q}$ are a standard basis pair and $\mathsf{Z}(\mathbf{u}, \mathbf{p}, \mathbf{q})$, by Lemma 11 we have that $x \in \mathbb{N}$ and $\mathbf{u} = x\mathbf{p}$. Also, since $\mathsf{YAX}(\mathbf{v}, \mathbf{p}, \mathbf{q})$ and $\mathsf{XAX}(\mathbf{w}, \mathbf{p}, \mathbf{q})$ hold, we have that $\mathbf{v} = y\mathbf{q}$ and $\mathbf{w} = z\mathbf{p}$. Now $\mathbf{q}+\mathbf{u} = \mathbf{q}+x\mathbf{p}$ lies on the line l_x passing through the point $(x, 1)$. Moreover, since $\mathsf{ESD}(\mathbf{q}+\mathbf{u}, \mathbf{v}+\mathbf{w})$ holds, $\mathbf{v}+\mathbf{w}$ also lies on l_x. But this means that the right-angled triangle A with vertices $\mathbf{0}$, $\mathbf{q}$ and $\mathbf{q}+\mathbf{u} = \mathbf{q}+x\mathbf{p}$ is similar to and parallel to the triangle B with vertices $\mathbf{0}$, $\mathbf{v} = y\mathbf{q}$ and $\mathbf{v}+\mathbf{w} = y\mathbf{q}+z\mathbf{p}$. Hence $z = xy$ completing the proof. $\square$

Lemma 16. *There is a formula $\mathsf{RTIMES}(x, y, z)$ in the additive language $\mathcal{L}_N^+$ which holds in $\mathbb{X}$ iff $z = xy$.*

Proof. Consider the following formulas of $\mathcal{L}_N^+$:

$$\begin{aligned}\mathsf{ZTIMES}(x, y, z) := \mathsf{NTIMES}(x, y, z) \vee \mathsf{NTIMES}(-x, y, -z) \vee \\ \mathsf{NTIMES}(x, -y, -z) \vee \mathsf{NTIMES}(-x, -y, z)\end{aligned}$$

$$\begin{aligned}\mathsf{QTIMES}(x,y,z) &:= \exists m\, n\, t \cdot n \neq 0 \wedge \\ &\qquad \mathsf{ZTIMES}(n,x,m) \wedge \mathsf{ZTIMES}(m,y,t) \wedge \mathsf{ZTIMES}(n,z,t)\\ \mathsf{RTIMES}(x,y,z) &:= \forall \epsilon \cdot \epsilon > 0 \Rightarrow \big(\exists \delta \cdot \delta > 0 \wedge \\ &\qquad \big(\forall r\, t \cdot |x-r| < \delta \wedge \mathsf{QTIMES}(r,y,t) \Rightarrow |z-t| < \epsilon\big)\big).\end{aligned}$$

By Lemma 15, in $\mathbb{X}$, $\mathsf{NTIMES}(x,y,z)$ defines the graph of the multiplication function restricted to $\mathbb{N}\times\mathbb{R}_{\geqslant 0}$. The predicate $\mathsf{ZTIMES}(x,y,z)$ therefore defines the graph of multiplication restricted to $\mathbb{Z}\times\mathbb{R}$. In the formula $\mathsf{QTIMES}(x,y,z)$, the matrix of the right-hand side of the definition asserts that $nx=m$ and that $my=t=nz$, so that, when $n\neq 0$, $z=(m/n)y=xy$, so $\mathsf{QTIMES}(x,y,z)$ defines the graph of multiplication restricted to $\mathbb{Q}\times\mathbb{R}$. By continuity, we have that $\mathsf{RTIMES}(x,y,z)$ defines the graph of multiplication without restriction completing the proof of the lemma. □

We write NS_+, NS^n_+, BS_+, etc. for the additive subtheories of NS, NS^n, BS, etc.

Theorem 17. *There is a primitive recursive reduction of second-order arithmetic to any of the theories* BS_+, BS^∞_+, NS_+, $\mathsf{NS}^{\mathbb{F}}_+$, NS^n_+, *and* NS^∞_+ *(*$n\geqslant 2$*).*

Proof. What we need to apply Theorem 3 is provided by part (ii) of Lemma 11 and Lemma 16 using Lemma 12 in the cases of NS^n_+ for $n>2$, BS^∞_+ and NS^∞_+. □

5.2. The many-one degrees of theories of normed spaces

Theorems 13 and 17 show that the decision problems for our theories of normed spaces and Banach spaces are at least as hard as that for the theory of second-order arithmetic. We now consider the converse problem of reducing the normed space and Banach space theories to theories of higher-order arithmetic.

As usual, writing $|A|$ for the cardinality of a set A, let $\aleph_0=|\mathbb{N}|$ be the first infinite cardinal and let $\mathfrak{c}=2^{\aleph_0}=|\mathbb{R}|$ be the cardinality of the continuum. If A is any non-empty finite or countably infinite set, the set $\mathbb{R}^A$ of real-valued functions on A has cardinality $\mathfrak{c}$. In particular, the set $\mathbb{R}^{\mathbb{N}}$ of countably infinite sequences of real numbers has cardinality $\mathfrak{c}$. If V is a vector space we write $|V|$ for the cardinality of its set of vectors. Note that $|V|$ is either 1 or at least $\mathfrak{c}$. We write $\mathbb{V}^{\mathfrak{c}}$ for some fixed vector space with a basis B of cardinality $\mathfrak{c}$, say $B=\{\mathbf{b}_x \mid x\in\mathbb{R}\}$. Clearly $|\mathbb{V}^{\mathfrak{c}}|\geqslant|B|=\mathfrak{c}$ and, conversely, as any element of $\mathbb{V}^{\mathfrak{c}}$ is a finite sum $\sum_{m=0}^{k} c_m\mathbf{b}_{x_m}$ for some $c_m, x_m\in\mathbb{R}$, $|\mathbb{V}^{\mathfrak{c}}|$ is at most $|(\mathbb{R}\times\mathbb{R})^{\mathbb{N}}|=\mathfrak{c}$. Thus a vector space has cardinality at most $\mathfrak{c}$ iff it is isomorphic to a subspace of $\mathbb{V}^{\mathfrak{c}}$. The following Skolem–Löwenheim theorem thus implies that any satisfiable first-order property of normed spaces or Banach spaces is satisfiable in a space given by equipping some subspace of $\mathbb{V}^{\mathfrak{c}}$ with a norm.

Theorem 18. *Let V be a real vector space. Then V has a subspace W with $|W|\leqslant\mathfrak{c}$ that is an elementary substructure of V, i.e., a sentence ϕ in the language $\mathcal{L}_N$ of normed spaces holds in V iff it holds in W. Moreover, W may be taken to be a Banach space if V is a Banach space.*

Proof. We will construct W using a certain function $F:\mathbb{N}\times\mathbb{R}^{\mathbb{N}}\times V^{\mathbb{N}}\to V$. Let us first show that for any such function there is a subset W of V of cardinality at most $\mathfrak{c}$ that is F-closed in the sense that $F[\mathbb{N}\times\mathbb{R}^{\mathbb{N}}\times W^{\mathbb{N}}]\subseteq W$. To see this, define a transfinite sequence of subsets W_α of V as follows, where α is any ordinal and λ is any limit ordinal:

$$\begin{aligned}W_0 &= \{\mathbf{0}\}\\ W_{\alpha+1} &= W_\alpha \cup F\big[\mathbb{N}\times\mathbb{R}^{\mathbb{N}}\times (W_\alpha)^{\mathbb{N}}\big]\\ W_\lambda &= \bigcup_{\alpha<\lambda} W_\alpha.\end{aligned}$$

Let $\aleph_1$ be the smallest uncountable cardinal and let $W=W_{\aleph_1}$. By transfinite induction, one may show that $|W_\alpha|\leqslant\mathfrak{c}$ for $\alpha\leqslant\aleph_1$, and so in particular $|W|\leqslant\mathfrak{c}$. Now if $(k,s,\mathbf{x})\in\mathbb{N}\times\mathbb{R}^{\mathbb{N}}\times W^{\mathbb{N}}$ then I claim $F(k,s,\mathbf{x})\in W$. For if α is the least ordinal such that $\mathbf{x}_m\in W_\alpha$ for all $m\in\mathbb{N}$, then $\alpha<\aleph_1$ (since α can be written as a countable union of countable ordinals and hence is countable). Thus $F(k,s,\mathbf{x})\in W_{\alpha+1}\subset W$ and W is indeed an F-closed subset of V of cardinality at most $\mathfrak{c}$.

To define the function F, let the formulas of $\mathcal{L}_N$ be enumerated as $\psi_1,\psi_2,\ldots$. We fix a total ordering on the variables of $\mathcal{L}_N$ and choose a vector variable $\mathbf{v}$, and then given $(k,s,\mathbf{x})\in\mathbb{N}\times\mathbb{R}^{\mathbb{N}}\times V^{\mathbb{N}}$, we define $F(k,s,\mathbf{x})$ as follows:

1. if $k=0$ and the $\mathbf{x}_m$ converge in V to a limit $\mathbf{p}$, we set $F(k,s,\mathbf{x})=\mathbf{p}$;
2. if $k>0$, consider the formula $\psi:=\exists\mathbf{v}\cdot\psi_k$ and let $x_0,\ldots,x_m$ and $\mathbf{v}_0,\ldots,\mathbf{v}_n$ list its free scalar and vector variables in order. We interpret x_i as s_i and $\mathbf{v}_j$ as $\mathbf{x}_j$. If ψ is true in V under this interpretation, then there is a $\mathbf{q}$ in V such that ψ_k becomes true if we extend the interpretation by interpreting $\mathbf{v}$ as $\mathbf{q}$, we choose such a $\mathbf{q}$ and set $F(k,s,\mathbf{x})=\mathbf{q}$;
3. in all other cases, we set $F(k,s,\mathbf{x})=\mathbf{0}$.

Now let $W \subseteq V$ be an F-closed subset of cardinality at most $\mathfrak{c}$ as constructed above. Clause 2 of the definition of F ensures that the Tarski–Vaught criterion applies so that W is an elementary substructure of V; see, e.g., Hodges [20]. In particular, W is a vector space over some subfield of $\mathbb{R}$. Clause 1 implies that the 1-dimensional subspaces of this vector space are metrically complete, so the field of scalars of W may be taken to be $\mathbb{R}$ so that W is a subspace of V. Finally, if V is a Banach space, clause 1 implies that W is also a Banach space. $\square$

It will simplify our syntactic constructions to extend the language $\mathcal{L}^2_A$ of second-order arithmetic as follows: first let $\mathcal{L}^2_{AR}$ be the result of adding to $\mathcal{L}^2_A$ a sort $\mathcal{R}$ for the real numbers together with function and predicate symbols for the operations of the ordered field $\mathbb{R}$ and for the injection $\iota : \mathcal{N} \to \mathcal{R}$ of $\mathbb{N}$ into $\mathbb{R}$; then let $\mathcal{L}^2_{AV}$ be obtained from $\mathcal{L}^2_{AR}$ by adding a sort $\mathcal{V}$ of vectors, together with function symbols for the vector space operations on $\mathcal{V}$ with scalars in $\mathcal{R}$ and for a function symbol $\gamma : \mathcal{V} \times \mathcal{R} \to \mathcal{R}$. The intended interpretation of $\mathcal{V}$ in $\mathcal{L}^2_{AV}$ is the vector space $\mathbb{V}^{\mathfrak{c}}$ with γ the operation that maps a pair $(\mathbf{v}, x)$ to the coefficient c_x of the basis element $\mathbf{b}_x$ in the expression of $\mathbf{v}$ as a linear combination of elements of the basis B. We choose the symbols so that the language $\mathcal{L}_V$ of vector spaces is a sublanguage of $\mathcal{L}^2_{AV}$.

A *standard model* of one of the languages $\mathcal{L}^2_A$, $\mathcal{L}^2_{AR}$ or $\mathcal{L}^2_{AV}$ is one in which (up to isomorphism) all the sorts and symbols of the language have their intended interpretations. In particular, in a standard model, the sort $\mathcal{P}$ is interpreted as the full powerset $\mathbb{P}(\mathbb{N})$ of the set of natural numbers. Let T^2_A, resp. T^2_{AV}, resp. T^2_{AR}, be the set of all sentences of $\mathcal{L}^2_A$, resp. $\mathcal{L}^2_{AV}$, resp. $\mathcal{L}^2_{AR}$, that are true in a standard model (and hence in all standard models). In the light of the following lemma, to reduce a decision problem to T^2_A, i.e., second-order arithmetic, it is sufficient to reduce it to T^2_{AV}.

Lemma 19. *There are primitive recursive reductions of T^2_{AV} and T^2_{AR} to the theory T^2_A of true sentences of second-order arithmetic.*

Proof. It is well-known that using suitable encodings, the real numbers may be constructed, e.g., via Dedekind cuts, as a definitional extension of second-order arithmetic; see [39]. Unwinding the definitions provides a primitive recursive reduction of T^2_{AR} to T^2_A. (The unwinding process requires occurrences of function symbols first to be unnested so that they can be replaced by predicates as in the proofs of Theorem 2 and Lemma 20.) So it suffices to give a primitive recursive reduction of T^2_{AV} to T^2_{AR}.

Now in $\mathcal{L}^2_{AR}$ we can encode the elements of sets such as $\mathbb{R} \times \mathbb{R}$, $\mathbb{R}^{\mathbb{N}}$, $(\mathbb{R} \times \mathbb{R})^{\mathbb{N}}$, etc. as real numbers. Given a vector $\mathbf{v} = \sum_{m=0}^{k} c_m \mathbf{b}_{x_m} \in \mathbb{V}^{\mathfrak{c}}$, we can arrange for the c_m to be non-zero and for the x_m to be listed in strictly increasing order, and then encode $\mathbf{v}$ as the real number that encodes the sequence s, with $s_m = (c_m, x_m), 0 \leqslant m \leqslant k$, and $s_m = (0, 0), m > k$. Using this encoding we can define the vector space operations on $\mathbb{V}^{\mathfrak{c}}$ together with the function γ. Unwinding these definitions gives the required primitive recursive reduction of T^2_{AV} to T^2_{AR}. $\square$

Now let $\mathcal{L}^2_{AN}$ be $\mathcal{L}^2_{AV}$ extended with a predicate symbol NRM of type $\mathcal{V} \times \mathcal{R}$. A *standard model* of a sentence of $\mathcal{L}^2_{AN}$ is to be one which extends a standard model of $\mathcal{L}^2_{AV}$, i.e., one in which all the sorts and symbols of $\mathcal{L}^2_{AV}$ have their intended interpretations while the interpretation of NRM is arbitrary.

Lemma 20. *There are primitive recursive functions, $\phi \mapsto \phi_N$ and $\phi \mapsto \phi_B$, which map sentences of the language $\mathcal{L}_N$ of normed spaces to sentences of $\mathcal{L}^2_{AN}$, such that the standard models of ϕ_N (resp. ϕ_B) comprise precisely those standard models in which $\mathsf{NRM}(\mathbf{v}, x)$ defines a norm on a subspace of $\mathbb{V}^{\mathfrak{c}}$ that provides a model (resp. Banach space model) of ϕ. Moreover ϕ has a model (resp. Banach space model) iff ϕ_N (resp. ϕ_B) has a standard model.*

Proof. There is a primitive recursive function mapping ϕ to a logically equivalent sentence ϕ_1 in which all occurrences of the norm operator are unnested, i.e., in which the norm operator only appears in atomic formulas of the form $\|\mathbf{v}\| = x$ where $\mathbf{v}$ and x are variables. Let ϕ_2 be the result of replacing each subformula $\|\mathbf{v}\| = x$ in ϕ_1 by $\mathsf{NRM}(\mathbf{v}, x)$. Then ϕ_2 is a sentence of $\mathcal{L}^2_{AN}$. Let ϕ_3 be the relativization of ϕ_2 to the domain of the relation defined by $\mathsf{NRM}(\mathbf{v}, x)$, i.e., obtain ϕ_3 from ϕ_2 by replacing subformulas of the form $\exists \mathbf{v} \cdot \psi$ by $\exists \mathbf{v} \cdot (\exists x \cdot \mathsf{NRM}(\mathbf{v}, x)) \wedge \psi$ and subformulas of the form $\forall \mathbf{v} \cdot \psi$ by $\forall \mathbf{v} \cdot (\exists x \cdot \mathsf{NRM}(\mathbf{v}, x)) \Rightarrow \psi$.

There is a sentence Q_N of $\mathcal{L}^2_{AN}$ asserting that $\mathsf{NRM}(\mathbf{v}, x)$ defines a relation that (i) is a partial function, (ii) has a domain that is closed under the vector space operations and (iii) satisfies the conditions for a norm on the vectors in its domain. As completeness may be defined using quantification over countably infinite sequences of vectors, which is available in $\mathcal{L}^2_{AN}$, there is a sentence Q_B of $\mathcal{L}^2_{AN}$ asserting that the metric given by $\mathsf{NRM}(\mathbf{v}, x)$ is complete. We take $\phi_N := \mathsf{Q}_N \wedge \phi_3$ and $\phi_B := \mathsf{Q}_N \wedge \mathsf{Q}_B \wedge \phi_3$.

Now if ϕ has a normed space model (resp. Banach space model), then by Theorem 18, it has a model that is isomorphic to a subspace W of $\mathbb{V}^{\mathfrak{c}}$ under some norm $\|\_\|$. Extending the standard interpretation of $\mathcal{L}^2_{AV}$ to interpret $\mathsf{NRM}(\mathbf{v}, x)$ as $\mathbf{v} \in W \wedge \|\mathbf{v}\| = x$ gives a standard model of ϕ_N (resp. ϕ_B). Conversely, a standard model of ϕ_N (resp. ϕ_B) gives a normed space model (resp. Banach space model) isomorphic to a subspace of $\mathbb{V}^{\mathfrak{c}}$ under the norm defined by the interpretation of $\mathsf{NRM}(\mathbf{v}, x)$. $\square$

Theorem 21. *There are primitive recursive reductions of the theories* $\mathsf{NS}^{\mathbb{F}}$ *and* NS^n, $n \in \mathbb{N}$, *to* T^2_{AV} *and hence to second-order arithmetic.*

Proof. Let a natural number n and a real number x be given. In T^2_{AV}, we can define the subspace $\mathbb{R}^n$ of $\mathbb{V}^c$ spanned by the $\mathbf{b}_m$, $1 \leqslant m \leqslant n, m \in \mathbb{N}$; we can define the subset $\mathbb{Q}^n$ of $\mathbb{R}^n$ comprising the points with rational coordinates; since $\mathbb{Q}^n$ is countable, we can view x as an encoding of an arbitrary subset $\mathbf{Q}^n_x$ of $\mathbb{Q}^n$; using the coefficient function γ, we can define the euclidean norm on $\mathbb{V}^c$ with respect to the basis B. Thus there is a formula $\Delta(n, x, \mathbf{p}, t)$ of $\mathcal{L}^2_{AV}$ that holds in a standard model iff every open disc in $\mathbb{R}^n$ centred on $\mathbf{p}$ meets both $t\mathbf{Q}^n_x$ and its complement $\mathbb{Q}^n \setminus t\mathbf{Q}^n_x$. But then if $\mathbf{Q}^n_x$ is the set $\mathbb{Q}^n \cap D$ of rational points in the unit disc D of a norm $\|\_\|$ on $\mathbb{R}^n$, $\Delta(n, x, \mathbf{p}, t)$ holds iff $\|\mathbf{p}\| = t$.

We complete the proof for $\mathsf{NS}^{\mathbb{F}}$, the proof for NS^n being very similar. As a sentence is valid iff its negation is unsatisfiable, it is sufficient to give a primitive recursive function $\phi \mapsto \phi_F$ from $\mathcal{L}_N$ to $\mathcal{L}^2_{AV}$ such that ϕ is satisfiable in a finite-dimensional normed space iff ϕ_F is true. Applying Lemma 20, we have a sentence ϕ_N of $\mathcal{L}^2_{AN}$ that has a standard model iff ϕ is satisfiable. Choose variables n of sort $\mathcal{N}$ and x of sort $\mathcal{R}$ that do not appear in ϕ_N and let ψ be the result of replacing each occurrence of $\mathsf{NRM}(\mathbf{p}, t)$ in ϕ_N by $\Delta(n, x, \mathbf{p}, t)$. Setting $\phi_F := \exists n\, x \cdot \psi$, ϕ_F holds in a standard model of $\mathcal{L}^2_{AV}$ iff there are $n \in \mathbb{N}$ and $x \in \mathbb{R}$ such that $\mathbf{Q}^n_x$ is a set of rational points whose closure is the unit disc of a norm on $\mathbb{R}^n$ and ϕ holds under this norm on $\mathbb{R}^n$. Since any n-dimensional normed space is isomorphic to one given by defining a norm on $\mathbb{R}^n$, ϕ_F is true iff ϕ is satisfiable. □

Using the terminology of recursion theory we have the following corollary concerning degrees of unsolvability; see, e.g., Rogers [36] for definitions.

Corollary 22. *The theories* $\mathsf{NS}^{\mathbb{F}} = \mathsf{BS}^{\mathbb{F}}$, $\mathsf{NS}^{\mathbb{F}}_+ = \mathsf{BS}^{\mathbb{F}}_+$ *and* $\mathsf{NS}^n = \mathsf{BS}^n$, $\mathsf{NS}^n_+ = \mathsf{BS}^n_+$, $n \geqslant 2$, *all have the same many-one degree as the theory* T^2_A *of second-order arithmetic.*

Proof. This is immediate from Theorems 17 and 21. □

Now let $\mathcal{L}^3_A$ be the language of third-order arithmetic. This is $\mathcal{L}^2_A$ extended with an additional sort $\mathcal{P}_2$ called “type 2” whose intended interpretation is $\mathbb{P}(\mathbb{P}(\mathbb{N}))$. $\mathcal{L}^3_A$ has a predicate symbol $\in$ of type $\mathcal{P} \times \mathcal{P}_2$ to denote the membership relation and a supply of type 2 variables $u = u_1, u_2, \ldots$, but we shall only need the first of these. A sentence of $\mathcal{L}^3_A$ is said to be Σ^2_1 (resp. Π^2_1) if it has the form $\exists u \cdot \psi(u)$ (resp. $\forall u \cdot \psi(u)$) where $\psi(u)$ contains no quantifiers over type 2 variables.

Theorem 23. *There are primitive recursive reductions of each of the theories* NS, BS, NS^∞ *and* BS^∞ *to the set of true* Π^2_1 *sentences.*

Proof. As with $\mathcal{L}^2_A$ we are free to work in a definitional extension $\mathcal{L}^3_{AV}$ of $\mathcal{L}^3_A$ that includes the language $\mathcal{L}_V$ of vector spaces (with $\mathbb{V}^c$ as the intended interpretation of the vector sort). Let a, $\mathbf{v}$ and x be variables of sort $\mathcal{P}$, $\mathcal{V}$ and $\mathcal{R}$ respectively. There is a formula $\mathsf{U}(a, \mathbf{v}, x)$ of $\mathcal{L}^2_{AV} \subseteq \mathcal{L}^3_{AV}$ with the indicated free variables that in a standard model of $\mathcal{L}^2_A$ defines the graph of a bijection mapping $a \in \mathbb{P}(\mathbb{N})$ to $(\mathbf{v}, x) \in \mathbb{V}^c \times \mathbb{R}$. This gives an encoding of all relations between $\mathbb{V}^c$ and $\mathbb{R}$, i.e., all subsets of $\mathbb{V}^c \times \mathbb{R}$, as type 2 sets.

To complete the proof, let us first consider NS. As a sentence is valid iff its negation is unsatisfiable, it suffices to give a primitive recursive function $\phi \mapsto \phi_1$ from the language $\mathcal{L}_N$ of normed spaces to the set of Σ^2_1 sentences such that ϕ is satisfiable iff ϕ_1 is true. Given a sentence ϕ in the language of normed spaces, apply Lemma 20, to give a sentence ϕ_N of $\mathcal{L}^2_{AN}$ that has a standard model iff ϕ is satisfiable. Let $\psi(u)$ be obtained from ϕ_N by replacing all instances of $\mathsf{NRM}(\mathbf{v}, x)$ by $\exists a \cdot a \in u \wedge \mathsf{U}(a, \mathbf{v}, x)$ and let ϕ_1 be $\exists u \cdot \psi(u)$. Then ϕ_1 is a Σ^2_1 formula that is true iff ϕ_N has a standard model. So ϕ_1 is true iff ϕ is satisfiable.

For BS, we use a primitive recursive function $\phi \mapsto \phi_2$ from $\mathcal{L}_N$ to the set of Σ^2_1 sentences such that ϕ is satisfiable in a Banach space iff ϕ_2 is true. The construction of ϕ_2 is identical to that of ϕ_1 except that we use the sentence ϕ_B from Lemma 20 rather than ϕ_N.

Finally, for NS^∞ and BS^∞, there is a formula $\mathsf{I}(u)$ of A^3 with no type 2 quantifiers which holds iff u encodes a relation between $\mathbb{V}^c$ and $\mathbb{R}$ whose domain is an infinite-dimensional subspace of $\mathbb{V}^c$. Relativization of ϕ_1 and ϕ_2 to $\mathsf{I}(u)$ gives the reductions required to complete the proof. □

We complete our study of the degrees of unsolvability of the normed space and Banach space theories by exhibiting a primitive recursive reduction of the set of true Π^2_1 sentences to the theories NS_+ and BS_+. To do this we need Banach spaces in which an arbitrary subset of the open interval $(0, 1)$ can be defined in a uniform way. We begin by considering the special case of a singleton set. So let $t \in (0, 1)$ be given and define points of $\mathbb{R}^2$ by $\mathbf{u} = (-1, -1)$, $\mathbf{v} = (1, -1)$ and $\mathbf{w} = (\frac{2}{t}, 0)$. Let $\mathbb{S}_t$ be $\mathbb{R}^2$ equipped with the norm $\|\_\|_t$ whose unit circle comprises the hexagon with vertices $\pm\mathbf{u}$, $\pm\mathbf{v}$ and $\pm\mathbf{w}$ (see Fig. 7). One finds using the ordinary euclidean norm $\|\_\|_e$ that $\|\mathbf{v} - \mathbf{u}\|_t = \frac{\|\mathbf{v}-\mathbf{u}\|_e}{\|\mathbf{w}\|_e} = \frac{2}{2/t} = t$. Let the line through $\mathbf{v}$ and $\mathbf{w}$ meet the y-axis at the point $\mathbf{p}$. Then the line segment $[\mathbf{p}, \mathbf{w}]$ is a translate of the line segment $[(\mathbf{p} - \mathbf{w})/2, (\mathbf{w} - \mathbf{p})/2]$

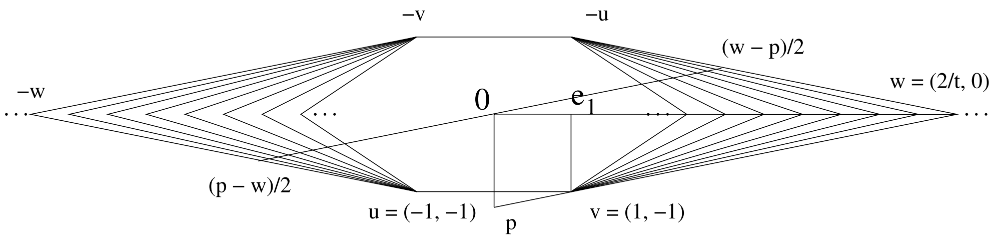

Fig. 7. The unit discs of the spaces $\mathbb{S}_t$.

which is a diameter of the unit disc in $\mathbb{S}_t$. So $\|\mathbf{w}-\mathbf{p}\|_t = 2$ and $\|\mathbf{w}-\mathbf{v}\|_t = 2 - \|\mathbf{v}-\mathbf{p}\|_t$. But the triangles $\mathbf{0pw}$ and $\mathbf{e}_1\mathbf{vw}$ are similar and so $\|\mathbf{v}-\mathbf{p}\|_t = \|\mathbf{w}-\mathbf{p}\|_t(\frac{\|\mathbf{v}-\mathbf{p}\|_e}{\|\mathbf{w}-\mathbf{p}\|_e}) = 2(\frac{\|\mathbf{e}_1\|_e}{\|\mathbf{w}\|_e}) = t$ whence $\|\mathbf{w}-\mathbf{v}\|_t = 2 - \|\mathbf{v}-\mathbf{p}\|_t = 2-t$. By symmetry, each edge of the hexagon that comprises the unit circle in $\mathbb{S}_t$ has length t or $2-t$ in the $\mathbb{S}_t$ norm.

Let us say that two vectors $\mathbf{p}$ and $\mathbf{q}$ in a normed space V are *adjacent* if $\mathbf{p}$ and $\mathbf{q}$ are distinct extreme points of the set $S_{\|\mathbf{p}\|}$ of vectors of length $\|\mathbf{p}\|$ and $\|\frac{1}{2}(\mathbf{p}+\mathbf{q})\| = \|\mathbf{p}\|$. This implies that the line segment $[\mathbf{p},\mathbf{q}]$ is the intersection of some affine line with the set $S_{\|\mathbf{p}\|}$. If $\mathbf{p}$ and $\mathbf{q}$ are adjacent unit vectors then $\|\mathbf{p}-\mathbf{q}\| \leqslant 2$ with equality iff $\mathbf{p}$ and $-\mathbf{q}$ are also adjacent unit vectors, in which case the linear transformation that maps $\mathbf{e}_1$ to $\mathbf{p}$ and $\mathbf{e}_2$ to $\mathbf{q}$ defines an isomorphism between $\mathbb{R}^2$ under the 1-norm and the subspace of V spanned by $\mathbf{p}$ and $\mathbf{q}$. Now consider the following formulas in $\mathcal{L}_N^+$, the first of which formalizes the notion of adjacency.

$$\mathsf{ADJ}(\mathbf{p},\mathbf{q}) := \mathsf{EP}(\mathbf{p}) \wedge \mathsf{EP}(\mathbf{q}) \wedge \mathbf{p} \neq \mathbf{q} \wedge \|\mathbf{p}\| = \|\mathbf{q}\| = \big\|(\mathbf{p}+\mathbf{q})/2\big\|$$

$$\mathsf{H}(\mathbf{u},\mathbf{v},\mathbf{w}) := \mathsf{ADJ}(\mathbf{u},\mathbf{v}) \wedge \mathsf{ADJ}(\mathbf{v},\mathbf{w}) \wedge \mathsf{ADJ}(\mathbf{w},-\mathbf{u}) \wedge$$
$$\|\mathbf{v}-\mathbf{u}\| < 2\|\mathbf{v}\| \wedge \|\mathbf{w}-\mathbf{v}\| < 2\|\mathbf{v}\| \wedge \|\mathbf{w}+\mathbf{u}\| < 2\|\mathbf{v}\|$$

$$\mathsf{T}(x) := \exists \mathbf{u}\,\mathbf{v}\,\mathbf{w} \cdot \|\mathbf{u}\| = 1 \wedge \mathsf{H}(\mathbf{u},\mathbf{v},\mathbf{w}) \wedge x = \|\mathbf{v}-\mathbf{u}\| < 1.$$

Clearly $\mathsf{ADJ}(\mathbf{u},\mathbf{v})$, $\mathsf{ADJ}(\mathbf{v},\mathbf{u})$ and $\mathsf{ADJ}(-\mathbf{u},-\mathbf{v})$ are all equivalent and so in any normed space, $\mathsf{H}(\mathbf{u},\mathbf{v},\mathbf{w})$ implies that the vectors $\mathbf{u},\mathbf{v},\mathbf{w},-\mathbf{w},-\mathbf{u},-\mathbf{w}$ are the vertices of a hexagon inscribed in the set $S_{\|\mathbf{u}\|}$ of vectors of length $\|\mathbf{u}\|$. $\mathsf{H}(\mathbf{u},\mathbf{v},\mathbf{w})$ also includes a condition on the length of the edges of this hexagon that will presently help us pick out elements of $\mathbb{S}_t$ when it is embedded in a larger space. Now in $\mathbb{S}_t$, if $\mathbf{u}$ and $\mathbf{v}$ are unit vectors and $\|\mathbf{v}-\mathbf{u}\| < 1$, $\mathsf{H}(\mathbf{u},\mathbf{v},\mathbf{w})$ can only hold if $\mathbf{u}$ and $\mathbf{v}$ are the end-points of one of the edges of S_1 whose length in the $\mathbb{S}_t$ norm is t, so in $\mathbb{S}_t$, $\mathsf{T}(x)$ defines the singleton set $\{t\}$.

We now need a generalization of the 1-sum construction that we used in the proof of Lemma 12. Let V_i, $i \in I$, be an arbitrary family of normed spaces and write $\|\_\|_i$ for the norm on V_i. If $\mathbf{f}$ is a member of $\prod_{i\in I} V_i$ and if J is a finite subset of I, let $n(\mathbf{f},J) = \sum_{j\in J}\|\mathbf{f}_j\|_j$. The *1-sum* $\sum_{i\in I} V_i$ comprises those $\mathbf{f}$ for which $n(\mathbf{f},J)$ is bounded as J ranges over all finite subsets of I. We define $\|\mathbf{f}\|$ to be the supremum of the $n(\mathbf{f},J)$. As is easily verified, $\sum_{i\in I} V_i$ is a normed space and is a Banach space iff the V_i are all Banach spaces. There is a natural isomorphism between the summand V_i and the subspace of $\sum_{i\in I} V_i$ comprising those $\mathbf{f}$ such that $\mathbf{f}_j = \mathbf{0}$ whenever $j \neq i$ and we may identify V_i with that subspace. Under this identification, the extreme points of the unit disc in $\sum_{i\in I} V_i$ comprise the union of the extreme points of the unit discs of the V_i. If $\mathbf{p} \in V_i$ and $\mathbf{q} \in V_j$ are unit vectors and $i \neq j$, then in the 1-sum, $\|\mathbf{p}-\mathbf{q}\| = 2$.

If T is any subset of the interval $(0,1)$, let $\mathbb{S}_T = \sum_{t\in T}\mathbb{S}_t$. Then $\mathbb{S}_T$ is the 1-sum of Banach spaces and hence is itself a Banach space. I claim that the formula $\mathsf{T}(x)$ that defines t in the space $\mathbb{S}_t$ defines T in the 1-sum $\mathbb{S}_T + V$ where V is any normed space whose unit circle contains no hexagons. For, assume that $\mathsf{T}(t)$ holds for some t. Then there are extreme points $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ of the unit disc in $\mathbb{S}_T + V$ such that $t = \|\mathbf{v}-\mathbf{u}\| < 1$ and $\|\mathbf{w}-\mathbf{v}\| < 2\|\mathbf{v}\| = 2$. Now as $\|\mathbf{v}-\mathbf{u}\|, \|\mathbf{w}-\mathbf{v}\| < 2$, $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ are either all in $\mathbb{S}_T$ or all in V (viewed as subspaces of $\mathbb{S}_T + V$), and as they lie on a hexagon contained in the unit circle they must all lie in $\mathbb{S}_T$. But then $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ must belong to the same summand of $\mathbb{S}_T$ and that summand must be $\mathbb{S}_t$, so $t \in T$. Conversely, if $t \in T$, then $\mathsf{T}(t)$ holds in $\mathbb{S}_t$, and then, as the extreme points in the unit disc of $\mathbb{S}_t$ are a subset of those of $\mathbb{S}_T + V$, $\mathsf{T}(t)$ must hold in $\mathbb{S}_T + V$.

Theorem 24. *There is a formula* $\mathsf{T}(x)$ *in* $\mathcal{L}_N$ *with the indicated scalar free variable such that*

(i) *in any normed space* $\mathsf{T}(x)$ *defines a subset of the interval* $(0,1)$ *and*
(ii) *for any set* $T \subseteq (0,1)$*, and any normed space* V *whose unit circle contains no hexagons,* $\mathsf{T}(x)$ *defines* T *in the* 1*-sum* $\mathbb{S}_T + V$.

Proof. Taking $\mathsf{T}(x)$ as defined above, we have already proved (ii), while (i) is immediate from the definition of $\mathsf{T}(x)$. □

Theorem 25. *There are primitive recursive reductions of the set of all true* Π_1^2 *sentences to each of the theories* NS_+, BS_+, NS_+^∞ *and* BS_+^∞.

Proof. It suffices to produce a primitive recursive function $\phi \mapsto \phi_A$ from the set of Σ^2_1 sentences to $\mathcal{L}^+_N$ such that (i) ϕ is true iff ϕ_A has a normed space model and (ii) whenever ϕ_A has a normed space model it also has an infinite-dimensional Banach space model. So let ϕ be a Σ^2_1 sentence $\exists u \cdot \psi(u)$.

We work in the 1-sum $\mathbb{X} + \mathbb{S}_T$ where $\mathbb{X}$ is the 2-dimensional normed space defined at the beginning of Section 5.1 and illustrated in Fig. 4 and $\mathbb{S}_T$ is as above for some $T \subseteq (0, 1)$. Consider the following formulas of $\mathcal{L}^+_N$:

$$\mathsf{EPX}(\mathbf{v}) := \mathsf{EP}(\mathbf{v}) \wedge \neg\exists \mathbf{u}\,\mathbf{w} \cdot \mathsf{H}(\mathbf{u}, \mathbf{v}, \mathbf{w})$$

$$\mathsf{X}(\mathbf{v}) := \exists \mathbf{u}\,\mathbf{w} \cdot \mathsf{EPX}(\mathbf{u}) \wedge \mathsf{EPX}(\mathbf{w}) \wedge \mathbf{v} = \mathbf{u} + \mathbf{w}.$$

In $\mathbb{X} + \mathbb{S}_T$, $\mathsf{EPX}(\mathbf{v})$ holds iff $\mathbf{v}$ is an extreme point of the disc of radius $\|\mathbf{v}\|$ in the summand $\mathbb{X}$ and so $\mathsf{X}(\mathbf{v})$ holds iff $\mathbf{v}$ is a sum of such extreme points, which is true iff $\mathbf{v} \in \mathbb{X}$. If, as in the proof of Lemma 12, we relativize the earlier definitions of the formulas $\mathsf{Nat}(x)$ and $\mathsf{RTIMES}(x, y, z)$ to $\mathsf{X}(\mathbf{v})$ then the resulting formulas will define the set of natural numbers and the graph of the multiplication function in $\mathbb{X} + \mathbb{S}_T$ just as in Theorem 17. As in Section 3, there are sentences Peano and Mult of $\mathcal{L}^+_N$ asserting that the relativized versions of $\mathsf{Nat}(x)$ and $\mathsf{RTIMES}(x, y, z)$ do indeed define the natural numbers and real multiplication respectively.

Let ψ_1 be obtained from $\psi(u)$ as follows: first, replace each subformula of the form $x^{\mathcal{P}} \in u$ by $\mathsf{T}(\frac{1}{3}(x^{\mathcal{R}} + 1))$ and translate all other formulas as in the reduction of second-order arithmetic of Theorem 2, using $\mathsf{D}(n, x)$ to represent sets of natural numbers as real numbers, using $\mathsf{S}(x)$ to single out canonical representatives and using the relativized $\mathsf{Nat}(x)$ as the predicate for the natural numbers; then, as in the proof of Theorem 3, eliminate multiplication using the relativized $\mathsf{RTIMES}(x, y, z)$. Now let $\phi_A := \psi_1 \wedge \mathsf{Peano} \wedge \mathsf{Mult}$. By construction ϕ_A contains no terms of the form $a\mathbf{v}$, so ϕ_A is indeed in $\mathcal{L}^+_N$.

We may now check conditions (i) and (ii). First, assume ϕ_A has a model, and in that model let $U = \{S \mid \mathsf{T}(\frac{1}{3}(\sharp S + 1))\}$ where $\sharp$ is the injection of $\mathbb{P}(\mathbb{N})$ into the interval $[0, 3/2]$ defined in Section 3. Then as Peano and Mult hold, $\psi(u)$ must hold in the standard model when u is interpreted as U, so ϕ, i.e., $\exists u \cdot \psi(u)$ is true. Conversely, if ϕ is true, so that $\psi(u)$ holds when u is interpreted as U say, then if we put $T = \{\frac{1}{3}(\sharp S + 1) \mid S \in U\} \cup (0, 1/3)$, ϕ_A is satisfied in the normed space $\mathbb{X} + \mathbb{S}_T$ (since if $\frac{1}{3}(x + 1) \in (0, 1/3)$ then $x < 0$ and $\mathsf{S}(x)$ is false). Now $\mathbb{X} + \mathbb{S}_T$ is a Banach space and is infinite-dimensional so if ϕ_A has a model it has an infinite-dimensional Banach space model. □

Corollary 26. *The theories* NS_+, BS_+, NS^∞_+, BS^∞_+, NS, BS, NS^∞ *and* BS^∞ *all have the same many-one degree as the set of all true* Π^2_1 *sentences.*

Proof. This follows immediately from Theorems 25 and 23. □

As a final remark on degrees of unsolvability, close analogues of the above results on normed spaces and Banach spaces hold for metric spaces: there is a Skolem–Löwenheim theorem stating that any (complete) metric space has an elementarily equivalent (complete) subspace of cardinality at most $\mathfrak{c}$; the theory of countable metric spaces is many-one equivalent to second-order arithmetic; and the theory of arbitrary metric spaces is many-one equivalent to the set of true Π^2_1 sentences. (For the analogue of the space $\mathbb{X} + \mathbb{S}_T$ in the proof of Theorem 25, choose $\mathbf{v} \in \mathbb{R}^2$ such that $d(\mathbf{v}, \mathbb{G}) \geqslant 2$ where $\mathbb{G}$ is the space of Theorem 5 and, for $T \subseteq (0, 1)$, let $\mathbb{H}_T := \{\mathbf{v}\} \cup \{\mathbf{u} \in \mathbb{R}^2 \mid d(\mathbf{u}, \mathbf{v}) - 1 \in T\}$. Then, in place of $\mathbb{X} + \mathbb{S}_T$, use $\mathbb{G} \cup \mathbb{H}_T$ and design the various formulas needed using the fact that $\mathbf{v}$ is the only isolated point.)

6. Quantifier elimination for theories of inner product spaces

The main idea of this section is that in the first-order theory of inner product spaces over $\mathbb{R}$ it should take at most k degrees of freedom to decide the validity of a formula with k vector variables. The key result implies that if a formula ϕ has free vector variables $\mathbf{v}_1, \ldots, \mathbf{v}_m$ and has k vector variables in all, then in all dimensions $\geqslant k$, ϕ is equivalent to a system of constraints on the inner products $\langle \mathbf{v}_i, \mathbf{v}_j \rangle$. The proof is via a process that eliminates vector quantifiers in favour of blocks of scalar quantifiers. It follows that to decide a sentence with k vector variables we need only decide it in $\mathbb{R}^n$ for $n = 0, 1, \ldots, k$ and that is easy after a simple syntactic transformation given a decision procedure for formulas that do not involve vectors, i.e., for the language of a real closed field.

In the paper that our title echoes, Tarski [42] gave the first quantifier elimination procedure for a real closed field and hence a decision procedure of the kind that we need. Apparently the first actual computer implementation of an algorithm for this problem was by Collins [12]. A relatively simple procedure due to Cohen and Hörmander [8,16,22] has been implemented by several people including one of the present authors. Collins's method of cylindrical algebraic decomposition has complexity exponential in the number of bound variables. The best known algorithms are exponential in the number of quantifier alternations (see Basu et al. [6]), but work on implementation of these algorithms is in its early stages. Since our syntactic transformations replace vector quantifiers by blocks of scalar quantifiers, these recent improvements are significant for the complexity of our decision procedure.

We write IP, resp., $\mathsf{IP}^{\mathbb{F}}$, resp., IP^∞ for the theories of real inner product spaces where the dimension is unconstrained, resp., constrained to be finite, resp., constrained to be infinite, and HS, $\mathsf{HS}^{\mathbb{F}}$ and HS^∞ for the theories of Hilbert spaces with the corresponding constraints on the dimension. By the well-known fact that finite dimensional inner product spaces are complete, $\mathsf{HS}^{\mathbb{F}} = \mathsf{IP}^{\mathbb{F}}$. We will show that all of these theories are decidable and that $\mathsf{IP} = \mathsf{IP}^{\mathbb{F}} = \mathsf{HS} = \mathsf{HS}^{\mathbb{F}}$ and that $\mathsf{IP}^{\infty} = \mathsf{HS}^{\infty}$.

Let us agree on some terminology and notation. Given a formula ϕ of $\mathcal{L}_I$, let $v(\phi)$ and $s(\phi)$ denote the sets of free vector variables and free scalar variables of ϕ respectively. If $\bar{\mathbf{v}} = (\mathbf{v}_1, \dots, \mathbf{v}_m)$ is a sequence of vector variables and $\bar{x} = (x_1, \dots, x_n)$ is a sequence of scalar variables, let us write $\phi(\bar{\mathbf{v}}, \bar{x})$ to indicate that $v(\phi) \subseteq \{\mathbf{v}_1, \dots, \mathbf{v}_m\}$ and $s(\phi) \subseteq \{x_1, \dots, x_n\}$. Let V be an inner product space. If ϕ is a sentence of $\mathcal{L}_I$, we write $V \models \phi$ to indicate that ϕ holds in V. More generally, if $\phi(\bar{\mathbf{v}}, \bar{x})$ is any formula in $\mathcal{L}_I$, and if $\bar{\mathbf{p}} \in V^m$ and $\bar{c} \in \mathbb{R}^n$, we write $V \models \phi(\bar{\mathbf{p}}, \bar{c})$ to indicate that ϕ holds in V if each $\mathbf{v}_i$ is interpreted as $\mathbf{p}_i$ and each x_j is interpreted as c_j. Note that if the formula ϕ contains no constants or variables of vector sort, then ϕ is a formula in the first-order language of an ordered field and, for any V, $V \models \phi(\emptyset, \bar{c})$ iff $\phi(\bar{c})$ holds in the ordered field $\mathbb{R}$.

For $k \in \mathbb{N}$, let us say that formulas $\phi_1(\bar{\mathbf{v}}, \bar{x})$ and $\phi_2(\bar{\mathbf{v}}, \bar{x})$ with the same free variables are *k-equivalent* iff for every inner product space V of dimension at least k, and every $\bar{\mathbf{p}} \in V^{\#(\bar{\mathbf{v}})}$ and every $\bar{c} \in \mathbb{R}^{\#(\bar{x})}$, $V \models \phi_1(\bar{\mathbf{p}}, \bar{c})$ iff $V \models \phi_2(\bar{\mathbf{p}}, \bar{c})$, i.e., ϕ_1 and ϕ_2 are equivalent in the theory of all spaces of dimension at least k. So, for example, the sentences $\exists \mathbf{v}\, \mathbf{w} \cdot \forall x \cdot \mathbf{v} \neq x\mathbf{w} \wedge \mathbf{w} \neq x\mathbf{v}$ and $\mathbf{0} = \mathbf{0}$ are 2-equivalent, but not 1-equivalent. Providing they have the same free variables, logically equivalent formulas are k-equivalent for any k.

If V is an inner product space and $\bar{\mathbf{p}} \in V^m$, recall that the *Gram matrix* of $\bar{\mathbf{p}}$ is the positive semidefinite symmetric $m \times m$ matrix $G = G(\bar{\mathbf{p}})$ with $G_{ij} = \langle \mathbf{p}_i, \mathbf{p}_j \rangle$. If $\bar{\mathbf{v}} = (\mathbf{v}_1, \dots, \mathbf{v}_m)$ is a sequence of vector variables, let us write $\overline{\mathsf{UG}}(\bar{\mathbf{v}})$ for the sequence of terms of $\mathcal{L}_I$ defined inductively by:

$$\overline{\mathsf{UG}}(\mathbf{v}_1) := \big(\langle \mathbf{v}_1, \mathbf{v}_1 \rangle\big)$$
$$\overline{\mathsf{UG}}(\mathbf{v}_1, \dots, \mathbf{v}_m) := \overline{\mathsf{UG}}(\mathbf{v}_1, \dots, \mathbf{v}_{m-1}) \frown \big(\langle \mathbf{v}_1, \mathbf{v}_m \rangle, \dots, \langle \mathbf{v}_m, \mathbf{v}_m \rangle\big)$$

where $\frown$ denotes concatenation. Thus $\overline{\mathsf{UG}}(\bar{\mathbf{v}})$ enumerates the upper triangle of the formal Gram matrix of $\bar{\mathbf{v}}$ by column. Let us say that a formula $\phi(\bar{\mathbf{v}}, \bar{x})$ is *special* iff it has the form $\psi(\overline{\mathsf{UG}}(\bar{\mathbf{v}}), \bar{x})$ where $\psi(w_1, \dots, w_{m(m+1)/2}, \bar{x})$ contains no variables or constants of vector sort, i.e., ψ is a formula in the language of an ordered field. Note that if ϕ and ψ are special, then so are $\neg\phi$, $\phi \circ \psi$ and $\exists x \cdot \phi$, where $\circ$ is any binary propositional connective, e.g., $\wedge$, $\vee$, $\Rightarrow$ or $\Leftrightarrow$.

Our main theorem will inductively transform a formula of $\mathcal{L}_I$ containing k vector variables into a k-equivalent special formula. The following two lemmas give the two main ingredients of the proof.

Lemma 27. *There is a primitive recursive function, $\phi \mapsto \phi^{\mathsf{R}}$, such that, for any formula ϕ of $\mathcal{L}_I$, ϕ^{R} is equivalent to ϕ in the theory of real inner product spaces and the only terms of vector sort in ϕ^{R} are variables occurring as operands of the inner product operator $\langle \_, \_ \rangle$. Moreover ϕ^{R} is quantifier-free if ϕ is.*

Proof. There is a primitive recursive p such that $p(\phi)$ results from ϕ by replacing each vector equation $\mathbf{a} = \mathbf{b}$ by the equivalent scalar equation $\langle \mathbf{a} - \mathbf{b}, \mathbf{a} - \mathbf{b} \rangle = 0$. I claim that there exists a primitive recursive q such that $q(\psi)$ results from ψ by repeatedly applying the following equations as left-to-right rewrite rules until no redexes remain.

$$\begin{aligned} \langle \mathbf{a}, \mathbf{0} \rangle &= 0 & \langle -\mathbf{a}, \mathbf{b} \rangle &= -\langle \mathbf{a}, \mathbf{b} \rangle \\ \langle \mathbf{0}, \mathbf{a} \rangle &= 0 & \langle \mathbf{a}, -\mathbf{b} \rangle &= -\langle \mathbf{a}, \mathbf{b} \rangle \\ \langle t\mathbf{a}, \mathbf{b} \rangle &= t\langle \mathbf{a}, \mathbf{b} \rangle & \langle \mathbf{a}, \mathbf{b} + \mathbf{c} \rangle &= \langle \mathbf{a}, \mathbf{b} \rangle + \langle \mathbf{a}, \mathbf{c} \rangle \\ \langle \mathbf{a}, t\mathbf{b} \rangle &= t\langle \mathbf{a}, \mathbf{b} \rangle & \langle \mathbf{a} + \mathbf{b}, \mathbf{c} \rangle &= \langle \mathbf{a}, \mathbf{c} \rangle + \langle \mathbf{b}, \mathbf{c} \rangle. \end{aligned}$$

Thus, if ψ contains no vector equations, $q(\psi)$ will contain no terms of vector sort other than variables occurring as operands of $\langle \_, \_ \rangle$. So given q, we may take $\phi^{\mathsf{R}} = q(p(\phi))$ to complete the proof. For the existence of q one can either apply a general result of Hofbauer [21] or use the following construction. Let the *weight* of a redex be the total number of constant and function symbols it contains. There is a primitive recursive f such that, if ψ contains a redex, then $f(\psi)$ results from ψ by applying one rule to a redex of maximal weight. Let $n(\psi)$ be the number of redexes of maximal weight in ψ and let $g(\psi) = f^{n(\psi)}(\psi)$. Now let $k(\psi)$ be 0 if ψ has no redexes and be the maximal weight of a redex in ψ otherwise and let $q(\psi) = g^{k(\psi)}(\psi)$. Then q is primitive recursive and $q(\psi)$ results from ψ by applying rewrite rules until no redexes remain. $\square$

Lemma 28. *Let M be a symmetric $m \times m$ matrix with real coefficients, let V be an inner product space of dimension at least m and let $\mathbf{p}_1, \dots, \mathbf{p}_{m-1} \in V$ be such that*

$$G(\mathbf{p}_1, \dots, \mathbf{p}_{m-1}) = (M_{ij})_{1 \leqslant i, j < m}$$

where $G(\mathbf{p}_1, \dots, \mathbf{p}_{m-1})$ is the Gram matrix of the $\mathbf{p}_i$. The following are equivalent:

(i) *there exists $\mathbf{p}_m \in V$ such that $M = G(\mathbf{p}_1, \dots, \mathbf{p}_{m-1}, \mathbf{p}_m)$;*
(ii) *there exist $b_1, \dots, b_m \in \mathbb{R}$ such that*

(a) $M_{im} = \sum_{j=1}^{m-1} b_j M_{ij}$ *for* $1 \leqslant i < m$,

(b) $M_{mm} = \sum_{i=1}^{m-1} \sum_{j=1}^{m-1} b_i b_j M_{ij} + b_m^2$.

Proof. For both parts of the proof, let W be the subspace spanned by the $\mathbf{p}_i$ with $1 \leqslant i < m$ and note that W is a proper subspace of V since $\dim(V) \geqslant m$.

(i) $\Rightarrow$ (ii): Given $\mathbf{p}_m \in V$ such that $G(\mathbf{p}_1, \ldots, \mathbf{p}_{m-1}, \mathbf{p}_m) = M$, there is a unit vector $\mathbf{c}$ orthogonal to W such that $\mathbf{p}_m$ lies in the subspace spanned by W and $\mathbf{c}$. Then we can write $\mathbf{p}_m = \sum_{i=1}^{m-1} b_i \mathbf{p}_i + b_m \mathbf{c}$ for some $b_i \in \mathbb{R}$. (a) and (b) follow for this choice of the b_i using the expression for $\mathbf{p}_m$ to expand the inner products $M_{im} = \langle \mathbf{p}_i, \mathbf{p}_j \rangle$, $1 \leqslant i \leqslant m$.

(i) $\Leftarrow$ (ii): Given $b_1, \ldots, b_m \in \mathbb{R}$ satisfying (a) and (b), choose a unit vector $\mathbf{c}$ orthogonal to W and let $\mathbf{p}_m = \sum_{i=1}^{m-1} b_i \mathbf{p}_i + b_m \mathbf{c}$. We must show that the equation $M_{ij} = \langle \mathbf{p}_i, \mathbf{p}_j \rangle$ holds for $1 \leqslant i, j \leqslant m$. But this is so by assumption when $1 \leqslant i, j < m$, by (a) when $1 \leqslant i < j = m$, by symmetry and (a) when $1 \leqslant j < i = m$ and by (b) when $i = j = m$. $\square$

We now give the main theorem of this section. In this theorem, we need to count the number of vector variables in a formula. This is to be done in a very frugal way, by ignoring variable binding and simply counting the number of distinct variable names that appear labelled with the vector sort: so that, for example, $(\forall \mathbf{v}\, \mathbf{w} \cdot \mathbf{v} + \mathbf{w} = \mathbf{0}) \Rightarrow (\forall \mathbf{v} \cdot \mathbf{v} = \mathbf{w})$ contains just two variables, $\mathbf{v}$ and $\mathbf{w}$.

Theorem 29. *There is a primitive recursive function, $\phi \mapsto \phi^{\mathsf{S}}$, such that, for any formula $\phi \in \mathcal{L}_I$ containing k vector variables counted in the sense described above, $\phi^{\mathsf{S}} \in \mathcal{L}_I$ is a special formula that is k-equivalent to ϕ.*

Proof. We will show by induction that every formula ϕ with k vector variables is k-equivalent to a special formula and it will be clear from the proof that a suitable special formula can be calculated as a primitive recursive function of ϕ.

We may replace $\forall \ldots \cdot \ldots$ by $\neg\exists \ldots \cdot \neg \ldots$ throughout, so the cases we have to consider are: (i) quantifier-free formulas (and hence in particular atomic formulas), (ii) logical negation, (iii) scalar existential quantification, (iv) vector existential quantification and (v) the binary propositional connectives.

(i) If ϕ is quantifier-free with free variables $\mathbf{v}_1, \ldots, \mathbf{v}_k$, Lemma 27 provides a quantifier-free formula ϕ^{R} that is 0-equivalent to ϕ and in which vector terms only occur in terms of the form $\langle \mathbf{v}_i, \mathbf{v}_j \rangle$. Replacing each occurrence of $\langle \mathbf{v}_i, \mathbf{v}_j \rangle$ in ϕ^{R} where $i > j$ by $\langle \mathbf{v}_j, \mathbf{v}_i \rangle$ then gives a special formula that is 0-equivalent and hence k-equivalent to ϕ for any k.

For steps (ii) to (iv), assume that ϕ is k-equivalent to a special formula, σ with the same free variables.

(ii) Like ϕ, $\neg\phi$ contains k vector variables and is k-equivalent to the special formula $\neg\sigma$.

(iii) Again like ϕ, $\exists x \cdot \phi$ contains k vector variables and is k-equivalent to the special formula $\exists x \cdot \sigma$.

(iv) If $\mathbf{v}$ does not appear free in ϕ, then $\exists \mathbf{v} \cdot \phi$ contains either k or $k+1$ vector variables and is logically equivalent to ϕ, and hence k-equivalent and so also $(k+1)$-equivalent to the special formula σ. If $\mathbf{v}$ does appear free in ϕ, then ϕ and $\exists \mathbf{v} \cdot \phi$ both contain k vector variables. Let $m = |v(\phi)|$ and $n = |s(\phi)|$. Let $\bar{\mathbf{v}} = (\mathbf{v}_1, \ldots, \mathbf{v}_m)$ enumerate $v(\phi) = v(\sigma)$ so that $\mathbf{v} \equiv \mathbf{v}_m$ and let $\bar{z} = (z_1, \ldots, z_n)$ enumerate $s(\phi) = s(\sigma)$. Since σ is special it has the form $\psi(\overline{\mathsf{UG}}(\bar{\mathbf{v}}), \bar{z})$, where $\psi(w_1, w_2, \ldots, w_{m(m+1)/2}, \bar{z})$ is a formula in the language of an ordered field. Let $\chi := \psi[\langle \mathbf{v}_1, \mathbf{v}_1 \rangle / w_1, \ldots, \langle \mathbf{v}_{m-1}, \mathbf{v}_{m-1} \rangle / w_{(m-1)m/2}]$ be the result of substituting the terms of the sequence $\overline{\mathsf{UG}}(\mathbf{v}_1, \ldots, \mathbf{v}_{m-1})$ for $w_1, \ldots, w_{(m-1)m/2}$ in ψ. I claim that $\exists \mathbf{v} \cdot \phi \equiv \exists \mathbf{v}_m \cdot \phi$ is k-equivalent to the special formula σ_1 defined as follows where the x_i and the y_i are fresh variables:

$$
\begin{aligned}
\sigma_1 \;:=\; & \exists x_1 \cdots x_m y_1 \cdots y_m \cdot \\
& \bigwedge_{i=1}^{m-1} x_i = \sum_{j=1}^{m-1} y_j \langle \mathbf{v}_i, \mathbf{v}_j \rangle \\
& \wedge x_m = \sum_{i=1}^{m-1} \sum_{j=1}^{m-1} y_i y_j \langle \mathbf{v}_i, \mathbf{v}_j \rangle + y_m^2 \\
& \wedge \chi[x_1 / w_{(m-1)m/2+1}, \ldots, x_m / w_{m(m+1)/2}].
\end{aligned}
$$

To see that $\exists \mathbf{v}_m \cdot \phi$ is indeed k-equivalent to σ_1, let V be an inner product space of dimension at least k, let $\mathbf{p}_1, \ldots, \mathbf{p}_{m-1} \in V$ and let $\bar{c} \in \mathbb{R}^n$. We have to show that $V \models (\exists \mathbf{v}_m \cdot \phi)(\mathbf{p}_1, \ldots, \mathbf{p}_{m-1}, \bar{c})$ iff $V \models \sigma_1(\mathbf{p}_1, \ldots, \mathbf{p}_{m-1}, \bar{c})$:

$\Rightarrow$: Assume $V \models (\exists \mathbf{v}_m \cdot \phi)(\mathbf{p}_1, \ldots, \mathbf{p}_{m-1}, \bar{c})$, so there is $\mathbf{p}_m \in V$, such that $V \models \phi(\mathbf{p}_1, \ldots, \mathbf{p}_m, \bar{c})$. Since ϕ and σ are k-equivalent, $V \models \sigma(\mathbf{p}_1, \ldots, \mathbf{p}_m, \bar{c})$, i.e., $V \models \psi(\overline{\mathsf{UG}}(\mathbf{p}_1, \ldots, \mathbf{p}_m), \bar{c})$. Applying Lemma 28 to the $\mathbf{p}_i$ and the Gram matrix $M = G(\mathbf{p}_1, \ldots, \mathbf{p}_m)$, we obtain $b_1, \ldots, b_m \in \mathbb{R}$ satisfying equations (a) and (b) of the lemma, so that if we interpret x_i as M_{im} and y_i as b_i, $1 \leqslant i \leqslant m$, the matrix of σ_1 holds, so that $V \models \sigma_1(\mathbf{p}_1, \ldots, \mathbf{p}_{m-1}, \bar{c})$ as required.

$\Leftarrow$: Assume $V \models \sigma_1(\mathbf{p}_1, \ldots, \mathbf{p}_{m-1}, \bar{c})$, so that there are $a_i, b_i \in \mathbb{R}$, $1 \leqslant i \leqslant m$, such that the matrix of σ_1 holds if we interpret the x_i as the a_i, the y_i as the b_i and $\mathbf{v}_1, \ldots, \mathbf{v}_{m-1}$ as $\mathbf{p}_1, \ldots, \mathbf{p}_{m-1}$. Let M be the $m \times m$ matrix with $M_{ij} = \langle \mathbf{p}_i, \mathbf{p}_j \rangle$, $1 \leqslant i, j < m$, and $M_{im} = M_{mi} = b_i$, $1 \leqslant i \leqslant m$. Then the assumptions of Lemma 28 hold for M as do equations (a) and (b) of the lemma, which thus gives us $\mathbf{p}_m \in V$ such that $M_{ij} = \langle \mathbf{p}_i, \mathbf{p}_j \rangle$, $1 \leqslant i, j \leqslant m$. The final conjunct in the matrix of σ_1 then implies that $V \models \psi(\overline{\mathsf{UG}}(\mathbf{p}_1, \ldots, \mathbf{p}_m), \bar{c})$, i.e. $V \models \sigma(\mathbf{p}_1, \ldots, \mathbf{p}_m, \bar{c})$. As ϕ and σ are k-equivalent, we must have $V \models \phi(\mathbf{p}_1, \ldots, \mathbf{p}_m, \bar{c})$, so that $V \models (\exists \mathbf{v}_m \cdot \phi)(\mathbf{p}_1, \ldots, \mathbf{p}_{m-1}, \bar{c})$ as required.

(v) At this point, the proof for formulas in prenex normal form would be complete. However, putting a formula into prenex normal form can cause an exponential explosion in the number of variables that it contains when counted in our frugal sense and this would make algorithms based on our results less efficient. So for the final step, assume that ϕ contains k vector variables and is k-equivalent to the special formula σ, while ϕ' contains k' vector variables and is k'-equivalent to the special formula σ'. Let $\circ$ be any binary propositional connective. It is easy to see that, $\phi \circ \phi'$ is $\max\{k, k'\}$-equivalent to $\sigma \circ \sigma'$. But if k'' is the number of vector variables in $\phi \circ \phi'$, we must have $k'' \geqslant \max\{k, k'\}$, so $\phi \circ \phi'$ is also k''-equivalent to the special formula $\sigma \circ \sigma'$. □

The only special feature of the field of real numbers used in the above proof is that it is *euclidean*, i.e., all positive elements have square roots (which is needed to ensure the existence of a unit vector in any given direction). Over a non-euclidean field, both the proof and the statement of the theorem break down: over the field of rational numbers, there is a countable infinity of distinct isomorphism classes of 1-dimensional inner product spaces indexed by square-free positive integers, the class corresponding to m being characterized by the sentence $\exists \mathbf{v} \cdot \langle \mathbf{v}, \mathbf{v} \rangle = m$.

The theorem immediately gives us an effective quantifier elimination procedure in the infinite-dimensional case:

Corollary 30. *There is a primitive recursive function, $\phi \mapsto \phi^{\mathsf{QE}}$, such that if $\phi \in \mathcal{L}_I$, $\phi^{\mathsf{QE}} \in \mathcal{L}_I$ is a quantifier-free formula that is equivalent to ϕ modulo either of the theories IP^∞ and HS^∞.*

Proof. First calculate the special formula ϕ^{S} given by the theorem; ϕ^{S} will be equivalent to ϕ in any infinite-dimensional inner product space and has the form $\psi(\overline{\mathsf{UG}}(\mathbf{v}_1, \ldots, \mathbf{v}_m), \bar{x})$ where the $\mathbf{v}_i$ are the free vector variables of ϕ and ψ is a formula in the language of an ordered field. Now apply quantifier elimination for real closed fields to ψ, giving an equivalent quantifier-free formula, χ, say, and put $\phi^{\mathsf{QE}} = \chi(\overline{\mathsf{UG}}(\mathbf{v}_1, \ldots, \mathbf{v}_m), \bar{x})$. □

It follows that IP^∞ are HS^∞ are both decidable and actually coincide. However, as we are also interested in the decision problem for IP, $\mathsf{IP}^{\mathbb{F}}$, HS and $\mathsf{HS}^{\mathbb{F}}$, we will take a different line, using the following theorem to justify an alternative approach.

Theorem 31. *Let ϕ be a sentence of $\mathcal{L}_I$ containing k vector variables and let V be any inner product space of* (*possibly infinite*) *dimension $d \geqslant k$. Then ϕ holds in V iff it holds in $\mathbb{R}^k$.*

Proof. By Theorem 29, ϕ is k-equivalent to a special formula, but a special formula with no free variables is just a sentence in the language of an ordered field and its truth is independent of the choice of vector space, so any space of dimension at least k, e.g., $\mathbb{R}^k$, will serve to test the truth of ϕ. □

Lemma 32. *There is a primitive recursive function that maps a sentence ϕ of $\mathcal{L}_I$ and a natural number n to a sentence $\phi|_n$ in the language of an ordered field such that $\mathbb{R}^n \models \phi \Leftrightarrow \phi|_n$, i.e., ϕ holds in $\mathbb{R}^n$ iff $\phi|_n$ holds in the ordered field $\mathbb{R}$.*

Proof. We describe a primitive recursive algorithm that constructs the sentence $\phi|_n$ and show that it holds in $\mathbb{R}$ iff ϕ holds in the standard n-dimensional inner product space $\mathbb{R}^n$, which proves the lemma.

If $n = 0$, $\phi|_0$ is obtained from ϕ by deleting all vector quantifiers, replacing all inner products by scalar 0 and replacing all vector equations by the scalar equation $0 = 0$. Evidently $\mathbb{R}^0 \models \phi$ iff $\mathbb{R}^0 \models \phi|_0$.

If $n \geqslant 1$, pick n fresh vector variables $\mathbf{b}_1, \ldots, \mathbf{b}_n$ and, for each vector variable $\mathbf{v}$ occurring in ϕ, pick n fresh scalar variables $x_1^{\mathbf{v}}, \ldots, x_n^{\mathbf{v}}$. Replace each vector quantifier $\forall \mathbf{v} \cdot$ (resp. $\exists \mathbf{v} \cdot$) in ϕ by the string of scalar quantifiers $\forall x_1^{\mathbf{v}} \cdots x_n^{\mathbf{v}} \cdot$ (resp. $\exists x_1^{\mathbf{v}} \cdots x_n^{\mathbf{v}} \cdot$) and replace all other occurrences of $\mathbf{v}$ by $x_1^{\mathbf{v}}\mathbf{b}_1 + \cdots + x_n^{\mathbf{v}}\mathbf{b}_n$. Let the resulting formula be $\phi_1(\mathbf{b}_1, \ldots, \mathbf{b}_n)$. Clearly, $\mathbb{R}^n \models \phi$ iff $\mathbb{R}^n \models_{\mathbb{R}^n} \phi_1(\bar{\mathbf{e}})$ where $\bar{\mathbf{e}} = (\mathbf{e}_1, \ldots, \mathbf{e}_n)$ is the standard basis for $\mathbb{R}^n$. By Lemma 27, ϕ_1 is equivalent to a special formula $\phi_2(\overline{\mathsf{UG}}(\mathbf{b}_1, \ldots, \mathbf{b}_n))$ where $\phi_2(w_1, \ldots, w_{n(n+1)/2})$ is a formula in the language of an ordered field. Writing the Kronecker symbol δ_{ij} to stand for the constant 1 when $i = j$ and the constant 0 otherwise, define $\phi|_n := \phi_2[\delta_{ij}/w_{n(i-1)+j} \mid 1 \leqslant i \leqslant j \leqslant n]$. We then have $\mathbb{R}^n \models \phi$ iff $\mathbb{R}^n \models \phi_1(\bar{\mathbf{e}})$ iff $\mathbb{R}^n \models \phi_2(\overline{\mathsf{UG}}(\bar{\mathbf{e}}))$ iff $\phi|_n$ holds in $\mathbb{R}$. □

In the construction of $\phi|_n$ in the above proof, an alternative way of eliminating vector variables from the formula ϕ_1 is to rearrange vector equations into the form $t_1\mathbf{b}_1 + \cdots + t_n\mathbf{b}_n = \mathbf{0}$ which may then be replaced by $t_1 = t_2 = \cdots = t_n = 0$ before applying the method of Lemma 27 to eliminate inner products. This is more efficient and also avoids introducing multiplication, which might be practically beneficial when working in the additive fragment of an extended language including a richer supply of vector constants.

In Section 2.5.2 we defined sentences $\mathsf{D}_{\leqslant n}$ for $n \in \mathbb{N}$ that hold in a vector space iff the space has finite dimension less than or equal to n. Let us define $\mathsf{D}_0 := \mathsf{D}_{\leqslant 0}$ and $\mathsf{D}_{n+1} := \mathsf{D}_{\leqslant n+1} \wedge \neg \mathsf{D}_{\leqslant n}$ so that the sentence D_n holds iff the dimension is exactly n. We use these sentences to reduce the theories of interest to the theory $\mathsf{IP}^{\leqslant k}$ of inner product spaces of dimension at most k.

Theorem 33. *Let ϕ be a sentence of $\mathcal{L}_I$ containing k vector variables; if $k = 0$, let $\phi^* := \phi|_0$, otherwise define ϕ^* by*

$$\phi^* := (\mathsf{D}_0 \wedge \phi|_0) \vee (\mathsf{D}_1 \wedge \phi|_1) \vee \cdots \vee (\mathsf{D}_{k-1} \wedge \phi|_{k-1}) \vee (\neg \mathsf{D}_{\leqslant (k-1)} \wedge \phi|_k).$$

Then ϕ^* is equivalent to ϕ in any of the theories IP, $\mathsf{IP}^{\mathbb{F}}$, IP^{∞}, HS, $\mathsf{HS}^{\mathbb{F}}$, *and* HS^{∞}.

Proof. Let V be any inner product space. If V has infinite dimension or finite dimension $d \geqslant k$, then D_n is false in V for $n \leqslant k-1$ and $\neg\mathsf{D}_{\leqslant k-1}$ is true, so ϕ^* is equivalent in V to $\phi|_k$. But, by Lemma 32, $\phi|_k$ is true iff ϕ is true in $\mathbb{R}^k$, and by Theorem 31, ϕ is true in V iff it is true in $\mathbb{R}^k$. If V has finite dimension $d < k$, then ϕ^* is equivalent to $\phi|_d$ which is valid iff ϕ holds in $\mathbb{R}^d$ iff ϕ holds in V, since V and $\mathbb{R}^d$ are isomorphic. So irrespective of the dimension of V, ϕ holds iff ϕ^* holds. Noting that our methods of proof make no assumptions about completeness this completes the proof of the theorem. □

Corollary 34. *For every sentence ϕ of $\mathcal{L}_I$ there is a subset D_ϕ of $\mathbb{N} \cup \{\infty\}$ such that ϕ holds in an inner product space V iff $\mathsf{dim}(V) \in D_\phi$. Moreover D_ϕ is either a finite subset of $\mathbb{N}$ or the complement of a finite subset of $\mathbb{N}$ and can be effectively computed from ϕ.*

Proof. First, calculate ϕ^* as in the theorem and then apply the quantifier elimination algorithm for the first-order theory of real arithmetic to determine the truth values of the sentences $\phi|_i$ that appear in ϕ^*. Now simplify to give either (i) a (possibly empty) disjunction of the form $\mathsf{D}_{i_1} \vee \cdots \vee \mathsf{D}_{i_m}$ or (ii) a disjunction of the form $\mathsf{D}_{i_1} \vee \cdots \vee \mathsf{D}_{i_m} \vee \neg\mathsf{D}_{\leqslant(k-1)}$ (where $k > i_m$ is the number of vector variables in ϕ). In both cases, the truth of the result is determined by a set D_ϕ of dimensions: in case (i), we have $D_\phi = \{i_1, \ldots, i_m\}$ which is a finite subset of $\mathbb{N}$, while in case (ii) D_ϕ is the complement in $\mathbb{N} \cup \{\infty\}$ of the finite subset $\{0, \ldots, k-1\} \backslash \{i_1, \ldots, i_m\}$. Let us represent D_ϕ as a pair (t, X), where $t \in \{0, 1\}$ and X is a finite set of natural numbers, D_ϕ being given by X when $t = 0$ and its complement when $t = 1$. Since the construction of ϕ^* and the $\phi|_i$ is primitive recursive, we have an effective procedure for computing the representation of D_ϕ. □

Corollary 35. *A class $\mathcal{C}$ of structures for the language $\mathcal{L}_I$ is axiomatizable (resp. recursively axiomatizable) iff it comprises all inner product spaces V such that $\mathsf{dim}(V) \in D$ for some $D \subseteq \mathbb{N} \cup \{\infty\}$ that is either finite or contains ∞ (resp. either finite or the complement of a recursively enumerable subset of $\mathbb{N}$).*

Proof. Recall that a class of structures for a language is said to be (recursively) axiomatizable iff it comprises all models of some (recursive) set of axioms. If A is any set of sentences of $\mathcal{L}_I$, then, by the previous corollary, V is a model of A iff $\mathsf{dim}(V) \in \bigcap_{\phi \in A} D_\phi$ where each D_ϕ is either a finite set of natural numbers or the complement in $\mathbb{N} \cup \{\infty\}$ of a finite set of natural numbers. A subset D of $\mathbb{N} \cup \{\infty\}$ can be written as such an intersection iff it is either a finite set of natural numbers or contains ∞.

The assertion about recursive axiomatizability is an easy exercise in recursion theory: in one direction, test for non-membership of D using an algorithm that on input d, enumerates the sentences of A checking for each sentence in turn whether it excludes models of dimension d; in the other direction, observe that a non-empty r.e. set of finite dimensions may be excluded by an r.e. set of axioms and then use the well-known trick of replacing the r.e. set $\phi_1, \phi_2, \phi_3, \ldots$ by the recursive set $\phi_1, \phi_1 \wedge \phi_2, \phi_1 \wedge \phi_2 \wedge \phi_3, \ldots$ to get a recursive axiomatization. □

Theorem 36. *The theories* IP, $\mathsf{IP}^{\mathbb{F}}$, IP^{∞}, HS, $\mathsf{HS}^{\mathbb{F}}$ *and* HS^{∞} *are all decidable. Moreover* $\mathsf{IP} = \mathsf{IP}^{\mathbb{F}} = \mathsf{HS} = \mathsf{HS}^{\mathbb{F}}$ *and* $\mathsf{IP}^{\infty} = \mathsf{HS}^{\infty}$.

Proof. By Corollary 34, given a sentence ϕ of $\mathcal{L}_I$, we can effectively calculate the set $D_\phi \subseteq \mathbb{N} \cup \{\infty\}$ of dimensions in which ϕ holds and for some finite $X \subseteq \mathbb{N}$, either $D_\phi = X$ or $D_\phi = (\mathbb{N} \cup \{\infty\}) \setminus X$. If $D_\phi = X$, then ϕ does not belong to any of the theories listed. If $D_\phi = (\mathbb{N} \cup \{\infty\}) \setminus X$, then ϕ certainly belongs to both IP^{∞} and HS^{∞}, while ϕ belongs to IP, $\mathsf{IP}^{\mathbb{F}}$, HS or $\mathsf{HS}^{\mathbb{F}}$ iff X is empty. Thus we have an effective procedure for deciding membership for each of the theories. Since the theories IP^{∞} and HS^{∞} have a common decision procedure they are equal and similarly IP, $\mathsf{IP}^{\mathbb{F}}$, HS and $\mathsf{HS}^{\mathbb{F}}$ are all equal. □

For $d \in \mathbb{N}$ there is exactly one inner product space of dimension d up to isomorphism. Corollary 34 implies that there is exactly one infinite-dimensional inner product space up to elementary equivalence. By contrast, it can be shown that, up to elementary equivalence, there are $\mathsf{c} = |\mathbb{R}|$ distinct d-dimensional normed spaces for each d, $2 \leqslant d \in \mathbb{N} \cup \{\infty\}$.

7. Decidable fragments of the theory of normed spaces

Although we have shown that the general theory of normed spaces is undecidable, there are some significant decidable fragments. In this section, we will find that the purely universal and purely existential fragments are both decidable via reductions to the first-order theory of the real numbers. The reduction for purely existential sentences is very simple, but for purely universal sentences, the reduction involves an interesting geometrical construction. In Section 8 we will find that the $\exists\forall$ and $\forall\exists$ fragments are undecidable, so these results are the best possible of their type.

Consider a sentence in the language of normed spaces that is in prenex normal form and contains no universal quantified vector variables: clearly such a sentence ϕ holds in all normed spaces iff it holds in the trivial normed space 0. We therefore obtain a decision procedure for valid sentences of this form by striking out all vector quantifiers, replacing all norm expressions by 0 and all vector equations by $0 = 0$ and then applying the decision procedure for the first-order theory of the real numbers. In particular, the set of valid purely existential sentences is decidable.

As we shall now see the set of true purely universal sentences in the language of normed spaces is also decidable, but the decision procedure and its verification are much less trivial: the crux of the argument lies in deciding satisfiability of a set of bounds on the norms of a finite set of vectors, so we start by considering how to define a norm satisfying a system of constraints.

A subset of X of a vector space V is said to be *symmetric* if $X = -X$ where $-X = \{-\mathbf{v} \mid \mathbf{v} \in X\}$. Given a subset Y of V we define the *symmetric convex hull* of Y, written $\mathsf{sconv}(Y)$, to be the intersection of the set of all symmetric convex sets containing Y. $\mathsf{sconv}(Y)$ is itself symmetric and convex and it is easy to verify that $\mathsf{sconv}(Y)$ is the convex hull of $Y \cup -Y$. If $\mathbf{v} \in \mathsf{sconv}(Y)$, then, by symmetry, $-\mathbf{v} \in \mathsf{sconv}(Y)$ and then, by convexity, the line segment $[-\mathbf{v}, \mathbf{v}]$ is contained in $\mathsf{sconv}(Y)$, i.e., $c\mathbf{v} \in \mathsf{sconv}(Y)$ for any c with $|c| \leqslant 1$.

Lemma 37. *Let $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ be a non-empty finite subset of a vector space. Then the symmetric convex hull of X is given by*:

$$\mathsf{sconv}(X) = \left\{ \sum_{i=1}^{n} c_i \mathbf{x}_i \,\middle|\, \sum_{i=1}^{n} |c_i| \leqslant 1 \right\}.$$

Proof. Write $D = \{\sum_{i=1}^n c_i \mathbf{x}_i \mid \sum_{i=1}^n |c_i| \leqslant 1\}$. It is easy to check that D is convex, symmetric and contains X, so $\mathsf{sconv}(X) \subseteq D$. Conversely, let $\mathbf{v} \in D$, so $\mathbf{v} = \sum_{i=1}^n c_i \mathbf{x}_i$ for some c_i where $c = \sum_{i=1}^n |c_i| \leqslant 1$. If $c = 1$, then $\mathbf{v}$ is a convex combination of the points $\pm\mathbf{x}_i$ and $\mathbf{v} \in \mathsf{sconv}(X)$ by the remarks above. If $c = 0$, then trivially $\mathbf{v} = \mathbf{0} \in \mathsf{sconv}(X)$. So assume $0 < c < 1$, so that $\mathbf{v} = c\sum_{i=1}^n (c_i/c)\mathbf{x}_i$ and we have $\sum_{i=1}^n |c_i/c| = 1$. Hence $\mathbf{v}$ can be written as $c\mathbf{w}$ where $|c| \leqslant 1$ and $\mathbf{w} \in \mathsf{sconv}(X)$ (by the case $c = 1$ just considered) so by the remarks above $\mathbf{v} \in \mathsf{sconv}(X)$. □

Lemma 38. *Let $Y = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ be a non-empty finite subset of a vector space V and let $D = \mathsf{sconv}(Y)$ be its symmetric convex hull. Then* (i) *D is the unit disc of a norm on the subspace W of V spanned by Y and* (ii) *if S is the unit circle for this norm, then*

$$D \setminus S = \left\{ \sum_{i=1}^{n} c_i \mathbf{x}_i \,\middle|\, \sum_{i=1}^{n} |c_i| < 1 \right\}.$$

Proof. For (i), as D is certainly convex, it will satisfy the criteria for the unit disc of a norm on W if it meets every line through the origin in W in a line segment $[-\mathbf{v}, \mathbf{v}]$ with $\mathbf{v} \neq 0$. By the Minkowski–Weyl theorem, D, which is the convex hull of a finite set of points, can be written as the intersection of a finite set of closed halfspaces. Hence if l is any line through the origin in W, $l \cap D$ is the intersection of l and a finite set of closed half-lines, and hence, as it is non-empty, bounded and symmetric about the origin, it must be the line segment, $[-\mathbf{v}, \mathbf{v}]$ for some $\mathbf{v}$. We have only to show that $\mathbf{v} \neq \mathbf{0}$. To see this let $\mathbf{w}$ be any point of $l \setminus \{\mathbf{0}\}$. Since $\mathbf{w} \in W$, there are c_i such that $\mathbf{w} = \sum_{i=1}^n c_i \mathbf{x}_i$. If we let $c = \sum_{i=1}^n |c_i|$, then $c \neq 0$, and by Lemma 37, $\mathbf{w}/c \in D$, but then $\mathbf{w}/c \in D \cap l = [-\mathbf{v}, \mathbf{v}]$ and as $\mathbf{w}/c \neq 0$ we must have $\mathbf{v} \neq 0$.

For (ii), note that $\mathbf{v} \in D \setminus S$ iff there is a $d > 1$ such that $d\mathbf{v} \in D$. By Lemma 37, $d\mathbf{v} \in D$ iff $d\mathbf{v}$ can be written as $\sum_{i=1}^n d_i \mathbf{x}_i$ with $\sum_{i=1}^n |d_i| \leqslant 1$ and this holds for $d > 1$ iff $\mathbf{v}$ can be written as $\sum_{i=1}^n c_i \mathbf{x}_i$ with $\sum_{i=1}^n |c_i| < 1$. □

Lemma 39. *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ and $\mathbf{y}_1, \ldots, \mathbf{y}_m$ be vectors in a vector space V. Then there exists a norm $\|\_\|$ on V such that $\|\mathbf{x}_i\| \leqslant 1$ for all i, $1 \leqslant i \leqslant n$, and $\|\mathbf{y}_j\| \geqslant 1$ for all j, $1 \leqslant j \leqslant m$, iff no $\mathbf{y}_k$ is expressible as $\mathbf{y}_k = \sum_{i=1}^n c_i \mathbf{x}_i$ with $\sum_{i=1}^n |c_i| < 1$.*

Proof. If a norm satisfies the stated bounds, then it is indeed impossible that any $\mathbf{y}_k = \sum_{i=1}^n c_i \mathbf{x}_i$ with $\sum_{i=1}^n |c_i| < 1$, for then by the triangle inequality $\|\mathbf{y}_k\| \leqslant \sum_{i=1}^n \|c_i \mathbf{x}_i\| = \sum_{i=1}^n |c_i| \|\mathbf{x}_i\| \leqslant \sum_{i=1}^n |c_i| < 1$, contradicting $\|\mathbf{y}_k\| \geqslant 1$.

Conversely, suppose no $\mathbf{y}_k$ is expressible as $\mathbf{y}_k = \sum_{i=1}^n c_i \mathbf{x}_i$ with $\sum_{i=1}^n |c_i| < 1$. By Lemma 38, we can define a norm $\|\_\|_0$ on the span V_0 of $\mathbf{x}_1, \ldots, \mathbf{x}_n$ with $D = \mathsf{sconv}(\{\mathbf{x}_1, \ldots, \mathbf{x}_n\})$ as its unit disc. Let V_1 be a complementary subspace of V_0, so that every $\mathbf{v} \in V$ is uniquely expressible as $\mathbf{v} = \mathbf{v}_0 + \mathbf{v}_1$ for $\mathbf{v}_0 \in V_0$ and $\mathbf{v}_1 \in V_1$. Let $\|\_\|_1$ be an arbitrary norm on V_1, e.g. defined using an inner product w.r.t. some basis. For any $B > 0$, the norm $\|\mathbf{v}_1\|_B = B\|\mathbf{v}_1\|_1$ is also a norm on V_1, and $\|\mathbf{v}\| = \|\mathbf{v}_0\|_0 + \|\mathbf{v}_1\|_B$ is a norm on V. I claim that for sufficiently large B, this satisfies the constraints $\|\mathbf{y}_j\| \geqslant 1$. First, if $\mathbf{y}_j \in V_0$, then this follows immediately since the assumption implies, by Lemma 38, that $\mathbf{y}_j$ is not in $\{\mathbf{w} \mid \|\mathbf{w}\|_0 < 1\}$. On the other hand, all the $\mathbf{y}_j \notin V_0$ can be written $\mathbf{y}_j = \mathbf{w}_j + \mathbf{z}_j$ for $\mathbf{w}_j \in V_0$, $\mathbf{z}_j \in V_1$ with $\mathbf{z}_j$ nonzero. To ensure $\|\mathbf{y}_j\| = \|\mathbf{w}_j\|_0 + B\|\mathbf{z}_j\|_1 \geqslant 1$, it suffices to choose $B > \max\{1/\|\mathbf{z}_1\|_1, \ldots, 1/\|\mathbf{z}_m\|_1\}$. □

Theorem 40. *Let $\mathbf{x}_1, \ldots, \mathbf{x}_n$ and $\mathbf{y}_1, \ldots, \mathbf{y}_m$ be vectors in a vector space V, and let $b_1, \ldots, b_n$ and $d_1, \ldots, d_m$ be real numbers such that $b_i \neq 0$ for some i, $1 \leqslant i \leqslant n$. Then there exists a norm $\|\_\|$ on V such that $\|\mathbf{x}_i\| \leqslant b_i$ for all i, $1 \leqslant i \leqslant n$, and $\|\mathbf{y}_j\| \geqslant d_j$ for all j, $1 \leqslant j \leqslant m$, iff the following conditions hold*:

- *For all* $1 \leqslant i \leqslant n$, $b_i \geqslant 0$;
- *For all* $1 \leqslant i \leqslant n$, *if* $b_i = 0$ *then* $\mathbf{x}_i = 0$;
- *No* $\mathbf{y}_j$ *is expressible as* $\mathbf{y}_j = \sum_{i=1}^n c_i \mathbf{x}_i$ *with* $\sum_{i=1}^n |c_i| b_i < d_j$.

Proof. If a norm satisfying the claimed inequalities exists, then all three properties follow immediately from the norm properties, the last one using the triangle inequality just as in the proof of Lemma 39.

Conversely, suppose the three properties hold. In order to construct a norm satisfying the inequalities, we can assume without loss of generality that all $b_i > 0$, because by the second property, if $b_i = 0$ then $\mathbf{x}_i = 0$ and so any norm at all satisfies $\|\mathbf{x}_i\| \leqslant b_i$. Similarly, we can assume that each $d_j > 0$ because if $d_j \leqslant 0$ the constraint $\|\mathbf{y}_j\| \geqslant d_j$ is automatically satisfied.

Define $\mathbf{u}_i = \mathbf{x}_i / b_i$ and $\mathbf{v}_j = \mathbf{y}_j / d_j$. Note that no $\mathbf{v}_j$ is expressible as $\mathbf{v}_j = \sum_{i=1}^n c_i \mathbf{u}_i$ with $\sum_{i=1}^n |c_i| < 1$, because then $\mathbf{y}_j = d_j \sum_{i=1}^n c_i \mathbf{u}_i = \sum_{i=1}^n (d_j c_i / b_i) \mathbf{x}_i$, and $\sum_{i=1}^n |d_j c_i / b_i| b_i = d_j \sum_{i=1}^n |c_i| < d_j$, contrary to the third condition. Therefore by Lemma 39, there is a norm on V satisfying $\|\mathbf{u}_i\| \leqslant 1$ for $1 \leqslant i \leqslant n$ and $\|\mathbf{v}_j\| \geqslant 1$ for $1 \leqslant i \leqslant m$. I.e., $\|\mathbf{x}_i\| \leqslant b_i$ for all $1 \leqslant i \leqslant n$ and $\|\mathbf{y}_j\| \geqslant d_j$ for all $1 \leqslant j \leqslant m$. $\square$

We can immediately obtain a simpler result if we seek conditions allowing us to set the specific values of the norms of a finite set of vectors:

Corollary 41. *Let* $\mathbf{x}_1, \ldots, \mathbf{x}_n$ *be vectors in a real vector space* V *and let* $b_1, \ldots, b_n$ *be real numbers. Then there exists a norm* $\|\_\|$ *on* V *such that* $\|\mathbf{x}_i\| = b_i$ *for all* i, $1 \leqslant i \leqslant n$, *iff*:

- *For all* i, $1 \leqslant i \leqslant n$, $b_i \geqslant 0$;
- *For all* i, $1 \leqslant i \leqslant n$, *if* $b_i = 0$ *then* $\mathbf{x}_i = 0$;
- *For each* k, $1 \leqslant k \leqslant n$ *there are no real numbers* $c_1, \ldots, c_n$ *such that some* $\mathbf{x}_k = \sum_{i=1}^n c_i \mathbf{x}_i$ *with* $\sum_{i=1}^n |c_i| b_i < b_k$.

Proof. The case when each $b_i = 0$ is evident. If some $b_i \neq 0$, then apply Theorem 40 with $m = n$, $\mathbf{x}_i = \mathbf{y}_i$ and $b_i = d_i$. $\square$

Corollary 42. *The set of valid purely universal sentences in the language of normed spaces is decidable.*

Proof. If σ is a purely universal sentence in prenex normal form $\forall \ldots \cdot \psi$, σ is true iff $\neg\psi$ is unsatisfiable. So it suffices to give a decision procedure for satisfiable quantifier-free formulas. So let ϕ be quantifier-free say with free variables given by $v(\phi) = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ and $s(\phi) = \{u_1, \ldots, u_m\}$. Introduce additional scalar variables $b_1, \ldots, b_k$, one b_i for each norm expression $\|\mathbf{y}_i\|$ appearing in ϕ. (Each such vector $\mathbf{y}_i$ can be written as $p_1\mathbf{x}_1 + \cdots + p_n\mathbf{x}_n$ for polynomials p_i, though the p_i may themselves involve other norm expressions.) Satisfiability of ϕ in a normed space is equivalent to satisfiability of $\phi' \wedge \bigwedge_{i=1}^k \|\mathbf{y}_i\| = b_i$, where ϕ' is ϕ with each $\|\mathbf{y}_i\|$ replaced by its corresponding b_i, in a bottom-up fashion so that ϕ' does not contain the norm operator. But by the corollary, this is equivalent to the satisfiability in a vector space of the following formula:

$$\begin{aligned}\phi'' := {} & \phi' \wedge \\ & \bigwedge_{i=1}^k b_i \geqslant 0 \wedge \\ & \bigwedge_{i=1}^k (b_i = 0 \Rightarrow \mathbf{y}_i = 0) \wedge \\ & \bigwedge_{i=1}^k \big(\forall c_1 \ldots c_k \cdot |c_1| b_1 + \cdots + |c_k| b_k < b_i \Rightarrow \mathbf{y}_i \neq c_1\mathbf{y}_1 + \cdots + c_k\mathbf{y}_k\big).\end{aligned}$$

The decision procedure of Theorem 36 applied to the existential closure of ϕ'' will then decide satisfiability of ϕ'' and hence of ϕ. $\square$

Note that, if the formula ϕ is satisfiable, then our methods give a norm on $\mathbb{R}^n$, whose unit disc may be taken to be a polyhedron, together with a satisfying assignment for ϕ in $\mathbb{R}^n$ under that norm. Thus, at least in principle, the above decision procedure can be extended to give a counter-example if the input purely universal sentence is false. It is also noteworthy that the only instances of multiplication introduced in the passage from ϕ to ϕ'' are in the last conjunct of ϕ''. For the case where the input sentence is purely additive, one can develop a more efficient algorithm using a parametrized linear programming technique.

If K is an ordered field, define a *normed space over* K to be a structure for the language $\mathcal{L}_N$ of normed spaces in which the scalar sort and its operations are interpreted in K and which satisfies the usual axioms for a norm. The proofs above

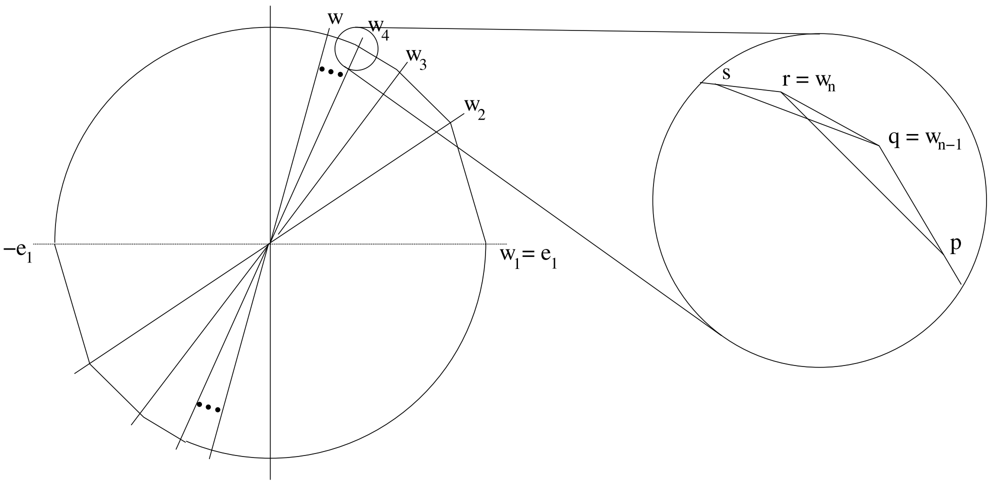

Fig. 8. The unit circle in the space $\mathbb{K}$ with a detail illustrating the predicate $\mathsf{A}(\mathbf{q},\mathbf{r})$.

go through over any real closed field K (for a proof of the Minkowski–Weyl theorem that does not appeal to separation properties that are only valid over $\mathbb{R}$, see, for example, Weyl [45]). We therefore have a decision procedure for the purely universal fragment of the theory of normed spaces over any real closed field K. As this decision procedure is independent of K, we may conclude that a universal sentence in $\mathcal{L}_N$ holds for all normed spaces over a real closed field K iff it holds for all real normed spaces.

8. The ∃∀ and ∀∃ fragments of the theory of normed spaces

In this section we shall see that the ∃∀ and ∀∃ fragments of the theory of normed spaces are both undecidable. Thus the results of Section 7 for the purely existential and purely universal fragments are the best of their type. The proofs given here do make use of multiplication, but the constructions they use have since been adapted to give undecidability results for the additive ∃∀ and ∀∃ fragments (over $\mathbb{R}$) and for theories of normed spaces over an arbitrary ordered field [4,5].

The plan of this section is as follows: we first prove undecidability for the ∃∀ fragment by giving a purely existential characterization of the natural numbers in a certain normed space $\mathbb{K}$ (cf. the proof of Theorem 9); then we prove undecidability of the ∀∃ fragment using a normed space $\mathbb{L}$ whose unit circle includes an encoding of a periodic function; finally, we show that a small adjustment to $\mathbb{L}$ allows us to prove undecidability of the set of all valid sentences of the form $\phi \Rightarrow \psi$ where ϕ and ψ are purely universal, which, up to a logical equivalence, covers undecidability for both the ∀∃ and ∃∀ fragments.

The first proof for the ∃∀ fragment is based on an extremely simple method for encoding the natural numbers in the unit disc of a 2-dimensional normed space.

Theorem 43. *There is a purely existential formula $\mathsf{N}(x)$ in the language of normed spaces such that for any $d \in \{2,3,4,\ldots\} \cup \{\infty\}$, there is a Banach space $\mathbb{K}^d$ of dimension d in which $\mathsf{N}(x)$ defines the natural numbers.*

Proof. We consider the case $d = 2$ first. Using the usual euclidean norm in the plane $\mathbb{R}^2$, define $\mathbf{w}_1 = \mathbf{e}_1$ and then, working anticlockwise around the unit circle, take $\mathbf{w}_n$ to be the unit vector with $\|\mathbf{w}_n - \mathbf{w}_{n-1}\|_e = \frac{1}{n!}$ for $n = 2, 3, \ldots$ as illustrated in Fig. 8 (but not to scale). Then $\|\mathbf{w}_n - \mathbf{w}_1\|_e < \sum_{n=2}^{\infty} \frac{1}{n!} = e - 2 < 1 < \sqrt{2} = \|\mathbf{e}_1 - \mathbf{e}_2\|_e$, and so the $\mathbf{w}_n$ are all in the north-east quadrant and tend to a limit $\mathbf{w}$. Evidently $\|\mathbf{w}\|_e = 1$ and we may define $\mathbb{K} = \mathbb{K}^2$ to be $\mathbb{R}^2$ with the norm $\|\_\|$ whose unit disc is the symmetric convex hull of $A \cup \{\mathbf{w}_1, \mathbf{w}_2, \ldots\}$ where A is the arc running anticlockwise from $\mathbf{w}$ to $-\mathbf{e}_1$. Note that this norm agrees with the euclidean norm on vectors in the north-west and south-east quadrants, so, in particular, if $\mathbf{p}$ and $\mathbf{q}$ are unit vectors in the north-east quadrant, $\|\mathbf{p} - \mathbf{q}\| = \|\mathbf{p} - \mathbf{q}\|_e$. Define predicates $\mathsf{A}(\mathbf{q},\mathbf{r})$ and $\mathsf{N}(x)$ as follows:

$$\begin{aligned}\mathsf{A}(\mathbf{q},\mathbf{r}) := \exists \mathbf{p}\,\mathbf{s} \cdot {} & \|\mathbf{p}\| = \|\mathbf{q}\| = \|\mathbf{r}\| = \|\mathbf{s}\| = 1 \wedge \\ & \big\|(\mathbf{p}+\mathbf{q})/2\big\| = \big\|(\mathbf{q}+\mathbf{r})/2\big\| = \big\|(\mathbf{r}+\mathbf{s})/2\big\| = 1 \wedge \\ & \big\|(\mathbf{p}+\mathbf{r})/2\big\| < 1 \wedge \big\|(\mathbf{q}+\mathbf{s})/2\big\| < 1\end{aligned}$$

$$\begin{aligned}\mathsf{N}(x) := \exists \mathbf{p}\,\mathbf{q}\,\mathbf{r} \cdot {} & \mathsf{A}(\mathbf{p},\mathbf{q}) \wedge \mathsf{A}(\mathbf{q},\mathbf{r}) \wedge \|\mathbf{p}-\mathbf{q}\| > \|\mathbf{q}-\mathbf{r}\| \wedge \\ & \|\mathbf{p}-\mathbf{q}\| = (x+4)\|\mathbf{q}-\mathbf{r}\|.\end{aligned}$$

For any $n \geqslant 3$, $\mathsf{A}(\mathbf{w}_{n-1}, \mathbf{w}_n)$ holds in $\mathbb{K}$ (choose $\mathbf{p} \in [\mathbf{w}_{n-2}, \mathbf{w}_{n-1})$ and $\mathbf{s} \in (\mathbf{w}_n, \mathbf{w}_{n+1}]$ as in Fig. 8). Conversely, assume the matrix of $\mathsf{A}(\mathbf{q}, \mathbf{r})$ holds for some $\mathbf{p}$, $\mathbf{q}$, $\mathbf{r}$ and $\mathbf{s}$: the conditions imposed imply that $\mathbf{p}$, $\mathbf{q}$, $\mathbf{r}$ and $\mathbf{s}$ are pairwise distinct and that the line segments $[\mathbf{p}, \mathbf{q}]$, $[\mathbf{q}, \mathbf{r}]$ and $[\mathbf{r}, \mathbf{s}]$ lie in the unit circle of $\mathbb{K}$; moreover, $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{r}$ cannot be collinear, otherwise we would have $\|(\mathbf{p}+\mathbf{r})/2\| = 1$ and similarly $\mathbf{q}$, $\mathbf{r}$ and $\mathbf{s}$ cannot be collinear; thus $\mathbf{q}$ and $\mathbf{r}$ must be adjacent isolated extreme points of the unit disc in $\mathbb{K}$ and, for some $n \geqslant 3$, we have $\pm\{\mathbf{q}, \mathbf{r}\} = \{\mathbf{w}_{n-1}, \mathbf{w}_n\}$. Since $\|\mathbf{w}_{n-1} - \mathbf{w}_n\| = \frac{1}{n!}$ for all $n \geqslant 2$, it follows that, $\mathsf{N}(x)$ holds iff for some $n \geqslant 3$, $\frac{1}{n!} = \frac{x+4}{(n+1)!}$, which holds (with $n = x+3$) iff $x \in \mathbb{N}$. Clearly $\mathsf{N}(x)$ is equivalent to a purely existential formula and the theorem is proved for $d = 2$.

For $d > 2$, let V be a Hilbert space of dimension $d-2$, and define $\mathbb{K}^d$ to be the 2-sum of $\mathbb{K}$ and V, i.e., the product vector space $\mathbb{K} \times V$ equipped with the norm defined by $\|(\mathbf{p}, \mathbf{v})\| = \sqrt{\|\mathbf{p}\|^2 + \|\mathbf{v}\|^2}$. That this is a norm making $\mathbb{K}^d$ into a Banach space is readily verified. We identify $\mathbb{K}$ with the subspace $\mathbb{K} \times 0$ of $\mathbb{K}^d$. It can be shown that if $\mathbf{a}$ and $\mathbf{b}$ are distinct unit vectors and if the line segment $[\mathbf{a}, \mathbf{b}]$ is contained in the unit sphere of $\mathbb{K}^d$, then $[\mathbf{a}, \mathbf{b}]$ is parallel to $\mathbb{K}$ (see [4] for a proof). Hence there are $\mathbf{p}, \mathbf{q} \in \mathbb{K}$, $\mathbf{v} \in V$ and $t \in (0, 1]$ such that $\mathbf{a} = (t\mathbf{p}, \mathbf{v})$, $\mathbf{b} = (t\mathbf{q}, \mathbf{v})$ and the line segment $[\mathbf{p}, \mathbf{q}]$ is contained in the unit circle of $\mathbb{K}$. Because $\mathsf{N}(x)$ only depends on the ratios of the distances between adjacent extreme points of the unit sphere, $\mathsf{N}(x)$ holds in $\mathbb{K}^d$ iff it holds in $\mathbb{K}$ and the proof is complete. □

Corollary 44. *Let $d \in \{2, 3, 4, \ldots\} \cup \{\infty\}$ and let $\mathcal{C}$ be any class of normed spaces that includes all Banach spaces of dimension d. The set of $\exists\forall$ sentences that are valid in $\mathcal{C}$ is undecidable.*

Proof. Just as in the proof of Theorem 9, use the existence of a structure in which a purely existential formula defines the subset $\mathbb{N}$ of $\mathbb{R}$ to reduce the satisfiability of systems of Diophantine equations to satisfiability in $\mathcal{C}$. □

Our next undecidability result concerns $\forall\exists$ sentences in theories of normed spaces. As in the proof of Theorem 9, given a quantifier-free formula $\psi(x_1, \ldots, x_k)$ in the language of arithmetic, we will exhibit an $\exists\forall$ sentence ψ_1 in the language of normed spaces that is satisfiable iff $\psi(x_1, \ldots, x_k)$ is satisfiable in $\mathbb{N}$. However, the quantifier structure of the sentence Peano no longer suits our purposes. Instead, we will design ψ_1 so that its models comprise spaces whose unit circle contains a representation of a periodic function on the set $\mathbb{R}_+$ of positive real numbers, which we then use to define $\mathbb{N}$. We begin by showing how a norm may be used to define a function on $\mathbb{R}_+$.

Consider a 2-dimensional normed space with a basis $\mathbf{e}_1$ and $\mathbf{e}_2$ where $\|\mathbf{e}_1\| = 1$. For $x \in (-1, 1)$, we have $\|x\mathbf{e}_1\| = |x| < 1$ and so by the remarks of Section 2.3.3, the vertical line $\{x\mathbf{e}_1 + y\mathbf{e}_2 \mid y \in \mathbb{R}\}$ must meet the unit circle in exactly two points, one in the upper half-plane and one in the lower. Thus with respect to the given basis, the part of the unit circle lying above the open line segment $(-\mathbf{e}_1, \mathbf{e}_1)$ forms the graph of a function.

Now consider the following formulas in which the vector variables $\mathbf{e}_1$ and $\mathbf{e}_2$ occur free in addition to the scalar variables listed as parameters:

$$\Gamma(x, y) :\equiv -1 < x < 0 \wedge 0 < y < 1 \wedge \|x\mathbf{e}_1 + y\mathbf{e}_2\| = 1$$

$$\mathsf{G}(s, t) :\equiv s > 0 \wedge t > 0 \wedge \|-(1+t)\mathbf{e}_1 + (1+s)t\mathbf{e}_2\| = (1+s)(1+t).$$

Thus, for $s, t \neq -1$, $\mathsf{G}(s, t)$ is equivalent to $\Gamma(\frac{-1}{1+s}, \frac{t}{1+t})$. Assume that $\mathbf{e}_1$ and $\mathbf{e}_2$ are vectors in some normed space V and that the following condition holds for all x and y:

$$\begin{aligned} \mathsf{Def} :\equiv\ & \|\mathbf{e}_1\| = \|\mathbf{e}_2\| = 1 \wedge \\ & (x\mathbf{e}_1 + y\mathbf{e}_2 = \mathbf{0} \Rightarrow x = y = 0) \wedge \\ & \big(|x| > 0 \wedge |y| > 0 \wedge \|x\mathbf{e}_1 + y\mathbf{e}_2\| = 1 \Rightarrow |x| < 1 \wedge |y| < 1\big). \end{aligned}$$

So $\mathbf{e}_1$ and $\mathbf{e}_2$ are unit vectors spanning a 2-dimensional subspace of V and, when we use them to define coordinates in that subspace, the unit circle is contained in the square with diagonal $[-\mathbf{e}_1 - \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2]$ and meets its boundary in the four points $\pm\mathbf{e}_1$ and $\pm\mathbf{e}_2$. This means that $\Gamma(x, y)$ will hold iff $y = \gamma(x)$ where γ is the function whose graph comprises the part of the unit circle lying strictly above the open line segment $(-\mathbf{e}_1, \mathbf{0})$, while $\mathsf{G}(s, t)$ will hold iff $t = g(s)$ where the graph of g is the image of the north-west quadrant of the unit circle under the continuous bijection $e : (-1, 0) \times (0, 1) \to \mathbb{R}_+ \times \mathbb{R}_+$ defined by $e(s, t) = (\frac{1+s}{-s}, \frac{t}{1-t})$. The condition Def ensures that $g(s)$ is well-defined for all $s \in \mathbb{R}_+$. Fig. 9 illustrates γ and g for various norms on $\mathbb{R}^2$.

Lemma 45. *For some positive integer M, there is a 2-dimensional normed space $\mathbb{L}$ containing vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ for which Def holds for every x and y, while $\mathsf{G}(s, t)$ holds iff $s > 0$ and $t = 2s + s^2 + \frac{1}{M}\sin(s)$.*

Proof. Define functions $g_r : \mathbb{R}_+ \to \mathbb{R}_+$ for $2 > r > 0$ by $g_r(s) = 2s + s^2 + r\sin(s)$. Under the bijection e, g_r corresponds to the function $\gamma_r : (-1, 0) \to (0, 1)$ where:

$$\gamma_r(x) = \frac{g_r(\frac{x+1}{-x})}{1 + g_r(\frac{x+1}{-x})}.$$

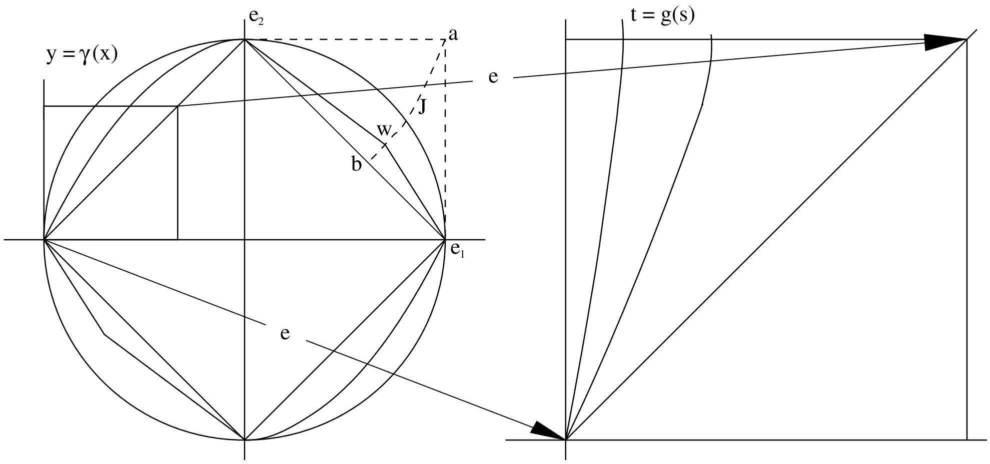

Fig. 9. The functions defined by $\Gamma(s,t)$ and $\mathsf{G}(s,t)$ under various norms on $\mathbb{R}^2$.

I claim that for all small enough $r>0$, γ_r is a concave function, i.e., the part of the plane lying below the graph of γ_r is convex. Assuming this, we can choose a positive integer M such that $\gamma=\gamma_{1/M}$ is concave. Noting that $\gamma(x)$ tends to 0 as x tends to -1 and to 1 from below as x tends to 0, we can extend γ to a concave function γ^* on $[-1,1]$ by taking $\gamma^*(-1)=0$ and $\gamma^*(x)=1-x$ for $0\leqslant x\leqslant 1$. Let $\mathbb{L}$ be $\mathbb{R}^2$ under the norm whose unit circle meets the upper half-plane in the graph of γ^*. Then in $\mathbb{L}$, Def holds for every x and y and $\Gamma(x,y)$ defines $\gamma=\gamma_{1/M}$, so that $\mathsf{G}(s,t)$ holds iff $s>0$ and $t=g_{1/M}(s)=2s+s^2+\frac{1}{M}\sin(s)$.

It remains to prove that γ_r is concave for all small enough $r>0$. Certainly γ_r is twice differentiable, and then, by standard results on concave functions, it is sufficient to show that the second derivative $\gamma_r''(x)$ is never positive for x in $(-1,0)$. Differentiating the formula for γ_r above twice gives:

$$\gamma_r'(x)=\frac{g_r'(\frac{x+1}{-x})}{x^2(1+g_r(\frac{x+1}{-x}))^2}$$

$$\gamma_r''(x)=\frac{g_r''(\frac{x+1}{-x})}{x^4(1+g_r(\frac{x+1}{-x}))^2}+\frac{2g_r'(\frac{x+1}{-x})}{-x^3(1+g_r(\frac{x+1}{-x}))^2}-\frac{2(g_r'(\frac{x+1}{-x}))^2}{x^4(1+g_r(\frac{x+1}{-x}))^3}.$$

Writing $s=\frac{x+1}{-x}$, so that $s>0$ and $\frac{1}{-x}=1+s$, and multiplying by the positive quantity $-x^3(1+g_r(s))^3$, we see that $\gamma_r''(x)$ has the same sign as $h_r(s)$ where:

$$\begin{aligned}h_r(s)&=\big(1+g_r(s)\big)\big[(1+s)g_r''(s)+2g_r'(s)\big]-2(1+s)\big(g_r'(s)\big)^2\\&=\big(1+2s+s^2+r\sin(s)\big)\big[(1+s)\big(2-r\sin(s)\big)+4+4s+2r\cos(s)\big]\\&\quad-2(1+s)\big(2+2s+r\cos(s)\big)^2.\end{aligned}$$

As $-1\leqslant\sin(s),\cos(s)\leqslant 1$, we have that $h_r(s)\leqslant p(s,r)$, where:

$$\begin{aligned}p(s,r)&=\big(1+2s+s^2+r\big)\big[(1+s)(2+r)+4+4s+2r\big]\\&\quad-2(1+s)(2+2s-r)^2\\&=p_0(s)+p_1(s)r+p_2(s)r^2.\end{aligned}$$

each $p_i(s)$ being a polynomial of degree at most 3 in s with constant coefficients, say $p_i(s)=p_{i0}+p_{i1}s+p_{i2}s^2+p_{i3}s^3$, $i=0,1,2$. Since $p_0(s)=p(s,0)=6(1+s)^3-8(1+s)^3=-2-6s-6s^2-2s^3$, each p_{0j} is negative. Let q_j be the coefficient of s^j in $p(s,r)$ so $q_j=p_{0j}+p_{1j}r+p_{2j}r^2$. Since $p_{0j}<0$, we may choose $\epsilon>0$ such that whenever $0<r<\epsilon$, $q_j<0$, $j=0,1,2,3$. But then if $0<r<\epsilon$, we find that $p(s,r)<0$ for all $s>0$ whence $\gamma_r''(x)$ is negative for all x in $(-1,0)$, since it has the same sign as the quantity $h_r(s)\leqslant p(s,r)<0$. Thus γ_r is concave for $0<r<\epsilon$ and the proof is complete. □

Let the space $\mathbb{L}$ and the positive integer M be as given by the lemma. In $\mathbb{L}$, the following formula then defines the graph of the positive half of the sine function when $\mathbf{e}_1$ and $\mathbf{e}_2$ are given their usual interpretations:

$$\mathsf{SIN}(s,t):=\mathsf{G}\left(s,2s+s^2+\frac{1}{M}t\right).$$

Now consider the following formulas:

$$\begin{aligned}\mathsf{Periodic} := {} & a > 0 \wedge \\ & \big(0 < s < 2a \Rightarrow \big(\mathsf{SIN}(s,0) \Leftrightarrow s = a\big)\big) \wedge \\ & \big(\mathsf{SIN}(s,t) \Rightarrow \mathsf{SIN}(s+a,-t)\big) \\ \mathsf{N}(x) := {} & \mathsf{SIN}\big((x+1)a, 0\big).\end{aligned}$$

In $\mathbb{L}$ with the usual interpretation of $\mathbf{e}_1$ and $\mathbf{e}_2$, $\mathsf{Periodic}$ holds for all s and t if we interpret a as π, in which case $\mathsf{N}(x)$ holds iff x is interpreted as a natural number. On the other hand, if V is any normed space and there are $\mathbf{e}_1, \mathbf{e}_2 \in V$ and $a \in \mathbb{R}$ such that Def and $\mathsf{Periodic}$ hold for all x, y, s and t, then $\mathsf{SIN}(s,t)$ must define the graph of a function on $\mathbb{R}_+$ whose zeroes comprise precisely the positive integer multiples of a, so that $\mathsf{N}(x)$ defines the natural numbers.

Theorem 46. *Let $d \in \{2, 3, 4, \ldots\} \cup \{\infty\}$ and let $\mathcal{C}$ be any class of normed spaces that includes all Banach spaces of dimension d. The set of $\forall\exists$ sentences that are valid in $\mathcal{C}$ is undecidable.*

Proof. We will prove the equivalent claim that the set of $\exists\forall$ sentences that are satisfiable in $\mathcal{C}$ is undecidable. Given a quantifier-free formula $\phi(x_1, \ldots, x_k)$ in the language of arithmetic, define:

$$\begin{aligned}\phi_1 := {} & \exists \mathbf{e}_1\, \mathbf{e}_2\, a\, x_1 \ldots x_k \cdot \forall x\, y\, s\, t \cdot \\ & \quad \mathsf{Def} \wedge \mathsf{Periodic} \wedge \mathsf{N}(x_1) \wedge \cdots \wedge \mathsf{N}(x_k) \wedge \phi(x_1, \ldots, x_k).\end{aligned}$$

Take $V = \mathbb{L} \times W$ where W is any vector space of dimension $d-2$ under any norm extending that of the factor $\mathbb{L}$; if $\phi(x_1, \ldots, x_k)$ is satisfiable in $\mathbb{N}$, ϕ_1 will be satisfiable in $\mathbb{L}$ and hence in V. Conversely, if ϕ_1 is satisfiable in some normed space, then under a satisfying assignment, the conditions Def and $\mathsf{Periodic}$ mean that $\mathsf{N}(x)$ must define $\mathbb{N}$ in V, so $\phi(x_1, \ldots, x_k)$ is satisfiable in $\mathbb{N}$. So, just as in the proof of Theorem 9, the existence of a decision procedure for $\exists\forall$ sentences satisfiable in $\mathcal{C}$ would contradict the undecidability of satisfiability for quantifier-free formulas in arithmetic. □

To state our final result on undecidability, let us say a sentence is $\forall \Rightarrow \forall$ if it has the form $A \Rightarrow B$ where A and B are purely universal. With a small adjustment to the construction used to prove Theorem 46, we now show that validity for $\forall \Rightarrow \forall$ sentences is undecidable. As $\forall \Rightarrow \forall$ sentences have both $\exists\forall$ and $\forall\exists$ equivalents, this provides an alternative proof for both Corollary 44 and Theorem 46.

Theorem 47. *Let $d \in \{2, 3, 4, \ldots\} \cup \{\infty\}$ and let $\mathcal{C}$ be any class of normed spaces that includes all Banach spaces of dimension d. The set of $\forall \Rightarrow \forall$ sentences that are valid in $\mathcal{C}$ is undecidable.*

Proof. If $d = 2$, let $\mathbb{L}$ be the normed space constructed in the proof of Lemma 45. Using the $\mathbb{L}$-norm, let $C = \{\mathbf{w} \mid \|\mathbf{e}_2 - \mathbf{w}\| = 1\}$ be the unit circle centred at $\mathbf{e}_2$ and consider the intersection $J = C \cap T$, where T is the triangle with vertices $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{a} = \mathbf{e}_1 + \mathbf{e}_2$ (see Fig. 9). J meets the perimeter of T at the vertex $\mathbf{a}$ with $\|\mathbf{e}_1 - \mathbf{a}\| = 1$ and at a point $\mathbf{b}$ on the edge $[\mathbf{e}_1, \mathbf{e}_2]$ with $t = \|\mathbf{e}_1 - \mathbf{b}\| < 1$ (since $1 < \|\mathbf{e}_2 - \mathbf{e}_1\| < 2$). J is a continuous curve and so $\|\mathbf{e}_1 - \mathbf{w}\|$ takes on all values in $[t, 1]$ as $\mathbf{w}$ ranges over J. Since T meets the unit disc of $\mathbb{L}$ in the edge $[\mathbf{e}_1, \mathbf{e}_2]$, it follows that there are $i/j \in \mathbb{Q}$ and $\mathbf{w} \in J \subseteq T$ such that $\|\mathbf{e}_2 - \mathbf{w}\| = 1 < \|\mathbf{w}\|$ and $\|\mathbf{e}_1 - \mathbf{w}\| = i/j < 1$. Let $\mathbb{L}_0$ be the normed space whose unit disc is the symmetric convex hull of the unit disc of $\mathbb{L}$ and such a $\mathbf{w}$. Then the $\mathbb{L}$-norm and the $\mathbb{L}_0$-norm agree in the north-west quadrant and so $\mathbb{L}$ and $\mathbb{L}_0$ define the same functions γ and g and assign the same lengths to the line segments that make up the north-east quadrant of the unit circle of $\mathbb{L}_0$. In $\mathbb{L}_0$, the following formula holds iff $\mathbf{p} = s\mathbf{e}_1$, $\mathbf{q} = s\mathbf{e}_2$ and $\mathbf{r} = s\mathbf{w}$ where $s = \pm 1$.

$$\begin{aligned}\mathsf{W}(\mathbf{p}, \mathbf{q}, \mathbf{r}) := {} & \|\mathbf{p}\| = \|\mathbf{q}\| = \|\mathbf{r}\| = \big\|(\mathbf{p}+\mathbf{r})/2\big\| = \big\|(\mathbf{q}+\mathbf{r})/2\big\| = 1 \wedge \\ & \|\mathbf{p} - \mathbf{r}\| = i/j \wedge \|\mathbf{q} - \mathbf{r}\| = 1 \wedge \big\|(\mathbf{p}+\mathbf{q})/2\big\| < 1.\end{aligned}$$

Also, the following formula is invariant under $\mathbf{v} \mapsto -\mathbf{v}$ and, when the free variables $\mathbf{e}_1$ and $\mathbf{e}_2$ are given their usual interpretation in $\mathbb{L}_0$, holds iff $x = \pi$.

$$\Pi(x) := x < 4 \wedge \mathsf{SIN}(x, 0).$$

Now, given a quantifier-free formula $\phi(x_1, \ldots, x_k)$ in the language of arithmetic, define sentences ψ and ρ as follows:

$$\psi := \forall \mathbf{e}_1\, \mathbf{e}_2\, \mathbf{w}\, a\, x\, y\, s\, t \cdot \mathsf{W}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{w}) \wedge \Pi(a) \Rightarrow \mathsf{Def} \wedge \mathsf{Periodic}$$

$$\rho := \forall \mathbf{e}_1\, \mathbf{e}_2\, \mathbf{w}\, a\, x_1 \ldots x_k \cdot \mathsf{W}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{w}) \wedge \Pi(a) \wedge \mathsf{N}(x_1) \wedge \cdots \mathsf{N}(x_k) \Rightarrow \neg\phi(x_1, \ldots, x_k).$$

By the above remarks on $\mathsf{W}(\mathbf{p}, \mathbf{q}, \mathbf{r})$ and $\Pi(x)$, ψ holds and $\mathsf{W}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{w}) \wedge \Pi(a)$ is satisfiable in $\mathbb{L}_0$. Also, in any normed space in which ψ holds, $\mathsf{N}(x)$ is true under an assignment that satisfies $\mathsf{W}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{w}) \wedge \Pi(a)$ iff $x \in \mathbb{N}$. Thus if ψ holds and

$\mathsf{W}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{w}) \wedge \Pi(a)$ is satisfiable, then ρ holds iff $\phi(x_1, \ldots, x_k)$ is unsatisfiable in $\mathbb{N}$. Thus $\psi \Rightarrow \rho$ is valid in a class of spaces including $\mathbb{L}_0$ iff $\phi(x_1, \ldots, x_k)$ is unsatisfiable and so a decision procedure for $\forall \Rightarrow \forall$ sentences that are valid in such a class would lead to a decision procedure for satisfiable quantifier-free sentences of arithmetic, which is impossible.

For $d > 2$, let V be a Hilbert space of dimension $d - 2$, let W be the 2-sum of $\mathbb{L}_0$ and V, and identify $\mathbb{L}_0$ with the subspace $\mathbb{L}_0 \times 0$ of W. As in the proof of Theorem 43, if a line segment $[\mathbf{u}, \mathbf{v}]$ lies in the unit sphere of W, then it is parallel to $\mathbb{L}_0$. Moreover, if also $\|\mathbf{u} - \mathbf{v}\| = 1$ then we must have that $\{\mathbf{u}, \mathbf{v}\} = \pm\{\mathbf{e}_2, \mathbf{w}\} \subseteq \mathbb{L}_0$. This means that the formula $\mathsf{W}(\mathbf{p}, \mathbf{q}, \mathbf{r})$ defines the same set of triples in W as it does in $\mathbb{L}_0$. The argument for $d = 2$ then shows that validity of $\forall \Rightarrow \forall$ sentences in any class of spaces including W is undecidable. □

9. Related work and concluding remarks

The reduction of second-order arithmetic to the theory of the real numbers augmented with a predicate symbol for the integers has been known since the 1960s if not before. In descriptive set theory, the main ideas of Section 3 are used to show that a subset of $\mathbb{R}^n$ is projective iff it is definable in the theory of the real numbers augmented with a predicate for the integers; see for example, Moschovakis [32], Theorem 8B.4 or Kechris [25, ex. 37.6]. However, we know of no published account of these ideas applied to problems of decidability.

Scott [37] considers *geometric relations*, i.e., relations such as "equidistant" that are defined on affine euclidean space in all dimensions and that are invariant under isometric embeddings. He works with single-sorted languages whose variables range over points and shows that a first-order sentence with $k + 1$ distinct variables holds for every interpretation of its relation symbols as geometric relations iff it holds in dimension k. This is clearly closely related to our Theorem 31. (Scott's $k + 1$ is our k because the constant vector $\mathbf{0}$ costs us one variable.) He applies his result to a formulation of euclidean geometry as a single-sorted theory with "between" and "equidistant" as primitive predicates and obtains decidability and related results for theories $\mathcal{E}$, $\mathcal{E}_m$ and $\mathcal{E}_\infty$ analogous to our IP, IP^m and IP^∞. Scott's proofs are based on semantic considerations that apply to all geometric relations, in contrast with our more algorithmic approach via a syntactic quantifier elimination procedure. In his single-sorted Tarski-style formalism, the emphasis is on geometry and the real numbers only arise implicitly as equivalence classes for the equidistance relation, while our two-sorted approach is closer to typical mathematical and engineering practice.

Before both our work and that of Kutz [28], Bondi [9] had proved the undecidability of the theory of metric spaces. Let $\mathcal{L}_B$ be the language of a single binary predicate, intended to be interpreted as a symmetric relation. Translated into our two-sorted framework, Bondi's proof defines a mapping $\sigma \mapsto \sigma_0$ from sentences in $\mathcal{L}_B$ to sentences in the language $\mathcal{L}_M$ of metric spaces such that σ_0 is valid if σ is valid and $\neg\sigma_0$ is valid if σ is finitely refutable. Undecidability follows from a theorem of Lavrov. The only metric spaces Bondi uses are finite subspaces of euclidean space, so her methods show that the theory of such metric spaces is undecidable. The methods of the present paper give more information about the many-one degree of the theory and show that the additive and the $\exists\forall$ fragments are undecidable. As far as we know, our Theorem 8 giving the decidability of the $\forall\exists$ fragment is the strongest known positive result on decision procedures for theories of metric spaces.

Bondi [10] considers normed spaces and inner product spaces (she actually writes "Hilbert spaces", but metric completeness plays no rôle in her proofs). She first proves the undecidability of the theory of normed spaces by a proof similar in structure to her proof for metric spaces sketched above. The method of the proof actually gives the undecidability of any class of normed spaces containing all finite-dimensional spaces, whereas the methods of the present paper give undecidability of any class of spaces containing all spaces of any given (possibly infinite) dimension $d > 1$. Bondi then turns to the decision problem for inner product spaces: she first gives recursive axiomatizations of the theory T of all non-trivial inner product spaces, of the theory T_0 of all infinite-dimensional inner product spaces and of the theories T_n of n-dimensional inner product spaces, $n = 1, 2 \ldots$. She shows that the theory T_0 is model complete and so complete, since any two models of T_0 contain isomorphic submodels. Being recursively axiomatizable and complete, T_0 is therefore decidable. This argument shows the correctness of a decision procedure that enumerates proofs rather than the more efficient procedures of Section 6 above. Writing T_{fin} for the theory of non-trivial finite-dimensional inner product spaces, Bondi goes on to argue that $T = T_{\text{fin}}$ and concludes using a lemma of Ershov that T is decidable. Unfortunately, there is a significant gap in her proof that $T = T_{\text{fin}}$: she claims that a certain sentence in T_0 must belong to T_n for sufficiently large n, but gives no proof of this. Her claim is true, but it is unclear how to prove it without appealing to Theorem 29 from the present paper. A precisely analogous situation in which the analogue of $T = T_{\text{fin}}$ fails can be reached by adding a predicate $\mathsf{D}(x)$ on scalars with the intended interpretation that $\mathsf{D}(x)$ hold in V iff $x \leqslant \dim(V)$. This gives extensions T', T'_0, etc. of the theories T, T_0, etc. $\mathsf{D}(x)$ can be defined by a recursive set of axioms and T'_0 and the T'_n can be seen to be complete using Bondi's arguments. However, $T' \neq T'_{\text{fin}}$, since $\exists x \cdot \neg\mathsf{D}(x)$ holds in an inner product space V iff V is finite-dimensional.

A vector space over the real field is a special case of a module over a ring. Theories of modules over rings have been widely studied, often with a view to applications in algebra. However, most of this work has concentrated on single-sorted theories in which quantification over the ring of scalars is not allowed, the action of the ring on the module being represented by function symbols f_x indexed by ring elements such that $f_x(\mathbf{v}) = x\mathbf{v}$ in the intended interpretations. With this formulation, the procedure of Baur and Monk gives quantifier elimination relative to a set of predicates that specialize to our dimension predicates D_n when the ring is a field. This procedure provides a powerful theoretical tool; see for example, Prest [34]. For modules over any Bézout domain, Van den Dries and Holly [44] give a quantifier elimination procedure for formulas in which free scalar variables are allowed. Their method is via a reduction to the single-sorted language over a ring of polynomials and it is unclear how it could be generalized to deal with scalar quantification.

Granger [17] considers the theory of vector spaces equipped with a bilinear form and proves a form of quantifier elimination for the natural two-sorted formulations using model-theoretic arguments. He gives an interesting discussion of two-sorted formulations that attempt to decouple the model theory of the underlying field from the model theory of vector spaces over it. These formulations lie somewhere between the single-sorted formulation of the Baur–Monk theorem and the two-sorted formulation adopted in the present work. Granger's conclusion is that such a decoupling is in some sense not possible.

As already mentioned in Section 8, our results on the undecidability of the ∀∃ and ∃∀ fragments of the theory of normed spaces can be strengthened to the additive case. Arthan [4] does this by adapting the constructions used here to prove Theorem 47 so that scalar multiplication becomes definable via a purely existential formula.

In the present paper, we have focused on the case when the field of scalars comprises the real numbers. However, all the results on inner product spaces and Hilbert spaces in Section 6 go through with the proofs unchanged for an arbitrary real closed field. As discussed at the end of Section 7, our positive decidability result for the universal fragment of the theory of real normed spaces can also be adapted to cover normed spaces over any real closed field.

One cannot hope to reduce second order arithmetic to a recursively axiomatizable theory like the theory of normed spaces over a real closed field. However, Arthan [5] gives a construction over an arbitrary ordered field of a 2-dimensional normed space that encodes the graph of natural number multiplication. Via a reduction of Robinson's theory Q, this gives the undecidability of the additive theory $\mathsf{NS}_{+}(\mathcal{C})$, and hence, *a fortiori*, the full theory $\mathsf{NS}(\mathcal{C})$ for normed spaces over any non-empty class of ordered fields $\mathcal{C}$ and similarly for the theories $\mathsf{NS}_{+}^{\infty}(\mathcal{C})$, $\mathsf{NS}_{+}^{n}(\mathcal{C})$, $1 < n \in \mathbb{N}$, and $\mathsf{NS}_{+}^{\mathbb{F}}(\mathcal{C})$ with the indicated constraints on dimension.

Kopperman [26,27] considers formalizations of Hilbert spaces and metric spaces in a family of infinitary languages, $L^{t}_{\pi,\varepsilon}$, where t amounts to a many-sorted signature and π and ε are cardinals. $L^{t}_{\pi,\varepsilon}$ has $\pi + \epsilon$ variables and admits conjunction and disjunction of any set X of formulas where $|X| < \pi$ and quantification over any set Y of variables where $|Y| < \epsilon$ (so the usual finitary language over a signature t is $L^{t}_{\omega,\omega}$). For t a signature appropriate for Hilbert spaces, he gives a result for $L^{t}_{\omega_1,\omega_1}$, redolent of our Corollary 35, stating that any formula is equivalent to a boolean combination of sentences D_n asserting the dimension is $n \in \mathbb{N} \cup \{\infty\}$. However, the infinitary languages are much more expressive than the languages we consider: in the case of separable metric spaces, one can encode the full metric structure of a countable dense subset in a single sentence. Thus, by contrast with our undecidability results, Kopperman proves quantifier elimination for separable complete metric spaces relative to a set of formulas that define the possible countable dense subsets. Unsurprisingly, Kopperman's methods of proof are quite different from those of the present work.

Special languages and logics for Banach spaces and similar structures have been widely studied, largely from a model-theoretic perspective, with applications in functional analysis in mind; see, for example, Henson and Iovino [18]. This work has typically involved logics that are weaker than full first-order logic, since metric completeness makes conventional model theory for Banach spaces less useful in the intended applications. Shelah and Stern [38] have demonstrated the problems with conventional model theory in this context using a construction with a similar flavour to our construction of a sentence that holds in all Banach spaces but is not valid in all normed spaces.

The work reported in the present paper was motivated by an interest in applying mechanized theorem-proving to problems in pure mathematics and engineering. For the potential applications, vector spaces and inner product spaces over the real field are important, and, as we have seen, they admit more powerful decision procedures than modules over an arbitrary ring. However, the complexity of these decision procedures and the undecidability of theories of normed spaces present some interesting challenges.

Acknowledgements

We thank the referee for a very constructive and thorough report, for helpful advice on the notation and presentation, particularly in Section 6, and for providing a neat improvement to the proof of Theorem 8. We are grateful to Angus MacIntyre, Dana Scott and the referee for valuable pointers to the literature. Finally, we are deeply indebted to Nadya Kuzmina for helping us to follow the referee's pointers by kindly undertaking to translate the papers by Irina Bondi from the original Russian.

Arthan's work was supported in part by UK EPSRC grant EP/F02309X/1.

References

[1] D. Amir, Characterizations of Inner Product Spaces, Operator Theory: Advances and Applications, vol. 20, Birkhäuser, 1986.
[2] N. Aronszajn, Caractérisation métrique de l'espace de Hilbert, des espaces vectoriels, et de certains groupes métriques, C. R. Acad. Sci., Paris 201 (1935) 811–813, 873–875.
[3] R. Arthan, On Aronszajn's criterion for Euclidean space, Amer. Math. Monthly 119 (5) (2012) 419–422.
[4] R.D. Arthan, Undecidability for the additive ∀ ⇒ ∀ fragment of the theory of normed spaces, http://arxiv.org/abs/1002.1381, 2010.
[5] R.D. Arthan, The decision problem for normed spaces over any class of ordered fields, http://arxiv.org/abs/1104.3293, 2011.
[6] S. Basu, R. Pollack, M.-F. Roy, Algorithms in Real Algebraic Geometry, Algorithms and Computation in Mathematics, vol. 10, Springer-Verlag, 2006.
[7] P. Bernays, M. Schönfinkel, Zum Entscheidungsproblem der mathematischen Logik, Math. Ann. 99 (1928) 401–419.

[8] J. Bochnak, M. Coste, M.-F. Roy, Real Algebraic Geometry, Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 36, Springer-Verlag, 1998.
[9] I.L. Bondi, The decision problem for metric spaces, Izv. Vysš. Učebn. Zaved. Matematika 128 (5) (1973) 24–27 (in Russian).
[10] I.L. Bondi, The decision problem for normed linear spaces and for Hilbert spaces, Izv. Vysš. Učebn. Zaved. Matematika 132 (5) (1973) 3–10 (in Russian).
[11] B.F. Caviness, J.R. Johnson (Eds.), Quantifier Elimination and Cylindrical Algebraic Decomposition, Texts and Monographs in Symbolic Computation, Springer-Verlag, 1998.
[12] G.E. Collins, Quantifier elimination for real closed fields by cylindrical algebraic decomposition, in: H. Brakhage (Ed.), Second GI Conference on Automata Theory and Formal Languages, in: Lecture Notes in Computer Science, vol. 33, Springer-Verlag, Kaiserslautern, 1976, pp. 134–183.
[13] P. Downey, Undecidability of Presburger arithmetic with a single monadic predicate letter, Technical Report 18-72, Center for Research in Computing Technology, Harvard University, 1972.
[14] J. Ferrante, C. Rackoff, A decision procedure for the first order theory of real arithmetic with order, SIAM J. Comput. 4 (1975) 69–76.
[15] P. Fontaine, Techniques for verification of concurrent systems with invariants, Ph.D. thesis, Institut Montefiore, Université de Liège, 2004.
[16] L. Gårding, Some Points of Analysis and Their History, University Lecture Series, vol. 11, American Mathematical Society/Higher Education Press, 1997.
[17] N. Granger, Stability, simplicity and the model theory of bilinear forms, Ph.D. thesis, University of Manchester, 1999.
[18] C.W. Henson, J. Iovino, Ultraproducts in analysis, in: Analysis and Logic, in: London Mathematical Society Lecture Notes, vol. 262, Cambridge University Press, 2002, pp. 1–113.
[19] L. Hodes, Solving problems by formula manipulation in logic and linear inequalities, Artificial Intelligence 3 (1972) 165–174.
[20] W. Hodges, Model Theory, Encyclopedia of Mathematics and its Applications, vol. 42, Cambridge University Press, 1993.
[21] D. Hofbauer, Termination proofs by multiset path orderings imply primitive recursive derivation lengths, Theoret. Comput. Sci. 105 (1) (1992) 129–140.
[22] L. Hörmander, The Analysis of Linear Partial Differential Operators II, Grundlehren der mathematischen Wissenschaften, vol. 257, Springer-Verlag, 1983.
[23] R.C. James, Orthogonality in normed linear spaces, Duke Math. J. 12 (1945) 291–302.
[24] P. Jordan, J. von Neumann, On inner products in linear, metric spaces, Ann. of Math. (2) 36 (3) (1935) 719–723.
[25] A.S. Kechris, Classical Descriptive Set Theory, Graduate Texts in Mathematics, vol. 156, Springer-Verlag, 1995.
[26] R. Kopperman, Application of infinitary languages to metric spaces, Pacific J. Math. 23 (1967) 299–310.
[27] R. Kopperman, The $L_{\omega_1\omega_1}$-theory of Hilbert spaces, J. Symbolic Logic 32 (1967) 295–304.
[28] O. Kutz, F. Wolter, H. Sturm, N.-Y. Suzuki, M. Zakharyaschev, Logics of metric spaces, ACM Trans. Comput. Log. 4 (2003) 260–294.
[29] M. Manzano, Extensions of First Order Logic, Cambridge Tracts in Theoretical Computer Science, vol. 19, Cambridge University Press, 1996.
[30] Y.V. Matiyasevich, Enumerable sets are Diophantine, Soviet Math. Dokl. 11 (1970) 354–358.
[31] J. Mok, A metric characterization of Hilbert spaces, Bull. Korean Math. Soc. 33 (1) (1996) 35–38.
[32] Y.N. Moschovakis, Descriptive Set Theory, Studies in Logic and the Foundations of Mathematics, vol. 100, North-Holland, 1980.
[33] M. Presburger, Über die Vollständigkeit eines gewissen Systems der Arithmetik ganzer Zahlen, in welchem die Addition als einzige Operation hervortritt, in: Sprawozdanie z I Kongresu metematyków slowiańskich, Warszawa 1929, 1930, pp. 92–101, 395, annotated English version by Stansifer [41].
[34] M. Prest, Model Theory and Modules, London Mathematical Society Lecture Notes, vol. 130, Cambridge University Press, 1988.
[35] M.O. Rabin, A simple method for undecidability proofs and some applications, in: Y. Bar-Hillel (Ed.), Logic and Methodology of Sciences, North-Holland, 1965, pp. 58–68.
[36] H. Rogers, Theory of Recursive Functions and Effective Computability, McGraw-Hill, 1967, republished by MIT Press, 1987.
[37] D. Scott, Dimension in elementary Euclidean geometry, in: The Axiomatic Method, with Special Reference to Geometry and Physics, in: Studies in Logic and the Foundations of Mathematics, North-Holland, 1959, pp. 53–67.
[38] S. Shelah, J. Stern, The Hanf number of the first order theory of Banach spaces, Trans. Amer. Math. Soc. 244 (1978) 147–171.
[39] S.G. Simpson, Subsystems of Second Order Arithmetic, Springer-Verlag, 1998.
[40] R.M. Smullyan, Gödel's Incompleteness Theorems, Oxford Logic Guides, vol. 19, Oxford University Press, 1992.
[41] R. Stansifer, Presburger's article on integer arithmetic: Remarks and translation, Technical Report CORNELLCS: TR84-639, Cornell University Computer Science Department, 1984.
[42] A. Tarski, A Decision Method for Elementary Algebra and Geometry, University of California Press, 1951, previous version published as a technical report by the RAND Corporation, 1948, prepared for publication by J.C.C. McKinsey. Reprinted in [11], pp. 24–84.
[43] A. Tarski, A. Mostowski, R.M. Robinson, Undecidable Theories, Studies in Logic and the Foundations of Mathematics, North-Holland, 1953, three papers: 'A General Method in Proofs of Undecidability', 'Undecidability and Essential Undecidability in Arithmetic' and 'Undecidability of the Elementary Theory of Groups'; all but the second are by Tarski alone.
[44] L. van den Dries, J. Holly, Quantifier elimination for modules with scalar variables, Ann. Pure Appl. Logic 57 (2) (1992) 161–179.
[45] H. Weyl, Elementare theorie der konvexen polyeder, Comment. Math. Helv. 7 (1935) 290–306.