Extracting Computable Bounds (and Algorithms) from Classical Existence Proofs: Girard Domains Enable Us to Go Beyond Local Compactness

Vladik Kreinovich[1] and Karen Villaverde[2]
[1]Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA, vladik@utep.edu
[2]Department of Mathematical Sciences
New Mexico State University
Las Cruces, NM 88003, USA, kvillave@nmsu.edu

Abstract

In classical mathematics, the existence of a solution is often proven indirectly, non-constructively, without an efficient method for constructing the corresponding object. In many cases, we can extract an algorithm from a classical proof: e.g., when an object is (non-constructively) proven to be unique in a locally compact space (or when there are two such objects with a known lower bound on the distance between them). In many other practical situations, a (seemingly) natural formalization of the corresponding practical problem leads to a non-compact set. In this paper, we show that often, in such situations, we can extract efficient algorithms from classical proofs – if we explicitly take into account (implicit) knowledge about the situation. Specifically, we show that if we consistently apply Heisenberg's operationalism idea and define objects in terms of directly measurable quantities, then we get a Girard-domain type representation in which a natural topology is, in effect, compact – and thus, uniqueness implies computability.

1 Introduction

In classical mathematics, the existence of a solution is often proven indirectly, non-constructively, without an efficient method for constructing the corresponding object. This non-constructivity was the main starting point for constructive mathematics, in which "there exists an x such as $P(x)$" is only considered to be true if we have an algorithm for constructing such an x.

Numerous papers have shown that in many cases, we can extract an algorithm from a classical proof: e.g., when an object is (non-constructively)

proven to be unique in a locally compact space, or when there are two such objects with a known lower bound on the distance between them, etc.; see, e.g., [27, 28, 29, 32, 33, 34, 41]; see also [46].

In many practical situations, we can use these results and get efficient algorithms. However, in many other practical situations, a (seemingly) natural formalization of the corresponding practical problem leads to a non-compact set. In this paper, we show that often, in such situations, we can extract efficient algorithms from classical proofs – if we explicitly take into account (implicit) knowledge about the situation.

Standard metrics on real numbers, vectors, and other locally compact spaces are usually directly physically motivated. On the other hand, on the (non-locally compact) spaces of higher-order objects (such as the space of all operators), the relation between standard metrics and physical meaning is indirect. We show that if we consistently apply Heisenberg's operationalism idea and define such objects in terms of directly measurable quantities, then we get a Girard-domain type representation [19, 20, 21, 57] in which a natural topology is, in effect, compact – and thus, uniqueness implies computability.

The structure of this paper is as follows. We start by describing the practical need for such solutions (in Section 2) and the algorithmic difficulty of the corresponding general problem (in Section 3). In Section 4, we recall the known results that for locally compact spaces, uniqueness implies algorithmic computability. In Section 5, we formulate the main problem and briefly describe how we plan to solve it. We start our solution with a general description of a measuring instrument: Girard-domain inspired motivations are given in Section 6; Section 7 contains the resulting definitions and their relation with Girard domains. Section 8 uses this formal definition of a measuring instrument to present a general description of a physical quantity – or, to me more precise, of the set of possible values of a physical quantity. A physically important case when the indistinguishability relation is transitive is analyzed in Section 9. In Section 10, we show how to naturally describe functions between thus defined quantities. Finally, Section 11 presents the main result of this paper: that uniqueness implies computability. The final section contains conclusions and plans for future work.

Some of these results were previously announced in [40, 38, 37].

2 Finding Solutions is Practically Important

In many real-life situations, we want to find the *best* decision, the *best* control strategy, etc. The corresponding problems are naturally formalized as *optimization* problems: we have a function $f(x_1, \ldots, x_n)$ of several variables, and we want to find the values $(x_1, \ldots, x_n)$ for which this function attains the largest (or the smallest) possible value.

Many numerical algorithms have been proposed for solving optimization problems. Unfortunately, many of these algorithms often end up in a *local* maximum instead of the desired global one.

- In some practical situations, e.g., in decision making, the use of local maximum simply degrades the quality of the decision but is not, by itself, disastrous.
- However, in some other practical situations, missing a global maximum or minimum may be disastrous.

Let us give two examples:

- In *chemical engineering*, global minima of the energy function often describe the stable states of the system. If we miss such a global minimum, the chemical reactor may go into an unexpected state, with possible serious consequences.
- In *bioinformatics*, the actual shape of a protein corresponds to the global minimum of the energy function. If we find a local minimum instead, we end up with a wrong protein geometry. As a result, if we use this wrong geometry as a computer simulation for testing recommendations on the medical use of chemicals, we may end up with medical recommendations which harm a patient instead of curing him.

For such applications, it is desirable to use *rigorous*, *automatically verified* methods of global optimization, i.e., methods which never discard an actual global maximum; for a survey of such methods, see, e.g., [23, 25].

In some real-life problems, we are not yet ready for optimization, e.g., because the problem has so many constraints that even finding *some* values $x = (x_1, \ldots, x_n)$ of the parameters x_i which satisfy all these constraints is an extremely difficult task. For such problems, we arrive at the problem of satisfying given constraints, e.g., solving a given system of equations. In many such problems, it is important not to miss a solution.

3 Finding Solutions is Algorithmically Difficult

How can we find this corresponding optimum or solution? In classical mathematics, the existence of an optimum or of a solution is often proven indirectly, non-constructively, without an efficient method for constructing the corresponding object.

This non-constructivity was the main starting point for *constructive mathematics*, in which "there exists an x such as $P(x)$" is only considered to be true if we have an algorithm for constructing such an x; see, e.g., [1, 4, 7, 8, 9, 10, 41, 42, 50].

The corresponding research has lead to many useful algorithms. However, in general, the problem is algorithmically undecidable: it is not possible, given a constructively defined system of equations, to algorithmically check whether this system has a solution. For example, it is well known that no algorithm is possible to check solvability of Diophantine equations, i.e., equations of the type $P(n_1, \ldots, n_k) = 0$, where P is a polynomial and $n_1, \ldots, n_k$ are integer-valued

variables. The solvability of this equation is clearly equivalent to the solvability of the following equation with real-valued variables

$$f(x_1, \ldots, x_k) \stackrel{\text{def}}{=} P^2(x_1, \ldots, x_n) + \sum_{i=1}^{k} \sin^2(\pi \cdot x_i) = 0 :$$

indeed, the only way for the left-hand side to be 0 is to have $\sin^2(\pi \cdot x_i) = 0$ – meaning that all x_i are integers – and to have $P(x_1, \ldots, x_k) = 0$. Thus, solvability of systems of equations is indeed algorithmically undecidable.

Similarly, it is not algorithmically possible to tell whether the global minimum of the function f is 0 or not. Thus, in general, global optimization is also not algorithmic.

4 For (Locally) Compact Spaces, Uniqueness Implies Algorithmic Computability: A Reminder

There is one important case in which the existence of a solution automatically leads to its algorithmic computability: the case of unique solutions on a (locally) compact space; see, e.g., [27, 28, 29, 32, 34, 41]. In order to formulate the corresponding result, we must recall some basic definitions of computable ("constructive") real numbers, computable metric spaces, and computable functions; see, e.g., [1, 4, 7, 8, 9, 10, 41, 42, 50]. In addition to presenting definitions, we will also recall motivations for these definitions. These motivations will be helpful in our later discussions.

Comment Readers who are familiar with the usual definitions of constructive real numbers, functions, etc., can skip the following subsections and go straight to the subsection "Uniqueness implies algorithmic computability" that contains the theorem about the computability of unique solutions.

4.1 Computable real numbers and sequences

What is a computable real number? Real numbers are a good model for many real-life quantities such as length, mass, time, etc. In real life, the values of these quantities come from measurements. Measurements are never 100% accurate; hence, each measurement result is an *approximate value* of the real number. Modern measuring instruments record these approximate values as (binary-)rational value. It is therefore reasonable to assume that all these approximations are real numbers.

It makes sense to say that a real number is computable if whatever accuracy we want, we can always (efficiently) get the approximation corresponding to the desired accuracy. An accuracy of k digits means that the corresponding

approximation r_k is 2^{-k}-close to the actual value x, i.e., that $|x - r_k| \le 2^{-k}$. Thus, we arrive at the following definition.

Definition 1. *A real number x is called* computable *if there exists an algorithm (program) that transforms an arbitrary natural number k into a rational number r_k which is 2^{-k}-close to x. It is said that this algorithm* computes *the real number x.*

When we say that a computable real number is given, we mean that we are given an algorithm that computes this real number.

Definition 2. *A sequence of real numbers $x_1, x_2, \ldots, x_n, \ldots$ is called* computable *if there exists an algorithm (program) that transforms arbitrary natural numbers n and k into a rational number r_{nk} which is 2^{-k}-close to x_n. It is said that this algorithm* computes *the sequence x_n.*

When we say that a computable sequence of real numbers is given, we mean that we are given an algorithm that computes this sequence.

4.2 Computable metric spaces and their computable points

A metric space (X, d) is a set on which we have a metric d, i.e., a function which assigns, to every two points $x, x' \in X$, a real number $d(x, x')$ called a *distance* between x and x'. How can we describe points of a computable metric space? Let us recall that a computable real number is described by its (rational) approximations. Similarly, it is reasonable to describe points of an arbitrary set X by their approximations.

From the computational viewpoint, each approximation can be represented in the computer, and thus, is encoded by a finite sequence of 0s and 1s. There are countably many such sequences, so we can describe these approximations as a sequence $x_1, \ldots, x_n, \ldots$ of points of X. Every element of the set X can be approximated, with arbitrary accuracy, by such approximations. Thus, every point from the metric space can be represented as a limit of some subsequence of $\{x_n\}$ – i.e., in topological terms, this subsequence must be *dense* in the original space X.

Thus, we arrive at the following definition:

Definition 3. *By a* computable metric space*, we mean a triple $(X, d, \{x_n\})$, where (X, d) is a metric space, $\{x_1, x_2, \ldots, x_n, \ldots\}$ is a dense subset of X, and there exists an algorithm that, given two natural numbers i and j, computes the distance $d(x_i, x_j)$.*

In other words, we have an algorithm that, given i, j, and an accuracy k, computes the 2^{-k}-rational approximation to $d(x_i, x_j)$. Similar to the previous examples, when we say that a computable metric space is given, we mean that we are given an algorithm that computes $d(x_i, x_j)$.

In particular, the set of all real numbers with a standard metric $d(x, x') = |x - x'|$ and all rational numbers as approximations $\{x_n\}$ is a computable metric space.

Similar to the definition of a computable real number, a computable point x in a metric space can be defined by the existence of an algorithm which returns the corresponding approximations to x:

Definition 4. *A point $x \in X$ of a computable metric space $(X, d, \{x_n\})$ is called* computable *if there exists an algorithm that transforms an arbitrary natural number k into a natural number i for which $d(x, x_i) \le 2^{-k}$. It is said that this algorithm* computes *the point x.*

When we say that a computable point is given, we mean that we are given an algorithm that computes this point.

It is easy to show that a distance between two computable points x and y is computable: indeed, to compute the distance $d(x, y)$ with an accuracy 2^{-k}, it is sufficient to compute the $2^{-(k+2)}$-approximations $\widetilde{x}$ and $\widetilde{y}$ to x and y, and then compute the $2^{-(k+1)}$-approximation $\widetilde{d}$ to the distance $d(\widetilde{x}, \widetilde{y})$. Then, $d(x, \widetilde{x}) \le 2^{-(k+2)}$ and $d(y, \widetilde{y}) \le 2^{-(k+2)}$, hence

$$|d(x, y) - d(\widetilde{x}, \widetilde{y})| \le 2^{-(k+2)} + 2^{-(k+2)} = 2^{-(k+1)}.$$

Since $|\widetilde{d} - d(\widetilde{x}, \widetilde{y})| \le 2^{-(k+1)}$, we thus have

$$|\widetilde{d} - d(x, y)| \le 2^{-(k+1)} + 2^{-(k+1)} = 2^{-k}.$$

4.3 Computable functions

Many real-life quantities x, y are related by an (efficiently computable) functional relation $y = f(x)$. For example, the volume V of a cube is equal to the cube of its linear size s: $V = f(s) = s^3$. This means that, once we know the linear size, we can compute the volume.

At every moment of time, we can only know an approximate value of the actual quality $x \in X$. Thus, to be able to compute $f(x)$ with a given accuracy 2^{-k}, we must:

- be able to tell with what accuracy we need to know x, and then
- be able to use the corresponding approximation to compute $f(x)$.

We thus arrive at the following definition.

Definition 5. *A function $f : X \to X'$ from a computable metric space $(X, d, \{x_n\})$ to a computable metric space $(X', d', \{x'_n\})$ is called* computable *if there exist two algorithms U_f and φ with the following properties:*

- *the algorithm φ takes a natural number k and produces a natural number $\ell = \varphi(k)$ such that $d(x, y) \le 2^{-\ell}$ implies that $d'(f(x), f(y)) \le 2^{-k}$;*

- *U_f takes two natural numbers n and k and produces a 2^{-k}-approximation to $f(x_n)$, i.e., a point x'_ℓ for which $d'(x'_\ell, f(x_n)) \le 2^{-k}$.*

In particular, one can easily show that for every computable point x_0, the function $x \to d(x, x_0)$ – which maps every other point $x \in X$ into a distance to x_0 – is a computable function.

4.4 Computable compact spaces

Let us recall that a metric space X is compact if and only if it is complete and totally bounded; see, e.g., [26, 59]. Here, *complete* means that every converging sequence of points X, i.e., every sequence y_n for which $d(y_n, y_m) \le 2^{-n} + 2^{-m}$ has a limit point in X, and *totally bounded* means that for every $\varepsilon > 0$, there exists a finite ε-net, i.e., a finite set of points $\{z_1, \ldots, z_N\}$ such that every point $x \in X$ is ε-close to one of the points z_i. It can be proven that:

- to check compactness, it is sufficient to check the existence of ε-nets only for some sequence of values $\varepsilon_n \to 0$, in particular, for $\varepsilon_n = 2^{-n}$;
- it is always possible to select an ε-net from the dense subset $\{x_n\} \subseteq X$; and
- to check that a given finite set F is indeed an ε-net, it is sufficient to check that every point x_n from the dense set is ε-close to one of the points of F.

Because of these results, the constructive analogues of the notion of compactness are usually formulated as the possibility to constructively design an 2^{-k}-net for a given k:

Definition 6. *A computable metric space $(X, d, \{x_n\})$ is called* a computable compact space *if there exists an algorithm that, given an arbitrary natural number k, returns a finite set of indices $F_k \subset \{1, 2, \ldots, n, \ldots\}$ such that for every i there is a $f \in F_k$ for which $d(x_i, x_f) \le 2^{-k}$.*

An important feature of computable compact spaces X is that for every computable function $f : X \to R$ from X to real numbers, it is possible to efficiently compute its maximum and its minimum.

Indeed, to compute $M \stackrel{\text{def}}{=} \max f(x)$ with the accuracy 2^{-k}, we must first use the fact that f is computable and find with what accuracy $2^{-\ell}$ we must compute x to be able to estimate $f(x)$ with the accuracy $2^{-(k+1)}$. Then, we use the fact that X is a computable compact space to find a finite $2^{-\ell}$-net. For each point x_i from this net, we compute the $2^{-(k+1)}$-approximation $\widetilde{f}(x_i)$ to $f(x_i)$. Then, $\widetilde{M} \stackrel{\text{def}}{=} \max \widetilde{f}(x_i)$ is the desired 2^{-k}-approximation to $M = \max f(x)$. Indeed, since $f(x_i) \ge \widetilde{f}(x_i) - 2^{-(k+1)}$, we have

$$M = \max f(x) \ge \max f(x_i) \ge \max \widetilde{f}(x_i) - 2^{-(k+1)} = \widetilde{M} - 2^{-(k+1)}.$$

On the other hand, since the values x_i form a $2^{-\ell}$-net, for every value x, there is an x_i for which $d(x, x_i) \le 2^{-\ell}$ and hence $|f(x) - f(x_i)| \le 2^{-(k+1)}$; hence,

$f(x) \le \max f(x_i) + 2^{-(k+1)}$ for all x and $M = \max f(x) \le \max f(x_i) + 2^{-(k+1)}$. Here, $f(x_i) \le \widetilde{f}(x_i) + 2^{-(k+1)}$ so

$$M \le \max \widetilde{f}(x_i) + 2^{-(k+1)} + 2^{-(k+1)}\widetilde{M} + 2^{-k}.$$

4.5 Uniqueness implies algorithmic computability

Now, we are ready to present the desired result.

Theorem 1. [27, 28, 29, 32, 34, 45] *There exists an algorithm U such that:*

- *U is applicable to an arbitrary computable function $f : X \to R$ that attains its maximum on a computable compact space X at exactly one point x,*
- *for every such function f, the algorithm U computes the global maximum point x.*

Comments This result was first proven in [43] for functions of one or several real variables defined on a bounded set. It was extended to general constructive compact spaces in [45]. It is worth mentioning that in [16, 17, 18], some of these results were extended to more general theorems describing a class of mathematical formulas for which classical validity automatically implies constructive (algorithmic) validity.

In Berger et al. [6], the constructive existence of a unique maximum point for a continuous realvalued function on a metric space is investigated in the spirit of reverse mathematics: specifically, this paper proves that a natural unique existence theorem is equivalent to the fan theorem.

The proof below (and the following proofs) uses *Markov's principle* (MP) [10, 42, 50] – according to which if it is not true that an algorithm does not stop, then we conclude that it does stop. It is known that if we prove an algorithm's correctness without using MP, then, from this proof, we can extract explicit bounds on the running time of this algorithm. It is therefore desirable to analyze when similar results hold without the MP. Some such results are given in [27, 28, 29].

Proof. Let us start by explaining the main idea of this proof. We want to compute the global maximum point, i.e., the value x_M for which $f(x_M) = M \stackrel{\text{def}}{=} \max f(x)$. The main idea behind this proof is that since the function f has a unique global maximum M, for every $\varepsilon > 0$, the maximum M_ε of $f(x)$ over all the points which are at distance $\ge \varepsilon$ from x_M is smaller than M. Indeed, the set $\{x : d(x, x_M) \ge \varepsilon\}$ is a closed subset of a compact set and therefore compact itself. The maximum of a continuous function on a compact set is always attained, so if $M_\varepsilon = M$, we would have a point whose distance from x_M is $\ge \varepsilon$ at which the maximum is also attained – thus contradicting the uniqueness.

To use this idea, it is desirable to take a ε-net $z_1, \ldots, z_N$, and compute the maxima M_i of the function $f(x)$ over ε-balls $\{x : d(x, z_i) \le \varepsilon\}$ centered in the

corresponding points z_i; these balls are also closed subsets of a compact set and thus, compact sets themselves. We compute these maxima with sufficient accuracy to distinguish between M and $M_\varepsilon < M$, Then, the balls for which the maximum is guaranteed to be smaller than for some other maximum are excluded as not possible locations of a global maximum. After this computation, all the balls which are completely located within the set $\{x : d(x, x_M) \geq \varepsilon\}$ will be dismissed. The only balls which can still contain global maximum are the ones which contain some points p at a distance $< \varepsilon$ from the (unknown) maximum location x_M. From $d(p, z_i) \leq \varepsilon$ and $d(p, x_M) < \varepsilon$, we now conclude that $d(z_i, x_M) \leq 2\varepsilon$. Thus, for all such balls,

$$d(z_i, z_j) \leq d(z_i, x_M) + d(z_j, x_M) \leq 4\varepsilon.$$

We do not know beforehand what is the difference between M and M_ε, so we do not know with what accuracy we need to compute the corresponding maxima. Thus, what we can do is compute these maxima M_i with higher and higher accuracy – until all the balls in which the maxima can still be located are 4ε-close to each other. Once we achieved this objective, the center z_i of each remaining ball is 2ε-close to the desired maximum location.

This general idea has to be slightly modified to become fully algorithmic. The first – simple – reason why need a modification is that we cannot compute the distance $d(z_i, z_j)$ exactly, we can only compute each distance with a certain accuracy – e.g., the same accuracy ε. In this case, if $d(z_i, z_j) \leq 4\varepsilon$, the ε-approximation to this distance is $\leq 5\varepsilon$.

The second – more complex – reason for the modification is that we need the balls to be computable compacts sets in order to be able to compute the maxima M_i. It is known, however, that while a closed subset of a compact set is always compact, a computable closed subset of a computable compact set is not always a computable compact set. Specifically, if $g : X \to R$ is a continuous mapping from a compact space X into real numbers – e.g., $g(x) = d(x, x_0)$ for a given point $x_0 \in X$ – then, for every real number $\alpha < \max g(x)$, the pre-image $\{x : g(x) \geq \alpha\}$ is also compact.

The following computable version of this result holds for computable functions (see, e.g., [7, 8]): if $g : X \to R$ is a computable mapping from a computable compact space X into real numbers, then, for every two rational numbers $r < r' \leq \max g(x)$, we can algorithmically produce a computable number $\alpha \in [r, r']$ for which the pre-image $\{x : g(x) \geq \alpha\}$ is also constructively compact (and the corresponding 2^{-k}-nets are also algorithmically produced).

In view of this result, for any given $\varepsilon = 2^{-k}$, we produce the ε-net $z_1, \ldots, z_N$. For each point z_i, we find a value $\alpha_i \in [\varepsilon, 2\varepsilon]$ for which the ball

$$B_i \stackrel{\text{def}}{=} \{x : d(x, z_i) \leq \alpha_i\}$$

of radius α_i with a center in z_i is a computable compact set.

For these sets, the fact that $M_i > M_\varepsilon$ means that the set

$$B_i \subseteq \{x : d(x, z_i) \leq 2\varepsilon\}$$

contains a point which is ε-close to x_M – and thus, that $d(z_i, x_M) \le 3\varepsilon$. So, for every two such sets, we have $d(z_i, z_j) \le 6\varepsilon$, and thus, an ε-approximation $\widetilde{d}(z_i, z_j)$ to $d(z_i, z_j)$ is smaller than or equal to 7ε.

Since each set B_i is a computable compact, we can compute the maximum $M_i \stackrel{\text{def}}{=} \max\{f(x) : x \in B_i\}$ of the function $f(x)$ over each ball B_i with arbitrary accuracy $2^{-\ell}$, i.e., we can compute rational numbers $m_{i\ell}$ which are $2^{-\ell}$-close to the actual M_i.

If for some i and j, we have $m_{i\ell} < m_{j\ell} - 2 \cdot 2^{-\ell}$, this means that $M_i < M_j$ and thus, the global maximum cannot be attained at B_i. Vice versa, if $M_i < M_j$, this means that $m_{i\ell} < m_{j\ell} - 2 \cdot 2^{-\ell}$ for sufficiently large ℓ and thus, after computing M_i and M_j with a sufficient accuracy, we will be able to confirm that $M_i < M_j$.

So, to find the desired approximation to x_M, we repeatedly, for $\ell = 1, 2, \ldots$, compute the values $m_{i\ell}$ and dismiss some balls until for all non-dismissed balls, we have $\widetilde{d}(z_i, z_j) \le 7\varepsilon$, where $\widetilde{d}(z_i, z_j)$ is the distance $d(z_i, z_j)$ computed with accuracy ε. At this stage, z_i is the desired 3ε-approximation to x_M.

We can perform this procedure for every ε, hence we have an algorithm for producing desired approximations to x_M. The theorem is proven. □

Similar results hold for roots (solutions) of a system of equations:

Definition 7. *By a* computable system of equations *we mean a system $f_1(x) = 0, \ldots, f_k(x) = 0$, where each of the functions f_i is a computable function on a computable compact set X.*

Theorem 2. [27, 28, 29, 32, 34] *There exists an algorithm U such that:*

- *U is applicable to an arbitrary computable system of equations which has exactly one solution, and*
- *for every such system of equations, the algorithm U computes its solution.*

Proof. This theorem immediately follows from our optimization result if we notice that the solution to the system is exactly the point x where a computable function $F(x) \stackrel{\text{def}}{=} f_1^2(x) + \ldots + f_k^2(x)$ attains its minimum – or, equivalently, where the negative function $f(x) \stackrel{\text{def}}{=} -F(x)$ attains its maximum. □

Comment. It is worth mentioning that in both results, uniqueness is important. For example, no algorithm U is possible:

- that is applicable to an arbitrary computable function $f : X \to R$ that attains its maximum on a computable compact space X at exactly two points x, and
- that computes one of the corresponding global maximum points x.

Similarly, no algorithm U is possible

- that is applicable to an arbitrary computable system of equations $f_1 = 0$, $\ldots$, $f_k = 0$ that has exactly two solutions x on a computable compact space X, and
- that computes one of the corresponding solutions x.

For proofs, see, e.g., [31, 32, 33, 34, 35, 36, 39, 41, 46].

5 Remaining Problem and What We Will Do

In many practical situations, we can efficiently apply the uniqueness-motivated algorithms to concrete mathematical problems; see, e.g., [29].

However, in many other practical situations, a (seemingly) natural formalization of the corresponding practical problem leads to a non-compact set. For example, many physical phenomena are described by *fields*, i.e., by functions. The space of all functions is not locally compact.

In quantum physics, possible states of a quantum system form a Hilbert space – which is not locally compact. Quantum observables are operators, and the space of all operators is also not locally compact. It is desirable to extend the uniqueness result to such situations.

Some extensions of this type are known. For example, in [51], it is proven that if a computable function $f : X \to R$ on a complete metric space has *uniformly at most one minimum*, i.e., if for every $\delta > 0$, there is an $\varepsilon > 0$ for which $f(x) < \inf f + \varepsilon$ and $f(y) < \inf f + \varepsilon$ implies $d(x, y) < \delta$, then we can efficiently find this minimum. In other words, if, in addition to an algorithm for computing f, we also have an algorithm that given δ produces the appropriate ε, then we can compute the location of the minimum.

In this paper, we consider a general situation, in which the algorithm producing ε from δ is not known.

Specifically, we show that often, in such situations, we can extract efficient algorithms from classical proofs – if we explicitly take into account (implicit) knowledge about the situation. Specifically, while the standard metrics on real numbers, vectors, and other locally compact spaces are usually directly physically motivated, on the (non-locally compact) spaces of higher-order objects (such as the space of all operators), the metric is more mathematically motivated.

We will show that if we consistently apply Heisenberg's operationalism idea and define such objects in terms of directly measurable quantities, then we get a Girard-domain type representation [19, 20, 21, 57] in which a natural topology is, in effect, compact – and thus, uniqueness implies computability.

Specifically, according to the operationalist idea, a quantity is what can be measured by measuring instruments. Each measuring instrument is imprecise, so based on a single instrument, we only have an approximate idea of the value of the quantity. To get a more adequate description, we must consider a sequence of more and more accurate measuring instruments. In the following sections, we will see how these ideas lead to a definition.

6 Towards a Description of a Measuring Instrument: Motivations

Plan In accordance with the above description, let us start with a description of a measuring instrument.

Every measuring device has only finitely many possible outcomes such as marks on a scale (for an analog device) or digital (usually binary) outputs (for a digital device). Not all the marks may be physically possible: for example, the velocity cannot exceed the speed of light, etc. To distinguish between physically possible and physically impossible marks, we need to have a *theory* that describes both this device and the measured quantity. We will show that, under reasonable assumptions, the existence of such a theory leads to an *algorithm* for producing a complete list of all physically possible outcomes. This list is the first step in describing a measuring device.

We have mentioned that measuring instruments are imprecise. Due to measurement uncertainty, when we apply the same measuring device to the same object twice, we may get different measurement results. Some pairs of possible outcomes (x, y) can therefore appear when measuring the same object twice. The information of which pairs can thus occur forms the second step in our description.

We need a theory The desired distinction between physically possible and physically impossible marks – i.e., crudely speaking, between physically possible and physically impossible values of the corresponding physical quantity – can only come from physics. In other words, to be able to provide such a distinction, we need a *physical theory*.

We need a theory that also describes a measuring device We need a theory that would describe both the physical quantity (or quantities) measured and/or influenced by the measuring device and the measuring device itself.

Why is it not enough to have a theory that describes the actual physical quantity? Because the measured value depends not only on the quantity but also on a measurement error introduced by the measuring device. For example, if we measure the temperature with a measuring device that has an accuracy 1 degree, then, in principle, it is still physically possible to have the measurement result -274 – when the actual measured temperature is close to the absolute zero value -273; however, the measured value -280 is impossible.

On the other hand, a thermometer with an accuracy ± 10 degrees can produce a reading of -280 – when the actual temperature is close to the absolute zero.

We want a theory that is "full" in some natural sense Our ultimate objective is to provide solutions to practical problems in which data processing is needed, problems related to engineering and applied science. In engineering and applied science, when we say that there is, say, a theory of flight, this usually means that there is a reasonably full (not partial) description of what can fly and

what cannot. There may be some computational difficulties in deciding how well a given complex airplane will fly, difficulties requiring the use of sophisticated supercomputers, but in principle, the problem is solved.

In other words, in engineering and applied science, by a theory, we usually mean a theory that is "full" in some natural sense. This is what we will mean by a theory in this paper.

Comment It is important to mention that in theoretical physics, the word "theory" also has a different meaning. When we say that, e.g., Einstein produced a theory of gravitation, it does not necessarily mean that his theory is capable of solving all problems related to gravitation. It is known that physical theories often have limitations beyond which they are not applicable; e.g., Einstein's theory of gravitation (General Theory of Relativity) cannot adequately describe the gravitational interaction between quantum objects. In other words, theories in fundamental physics are often *partial*.

There is a reason why in engineering, theories are usually full, while in fundamental science, theories are usually partial. Indeed, the class of objects that are of interest for an engineering theory consists of objects that are built by us; we have control over which object to build and to design – and we are usually limiting ourselves to objects whose behavior we can predict. If a physical process is not yet well known and well studied, and its results are therefore unpredictable, an engineer would not use this process in designing, say, an engine. In contrast, in general physics, we want to study objects given by nature; we do not have a choice to dismiss some processes just because we do not know how they function.

Thus, it makes sense to assume that in engineering and applied science, we have a full theory – while in fundamental physics, theories are rarely full in this sense.

A seemingly natural definition of a full theory is not always adequate At first glance, it may look like a full theory is a one that provides an answer to all relevant questions. This is exactly what is meant by a "decidable theory" in mathematical logic: a theory in which, for every statement, either this statement or its negation are deducible from the theory.

If we define fullness in this manner, then, for example, if we want to know whether a given mark on a scale of the measuring device is physically possible, then this theory should be able to tell us whether it is physically possible or not.

Unfortunately, this seemingly natural definition does not always capture the physical meaning of the notion of a full theory, the meaning that we are trying to capture in this section. Let us explain the problem with the standard definition on the example of a theory in which there is an upper (or lower) bound on the physical quantity – e.g., the speed of light is the upper bound on velocity, and the absolute zero temperature is the lower bound for temperature.

From the physical viewpoint, when we say that a theory is full, we mean, in

particular, that this theory should enable us to compute the exact value of this bound b.

What does it mean to compute b? As we have mentioned, what people usually understand by computing a real number is that for any given accuracy 2^{-k}, we must be able to produce a rational number r_k that is 2^{-k}-close to b; in other words, that we have an algorithm that, given k, produces a rational number r_k for which $|r_k - b| \le 2^{-k}$.

A full theory should produce an algorithm for computing the bound b, i.e., b should be a computable real number. Similarly, a full theory should compute the exact upper bound Δ on the measurement error of this measuring instrument; hence, Δ should also be a computable real number.

Once we know b and Δ, which marks are physically possible? Without losing generality, let us consider the case of the upper bound. By definition of the measurement error Δx as the difference $\Delta x \stackrel{\text{def}}{=} \widetilde{x} - x$ between the measured and the actual values of the physical quantity, the measured value $\widetilde{x}$ can be represented as the sum of the actual value x and the measurement error Δx. Thus, the largest possible value of $\widetilde{x}$ corresponds to the case when both x and Δ attain their largest possible values. The largest possible value of x is b, and the largest possible value of Δx is Δ. Therefore, the largest possible value of $\widetilde{x}$ is equal to $b + \Delta$.

It is known that the sum of computable real numbers is computable, so $b + \Delta$ is a computable real number. Suppose now that a given mark on a scale corresponds to the value v of the measured physical quantity; usually, this value v is a rational number. Then, v is physically possible if and only if $v \le b + \Delta$. Thus, to be able to tell whether a given mark on a scale is physically possible or not, we must be able to tell whether a given computable number (in our case, $b + \Delta$) is larger or equal than a given rational number (in our case, v).

Alas, it is known that this problem is, in general, not algorithmically solvable [1, 4, 7, 8, 9, 10, 41, 42, 50]. So, we cannot simply assume that a "full" theory will always tell us which marks are physically possible and which are not. Instead, we must provide a new definition of a full theory and analyze what this definition entails.

Before we do that, let us first describe what we mean by a "theory" in the first place.

What exactly is a theory In general, in mathematics, to describe a theory means to describe a finite list of its basic statements (called *axioms*) and a finite list of *deduction rules* that enable us to deduce new statements from the ones that we have already proven; see, e.g., [3, 13, 44, 52].

Once we have such a theory, we can algorithmically produce all the statements deducible in this theory: we start with the axioms and apply deduction rules again and again. In this way, we can first produce all the statements that can be derived by a single application of deduction rules, then all the statements that can be produced by two applications of deduction rules, etc. Eventually, each deducible statement will thus be produced.

In theory of computing, a set whose elements can be enumerated by an algorithm is called *recursively enumerable* (r.e., for short). So, when we say that we have a theory, we mean that the set of all statements deduced from this theory is r.e.

R.e. sets S are also called *semi-decidable* (see, e.g., [2, 44]) because the enumerating algorithm provides us with a "semi-decision" procedure for deciding whether a given element x belongs to the set S: We simply generate all the elements of the set S one by one, and compare the given element x with each newly generated element from S. Once we find an element from S that is equal to x, we stop and return the answer "yes". Clearly, if this algorithm returns the answer "yes", it means that $x \in S$; vice versa, if $x \in S$, then this procedure always returns the answer "yes". This procedure is a "semi-decision" procedure because whatever answer it produces is correct, and it is guaranteed to provide a correct answer when the element x actually belongs to S, but it is not required to produce any answer when x is not actually in S.

What kind of statements are we allowing A physical theory usually describes properties of physically possible states and their transitions: e.g., when we describe the time changes, we must be sure that the energy is preserved, that the moments are preserved, that the overall electric charge is preserved, etc.

If a transition from a given state s into another given state s' is physically *impossible*, we will eventually find it out – by testing all preservation properties.

On the other hand, if the transition is physically *possible* – in the sense that all conserved quantities have the same values in s and s' and therefore, hopefully, there exists an operation that actually transforms s into s' – then we may not be able to deduce this possibility from the theory: there may be infinitely many conserved quantities and during a finite period of time (no matter how long), we can only test finitely many of them, leaving others un-tested.

Similarly, a theory usually describes conditions under which a mathematically defined state is physically possible: the velocity of all the particles cannot exceed the speed of light, the electric charge of each particle must be equal to a whole number of charge quanta, the temperature in each region cannot be smaller than the absolute zero, etc. If a mathematically defined state s is physically *impossible*, we will eventually find it out – by testing all the conditions that a physically possible state must satisfy.

On the other hand, if a mathematically defined state s is physically *possible* – in the sense that all the necessary conditions are satisfied and therefore, hopefully, there exists an operation that actually generates this state s – then we may not be able to deduce this possibility from the theory: there may be infinitely many conditions and during a finite period of time (no matter how long), we can only test finitely many of them, leaving others un-tested.

In our case, we talk about possible post-measurement states of a measuring device. From this viewpoint, it is reasonable to assume that the set of statements (potentially) covered by a theory includes statements of the type "an outcome v is physically impossible". (It may also happen that a particular theory is able

to deduce statements of the type "a mark v is possible" as well, but since we want the most general description, we will not require these statements to be covered by the theory.)

We want the theory to be *correct*; therefore, if the theory implies that an outcome v is not physically possible, then v should indeed be a physically impossible outcome. Because of this, if a theory implies that an outcome v is physically impossible, we will call such an outcome *provably impossible*.

As a result, we conclude that *the set of provably impossible outcomes is r.e. (semi-decidable).*

What exactly is a full theory Intuitively, a physical theory is full if whatever is not prohibited by this theory actually occurs in real life.

In other words, we require that *if something never happens in real life, the theory must have an explanation for it not happening*, i.e., in precise terms, the theory entails the statement that this particular event never happens.

In other words, if some transformation is never observed, and there is no known law that would prevent this transformation from happening, the physicists usually imply that there must be a (yet unknown) additional law that prevents such transformations – i.e., in other words, that the existing theory is not full in the physical sense; see, e.g., [14].

For example, from the viewpoint of the chemical composition, it is possible for water to dissolve into hydrogen and oxygen: $2H_20 \to 2H_2 + O_2$; however, if we place water in a bowl, it will never dissolve by itself. A physicist will thus conclude that since this dissolution never happens, there must be an explanation for this – and there actually is such an explanation: the total energy of two water molecules $2H_20$ is smaller than the sum of the energies of two hydrogen molecules $2H_2$ and one oxygen molecule O_2, so this transformation is prohibited by the energy conservation law.

Another example: from the viewpoint of pure particle mechanics, nothing prevents all the air molecules in a room to concentrate in one half of it, leaving vacuum in the other half. However, this never happens. This means that we must have a physical explanation why it never happens, and there is such an explanation – the second law of thermodynamics that prohibits such transformations.

In short, physicists believe that if some event never happens, there must be a theoretical explanation for this event never happening. So, if a theory T provides a full description of a given physical phenomenon, and some event e never happens, this means that the theory T must entail a statement stating that this event e never happens.

How does this apply to measuring devices? Suppose, for example, that we have a measuring device with a scale. Suppose that we use this measuring device again and again and a certain mark never occurs as a result of this measurement. Example: we measure the velocity of different particles from the cosmic rays, and the value $v =$400,000 km/s never occurs. Since this mark never occurs, a full theory T must have an *explanation* for it, i.e., in precise terms, this theory

T must imply that this mark never occurs.

The existence of a full theory makes the set of all physically possible outcomes decidable Let us show that the existence of a full theory T enables us to tell, for each outcome v, whether v is physically possible or not.

We want to know whether v will occur in one of the measurements, or whether it never happens.

If the outcome v never happens, then, according to our "definition" of a full theory, T must imply that v never happens. So, if we start with the axioms of the theory T and generate all its conclusions one by one (as we described earlier), we will eventually come up with the statement that v never happens.

If the outcome v will occur, then we can apply the measuring device to all possible objects, and eventually, we will get v. How can we apply the measuring device to all possible objects? We can describe each possible object of measurement by a text in a natural language, a text instructing us how we can find (or generate) this object. We can enumerate all such texts, and apply all such instructions. To be more precise, we can apply the instructions from the first such text for 1 hour, then the instructions from the first two texts for 2 hours, etc. Eventually, we will thus implement each possible object.

So, if we apply the measuring instrument to different objects generated in this way, we eventually get v. (If, by a measuring device, we mean a population of sensors, then we must apply all possible sensors from this population to every object.)

Let us now show how we can check whether v happens or not. We have two algorithmic processes:

- the process that generates all possible statements proven in a theory T, and
- the process that generates all possible objects of measurement.

To check whether an outcome v is possible or not, we simultaneously launch both algorithmic processes and wait until v appears in one of them:

- If an outcome v never occurs (i.e., is physically impossible), then the statement that v never occurs will appear in the first process;
- If an outcome v will occur (i.e., is physically possible), then v will appear as a result of the second process.

Thus, eventually, we will have v as a result of one of these processes – and depending on which one, we will be able to tell whether this outcome v is physically possible or not.

Comment In mathematical terms, the existence of the first algorithmic process means that the set of all physically *impossible* outcomes is r.e. (semi-decidable). The existence of the second algorithmic process means that the set of all physically *possible* outcomes is also r.e. (semi-decidable). It is well known (see, e.g.,

[44]) that if a set is r.e. and its complement is r.e., then this set is algorithmically decidable, i.e., there exists an algorithm that checks, for each outcome v, whether this outcome is physically possible or not. This is exactly what we have just shown.

Conclusion: decidable set of physically possible outcomes We have shown that, under reasonable assumptions, the existence of a full theory leads to an algorithm for producing a complete finite list of all physically possible outcomes. This list is the first step in describing a measuring device.

Comment Since we can always algorithmically check whether each outcome is possible or not, we can always assume that the list of physically possible outcomes is given from the very beginning. Thus, in the following text, we will only consider physically possible outcomes.

How do we describe uncertainty: main idea So, we have a set of possible outcomes. On this level, how do we describe uncertainty? Due to measurement uncertainty, when we apply the same measuring device to the same object twice, we may get different measurement results.

This is a known fact for anyone who has ever measured anything with a real measuring device, be it current or voltage or temperature.

For example, suppose that we measure temperature by using a measuring device with an accuracy of ± 2 degrees. If the actual temperature is, say, 36.6, then, due to measurement errors, the measurement result can take any value from $36.6 - 2 = 34.6$ to $36.6 + 2 = 38.6$. If this measuring instrument has marks corresponding to $0, 1, 2, \ldots, 100$ degrees, then for the actual temperature of 36.6, the only possible outcomes are the ones that lie within the interval $[34.6, 38.6]$, i.e., the marks 35, 36, 37, and 38.

Some pairs of outcomes are compatible (close), some are not We have already mentioned that due to measurement uncertainty, multiple measurement of the same object can lead to different outcomes.

For example, if we measure temperature with the accuracy of 2 degrees, and possible outcomes are 0, 1, 2, etc., then the outcome 0 can happen if the actual temperature is in the interval $[-2, 2]$; the outcome 1 can happen if the actual temperature is in the interval $[-1, 3]$, etc.

When the outcomes are very different, e.g., 0 and 5, the corresponding intervals $[-2, 2]$ and $[3, 7]$ do not intersect – which means that there is no actual value for which repeated measurements would produce 0 and 5.

When the outcomes are close, e.g., 1 and 3, then the corresponding intervals $[-1, 3]$ and $[1, 5]$ have a non-empty intersection – which means that there are actual values for which repeated measurements can produce 1 and 3 – namely, all the values from this intersection.

It is therefore natural to say that the physically possible outcomes x and y are *compatible*, or *close* if for some object, multiple measurements can lead to x

and y. We will denote compatible outcomes by $\sim$. For example, for the above thermometer, $1 \sim 3$ but $0 \not\sim 5$.

We can also say that in this case, the outcomes $x \sim y$ are *indistinguishable* by the given measuring device: if we got x as an outcome for one object, and y as an outcome for another object, then we cannot guarantee that these objects are different: they may be identical.

Historical comment The notion of indistinguishability was first proposed by H. Poincaré in [48]. In the modern times, it was re-introduced by E. C. Zeeman in [60]. For an overview of known results about this notion, see, e.g., [49, 53, 54, 55].

The existence of a full theory makes the set of all compatible pairs of outcomes decidable How do we tell which pairs of outcomes are compatible and which are not? As in the description of possible outcomes, the only way to distinguish between compatible and non-compatible pairs is to have a physical theory that describes both the measured quantity and the measuring device.

Similarly to the description of possible outcomes, we can argue that this theory should be able to make predictions about incompatibility of different pairs. We can also argue that this theory should be full in the sense that if some pair (x, y) never occurs in a repeated measurement, then the theory should have an explanation for it never occurring, i.e., the theory should entail that this pair never occurs. If it does occur, then we will be able to see it in one of the actual measurements.

Similarly to the description of possible outcomes, we can therefore conclude that the set of all compatible pairs is decidable.

Conclusion: decidable set of compatible pairs of outcomes We have shown that, under reasonable assumptions, the existence of a full theory leads to an algorithm for producing a complete finite list of all compatible pairs of outcomes. Thus, at the current (second) step of describing a measuring device, we can describe it as a pair $\langle X, \sim \rangle$, where X is a finite set (of all physically possible outcomes), and $\sim$ is a symmetric and reflexive relation on X describing which outcomes are compatible.

7 Resulting Description of a Measuring Instrument and Its Relation to Binary Girard Domains

From the mathematical viewpoint, a pair $\langle X, \sim \rangle$ is nothing else but a finite *graph*. In domain theory, graphs – also called *webs* – form the basis of *binary Girard domains*; see, e.g., [19, 20, 21, 57]. From this viewpoint, we describe measuring devices by using binary Girard domains.

Definition 8. *A* binary domain *is a pair $\langle X, \sim\rangle$, where X is a finite set, and $\sim$ is a symmetric reflexive relation on X.*

This description is useful because many constructions relevant for measuring instruments can be naturally described in terms of Girard domains – e.g., if we consider a pair of independent measuring instruments $\langle X, \sim\rangle$ and $\langle Y, \sim\rangle'$ as a single measuring device, then its set of possible outcomes is the Cartesian product $X \times X'$ and the corresponding compatibility relation $(x, y) \sim (x', y')$ takes the form "$x \sim x'$ and $y \sim' y'$" – i.e., we have a tensor product of the corresponding binary Girard domains.

Comment In graph terms, a graph in which vertices correspond to sets and two vertices are connected by an edge if and only if the corresponding sets intersect is called an *intersection graph* [22, 47]. For the case when sets are intervals, the intersection graph is called an *interval graph*; see, e.g., [15].

Examples Let us first show how the standard interval uncertainty fits into this general picture. For example, suppose that we measure temperature with the accuracy 1°, and the scale consists of the values $X = \{0, 1, 2, 3 \ldots, T\}$. Here, e.g., $\tilde{t} = 0$ means that the actual temperature t is in the interval $[-1, 1]$; so, $a \sim b$ if the corresponding intervals $[a - 1, a + 1]$ and $[b - 1, b + 1]$ intersect, i.e., if $|a - b| \le 2$.

An even simpler example comes from counting. Every actual counting device has a limitation of how many objects we can count, so here, $X = \{1, 2, \ldots, n, \text{many}\}$, where $1, 2, \ldots, n$ are counting results, and the additional result "many" indicates that we have exhausted the given device, and there are still objects to count. For the results from this set X, $a \sim b$ if and only if $a = b$.

Yet another example comes from "yes"-"no" questions; here, possible results are "false" (usually denoted by 0), "true" (usually denoted by 1), and "unknown" (we will denote it by U). Here, $X = \{0, 1, U\}$, and the relation $\sim$ has the form $0 \sim U$ and $U \sim 1$. Indeed, if we do not know the truth value (U), then, in reality, the answer may turn out to be "true" (hence $U \sim 1$) or "false" (hence, $U \sim 0$). However, once we know the answer "true", we cannot get the answer "false", hence $0 \not\sim 1$.

8 From a Description of a Single Measuring Instrument to a General Description of a Physical Quantity

A single measurement only leads to an *approximate* value of the measured quantity. To describe the *actual* value of the measured quantity, we must consider a sequence of more and more accurate measuring instruments.

Let X_k describe the results of first k measurements. Then, for every $k < l$, there exists a natural "forgetful functor" $\pi_{lk} : X_l \to X_k$ that simply erases the

results of the last $l - k$ measurements.

What are the natural properties of this mapping? If a' and b' can occur at the same object, i.e., if $a' \sim' b'$, then for this same object, the partial results $\pi(a')$ and $\pi(b')$ are also possible, i.e., $\pi(a') \sim \pi(b')$. Similarly, if partial results a and b can occur for the same object, i.e., if $a \sim b$, then by performing additional measurements, we get some values $a' \sim' b'$ for which $\pi(a') = a$ and $\pi(b') = b$. Thus, we arrive at the following definition.

Definition 9. *Let $\langle X, \sim\rangle$ and $\langle X', \sim'\rangle$ be finite domains. A mapping $\pi : X' \to X$ is called a* projection *if it satisfies the following two properties:*

(i) if $a' \sim' b'$, then $\pi(a') \sim \pi(b')$;

(ii) if $a \sim b$, then $\exists a', b'$ such that $\pi(a') = a$, $\pi(b') = b$, and $a' \sim' b'$.

A physical quantity can be described by a sequence of more and more accurate measurements $X_1 \stackrel{\pi_{2,1}}{\leftarrow} X_2 \stackrel{\pi_{3,2}}{\leftarrow} X_3 \stackrel{\pi_{4,3}}{\leftarrow} \ldots$ A value of a quantity is described by the measurement results $x_i \in X_i$ corresponding to different measurements – these results have to all be compatible, i.e., we must have $\pi_{lk}(x_l) = x_k$ for all $k < l$.

It is possible that two sequences

$$a = (a_1, a_2, \ldots, a_n, \ldots)$$

and

$$b = (b_1, b_2, \ldots, b_n, \ldots)$$

describe the exact same value – when all intermediate results are compatible: $a_i \sim_i b_i$ for all i. (This is similar to the fact that $0.999\ldots$ and $1.000\ldots$ is the same real number.)

We thus arrive at the following definition.

Definition 10.

- *By a* (physical) quantity*, we mean a sequence X of binary domains $\langle X_i, \sim_i \rangle$ with projections: $X_1 \stackrel{\pi_{2,1}}{\leftarrow} X_2 \stackrel{\pi_{3,2}}{\leftarrow} X_3 \stackrel{\pi_{4,3}}{\leftarrow} \ldots$*
- *A quantity is called* computable *if there exists an algorithm which generates the corresponding domains and projections.*
- *By a* value *of a quantity X, we mean a sequence $x = (x_1, x_2, \ldots)$, where $x_i \in X_i$ and $\pi_{lk}(x_l) = x_k$ for all $k < l$.*
- *A value $x = (x_1, x_2, \ldots)$ of a computable quantity X is called* computable *if there exists an algorithm which, given i, returns x_i.*
- *The set X of all values is called a* projective limit *of the sequence X_i.*
- *Two values $a, b \in X$ are called* indistinguishable *if $a_i \sim_i b_i$ for all i.*

Comment Projective limits of finite graphs have been introduced and analyzed by M. Smyth; see, e.g., [55].

On the set X of possible values of the quantity, we can naturally define neighborhoods and limits:

Definition 11.

- *By a* neighborhood *of a value x we mean a set $N_n(x) \stackrel{\text{def}}{=} \{b \,|\, b \sim_n x\}$ for some integer n.*
- *If $a^{(k)}$ is a sequence of elements from X, then we say that $a^{(k)}$* tends to a limit *a (denoted $a^{(k)} \to a$) if $\forall n\, \exists m\, \forall k > m\, (a_n^{(k)} \sim_n a)$.*

Examples It is easy to see that if we consider interval-related sets $X_i = \{-p_i/q_i, -(p_i-1)/q_i, \ldots, (p_i-1)/q_i, p_i/q_i\}$ (where $k/q_i \sim_i (k+1)/q_i$) with a better and better accuracy ($q_i \to \infty$) and broader and broader span ($p_i/q_i \to \infty$), then the corresponding set X is the set of all real numbers (+ two extra values $-\infty$ and $+\infty$) with a natural topology.

If we start with n-dimensional "boxes", we naturally end up with the set R^n.

For "yes"-"no" questions, if one measurement does not lead to a definite answer (i.e., if the answer is U), we can perform a more accurate measurement; as a result, we may get a definite answer, i.e., we may get a sequence of answers $U0$ or $U1$, or we may still get "unknown" – i.e., the sequence UU. So, after two measurements, we have 5 possible results: $X_2 = \{0, 1, U0, U1, UU\}$. In the set X_2, all "yes" answers (0 and $U0$) can happen in the same state, so $0 \sim U0$; it is also possible that in the same state, sometimes, the answer is "yes", and sometimes, the answer is still unknown, so $0 \sim UU$ and $U0 \sim UU$. The natural projection from X_2 to X_1 simply deleted the second answer: e.g., $\pi_{2,1}(U0) = U$. So, here:

- X_1: $0 \sim_1 U$, $U \sim_1 1$;
- X_2: $0 \sim_2 U0$, $0 \sim_2 UU$, $U0 \sim_2 UU$, $1 \sim_2 U1$, $1 \sim_2 UU$, $U1 \sim_2 UU$;

etc. Thus, the projective limit X consists of three different elements 0, 1, and U, with the relation $0 \sim U$ and $U \sim 1$.

It is easy to show that in some sense, every such limit is compact:

Proposition 1. *For every projective limit X, every sequence $a^{(k)}$ has a convergent subsequence.*

Proof Since the set X_1 is finite, and there are infinitely many elements $a^{(k)}$ in the sequence, there exist at least one value $x \in X_1$ for which infinitely many elements $a^{(k)}$ have $a_1^{(k)} = x$. We can therefore consider a subsequence consisting of such elements. Let us fix the first element in this new subsequence. There are infinitely many elements in the remaining part of the subsequence, and only finitely many elements in X_2. Thus, we can select a sub-subsequence in which all elements but one have the same value of a_2, etc. As a result, we get a convergent subsequence.

Discussion For example, for real numbers, instead of the set R, we have a compactification $R \cup \{-\infty, +\infty\}$.

Compactness is important for solving *inverse problems*; see, e.g., [56]. The main reason why we have measurements is that we want to reconstruct the actual values of the measured quantities. In general, we observe $f(x)$ for some continuous $f : X \to Y$, and we want to reconstruct x. For example, we want to reconstruct an image x, but what we observe is an image $f(x)$ distorted by the inaccuracies of the lens. The problem is that even in the presence of noise, when the mapping f is 1-1, the function f^{-1} is often discontinuous, so a small measurement error y can lead to a large error in reconstructing x. A known solution is to restrict ourselves to compact sets X because for compact sets, the inverse f^{-1} to a continuous mapping is continuous as well.

The problem is that, e.g., the set X of all images is not compact under standard mathematical metrics such as L^2 or L^∞. Our result shows that this set is compact if we consider a topology that naturally comes from measurements.

Let us illustrate this result on a simplified version of quantum physics. In quantum physics (see, e.g., [14]), states are elements of the complex-valued L^2-space, e.g., of the space of all complex-valued functions $\psi(x)$ defined on the 3-D space R^3 for which $\int |\psi(x)|^2\,dx = 1$. A natural distance $d(\psi, \psi')$ on this space can be defined as $d^2(\psi, \psi') = \int |\psi(x) - \psi'(x)|^2\,dx$.

To be more precise, in quantum physics, for every real number α, the functions $\psi(x)$ and $e^{\mathrm{i}\cdot\alpha} \cdot \psi(x)$ (where $\mathrm{i} \stackrel{\mathrm{def}}{=} \sqrt{-1}$) represent the same physical state. Thus, a physically meaningful distance $d(\psi, \psi')$ between the physical states represented by the functions $\psi(x)$ and $\psi'(x)$ is usually defined as $d'(\psi, \psi') = \min\limits_{\alpha,\alpha'} d(e^{\mathrm{i}\cdot\alpha} \cdot \psi, e^{\mathrm{i}\cdot\alpha'} \cdot \psi')$.

From the physical viewpoint, each measurable quantum quantity can be described by an operator $A : L^2 \to L^2$. After an (ideal) measurement, the state transforms into one of the eigenvectors $e_1(x), \ldots, e_n(x), \ldots$ of this operator, and the corresponding eigenvalue λ_i is returned as a measurement result. The eigenvectors form an orthonormal basis in L^2, and the probability of returning λ_i is equal to $|\psi_i|^2$, where ψ_i is the coefficient in the expansion of the original state w.r.t. this basis: $\psi(x) = \sum\limits_{i=1}^{\infty} \psi_i \cdot e_i(x)$.

The L^2-distance between the states $\psi = (\psi_1, \psi_2, \ldots)$ and $\psi' = (\psi'_1, \psi'_2, \ldots)$ takes the form $d^2(\psi, \psi') = \sum |\psi_i - \psi'_i|^2$. The coefficients satisfy the condition $\sum\limits_{i=1}^{\infty} |\psi_i|^2 = 1$, hence $|\psi_i| \le 1$ for all i.

To avoid complexities related to the equivalence between $e^{\mathrm{i}\cdot\alpha} \cdot \psi(x)$ and $\psi(x)$, let us restrict ourselves to the states in which all the values ψ_i are non-negative real numbers (i.e., for which $0 \le \psi_i \le 1$). In this case, each value ψ_i coincides with $|\psi_i|$ and can thus be directly determined by measuring the corresponding probability $p_i = |\psi_i|^2$.

The above description corresponds to idealized measurements. In practice, as we have mentioned, each measuring instrument has limit on the values that it can measure. Thus, based on its measurements, we only get information

about the states for which the corresponding eigenvalues are within these limits. Usually, within each interval, there are only finitely many eigenvalues. So, based on each measurement, we only get information about finitely many values $|\psi_i|^2$.

Also, due to measurement errors – and due to the finite size N of the sample – we can only get an approximate value of each probability $|\psi_i|^2$. In statistical estimates, the accuracy of determining a statistical characteristic based on a sample of size N is $\approx 1/\sqrt{N}$.

To get more and more accurate description of the state, we can take measuring instruments that cover larger and larger range – and at the same time, we increase the sample size to get more and more accurate values of the probabilities p_i – and, thus, of the desired values ψ_i.

For simplicity, let us assume that for each new measuring instrument, we cover one more eigenvalue and double the sample size. This means that the n-th measuring instrument measures the first n values $\psi_1, \ldots, \psi_n$ with accuracy $\frac{1}{\sqrt{N}} = \frac{1}{\sqrt{2^n}} = 2^{-n/2}$. Thus, the two states $\psi = (\psi_1, \psi_2, \ldots)$ and $\psi' = (\psi'_1, \psi'_2, \ldots)$ are equivalent relative to the first n measurements – i.e., in our notations, $\psi \sim_n \psi'$ – if and only if $\max\limits_{1 \le i \le n} |\psi_i - \psi'_i| \le 2^{-n/2}$.

Then, we define a topology by neighborhoods $\{\psi' : \psi' \sim_n \psi\}$. With respect to this topology, the set S of all physical states is *compact*: indeed, for every $\varepsilon = 2^{-n/2}$, a finite ε-net in the n-dimensional sphere

$$\{(\psi_1, \ldots, \psi_n) : \sum_{i=1}^{n} \psi_i^2 = 1\}$$

is also an ε-net for the set S. So, in this topology, uniqueness indeed implies algorithmic computability.

Mathematical comment The new topology can be described by a metric. Indeed, the above similarity relation $\sim_n$ is "transitive" in the following sense: since $2^{-n'/2} + 2^{-n''/2} \le 2 \cdot 2^{-\min(n', n'')/2} = 2^{-n/2}$, where $n \stackrel{\text{def}}{=} \min(n', n'') - 2$, we can conclude that if $\psi \sim_{n'} \psi'$ and $\psi'_n \sim_{n''} \psi''$, then $\psi \sim_n \psi''$ for $n = \min(n', n'') - 2$.

If $\psi \sim_n \psi'$ for all n, this means that for every i, the values ψ_i and ψ'_i coincide and thus, that the states ψ and ψ' are equal. Hence, if the states ψ and ψ' differ, then eventually, from the results of these measurements, we will detect these differences. The smaller the difference between the states, the more accurate measuring instrument is needed to distinguish between them. We can therefore gauge the difference between the two states ψ and ψ' by the value $n(\psi, \psi')$ – the smallest integer n for which the n-th measuring instrument can distinguish between ψ and ψ', i.e., for which $\max\limits_{1 \le i \le n} |\psi_i - \psi'_i| > 2^{-n/2}$.

One can check that the expression $\rho(\psi, \psi') \stackrel{\text{def}}{=} 2^{-n(\psi, \psi')/2}$ is a metric corresponding to the above topology.

Physical comment Possible physical consequences of this idea are discussed, e.g., in [30].

9 Physically Important Case when Indistinguishability is Transitive

From the physical viewpoint, it is desirable to require that the "indistinguishability" relation $\sim$ on the set X of possible values of a physical quantity is an equivalence relation, i.e., that it is reflexive ($a \sim a$), symmetric, and transitive.

The above defined relation is always reflexive and symmetric, but it is not always transitive: e.g., for the "yes"-"no" measurements, $0 \sim U$, $U \sim 1$, but $0 \not\sim 1$. How can we describe sets with a transitive relation?

Proposition 2. *For each quantity X, the following two conditions are equivalent to each other:*

- *the indistinguishability relation $\sim$ is transitive, i.e., for all $a, b, c \in X$, if $a \sim b$ and $b \sim c$, then $a \sim c$;*
- *for every i, there exists a j for which, if $a_j \sim_j b_j$ and $b_j \sim_j c_j$, then $a_i \sim_i c_i$.*

Proof It is easy to prove that the second condition implies the first one.

Indeed, let $a \sim b$ and $b \sim c$. We want to prove that $a \sim c$. By definition, $a \sim c$ means that $a_i \sim_i c_i$ for all i. Let us pick any i and prove that $a_i \sim_i c_i$. Indeed, due to the second condition, there exists a j for which $a_j \sim_j b_j$ and $b_j \sim_j c_j$ imply that $a_i \sim_i c_i$. By definition of $\sim$, the condition $a \sim b$ means that $a_k \sim_k b_k$ for all k, in particular, for $k = j$. Thus, we have $a_j \sim_j b_j$, and similarly, we have $b_j \sim_j c_j$, hence we can indeed conclude that $a_i \sim_i b_i$.

Let us now prove that the first condition implies the second one.

We will prove it by contradiction. Indeed, if no such j exists, this means that for every j, there exist values $a^{(j)}$, $b^{(j)}$, and $c^{(j)}$ for which $a_j^{(j)} \sim_j b_j^{(j)}$, $b_j^{(j)} \sim_j c_j^{(j)}$, but $a_i^{(j)} \not\sim_i c_i^{(j)}$. By taking a convergent subsequence, in the limit, we get sequences a, b, and c, for which $a \sim b$, $b \sim c$, but $a \not\sim c$ – specifically, $a_i \not\sim_i c_i$.

Let us describe this construction in detail. The construction consists of Steps $1, 2, \ldots$ On each Step k, we will have a set of triples $(a^{(j)}, b^{(j)}, c^{(j)})$ corresponding to all the values j from some infinite set of indices $J_k \subseteq N$. We start with the class of all such triple $J_0 = N$. At Step k, we define the values a_k, b_k, and c_k for this k, and further limit the class of the triples to the class J_k of all the triples from J_{k-1} that are consistent with the selected values a_k, b_k, and c_k in the sense that $a_k^{(j)} = a_k$, $b_k^{(j)} = b_k$, and $c_k^{(j)} = c_k$.

Let us now describe how exactly we can perform Step k. For each of the infinitely many triples $(a^{(j)}, b^{(j)}, c^{(j)})$, $j \in J_{k-1}$, the k-th projection $(a_k^{(j)}, b_k^{(j)}, c_k^{(j)})$

is an element of the finite set $X_k \times X_k \times X_k = X_k^3$ of all possible triples of outcomes of the device I_k. Thus, at least one of these triples is repeated infinitely many times in the sequence $(a^{(j)}, b^{(j)}, c^{(j)})$. We will pick one of such infinitely repeated triples as the triple (a_k, b_k, c_k). This way, we guarantee that the selected set J_k is still infinite.

Let us show that the resulting mappings a, b, and c indeed belong to the set X, i.e., indeed satisfy the projection property. It is sufficient to prove it for a. Indeed, on each Step k, we restrict ourselves to indices $j \in J_k$ for which $a_k^{(j)} = a_k$. Thus, on the next Step $k+1$, we only consider such values $a^{(j)}$. The value a_{k+1} selected on the next step is equal to $a_{k+1}^{(j)}$ for one of the indices $j \in J_k$. By definition of X, we have $\pi_{k+1,k}(a_{k+1}^{(j)}) = a_k^{(j)}$, hence $\pi_{k+1,k}(a_{k+1}) = a_k$.

Let us show that for the corresponding elements, $a \sim b$ and $b \sim c$. It is sufficient to prove it for a and b. We need to prove that for every k, we have $a_k \sim_k b_k$. Indeed, for every k, we selected the triple (a_k, b_k, c_k) as a triple that appear in a sequence $(a_k^{(j)}, b_k^{(j)}, c_k^{(j)})$ for infinitely many different indices $j \in J_{k-1}$. Since there are infinitely many such indices, at least one of these indices j must be $\geq k$. For this j, we have $a_k = a_k^{(j)}$ and $b_k = b_k^{(j)}$. By definition of a triple $(a^{(j)}, b^{(j)}, c^{(j)})$, we have $a_j^{(j)} \sim_j b_j^{(j)}$. Since $k \leq j$, we can thus conclude that $\pi_{j,k}(a_j^{(j)}) \sim_k \pi_{j,k}(b_j^{(j)})$.

Since $a^{(j)}$ and $b^{(j)}$ are elements of the projective limit, the value $\pi_{j,k}(a_j^{(j)})$ is equal to $a_k^{(j)}$, i.e., to a_k. Similarly, $\pi_{j,k}(a_j^{(j)}) = b_k$, so we indeed conclude that $a_k \sim_k b_k$ for all k.

To complete the proof, we must now prove that $a_i \not\sim_i c_i$. Indeed, according to our construction, there exists a j for which $a_i = a_i^{(j)}$, $b_i = b_i^{(j)}$, and $c_i = c_i^{(j)}$. By definition of the triples $(a^{(j)}, b^{(j)}, c^{(j)})$, for each of these triples, we have $a_i^{(j)} \not\sim_i b_i^{(j)}$, hence $a_i \not\sim_i c_i$.

The proposition is proven.

If indistinguishability is transitive, then from each i, we can algorithmically find j for which $a_j \sim_j b_j$ and $b_j \sim_j c_j$ imply $a_i \sim_i c_i$ The above proof is not constructive: it proves the existence of the desired j not by explicitly constructing such a j but by reduction to a contradiction (i.e., by showing that if such a j does not exist, then we get a contradiction).

We can show, however, that for computable quantities, there is an *algorithmic* way of finding such a j. First, it is easy to show that if this implication is true for $j < i$, it is thus true for all larger j as well. Thus, it is sufficient to consider the case when $j \geq i$.

In this case, the condition $a_i \sim_i c_i$ can be described in terms of a_j and c_j, as $\pi_{j,i}(a_j) \sim_i \pi_{j,i}(c_j)$.

For each j, there are only finitely many possible pairs of triples $(a_j, b_j, c_j) \in X_j^3$, and the relations $\sim_j$ and $\sim_i$ are algorithmically decidable. Therefore, by testing all possible triples, we can algorithmically check whether for this j, $a_j \sim_j b_j$ and $b_j \sim_j c_j$ indeed imply $a_i \sim_i c_i$.

If the implication is valid, we got our j. If not, we increase j by 1 and test again, etc., until we find the desired value j.

10 Pre-Functions (Relations) and Functions

Once we have a description of the sets A and B as quantities (i.e., as projective limits), how can we describe, in these terms, the set of all functions from A to B? For example, if we know how to describe time t and how to describe spatial coordinate x, how can we then describe a trajectory $x(t)$, i.e., a function that maps t into x?

In physical terms, a function $f : A \to B$ means that, once we know an approximation a_n to a, we can find some approximation $b_m = F(a_m)$ to b. There are two natural requirements:

- that if the approximations a and a' are compatible, i.e., if there is an object which can lead to both measurement results a and a', then the corresponding values $F(a)$ and $F(a')$ should also be compatible – since they can both come from this same object, and
- that if we add more information to a, we should get more information about $f(a)$.

Thus, we arrive at the following definitions:

Definition 12. *We say that an approximation a_k is* more informative *than an approximation a_l (and denote it $a_k \subseteq a_l$) if $a_l = \pi_{l,k}(a_k)$.*

Definition 13. *Let A and B be two quantities (projective limits). By a* pre-function

$$F : A \to B,$$

we mean a mapping from $\cup A_n$ to $\cup B_n$ such that:

- *$a \sim a'$ implies $F(a) \sim F(a')$;*
- *if $a \subseteq a'$, then $F(a) \subseteq F(a')$.*

A pre-function $F : A \to B$ between computable quantities A and B is called computable *if the mapping F is computable.*

We call this mapping a *pre-function*, not a function, because the results may not converge: It is quite possible that for some element $a = (a_1, a_2, \ldots, a_n, \ldots)$, the approximations a_n provide better and better approximations to a – until we are left with a single element a – but we may still have $F(a_1) = F(a_2) = \ldots = F(a_n) = \ldots$ with several elements inside this common approximation $F(a_i)$. In mathematical terms, such pre-functions are *relations*, not functions, since to the same element a, we can put into correspondence several different values $F(a)$.

For a pre-function to be a function, we need to make sure that for every element $a \in A$, there is only one element from B which can be viewed as the image of F.

Definition 14. *We say that a pre-function $F : A \to B$ is* defined *at an element $a = (a_1, a_2, \ldots, a_n, \ldots)$ if from a sequence $(F(a_1), F(a_2), \ldots, F(a_n), \ldots)$, we can extract a subsequence which is an element of B, i.e., if for every m, there exists an n such that $b_m \stackrel{\text{def}}{=} F(a_n) \in B_m$. The resulting element $b = (b_1, b_2, \ldots, b_m, \ldots)$ is called the* value *of the pre-function F on an element a and denoted by $f(a)$.*

Definition 15. *We say that a pre-function $F : A \to B$ is a* function *if it is defined for all elements of A.*

How can we describe pre-functions which are functions?

Proposition 3. *A pre-function $F : A \to B$ is a function if and only if for every m, there exists an n such that for every $a_n \in A_n$, we have $F(a_n) \in B_m \cup B_{m+1} \cup \ldots$*

Proof This statement can be proven by using compactness and reduction to a contradiction.

In other words, a pre-function is a function if for every desired accuracy level m, we can determine beforehand with what accuracy we must compute a to get $F(a)$ with the desired accuracy. This is what is usually called continuity, so in this description, every function is continuous in this sense.

Comment Similar to transitivity, if a computable pre-function is a function, then based on m, we can algorithmically find n for which $a_n \in A_n$ implies that $F(a_n)$ gives the result with accuracy level m or higher.

Similarly to the computable functions on computable compacts, we can algorithmically compute the maximum of each function $f : X \to R$:

Proposition 4. *There exists an algorithm that, given a computable quantity (projective limit) X and a computable function $F : X \to R$ from X to real numbers R, returns the maximum value $M = \max f(x)$ of this function.*

Proof First, we can show that the set of values $f(x)$ of the function f is bounded. Indeed, if this set was, e.g., unbounded from above, then for every n, we would have values $x^{(n)}$ for which $f(x^{(n)}) \geq n$. From the sequence $x^{(n)}$, we can extract a convergent subsequence $x^{(n_k)} \to x$. For its limit x, we get $f(x) = \lim f(x^{(n_k)}) = +\infty$, which contradicts our assumption that $f(x) \in R$ for all x.

Since the range of the function $f(x)$ is a bounded set, we can view the corresponding values $F(x_k)$ as simply rational numbers, and the relation $a \sim_n b$ between these rational numbers as simply $|a - b| \leq 2^{-n}$. Since the pre-function F is a function, for every given m, there exists n for which, for every $x_n \in X_n$, the value $F(x_n)$ is at a level m or higher. Thus, for every $x \in X$, we have $|f(x) - F(x_n)| \leq 2^{-m}$. Hence, the maximum of the corresponding finitely many

values $f(x_n)$, $x_n \in X_n$, is 2^{-m}-close to the maximum $M = \max f(x)$ over all possible $x \in X$.

So, to compute M with a given accuracy 2^{-m}, we first algorithmically find n for which $F(x_n)$ is always of level m ro higher, and then take, as the desired approximation $\widetilde{M}$ to M, the largest of the finitely many values $F(x_n)$ for all $x_n \in X_n$. The proposition is proven.

Comment For functions with possibly infinite values, we can also algorithmically get a maximum – as a computable element of the corresponding projective limit.

11 Main Result: Uniqueness Implies Computability

Let us now prove that for newly defined sets of values of physical quantities, uniqueness implies algorithmic computability.

Definition 16. *Let $F : X \to R$ be a function from a quantity (projective limit) X. We say that it* attains a zero *at a unique point if $f(x) = f(x') = 0$ (classically) implies $x \sim x'$.*

Theorem 3. *There exists an algorithm A which is applicable to an arbitrary computable function $F : X \to R$ from a computable transitive quantity X to real numbers which attains a 0 at a unique point x, and which returns the "zero" point $z = A(F) \in X$ for which $f(z) = 0$.*

Definition 17. *Let $F : X \to R$ be a function from a quantity X. We say that it* attains a maximum *at a unique point if $f(x) = f(x') = \max f(y)$ (classically) implies $x \sim x'$.*

Theorem 4. *There exists an algorithm A which is applicable to an arbitrary computable function $F : X \to R$ from a computable transitive quantity X to real numbers which attains a maximum at a unique point x, and which returns the point $x_M = A(F) \in X$ for which $f(x_M) = \max f(x)$.*

Proof A function $f(x)$ attains its maximum if and only if the auxiliary function $f'(x) \stackrel{\text{def}}{=} f(x) - \max f(y)$ attains a 0. Since the maximum $\max f(y)$ is computable, we can thus reduce the second problem to the first one. So, it is sufficient to consider the first problem – of finding a value z for which $f(z) = 0$.

Let us show how for a given k, we can find z with accuracy of level k, i.e., how we can find an approximation a for which $a \sim_k z$. Since X is transitive, there exists an n such that $a \sim_n b$ and $b \sim_n c$ imply that $a \sim_k c$.

The main idea of this proof is the same as for computable functions on a computable compact set: that eventually, all the points with large enough values will be concentrated in a small vicinity of the desired zero point. Let us describe this idea in detail.

Uniqueness means that $f(x) = 0$ implies $x \sim z$. In particular, $f(x) = 0$ implies that $x \sim_n z$. Let us prove that there exists an integer m such that $f(x) \sim_m 0$ – i.e., $|f(x)| \le 2^{-m}$ – implies $x \sim_n z$. Indeed, if no such m existed, then we would have a sequence of elements $x^{(m)}$ for which $f(x^{(m)}) \sim_m 0$ but $x^{(m)} \not\sim_n z$. From the sequence $x^{(m)}$, we can extract a convergent subsequence. For the limit x of this subsequence, we would get $f(x) = 0$ but $x \not\sim_n z$ – a contradiction.

Let m be the integer whose existence we have just proven. Since f is a function, there exists an ℓ such that for every x, we have $|f(x) - F(x_\ell)| \le 2^{-(m+1)}$. In particular, for $x = z$, we have $f(z) = 0$ hence $|F(z_\ell)| \le 2^{-(m+1)}$. Vice versa, if $|F(x_\ell)| \le 2^{-(m+1)}$ for some $x_\ell \in X_\ell$, this means that

$$|f(x)| \le |f(x) - F(x_\ell)| + |F(x_\ell)| \le 2^{-m}$$

and hence, $x \sim_n z$, i.e., $x_\ell \sim_n z_\ell$.

So, there exists an ℓ for which there is a $y_\ell \in X_\ell$ that satisfies the following two properties:

- $|F(y_\ell)| \le 2^{-(m+1)}$,
- y_ℓ is n-close to every other $x_\ell \in X_\ell$ for which $|F(x_\ell)| \le 2^{-(m+1)}$ (i.e., $x_\ell \sim_n y_\ell$).

We can find this ℓ by trying all possible natural numbers $\ell = 1, 2, \ldots$

For the found ℓ, the above two properties hold for $y_\ell = z_\ell$, but they may also hold for some other value $y_\ell \ne z_\ell$. In this case, we have $y_\ell \sim_n z_\ell$ and $z_\ell \sim_n z$, hence, due to transitivity, $y_\ell \sim_k z$. Thus, this y_ℓ is indeed the desired approximation to the root z. The theorem is proven.

Comment In our formulation of the theorems, we assume that the uniqueness is classically true, i.e., that $f(x) = f(x') = 0$ implies $x \sim x'$ in the classical (non-constructive) sense. If instead, we assume that this implication is *constructively* true, then this assumption, in effect, provides us with a "modulus of uniqueness", i.e., with the value m for which $f(x) \sim_m 0 \to x \sim_n z$. By applying this idea systematically, we can avoid exhaustive search and get algorithms which are valid without Markov's principle – and which are, thus, hopefully, more feasible.

This idea can be also applied to other similar results: e.g., to our proof that if indistinguishability is transitive, then from each i, we can algorithmically find j for which $a_j \sim_j b_j$ and $b_j \sim_j c_j$ imply $a_i \sim_i c_i$. Specifically, if transitivity is understood constructively, then this transitivity condition, in effect, explicitly provides us with the desired integer j – without the need to use exhaustive search and Markov's principle.

Philosophical comment: optimism is justified Let us summarize what we have proven:

- the problem of finding a unique solution is algorithmically solvable, while

- if we have at least two solutions, no general algorithm is possible.

This result can be viewed as a foundation of *optimism*; see, e.g., [36, 41].

Let us illustrate this idea on the example of *historical processes*. It often happens in history that a country is in a bad situation; it may be that a tyrant is ruining it, it may be that the enemies are devastating it. As times goes by, the situation gets gloomier and gloomier; it seems that there are fewer and fewer chances of survival, and then, in the darkest hour, when there seems to be the only easy-to-miss chance of survival, a miracle happens and this only possible way out is indeed followed.

It often happens that the army wins only after it was almost defeated. It often happens that the personal enlightment comes only after a person has plunged into despair. It often happens that a proof of the theorem comes only after unsuccessful attempts have almost led to despair. The above results explain these "miracles": according to these results, in general, it is much easier to find a way out when there is the only way out left, than to find it when there were still several possible.

This result teaches us to be *optimistic*: if the situation gets gloomier and gloomier we should not despair but try harder, and then, hopefully, when there will be only one way out left, we will find it!

12 Conclusions and Future Work

Conclusions In many practical problems, we are interested in solving an optimization problem or to solve a solution to a system of equations. Often, the existence of a solution is proven indirectly, non-constructively, without an efficient method for constructing the corresponding object. In many cases, we can extract an algorithm from a classical proof: e.g., when an object is (non-constructively) proven to be unique in a locally compact space (or when there are two such objects with a known lower bound on the distance between them). In many other practical situations, a (seemingly) natural formalization of the corresponding practical problem leads to a non-compact set.

We have shown that if we consistently apply Heisenberg's operationalism idea and define objects in terms of directly measurable quantities, then we get a Girard-domain type representation in which a natural topology is, in effect, compact – and thus, uniqueness implies computability.

First direction of future work: from "from-scratch" definitions to a more systematic approach In the above text, we assumed a "from-scratch" approach to describing every set of possible values of a physical quantity. In this approach, we need to know how exactly to measure this quantity.

Fundamental physical quantities are indeed usually defined in this way, by describing the measurement procedures which enable us to estimate the values of these quantities. However, many other physical quantities are defined *indirectly*, in terms of already defined quantities. For example, a 3-D point can be described

as a triple of coordinates, an electric field can be described as a function that maps a 3-D (or 4-D) into a value of the strength and direction of the filed at this point, etc. For such indirectly defined quantities, it is desirable to avoid a lengthy "from-scratch" procedure and to be able to instead utilize the (Girard domain-related) descriptions of the related simpler quantities. The possibility of such a utilization comes from the fact that in Girard domains, there are natural descriptions of domain of pairs, domains of functions, etc.

We have already mentioned that in some cases, such a utilization is indeed possible: e.g., for the set of pairs (Cartesian product), which is naturally represented by a tensor product of the corresponding domains. It is desirable to generalize these examples into a systematic way of such utilization.

Second direction of future work: from general (possibly inefficient) algorithms to efficient (practical) algorithms In the above text, we were interested in checking whether there is an algorithm for solving a given practical problem. While it is known that in general, e.g., the optimization problem is not algorithmically solvable, an algorithmic solution exists when the solution is unique.

However, the existence of an algorithm does not mean that the problem is always practically solvable: an algorithm may require un-realistic exponential (or even hyper-exponential) computation time. It is therefore desirable not only to prove the existence of an algorithm, but to try to come up with *efficient* algorithms for solving the desired problems.

One can show that, e.g., the general problem of solving a system of polynomial equations is NP-hard; it is even NP-hard for quadratic equations (see, e.g., [41] and references therein). Similarly, we can prove that the problem of finding the *unique* solution to a system of equations (or the unique point where the maximum is attained) is as complicated as the problem of finding the *unique satisfying vector* for a given propositional formula. The latter problem (it is usually denoted by USAT, from *u*nique *sat*isafiability) is known to be "almost" NP-hard in the sense that every other problem from the class NP can be *probabilistically* reduced to USAT; so if we were able to solve all the instances of USAT in polynomial time, we would have a *probabilistic* polynomial-time algorithm that solves almost all instances of all problems from the class NP; see, e.g., [24, 58]. Note, however, that in [5], arguments are given that this problem may not be NP-hard.

Third direction of future work: from binary domains (graphs) to simplicial complexes To get a better description of a measuring instrument, it may be desirable to know not only which pairs are "compatible", but also which triples etc. can come from the same object. If $a \sim b$, $b \sim c$, and $a \sim c$, then for some measuring instruments, all three values are possible outcomes for some object, while for others, no single object can lead to these three outcomes.

Informally, let us say that a set $S \subseteq X$ is *compatible* if for some object, all values from S are possible. Then, a measuring instrument can be represented as

a pair $\langle X, \mathcal{S}\rangle$, where $X \subseteq \mathcal{S} \subseteq 2^X$ is the class of all compatible sets. Clearly, if a set S is compatible then each subset of S is compatible too. In mathematical terms, such a pair is called a *simplicial complex*: X is the set of vertices, and $\mathcal{S}$ is the set of faces. In precise terms, a *simplicial complex* is defined as a pair $\langle X, \mathcal{S}\rangle$, where $X \subseteq \mathcal{S} \subseteq 2^X$.

In a similar context, simplicial complexes have been introduced and analyzed by T. Ehrhard – as part of his study of hypercoherences [11, 12] – and by M. Smyth [55]. It is desirable to use the results from [11, 12, 55] to extend our approach and our results to simplicial complexes.

Acknowledgments

This work was supported in part by NSF grants HRD-0734825, EAR-0225670, and EIA-0080940, by Texas Department of Transportation grant No. 0-5453, by the Japan Advanced Institute of Science and Technology (JAIST) International Joint Research Grant 2006-08, and by the Max Planck Institut für Mathematik.

The authors are thankful to all the participants of the Conference on the Methods of Proof Theory in Mathematics (Bonn, Germany, June 3–10, 2007), especially to Ulrich Kohlenbach and Grigory Mints, for valuable discussions of the first draft of this paper.

References

[1] O. Aberth, *Precise Numerical Analysis Using C++*, Academic Press, New York, 1998.

[2] A. Aho, J. Hopcroft, and J. Ulman, *Data Structures and Algorithms*, Addison-Wesley, Boston, Massachusetts, 1983.

[3] J. Barwise (ed.), *Handbook of Mathematical Logic*, North-Holland, Amsterdam, 1977.

[4] M. J. Beeson, *Foundations of constructive mathematics*, Springer-Verlag, N.Y., 1985.

[5] R. Beigel, H. Buhrman, and L. Fortnow, "NP might not be as easy as detecting unique solutions", In: *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing (STOC)*, 1998, pp. 203–208.

[6] J. Berger, D. Bridges, and P. Schuster, "The fan theorem and unique existence of maxima", *J. Symbolic Logic*, 2006, Vol. 71, No. 2, pp. 713–720.

[7] E. Bishop, *Foundations of Constructive Analysis*, McGraw-Hill, 1967.

[8] E. Bishop and D. S. Bridges, *Constructive Analysis*, Springer, N.Y., 1985.

[9] D.S. Bridges, *Constructive Functional Analysis*, Pitman, London, 1979.

[10] D. S. Bridges and S. L. Vita, *Techniques of Constructive Analysis*, Springer-Verlag, New York, 2006.

[11] T. Ehrhard, "Hypercoherences: a strongly stable model of linear logic", *Mathematical Structures in Computer Science*, 1993, pp. 365–385.

[12] T. Ehrhard, "Parallel and serial hypercoherences", *Theoretical Computer Science*, 2000, Vol. 247, pp. 39–81.

[13] H. B. Enderton, *A Mathematical Introduction to Logic*, Academic Press, New York, 1972.

[14] R. P. Feynman, R. B. Leighton, and M. Sands, *Feynman Lectures on Physics*, Addison-Wesley, Boston, Massachusetts, 2005.

[15] P. C. Fishburn, *Interval Orders and Interval Graphs*, Wiley, New York, 1985.

[16] M. G. Gel'fond, "On constructive pseudofunctions", *Proceedings of the Leningrad Mathematical Institute of the Academy of Sciences*, 1969, Vol. 16, pp. 20–27 (in Russian), English translation: *Seminars in Mathematics*, Published by Consultants Bureau (New York-London), 1971, Vol. 16, pp. 7–10.

[17] M. G. Gel'fond, "Relationship between the classical and constructive developments of mathematical analysis", *Proceedings of the Leningrad Mathematical Institute of the Academy of Sciences*, 1972, Vol. 32, pp. 5–11 (in Russian), English translation: *Journal of Soviet Mathematics*, 1976, Vol. 6, No. 4, pp. 347–352.

[18] M. G. Gelfond, *Classes of formulas of classical analysis which are consistent with the constructive interpretation*, PhD Dissertation, Leningrad Mathematical Institute of the Academy of Sciences, 1975 (in Russian).

[19] J.-Y. Girard, "The system F of variable types, fifteen years later", *Theoretical Computer Science*, 1986, Vol. 45, No. 2, pp. 159–192.

[20] J.-Y. Girard, "Linear logic", *Theoretical Computer Science*, 1987, Vol. 50, No. 1, pp. 1–102.

[21] J.-Y. Girard, P. Taylor, and Y. Lafont, *Proofs and Types*, Cambridge University Press, Cambridge, 1989.

[22] M. C. Golumbic, *Algorithmic Graph Theory and Perfect Graphs*, Academic Press, New York, 1980.

[23] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control and Robotics*, Springer-Verlag, London, 2001.

[24] D. S. Johnson, "The NP-completeness column", *Journal of Algorithms*, 1985, Vol. 6, pp. 291–305.

[25] R. B. Kearfott, *Rigorous Global Search: Continuous Problems*, Kluwer, Dordrecht, 1996.

[26] J. L. Kelley, *General Topology*, Springer Verlag, Berlin-Heidelberg-New York, 1975.

[27] U. Kohlenbach. *Theorie der majorisierbaren und stetigen Funktionale und ihre Anwendung bei der Extraktion von Schranken aus inkonstruktiven Beweisen: Effektive Eindeutigkeitsmodule bei besten Approximationen aus ineffektiven Eindeutigkeitsbeweisen*, Ph.D. Dissertation, Frankfurt am Main, 1990.

[28] U. Kohlenbach, Effective moduli from ineffective uniqueness proofs. An unwinding of de La Vallée Poussin's proof for Chebycheff approximation, *Annals for Pure and Applied Logic*, 1993, Vol. 64, No. 1, pp. 27–94.

[29] U. Kohlenbach, *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*, Springer Verlag, Berlin-Heidelberg, 2008.

[30] O. M. Kosheleva and V. Kreinovich, *On the Algorithmic Problems of a Measurement Process,* Research Reports in Philosophy of Physics, University of Toronto, Ontario, Canada, Department of Philosophy, No. 5, 1978, 63 pp.

[31] V. Kreinovich, *Complexity measures: computability and applications*, Master Thesis, Leningrad University, Department of Mathematics, Division of Mathematical Logic and Constructive Mathematics, 1974 (in Russian).

[32] V. Kreinovich, Uniqueness implies algorithmic computability, *Proceedings of the 4th Student Mathematical Conference*, Leningrad University, Leningrad, 1975, pp. 19–21 (in Russian).

[33] V. Kreinovich, Reviewer's remarks in a review of D. S. Bridges, *Constrictive functional analysis*, Pitman, London, 1979; *Zentralblatt für Mathematik*, 1979, Vol. 401, pp. 22–24.

[34] V. Kreinovich, *Categories of space-time models*, Ph.D. dissertation, Novosibirsk, Soviet Academy of Sciences, Siberian Branch, Institute of Mathematics, 1979 (in Russian).

[35] V. Kreinovich, "Unsolvability of several algorithmically solvable analytical problems", *Abstracts Amer. Math. Soc.*, 1980, Vol. 1, No. 1, p. 174.

[36] V. Kreinovich, *Philosophy of Optimism: Notes on the possibility of using algorithm theory when describing historical processes*, Leningrad Center for New Information Technology "Informatika", Technical Report, Leningrad, 1989 (in Russian).

[37] V. Kreinovich, "Physics-motivated ideas for extracting efficient bounds (and algorithms) from classical proofs: beyond local compactness, beyond uniqueness", *Abstracts of the Conference on the Methods of Proof Theory in Mathematics*, Max-Planck Institut für Mathematik, Bonn, Germany, June 3–10, 2007, p. 8.

[38] V. Kreinovich, G. P. Dimuro, and A. C. da Rocha Costa, *From Intervals to? Towards a General Description of Validated Uncertainty*, Catholic University of Pelotas, Brazil, Technical Report, January 2004.

[39] V. Kreinovich and R. B. Kearfott, "Computational complexity of optimization and nonlinear equations with interval data", *Abstracts of the Sixteenth Symposium on Mathematical Programming with Data Perturbation*, The George Washington University, Washington, D.C., May 26–27, 1994.

[40] V. Kreinovich, O. Kosheleva, S. A. Starks, K. Tupelly, G. P. Dimuro, A. C. da Rocha Costa, and K. Villaverde, "From Intervals to Domains: Towards a General Description of Validated Uncertainty, with Potential Applications to Geospatial and Meteorological Data", *Journal of Computational and Applied Mathematics*, 2007, Vol. 199, No. 2, pp. 411–417.

[41] V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational complexity and feasibility of data processing and interval computations*, Kluwer, Dordrecht, 1998.

[42] B. A. Kushner, *Lectures on Constructive Mathematical Analysis*, Amer. Math. Soc., Providence, Rhode Island, 1984.

[43] D. Lacombe, "Les ensembles récursivement ouvert ou fermés, et leurs applications à l'analyse récurslve", *Compt Rend.*, 1957, Vol. 245, No. 13, pp. 1040–1043.

[44] L. R. Lewis and C. H. Papadimitriou, *Elements of the Theory of Computation*, Prentice-Hall, Englewood Cliffs, New Jersey, 1981.

[45] V. A. Lifschitz, "Investigation of constructive functions by the method of fillings", *J. Soviet Math.*, 1973, Vol. 1, pp. 41–47.

[46] L. Longpré, V. Kreinovich, W. Gasarch, and G. W. Walster, "m solutions good, $m-1$ solutions better", *Applied Math. Sciences*, 2008, Vol. 2, No. 5, pp. 223–239.

[47] T. A. McKee and F. R. McMorris, *Topics in Intersection Graph Theory*, SIAM, Philadelphia, 1999.

[48] H. Poincaré, *Le Valeur de la Science*, Flammarion, Paris, 1905; English translation in H. Poincaré, *The Foundations of Science: Science and Hypothesis, The Value of Science, Science and Method*, The Science Press, 1913; available at http://www.archive.org/details/foundationsscie01poingoog

[49] T. Poston, *Fuzzy Geometry*, Ph.D. Thesis, University of Warwick, UK, 1971.

[50] M. Pour-El and J. Richards, *Computability in Analysis and Physics*, Springer-Verlag, New York, 1989.

[51] P. Schuster, "Problems, solutions, and completions", *Journal of Logic and Algebraic Programming*, 2010, Vol. 79, No. 1, pp. 84–91.

[52] J. R. Shoenfield, *Mathematical Logic*, Addison-Wesley, Boston, Massachusetts, 1967.

[53] M. B. Smyth, "Semi-metrics, closure spaces and digital topology", *Theoretical Computer Science*, 1995, Vol. 151, pp. 257–276.

[54] M. B. Smyth, "Topology and tolerance", *Electronic Notes in Theoretical Computer Science*, 1997, Vol. 6, p. 326.

[55] M. B. Smyth and R. Tsaur, "AFPP vs FPP the link between almost fixed point properties of discrete structures and fixed point properties of spaces", *Applied Categorical Structures*, 2003, Vol. 11, pp. 95–116.

[56] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posed Problems*, W. H. Whinston & Sons, Washington, D.C., 1977.

[57] A. S. Troelstra, *Lectures on Linear Logic*, CSLI, Stanford, California, 1992.

[58] L. G. Valiant and V. V. Vazirani, "NP is as easy as detecting unique solutions", *Theoretical Computer Science*, 1986, Vol. 47, pp. 85–93.

[59] S. Willard, *General Topology*, Dover Publ., New York, 2004.

[60] E. C. Zeeman, "The topology of the brain and visual perception", in: K. Fort, ed., *Topology of 3-Manifolds and Related Topics*, Prentice-Hall, Englewood Cliffs, New Jersey, 1962, pp. 240–256.