

# Free Choice in Quantum Theory: A $p$ -adic View

Vladimir Anashin <sup>1,2</sup> 

<sup>1</sup> Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Leninskie Gory 1, 119991 Moscow, Russia; vladimir.anashin@cs.msu.ru or anashin@iisi.msu.ru

<sup>2</sup> Federal Research Center 'Information and Control' Russian Academy of Sciences, 119333 Moscow, Russia

**Abstract:** In this paper, it is rigorously proven that since observational data (i.e., numerical values of physical quantities) are rational numbers only due to inevitably nonzero measurements errors, the conclusion about whether Nature at the smallest scales is discrete or continuous, random and chaotic, or strictly deterministic, solely depends on experimentalist's free choice of the metrics (real or  $p$ -adic) he chooses to process the observational data. The main mathematical tools are  $p$ -adic 1-Lipschitz maps (which therefore are continuous with respect to the  $p$ -adic metric). The maps are exactly the ones defined by sequential Mealy machines (rather than by cellular automata) and therefore are causal functions over discrete time. A wide class of the maps can naturally be expanded to continuous real functions, so the maps may serve as mathematical models of open physical systems both over discrete and over continuous time. For these models, wave functions are constructed, entropic uncertainty relation is proven, and no hidden parameters are assumed. The paper is motivated by the ideas of I. Volovich on  $p$ -adic mathematical physics, by G. 't Hooft's cellular automaton interpretation of quantum mechanics, and to some extent, by recent papers on superdeterminism by J. Hance, S. Hossenfelder, and T. Palmer.

**Keywords:** automaton; letter-to-letter transducer; sequential Mealy machine;  $p$ -adic 1-Lipschitz map; interpretation of quantum mechanics; Planck's scale; experimenter's free choice

## 1. Introduction

The main goal of the current paper is to prove some of results which were announced without proofs in [1], namely, to prove rigorously mathematical statements which show that an experimentalist's conclusions about whether Nature on the smallest of scales is discrete or continuous [2], random and chaotic, or strictly deterministic [3] solely depends on the experimentalist's free choice of the metrics he chooses to process the measurement data which basically are rational numbers due to inevitably nonzero measurement errors. It should be stressed that the said statements are not types of free-will theorems in quantum mechanics since the statements are about how the data obtained during experiments are postprocessed rather than about how an experimentalist chooses the measurement setting during experiments. This is a crucial difference between results of the current paper and, for example, a Conway–Kochen strong free will theorem [4]. *In order to distinguish between these two faces of experimentalist's freedom, in this paper, the two terms "free choice" and "free will" are used, and they are not interchangeable.*

There is some resemblance between the meanings of terms used in the invariant set theory [3,5] (within which a  $p$ -adic metric is briefly mentioned) and in the current paper; however, the current paper discusses a mathematical model for postprocessing of measurement data rather than broader physical theories.

The paper is inspired by the ideas of I. Volovich who, in collaboration with V. Vladimirov in the 1980s laid the cornerstone of contemporary  $p$ -adic mathematical physics [6]. The paper is motivated also by the ideas of G. 't Hooft who initiated the development of the *cellular automaton interpretation of quantum mechanics* [7] which is based on a suggestion that on some basic level there is no intrinsic randomness in nature.



**Citation:** Anashin, V. Free Choice in Quantum Theory: A  $p$ -adic View. *Entropy* **2023**, *25*, 830. <https://doi.org/10.3390/e25050830>

Academic Editor: Wilson A. Zuniga-Galindo

Received: 9 May 2023  
Accepted: 17 May 2023  
Published: 22 May 2023



**Copyright:** © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

More formally, the paper introduces a wide class of functions, each of which can be regarded as a continuous (and sometimes as a chaotic, having positive entropy) real function over continuous real time with respect to real metric and which simultaneously is strictly deterministic (and a nonchaotic, having zero entropy) causal function over discrete time with respect to the  $p$ -adic metric for every  $p > 1$ . By the common definition, causal functions are the mappings which can be performed by automata but only those automata which are the so-called letter-to-letter transducers (or, sequential Mealy machines whose sets of states are not necessarily finite) over a  $p$ -letter alphabet rather than by cellular automata on which G. 't Hooft's interpretation is based. These classes of automata differs both from algorithmic and physical points of view. From the algorithmic point of view, letter-to-letter transducers can be judged as the least powerful computers compared to cellular automata which are the most powerful ones. Any algorithm (i.e., any general recursive function) can be implemented on a suitable cellular automaton since the class of all cellular automata is Turing-complete [8,9], whereas algorithms which can be implemented by the transducers are necessarily primitive recursive functions, and moreover, constitute a small class of primitive recursive functions; see the end of Section 3.3. From a physical point of view, the sequential machines are models of *open systems* whereas cellular automata are models of *isolated systems*. In contrast to a sequential machine, a cellular automaton updates its states according only to a fixed local rule which does not depend on input, whereas the next state of a sequential machine depends both on input information and on a current state; the sequential machine produces output information which also depends both on input information and on the current state. Throughout this paper, the term **automaton** refers to a sequential Mealy machine with a potentially infinite number of states; for a formal definition of the latter machine see Definition 2. In what follows, types of automata different from the said Mealy machines are mentioned with respective adjectives, e.g., "cellular automaton" and "push-down automaton".

The paper is organised as follows:

- In Section 2, we recall a formal definition of causal function over discrete time (cf., Definition 1). The very term "causality" is based on the notion of time; this is why in the paper, "time" as a measurable physical entity is a central theme: time may be either discrete (e.g., Planck time) or continuous (e.g., real time) at respective "ends of scale". In this paper, we generally advocate that these cases are indistinguishable by measurements and actually are subject to an experimentalist's free choice of metric with respect to which he processes the numerical values of the experimental physical data. After the formal definition of causality over discrete time, we introduce as postulates statements of I. Volovich on indistinguishability using measurements of physical quantities between rational and irrational values and of G. 't Hooft on the nonexistence of randomness in Nature; then, we formalise the notion a "physical law" as a function which is consistent with these postulates, cf., Conditions 1.
- In Section 3, we review some notions and facts from  $p$ -adic analysis and from automata theory which will be needed further in the paper.
- In Section 4, we introduce one of the main notions of the paper; that is, the real causal functions which are the functions that are continuous both with respect to a real metric and to the  $p$ -adic metric; i.e., causal functions which reside simultaneously in two worlds, Archimedean and non-Archimedean. The main results described in these sections are as follows:
  - Theorem 5 completely describes the class of functions that satisfy Conditions 1; i.e., those which are completely consistent both with Volovich postulates and with 't Hooft causality postulate. We interpret this theorem to be a manifestation of the observer's freedom to conclude whether Nature on the smallest of scales is discrete or continuous since the conclusion depends solely on the observer's free choice of metric with respect to which the observer processes the measured numerical data.
  - In Section 4.4, we argue that the observer's conclusion as whether Nature is basically random and chaotic or totally predictable and deterministic also depends

solely on the observer's free choice of metric with respect to which the observer processes the measured numerical data; namely, we show that maps which are chaotic with respect to the real metric are strictly deterministic and predictable with respect to  $p$ -adic metric, irrespective to which common definition of chaos is used.

- In Section 4.5, we argue that Conditions 1 may be too restrictive from the physical point of view and relax the conditions, letting them hold only for some prime  $p$  rather than for all primes. This way we introduce a notion of a  $p$ -consistent function, show that the class of  $p$ -consistent functions is much wider than the class of completely consistent ones (Theorem 7), prove hologram-like property (Theorem 8) which shows that global behaviour of  $p$ -consistent functions is completely defined by their local behaviour, and then prove that wide classes of physically important functions (such as continuous real functions, real functions that vanish at infinity,  $n$ -th power integrable functions, wave functions) can be uniformly approximated by infinitely differentiable  $p$ -consistent functions; see Theorem 9. This theorem is yet one more piece of evidence supporting the notion that observer's conclusion on discreteness, continuity, and reversibility of time solely depends on the observer's free choice of metric. Finally, in this subsection, we prove Theorem 10 which yields that smooth  $p$ -consistent functions related to systems having a *finite number of states* are necessarily affine; this theorem may demonstrate where the *linearity* of operators used in quantum theory is rooted.
- In Section 5, we argue that “continuous” and “discrete” models of physical world “meet each other in the middle of the scales”, and the wave function is the “meeting point”. The specifics of this section are as follows:
  - In Section 5.1, we formalise what is meant by “measurements at each end of the scale” by introducing two observers, *Big-endian* and *Little-endian*, that perform measurements at respective ends, macro and micro.
  - In Section 5.2, we introduce a  $p$ -adic model of the instrument which measure and indicates time, a  *$p$ -adic clock*, and a respective notion of  $p$ -adic time, which is time by the Little-endian's clock. Then, we outline (Theorem 11), which proves that there exists a unique clock which is the same for Little-endian and for Big-endian, the *universal clock*. We argue that the known effect in quantum theory of indistinguishability of which of two event happens earlier than does another one may be rooted in the fact that  $p$ -adic time cannot be ordered, i.e., that in contrast to the ring of integers  $\mathbb{Z}$ , the ring  $\mathbb{Z}_p$  of  $p$ -adic integers cannot be ordered. Therefore, the existence or nonexistence of the “time arrow” is again subject to free choice of the metric by the experimentalist.
  - Section 5.3 describes the base on which the construction of wave function is founded. The section describes, in formal terms, the process of finding cluster points for experimental points in Euclidean space and constructing a smooth line (or surface) on which these cluster points fall. In the subsection, we mostly refer to results which were published earlier in [10,11] and interpret these as the models of physical systems having either discrete or continuous spectra. Based on these results, we argue that chaos is either immanent to continuous time models or emerges as a result of sufficiently long evolution of a physical system in discrete time models.
  - In Section 5.4, we construct two types of wave functions, the sharp one for Little-endian, with respect to discrete time, and the fuzzy one for Big-endian, with respect to continuous time. The fuzzy wave function can be approximated by sharp wave functions with any desirable accuracy, so this is again a subject to the free choice of the experimentalist regarding the type of wave function which depends on the experimentalist's free choice of metric. We show then in (Theorem 18), that under a reasonable finiteness assumption, the fuzzy 1-dimensional wave function is actually a sharp  $N$ -dimensional wave function over discrete 2-adic

- time. Here, as an extra mathematical tool, we use  $\beta$ -expansions of numbers; the  $\beta$ -expansions were originally introduced in [12,13].
- In Section 5.5, we formally derive a time-energy uncertainty relation in *entropic* form. Here, we use yet one more extra mathematical tool, the theory of prefix codes. All necessary notions, results, and proper references of this theory are given in the subsection. We stress that *no hidden variables are assumed*, and the uncertainty relation holds both for the Big-endian and for the Little-endian.
  - We conclude in Section 6. Here, we state that basically the results of the paper may be treated as information–theoretic and remark that the paper highlights J. Wheeler’s “it from bit” doctrine [14] since the final results on wave functions, especially Theorem 18, show that “it” is “from bit” indeed: both sharp and fuzzy wave functions actually turn out to be 2-adic 1-Lipschitz functions, i.e., automata functions over the alphabet  $\{0, 1\}$ .

## 2. Formalisation

I. V. Volovich, in his numerous papers, books, talks, etc., has stated, many times, the following postulates (further referred to as *Volovich postulates*) on which  $p$ -adic mathematical physics is founded:

- (i) Only rational numbers can be observed; irrational numbers cannot.
- (ii) Distances smaller than Planck length cannot be measured.
- (iii) Fundamental physical laws should be invariant with respect to a change of number field.

According to Ostrowski’s theorem, every nontrivial absolute value on the rational numbers  $\mathbb{Q}$  is equivalent to either the usual real absolute value or a  $p$ -adic absolute value (c.f., e.g., [15] [Theorem 10.1]). Then, to ensure the limits of convergent sequences over a field belong to the field; *the mentioned number fields must be the fields  $\mathbb{Q}_p$  of  $p$ -adic numbers or the field  $\mathbb{R}$  of real numbers* since these fields are the only completions of the field  $\mathbb{Q}$  with respect to absolute values on  $\mathbb{Q}$ . Of course, the fields can be complete extensions of the fields  $\mathbb{Q}_p$  and  $\mathbb{R}$  like, e.g., the fields of complex  $p$ -adic numbers  $\mathbb{C}_p$  or a field of “ordinary” complex numbers  $\mathbb{C}$ , but  $\mathbb{Q}_p$  and  $\mathbb{R}$  are the only “smallest” fields which satisfy the third Volovich postulate.

G. ‘t Hooft in his book *The Cellular Automaton Interpretation of Quantum Mechanics* [7] makes the following claim (further referred to as the *‘t Hooft causality postulate*) which is fundamental for the cellular automaton interpretation of quantum mechanics:

It may well be that, at its most basic level, there is no randomness in Nature, no fundamentally statistical aspect to the laws of evolution. Everything, up to the most minute detail, is controlled by invariable laws. Every significant event in our universe takes place for a reason, it was caused by the action of physical law, not just by chance. This is the general picture conveyed by this book.

To be consistent with this postulate, a physical system must be *causal*; that is, the “effect”, which is the reaction of the system to a “cause”, i.e., to an impact the system has been exposed, must be a function of the “cause” and of the “state” of the system. However, the very notion of causality is based on the notion of “time” which must be a totally ordered set since the “effect” cannot happen earlier than can the “cause” whose function the “effect” is. It is impossible to experimentally distinguish rational numbers from real numbers (cf. Volovich first postulate); therefore it is reasonable to assume that “time” is a totally ordered countable set. It is well known that any totally ordered countable set  $T$  is order-isomorphic to a subset of  $\mathbb{Q}$  (c.f., e.g., [16]) with respect to the natural order  $\leq$  on  $\mathbb{Q}$ . Time  $T$  is called *continuous* if the ordering of elements in  $T$  is dense; i.e., given  $t_1, t_2 \in T$  there exists  $t_3 \in T$  such that  $t_1 < t_3 < t_2$ . Time  $T$  is called *discrete*, and if given any  $t_1, t_2 \in T$ , there is not more than a finite  $t_3 \in T$  such that  $t_1 < t_3 < t_2$ .

“Continuous” physical models are based on the assumption that any temporal/spatial interval can be divided into smaller intervals ad infinitum. The “discrete” models assume that spacetime should somehow be “quantized” at the smallest of scales; i.e., there exist

the smallest spatial/temporal intervals which can not be divided into smaller ones, [2]. In the latter case, it would be reasonable to try to construct a mathematical theory assuming that total amount of these “indivisible” values can be increased ad infinitum. In the both cases, as well as in respective physical theories, the “infinity” simply stands for a value which is extremely small (or extremely large) compared to a given value so that calculations involving the notion of infinity result in values which agree with respective measured values up to a small real number, the error. Therefore, if theories of either type adequately describe physical reality at respective “ends of scale”, the theories must “meet one another somewhere in the middle of the scale”.

The discreteness implies that the indivisible intervals are respective units, i.e., take values of 1; moreover, both “cause” and “effect” are sequences of “elementary causes” and “elementary effects” which happen at discrete time instants  $0, 1, 2, \dots$

Actually, the “time unit” is the longest temporal interval within which it is impossible for an observer to determine whether any two events are simultaneous or not; i.e., which of the two events happens earlier/later than another one does. In other words, an “event” is like a film consisting of frames where each frame is a static picture, but the sequence of the pictures produces a movie on a screen which the audience of the cinema sees as dynamical process. Thus, the “elementary event” (“elementary cause”, “elementary effect”) is an event that lasts exactly one time unit similar to a momentary splash for which the moment when it begins is undistinguishable from the moment when it finishes.

We recall a notion of causal function over discrete time in terms of general system theory, c.f., e.g., [17,18].

**Definition 1** (Causality over discrete time). Causal functions over discrete time  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$  are exactly the functions  $f$  which satisfy the following conditions:

- (i) The domain (the “causes”) and range (the “effects”) of  $f$  are, accordingly, all sequences  $\mathbf{a} = (a_i)_{i \in \mathbb{N}_0}$  and  $\mathbf{b} = (b_i)_{i \in \mathbb{N}_0}$  over respective sets  $\mathcal{A}$ , the “elementary causes”, and  $\mathcal{B}$ , the “elementary effects”;
- (ii) If  $f(\mathbf{a}) = (b_i)_{i \in \mathbb{N}_0}$ , then  $b_i$  does not depend on  $a_{i+1}, a_{i+2}, \dots$ , for all  $i \in \mathbb{N}_0$ .

In other words, the function  $f$  is causal if and only if there exists a sequence  $(\varphi_i^f)_{i=0}^\infty$  of maps  $\varphi_i^f: \mathcal{A}^{i+1} \rightarrow \mathcal{B}$ , ( $i \in \mathbb{N}_0$ ), such that

$$f(\mathbf{a}) = (\varphi_i^f(a_0, \dots, a_i))_{i \in \mathbb{N}_0} \tag{1}$$

It is reasonable to assume that both sets  $\mathcal{A}$  of “elementary causes” and  $\mathcal{B}$  of “elementary effects” contain at least two elements and, moreover, that the sets are finite since no physical objects are known which have been proven to be infinite in some natural meaning: *Infinity is a mathematical rather than a physical notion which is used in mathematical calculations in order to find good estimates of physical values since the values can be measured with a nonzero error only.* From this **finiteness assumption**, it follows that the causal functions are exactly the mappings which are produced by a special class of automata, the letter-to-letter transducers (or, sequential machines) which transform input sequences  $\mathbf{a} = (a_i)_{i \in \mathbb{N}_0}$  of elementary causes into output sequences  $\mathbf{b} = (b_i)_{i \in \mathbb{N}_0}$  of elementary effects so that (ii) is satisfied (cf., e.g., [19], a classical monograph on automata theory). Note that condition (ii) is just a Lipschitz condition with a constant 1 with respect to the natural non-Archimedean metric  $d$  on sequences. The metric  $d$  can be defined as follows: given two sequences,  $\mathbf{c} = (c_i)_{i \in \mathbb{N}_0}$  and  $\mathbf{c}' = (c'_i)_{i \in \mathbb{N}_0}$ , over the same finite set,  $d(\mathbf{c}, \mathbf{c}') = p^{-n}$ , where  $n = \max\{i \in \mathbb{N}: c_i = c'_i\}$  if such  $n$  exists, and  $d(\mathbf{c}, \mathbf{c}') = 0$  if  $c_i = c'_i$  for all  $i \in \mathbb{N}_0$  (here  $p > 1$  is arbitrary real number). In this paper, we mostly consider the case when  $\mathcal{A}$  and  $\mathcal{B}$  are a finite  $p$ -element set  $\mathbb{F}_p$  where  $p$  is a prime number (the latter restriction is more a technical one imposed in order to not overload statements). This way, we may assume that  $\mathbb{F}_p$  is a finite  $p$ -element field and that the infinite sequences  $(c_i)_{i \in \mathbb{N}_0}$  (where  $c_i \in \mathbb{F}_p$ ) constitute the space  $\mathbb{Z}_p$  of  $p$ -adic integers under a natural one-to-one correspondence between the infinite sequences and canonical

representations of  $p$ -adic integers  $\sum_{i=0}^{\infty} c_i p^i$ . In the case when  $p > 1$  is not a prime number, the sequences also may be put in one-to-one correspondence with the space  $\mathbb{Z}_p$  of  $p$ -adic integers since the latter spaces are defined for all  $p = 2, 3, 4, \dots$ , and not necessarily only for prime  $p$ ; see, e.g., [20].

Physical models, loosely speaking, describe functions  $f$  which are “physical laws” that express dependencies of physical quantities on other physical quantities; therefore, if time is one of these quantities, it is reasonable to assume causality, i.e., the functions  $f$  are causal. Let us express more formally the conditions the functions  $f$  must meet in order to be consistent both with Volovich postulated and ‘t Hooft causality postulate.

In order to be consistent with Volovich postulates, the following conditions should be satisfied.

- As only rational numbers can be measured, the functions  $f$ , i.e., the closed forms of physical laws which can be experimentally verified, must be mappings of rational numbers to rational numbers; i.e., the functions  $f$  must take rational values when values of variables are rational.
- In order to study functions  $f$  when values of variables are “very large” or “very small” with respect to some reasonable metric, one has to expand the laws from the field of rational numbers  $\mathbb{Q}$  to a bigger field which is complete with respect to that metric; therefore, this bigger field can only be the field of real numbers  $\mathbb{R}$  and/or  $p$ -adic fields  $\mathbb{Q}_p$  for primes  $p = 2, 3, 5, 7, 11, \dots$ ; however, in order to be invariant with respect to the change of the number field, a restriction to  $\mathbb{Q}$  of any such expansion of  $f$  to a bigger field  $\mathbb{F} \supset \mathbb{Q}$  must be the same irrespective to which field  $\mathbb{F}$  was used for in the expansion, whether  $\mathbb{F} = \mathbb{R}$  or  $\mathbb{F} = \mathbb{Q}_p$ .

Further, to be consistent also with the ‘t Hooft causality postulate, the functions  $f$  should be causal; however, as it has been argued before, the “time” with respect to which the functions are causal must be order-isomorphic to a subset of  $\mathbb{Q}$ . However, since, according to the Volovich postulates, no temporal interval smaller than Planck’s time can be measured, the temporal intervals can only be multiples of Planck’s time; therefore, the “time” over which the functions  $f$  are causal must be order-isomorphic to a subset of  $\mathbb{Z}$ . Thus, up to order isomorphism, the time scale is either  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$  or  $\mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$  depending on whether the “beginning of time” exists or does not exist. According to the contemporary physical picture of the universe, it is reasonable to assume that the “beginning of time” exists; thus, the time scale must be  $\mathbb{N}_0$ , up to order isomorphism. However, causal functions over the discrete time  $\mathbb{N}_0$  can be treated as  $p$ -adic 1-Lipschitz functions whose domain and range are  $p$ -adic integers  $\mathbb{Z}_p$  rather than the whole field  $\mathbb{Q}_p$ , c.f., the reasoning which follows Definition 1; thus, as the “common part” of  $\mathbb{Z}_p$  and  $\mathbb{R}$  (which we further denote via  $\mathbb{Z}_p \cap \mathbb{Q}$ ) are rational  $p$ -adic integers, i.e., the irreducible fractions, whose denominators are coprime to  $p$ , to be consistent with Volovich postulates, the causal functions must take values from  $\mathbb{Z}_p \cap \mathbb{Q}$  rather than from the whole  $\mathbb{Q}$ ; moreover, the functions must be expandable to the whole field  $\mathbb{R}$  since  $\mathbb{Z}_p \cap \mathbb{Q}$  is a dense subset of  $\mathbb{R}$ . Finally, we can specify the formal properties the functions  $f$  must share in order to be consistent both with Volovich postulates and with the ‘t Hooft causality postulate, as follows:

**Condition 1** (Complete consistency). A (univariate) continuous real function  $f: \mathbb{R} \rightarrow \mathbb{R}$  which is consistent with both Volovich postulates and the ‘t Hooft causality postulate must share the properties listed below.

- (i) For every prime  $p$ , the restriction  $f|_{\mathbb{N}_0}$  must be a causal function over discrete time  $\mathbb{N}_0$ ; i.e., the restriction  $f|_{\mathbb{N}_0}$  must satisfy a  $p$ -adic Lipschitz condition with a constant 1. That is, for all  $m, n \in \mathbb{N}_0$ , there must hold the inequality

$$d_p(f|_{\mathbb{N}_0}(m), f|_{\mathbb{N}_0}(n)) \leq d_p(m, n),$$

where  $d_p$  is the  $p$ -adic metric.

- (ii) Since  $\mathbb{N}_0$  is a dense subset in  $\mathbb{Z}_p$ , by (i), for every prime  $p$ , there exists a unique extension of  $f|_{\mathbb{N}_0}$  to the function  $f_p: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  which satisfies a Lipschitz condition with a constant of 1 with respect to the  $p$ -adic metric  $d_p$ . Therefore, to be invariant with respect to the change of the field, the function  $f: \mathbb{R} \rightarrow \mathbb{R}$  must act on the set  $\mathbb{Z}_p \cap \mathbb{Q}$  of all  $p$ -adic rational integers exactly as the function  $f_p$  does; that is, for every prime  $p$ , the restriction  $f|_{\mathbb{Z}_p \cap \mathbb{Q}}$  on rational  $p$ -adic integers  $\mathbb{Z}_p \cap \mathbb{Q}$  must coincide with the restriction  $f_p|_{\mathbb{Z}_p \cap \mathbb{Q}}$  on  $\mathbb{Z}_p \cap \mathbb{Q}$ :

$$f|_{\mathbb{Z}_p \cap \mathbb{Q}}(r) = f_p|_{\mathbb{Z}_p \cap \mathbb{Q}}(r), \text{ for all } r \in \mathbb{Z}_p \cap \mathbb{Q}.$$

**Note 1.** Condition 1 (ii) immediately implies that  $f(\mathbb{Z}_p \cap \mathbb{Q}) \subset \mathbb{Z}_p \cap \mathbb{Q}$  for all prime  $p$ ; thus, necessarily

$$f(\mathbb{Z}) \subset \bigcap_{p \text{ prime}} (\mathbb{Z}_p \cap \mathbb{Q}) = \mathbb{Z}.$$

The questions which immediately arise are whether there exist functions  $f$  which satisfy the conditions, and if such functions do exist, what are these functions. In Section 4, we show that functions which meet Conditions 1 do exist and constitute a class of all polynomials over  $\mathbb{Q}$  of a special type (the class contains, e.g., all polynomials over  $\mathbb{Z}$ ); see Theorem 5. Moreover, the functions turn out to be causal with respect to all finite alphabets and not necessarily with respect to  $p$ -symbol alphabets for prime  $p$ . This implies in particular that the answer to the commonly asked question about  $p$ -adic mathematical physics concerning what  $p$  should be chosen by an experimenter in order to make the theory consistent with the observations is as follows: the choice of  $p$  is absolutely free if causality, discreteness at Planck’s scale, and invariance with respect to the change of the number field are assumed.

We stress that the functions  $f$  which satisfy Conditions 1 are causal for all  $p$ -symbol alphabets and for all prime  $p$ , and hence, for all finite alphabets. In our view, the latter property appears to be too restrictive (and somewhat nonphysical, cf., the reasoning concerning the finiteness assumption above) since Planck’s scale includes a finite number of physical quantities (time, length, etc.) rather than an infinite number. Thus, it is reasonable to assume that Conditions 1 hold only for a finite set of primes; this implies that the functions  $f$  are causal with respect to finite alphabets, the prime power decompositions of the number of elements of which include only powers of primes from that set. The study of this class of functions can be reduced to cases containing only one prime,  $p$ . We show that if in the statement of Conditions 1, a prime  $p$  is fixed and that “for every prime  $p$ ” is replaced by “for the prime  $p$ ” then there exist functions  $f$  that satisfy the Conditions, which are continuous real functions on  $\mathbb{R}$  but which are not rational functions over  $\mathbb{Z}$ ; i.e., are not of the form  $u(x)/v(x)$  where  $u(x), v(x) \in \mathbb{Z}[x]$  are polynomials with integer coefficients; see Theorem 7. Note also that under such a restatement of the Conditions,  $f(\mathbb{Z})$  is not necessarily a subset of  $\mathbb{Z}$  but only a subset of  $\mathbb{Z}_p \cap \mathbb{Q}$ , cf., Note 1.

### 3. Preliminaries

We review some notions and facts from  $p$ -adic analysis and from automata theory which will be needed further in the paper.

#### 3.1. A Few Words about Words

An alphabet is just a finite nonempty set  $\mathcal{A}$ ; further in the paper, typically  $\mathcal{A} = \{0, 1, \dots, p - 1\}$ , where  $p > 1$  is an integer (mostly, but not always,  $p$  is a prime). Elements of  $\mathcal{A}$  are called symbols, or letters. By this definition, a word of length  $n$  over alphabet  $\mathcal{A}$  is a finite sequence (stretching from right to left)  $\alpha_{n-1} \cdots \alpha_1 \alpha_0$ , where  $\alpha_{n-1}, \dots, \alpha_1, \alpha_0 \in \mathcal{A}$ . The number  $n$  is called the length of the word  $w = \alpha_{n-1} \cdots \alpha_1 \alpha_0$  and is denoted via  $\Lambda(w)$ . The empty word  $\phi$  is a sequence of length 0; that is, the one that contains no symbols. Given a word  $w = \alpha_{n-1} \cdots \alpha_1 \alpha_0$ , any word  $v = \alpha_{k-1} \cdots \alpha_1 \alpha_0, k \leq n$ , is called a prefix of the word  $w$ , whereas any word  $u = \alpha_{n-1} \cdots \alpha_{i+1} \alpha_i, 0 \leq i \leq n - 1$  is called a suffix of the word  $w$ . Every word  $\alpha_j \cdots \alpha_{i+1} \alpha_i$  where  $n - 1 \geq j \geq i \geq 0$  is called a subword of the word

$w = \alpha_{n-1} \cdots \alpha_1 \alpha_0$ . Given words  $a = \alpha_{n-1} \cdots \alpha_1 \alpha_0$  and  $b = \beta_{k-1} \cdots \beta_1 \beta_0$ , the concatenation  $ab$  is the following word (of length  $n + k$ ):

$$ab = \alpha_{n-1} \cdots \alpha_1 \alpha_0 \beta_{k-1} \cdots \beta_1 \beta_0.$$

Given a word  $w$ , its  $k$ -times concatenation is denoted via  $(w)^k$

$$(w)^k = \underbrace{ww \dots w}_{k \text{ times}}.$$

We denote using  $\mathcal{W} = \mathcal{W}(\mathcal{A})$  the set of all nonempty words over  $\mathcal{A} = \{0, 1, \dots, p - 1\}$  and using  $\mathcal{W}_\phi$  the set of all words including the empty word  $\phi$ . In the sequel, the set of all  $n$ -letter words over the alphabet  $\mathcal{A}$ , we denote as  $\mathcal{W}_n$ ; thus,  $\mathcal{W} = \cup_{n=1}^\infty \mathcal{W}_n$ . To every word  $w = \alpha_{n-1} \cdots \alpha_1 \alpha_0$ , we put into the correspondence a non-negative integer  $\text{num}(w) = \alpha_0 + \alpha_1 \cdot p + \cdots + \alpha_{n-1} \cdot p^{n-1}$ . Thus,  $\text{num}$  maps the set  $\mathcal{W}$  of all the nonempty finite words over the alphabet  $\mathcal{A}$  onto the set  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$  of all non-negative integers. We will also consider a map  $\rho$  of the set  $\mathcal{W}$  into the real unit half-open interval  $[0, 1)$ ; the map  $\rho$  is defined as follows: given  $w = \beta_{r-1} \dots \beta_0 \in \mathcal{W}$ , put

$$\rho(w) = \text{num}(w) \cdot p^{-\Lambda(w)} = \frac{\beta_0 + \beta_1 p + \cdots + \beta_{r-1} p^{r-1}}{p^r} = 0.\beta_{r-1} \dots \beta_0 \in [0, 1). \quad (2)$$

We also use the notation  $0.w$  for  $0.\beta_{r-1} \dots \beta_0$ .

Along with finite words, we also consider one-side infinite words over the alphabet  $\mathcal{A}$ ; these are the infinite sequences of the form  $\dots \alpha_2 \alpha_1 \alpha_0$  where  $\alpha_i \in \mathcal{A}, i \in \mathbb{N}_0$ . In this paper, we may write one-side infinite words either stretching from left to right or from right to left when convenient, i.e., both  $\alpha_0 \alpha_1 \alpha_2 \dots$  and  $\dots \alpha_2 \alpha_1 \alpha_0$  denote the same word. For finite words, we may also use both notations, left and right, and the order of indices of letters in the word shows which of the two notations is used. For infinite words, notions of prefix, suffix, and subwords are defined in the same way as they are for finite words; note that suffixes are always infinite words whilst prefixes and subwords are always finite words. Let an infinite word  $w$  be eventually periodic; that is, let

$$w = \dots \beta_{t-1} \beta_{t-2} \dots \beta_0 \beta_{t-1} \beta_{t-2} \dots \beta_0 \alpha_{r-1} \alpha_{r-2} \dots \alpha_0$$

for  $\alpha_i \beta_j \in \mathcal{A}$ ; then, the subword  $\beta_{t-1} \beta_{t-2} \dots \beta_0$  is called a *period* of the word  $w$ , and the suffix  $\alpha_{r-2} \dots \alpha_0$  is called the *preperiod* of the word  $w$ . Note that a preperiod may be an empty word, while a period cannot. We ultimately write the periodic word  $w$  as  $w = (\beta_{t-1} \beta_{t-2} \dots \beta_0)^\infty \alpha_{r-1} \alpha_{r-2} \dots \alpha_0$ .

### 3.2. *p*-adic Integers

We briefly recall some very basic facts about *p*-adic integers referring the reader to any monograph on *p*-adic analysis (e.g., to [20]) for deeper introduction to the subject. Let  $p > 1$  be an integer. A *p*-adic integer  $z \in \mathbb{Z}_p$  can be uniquely represented by a *canonical form*  $z = \sum_{i=0}^\infty \zeta_i p^i$ , where  $\zeta_i \in \{0, 1, \dots, p - 1\}, (i = 0, 1, 2, \dots)$ . Thus, to every infinite sequence  $\mathbf{z} = (\zeta_i)_{i=0}^\infty$ , we put into a correspondence a *p*-adic integer represented by a respective canonical form. The sequences  $\mathbf{z}$  may also be treated as (one-side) infinite words over the alphabet  $\{0, 1, \dots, p - 1\}$ ; thus, we now can expand a mapping  $\text{num}$  to the set  $\mathbf{W}_p$  of all infinite sequences over  $\{0, 1, \dots, p - 1\}$  so that  $\text{num}(\mathbf{z}) = \sum_{i=0}^\infty \zeta_i p^i \in \mathbb{Z}_p$ . The so defined mapping  $\text{num}: \mathbf{W}_p \rightarrow \mathbb{Z}_p$  is one-to-one; thus, in what follows, we will not distinguish when necessary between *p*-adic integers, (one-side) infinite sequences over  $\{0, 1, \dots, p - 1\}$ , and infinite words over the alphabet  $\{0, 1, \dots, p - 1\}$ .

The sequences  $\mathbf{z}$  which contain only finitely many nonzero terms correspond to non-negative integers from  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$  represented by their base-*p* expansions; the sequences  $\mathbf{z}$  which contain only finitely many terms not equal to  $p - 1$  correspond to negative integers  $-\mathbb{N} = \{-1, -2, -3, \dots\}$ . The sequences  $\mathbf{z}$  which are ultimately periodic

correspond to rational  $p$ -adic integers  $z \in \mathbb{Z}_p \cap \mathbb{Q}$ ; i.e., to rational numbers which can be represented by irreducible fractions  $u/v$  whose denominators  $v$  are coprime to  $p$ . Any  $z \in \mathbb{Z}_p \cap \mathbb{Q}$  can be represented as  $z = c + d/(p^t - 1)$  where  $c \in \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\}$ ,  $t \in \mathbb{N}$ ,  $d \in \{0, 1, \dots, p^t - 2\}$ .

The rational  $p$ -adic integers constitute a subring  $\mathbb{Z}_p \cap \mathbb{Q}$  of  $\mathbb{Z}_p$  which is a dense subset of  $\mathbb{Z}_p$  with respect to the  $p$ -adic metric. The metric is induced by the  $p$ -adic absolute value  $|z|_p$  which is equal to  $p^{-\text{ord}_p z}$ , where  $\text{ord}_p z$  is the length of the longest zero-prefix (the prefix which consists of zeros only) of  $z$  if  $z \neq 0$ , and  $|0|_p = 0$  by definition.

Given  $n \in \mathbb{N} = \{1, 2, 3, \dots\}$  and a canonical expansion  $z = \sum_{i=0}^{\infty} \alpha_i p^i$  for  $z \in \mathbb{Z}_p$ , we further denote  $z \bmod p^n = \sum_{i=0}^{n-1} \alpha_i p^i \in \mathbb{N}_0$ . The mapping  $\text{mod } p^n: z \mapsto z \bmod p^n$  can be treated as a ring epimorphism of  $\mathbb{Z}_p$  onto the residue ring  $\mathbb{Z}/p^n\mathbb{Z}$ , under a natural representation of elements of the residue ring by the least non-negative residues  $\{0, 1, \dots, p^n - 1\}$ . Given  $n \in \mathbb{N}$ , the base- $p$  expansion of  $n$  is a finite word over  $\mathbb{F}_p$  whose length is  $\lfloor \log_p n \rfloor + 1$ . As the base- $p$  expansion of 0 is a one-letter word (namely, 0), in what follows, we assume that  $\lfloor \log_p 0 \rfloor = 0$ . We stress that when considering words corresponding to numbers, for numbers  $0, 1, 2, \dots$ , we distinguish their base- $p$  expansions from their canonical  $p$ -adic representations: the latter are treated as infinite words rather than as finite words. We also stress that the mapping  $\text{num}: \mathcal{W} \rightarrow \mathbb{N}_0$  is a surjection but not one-to-one, whereas the mapping  $\text{num}: \mathbf{W}_p \rightarrow \mathbb{Z}_p$  is one-to-one. In what follows, it always will be clear from the context what domain of  $\text{num}$  is considered.

A probability measure  $\mu$  on  $\mathbb{Z}_p$  can be defined as follows: elementary  $\mu$ -measurable sets are balls  $\mathbf{B}_{p^{-r}}(a) = \{b \in \mathbb{Z}_p: b \equiv a \pmod{p^r}\} \subset \mathbb{Z}_p$ , where  $a \in \mathbb{Z}_p$ ,  $r \in \mathbb{N}$ ; put  $\mu(\mathbf{B}_{p^{-r}}(a)) = p^{-r}$ . As the balls are simultaneously open and closed in topology induced by the  $p$ -adic absolute value  $|\cdot|_p$  and as every two balls are either disjoint or one of them contains another one, the balls constitute a base of sigma-algebra which define a sigma-additive measure  $\mu$  on  $\mathbb{Z}_p$ . Actually, this measure  $\mu$  is a Haar measure normalised so that  $\mu(\mathbb{Z}_p) = 1$ . The measure  $\mu$  is a Borel measure; that is, every open subset is  $\mu$ -measurable (hence, every closed subset is  $\mu$ -measurable as well). The measure  $\mu$  is regular; that is, for any  $\mu$ -measurable subset  $A \subset \mathbb{Z}_p$

$$\mu(A) = \sup\{\mu(S) : S \subset A, S \text{ is closed in } \mathbb{Z}_p\} = \inf\{\mu(S) : S \supset A, S \text{ is open in } \mathbb{Z}_p\}$$

Thus,  $\mathbb{Z}_p$  is a totally disconnected compact metric space whose metric is induced by the  $p$ -adic absolute value  $|\cdot|_p$  and a probability space with respect to the measure  $\mu$ . Note that the probability measure agrees with the metric; i.e., any function  $\mathbb{Z}_p \rightarrow \mathbb{Z}_p$  that is continuous with respect to the metric is measurable:  $f^{-1}(S)$  is  $\mu$ -measurable once  $S \subset \mathbb{Z}_p$  is  $\mu$ -measurable. Also note that the  $p$ -adic metric  $d_p(a, b) = |a - b|_p$  (where  $a, b \in \mathbb{Z}_p$ ) is non-Archimedean; that is, the triangle inequality holds for that metric in a stronger form:

$$|a - b|_p \leq \max\{|a - c|_p, |c - b|_p\}, \text{ for all } a, b, c \in \mathbb{Z}_p$$

In a similar way, the metric and probability measure can be defined for spaces  $\mathbb{Z}_p^n = \underbrace{\mathbb{Z}_p \times \dots \times \mathbb{Z}_p}_n$ , but in this paper, this  $n$ -dimensional space is mentioned only briefly

in appropriate places; in order not to overload the exposition, we limit our “working space” to  $\mathbb{Z}_p$ .

### 3.3. Systems, Transducers, Automata, Sequential Machines

Terminology in automata theory is somewhat diverse; in order to avoid a misunderstanding of the basic notions, we state them below.

**Definition 2** (System, transducer, automaton, sequential machine). A (discrete) system (or a system with discrete time  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ ) is a 5-tuple  $\mathfrak{A} = \langle \mathcal{I}, \mathcal{S}, \mathcal{O}, S, O \rangle$  where

- $\mathcal{I}$  is a nonempty finite set, the input alphabet;

- $\mathcal{O}$  is a nonempty finite set, the output alphabet;
- $\mathcal{S}$  is a nonempty (possibly, infinite) set of (epistemic) states;
- $S: \mathcal{J} \times \mathcal{S} \rightarrow \mathcal{S}$  is a state transition function;
- $O: \mathcal{J} \times \mathcal{S} \rightarrow \mathcal{O}$  is an output function.

The system is called *autonomous* if neither  $S$  nor  $O$  depend on input letters (that is, if  $S: \mathcal{S} \rightarrow \mathcal{S}$ ,  $O: \mathcal{S} \rightarrow \mathcal{O}$ ); otherwise, the system is called *nonautonomous*. A subsystem  $\mathfrak{A}'$  of  $\mathfrak{A}$  is a system  $\langle \mathcal{J}, \mathcal{S}', \mathcal{O}, S, O \rangle$  such that  $\mathcal{O} \neq \mathcal{S}' \subset \mathcal{S}$  and  $S(\chi, s') \in \mathcal{S}'$  for all  $\chi \in \mathcal{J}, s' \in \mathcal{S}'$ . A subsystem is called *minimal* if it has no subsystems other than itself. An *initial automaton* (or in other terminology, a *letter-to-letter transducer* [21], a *Mealy sequential machine* [19], an *initial synchronous automaton* [22])  $\mathfrak{A}(s_0)$  is a system where one of the states,  $s_0 \in \mathcal{S}$ , is fixed;  $s_0$  is called the *initial state*.

In what follows, the term *automaton* stands for an initial automaton; the subsystems of the automata are also called *subautomata*. A noninitial state  $s \in \mathcal{S}$  is called *reachable* (or, *accessible*) if there exists a finite sequence  $\chi_0, \chi_1, \dots, \chi_{N-1} \in \mathcal{J}$  such that  $S(\chi_{N-1}, s_{N-1}) = s$ , where  $s_i = S(\chi_{i-1}, s_{i-1}), i = 1, 2, \dots, N - 1$ ; i.e., if there exists a *path from the initial state  $s_0$  to  $s$  of finite length  $N$* .

An automaton  $\mathfrak{A}$  determines a unique map  $f_{\mathfrak{A}}: \dots \chi_2 \chi_1 \chi_0 \mapsto \dots \xi_2 \xi_1 \xi_0$  from the set  $\mathbf{W}(\mathcal{J})$  of all (one-side) infinite words over the alphabet  $\mathcal{J}$  to the set  $\mathbf{W}(\mathcal{O})$  of all (one-side) infinite words over the alphabet  $\mathcal{O}$ , as follows: at time instant  $i = 0$ , the automaton, being in the state  $s_0$ , accepts the first input letter  $\chi_0$ , updates its state to a newer state  $s_1 = S(\chi_0, s_0)$ , and produces an output letter  $\xi_0 = O(\chi_0, s_0)$ ; at the next time instant  $i = 1$ , the automaton accepts  $\chi_1$ , updates its state to  $s_2 = S(\chi_1, s_1)$ , and produces an output letter  $\xi_1 = O(\chi_1, s_1)$  etc. Therefore,  $\xi_i = \varphi_i(\chi_0, \dots, \chi_i)$ , where  $\varphi_i: \mathcal{J}^i \rightarrow \mathcal{O}$  is a uniquely determined sequence of maps. The mapping  $f_{\mathfrak{A}}$  is called an *automaton function* of the automaton  $\mathfrak{A}$ ; clearly, the mapping is causal. It is well known that the converse is also true: *every causal mapping  $f: \mathcal{J} \rightarrow \mathcal{O}$  is an automaton function of a suitable automaton  $\mathfrak{A}_f$*  (see, e.g., [19] [Chapter IV, Theorem 8.2]). This is why for the rest of this paper we use the terms *causal function, automaton map, automaton function, automatic function, and 1-Lipschitz function* as synonyms.

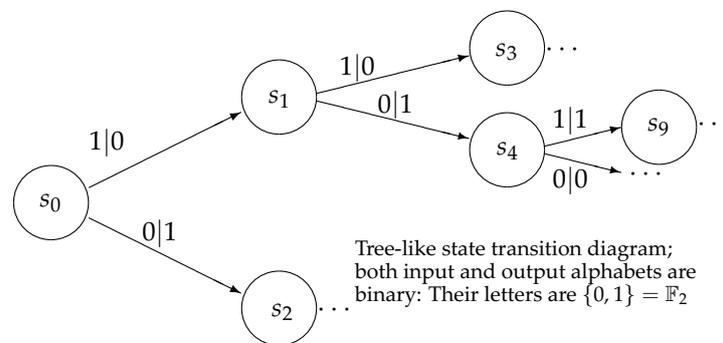
For instance, take a prime number  $p$  and consider an automaton whose input (respectively, output) alphabet is  $m$ -tuple  $(\alpha_1, \dots, \alpha_m) \in \mathbb{F}_p^m = \mathcal{J}$  (respectively,  $n$ -tuple from  $\mathbb{F}_p^n = \mathcal{O}$ ); then, the *automaton function is a map  $\mathbb{Z}_p^m \rightarrow \mathbb{Z}_p^n$  which satisfies a Lipschitz condition with a constant 1* (further, *1-Lipschitz for brevity*) *with respect to the  $p$ -adic metric* which is defined by the  $p$ -adic absolute value  $|(z_1, \dots, z_k)|_p = \max\{|z_1|_p, \dots, |z_k|_p\}$  on  $\mathbb{Z}_p^k$  (here  $z_j = \sum_{i=0}^{\infty} \alpha_{ji} p^i \in \mathbb{Z}_p, \alpha_{ji} \in \mathbb{F}_p, j = 1, 2, \dots, k$ ). Moreover, *every 1-Lipschitz map  $f: \mathbb{Z}_p^m \rightarrow \mathbb{Z}_p^n$  is an automaton function of a suitable automaton  $\mathfrak{A}_f$* . Note that it is convenient sometimes to consider automata whose input/output alphabets' cardinalities  $\#\mathcal{J}, \#\mathcal{O}$  are *multiplicatively dependent* (i.e., such that  $\#\mathcal{J}, \#\mathcal{O}$  are powers of some integer  $p > 1$ ) as automata having multiple inputs/outputs; i.e., to consider the 1-Lipschitz map  $f: \mathbb{Z}_p^m \rightarrow \mathbb{Z}_p^n$  as an automaton function of an automaton having  $m$  input channels and  $n$  output channels, each channel over a  $p$ -symbol alphabet. That is, the automaton function in this case is a multivariate map over infinite words over a  $p$ -symbol alphabet. In what follows, we will refer to such a case as to *multivariate*.

It is clear that a composition of automaton functions is an automaton function of an automaton which is a sequential composition of respective automata. For automata (and for their functions), the Cartesian product and Kronecker product can also be defined, but we do not need these constructions within the scope of the current paper.

Given  $f$ , the automaton  $\mathfrak{A}_f$  is not unique in the meaning of Definition 2: There are infinitely many different automata (i.e., the ones whose sets of epistemic states are different, whose state transition functions are different, whose output functions are different) whose automaton function is  $f$ . Therefore, an *observer can only make guesses about the "internal structure" of the system by observing pairs of "causes and effects", i.e., pairs  $(z, f(z)), z \in \mathbb{Z}_M$ ; moreover, the equivalent states are indistinguishable for the observer*. However, given  $f$  there exists a unique automaton whose automaton function is  $f$  and whose set of states  $\mathcal{S}$  is

the “smallest”. Call the two states  $s_i, s_j \in S$  of the automaton  $\mathfrak{A}$  *equivalent*; if whenever  $s_i, s_j$  are taken as initial states, the word mappings performed by either of the two initial automata are equal to one to another; i.e., if the input words are equal one to another, then the corresponding output words are also equal one to another. Factorising the state set of the automaton  $\mathfrak{A}$  by the equivalence relation, we obtain an automaton having no equivalent nonequal states whose automaton function is  $f_{\mathfrak{A}}$ . An automaton function  $f_{\mathfrak{A}}$  is called *finite* if it can be produced by an automaton whose set of states is finite; that is, the factor set by the equivalence relation is finite.

It is convenient to represent automata by their *state transition diagrams* (or *Moore diagrams*), which are directed graphs (the *digraphs*) whose vertices are states and whose arrows are state transitions, with the arrows labelled by *input letter|output letter*. Given an automaton function  $f: \mathbb{Z}_p^m \rightarrow \mathbb{Z}_p^n$ , there exists an automaton whose automaton function is  $f$  and whose state transition diagram is an infinite tree such that each vertex (i.e., a state) has exactly  $p^m$  outgoing arrows which go to  $p^m$  different vertices, cf., Figure 1 which depicts a state transition diagram of an automaton whose automaton function is  $f: \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$ .



**Figure 1.** State transition diagram of a 2-adic automaton. Label  $\alpha|\beta$  of the arrow that goes from the state  $s_i$  to the state  $s_j$  means that if the automaton is in the state  $s_i$  and obtains  $\alpha$  as the input symbol, it changes its state to  $s_j$  and produces  $\beta$  as the output symbol.

The automaton function of the automaton whose state transition diagram is depicted by Figure 1 is  $f(z) = z + 1$  ( $z \in \mathbb{Z}_2$ ), the 2-adic odometer. The *reduced* state transition diagram (which is obtained by factorisation with the equivalence relation defined earlier) is a digraph having only two vertices, cf., Figure 2. The automaton whose state transition diagram is depicted as in Figure 2 has the same automaton function  $f(z) = z + 1$  on  $\mathbb{Z}_2$ ; thus, the function  $f$  is a finite automaton function since it is produced by an automaton having only two states. Note that a *finite automaton is minimal if and only if its state transition diagram is a strongly connected digraph*; i.e., given any two vertices, there is a path connecting the vertices. The 2-adic odometer, therefore, has the only minimal subautomaton, the one whose set of states consists of the only state  $s_1$ .

Recall that a *path* in a digraph is a (finite or infinite) sequence of arrows  $\vec{a}_0, \vec{a}_1, \dots$  such that for every pair  $\vec{a}_j, \vec{a}_{j+1}$  of the arrows there is a state  $s$  such that the arrow  $\vec{a}_j$  goes to  $s$  and  $\vec{a}_{j+1}$  goes from  $s$ . In a state transition diagram of an automaton having input alphabet  $\mathcal{A}$ , to every path there corresponds a word  $\chi_0\chi_1\dots$  over  $\mathcal{A}$  where  $\chi_j$  are input letters, the ones which occupy the first positions in the label  $\alpha|\beta$  of the arrow: if  $\mathcal{A} = \{0, 1, \dots, p - 1\}$  then to every path  $\vec{a}_0 \vec{a}_1 \dots$  that starts from the initial state  $s_0$ , there corresponds the  $p$ -adic integer  $\chi_0 + \chi_1 p + \dots + \chi_{k-1} p^{k-1} + \dots$  where  $\chi_j|\cdot$  is a label which marks the arrow  $\vec{a}_j, j = 0, 1, 2, \dots$ . Simply speaking, the word  $\chi_0\chi_1\dots$  is an input word such that when an automaton is fed by that word, the automaton updates its states  $s_0 \rightarrow s_1 \rightarrow s_2 \rightarrow \dots$  where  $s_j$  is a state from which the arrow  $\vec{a}_j$  starts and  $s_{j+1}$  is a state to which the arrow  $\vec{a}_j$  goes; thus, the states  $s_j, s_{j+1}$  are connected by the arrow  $\vec{a}_j$  which goes from  $s_j$  to  $s_{j+1}$  and which is labelled as  $\chi_j|\cdot$ .

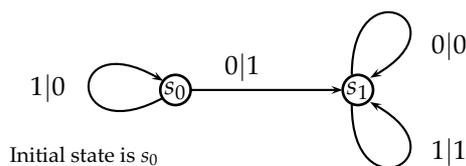


Figure 2. Reduced state transition diagram of the 2-adic odometer.

The statement of the following proposition is well known; see, e.g., [10]:

**Proposition 1** (Finite and nonfinite automata functions). Both addition  $+: \mathbb{Z}_p^2 \rightarrow \mathbb{Z}_p$  and multiplication  $\cdot: \mathbb{Z}_p^2 \rightarrow \mathbb{Z}_p$  are automata functions; addition is a finite automaton function, whereas multiplication is not. A constant map  $f: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  is a finite automaton function if and only if  $f(z) = \text{const} \in \mathbb{Z}_p \cap \mathbb{Q}$  for all  $z \in \mathbb{Z}_p$ . An affine map  $f(z) = az + b, (z \in \mathbb{Z}_p)$  is a finite automaton function if and only if  $a, b \in \mathbb{Z}_p \cap \mathbb{Q}$ .

Automata functions of automata whose input/output alphabets are  $\mathbb{F}_p$  can be explicitly represented via Mahler series. Recall that if  $p > 1$  is an integer (which is not necessarily a prime), then every function  $f: \mathbb{N}_0 \rightarrow \mathbb{Z}_p$  (or, respectively,  $f: \mathbb{N}_0 \rightarrow \mathbb{Z}$ ) has the only Mahler expansion; that is, has a unique representation via the so-called Mahler (interpolation) series [20]:

$$f(x) = \sum_{i=0}^{\infty} a_i \binom{x}{i}, \tag{3}$$

where  $a_i \in \mathbb{Z}_p$  (respectively,  $a_i \in \mathbb{Z}$ ),  $i = 0, 1, 2, \dots$ , and

$$\binom{x}{i} = \frac{x(x-1) \cdots (x-i+1)}{i!}$$

for  $i = 1, 2, \dots$ ;

$$\binom{x}{0} = 1,$$

by definition. The following reciprocity relations hold:

$$a_i = \sum_{j=0}^i (-1)^j \binom{i}{j} f(i-j), \quad i = 0, 1, 2, \dots \tag{4}$$

The function  $f: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  represented by series (3) is continuous with respect to the  $p$ -adic metric if and only if  $a_i$  tends  $p$ -adically to 0 as  $i$  tends to infinity.

To represent functions of several variables, one may use interpolation series of the following form:

$$f(x_1, \dots, x_n) = \sum_{(i_1, \dots, i_n) \in \mathbb{N}_0^n} a_{i_1, \dots, i_n} \binom{x_1}{i_1} \binom{x_2}{i_2} \cdots \binom{x_n}{i_n}; \tag{5}$$

Here,  $a_{i_1, \dots, i_n} \in \mathbb{Z}_p$ . As the map  $f: \mathbb{Z}_p^n \rightarrow \mathbb{Z}_p$  is an automaton function (of the automaton having  $n$  inputs and one output over a  $p$ -symbol alphabet  $\mathbb{F}_p$ ), the following Theorem 1 completely describes the automaton functions. Note that  $\lfloor \log_p i \rfloor$  is the smallest integer which does not exceed  $\log_p i$ ; thus,  $\lfloor \log_p i \rfloor$  is reduced by 1 number of digits in the base- $p$  expansion of  $i \in \mathbb{N}_0$ ; thus,  $\lfloor \log_p 0 \rfloor = 0$ .

**Theorem 1** ([23] [Theorem 3.53]). A function  $f: \mathbb{Z}_p^n \rightarrow \mathbb{Z}_p$  represented by the Mahler expansion (5) is 1-Lipschitz (with respect to the  $p$ -adic metric) if and only if

$$|a_{i_1, \dots, i_n}|_p \leq p^{-v(i_1, \dots, i_n)},$$

where  $v(i_1, \dots, i_n) = \max\{\lfloor \log_p i_k \rfloor : k = 1, 2, \dots, n\}$ .

In particular, a univariate function  $f: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  represented by the Mahler expansion (3) is 1-Lipschitz if and only if

$$|a_i|_p \leq p^{-\lfloor \log_p i \rfloor}$$

for all  $i = 1, 2, \dots$ . In other words, a function  $f: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  is automatic if and only if it can be represented as

$$f(x) = \sum_{i=0}^{\infty} c_i p^{\lfloor \log_p i \rfloor} \binom{x}{i}, \tag{6}$$

for suitable  $c_i \in \mathbb{Z}_p; i = 0, 1, 2, \dots$

**Note 2.** The series (6) converges uniformly on  $\mathbb{Z}_p$ . Given a 1-Lipschitz function  $f: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$ , the representation (2) is unique.

There are explicit representations of automaton functions in other terms (e.g., via van der Put series, digital derivatives) which are not needed within the scope of the paper; an interested reader is referred to an expository paper [24]. Additionally, it is worth noting that Moore sequential machines are initial automata whose output function depends only on states, cf. Definition 2, but it is well known that the latter machines are equivalent to Mealy machines in the following meaning: under the assumption that an output of a Moore machine at initial state is an empty symbol (i.e., no output), then the classes of causal functions represented by Mealy machines and by Moore machines coincide; however, to represent a causal function via a state transition diagram of a Moore machine, one needs more states compared to the diagram of the respective Mealy machine. This is why in the rest of the paper, the example state transition diagrams are given for Mealy machines although, from a physical point of view, it might be more natural to deal with Moore machines since they appear to be defined on Markov chains whilst Mealy machines are not, as the output of Moore machines formally depends only on states rather than on arrows reaching the states; however, this view is misleading since Mealy machines do exactly what Moore machines do.

Finally, automaton functions is the concept which illuminates the sharp difference between the two approaches, the 't Hooft's one based on cellular automata and ours based on letter-to-letter transducers: the class of functions computed by the transducers is much smaller than the class of functions computed by cellular automata. To exemplify this, consider one more type of transducer, the letter-to-word transducer (or, asynchronous initial automata, [22]) whose output function is  $\mathcal{J} \times \mathcal{S} \rightarrow \mathcal{W}_\phi$  rather than  $\mathcal{J} \times \mathcal{S} \rightarrow \mathcal{O}$  and where  $\mathcal{W}_\phi$  is the set of all finite words (including the empty word  $\phi$ ) over the output alphabet  $\mathcal{O}$ , c.f., Definition 2. In the case when  $\mathcal{J} = \{0, 1, \dots, p - 1\}$ , an asynchronous initial automaton, produces a map  $\mathbb{Z}_p \rightarrow \mathbb{Z}_p$  that can be constructed by an analogy with the synchronous case; then, the maps  $\mathbb{Z}_p \rightarrow \mathbb{Z}_p$ , which are automaton functions of nondegenerate synchronous initial automata, constitute the class of all functions that are continuous with respect to the  $p$ -adic metric, c.f., [22] [Theorem 2.4]. Therefore, these functions are defined by the maps  $\mathbb{N}_0 \rightarrow \mathbb{N}_0$  as  $\mathbb{N}_0$  is dense in  $\mathbb{Z}_p$  with respect to the  $p$ -adic metric. The automata functions of initial synchronous automata are all of the form (6), so if  $f$  is an automaton function

of a synchronous automaton such that  $f: \mathbb{N}_0 \rightarrow \mathbb{N}_0$ , then necessarily  $c_i p^{\lfloor \log_p i \rfloor} \in \mathbb{Z}$  for all  $i \in \mathbb{N}_0$  as the value of  $a_i$  for every  $i$  can be calculated by using (4). Therefore, from the algorithmic point of view,  $f$  is a primitive recursive function. In a similar way, it can be shown that the functions  $\mathbb{N}_0 \rightarrow \mathbb{N}_0$  which are automaton functions of nondegenerate asynchronous automata are also a primitive recursive function since they can be uniquely expanded to continuous  $p$ -adic functions  $\mathbb{Z}_p \rightarrow \mathbb{Z}_p$  and thus are of the form (3). However, a class of cellular automata is Turing-complete; therefore, the automaton functions of cellular automata (which can be defined for these automata as well) constitute the class of all general recursive functions; hence, they are not even everywhere defined on  $\mathbb{N}_0$ , let alone  $p$ -adic continuity or 1-Lipschitzness. In other words, one may say that the class

of automata functions of initial synchronous automata is the smallest class of automata functions, whereas the class of automata functions of cellular automata is the largest one.

### 3.4. On the Dynamics of Causal Functions

Here, we briefly recall some facts about the dynamics of automaton functions following [23]; i.e., on the dynamics of the  $p$ -adic 1-Lipschitz functions. The dynamics arises quite naturally since the automaton function of a sequential composition of automata is a composition of automaton functions. In addition, we recall from [25] a few general notions and facts from dynamical system theory which will be needed in subsequent steps.

A map  $F: \mathbb{S} \rightarrow \mathbb{Y}$  from a measure space  $\mathbb{S}$  into a measure space  $\mathbb{Y}$  endowed with probability measures  $\mu$  and  $\nu$ , respectively, is said to be *measure-preserving* if  $\mu(F^{-1}(S)) = \nu(S)$  for each measurable subset  $S \subset \mathbb{Y}$ ; in the case when  $\mathbb{S} = \mathbb{Y}$  and  $\mu = \nu$ , a measure-preserving map  $F$  is said to be *ergodic* if given a measurable subset  $S$  such that  $F^{-1}(S) = S$ , either  $\mu(S) = 1$  or  $\mu(S) = 0$ ; the map  $F$  is called *weak mixing* if for any two measurable sets  $A, B$ , there exists a sequence  $n_k \rightarrow \infty$  over  $\mathbb{N}_0$  such that  $\mu(F^{-n_k}(A) \cap B) \rightarrow \mu(A)\mu(B)$  as  $k \rightarrow \infty$ . If  $n_k = k$ , the weak mixing is called strong mixing. Weak mixing implies ergodicity but is a stronger condition than is ergodicity: the map  $F$  is weak mixing if and only if the map  $(x, y) \mapsto (F(x), F(y))$  of  $\mathbb{S} \times \mathbb{S}$  into  $\mathbb{S} \times \mathbb{S}$  is ergodic.

**Example 1** (Trivial although important). *Let  $\mathbb{S}$  be a finite set,  $\#\mathbb{S} = N$ , which is endowed with a uniform probability measure  $\mu$ : given  $A \subset \mathbb{S}$ ,  $\#A = M$ , we put  $\mu(A) = \frac{M}{N}$ . A transformation  $f$  on  $\mathbb{S}$  is measure-preserving if and only if  $f$  is bijective, i.e., if  $f$  is a permutation on  $\mathbb{S}$ . The map  $f$  is ergodic if and only if this permutation consists of a single cycle, i.e., if it is transitive on  $\mathbb{S}$ .*

**Definition 3** (Topological transitivity). *Given a topological space  $\mathbb{X}$  and a continuous mapping  $f: \mathbb{X} \rightarrow \mathbb{X}$ , the mapping  $f$  (as well as the respective dynamical system) is called topologically transitive if there exists a dense orbit of  $f$ ; that is, if there exists  $x \in \mathbb{X}$  such that the set of iterations  $\{f^i(x) \mid i \in \mathbb{N}_0\}$  is everywhere dense in  $\mathbb{X}$ . A dynamical system is called minimal if every orbit is dense.*

There is another (generally, nonequivalent to the above) definition of topological transitivity: the map  $f$  is called topologically transitive if for every pair of nonempty open sets  $U, V \subset \mathbb{X}$ , there exists a non-negative integer  $\ell$  such that  $f^\ell(U) \cap V \neq \emptyset$ . However, as in the sequel we deal with spaces  $\mathbb{X} = \mathbb{Z}_p^n$ ,  $n \in \mathbb{N}$ , the two definitions are equivalent since the spaces have no isolated points and are separable and of second category.

**Definition 4** (Unique ergodicity). *A mapping  $f: \mathbb{S} \rightarrow \mathbb{S}$  is called uniquely ergodic if there exists a unique  $f$ -invariant probability measure  $\mu$  on  $\mathbb{S}$ ; i.e., such that  $f$  is ergodic with respect to  $\mu$ .*

**Proposition 2** ([25] [Corollary 4.3.6]). *A minimal isometry of a compact metric space is uniquely ergodic.*

Given a 1-Lipschitz function  $f: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$ , a map  $f \bmod p^k: z \mapsto f(z) \bmod p^k$  is a well-defined map of the residue ring  $\mathbb{Z}/p^k\mathbb{Z}$  into itself, cf., Section 3.2. This map is called an *induced function modulo  $p^k$* . The function induced modulo  $p^k$  by a 1-Lipschitz function  $F: \mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$  can be defined by analogy.

**Definition 5** (Bijectivity and transitivity modulo  $p^k$ ). *A 1-Lipschitz function  $F: \mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$  is said to be a bijective modulo  $p^k$  (respectively, a transitive modulo  $p^k$ ) whenever the induced function  $F \bmod p^k: (\mathbb{Z}/p^k\mathbb{Z})^n \rightarrow (\mathbb{Z}/p^k\mathbb{Z})^n$  is bijective (respectively, transitive).*

In what follows, if the measure is not specified explicitly, measure preservation and ergodicity are defined with respect to the Haar probability measure on  $\mathbb{Z}_p^n$ , cf., Section 3.2. The following Theorem and Proposition are proven in [23] [Chapter 4].

**Theorem 2** (Main ergodic theorem for 1-Lipschitz  $p$ -adic dynamics). *A 1-Lipschitz function  $F: \mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$  is measure-preserving (or, accordingly, ergodic) if and only if it is bijective, (or, accordingly, transitive) modulo  $p^k$  for all  $k = 1, 2, 3, \dots$*

**Proposition 3.** *A function  $F: \mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$  is measure-preserving and 1-Lipschitz if and only if it is an isometry of  $\mathbb{Z}_p^n$  onto itself. A measure-preserving 1-Lipschitz function  $F$  is ergodic if and only if it has a dense orbit; moreover, all orbits of ergodic 1-Lipschitz function  $F: \mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$  are dense.*

The space  $\mathbb{Z}_p^n$  is a probability space and a metric (and thus topological) space. Therefore, for a continuous function  $\mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$ , one can define a *metric entropy* (related to the probability) and a *topological entropy* (related to the topology). In general, given  $F$ , these entropies may differ. However, for 1-Lipschitz functions  $F$ , both entropies coincide and are 0. Indeed, it is known that if  $G: \mathbb{X} \rightarrow \mathbb{X}$  is an isometry of a compact metric space  $\mathbb{X}$  onto itself, then the topological entropy of  $G$  is 0, cf., e.g., [26] [Exercise 6.3]. Yet, the variational principle for the topological entropy necessitates that the topological entropy of a continuous transformation  $G$  of a compact metric space  $\mathbb{X}$  is a supremum of all metric entropies of  $G$  with respect to  $G$ -invariant measures on  $\mathbb{X}$ , cf., [26] [Theorem 6.8.1]; this proves the claim. Moreover, from Proposition 3 it follows that given a 1-Lipschitz ergodic map  $F: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$ , the map  $F \times F: (x, y) \mapsto (F(x), F(y))$  of  $\mathbb{Z}_p^2 \rightarrow \mathbb{Z}_p^2$  is never ergodic since an orbit which starts from  $(z, z) \in \mathbb{Z}_p^2$  is never dense in  $\mathbb{Z}_p^2$ , so  $F$  is never weak mixing.

We summarize as follows:

- A function  $F: \mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$  is measure-preserving and 1-Lipschitz if and only if it is isometric.
- A 1-Lipschitz function  $F: \mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$  is isometric if and only if it is bijective; i.e., if and only if the respective automaton is *time-reversible*: an automaton  $\mathfrak{A}$  whose automaton function is  $F$  is called *time-reversible* if there exists an automaton  $\mathfrak{B}$  whose automaton function is  $G$  and such that  $G = F^{-1}$ , i.e., the composition  $G(F)$  is an identity map  $\mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$ . The time-reversibility is also called *automaton weak invertibility*, [27].
- All 1-Lipschitz functions  $\mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$  have zero topological entropy (thus, zero metric entropy).
- All 1-Lipschitz ergodic maps  $F: \mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$  are uniquely ergodic.
- None of the 1-Lipschitz ergodic maps  $F: \mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$  is weak mixing.
- Every orbit of every 1-Lipschitz ergodic map  $F: \mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$  is dense.

When  $n = 1$  the following is true [28] [Theorem 6]:

**Theorem 3.** *Let  $f: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  be surjective and 1-Lipschitz. The following propositions are equivalent:*

- (i)  $f$  is minimal;
- (ii)  $f$  is conjugate to the translation  $\tau: x \mapsto x + 1$  on  $\mathbb{Z}_p$ ;
- (iii)  $f$  is uniquely ergodic;
- (iv)  $f$  is ergodic.

In subsequent steps, we will need the following sufficient conditions of measure-preservation/ergodicity for 1-Lipschitz functions  $\mathbb{Z}_p \rightarrow \mathbb{Z}_p$ , [23] [Lemma 4.41]:

**Lemma 1.** *Given a 1-Lipschitz function  $f: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  and  $p$ -adic integers  $c, d, c \not\equiv 0 \pmod{p}$ , the function  $g(x) = d + cx + p \cdot f(x)$  is 1-Lipschitz measure-preserving and the function  $h(x) = c + x + p \cdot \Delta f(x)$  is 1-Lipschitz ergodic. (Here,  $\Delta$  is a difference operator  $\Delta f(x) = f(x + 1) - f(x)$  by definition.)*

#### 4. Completely Consistent Functions

Causal functions over discrete time  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$  from Section 2 are the maps  $f$  from the set  $\mathbf{W}(\mathcal{J})$  of all infinite words over a finite alphabet  $\mathcal{J}$  to the set  $\mathbf{W}(\mathcal{O})$  of all

infinite words over a finite alphabet  $\mathcal{O}$  which are 1-Lipschitz with respect to standard non-Archimedean metric  $d$  on the words,  $d(f(\mathbf{u}), f(\mathbf{v})) \leq d(\mathbf{u}, \mathbf{v})$ , or, which is the same if and only if words  $f(\mathbf{u})$  and  $f(\mathbf{v})$  have a common prefix of length at least  $k$  whenever respective words  $\mathbf{u}$  and  $\mathbf{v}$  have a common prefix of length  $k$ .

As a composition of automaton functions is an automaton function, the following example introduces an important class of functions which are automaton functions for every  $p$  (by Proposition 1) and, moreover, which at the same time can be considered as continuous real functions.

**Example 2** (Polynomials over  $\mathbb{Z}$  are automata functions). *A polynomial map  $f: z \mapsto f(z)$  where  $f(x) \in \mathbb{Z}[x]$  is an automaton function;  $f$  is never a finite automaton function if  $\deg f \geq 2$ .*

That is, as the set  $\mathbb{Z}_p \cap \mathbb{Q}$  of all rational  $p$ -adic integers is dense both in  $\mathbb{Z}_p$  with respect to the  $p$ -adic metric on  $\mathbb{Z}_p$  for every  $p$  and with respect to usual real metric on  $\mathbb{R}$ , the map induced by a polynomial  $f \in \mathbb{Z}[x]$  is well-defined both on  $\mathbb{Z}_p$  for all  $p$  and on  $\mathbb{R}$ ; i.e., the map  $f: z \mapsto f(z)$ , ( $z \in \mathbb{Z}_p \cap \mathbb{Q}$ ), can be uniquely extended both to continuous maps  $f: u \mapsto f(u)$ , ( $u \in \mathbb{Z}_p$ ), for all  $p$ , and to a continuous map  $f: y \mapsto f(y)$ , ( $y \in \mathbb{R}$ ). This is because any polynomial map is a composition of additions and multiplications, and these operations are well-defined and continuous both on all  $\mathbb{Z}_p$  and on  $\mathbb{R}$  with respect to corresponding metrics and agree on  $\mathbb{Z}_p \cap \mathbb{Q}$ .

#### 4.1. Universally Causal Functions

The maps  $f: \mathbb{N}_0 \rightarrow \mathbb{Z}$  defined by polynomials over  $\mathbb{Z}$  are examples of functions which we call *universally causal*; these are the functions which, loosely speaking, are causal with respect to all finite alphabets  $\mathcal{A}$  and  $\mathcal{B}$  such that  $\#\mathcal{A} = \#\mathcal{B} = r$  for whatever  $r \in \{2, 3, 4, \dots\}$  is taken. Here is a formal definition.

**Definition 6** (Universally causal functions). *A causal function  $f: (a_i)_{i=0}^\infty \mapsto (\varphi_i^f(a_i))_{i=0}^\infty$  whose domain is all sequences  $\mathbf{a} = (a_i)_{i=0}^\infty$  over  $\mathcal{A}$  and whose codomain is all sequences  $\mathbf{b} = (b_i)_{i=0}^\infty$  over  $\mathcal{B}$  (see Section 1) is called universally causal if  $\#\mathcal{A} = \#\mathcal{B} = r > 1$ , and there exist bijections  $\alpha: \mathcal{A} \leftrightarrow \{0, 1, \dots, r - 1\}$  and  $\beta: \mathcal{B} \leftrightarrow \{0, 1, \dots, r - 1\}$  such that the induced map  $\tilde{f}: \mathbb{Z}_r \rightarrow \mathbb{Z}_r$  defined by  $\tilde{f}: \sum_{i=0}^\infty \alpha(a_i)r^i \mapsto \sum_{i=0}^\infty \beta(\varphi_i^f(a_0, \dots, a_i))r^i$ , c.f., (1), satisfies the following conditions:*

- (i)  $\tilde{f}(\mathbb{N}_0) \subset \mathbb{Z}$ , where  $\mathbb{N}_0$ , the rational non-negative integers, are all  $r$ -adic integers whose canonical  $r$ -adic representations contain only a finite number of nonzero terms; and  $\mathbb{Z}$ , the rational integers, are either non-negative rational integers or negative rational integers. The latter are all  $r$ -adic integers whose canonical  $r$ -adic representations contain only a finite number of terms other than  $(r - 1)r^i$ .
- (ii)  $\tilde{f}(m) \equiv \tilde{f}(n) \pmod{q}$  once  $m \equiv n \pmod{q}$ , where  $m, n, q \in \mathbb{N}_0$ ,  $q > 1$ .

The class of universally causal functions is much wider than that of functions defined by polynomials over  $\mathbb{Z}$ . Actually, up to the bijections  $\alpha, \beta$ , the universally causal functions constitute a class of the so-called *pseudo-polynomials*, Ref. [29] or *universal functions* [30]; these are maps  $g: \mathbb{N}_0 \rightarrow \mathbb{Z}$  which satisfy (ii) from Definition 6.

**Theorem 4** (On pseudo-polynomials). *A map  $g: \mathbb{N}_0 \rightarrow \mathbb{Z}$  is a pseudo-polynomial if and only if  $g$  can be represented as*

$$g(z) = c_0 + \sum_{i=1}^\infty c_i \cdot \text{lcm}\{1, 2, \dots, i\} \cdot \binom{z}{i} = c_0 + \sum_{i=1}^\infty c_i \cdot e^{\psi(i)} \cdot \binom{x}{i}, \tag{7}$$

where  $c_i \in \mathbb{Z}$ ,  $\text{lcm}\{1, 2, \dots, i\}$  is the least common multiple of the numbers  $1, 2, \dots, i$ , and  $\psi(i) = \sum_{q \leq i, q \text{ prime}} \lfloor \log_q i \rfloor \ln q$  is the second Chebyshev function,  $i = 1, 2, \dots$  (recall that  $\psi(i) = i + o(i)$ ).

In the literature, often only the functions of the form (7) which are not polynomials are called pseudo-polynomials, but in the current paper, we call pseudo-polynomials all functions of that form. The class of pseudo-polynomials is wide and is a subject of study for a number theorists, who focus mostly on Ruzsa’s conjecture, which is about the sufficient conditions for when a pseudo-polynomial is a polynomial; see, e.g., [31]. Classical examples of pseudo-polynomials which are not polynomials are  $\sum_{i=0}^{\infty} x^i$  and  $\sum_{i=0}^{\infty} (-1)^i x^i$ , where  $x^i$  is the  $i$ -th falling factorial power,  $x^i = x(x - 1) \cdots (x - i + 1)$ , if  $i > 0$  and  $x^0 = 1$ .

**Note 3.** *The following is noteworthy.*

- *Even if all but a finite number of  $c_i$  in (7) are 0, i.e., if  $g$  is a polynomial, then  $g$  is not necessarily a polynomial with integer coefficients, although  $g$  is polynomial over  $\mathbb{Q}$ . For instance, put  $c_4 = 1$  and put  $c_i = 0$  for  $i \neq 4$ .*
- *If all but a finite number of  $c_i$  are 0, the function  $g$  is well-defined on  $\mathbb{R}$ ; that is,  $f$  can be uniquely expanded to a map  $\mathbb{R} \rightarrow \mathbb{R}$  which is continuous with respect to the real metric.*
- *For every  $p > 1$ , the map  $g$  can be uniquely expanded to 1-Lipschitz (thus, automatic) map  $\mathbb{Z}_p \rightarrow \mathbb{Z}_p$ , cf., (6) from Theorem 1.*

4.2. *The Main Theorem on Complete Consistency*

Therefore, polynomials of the form (7) satisfy Conditions 1 for every prime  $p$ . It turns out that the converse statement is also true. Note that a function which satisfies the conditions for all prime  $p$  must be universally causal, i.e., it must be a pseudo-polynomial; however, the only pseudo-polynomials which are well-defined on  $\mathbb{R}$  are polynomials since if an infinite number of  $c_i$  in (7) are nonzero then the series diverges at; for example,  $z = -1$  as the common term at  $z = -1$  is  $(-1)^i c_i \cdot \text{lcm}\{1, 2, \dots, i\}$  and thus does not go to 0 as  $i \rightarrow \infty$ . However, this argument does not prove the converse claim since, for instance, if  $g$  is a pseudo-polynomial which is not a polynomial, then the composition  $g(z^2)$  is also a pseudo-polynomial, but the map  $z \mapsto g(z^2)$  is well-defined on  $\mathbb{Z}$ . Nonetheless, the following theorem holds true.

**Theorem 5** (Functions which satisfy Conditions 1). *A continuous function  $f: \mathbb{R} \rightarrow \mathbb{R}$  satisfies Conditions 1 if and only if  $f$  is a polynomial of the form (7); i.e., when all but a finite number of  $c_i$  in (7) are zero.*

**Proof.** According to Theorem 4, every polynomial  $g$  over  $\mathbb{Q}$  of the form (7) satisfies (i) from Conditions 1, cf., (ii) of Definition 6. Therefore,  $g$  also satisfies (ii) from Conditions 1 since  $g(\mathbb{Q}) \subset \mathbb{Q}$  as  $g$  is a polynomial over  $\mathbb{Q}$ .

To prove the converse claim, note that the map  $u: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  is 1-Lipschitz if and only if  $\Delta^i u(z)/i \in \mathbb{Z}_p$  for all  $z \in \mathbb{Z}_p$  and all  $i \in \mathbb{N}$ , cf., [23] [Proposition 3.38] or [32] [Proposition 3.1]. Here,  $\Delta$  is the (forward) difference operator, i.e.,  $\Delta^1 u(z) = \Delta u(z) = u(z + 1) - u(z)$ ,  $\Delta^{i+1} u(z) = \Delta(\Delta^i u(z))$ . Therefore, we have the following:

$$\frac{\Delta^i f(z)}{i} \in \mathbb{Z}, \text{ for all } z \in \mathbb{Z}, i \in \mathbb{N}. \tag{8}$$

Further, from (ii) of Conditions 1, it follows (by Note 1) that  $\frac{1}{h} \cdot (f(z + h) - f(z)) = q(z, h) \in \mathbb{Q}$  for all  $z, h \in \mathbb{Q}$  and  $h \neq 0$  since  $z, h \in \mathbb{Z}_p \cap \mathbb{Q}$  for all but not more than a finite number of primes  $p$ . However,  $f(z + h) = \sum_{i=0}^{\infty} \binom{h}{i} \Delta^i f(z)$  where the series converges  $p$ -adically for all but not more than a finite number of primes  $p$  as  $\Delta^i f(z)$  tends  $p$ -adically to 0 according to Theorem 1; cf., (i) of Conditions 1. Thus, the series converges to some  $q'(z, h) \in \mathbb{Q}$  by (ii) of Conditions 1, and, therefore, the series converges in  $\mathbb{R}$  to that rational number  $q'(z, h)$ . We have

$$\frac{1}{h}(f(z + h) - f(z)) = \frac{1}{h} \sum_{i=1}^{\infty} \binom{h}{i} \Delta^i f(z) = \sum_{i=1}^{\infty} \binom{h-1}{i-1} \frac{\Delta^i f(z)}{i}, \tag{9}$$

where the series in the right hand part converges in  $\mathbb{R}$  to the rational number  $q(z, h) \in \mathbb{Q}$ ; therefore, the absolute value  $|\binom{h-1}{i-1} \frac{\Delta^i f(z)}{i}|$  must tend to 0 in  $\mathbb{R}$  as  $i \rightarrow \infty$ . Represent the following:

$$\binom{h-1}{i-1} = \left(\frac{h}{1} - 1\right) \left(\frac{h}{2} - 1\right) \cdots \left(\frac{h}{i-1} - 1\right)$$

From here, it follows that

$$\left|\binom{h-1}{i-1}\right| \geq \left|1 - \frac{h}{i-1}\right|^{i-1} > 0 \text{ for all } i = 2, 3, \dots; -1 < h < 0$$

As for  $-1 < h < 0$  rational, it holds

$$\lim_{i \rightarrow \infty} \left|1 - \frac{h}{i-1}\right|^{i-1} = e^{-h} > 0,$$

from the convergence of the series in the right hand part of (9), and it follows necessarily that  $\lim_{i \rightarrow \infty} \frac{\Delta^i f(z)}{i} = 0$ ; therefore, according to (8), given  $z \in \mathbb{Z}$ , then  $\Delta^i f(z) = 0$  for all sufficiently large  $i$ . In particular,  $\Delta^i f(0) = 0$  for all sufficiently large  $i$ . As  $f(x) = \sum_{i=0}^{\infty} \binom{x}{i} \Delta^i f(0)$  and, in view of Conditions 1 (i), the series in the right hand part converges  $p$ -adically in  $\mathbb{Z}_p$ , then, according to Note 2, we finally conclude that  $f$  is a polynomial over  $\mathbb{Q}$ ; hence, a polynomial of the form (7).  $\square$

**Definition 7** (Totally consistent functions). *Further in the paper, functions described by Theorem 5 are called totally consistent;  $\mathcal{C}(\mathbb{R})$  denotes the class of all totally consistent functions.*

**Note 4.** *In view of Theorem 1, the statement of Theorem 5 holds true for continuous real functions  $\mathbb{R}^m \rightarrow \mathbb{R}^n$  as well. The proof is a minor modification of the proof of the said theorem and thus is omitted.*

**Note 5.** *From the proof of Theorem 5, it follows that relaxation of Conditions 1 to functions  $f$  whose domain contains a real interval rather than coincides with the whole  $\mathbb{R}$  does not widen the class of functions.*

### 4.3. The Free Choice of Discreteness/Continuity

We stress once again that in the measurement of values of physical quantities, the rational  $p$ -adic integers  $\mathbb{Z}_p \cap \mathbb{Q}$  are indistinguishable from rational numbers  $\mathbb{Q}$  since every real number can be approximated by a rational  $p$ -adic integer with any desirable accuracy. Note also that polynomials over  $\mathbb{Z}$  are totally consistent; cf. Example 2. The theorem by M. I. Chlodovsky states that a continuous real-valued function on a real interval which does not contain integers can be uniformly approximated by polynomials over  $\mathbb{Z}$  [33,34]. Therefore, according to Theorem 5, any continuous real function on the real interval  $[\alpha, \beta]$  where  $0 < \alpha < \beta < 1$  can be uniformly approximated (with respect to a real metric) by completely consistent functions, i.e., by functions from  $\mathcal{C}_{\text{all primes}}(\mathbb{R})$ . On the other hand, Theorems 1 and 5 imply that any  $p$ -adic 1-Lipschitz function  $f: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  can be uniformly approximated (with respect to the  $p$ -adic metric) by completely causal functions, regardless of which prime  $p$  is taken.

Indeed, according to (6), the function  $f$  can be represented by the Mahler expansion  $f(z) = \sum_{i=0}^{\infty} b_i p^{\lfloor \log_p i \rfloor} \binom{z}{i}$  where  $b_i \in \mathbb{Z}_p$ . According to Theorem 5, given  $n \in \mathbb{N}$ , we must find a polynomial  $g(x) = c_0 + \sum_{i=1}^{\infty} c_i \cdot \text{lcm}\{1, 2, \dots, i\} \cdot \binom{x}{i}$  where all  $c_i \in \mathbb{Z}$  such that  $f(z) \equiv g(z) \pmod{p^n}$  for all  $z \in \mathbb{Z}_p$ . As  $\text{lcm}\{1, 2, \dots, i\} = \prod_{(q)} q^{m_{i,q}}$  where  $q^{m_{i,q}}$  is the largest power of a prime  $q$  that does not exceed  $i$ , then  $m_{i,q} = \lfloor \log_q i \rfloor$ , and therefore  $c_i \cdot \text{lcm}\{1, 2, \dots, i\} = c_i p^{\lfloor \log_p i \rfloor} a_i$  where  $a_i = \text{lcm}\{1, 2, \dots, i\} / p^{\lfloor \log_p i \rfloor}$  is in  $\mathbb{Z}$  and is coprime to  $p$ . Hence, given  $b_i \in \mathbb{Z}_p$ , a congruence  $b_i \equiv c_i a_i \pmod{p^n}$  has an integer solution  $c_i \in \mathbb{Z}$ . Put  $c_i = 0$  for all  $i$  such that  $\lfloor \log_p i \rfloor \geq n$ , and let  $c_0 \in \mathbb{Z}$  be the least non-negative residue of  $b_0 \in \mathbb{Z}_p$  modulo  $p^n$ . Then, the so-defined polynomial  $g$  is the one we need.

All the considerations already outlined in this paper may be taken as evidence in favour of the following plausible statement which answers the question to which the whole book [2] is devoted:

**Interpretation 1** (Observer's free choice of discreteness/continuity). *Due to the inevitable nonzero error in the measurements of values of physical quantities, an observer's conclusion whether Nature on the smallest of the scales is discrete or continuous completely depends on the observer's free choice of metric with respect to which the observer processes the measured numerical data. Moreover, the very "degree of the discreteness", the number  $p$ , is subject to observer's free choice.*

#### 4.4. The Free Choice of Chaoticity/Predictability

The next important question which should be addressed is related to the 't Hooft causality postulate and can be posed as follows: Can an observer determine through numerical observational data whether Nature on the smallest of scales is random or absolutely predictable? In what follows, the second term is understood as causality, i.e., if an observer probes a system by exposing it to some impacts, reactions of the system coincide whenever impacts coincide up to a precision of measurement equipment; that is, *the same causes imply same effects*, so the behaviour of the system is completely predictable since a cause results in a unique effect within the measurement precision. The *randomness means that the "same" causes may result in different effects*. Specifically, causes whose numerical values are indistinguishable in measurement since the values coincide up to the precision of measurement equipment may result in effects which are distinguishable by measurement, i.e., differences of numerical values of respective effects exceed the measurement error. This is why we treat what follows randomness as chaos in a broad meaning since the definitive feature of chaos is its extreme sensitivity to negligible distortions/perturbations.

Recall that there are many nonequivalent mathematical notions of chaos; see, e.g., the expository paper [35]. One of the most common of these definitions is in the work of R. L. Devaney [36] [Definition 8.5] which reads as follows:

**Definition 8** (Devaney's chaos on metric spaces). *Let  $F: \mathbb{X} \rightarrow \mathbb{X}$  be a continuous function on a metric space  $\mathbb{X}$  equipped with a metric  $d$ . The function  $F$  is said to be chaotic if it satisfies the following three conditions:*

- (i) Sensitive dependence on initial conditions: *There is  $\delta > 0$  such that, for any  $x \in \mathbb{X}$  and any neighbourhood  $A \subset \mathbb{X}$  of  $x$ , there exists  $y \in A$  and  $n \in \mathbb{N}_0$  such that  $d(F^n(x), F^n(y)) > \delta$ .*
- (ii) Topological transitivity: *Given any pair of open subsets  $U, V \subset \mathbb{X}$ , there exists  $k \in \mathbb{N}$  such that  $F^k(U) \cap V \neq \emptyset$ .*
- (iii) Density of periodic points: *The set of all periodic points of  $F$  is dense in  $\mathbb{X}$  (a point  $x \in \mathbb{X}$  is called periodic if  $F^k(x) = x$  for some  $k \in \mathbb{N}$ ).*

It is known that conditions (i)–(iii) are *not independent*. In [37], it is proven that sensitive dependence on the initial conditions is a redundant element in Devaney's definition because it follows from topological transitivity and denseness of the periodic points; in [38], it is shown by construction of counter examples, that neither topological transitivity nor denseness of the periodic points follow from the remaining two properties. In [39], it is proven that chaos, according to Devaney's definition, may exist in bounded but noncompact spaces without any nonperiodic orbits. For bounded metric spaces, however, the following theorem is true:

**Theorem 6** (C. Knudsen, [39]). *Let  $F, \mathbb{X}, d$  be as that in Definition 8; let  $\mathbb{X}$  be bounded; let  $f = F|_{\mathbb{Y}}$  be a restriction of  $F$  to a dense subset  $\mathbb{Y}$  of  $\mathbb{X}$ . Then, we obtain the following:*

- $F: \mathbb{X} \rightarrow \mathbb{X}$  is topologically transitive, if and only if  $f: \mathbb{Y} \rightarrow \mathbb{Y}$  is topologically transitive;
- $F: \mathbb{X} \rightarrow \mathbb{X}$  exhibits sensitive dependence on the initial conditions, if and only if  $f: \mathbb{Y} \rightarrow \mathbb{Y}$  exhibits sensitive dependence on the initial conditions.

The following definition of chaos on a bounded metric space is from Knudsen.

**Definition 9** (Knudsen’s chaos on bounded metric spaces [39]). *Let  $F$  be a continuous transformation of a bounded metric space  $\mathbb{X}$ . If  $F$  has a dense orbit in  $\mathbb{X}$  and if  $F$  exhibits sensitive dependence on the initial conditions, then  $F$  is said to be chaotic.*

We stress that to the best of our knowledge, all definitions of chaos on metric spaces contain sensitive dependence on initial conditions as an inherent property; other conditions vary, but the sensitive dependence condition is always present, [40]. For other various types of chaos on compact metric spaces  $\mathbb{X}$ , see [35]. We only mention that a continuous map  $F: \mathbb{X} \rightarrow \mathbb{X}$  is called *topologically chaotic* if topological entropy of  $F$  is positive. The topological chaos implies *Li-Yorke chaos*, which is yet one more widely known type of chaos, for whose definition the reader is referred to [35]. In addition, positive topological entropy implies *distributional chaos* of type DC2, [41]. Chaos can also be defined in terms of measure-preserving transformations of measure spaces rather than of metric spaces; see [41].

The “chaos-like” behaviour may also be expressed in terms of “blending capability” which we first illustrate by an example taken from [42]. If in a cocktail shaker of volume 1 there are 10 shares of gin and 90 shares of vermouth then, after *ergodic* shaking, in every volume  $V$  of the shaker there will be 10 shares of gin and 90 shares of vermouth *on average*, whereas after *strong-mixing* shaking, in every  $V$  there will be approximately 10 shares of gin and 90 shares of vermouth; after *weak-mixing* shaking, *in every  $V$  with the exception of some rare instants* there will be 10 shares of gin and 90 shares of vermouth. Formally, a measure-preserving transformation  $F$  is by definition strong mixing if  $\lim_{n \rightarrow \infty} \mu(F^{-n}(A) \cap B) = \mu(A)\mu(B)$  for every  $\mu$ -measurable subsets  $A, B$ . Thus, if  $\mu$  is a probability measure, the strong-mixing transformation, after being applied a sufficiently large number of times, makes any two “events”  $A, B$  “independent” in the probabilistic meaning. As mentioned in Section 3.4, a 1-Lipschitz measure-preserving map can be neither strong nor weak mixing; only the ergodicity is possible.

Finalising the considerations of chaos, we claim that 1-Lipschitz functions  $F: \mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$  are deterministic and nonchaotic with respect to chaos of any type. Indeed, due to the 1-Lipschitzness, these functions exhibit no sensitive dependence on initial conditions, and their topological entropy is zero; hence, any metric entropy is zero; cf. Section 3.4. Moreover, as measure-theoretical chaos is defined only for measure-preserving maps, and as a 1-Lipschitz map  $F: \mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$  preserves the Haar probability measure if and only if it is an isometry, it can be easily shown that  $F$  is chaotic with respect to no type of measure-theoretic chaos defined in [41]. One may say, therefore, that totally consistent functions (see Definition 6) are the best candidates to be called *superdeterministic*. The latter term also must not be treated in the meaning which is common for physical theories [5] but rather as a mathematical notion to stress the “extremely nonchaotic” behaviour of the functions.

On the other hand, one may also say that totally consistent functions are similar to *Ianus Bifrons*: being deterministic with respect to a  $p$ -adic metric for every  $p$ , the totally consistent functions can nevertheless be chaotic if considered as real functions on a real interval. Let us consider an illustrative example.

A well-known “canonical” example of real chaotic maps, the logistic map  $L(x) = 2x(1 - x)$ , maps a real closed interval  $[0, 1]$  to  $[0, 1]$ . The map  $L$  has positive entropy  $\log 2$ . On the other hand,  $L$  is a polynomial with integer coefficients; hence, it is a totally consistent function, thus its entropy as a  $p$ -adic 1-Lipschitz map  $z \mapsto 2z(1 - z)$  (both topological and metric with respect to Haar probability measure) is 0, and  $L: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  is not sensitive to initial conditions. The map  $L$  on  $\mathbb{Z}_2$  is not measure-preserving with respect to the Haar probability measure on  $\mathbb{Z}_2$ ; it has the only point of attraction (namely, 0) to which all orbits converge; thus,  $L$  is not topologically transitive on  $\mathbb{Z}_2$ .

However, the map  $L$  is *ergodic on the 3-adic sphere  $\mathbf{S}_{1/27}(0)$  of radius  $1/27$  centred at 0* since 0 is a fixed point of  $L$  and  $L'(0) = 2$  is a generator of the group of units modulo 9;

see [23] [Theorem 4.79] or [43] [Theorem 5.7]. Specifically, the sphere  $S_{1/27}(0)$  is a disjoint union of two 3-adic balls  $B_{1/81}(27)$  and  $B_{1/81}(54)$ , the sphere is invariant under the action of  $L$  on  $\mathbb{Z}_3$ , and the sphere is measurable with respect to the Haar probability measure on  $\mathbb{Z}_3$ . Thus, the probability measure on  $\mathbb{Z}_3$  induces a probability measure on  $S_{1/9}(0)$  with respect to which the action of  $L$  on the sphere is measure-preserving and ergodic. The set of all rational 3-adic numbers from  $S_{1/27}(0)$  which lie in the real closed interval  $[0, 1]$  is dense in  $[0, 1]$  with respect to the real metric. Therefore, as 3-adic rational integers are indistinguishable from real numbers by measurement due to inevitable nonzero error, the map  $L$  can be judged as measure-preserving and ergodic.

Now consider the map  $L$  on the 3-adic sphere  $S_{1/27}(1)$ . The sphere is a disjoint union of balls  $B_{1/81}(28)$  and  $B_{1/81}(55)$ . The sphere is invariant under the action of  $L$  on  $\mathbb{Z}_3$  and is measurable with respect to the probability measure on  $\mathbb{Z}_3$ . The map  $L$  on  $S_{1/27}(1)$  is *measure-preserving with respect to the induced probability measure but is not ergodic* by the criterion of ergodicity on  $p$ -adic spheres (see [23] [Theorem 4.79] or [43] [Theorem 5.7]) since  $L'(1) = -2$  is not a generator of the group of units modulo 9. The set  $S_{1/27}(1) \cap [0, 1]$  is dense in  $[0, 1]$  with respect to the real metric; therefore, by the reasoning similar to that as above, the map  $L$  can be judged as measure-preserving but not ergodic.

Finally, the map  $L: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  is *measure-preserving for no  $p$*  as  $L$  is not a bijective modulo  $p$ ; cf. Theorem 2. However, the set  $\mathbb{Z}_p \cap \mathbb{Q} \cap [0, 1]$  is also dense both in  $\mathbb{Z}_p$  and in  $[0, 1]$  with respect to the  $p$ -adic and to the real metrics accordingly, so the values of the map  $L$  that takes on  $\mathbb{Z}_p \cap \mathbb{Q} \cap [0, 1]$  define a unique map both on  $[0, 1]$  and on  $\mathbb{Z}_p$  for every  $p$ . However, an observer’s measurement data may only be rational numbers due to the inevitable nonzero measurement error, and any rational number from  $[0, 1]$  can be approximated with arbitrarily high accuracy (with respect to the real metric) by numbers from  $\mathbb{Z}_p \cap \mathbb{Q} \cap [0, 1]$  regardless of whichever  $p$  is taken. In other words, numbers from  $\mathbb{Z}_p \cap \mathbb{Q} \cap [0, 1]$  (as well as from  $S_{1/27}(1) \cap [0, 1]$ , or from  $S_{1/27}(0) \cap [0, 1]$ ) are indistinguishable from numbers in  $\mathbb{Q} \cap [0, 1]$  and from numbers in  $[0, 1]$  by measurements due to nonzero measurement error, but the choice of metric (and of the dense subset) with respect to which the measured numbers are processed is crucial for the observer’s conclusion whether the obtained data are completely random or satisfy a strictly deterministic law.

All these facts can be judged as evidence in favour of the following assertion.

**Interpretation 2** (Observer’s free choice of determinism/randomness). *Due to the inevitable nonzero error in measurements of values of physical quantities, an observer’s conclusion as to whether Nature on the smallest of the scales is superdeterministic or random completely depends on the observer’s free choice of metric with respect to which the observer processes the measured numerical data.*

#### 4.5. $p$ -Consistent Functions

In view of the *finiteness assumption* (cf. the text which follows Definition 1), Conditions 1 may appear to be too restrictive since according to physical reasons, the number of “elementary causes” and “elementary effects” cannot be arbitrarily large; therefore, it does not exceed some  $p$ . This is a motivation to introduce the following class of causal functions, the (univariate)  $p$ -consistent functions: given a prime  $p$ , we denote via  $\mathcal{C}_p(\mathbb{R})$  the class of all continuous (with respect to the usual metric on  $\mathbb{R}$ ) functions  $\check{f}: \mathbb{R} \rightarrow \mathbb{R}$  such that the following conditions are satisfied:

- (i)  $\check{f}(\mathbb{Z}_p \cap \mathbb{Q}) \subset \mathbb{Z}_p \cap \mathbb{Q}$ ;
- (ii) There exists a  $p$ -adic 1-Lipschitz function  $f: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  such that the following are obtained:
  - $f(\mathbb{Z}_p \cap \mathbb{Q}) \subset \mathbb{Z}_p \cap \mathbb{Q}$
  - $f(z) = \check{f}(z)$  for every  $z \in \mathbb{Z}_p \cap \mathbb{Q}$

The multivariate  $p$ -consistent functions  $\mathbb{R}^m \rightarrow \mathbb{R}^n$  can be defined similarly.

Loosely speaking, the functions from  $\mathbb{C}_p(\mathbb{R})$  “are living simultaneously in two worlds”, the Archimedean one and the non-Archimedean one: any  $\check{f} \in \mathbb{C}_p(\mathbb{R})$  defines a unique 1-Lipschitz (i.e., automaton) function  $f: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  since  $\mathbb{Z}_p \cap \mathbb{Q}$  is dense in  $\mathbb{Z}_p$  with respect to  $p$ -adic metric, and vice versa, any  $f$  defines a unique continuous real function  $\check{f}: \mathbb{R} \rightarrow \mathbb{R}$  since  $\mathbb{Z}_p \cap \mathbb{Q}$  is dense in  $\mathbb{R}$  with respect to the real metric (this is why in what follows, we use the same symbol  $f$  for  $\check{f}$  as well).

The functions from  $\mathbb{C}_p(\mathbb{R})$  may suit the best for physical modelling of causal dependencies both at the macro- and micro- scales since values of  $f \in \mathbb{C}_p(\mathbb{R})$  on, e.g.,  $\mathbb{N}_0$ , completely define the function  $f$  on  $\mathbb{R}$ .

From this definition, it immediately follows that any function from  $\mathbb{C}_p(\mathbb{R})$  can be represented via a Mahler series (6) where all  $c_i$  are in  $\mathbb{Z}_p \cap \mathbb{Q}$ ; the series converges both on  $\mathbb{R}$  and on  $\mathbb{Z}_p$  with respect to the real and, accordingly, to the  $p$ -adic metric. It would be interesting to find necessary and sufficient conditions on the coefficients  $c_i$  when the series (6) defines a  $\mathbb{C}_p(\mathbb{R})$ -function. The general conditions are not yet known, but nevertheless it is clear that the class  $\mathbb{C}_p(\mathbb{R})$  is rich; for instance, it contains not only polynomials over  $\mathbb{Z}_p \cap \mathbb{Q}$  but also some rational functions.

**Example 3.** Given polynomials  $u, v \in \mathbb{Z}[x]$  such that  $v(z) \not\equiv 0 \pmod{p}$  for all  $z \in \mathbb{Z}_p$  and  $v(x) \neq 0$  for all  $x \in \mathbb{R}$ , the rational function  $f(x) = u(x)/v(x)$  is in  $\mathbb{C}_p(\mathbb{R})$ . The rational functions  $f$  are differentiable with respect to both the  $p$ -adic metric and the real metric; moreover,  $\check{f}' = f'$  everywhere on  $\mathbb{Z}_p \cap \mathbb{Q}$  and  $f' \in \mathbb{C}_p(\mathbb{R})$ ; c.f., [44] or [23] [Section 3.10.2].

For  $k \in \mathbb{N}$ , we denote as  $\mathbb{C}_p^k(\mathbb{R})$  (respectively via  $\mathbb{C}_p^\infty(\mathbb{R})$ ) the subclass of all functions which are  $k$ -times (respectively, infinitely many times) differentiable with respect to both  $p$ -adic and real metric, whose derivatives are also in  $\mathbb{C}_p(\mathbb{R})$ . Put  $\mathbb{C}_p^0(\mathbb{R}) = \mathbb{C}_p(\mathbb{R})$ . It is natural to ask, therefore, whether there exist functions in  $\mathbb{C}_p(\mathbb{R})$  which are not rational functions. The answer is affirmative.

**Theorem 7.** There exist functions in  $\mathbb{C}_p^\infty(\mathbb{R})$  which are not rational functions.

**Proof.** The theorem can be proven by employing ideas from [45,46]. The set  $\mathbb{Z}_p \cap \mathbb{Q}$  is countable; let us enumerate its elements as  $z_1, z_2, \dots$ . Define by simultaneous induction a sequence of functions  $g_0, g_1, g_2, \dots$  and integers  $m_0 < m_1 < m_2 < \dots$  as follows: put  $g_0(x) = 0, m_0 = 1$ . For  $n \geq 1$ , consider the following polynomial over the ring  $\mathbb{Z}_p \cap \mathbb{Q}$ :

$$h_n(x) = \prod_{i=1}^n (r_i - x) = a_{n,0} + a_{n,1}x + \dots + a_{n,n-1}x^{n-1} + (-1)^n x^n.$$

Put

$$g_n(x) = \frac{p^n h_n(x)}{(p^{2n} + 1) \lceil |a_{n,0}| + |a_{n,1}|m_{n-1} + |a_{n,n-1}|m_{n-1}^{n-1} + m_{n-1}^n \rceil_p}$$

where  $|\cdot|$  is the real absolute value and  $\lceil r \rceil_p$  for  $r \in \mathbb{R}$  is the smallest  $\ell \in \mathbb{N}$  such that  $r \leq \ell$  if  $\ell \not\equiv 0 \pmod{p}$ , or  $\lceil r \rceil_p = \ell + 1$  if  $\ell \equiv 0 \pmod{p}$ . Then,  $g_n(x)$  is a polynomial over  $\mathbb{Z}_p \cap \mathbb{Q}$  and  $\|g_n(c)\| < p^{-n}$  for every  $c \in \mathbb{C}$  whose complex absolute value  $\|c\| \leq m_{n-1}$ .

Now, if  $n$  is even, let  $m_n$  be first integer larger than  $m_{n-1}$  such that  $\sum_{i=0}^n g_i(m_n) \geq 2$ . If  $n$  is odd, let  $m_n$  be the first integer larger than  $m_{n-1}$  such that  $\sum_{i=0}^n g_i(m_n) \leq -2$ . Since the leading coefficient of  $h_n(x)$  is  $(-1)^n$ , these conditions are always true if  $m_n$  is large enough. After defining all  $g_n(x)$ , put  $g(x) = \sum_{i=1}^\infty g_i(x)$ . Then, the following is true:

- The sum  $g(x) = \sum_{i=1}^\infty g_i(x)$  converges uniformly in the open complex disk  $D_{m_n}(0)$  of radius  $m_n$  centred at 0 for all  $n \in \mathbb{N}$  because, except for the first  $n$  terms, every term  $g_i(c)$  is bounded absolutely by  $p^{-i}$ , and the sum of these converges.
- Because the uniform convergence in an open subset of  $\mathbb{C}$  preserves analyticity, the function  $g$  is analytic on every  $D_{m_n}(0)$  and so also on the whole  $\mathbb{C}$ .
- $g(\mathbb{Z}_p \cap \mathbb{Q}) \subset \mathbb{Z}_p \cap \mathbb{Q}$  since for every  $z = z_k \in \mathbb{Z}_p \cap \mathbb{Q}$ , all  $g_j(z) = 0$  for  $j \geq k$ .

- In the sequence  $(g(m_k))_{k=1}^\infty$ , the terms having odd indices  $k$  are less than  $-1$ , whereas the terms having even indices  $k$  are greater than  $1$  since  $\sum_{i=0}^n g_i(m_n) \geq 2$  for even  $n$  and  $\sum_{i=0}^n g_i(m_n) \leq -2$  for odd  $n$ , and the remaining terms in  $g(m_n)$  cannot change the whole sum for more than  $1$ .

Thus, the function  $g$  is well-defined on the whole  $\mathbb{R}$ ;  $g$  is a continuous function with respect to the real metric, and according to the intermediate value theorem, the function  $g$  has a zero between  $m_n$  and  $m_{n+1}$  for all sufficiently large  $n \in \mathbb{N}$ . Therefore,  $g$  has infinitely many zeroes in  $\mathbb{R}$  and thus cannot be of the form  $u(x)/v(x)$ , where  $u(x), v(x) \in \mathbb{Z}[x]$ . The function  $g$  according to this construction is a complex analytic function which is analytic on the whole  $\mathbb{C}$ ; thus, the restriction of  $g$  on  $\mathbb{R}$  is a function  $\mathbb{R} \rightarrow \mathbb{R}$  which is infinitely many times differentiable everywhere in  $\mathbb{R}$ , and each derivative is continuous with respect to real metric and thus is uniquely defined by its values on  $\mathbb{Z}_p \cap \mathbb{Q}$ , as  $\mathbb{Z}_p \cap \mathbb{Q}$  is dense in  $\mathbb{R}$  with respect to the real metric.

On the other hand, given  $a, b \in \mathbb{Z}_p \cap \mathbb{Q}$ ,  $a \neq b$ , there are unique  $k, n \in \mathbb{N}$  such that  $a = z_k, b = z_n$  with respect to the numeration of numbers in  $\mathbb{Z}_p \cap \mathbb{Q}$ . If  $n > k$ , then  $g(a) = \sum_{i=0}^{n-1} g_i(a)$ ,  $g(b) = \sum_{i=0}^{n-1} g_i(b)$ ; therefore,  $|g(a) - g(b)|_p \leq |a - b|_p$  since  $\sum_{i=0}^{n-1} g_i(x)$  is a polynomial over  $\mathbb{Z}_p \cap \mathbb{Q}$ ; thus, a unique continuation of  $g$  to the whole  $\mathbb{Z}_p$  is a  $p$ -adic 1-Lipschitz function. Let  $\bar{g}_k = g \bmod p^k$  be a polynomial over  $\mathbb{N}_0$  obtained by the reduction modulo  $p^k$  of the function  $g$  (it is clear that then  $\deg \bar{g}_k \leq k$  via the construction of  $g$ ). Then, the function  $g: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  can be uniformly approximated by the polynomials  $\bar{g}_k$  with respect to the  $p$ -adic sup-norm which is defined as follows: Given a  $p$ -adic 1-Lipschitz functions  $u: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$ , the  $p$ -adic sup-norm is  $\max\{|u(z)|_p : z \in \mathbb{Z}_p\}$ . Therefore the function  $g: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  is a  $\mathcal{B}$ -function, the Stone–Weierstrass completion of the polynomials over  $\mathbb{N}_0$  with respect to the said  $p$ -adic sup-norm; thus,  $g$  is infinitely many times differentiable with respect to the  $p$ -adic metric, all derivatives are  $\mathcal{B}$ -functions, and thus the derivatives are uniquely defined by their values on  $\mathbb{Z}_p \cap \mathbb{Q}$ , as  $\mathbb{Z}_p \cap \mathbb{Q}$  is dense in  $\mathbb{Z}_p$  with respect to the  $p$ -adic metric; see [44] [Proposition 4.4.] or [23] [Section 3.10.2, Proposition 3.59].

Therefore,  $g$  is infinitely many times differentiable both on  $\mathbb{R}$  and on  $\mathbb{Z}_p$ , and the values of the derivatives both with respect to the real and to the  $p$ -adic metric coincide on  $\mathbb{Z}_p \cap \mathbb{Q}$ . This finally proves that  $g$  is a  $\mathcal{C}_p^\infty(\mathbb{R})$ -function.  $\square$

The  $\mathcal{C}_p(\mathbb{R})$ -functions exhibit a sort of “hologram-likeness”. The values a  $\mathcal{C}_p(\mathbb{R})$ -function takes on arbitrarily small real interval, completely define the function on  $\mathbb{R}$  and on  $\mathbb{Z}_p$ . Recall that a complete hologram can be restored from a small piece of a holography plate.

**Theorem 8** (Hologram-likeness of  $\mathcal{C}_p(\mathbb{R})$ -functions). *Let  $f, g \in \mathcal{C}_p(\mathbb{R})$  and let  $(\alpha, \beta) \subset \mathbb{R}$  be any open interval; then  $f = g$  if and only if  $f(x) = g(x)$  for all  $x \in (\alpha, \beta) \subset \mathbb{R}$  (equivalently, for all  $x \in (\alpha, \beta) \cap \mathbb{Z}_p \cap \mathbb{Q}$ ).*

**Proof.** For  $n \in \mathbb{N}, d \in \{1, \dots, p^n - 2\}$  put  $z_{n,d} = \frac{d}{1-p^n}$ ; then,  $z_{n,d} \in (-1, 0) \subset \mathbb{R}, d \in \mathbb{Z}_p \cap \mathbb{Q}$ . If  $d = \zeta_0 + \zeta_1 p + \dots + \zeta_{n-1} p^{n-1}$  is the base- $p$  expansion of  $z$  then

$$z_{n,d} = \zeta_0 + \zeta_1 p + \dots + \zeta_{n-1} p^{n-1} + \zeta_0 p^n + \zeta_1 p^{n+1} + \dots + \zeta_{n-1} p^{2n-1} + \dots \tag{10}$$

is a  $p$ -adic canonical form of  $z_{n,d}$  as  $(1 - p^n)^{-1} = 1 + p^n + p^{2n} + p^{3n} + \dots \in \mathbb{Z}_p$ ; cf., Section 3.2. From 10, it immediately follows that the set  $\mathcal{Z}$  of all these  $z_{n,d}$  is dense in  $\mathbb{Z}_p$ . Therefore,  $f = g$  on  $\mathbb{Z}_p$  if and only if  $f = g$  on  $\mathcal{Z}$ , but  $f = g$  on  $\mathbb{Z}_p$  if and only if  $f = g$  on  $\mathbb{R}$ .

Let  $\alpha, \beta \in \mathbb{Z}_p \cap \mathbb{Q}$ ,  $\alpha < \beta$ . Put  $\gamma = -1$  if  $\alpha - \beta \leq -1$ ; let  $\gamma = \frac{1}{1-p^t}$  be such that  $0 > \gamma > \alpha - \beta$  if  $\alpha - \beta > -1$  for a suitable  $t \in \mathbb{N}$ . It is clear from what we have already proven that  $f(x) = g(x)$  for all  $x \in \mathbb{R}$  if and only if  $f(x) = g(x)$  for all  $x \in (\gamma^{-1}(\beta - \alpha), 0)$ , as  $(\gamma^{-1}(\beta - \alpha), 0) \supset (-1, 0)$ .  $\square$

**Interpretation 3** (“Causality” vs. “locality”). The proof of Theorem 8 shows that the “local” behaviour of  $\mathcal{C}_p(\mathbb{R})$ -functions completely defines their “global” behaviour. Given values of  $\mathcal{C}_p(\mathbb{R})$ -function takes on an arbitrarily small neighbourhood of an arbitrary point, the values the function takes at all other points can be “restored uniquely”. If the points of  $\mathbb{R}$  are treated as “positions” and values of the function as “measurement data” of a physical system to which the function is ascribed, then the data an observer obtains by probing a system in a given position let him completely predict values of physical quantities obtained by measurements at all other positions.

This property of  $\mathcal{C}_p(\mathbb{R})$ -functions is especially important since various classes of real functions can be approximated by  $\mathcal{C}_p(\mathbb{R})$ -functions.

**Theorem 9** (Approximations of real functions by  $\mathcal{C}_p(\mathbb{R})$ -functions).

- (i) Any continuous function  $g: [a, b] \rightarrow \mathbb{R}$  can be uniformly approximated on  $[a, b] \subset \mathbb{R}$  by  $\mathcal{C}_p^\infty(\mathbb{R})$ -functions.
- (ii) Any continuous function  $g: [a, b] \rightarrow \mathbb{R}$  can be uniformly approximated on  $[a, b] \subset \mathbb{R}$  by  $\mathcal{C}_p^\infty(\mathbb{R})$ -functions which are automaton functions of time-reversible automata.
- (iii) Any continuous function  $g: [a, b] \rightarrow \mathbb{R}$  can be uniformly approximated by ergodic automata functions from  $\mathcal{C}_p^\infty(\mathbb{R})$ .
- (iv) Any continuous function  $g: \mathbb{R} \rightarrow \mathbb{R}$  that vanishes at infinity can be uniformly approximated on  $\mathbb{R}$  by  $\mathcal{C}_p^\infty(\mathbb{R})$ -functions. (recall that a continuous function  $g: \mathbb{R} \rightarrow \mathbb{R}$  vanishes at infinity, if, for every  $\varepsilon > 0$ , there exists a compact set  $K \subset \mathbb{R}$  such that  $|g(x)| < \varepsilon$  for all  $x \in \mathbb{R} \setminus K$ ).
- (v) Any continuous function  $g: \mathbb{R} \rightarrow \mathbb{R}$  that vanishes at infinity can be uniformly approximated on  $\mathbb{R}$  by  $\mathcal{C}_p^\infty(\mathbb{R})$ -functions which are automaton functions of time-reversible automata.
- (vi) Any square-integrable function  $g: \mathbb{R} \rightarrow \mathbb{R}$  (and moreover, any function  $g: \mathbb{R} \rightarrow \mathbb{R}$  that is integrable with its  $n$ -th power for some  $n \in \mathbb{N}$ ) can be uniformly approximated by  $\mathcal{C}_p^\infty(\mathbb{R})$ -functions which are automaton functions of time-reversible automata.

**Proof.** The class  $\mathcal{C}_p^\infty(\mathbb{R})$  contains all polynomial functions over  $\mathbb{Z}$ . Chlodovsky theorem yields that a continuous real-valued function, which is defined on a real interval that does not contain an integer, can be uniformly approximated by polynomials over  $\mathbb{Z}$  [33,34]. If the interval  $[a, b]$  contains integers, take  $\alpha, \beta \in \mathbb{Z}_p \cap \mathbb{Q}$  such that the interval  $[a', b'] = [\alpha a + \beta, \alpha b + \beta]$  contains no integers (e.g., take  $m \in \mathbb{N}$  such that  $|b - a| < p^m - 1$  and put  $\alpha = 1/(p^m - 1)$ ). Given a continuous function  $g: [a, b] \rightarrow \mathbb{R}$ , the function  $g(\alpha^{-1}(x - \beta))$  can be uniformly approximated by polynomials  $u_i(x) \in \mathbb{Z}[x]$  on  $[a', b']$  by Chlodovsky’s theorem; thus, the function  $g$  can be uniformly approximated by polynomials  $u_i(\alpha x + \beta)$  on  $[a, b]$ . However,  $\hat{u}_i(x) = u_i(\alpha x + \beta)$  is a polynomial over  $\mathbb{Z}_p \cap \mathbb{Q}$  in variable  $x$  since  $\alpha, \beta \in \mathbb{Z}_p \cap \mathbb{Q}$ ; thus,  $\hat{u}_i \in \mathcal{C}_p^\infty(\mathbb{R})$ . This proves claim (i).

To prove claim (ii), consider the function  $\tilde{g}(x) = \frac{g(x)-x}{p}$ . In view of (i), since  $\tilde{g}(x)$  is continuous on  $[a, b]$ ,  $\tilde{g}$  can be uniformly approximated by  $\mathcal{C}_p^\infty(\mathbb{R})$ -functions  $u_i$ ; thus,  $g$  can be uniformly approximated by functions  $x + pu_i(x)$  which are also in  $\mathcal{C}_p^\infty(\mathbb{R})$ . However, given any 1-Lipschitz function  $u: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$ , the function  $z + pu(z)$  is 1-Lipschitz measure-preserving according to Lemma 1. Thus, the  $\mathcal{C}_p^\infty(\mathbb{R})$ -functions  $x + pu_i(x)$  are automata functions of time-reversible automata.

To prove claim (iii), note that the function  $g$  can be uniformly approximated by polynomials  $w_j(x)$  over  $\mathbb{Z}_p \cap \mathbb{Q}$ ; c.f., the proof of (i). Then, the difference equation  $\frac{w_j(x)-x-1}{p} = \Delta \tilde{w}_j(x)$  has a solution  $\tilde{w}_j(x)$  which is a polynomial since  $\frac{w_j(x)-x-1}{p}$  is. These  $\tilde{w}_j(x)$  can be uniformly approximated by polynomials  $u_{ji}(x)$  over  $\mathbb{Z}_p \cap \mathbb{Q}$ ; c.f., the proof of (i). Therefore,  $g$  can be uniformly approximated by polynomials over  $\mathbb{Z}_p \cap \mathbb{Q}$  of the form  $1 + x + p \cdot \Delta u(x)$  which are all ergodic according to Lemma 1.

To prove claim (iv), consider functions of the form  $\frac{ru(x)}{1+pv(x)^2}$  where  $r \in \mathbb{Z}_p \cap \mathbb{Q}$ ,  $u(x), v(x) \in \mathbb{Z}[x]$ , and  $\deg u(x) \leq \deg v(x)$ . All these functions vanish at infinity, are  $\mathcal{C}_p^\infty(\mathbb{R})$ -functions (c.f., Example 3) and separate points. Therefore, the  $\mathbb{R}$ -algebra  $\mathbf{A}$  generated by the set  $A$  of all these functions satisfies conditions of the Stone–Weierstrass theorem for locally compact spaces, i.e., the algebra is dense with respect to the topology of the uniform convergence in the Banach algebra of all real-valued continuous functions on  $\mathbb{R}$  which vanish at infinity. However, the set  $A$  is dense in  $\mathbf{A}$ .

In order to prove claim (v), note that in view of the proof of claim (iv), it suffices to approximate uniformly on  $\mathbb{R}$  the functions of the form  $h(x) = \frac{ru(x)}{1+pv(x)^2}$ , where  $r \in \mathbb{Z}_p \cap \mathbb{Q}$ ,  $u(x), v(x) \in \mathbb{Z}[x]$ , and  $1 \leq \deg u(x) \leq \deg v(x)$ , by  $\mathcal{C}_p^\infty(\mathbb{R})$ -functions which are automaton functions of time-reversible automata. Represent  $h(x) = \frac{x}{1+pv(x)^2} + \frac{p\tilde{u}(x)}{p(1+pv(x)^2)}$ , then  $\tilde{u}(x) \in \mathbb{Z}_p[x]$ ,  $\deg \tilde{u}(x) \leq \deg v(x)$ . Given  $c_i \in \mathbb{Z}_p \cap \mathbb{Q}$ , the function  $\frac{x}{1+pv(x)^2} + c_i \cdot \frac{p\tilde{u}(x)}{1+pv(x)^2}$  vanishes at infinity; moreover, this function is a  $\mathcal{C}_p^\infty(\mathbb{R})$ -function, and it is a measure-preserving 1-Lipschitz function  $\mathbb{Z}_p \rightarrow \mathbb{Z}_p$  since it is bijective modulo  $p$ , and its derivative modulo  $p$  vanishes nowhere; c.f., [44] [Corollary 3.3] or [23] [Theorem 4.45]. Taking a sequence  $(c_i)_{i=0}^\infty$  over  $\mathbb{Z}_p \cap \mathbb{Q}$  that converges to  $1/p$  in  $\mathbb{R}$ , we conclude that the function  $h(x)$  can be uniformly approximated on  $\mathbb{R}$  by  $\mathcal{C}_p^\infty(\mathbb{R})$ -functions which are measure-preserving 1-Lipschitz functions  $\mathbb{Z}_p \rightarrow \mathbb{Z}_p$ ; that is, automaton functions of time-reversible automata.

It is well known that functions which are integrable with their  $n$ -th powers, for some  $n \in \mathbb{N}$ , can be uniformly approximated by Schwartz functions; but the latter are smooth and vanish at infinity. With (v), this proves claim (vi) and the theorem.  $\square$

**Example 4.** Wave function  $\Psi(x, t)$  vanishes at infinity since it must satisfy the condition  $\lim_{t \rightarrow \pm\infty} \Psi(x, t) = 0$ ; see, e.g., [47] [Section 1.4]. Thus, wave functions can be uniformly approximated by automaton functions of time-reversible automata.

**Interpretation 4** (Observer’s free choice of arrow of time). Due to the inevitable nonzero error in measurements of values of physical quantities, an observer’s conclusion on the direction of “arrow of time” completely depends on the observer’s free choice of metric with respect to which the observer processes the measured numerical data: according to claims (ii) and (v)–(vi) of Theorem 9, “causes” can be recovered from “effects”, with any desirable accuracy. The “entropic arrow of time” also depends on the choice of metric since the value of entropy does as well; c.f., Section 4.4.

**Theorem 10** (On finite automata  $\mathcal{C}_p^1(\mathbb{R})$ -functions). Let a finite automaton function  $f \in \mathcal{C}_p^1(\mathbb{R})$ ; i.e., let  $f$  be differentiable both over  $\mathbb{R}$  and over  $\mathbb{Z}_p$ ; let  $f' \in \mathcal{C}_p(\mathbb{R})$ . Then,  $f$  is an affine function over  $\mathbb{Z}_p \cap \mathbb{Q}$ ; i.e.,  $f(x) = ax + b$  for suitable  $a, b \in \mathbb{Z}_p \cap \mathbb{Q}$ . Vice versa, all these affine functions are finite automaton functions from  $\mathcal{C}_p^\infty(\mathbb{R})$ .

**Proof.** Given a 1-Lipschitz function  $f$  for  $n \in \mathbb{N}_0, k \geq \lfloor \log_p n \rfloor + 1$ , consider functions  $f_{n,k} : \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  which are defined as follows:

$$f_{n,k}(z) = \frac{1}{p^k} \left( f(n + p^k z) - (f(n) \bmod p^k) \right) = \frac{f(n + p^k z) - f(n)}{p^k} - \frac{f(n) - f(n) \bmod p^k}{p^k} = \frac{f(n + p^k z) - f(n)}{p^k} - f_{n,k}(0), \quad (11)$$

for all  $z \in \mathbb{Z}_p$ . The function  $f$  is an automaton function of a finite automaton if and only if the collection  $\mathcal{F}$  of function  $f_{n,k}$  (where  $n \in \mathbb{N}_0, k \in \mathbb{N} = \{1, 2, 3, \dots\}, k \geq \lfloor \log_p n \rfloor + 1$ ) contains only a finite number of pairwise distinct functions. Note that  $f_{n,k}$  is the automaton function that corresponds to the automaton  $\mathfrak{A}(s(n_k)) = \langle \mathbb{F}_p, \mathcal{S}, \mathbb{F}_p, \mathcal{S}, O, s(n_k) \rangle$ , where  $s(n_k) \in \mathcal{S}$  is the state the automaton  $\mathfrak{A} = \mathfrak{A}_f = \langle \mathbb{F}_p, \mathcal{S}, \mathbb{F}_p, \mathcal{S}, O, s_0 \rangle$  reaches after it has been fed by the input word  $n_k$  (of length  $p^k$ ) that corresponds to the base- $p$  expansion of  $n$  (so the word  $n_k$

may contain some leading zeros that correspond to higher order digits of the expansion). That is, there are  $N, K \in \mathbb{N}$  such that for every  $n \in \mathbb{N}_0, k \in \mathbb{N}$ , one finds  $\check{n} \leq N, \check{k} \leq K$  such that  $f_{n,k}(z) = f_{\check{n},\check{k}}(z)$  for all  $z \in \mathbb{Z}_p$ .

Take  $z_k = \frac{h}{p^{2k-1}}$  where  $h \in \mathbb{N}$ . Note that  $z_k \in \mathbb{Z}_p \cap \mathbb{Q}$  and that  $\lim_{k \rightarrow \infty} p^k z_k = 0$  both with respect to real metric and to the  $p$ -adic metric. Then,

$$\frac{f_{n,k}(z_k)}{z_k} = \frac{f(n + p^k z_k) - f(n)}{p^k z_k} = \frac{f(n) - f(n) \bmod p^k}{p^k z_k}, \tag{12}$$

where  $\lim_{k \rightarrow \infty} \frac{f(n+p^k z_k) - f(n)}{p^k z_k} = f'(n)$  both with respect to the real metric and to the  $p$ -adic metric. Thus, from (11) it follows that  $\lim_{k \rightarrow \infty}^p (f_{n,k}(u) - f_{n,k}(0)) = u f'(n)$  for every  $u \in \mathbb{Z}_p, n \in \mathbb{N}_0$ . However,  $f'(n)$  is a derivative at  $n \in \mathbb{N}_0$  both with respect to the real metric and to the  $p$ -adic metric; however,  $f'(n)$  may take only a finite number of values due to the finiteness of the number of pairs  $n, k$  which enumerate pairwise distinct  $f_{n,k}$ . Therefore,  $f'(z)$  may take not more than a finite number of values on  $\mathbb{Z}_p$  since any  $z \in \mathbb{Z}_p$  is a  $p$ -adic limit of some sequence over  $\mathbb{N}_0$ , and  $f'$  is a continuous function  $\mathbb{Z}_p \rightarrow \mathbb{Z}_p$  according the conditions of the theorem. Hence,  $f'$  may take not more than a finite number of values on  $\mathbb{Z}_p \cap \mathbb{Q}$  and thus on  $\mathbb{R}$  since  $\mathbb{Z}_p \cap \mathbb{Q}$  is dense in  $\mathbb{R}$  and since  $f'$  is a continuous real function according the conditions of the theorem. Therefore, the derivative  $f'$  is a constant function over  $\mathbb{R}$  and thus over  $\mathbb{Z}_p \cap \mathbb{Q}$  and over  $\mathbb{Z}_p$ ; that is,  $f(x) = ax + b$  for some  $a, b \in \mathbb{Z}_p \cap \mathbb{Q}$ . Proposition 1 proves the converse claim of the theorem.  $\square$

**Note 6.** The theorem remains true for multivariate  $\mathbb{C}_p(\mathbb{R})$ -maps  $F: \mathbb{Z}_p^m \rightarrow \mathbb{Z}_p^n$  as well: affine maps over  $\mathbb{Z}_p \cap \mathbb{Q}$  are the only maps which satisfy the multivariate version of Theorem 10. This can be proven by a similar argument, the details of which are omitted.

**Interpretation 5** (Finiteness implies linearity). This result may serve as a sort of hint as to why the mathematical formalism of quantum mechanics is the theory of linear operators over Hilbert space. As all “real-world” systems have a finite number of states, then when the duration of the temporal interval measured in the smallest (say, Planck) time units becomes comparable to the number of states, the finiteness reveals itself as the linearity.

### 5. In the Middle of the Scales

In Section 1, we conjectured that if both “continuous” and “discrete” theories adequately describe physical reality at respective “ends of the scale”, the theories must “meet one another somewhere in the middle of the scale”. In this Section, we argue that the “meeting point in the middle of the scale” is the wave function. To do this, we first need to formalise the notion of an observer; actually, we will consider observers of two kinds, each for the respective ends of the scale.

#### 5.1. Observation and Measurement at the Ends of the Scale

To begin with, let us introduce two types of observers, the *Big-endian* and the *Little-endian*. The names of the two observers are more related to big-end and little-end orders the bytes of representation of a number are read in computer science and less with *Gulliver’s Travels* by Jonathan Swift. Given a large non-negative number having a very long base- $p$  expansion, the Big-endian is capable of observing only the highest order digits of the expansion, i.e., he knows the order of magnitude of the number and (up to a nonzero error) a mantissa since the Big-endian is not able to see the rightmost digits of the numbers. Conversely, the Little-endian sees the rightmost digits of the number, starting with the smallest order digit, but has no idea what are the leftmost digits and the order of magnitude of the number (although he assumes that the order is finite but very large). One may call the Big-endian a macro-observer and the Little-endian a micro-observer. However, both observers measure observable values which are rational  $p$ -adic integers. As already mentioned, real numbers are indistinguishable during measurements from rational  $p$ -adic

integers  $\mathbb{Z}_p \cap \mathbb{Q}$  due to the inevitable nonzero measurement error with respect to real metrics. This is why we assume that numerical values of observable are in  $\mathbb{Z}_p \cap \mathbb{Q}$ , and the Little-endian sees the first terms of the canonical  $p$ -adic expansion of the observable value, whereas Big-endian sees the highest order digits of the base- $p$  expansion of the same value as of a real number. We explain this more formally.

It is known (see, e.g., [10]) that  $z \in \mathbb{Z}_p \cap \mathbb{Q}$  if and only if  $z$  can be represented as  $z = c + \frac{d}{p^t - 1}$  for some  $t \in \mathbb{N}$  and  $d \in \{1, 2, \dots, p^t - 1\}$ ,  $c \in \mathbb{Z}$ , or, if and only if the  $p$ -adic canonical representation of  $z$  is eventually periodic

$$z = \alpha_0 + \alpha_1 p + \dots + \alpha_{r-1} p^{r-1} + (\beta_0 + \beta_1 p + \dots + \beta_{t-1} p^{t-1}) p^r + (\beta_0 + \beta_1 p + \dots + \beta_{t-1} p^{t-1}) p^{r+t} + (\beta_0 + \beta_1 p + \dots + \beta_{t-1} p^{t-1}) p^{r+2t} + \dots \quad (13)$$

for suitable  $\alpha_j, \beta_i \in \{0, 1, \dots, p - 1\}$ ,  $r \in \mathbb{N}_0$ ,  $t \in \mathbb{N}$  (the sum  $\alpha_0 + \alpha_1 p + \dots + \alpha_{r-1} p^{r-1}$  is absent in the above expression once  $r = 0$ ). In this case, the base- $p$  representation of the fractional part of  $z$  as of a real number is as follows:

$$z \bmod 1 = 0.(\hat{\beta}_{t-1-\bar{r}} \hat{\beta}_{t-2-\bar{r}} \dots \hat{\beta}_0 \hat{\beta}_{t-1} \hat{\beta}_{t-2} \dots \hat{\beta}_{t-\bar{r}})^\infty \bmod 1, \quad (14)$$

where  $\hat{\beta} = p - 1 - \beta$  for  $\beta \in \{0, 1, \dots, p - 1\}$ , and  $\bar{r}$  is the least non-negative residue of  $r$  the modulo  $t$  if  $t > 1$  or  $\bar{r}$  is zero otherwise.

To illustrate what Big-endian observations and Little-endian observations are, let  $r = 1$ ,  $t \gg 1$ ,  $\alpha_0 = 1$ ,  $\beta_{t-2} = \beta_{t-1} = 0$ ; thus, both “-endians” measure physical quantity  $z$  that takes values in  $[0, 1]$ . Then, as none of the observers is able to measure the value  $z$  with a nonzero error, the Big-endian will obtain only the digits  $\hat{\beta}_{t-2}, \hat{\beta}_{t-3}, \dots, \hat{\beta}_{t-n}$  for some  $n < t$ ; meanwhile, the Little-endian will obtain  $\beta_0, \beta_1, \dots, \beta_m$  for some  $m < t - 1$ . Thus, the only information about  $z$  which possibly is common for both observers is the values  $\beta_{t-\ell}, \beta_{t-\ell-1}, \dots, \beta_k$  for some  $k > 0$ ,  $\ell > 1$ . The two observers may communicate with each other and thus make only common guesses about what  $z$  is. Moreover, both do not know what  $t$  is; therefore, as  $t \gg 1$ , the only thing that both observers may know for sure is that  $1 \geq z \geq 1 - 1/p$ . Note that there are no “hidden variables” in this scenario since both observers may unboundedly increase the precision of their measurement despite neither being able to measure quantities with a nonzero error.

### 5.2. $p$ -adic Clocks

In this section, we introduce a  $p$ -adic model of the instrument which measure and indicates time, a  $p$ -adic clock; then, we prove that there exist only one clock, which is the same for all Little-endians and Big-endian, the *universal clock*.

A timekeeping element of the contemporary physical clock is a harmonic oscillator of a particular frequency, which is assumed to be a positive integer showing the number of periods per unit interval; therefore, the shortest time interval which can be measured is a reciprocal of the frequency. In order to measure the value of time elapsed, one merely counts the number of periods from one moment of time to another and represents this non-negative integer in some base, say,  $p$ , where  $p$  is the frequency of the oscillator. In what follows, we assume that  $p$  is a prime as to not overload the exposition with unimportant technical details. Thus, a model of such clock can be represented using the  $p$ -adic odometer, a dynamical system  $f = \tau_p: z \mapsto z + 1$  on the space of  $p$ -adic integers  $\mathbb{Z}_p$ . If the initial point is  $x_0 \in \mathbb{Z}_p$ , e.g.,  $x_0 = 0$ , put  $f^0(x_0) = x_0, f^1(x_0) = f(x_0) = x_1, \dots, f^i(x_0) = f(f^{i-1}(x_0)) = x_i$ , and then the base- $p$  expansion of  $x_i$  represents the time elapsed,  $i = \sum_{j=0}^{\lfloor \log_p i \rfloor + 1} \chi_j^i \cdot p^j$ , where  $\chi_j^i = \delta_j(i) \in \{0, 1, \dots, p - 1\}$  is the  $j$ -th digit of the base- $p$  expansion of  $i$ . In loose terms, the  $p$ -adic clock is simply a counter whose face consists of windows; at each time moment  $i$ , each  $j$ -th window shows  $\delta_j(i)$ . It is convenient to assume that the number of windows is infinite to have the time elapsed be unrestricted; thus, we obtain the dynamical system  $\tau_p$  on  $\mathbb{Z}_p$ . Note that the initial state  $x_0$  may be taken arbitrarily and not necessarily as  $x_0 = 0$ ;

then, to get the base- $p$  representation of time elapsed since the initial moment, one has to perform subtraction  $x_i - x_0$  in  $\mathbb{Z}_p$ . The  $p$ -adic clock is depicted in Figure 3. To the right, the content of the registry is similar to a standard representation of time in decimal (rather than  $p$ -ary) fractions of a second (millisecond, microsecond, nanosecond, ...) with Planck time at the rightmost position; meanwhile, to the left are decimal multiples of a second (petasecond, exasecond, ...).

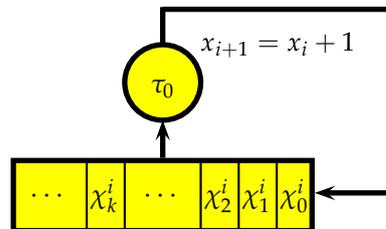


Figure 3. The  $p$ -adic clock.

Speaking loosely, the registry in Figure 3 is like a face of a mechanical counter consisting of cogwheels. The period of the sequence of states of the rightmost cell of the registry (which can be judged as the rightmost cogwheel) is  $p$ , the period of the sequence of states of the second rightmost cell is  $p^2$  since the figure in that cell changes once in a period of the rightmost cell, etc. The latter property is a definitive property of an ergodic transformation on  $\mathbb{Z}_p$ ; cf. Theorem 2. Therefore, all ergodic 1-Lipschitz transformations on  $\mathbb{Z}_p$  should be considered to be clocks, cf. (ii) of Theorem 3, as they can be "adjusted" one to another since they all are conjugate to the  $p$ -adic odometer.

If the initial state of the odometer is taken to be 0 (i.e., each cell of the registry depicted by Figure 3 is 0), then after  $n \in \mathbb{N}$  time units elapse, the registry will contain the base- $p$  expansion of the number  $n$  since  $\tau_p^n(0) = n$ . Let us now take any ergodic 1-Lipschitz map  $f: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$ , and any  $t \in \mathbb{Z}_p$  and any sequence  $(n_i)_{i=0}^\infty$  over  $\mathbb{N}_0$  which converges  $p$ -adically to  $t$  (such a sequence exists as  $\mathbb{N}_0$  is dense in  $\mathbb{Z}_p$ ). It turns out then that for any  $z \in \mathbb{Z}_p$ , the  $p$ -adic limit  $\lim_{i \rightarrow \infty}^p f^{n_i}(z)$  exists; denote this limit via  $f^t(z)$ , then  $(z; t) \mapsto f^t(z)$  is a 1-Lipschitz map  $\mathbb{Z}_p^2 \rightarrow \mathbb{Z}_p$  which is measure-preserving with respect to  $t$ ; see [23] [Propositions 4.87–4.88, 4.90]. Therefore,  $p$ -adic time  $t$  is well-defined. For instance (see [23] [Example 4.89]), given an ergodic affine map  $f(z) = az + b$  on  $\mathbb{Z}_p$ , the two-variate function  $f^t(z)$  is of the form  $f^t(z) = bt + z$  if  $a = 1$ , and

$$f^t(z) = b \cdot \frac{a^t - 1}{a - 1} + a^t z,$$

if  $a \neq 1$ . Note that if the affine map  $z \mapsto az + b$  is ergodic then  $b \not\equiv 0 \pmod{p}$  and  $a \equiv 1 \pmod{p}$  (see [23] [Theorem 4.36]); thus, both  $a^t$  and  $\frac{a^t - 1}{a - 1}$  are well-defined  $p$ -adic integers for every  $t \in \mathbb{Z}_p$ .

The problem which immediately arises is that  $p$ -adic time  $t$  is well-defined for every  $t \in \mathbb{Z}_p$ , but if  $q$  is a prime number distinct from  $p$ , the  $p$ -adic time  $t$  may be meaningless for a  $q$ -adic observer, the  $q$ -adic Little-endian, not to mention the Big-endian. Fortunately, however, there is a clock (and therefore time) which is common both for all Little-endians and Big-endian. This clock/time is unique up to the direction of the time arrow. It is clear that the clock, which is common for all  $p$ -adic Little-endians and Big-endian, must be a totally consistent function. The following theorem holds:

**Theorem 11.** *Totally consistent functions which are measure-preserving for all prime  $p$  are exactly the functions  $x \mapsto \pm x + c$ , where  $c \in \mathbb{Z}$ ; only the functions  $\tau_\pm(x) = x \pm 1$  are ergodic for all prime  $p$ .*

This means that the “universal clock” is a standard odometer which runs forward ( $\tau_+(x) = x + 1$ ) or backward ( $\tau_-(x) = x - 1$ ).

**Proof of Theorem 11.** According to Theorem 5, any totally consistent function  $g$  is a polynomial; therefore, to be measure-preserving on  $\mathbb{Z}_p$ ,  $g$  must be (1) bijective modulo  $p$  and (2) its derivative  $g'(x)$  must vanish modulo  $p$  nowhere for all prime  $p$ ; see, e.g., [23] [Theorem 4.45]. As  $g$  is 1-Lipschitz on  $\mathbb{Z}_p$  for all prime  $p$ , and  $g'(x)$  is a polynomial, the derivative exists and takes values from  $\mathbb{Z}_p$  for all prime  $p$ ; hence,  $g'(\mathbb{Z}) \subset \mathbb{Z}$ . Therefore, (as  $g'(x)$  is a polynomial), condition (2) implies that  $g'(\mathbb{Z}) \in \{1, -1\}$ , which means that  $g'$  is a constant,  $\pm 1$ . This means that  $g$  is the affine function, namely, either  $g(x) = -x + c$  or  $g(x) = x + c$  for some  $c \in \mathbb{Z}$  (since  $g(0)$  must be an integer as  $g(\mathbb{Z}) \subset \mathbb{Z}$  due to total consistency). This proves the claim concerning measure-preservation.

The ergodicity claim follows from the ergodicity criterion for affine maps  $z \mapsto az + b$  which implies that if the map is ergodic on  $\mathbb{Z}_p$  then  $a \equiv 1 \pmod{p}$  and  $b \not\equiv 0 \pmod{p}$ ; see [23] [Theorem 4.36]. As these conditions must hold for all prime  $p$ , we conclude that  $a = 1$  and  $b \in \{1, -1\}$ .  $\square$

**Interpretation 6** (Free choice of temporal ordering at the smallest of scales). *The only clock that is common for “both ends of the scale” is the standard odometer  $\tau(t) = t_0 + t$  which shows the time  $t - t_0 \in \mathbb{R}$  elapsed since the moment  $t_0 \in \mathbb{R}$ . All observers acquire the value of the time elapsed up to a nonzero error with respect to the corresponding metrics. Therefore, in a contrast to a real observer (the Big-endian) the  $p$ -adic observers (the Little-endians) generally cannot determine with the “time stamps” of events which one of the two events happened earlier and which one later since there is no order on the field of  $p$ -adic numbers which agrees with field operations.*

**Note 7.** *It is known that generally there is no ordering of events in quantum mechanics; see, e.g., [48].*

### 5.3. Digitalization

Initial automaton is a model of a (generally open) physical system prepared in some fixed state; the system is exposed by an experimenter to a time series of “elementary impacts” and thus produces the time series of “elementary reactions”. The impacts/reactions occurs at discrete instants of time since time is assumed to be discrete; for example, at Planck’s scale, the smallest time interval is Planck time  $5.391247(60) \times 10^{-44}$  s. Concrete values of that smallest time interval depend on the process which is modelled (e.g., in smart contracts of digital economy the smallest interval is usually assumed to be 24 h) and are not specified; the definitive feature of the model is that “time flow” consists of “indivisible time intervals”.

The experimenter prepares a number of identical systems in the same state and probes them by exposing them to different impacts, observing reactions and thus obtaining a number of experimental points ( $\langle \text{impact} \rangle; \langle \text{reaction} \rangle$ ), where  $\langle \star \star \star \rangle$  are measured values of components of the impact–reaction pair. The experimenter then treats any measured value as a real number up to a nonzero real error.

In order to not overload the exposition, in what follows we consider a one-dimensional case mostly when the values  $\langle \text{impact} \rangle$  and  $\langle \text{reaction} \rangle$  are numbers rather than vectors. Up to normalisation, we may assume that the measured numerical values are all in the unit real interval  $[0, 1]$ ; thus, the experimenter obtains a number of experimental points in the real unit square  $[0, 1] \times [0, 1] = \mathbb{I}^2 \subset \mathbb{R}^2$ . Namely, given an automaton  $\mathfrak{A}$ , let  $f = f_{\mathfrak{A}}: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  be its automaton function (i.e., a 1-Lipschitz map). Consider a subset  $\mathcal{E}(f)$  of all the following points of the Euclidean unit square  $\mathbb{I}^2 = [0, 1] \times [0, 1] \subset \mathbb{R}^2$ :

$$e_k^f(z) = \left( \frac{z \bmod p^k}{p^k}; \frac{f(z) \bmod p^k}{p^k} \right) \in \mathbb{I}^2,$$

$z \in \mathbb{Z}_p, k = 1, 2, \dots$ . Here,  $z \bmod p^k = \sum_{i=0}^{k-1} \chi_i p^i$  if  $z \in \mathbb{Z}_p$  is represented by its canonical form  $z = \sum_{i=0}^{\infty} \chi_i p^i, (\chi_i \in \mathbb{F}_p; i = 0, 1, 2, \dots)$ . Note that  $f(x) \bmod p^k$  corresponds to a  $k$ -letter output word  $\zeta_{k-1} \dots \zeta_1 \zeta_0$  of the automaton which is fed by the  $k$ -letter input word  $\chi_{k-1} \dots \chi_1 \chi_0$  which corresponds to  $x \bmod p^k$ ; cf. Figure 4.

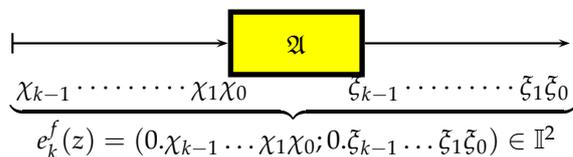


Figure 4. A point in the unit square  $\mathbb{I}^2 \subset \mathbb{R}^2$  produced by the automaton  $\mathfrak{A}$ .

Further, although all the word lengths  $k$  are finite, the clustering is equivalent to sending  $k \rightarrow \infty$ . Therefore, the clustering is equivalent to taking limit points of the closure  $\mathcal{P}(f)$  of the set  $\mathcal{E}(f)$  with respect to the standard topology of  $\mathbb{R}^2$ . We call  $\mathcal{P}(f)$  a plot of  $f$ . Speaking very loosely, the plot is a picture the experimenter obtains as an output of the experiment which consists of a number of individual probes of a physical system which is prepared in the same state before each probe. Note that the set of cluster points of the pictures for both experimenters, the Little-endian and the Big-endian, obtained as result of the experiment look very similar for the both since Little-endian makes the word lengths as long as possible to construct the cluster points while Big-endian is only capable of obtaining the points which correspond to sufficiently long words, i.e., the points which are close to the cluster points. This fact is crucial for the future construction of wave function by the both experimenters as well as for the uncertainty relation on which the both agree.

Let us describe this procedure more formally. For  $s = \sum_{j=-k}^{\infty} \zeta_j p^j \in \mathbb{Q}_p, (\zeta_j \in \{0, 1, \dots, p-1\}, j \in \mathbb{Z})$ , let  $[s]_p = \zeta_0 + \zeta_1 p + \zeta_2 p^2 + \dots \in \mathbb{Z}_p$  and  $\{s\}_p = \zeta_{-k} p^{-k} + \dots + \zeta_{-1} p^{-1}$  be the integral and fractional parts of  $s$ , respectively. Recall that any complex character of additive group  $\mathbb{Q}_p^+$  of the field  $\mathbb{Q}_p$  of  $p$ -adic numbers is of the form  $\chi_r(s) = e^{2\pi i \{sr\}_p}$ , where  $r \in \mathbb{Q}_p$ ;  $\chi_r$  is a continuous group epimorphism into the group of complex roots of unity (which is isomorphic to the group  $\mathbb{Q}^+ / \mathbb{Z}^+$ ). Take  $r = 1$ , denote  $\chi_1$  via  $\chi$ ; given a 1-Lipschitz map  $f: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$ , consider the mappings

$$\check{f}_k: e^{2\pi i \{p^{-k}z\}_p} \mapsto e^{2\pi i \{p^{-k}f(z)\}_p}, \quad (z \in \mathbb{Z}_p),$$

for all  $k \in \mathbb{N}_0$ . As every  $\check{f}_k$  maps points of the unit circle  $\mathbb{S}$  into points of  $\mathbb{S}$ , the pairs  $(e^{2\pi i \{p^{-k}z\}_p}; e^{2\pi i \{p^{-k}f(z)\}_p})$  constitute a set of points on the unit torus  $\mathbb{T}^2 = \mathbb{S} \times \mathbb{S}$ . The unit square  $\mathbb{I}^2$  is a universal cover of the torus  $\mathbb{T}^2$ ; this way, the points  $e_k^f(z) \in \mathbb{I}^2$  are identified with the points  $(e^{2\pi i \{p^{-k}z\}_p}; e^{2\pi i \{p^{-k}f(z)\}_p}) \in \mathbb{T}^2$ , and in what follows, we do not differ between the point sets and speak either of the points on the surface of the torus  $\mathbb{T}^2$  or on the square  $\mathbb{I}^2$ , whichever is more convenient.

**Definition 10** (Plots of automata). Given an automaton  $\mathfrak{A}$ , let  $f = f_{\mathfrak{A}}: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  be the automaton function. The closure  $\mathcal{P}(f) = \mathcal{P}(\mathfrak{A})$  of all the points  $e_k^f = (\frac{z \bmod p^k}{p^k}; \frac{f(z) \bmod p^k}{p^k})$  in the square  $\mathbb{I}^2$  (or of all the points  $(e^{2\pi i \{p^{-k}z\}_p}; e^{2\pi i \{p^{-k}f(z)\}_p})$  in the torus  $\mathbb{T}^2$ ), where  $k \in \mathbb{N}, z \in \mathbb{Z}_p$  is called a (one-dimensional) plot of the automaton  $\mathfrak{A}$  or, similarly, of the automaton function  $f = f_{\mathfrak{A}}$ . The set  $\mathcal{P}'(f) = \mathcal{P}'(\mathfrak{A})$  of all the limit points of the plot, the derived set of the set  $\mathcal{P}(f) = \mathcal{P}(\mathfrak{A})$ , is called the limit plot of the automaton  $\mathfrak{A}$  (of the automaton function  $f_{\mathfrak{A}}$ ).

Recall that the limit point, accumulation point, or cluster point is a synonymic notion of the point such that every neighbourhood of which contains points other than that point. Recall also that the derived set of a closed set is also closed; thus,  $\mathcal{P}'(f) = \mathcal{P}'(\mathfrak{A})$  is closed. Being closed, the set  $\mathcal{P}(\mathfrak{A})$  is measurable with respect to the Lebesgue measure on  $\mathbb{R}^2$ ; denote as

$\alpha(\mathfrak{A}) = \alpha(f)$  the measure of  $\mathcal{P}(\mathfrak{A})$ . Respective notions for the general  $n$ -dimensional case,  $n > 1$ , are defined as follows: for  $z \in \mathbb{Z}_p, k, n \in \mathbb{N}, n > 1$  denote

$$e_{k,n}^f(z) = \left( \frac{z \bmod p^k}{p^k}, \frac{f(z) \bmod p^k}{p^k}, \dots, \frac{f^{n-1}(z) \bmod p^k}{p^k} \right) \in \mathbb{I}^n \subset \mathbb{R}^n.$$

The respective notation in this case is  $\mathcal{P}_n(f) = \mathcal{P}_n(\mathfrak{A}), \mathcal{P}'_n(f) = \mathcal{P}'_n(\mathfrak{A}), \alpha_n(\mathfrak{A})$ , etc. We usually omit the index  $n$  when  $n = 2$ .

**Theorem 12** (The automata 0-1 law, [49]). *Given the arbitrary automaton  $\mathfrak{A}$ , the following alternative holds: either  $\alpha(\mathfrak{A}) = 0$  (equivalently,  $\mathcal{P}(\mathfrak{A})$  is nowhere dense in  $\mathbb{I}^2$ ), or  $\alpha(\mathfrak{A}) = 1$  (equivalently,  $\mathcal{P}(\mathfrak{A}) = \mathbb{I}^2$ ).*

**Note 8.** Recall that nowhere dense sets can nevertheless have positive Lebesgue measures, for instance, the “fat” Cantor sets (e.g., the Smith-Volterra-Cantor set), which are also known as  $\epsilon$ -Cantor sets; see e.g., [50]; however, this is not the case for the set  $\mathcal{P}(\mathfrak{A})$ . The Lebesgue measure of this set is 0 if and only if it is nowhere dense.

Theorem 12 is true in the multidimensional case as well. We will say briefly that a 1-Lipschitz map  $f: \mathbb{Z}_p^n \rightarrow \mathbb{Z}_p^n$  (or respective automaton whose automaton function is  $f$ ) is *measure-0* in dimension  $n$  if  $\alpha_n(f) = 0$ , and *measure-1* otherwise. It turns out that all polynomials over  $\mathbb{Z}$  whose degree is greater than 1 are *measure-1* in all dimensions. Actually, for  $f \in \mathbb{Z}[x]$ , a much stronger result is true: if  $\deg f \geq 2$ , then the distribution of points  $e_{k,n}^f(z)$  in the unit hypercube  $\mathbb{I}^n$  tends to uniform as  $k \rightarrow \infty$ , for every  $n \in \{2, 3, 4, \dots\}$ . Specifically, the following theorem holds:

**Theorem 13** ([11]). *Let  $f$  be a polynomial over  $\mathbb{Z}$ ,  $\deg f \geq 2$ . Then, the sequence  $(e_{k,n}^f(z \bmod p^k))_{k=1}^\infty$  of random vectors weakly converges as  $k \rightarrow \infty$  to a random vector having a continuous uniform distribution in  $[0, 1)^n$ .*

Theorem 13 may be interpreted as showing another way by which chaos emerges.

**Interpretation 7** (Emergence of chaos: The two ways).

- **1-st:** Chaos emerges from infinite “chaotic sequences” such as random real numbers by iterating them via Bernoulli-shift-like mappings, logistic mappings, etc; that is, when it is assumed a priori that “chaos does exist immanently”.
- **2-nd:** Chaos emerges from the “lack of knowledge what elementary causes happened at the very beginning”; that is, if a Big-endian observer is incapable of determining what the digits  $\xi_0, \xi_1, \dots, \xi_{k-1}, \dots$  are of the input  $z = \sum_{j=0}^\infty \xi_j p^j \in \mathbb{Z}_p$  of the causal function  $f$  if  $k$  is small enough.

Note that in the second case, the Little-endian observer is capable of determining the digits  $\xi_k$  if  $k$  is “not too large”, so these digits are *not* hidden parameters. Nonetheless, further in the paper, we show that a specific uncertainty relation holds both for the Little-endian and Big-endian observers.

Note also that polynomials over  $\mathbb{Z}$  whose degrees are greater than 1 are automaton functions of infinite automata; cf., Example 2. However, any automaton function  $f: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  of an infinite automaton can be uniformly approximated on  $\mathbb{Z}_p$  by automaton functions of finite automata, for instance, by the functions  $f_n: z \mapsto f(z) \bmod p^n$ . This fact, together with the finiteness assumption of Section 2, emphasises a distinguished role the finite automata play in further considerations; thus, we now pay special attention to finite automata.

**Theorem 14** (see [23] [Section 11.1.2]). *Finite automata are measure-0 in all dimensions.*

**Example 5.** Automata may be infinite and measure-0; constants may be measure-1:

- The automaton whose automaton function is  $f(z) = z + (z^2 \text{OR}(-\frac{1}{3}))$ , ( $z \in \mathbb{Z}_2$ ), is infinite and measure-0. Here, OR is bit-by-bit logical  $\vee$  with no carries to higher order bits; that is, if  $z = \sum_{j=0}^{\infty} \zeta_j 2^j$ , then  $z \text{OR}(-\frac{1}{3}) = \sum_{j=0}^{\infty} \zeta_j 2^{2j}$  as  $-\frac{1}{3} = \sum_{j=0}^{\infty} 2^{2j}$  is a canonical 2-adic representation of  $-\frac{1}{3} \in \mathbb{Z}_2 \cap \mathbb{Q}$ .
- The automaton whose automaton function is  $f(z) = C$  where  $C$  is a  $p$ -adic integer whose canonical representation corresponds to a Champernowne word is a measure-1 automaton. Recall that a Champernowne word is a word obtained via concatenation of the base- $p$  expansions of numbers 1, 2, 3, 4, 5, 6, ...; for instance, the 2-adic Champernowne word is 101110011010111...

In short, Theorems 12 and 14 imply that plots of finite automata cannot contain “figures” but may contain “lines”. These lines are of the utmost importance in further considerations since they may naturally be treated as “experimental curves” obtained by probing a physical system both by Little-endian and Big-endian observers. It turns out that smooth lines from limit plots of finite automata are windings of torus; therefore, the lines may be treated as sine waves, so the smooth lines in the limit plot of a finite automaton constitute a collection of sine waves. Moreover, the waves are limit plots of finite affine automata. Now, we express these facts rigorously.

Recall that a knot is a smooth embedding of a circle  $\mathbb{S}$  into  $\mathbb{R}^3$  and a link is a smooth embedding of several disjoint circles in  $\mathbb{R}^3$ ; cf. [51]. We will consider only special types of knots and links, namely, torus knots and torus links. Informally, a torus knot is a smooth closed curve without intersections which lies completely in the surface of a torus  $\mathbb{T}^2 \subset \mathbb{R}^3$ , and a link (of torus knots) is a collection of (possibly knotted) torus knots; see, e.g., [52] [Section 26] for formal definitions.

We also need a notion of a winding of a torus. Formally, a winding of a torus is any geodesic on a torus. Recall that geodesics on torus  $\mathbb{T}^2$  are images of straight lines in  $\mathbb{R}^2$  under the mapping  $(x; y) \mapsto (x \bmod 1; y \bmod 1)$  of  $\mathbb{R}^2$  onto  $\mathbb{T}^2 = \mathbb{R}^2 / \mathbb{Z} \times \mathbb{Z}$ ; cf., e.g., [53] [Section 5.4].

**Definition 11** (Winding of the torus). A winding of the torus is an image of a straight line in  $\mathbb{R}^2$  under the map  $\text{mod}1: (x; y) \mapsto (x \bmod 1; y \bmod 1)$  of the Euclidean plane  $\mathbb{R}^2$  onto the 2-dimensional real torus  $\mathbb{T}^2 = \mathbb{R}^2 / \mathbb{Z} \times \mathbb{Z} = \mathbb{S} \times \mathbb{S} \subset \mathbb{R}^3$ . If the line is defined by the equation  $y = ax + b$ , we say that  $a$  is a slope of the winding  $\mathbf{C}(a, b)$ . We denote via  $\mathbf{C}(\infty, b)$  a winding which corresponds to the line  $x = b$ , the meridian, and say that the slope is  $\infty$  in this case. Windings  $\mathbf{C}(0, b)$  of slope 0 (i.e., the ones that correspond to straight lines  $y = b$ ) are called parallels.

In dynamics, windings of torus  $\mathbb{T}^2$  are viewed as orbits of linear flows on the torus; that is, of dynamical systems on  $\mathbb{T}^2$  defined by a pair of differential equations of the form  $\frac{dx}{dt} = \beta; \frac{dy}{dt} = \alpha$  on  $\mathbb{T}^2$  and thus by a pair of parametric equations  $x = (\beta t + \tau) \bmod 1; y = (\alpha t + \sigma) \bmod 1$  in Cartesian coordinates; cf., e.g., [54] [Section 4.2.3].

**Note 9.** It is well known that a winding defined by the straight line  $y = ax + b$  is dense in  $\mathbb{T}^2$  if and only if  $-\infty < a < +\infty$  and the slope  $a = \frac{\alpha}{\beta}$  is irrational; see, e.g., [54] [Proposition 4.2.8] or [53] [Section 5.4].

Theorem 15 which follows states that  $C^2$ -smooth lines (i.e., those which are twice differentiable and have continuous second derivatives) in  $\mathcal{P}(f_{\mathfrak{A}})$  are windings of the torus  $\mathbb{T}^2$  provided the automaton  $\mathfrak{A}$  is finite; cf., Figures 5 and 6.

**Theorem 15** ([10]). Let  $f: \mathbb{Z}_p \rightarrow \mathbb{Z}_p$  be an automaton function of a finite automaton; let  $g$  be a  $C^2$ -function with domain  $[a, b] \subset [0, 1) \subset \mathbb{R}$  and range  $[0, 1) \subset \mathbb{R}$ . Let the graph  $\mathbf{G}(g) = \{(x; g(x)): x \in [a, b]\}$  of the function  $g$  lie completely in  $\mathcal{P}(f)$ . Then, there exist  $a, b \in \mathbb{Q} \cap \mathbb{Z}_p$  such that  $g(x) = (ax + b) \bmod 1$  for all  $x \in [a, b]$ ; moreover, there is a winding of the torus

$\mathbb{T}^2$  which lies completely in  $\mathcal{P}(f)$  and which contains the graph  $\mathbf{G}(g)$  of the function  $g$ . There are not more than a finite number of pairwise distinct windings of the unit torus  $\mathbb{T}^2$  in  $\mathcal{P}^2(f)$ ; all of these are images of real affine functions  $x \mapsto ax + b$  for  $a, b \in \mathbb{Z}_p \cap \mathbb{Q}$  under the mapping  $\text{mod } 1: \mathbb{R}^2 \rightarrow \mathbb{T}^2$ .

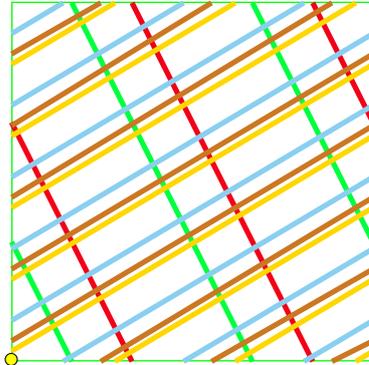


Figure 5. Limit plot in  $\mathbb{R}^2$  of an automaton having two affine subautomata.

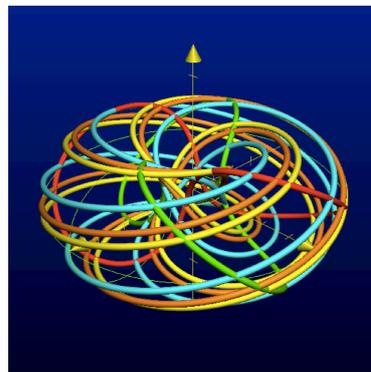


Figure 6. Limit plot of the same automaton on the torus  $\mathbb{T}^2$  in  $\mathbb{R}^3$ .

**Note 10.** The  $C^2$ -smoothness condition can be relaxed:  $C^1$ -smoothness is sufficient to ensure the affinity; see [55].

Although Theorem 15, after proper restatement, holds for  $m$ -variate 1-Lipschitz maps  $f: \mathbb{Z}_p^m \rightarrow \mathbb{Z}_p^m$  as well, see [10], we restrict considerations in the rest part of the paper mostly by a univariate case for simplicity.

The torus link which is a limit plot of a finite automaton affine function  $f: z \mapsto az + b$  on  $\mathbb{Z}_p$  is completely described by the following theorem:

**Theorem 16 ([10]).** Given a finite automaton affine function  $f: z \mapsto az + b$  on  $\mathbb{Z}_p$ , (i.e., such that  $a, b \in \mathbb{Z}_p \cap \mathbb{Q}$ ), represent  $a, b$  as irreducible fractions:  $a = \frac{\alpha}{\beta}; b = \frac{\alpha'}{\beta'}$ , where  $\alpha, \beta, \alpha', \beta' \in \mathbb{Z}, \beta, \beta' \not\equiv 0 \pmod{p}$ . Then, the limit plot  $\mathcal{P}^l(f)$  on the torus  $\mathbb{T}^2$  is a torus link which consists of  $N$  torus windings whose slope is  $a$ , where  $N = \text{mult}_p \frac{\beta'}{d}$  is a multiplicative order of  $p$  modulo  $\frac{\beta'}{d}, d = \text{gcd}(\beta, \beta')$  is the greatest common divisor of  $\beta, \beta'$ , and  $N = 1$  if  $\frac{\beta'}{d} = 1$ . Every torus winding is a graph of the complex-valued function  $\psi(\rho, k): \mathbb{R} \rightarrow \mathbb{C}$  on the torus  $\mathbb{T}^2$  for a suitable  $k = 0, 1, \dots, \text{mult}_p \frac{\beta'}{d} - 1$ , where  $\psi(\rho, k) = e^{i(\frac{\alpha}{\beta}\rho - 2\pi p^k \frac{\alpha'}{\beta'})}, (\rho \in \mathbb{R})$ .

In cylindrical coordinates, every torus winding  $x \mapsto ax + b$  of a torus that is obtained by revolving around  $Z$ -axis of a circle that is coplanar with the axis and has radius  $r$  and

a centre at the distance  $R$  from the origin can be represented by the following parametric equations

$$\begin{bmatrix} r_0 \\ \theta \\ z \end{bmatrix} = \begin{bmatrix} R + r \cos(ax + b) \\ x \\ r \sin(ax + b) \end{bmatrix}, \quad x \in \mathbb{R}. \tag{15}$$

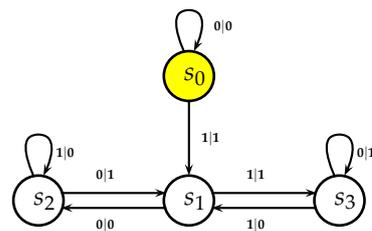
If  $a \in \mathbb{Z}_p \cap \mathbb{Q}$ , then  $a$  is irreducible fraction  $\alpha/\beta$  where  $\alpha, \beta \in \mathbb{Z}$  and  $p \nmid \beta$ ; then, corresponding winding winds  $\beta$  times around the  $Z$ -axis and  $|\alpha|$  times around a circle in the interior of the torus, whereas the sign of  $\alpha$  determines whether the rotation is clockwise or counter-clockwise. Hence, “physical meaning” that can be ascribed to the coefficient  $a = \frac{\alpha}{\beta}$  of the affine map  $z \mapsto az + b$ , ( $z \in \mathbb{Z}_p$ ), which is a finite automaton function of affine automaton if and only if  $a, b \in \mathbb{Z}_p \cap \mathbb{Q}$ , is *frequency* (or, as a *wavenumber*, under a proper choice of units). The choice of sign  $+$  or  $-$  depends only on what direction of rotation is assumed to be “positive” or “negative”; thus, *polarization* and *spin* can be ascribed to the sign of  $a$  in relevant models.

Theorem 16 in view of representation (15) implies that the limit plot of a finite automaton whose function is  $z \mapsto az + b$ , (where  $a, b \in \mathbb{Z}_p \cap \mathbb{Q}$ ,  $z$  runs over  $\mathbb{Z}_p$ ) is in one-to-one correspondence to a complex-valued function  $\psi: \mathbb{R} \times \mathbb{N}_0 \rightarrow \mathbb{C}$ :

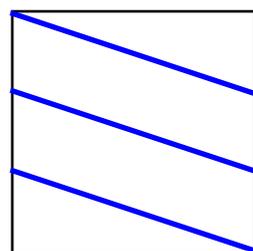
$$\psi(x, k) = e^{i(ax - 2\pi p^k b)}, \quad \text{where } x \in \mathbb{R}, k \in \mathbb{N}_0 \tag{16}$$

It is worth noting that the function  $\psi(x, k)$  is well-defined for all  $k \in \mathbb{Z}$  since  $p$  is the invertible modulo  $\beta^l/d$  and thus  $e^{-2\pi i p^k b}$  is well defined for every  $k \in \mathbb{Z}$ ; cf., Theorem 16.

**Note 11.** According to Theorem 16, different affine functions  $z \mapsto az + b$  may have identical limit plots. For instance, all the functions  $f(z) = z + c$  where  $c \in \mathbb{Z}_p \cap \mathbb{Q}$  have identical limit plots which correspond to the function  $\psi(x) = e^{ix}$ . Note also that whenever a limit plot of a finite automaton  $\mathfrak{A}$  is the same as that of the finite automaton whose automaton function  $f$  is affine,  $f(z) = az + b$ , there exist a minimal subautomaton of  $\mathfrak{A}$  (i.e., the one having no subautomata other than itself) which has exactly the same limit plot; see Figures 7 and 8. A finite automaton is minimal if and only if its reduced state transition diagram is totally connected: Given two states  $s, t \in S$ , there is finite word  $w$  such that when the automaton in state  $s$  accepts the word  $w$ , the automaton changes its state to  $t$ . If an automaton reaches a state which belongs to its (minimal) subautomaton, the automaton will never reach a state which does not belong to the subautomaton.



**Figure 7.** The automaton function is  $z \mapsto -\frac{1}{3}z$ ; the minimal subautomaton function is  $z \mapsto -\frac{1}{3}z - \frac{2}{3}$ ; ( $z \in \mathbb{Z}_2$ ),  $s_0$  and  $s_1$  are respective initial states.



**Figure 8.** Limit plots of the automaton and of its minimal subautomaton coincide

**Example 6 (Limit plots of the automata).** Figures 9 and 10 show the limit plot of a constant function which is an automaton function of finite autonomous automaton; autonomous automata may be judged as models of either isolated or closed physical systems. Parallel lines shown by Figure 9 may be ascribed to energy levels.

The remaining examples are nonautonomous automata; these can serve as models of open physical systems. Figures 9 and 10 depict limit plots produced of an autonomous automaton whose state transition diagram depicts Figure 11. Figures 12 and 13 show the limit plot of an automaton having two minimal subautomata; the state transition diagram of the automaton is shown in Figure 14.

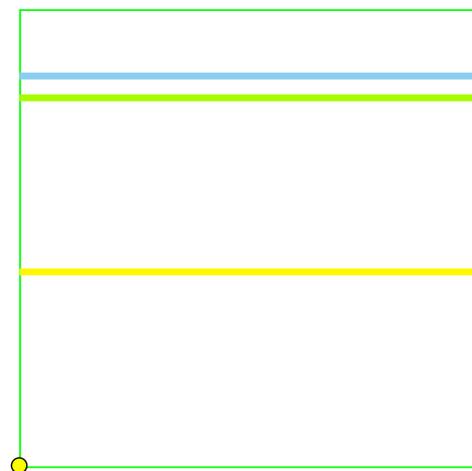
Figure 15 represents a plot of a finite automaton which approximates a measure-1 (and thus infinite) automaton whose automaton function is  $z \mapsto 1 + 3z + 2z^2$ , ( $z \in \mathbb{Z}_2$ ). Note the pronounced straight lines in the plot; these lines constitute the limit plot of a minimal subautomaton.

Figure 16 depicts a plot of a measure-0 (but infinite) automaton which has the only minimal finite affine subautomaton; the automaton function of the latter subautomaton is  $z \mapsto 5z$ , ( $z \in \mathbb{Z}_2$ ). The limit plot of the latter automaton are red lines; cf., Figure 12; the state transition diagram is the lower part of the diagram shown in Figure 14.

Basically, the limit plot of a finite automaton whose minimal subautomata are affine consists of families of parallel straight lines in the unit square or, respectively, of links of the torus windings whose slopes are in  $\mathbb{Z}_p \cap \mathbb{Q}$ ; cf., Figures 5, 6, 12, and 13. The the minimal subautomata from the first example “exhibit nonzero phase shifts”, while for the ones from the second example, the “phase shifts” are 0. Both examples are automata having two minimal affine subautomata. The minimal subautomata from the first example (Figures 5 and 6) have limit plots defined by the functions  $f_1(z) = -2z + \frac{1}{3}$  (red and green windings) and  $f_2(z) = \frac{3}{5}z + \frac{2}{7}$ , (yellow, brown, and blue windings), respectively,  $z \in \mathbb{Z}_2$ . The minimal subautomata from the second example (Figures 12 and 13) have limit plots defined by the respective functions  $z \mapsto 3z$  (blue lines) and  $z \mapsto 5z$  (red lines),  $z \in \mathbb{Z}_2$ .

The limit plot of a finite affine automaton whose automaton function is  $z \mapsto az + b$  in the unit square  $\mathbb{I}^2$  consists of parallel straight lines with slope  $a = \alpha/\beta \in \mathbb{Z}_p \cap \mathbb{Q}$ ; thus, the plot may be considered not only on the torus obtained by “gluing together” opposite sides of the square but also on a cylinder obtained by “gluing together” only a pair of opposite sides of the square. This way, one obtains solenoid rather than a torus link. This representation of a limit plot is also convenient in some cases. For instance, Figures 17 and 18 depict the limit plot of the automaton whose automaton function is  $f(z) = ((z\text{AND}1) - ((\text{NOT}(z))\text{AND}1)) \cdot z$ , where AND and NOT are respectively bitwise logical “and” and bitwise logical “not” operations on base-2 expansions of numbers (with no carries), while “.” and “−” are usual multiplication and subtraction of numbers (with carries).

Figure 19 represents the state transition diagram of a general automaton all whose minimal automata are finite and affine.



**Figure 9.** Limit plot of the function  $f(z) = 2/7$  ( $z \in \mathbb{Z}_2$ ), in  $\mathbb{I}^2$ .

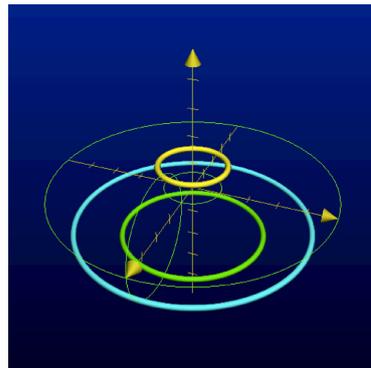


Figure 10. Limit plot of the same function on the torus  $\mathbb{T}^2$ .

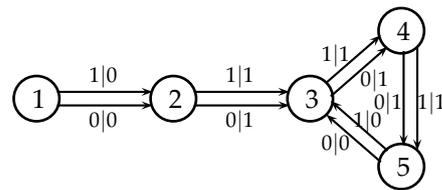


Figure 11. State transition diagram of the autonomous automaton whose automaton function  $f: \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$  is a constant:  $f(z) = 2/7, (z \in \mathbb{Z}_2)$ . State 1 is initial.

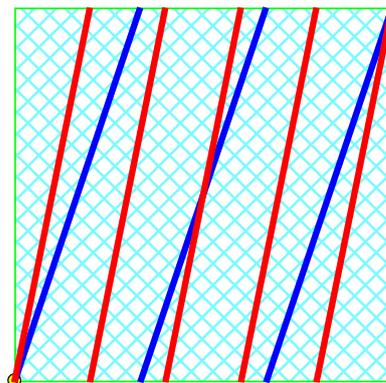


Figure 12. Limit plot of the automaton having two subautomata whose functions are  $z \mapsto 3z$  and  $z \mapsto 5z, (z \in \mathbb{Z}_2)$ .

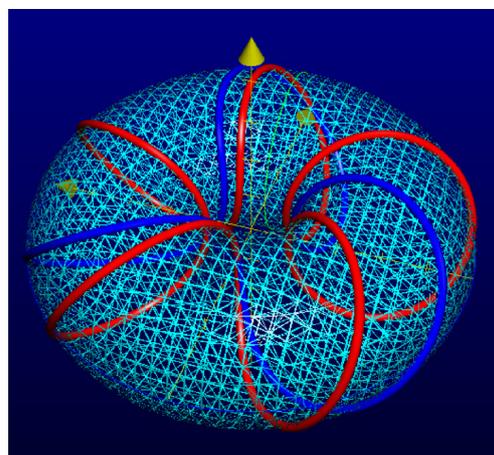
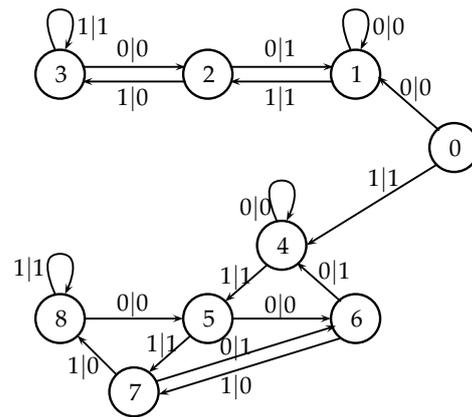
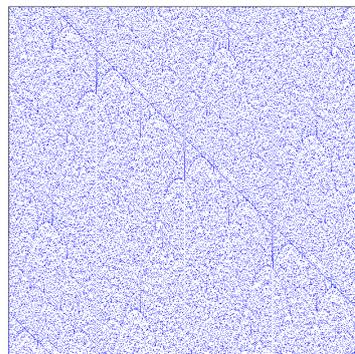


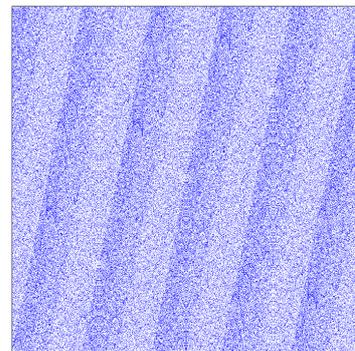
Figure 13. Limit plot of the same automaton on the torus  $\mathbb{T}^2 \subset \mathbb{R}^3$ . The surface of the torus is made visible by cross-hatching.



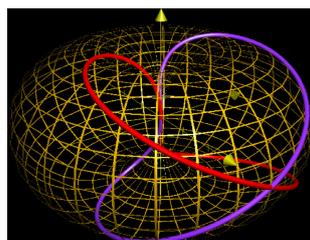
**Figure 14.** State transition diagram of the automaton having two minimal subautomata whose automata functions are  $z \mapsto 3z$  and  $z \mapsto 5z, z \in \mathbb{Z}_2$ . The initial state is 0.



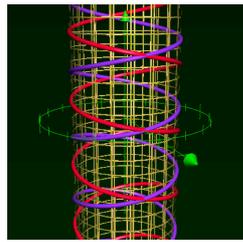
**Figure 15.** Plot of a finite automaton which is an approximation of a measure-1 automaton whose automaton function is  $z \mapsto 1 + 3z + 2z^2, (z \in \mathbb{Z}_2)$ .



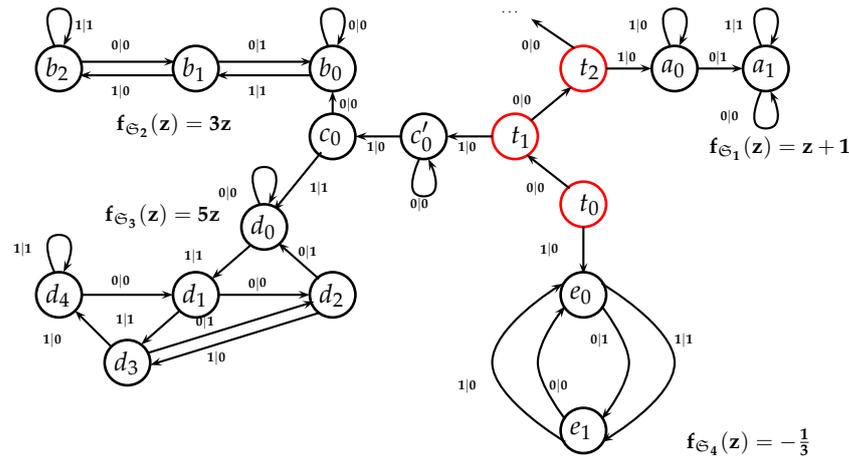
**Figure 16.** Plot of a measure-0 automaton having the only minimal subautomaton whose automaton function is  $z \rightarrow 5z, (z \in \mathbb{Z}_2)$ .



**Figure 17.** Limit plot of a finite automaton whose automaton function is  $z \mapsto ((z \text{ AND } 1) - ((\text{NOT}(z)) \text{ AND } 1)) \cdot z, (z \in \mathbb{Z}_2)$  on the (horn) torus.



**Figure 18.** Solenoid that is a limit plot of the automaton having the same automaton function  $f(z) = ((z \text{AND} 1) - ((\text{NOT}(z)) \text{AND} 1)) \cdot z, (z \in \mathbb{Z}_2)$ .



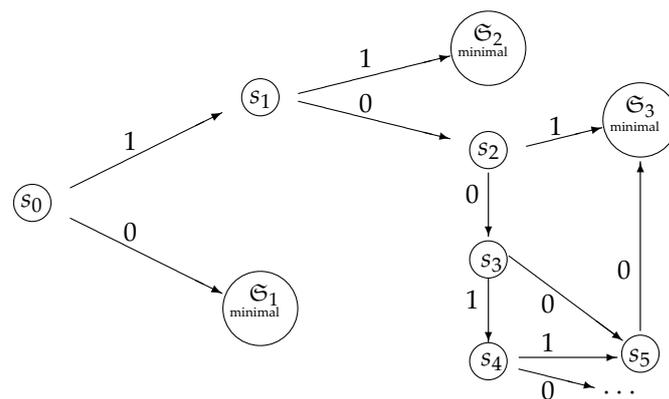
**Figure 19.** General automaton whose minimal subautomata are all finite and affine

#### 5.4. Wave Functions Emerging from Automata

This section discusses the main notion of quantum theory, the wave function. Our goal is to derive wave functions from causal functions; that is, from automata. Functions (16) are building blocks of the construction of the wave function on the base of causal maps. To begin, we briefly outline the general idea of the construction.

Recall that the reduced state transition diagram of a finite automaton is a digraph in which each path ultimately reaches a minimal subautomaton. There are no outgoing paths from subautomata. By feeding the automaton with random long words, to each minimal subautomaton we assign a probability for when the automaton reaches states which belong to the subautomaton; cf., Figure 20. Let automaton  $\mathfrak{A}$  be such that, being fed by random long words, the automaton at some finite step reaches, with a probability 1, a state which belongs to a minimal automaton which is finite and affine. The limit plot of every such subautomaton is described by a complex-valued function of the form (16).

To every minimal subautomaton that is finite and affine we ascribe its limit plot. There are only countably many such limit plots since there are only countably many such affine functions  $\mathbb{Z}_p \rightarrow \mathbb{Z}_p$  that are automata functions of these subautomata: Due to the finiteness of the subautomata, coefficients of these affine functions must belong to the set  $\mathbb{Z}_p \cap \mathbb{Q}$  which is countable. As every two minimal subautomata have no common states due to the minimality and as to every minimal subautomaton it is assigned a probability of reaching the subautomaton, to every limit plot one assigns a probability to “observe” that limit plot in the experiment, i.e., to obtain accumulation points in the unit square which constitute that limit plot. The probability is equal to a sum of all probabilities to reach the minimal subautomata having that plot. Therefore, these probabilities constitute a distribution assigned to the automaton; a *characteristic function* of that distribution is a (generally infinite) series whose terms are functions  $\psi(x, k) = e^{i(ax - 2\pi p^k b)}$  multiplied by values of respective probabilities; cf., (16) (there is a vast literature on characteristic functions of probability distributions; see, e.g., [56]). We argue that this characteristic function of the distribution may be treated as a wave function.



**Figure 20.** Example state transition diagram of 2-adic automaton having minimal subautomata (output symbols of labels of arrows are omitted).  $s_0$  is the initial state. The respective probabilities of reaching subautomata  $\mathfrak{S}_1$ ,  $\mathfrak{S}_2$ , and  $\mathfrak{S}_3$  are  $1/2$ ,  $1/4$ , and  $11/64 = 1/8 + 1/32 + 1/64$ .

Proceeding to a formal rigorous construction, let us review a few preliminary conventions:

- We do not distinguish affine automata whose limit plots coincide, so the actual probability distribution related to the automaton is distribution of classes of finite affine subautomata having coinciding limit plots;
- We use terms “ $p$ -adic integer”, “infinite word over  $p$ -symbol alphabet”, and “infinite path in a state transition diagram” as synonyms; see Sections 3.1–3.3.

A word of caution: there is a one-to-one correspondence between all paths of length  $k$  in the state transition diagram and all numbers from  $\{0, 1, \dots, p^k - 1\}$ ; however, to every number from  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ , there corresponds an infinite number of paths: Every such path has a prefix which is simply a base- $p$  expansion of a number and a suffix which consists of zeros only; cf., Section 3.1.

Given an automaton  $\mathfrak{A}$ , let  $\mathfrak{S}$  be its subautomaton. Let  $W(\mathfrak{S})$  be the set of all infinite paths starting from the initial state of  $\mathfrak{A}$  in a state transition diagram of  $\mathfrak{A}$  which reach states of  $\mathfrak{S}$  at finite steps. Note that if a path  $w$  reaches  $\mathfrak{S}$  at  $k$ -th step, then all paths which correspond to infinite words having the same prefix of length  $k$  reach  $\mathfrak{S}$  at the  $k$ -th step; therefore, the  $p$ -adic integers which correspond to these paths constitute a  $p$ -adic ball of radius  $p^{-k}$ . Therefore, all  $p$ -adic integers that correspond to infinite paths which reach the subautomaton  $\mathfrak{S}$  at finite steps constitute a disjoint union  $\mathbf{B}(\mathfrak{S})$  of balls of nonzero radii; hence,  $\mathbf{B}(\mathfrak{S})$  is a  $\mu$ -measurable subset of  $\mathbb{Z}_p$  with respect to the Haar measure on  $\mathbb{Z}_p$  which is normalised so that  $\mu(\mathbb{Z}_p) = 1$ . This way to  $\mathfrak{S}$  is assigned a probability  $\mu(\mathfrak{S}) = \mu(\mathbf{B}(\mathfrak{S}))$ .

Note that the set  $W(\mathfrak{S})$  does not depend on a concrete state transition diagram of the automaton  $\mathfrak{A}$ , but to be more definite, one may assume that the state transition diagram of the automaton is reduced; thus, given an automaton function, the reduced state transition diagram of respective automaton is unique; cf., Section 3.3. In this case, some care should be taken speaking of paths since some arrows in the reduced state transition diagram may actually be loops; see, e.g., Figure 19. The paths (which we write from left to right) that begin at the initial state  $t_0$  and have prefixes 0111, 01011, 010011, 0100011, ... all reach the subautomaton  $\mathfrak{S}_3$  on the fourth, fifth, sixth, seventh, ... steps respectively, so the probability to reach the subautomaton  $\mathfrak{S}_3$  is  $1/16 + 1/32 + 1/64 + 1/128 + \dots = 1/8$  and  $\mathbf{B}(\mathfrak{S}_3)$  is a disjoint union of balls  $\mathbf{B}_{1/16}(14)$ ,  $\mathbf{B}_{1/32}(26)$ ,  $\mathbf{B}_{1/64}(50)$ , ...,  $\mathbf{B}_{1/2^k}(2 + 3 \cdot 2^{k-2})$ , ... where  $k = 4, 5, 6, \dots$

Given two minimal subautomata  $\mathfrak{S}$  and  $\mathfrak{T}$  of the automaton  $\mathfrak{A}$  that are finite and affine, by virtue of the minimality one has  $\mathbf{B}(\mathfrak{S}) \cap \mathbf{B}(\mathfrak{T}) = \emptyset$ ; thus, the probability that a random infinite path starting from the initial state reaches at a finite step some minimal subautomaton of the automaton  $\mathfrak{A}$  is the sum  $\sum \mu(\mathbf{B}(\mathfrak{S}))$  taken over all minimal subautomata  $\mathfrak{S}$  which are finite and affine. We call an automaton  $\mathfrak{A}$  *ultimately affine* if the probability is 1. Note that if an ultimately affine automaton is infinite, then, according to König’s lemma (also known as Beth’s tree theorem) [57], there are infinite paths that never reach states

belonging to these minimal subautomata. These paths constitute a  $\mu$ -measurable subset in  $\mathbb{Z}_p$  but the measure of the subset is 0 since the subset is a complement to a countable union of balls whose measure is 1. For instance, the path 111... in the state transition diagram depicted by Figure 2 never reaches a minimal subautomaton (which has only one state, namely,  $s_1$ ), but all other paths reach the subautomaton at finite steps, so the probability to reach that minimal subautomaton is 1.

**Definition 12** (Plot equivalence of automata). *Call the finite affine automata  $\mathfrak{S}$  and  $\mathfrak{T}$  plot equivalent  $\mathfrak{S} \equiv_{\mathcal{P}} \mathfrak{T}$  if their respective functions  $\psi: \mathbb{R} \times \mathbb{Z} \rightarrow \mathbb{C}$  defined by (16) coincide; that is, if their limit plots coincide,  $\mathcal{P}'(\mathfrak{S}) = \mathcal{P}'(\mathfrak{T})$ , i.e., if the limit plots are links of the same number of torus windings with a common slope.*

Given  $a, b \in \mathbb{Z}_p \cap \mathbb{Q}$ , denote via  $\mathfrak{S}_{a,b}$  an automaton whose automaton function is  $z \mapsto az + b$ . Let  $[\mathfrak{S}_{a,b}]$  be the set of all minimal subautomata of  $\mathfrak{A}$  that are plot-equivalent to  $\mathfrak{S}_{a,b}$ . By virtue of the minimality, given  $\mathfrak{S}, \mathfrak{T} \in [\mathfrak{S}_{a,b}]$ , the subautomata  $\mathfrak{S}$  and  $\mathfrak{T}$  have no common states; therefore,  $\mathbf{B}(\mathfrak{S}) \cap \mathbf{B}(\mathfrak{T}) = \emptyset$ ; that is, the probability

$$q_{[\mathfrak{S}_{a,b}]} = \sum_{\mathfrak{S} \in [\mathfrak{S}_{a,b}]} \mu(\mathbf{B}(\mathfrak{S}))$$

is well-defined. Given  $a, b \in \mathbb{Z}_p \cap \mathbb{Q}$ , the equivalence relation  $\equiv_{\mathcal{P}}$  induces an equivalence relation on the set of all pairs  $(a; b) \in (\mathbb{Z}_p \cap \mathbb{Q}) \times (\mathbb{Z}_p \cap \mathbb{Q})$  which we denote by the same symbol, i.e.,  $(a; b) \equiv_{\mathcal{P}} (c; d)$  if and only if  $\mathfrak{S}_{a,b} \equiv_{\mathcal{P}} \mathfrak{S}_{c,d}$ .

Let  $\text{Spec}(\mathfrak{A})$  be the set of all equivalence classes defined by minimal subautomata of  $\mathfrak{A}$  which are finite and affine. Then, the series

$$\Psi_{\mathfrak{A}}(\rho, k) = \sum_{[\mathfrak{S}_{a,b}] \in \text{Spec}(\mathfrak{A})} q_{[\mathfrak{S}_{a,b}]} e^{i(a\rho - 2\pi p^k b)} \tag{17}$$

converges absolutely for all  $\rho \in \mathbb{R}, k \in \mathbb{Z}$  and therefore defines a complex-valued function  $\Psi_{\mathfrak{A}}(\rho, k)$ . Call the function  $\Psi_{\mathfrak{A}}$  a *sharp wave function* assigned to the automaton  $\mathfrak{A}$ .

**Theorem 17** (On automata having a prescribed wave function). *Given non-negative real numbers  $q_1, q_2, \dots$  such that  $\sum_{j=1}^{\infty} q_j = 1$  and finite affine automata  $\mathfrak{S}_j = \mathfrak{S}_{a_j, b_j}$ , ( $a_j, b_j \in \mathbb{Z}_p \cap \mathbb{Q}$ ,  $j = 1, 2, \dots$ ) which are pairwise plot-nonequivalent, there exists an ultimate affine automaton  $\mathfrak{A}$  such that  $\text{Spec}(\mathfrak{A}) = \{[\mathfrak{S}_j]: j = 1, 2, \dots\}$ ,  $q_j = q_{[\mathfrak{S}_{a_j, b_j}]}$ , and  $\Psi_{\mathfrak{A}}(\rho, k) = \sum_{j=1}^{\infty} q_j e^{i(a_j \rho - 2\pi p^k b_j)}$ .*

To prove the theorem we require a lemma.

**Lemma 2** (All discrete random variables can be modelled on  $\mathbb{Z}_p$ ). *Given convergent series  $\sum_{j=0}^{\infty} q_j = 1$  of positive real numbers  $q_j \in \mathbb{R}_{\geq 0}$  there exist pairwise disjoint open sets  $W_j \subset \mathbb{Z}_p$  such that the normalised Haar measure  $\mu$  of  $W_j$  is  $q_j$ ,  $j = 0, 1, 2, \dots$*

**Proof of Lemma 2.** Most likely, the lemma is known, but as the author is aware of no proper reference, a proof follows. Consider the *Monna map*  $\text{mon}(z) = \sum_{i=0}^{\infty} \alpha_i p^{-i-1} = 0.\alpha_0\alpha_1\alpha_2\dots \in [0, 1] \subset \mathbb{R}$  where  $z = \sum_{i=0}^{\infty} \alpha_i p^i$  is a  $p$ -adic canonical expansion of  $z \in \mathbb{Z}_p$ . Note that  $\text{mon}(\mathbf{B}_{1/p^k}(a)) = [0.\alpha_0\alpha_1\dots\alpha_{k-1}, 0.\alpha_0\alpha_1\dots\alpha_{k-1} + p^{-k}] \subset [0, 1]$ , where  $a = \sum_{i=0}^{k-1} \alpha_i p^i \in \mathbb{Z}_p$ ; that is, the Monna map  $\text{mon}$  maps  $p$ -adic balls  $\mathbf{B}_{1/p^k}(a) \subset \mathbb{Z}_p$  of radii  $1/p^k$  centred at  $a \in \mathbb{Z}_p$  onto closed subintervals of length  $1/p^k$  of the unit interval  $[0, 1]$ ; note that  $\lambda(\text{mon}(\mathbf{B}_{1/p^k}(a))) = \mu(\mathbf{B}_{1/p^k}(a))$  where  $\mu$  is the Haar measure on  $\mathbb{Z}_p$  normalised so that  $\mu(\mathbb{Z}_p) = 1$ , and  $\lambda$  is Lebesgue measure on the unit real interval  $[0, 1]$ , i.e., the length of the closed interval.

Split the unit interval  $[0, 1]$  into pairwise disjoint open intervals  $Q_j$  such that the length of the  $j$ -th interval  $Q_j$  is  $q_j$ ; namely, let  $Q_1 = (0, q_1)$ ,  $Q_2 = (q_1, q_1 + q_2)$ ,  $Q_3 = (q_1 + q_2, q_1 + q_2 + q_3)$ , ...; then,  $Q = \bigcup_{j=1}^{\infty} Q_j$  is  $\lambda$ -measurable and  $\lambda(Q) = 1$ .

For each  $Q_j$  let  $B_j$  be a set of all balls of nonzero radii such that  $\text{mon}(\mathbf{B}) \subset Q_j$  for every  $\mathbf{B} \in B_j$ . As any two  $p$ -adic balls either disjoint or one is a subset of another one, the set  $B_j$  is a countable disjoint union of balls of nonzero radii. Thus,  $B_j$  is open as each  $p$ -adic ball of nonzero radius is clopen; hence,  $B_j$  is  $\mu$ -measurable. As every point from  $Q_j$  lies in mon-image of some ball from  $B_j$ , we conclude that  $\mu(B_j) = q_j$  and  $\mu(\bigcup_{j=1}^{\infty} B_j) = \sum_{j=1}^{\infty} \mu(B_j) = 1$  as  $B_j \cap B_k = \emptyset$  when  $j \neq k$  by the construction.  $\square$

**Proof of Theorem 17.** This proof follows immediately from the proof of Lemma 2. Every  $B_j, j = 1, 2, \dots$  is a countable disjoint union of balls  $\mathbf{B}_{1/p^{r_{jm}}}(a_{jm}), m = 1, 2, 3, \dots$ , centred at  $a_{jm} = \sum_{k=0}^{r_{jm}-1} \alpha_{j,m,k} p^k \in \mathbb{Z}_p$ . Let branches of a  $p$ -adic tree be  $\alpha_{j,m,0} \alpha_{j,m,1} \dots \alpha_{j,m,r_{jm}-1}$ , and let leafs be  $\mathbf{B}_{1/p^{r_{jm}}}(a_{jm}), j, m = 1, 2, \dots$ . In this digraph, replace all leafs  $\mathbf{B}_{1/p^{r_{jm}}}(a_{jm})$  with state transition diagrams of automata  $\mathfrak{S}_m \in [\mathfrak{S}_j]$ . Thus, the constructed digraph is a state transition diagram of the automaton  $\mathfrak{A}$  which is the ultimate affine and such that  $\Psi_{\mathfrak{A}}(\rho, k) = \sum_{j=1}^{\infty} q_j e^{i(a_j \rho - 2\pi p^k b_j)}$ .  $\square$

**Note 12.** From the proof of Theorem 17 it follows that the ultimate affine automaton may be either measure-0 or measure-1. The first case occurs when, for example, the series  $\sum_{j=0}^{\infty} q_j$  is finite; therefore the automaton  $\mathfrak{A}$  is finite and thus measure-0. The measure-1 case occurs when, for example, all coefficients  $a_j \in \mathbb{Z}_p \cap \mathbb{Q}$  constitute a dense subset in  $\mathbb{R}$  and all  $b_j = 0$ .

In what follows, we will need a slightly generalised version of Lemma 2:

**Corollary 1 (Generalized Lemma 2).** Given convergent series  $\sum_{j=0}^{\infty} q_j = q \leq 1$  of positive real numbers  $q_j \in \mathbb{R}_{\geq 0}$ , there exist pairwise disjoint open sets  $W_j \subset \mathbb{Z}_p$  such that the normalized Haar measure  $\mu$  of  $W_j$  is  $q_j, j = 0, 1, 2, \dots$

**Proof of Corollary 1.** Take  $[0, q]$  instead of  $[0, 1]$  in the proof of Lemma 2 and modify the argument in an obvious way.  $\square$

Sharp wave functions may be considered as wave functions with respect to *discrete time* since the map  $e^{2\pi i b} \mapsto e^{2\pi i p^k b}$  is equivalent to a  $k$ -digit *shift* of the base- $p$  representation of  $b$  and a reduction modulo 1 of the resulting number. As  $k$  is the order of time elapsed (and is measured by  $p$ -adic clock see Section 5.2 and Figure 3) since the moment the automaton reaches a state from its minimal affine subautomaton whose automaton function is  $z \mapsto az + b$ , a sharp wave function may be judged as the one the Little-endian can construct by observing reactions of a physical system at the smallest of scales.

We argue that a wave function with respect to *continuous time* can also be constructed by using ultimate affine automata. The core idea of the construct is using the beta representations of numbers rather than the base- $p$  expansions. The beta representations of real numbers were first introduced by A. Rényi in 1957 and since then have attracted substantial attention in ergodic theory and symbolic dynamics; see, e.g., monograph [21].

Recall that given real  $\beta > 1$ , a  $\beta$ -representation of real  $b \geq 0$  is an infinite word  $\chi_0 \chi_1 \dots$  over the alphabet  $\mathcal{B} = \{0, 1, \dots, \lfloor \beta \rfloor\}$  such that  $b = \sum_{j=-k}^{\infty} \chi_{k+j} \beta^{-k-j}$ . Note that we consider  $\beta$ -representations of real  $b \geq 0$  and not only of real  $b \in [0, 1]$  as in [21]. Of course, in (17), we always may assume that  $b \in [0, 1]$ ; however, to assign real numbers to paths in state transition diagrams of automata we need beta representations of numbers from  $\mathbb{N}_0$  which then are converted into real numbers in a way similar to what we used in Section 5.3 by exploiting  $p$ -adic representations.

Specifically, we first use  $\beta$  instead of  $p$ . Thus, each arrow in a state transition diagram of the automaton whose input and output alphabets are  $\mathcal{B}$ , is labelled by a pair  $\chi|\xi$ , where  $\chi, \xi \in \mathcal{B}$ ; for an infinite path which starts from an initial state, there corresponds an infinite

word  $w = \chi_0\chi_1 \cdots$  over alphabet  $\mathcal{B}$ ; for  $w$ , we place a corresponding  $(\lfloor \beta \rfloor + 1)$ -adic integer  $\sum_{j=0}^{\infty} \chi_j(\lfloor \beta \rfloor + 1)^j$ . To construct a plot, we convert these  $(\lfloor \beta \rfloor + 1)$ -adic integers into sequences of real numbers  $\chi_0\beta^{-1}, \chi_1\beta^{-1} + \chi_0\beta^{-2}, \chi_2\beta^{-1} + \chi_1\beta^{-2} + \chi_0\beta^{-3}, \dots$ , thus obtaining points  $(\chi_{k-1}\beta^{-1} + \cdots + \chi_0\beta^{k-2}; \xi_{k-1}\beta^{-1} + \cdots + \xi_0\beta^{k-2}) \in \mathbb{R}^2$ . To put it in other words, we simply use  $\beta$ -representations for input/output words of the automaton  $\mathfrak{A}$  when constructing a plot of the automaton, but the automaton function is still a 1-Lipschitz map from  $(\lfloor \beta \rfloor + 1)$ -adic integers to  $(\lfloor \beta \rfloor + 1)$ -adic integers. This way, we construct a sharp wave function  $\Psi_{\mathfrak{A}}(\rho, k) = \sum_{[\mathfrak{S}_{a,b}] \in \text{Spec}(\mathfrak{A})} q_{[\mathfrak{S}_{a,b}]} e^{i(a\rho - 2\pi(\lfloor \beta \rfloor + 1)^k b)}$  (cf., (17)), which is a well-defined complex valued-function of  $\rho \in \mathbb{R}$  and  $k \in \mathbb{Z}$ ; then, we replace  $(\lfloor \beta \rfloor + 1)$  by  $\beta$  in the formula, thus resulting in another complex-valued function of  $\rho \in \mathbb{R}$  and  $k \in \mathbb{Z}$ . The crucial point is that if  $1 < \beta \ll 2$ , i.e., if  $\beta = 1 + \tau$  where  $0 < \tau \ll 1$ , then  $\beta^k = (1 + \tau)^k \approx 1 + k\tau$ . When  $\tau$  is small (e.g., if  $\tau = 5.391247(60) \times 10^{-44}$  s, the Planck time) then for the Big-endian observer who is incapable of performing measurements with that accuracy (which is currently only about  $10^{-20}$  s),  $k\tau \in \mathbb{R}$  is indistinguishable from continuous time. Thus, we obtain a *fuzzy wave function*

$$\Psi_{\mathfrak{A}}(\rho, t) = \sum_{[\mathfrak{S}_{a,b}] \in \text{Spec}(\mathfrak{A})} q_{[\mathfrak{S}_{a,b}]} e^{i(a\rho - 2\pi t b)} \tag{18}$$

which is ascribed to the automaton  $\mathfrak{A}$ . The function is well-defined for all  $\rho, t \in \mathbb{R}$  since the series converges absolutely. From this point, the *sharp wave function (which is a discrete time function) can be viewed as an approximation of a fuzzy wave function (which is a continuous time function)*. Note that since  $\lfloor \beta \rfloor = \lfloor 1 + \tau \rfloor = 1$ , i.e.,  $\mathcal{B}$  is a 2-letter alphabet, then necessarily  $p = 2$ ; see sharp wave function Formula (17).

The term “approximation” here is not rigorous (although some hint is already given by Example 4); to prove this statement with a full rigour is a separate problem which will be considered in the future. In the current paper, we only find an exact representation for  $\beta = 1 + \tau$  under the *finiteness assumption* of Section 2, but before doing this, we illustrate the usage of that  $\beta$ -representation using the analogy of film which is discussed in Section 2. Each frame of a film contains a number of details, but to cause an illusion of motion to a viewer, only a small share of the whole number of details is changed from one frame to the next frame; the smaller the share is, the slower the motion appear to the a viewer. For a Little-endian viewer, the share is  $p - 1$  since he uses the base- $p$  representation of numbers; in the case when the share is  $\tau$ , one has the  $(1 + \tau)$ -representation. If  $0 < \tau \ll 1$ , we have the case of a Big-endian viewer.

It is important to stress that *to represent numbers from  $\mathbb{N}_0$  in the base  $\beta$ , we use only non-negative powers of  $\beta$  in order to guarantee the uniqueness of  $\beta$ -representation for each number from  $\mathbb{N}_0$  since if negative powers of  $\beta = 1 + \tau$  when  $\tau \ll 1$  are allowed in  $\beta$ -representations, then every number from  $(0, \tau^{-1})$  has a continuum of distinct  $\beta$ -representations provided  $\tau < \frac{\sqrt{5}-1}{2}$  [58]. However, in such a case, the very problem of assigning a number to a finite path in a state transition diagram becomes ill-posed. Under said convention, the following theorem is true:*

**Theorem 18** (Finiteness assumption implies  $\beta = \sqrt[N]{2}$ ). *Let  $1 < \beta < 2$ . If an automaton that performs the addition of  $\beta$ -representations of numbers from  $\mathbb{N}_0$  is finite then necessarily  $\beta = \sqrt[N]{2}$  for some  $N \in \mathbb{N}$ . For each  $N \in \mathbb{N}$ , the addition of numbers from  $\mathbb{N}_0$  that are represented by  $\sqrt[N]{2}$ -representations can be performed with a finite automaton.*

**Proof.** Number 1 admits the only  $\beta$ -representation  $1 = 1 + 0 \cdot \beta + 0 \cdot \beta^2 + \cdots$  in non-negative powers of  $\beta$  as  $\beta > 1$ . A finite automaton ultimately maps periodic sequences onto ultimately periodic sequences; therefore, if a finite automaton that maps pairs of

infinite words into infinite words over the alphabet  $\mathcal{B} = \{0, 1\}$  and performs  $1 + 1 = 2$ , then necessarily

$$2 = \alpha_0 + \alpha_1\beta + \dots + \alpha_{n-1}\beta^{n-1} + (\gamma_0 + \gamma_1\beta + \dots + \gamma_{s-1}\beta^{s-1})\beta^n + (\gamma_0 + \gamma_1\beta + \dots + \gamma_{s-1}\beta^{s-1})\beta^{2n} + \dots = \alpha_0 + \alpha_1\beta + \dots + \alpha_{n-1}\beta^{n-1} + (\gamma_0 + \gamma_1\beta + \dots + \gamma_{s-1}\beta^{s-1})\beta^n(1 + \beta^n + \beta^{2n} + \dots),$$

where  $\alpha_i, \gamma_j \in \{0, 1\}$ . As the series  $1 + \beta^n + \beta^{2n} + \dots$  diverges, then all  $\gamma_j = 0$ ; hence,

$$2 = \alpha_0 + \alpha_1\beta + \dots + \alpha_{k-1}\beta^{k-1} + \beta^k, \tag{19}$$

for suitable  $k \leq n - 1, \alpha_i \in \{0, 1\}$ . If  $\alpha_0 = 1$ , then the right-hand side of (19) is not equal to the left-hand side; therefore,  $\alpha_0 = 0$ , and by substituting  $\beta = 1 + \tau$  and collecting terms of positive degrees in  $\tau$  we obtain the following (by binomial theorem):

$$2 = \alpha_1\beta + \dots + \alpha_{k-1}\beta^{k-1} + \beta^k = (\alpha_1 + \dots + \alpha_{k-1} + 1) + \tau u(\tau),$$

where  $u(x)$  is a polynomial of variable  $x$  whose coefficients are in  $\mathbb{N}_0$ . Hence,  $1 = \alpha_1 + \dots + \alpha_{k-1} + \tau u(\tau)$ , where  $\alpha_j \in \{0, 1\}, j = 1, 2, \dots, k - 1$ .

If  $u(x)$  is a nonzero polynomial, then  $\tau u(\tau) > 0$ ; thus, as  $\alpha_1 + \dots + \alpha_{k-1} \in \mathbb{N}_0$ , we must conclude that  $\alpha_1 + \dots + \alpha_{k-1} = 0$ : Otherwise, the right-hand side in  $1 = \alpha_1 + \dots + \alpha_{k-1} + \tau u(\tau)$  is strictly greater than is the left-hand side. Therefore, all  $\alpha_j = 0$  and thus  $2 = \beta^k$ , i.e.,  $\beta = \sqrt[k]{2}$ .

If  $u(x)$  is a zero polynomial, then necessarily  $\alpha_1 + \dots + \alpha_{k-1} = 1$ . Therefore, there must be exactly one nonzero  $\alpha_j$ ; hence,  $2 = \beta^j + \beta^k$ , where  $0 < j < k$ . However,  $2 \neq \beta^j + \beta^k$  since  $\beta > 1$ ; so we get a contradiction.

The converse statement of the theorem is obvious since the addition of numbers represented by  $\sqrt[N]{2}$ -expansions is an “addition with carry to the  $N$ -th digit”; for example, when  $N = 2$  one has

$$\begin{array}{cccccccc} \dots 1 & 1 & 1 & 1 & 1 & 1 & 1 & = & - & \sqrt{2} & - & 1 \\ + & & & & & & & & & & & \\ \dots 0 & 0 & 0 & 0 & 0 & 0 & 1 & = & 1 \\ \hline \dots 0 & 1 & 0 & 1 & 0 & 1 & 0 & = & - & \sqrt{2} \end{array}$$

□

It is worth warning the reader that Theorem 18 is *not* about the calculation of Planck time, whose value depends on the choice of units. In short, *Theorem 18 is about how much information one needs to have both worldviews, that of the Little-endian and the Big-endian, agree.* Specifically, Theorem 18 implies that the fuzzy wave function is the one which corresponds to an automaton over a  $2^N$ -symbol alphabet; that is, to the automaton whose function is  $f: \mathbb{Z}_2^N \rightarrow \mathbb{Z}_2^N$ , i.e., a  $N$ -variate 2-adic 1-Lipschitz map; see Section 3.3. Actually,  $f$  is a 1-Lipschitz map  $\mathbb{Z}_2(\sqrt[N]{2}) \rightarrow \mathbb{Z}_2(\sqrt[N]{2})$ , where  $\mathbb{Z}_2(\sqrt[N]{2})$  is the ring of integers of the field  $\mathbb{Q}_2(\sqrt[N]{2})$ ; we leave further discussion of theory to future papers.

We remind the reader that for multivariate  $p$ -adic 1-Lipschitz maps, most theorems that have been proven or mentioned in this paper hold true; in particular, Theorem 15 holds true. Given a real function  $G: H \rightarrow \mathbb{R}^n$  whose domain is  $H \subset \mathbb{R}^m$ , by the *graph of the function* (on the torus  $\mathbb{T}^{m+n}$ ), we mean the point subset  $\mathbf{G}_H(g) = \{(\vec{x} \bmod 1; G(\vec{x}) \bmod 1) : \vec{x} \in H\} \subset \mathbb{T}^{m+n}$ . Note that if  $\vec{y} = (y_1; \dots; y_k) \in \mathbb{R}^k$ , then  $\vec{y} \bmod 1$  stands for  $(y_1 \bmod 1; \dots; y_k \bmod 1)$ .

**Theorem 19** ([10]). Let  $\mathfrak{A}$  be a finite automaton over the alphabet  $\{0, 1, \dots, p - 1\}$ , let  $\mathfrak{A}$  have  $m$  inputs and  $n$  outputs, and let  $G = (G_1; \dots; G_n): [\vec{a}, \vec{b}] = [a_1, b_1] \times \dots \times [a_m, b_m] [0, 1]^n$  (where  $[a_k, b_k] \subset [0, 1)$ ,  $G_i: [\vec{a}, \vec{b}] \rightarrow [0, 1)$ ,  $k = 1, 2, \dots, m$ ) be a two-times differentiable function such that all its second partial derivatives are continuous on  $[\vec{a}, \vec{b}]$ . If  $\mathbf{G}(G)$  is a subset in a plot  $\mathcal{P}(\mathfrak{A}) \subset \mathbb{T}^{m+n}$  of the automaton  $\mathfrak{A}$ , then there exist an  $m \times n$  matrix  $\mathbf{D} = (d_{kj})$  and a vector  $\vec{c} = (c_1; \dots; c_n)$  such that  $d_{kj} \in \mathbb{Q} \cap \mathbb{Z}_p$ ,  $c_j \in \mathbb{Q} \cap \mathbb{Z}_p \cap [0, 1)$  ( $k = 1, 2, \dots, m$ ;  $j = 1, 2, \dots, n$ ) and  $G(\vec{x}) = (\vec{x}\mathbf{D} + \vec{c}) \bmod 1$  for all  $\vec{x} \in [\vec{a}, \vec{b}]$ . There are not more than a finite number of  $\mathbf{D}$  and  $\vec{c}$  such that  $d_{kj} \in \mathbb{Q} \cap \mathbb{Z}_p$ ,  $c_j \in \mathbb{Q} \cap \mathbb{Z}_p \cap [0, 1)$  ( $k = 1, 2, \dots, m$ ;  $j = 1, 2, \dots, n$ ) and  $\mathbf{G}_{[\vec{a}, \vec{b}]}((\vec{x}\mathbf{D} + \vec{c}) \bmod 1) \subset \mathcal{P}(\mathfrak{A})$  for some  $[\vec{a}, \vec{b}] \subset [0, 1)^m$ ; moreover, if  $\mathbf{G}_{[\vec{a}, \vec{b}]}(\vec{x}\mathbf{A} + \vec{c}) \subset \mathcal{P}(\mathfrak{A})$  for some  $[\vec{a}, \vec{b}] \subset [0, 1)^m$  then  $\mathbf{G}_{\mathbb{R}^m}((\vec{x}\mathbf{D} + \vec{c}) \bmod 1) \subset \mathcal{P}(\mathfrak{A}) \subset \mathbb{T}^{n+m}$ .

The theorem implies that in the multivariate case, the sharp wave function is of the following form:

$$\Psi_{\mathfrak{A}}(\vec{x}, r) = \sum_{[\mathfrak{S}_{\mathbf{A}, \vec{b}}] \in \text{Spec}(\mathfrak{A})} q_{[\mathfrak{S}_{\mathbf{A}, \vec{b}}]} e^{i(\vec{x}\mathbf{A} - 2\pi p^r \vec{b})}; (\vec{x} \in \mathbb{R}^m; \vec{b} \in \mathbb{R}^n; r \in \mathbb{Z}).$$

Therefore, Theorem 18 implies that a univariate fuzzy wave function is actually a multivariate sharp wave function; however, it is for a large number of dimensions. For instance, if  $\sqrt[2]{2} = 1 + \tau$  where  $\tau$  is of order of Planck time, then  $N \approx \frac{\ln 2}{\tau} \approx 10^{43}$ ; that is, the automaton function of respective automaton is a 1-Lipschitz map  $\mathbb{Z}_2^{10^{43}} \rightarrow \mathbb{Z}_2^{10^{43}}$ . This means that the matrices  $\mathbf{A}$  in the above formula for the sharp wave function  $\Psi_{\mathfrak{A}}(\vec{x}, r)$  are  $10^{43} \times 10^{43}$ ; that is, each of the matrices contains more entries than the number of atoms in the universe. An infinite-dimensional space is an adequate model for a  $10^{43}$ -dimensional space; this is why both the Big-endian and Little-endian would agree that wave functions “live” in Hilbert spaces. We postpone to a future paper more rigorous statements and proofs on how pure and fuzzy wave functions are related one to another; here, we only explain why both functions, which may be judged as “physical”, are elements of Hilbert space  $\ell^2(\text{Spec}(\mathfrak{A}))$  of square-summable complex sequences whose terms are indexed by elements of the set  $\text{Spec}(\mathfrak{A})$  (which is countable) since a “physical” wave function must be square-summable and the sum of squares of probability amplitudes must be 1. Recall that any separable Hilbert space is metrically isomorphic to  $\ell^2$  and that the Fourier transform on the circle is such an isomorphism between the Hilbert space of square-integrable functions on  $[0, 1] = \mathbb{I}$  and the space  $\ell^2(\mathbb{Z})$  of square-summable complex sequences whose terms are enumerated by integers. It is not difficult to construct sharp wave functions which can be judged as “physical” with this meaning. Indeed, take any sequence  $q_1, q_2, \dots$  of positive real numbers such that  $\sum_{j=1}^{\infty} q_j = 1$ , and the series  $\sum_{j=1}^{\infty} \sqrt{q_j}$  of positive square roots converges; by using Theorem 17, construct the automaton  $\mathfrak{A}$ . Then, function  $\sum_{j=1}^{\infty} \sqrt{q_j} e^{i(a_j p - 2\pi p^k b_j)}$  is the one we are seeking.

We finalise the subsection with the following interpretation.

**Interpretation 8** (Discrete spectrum; continuous spectrum). *The measure-0 ultimate affine automata may be treated as models of physical systems having discrete (energy, frequency, ...) spectra, while measure-1 ultimate affine automata may be treated as models of physical systems having continuous spectra.*

### 5.5. Uncertainty

In this subsection, we formally derive an uncertainty relation which holds for wave functions of automata. We stress, once again, that *despite the Little-endian being capable of performing observation at the smallest scale and the Big-endian not being able to do so, the*

uncertainty relation, which can be treated as a time-energy uncertainty, holds for both observers, i.e., for Little-endian as well as for Big-endian; thus, no hidden parameters are assumed.

The uncertainty relation we are going to deduce is an entropic one. A number of research papers have been devoted to discussing entropic uncertainty relations; see, e.g., the expository paper [59] and the references therein. The entropic uncertainty relation derived below is of a novel type since it relates the time during which a system reaches a “pure state” that can be ascribed to a minimal affine subautomaton and the state (i.e., an element of  $\text{Spec}(\mathfrak{A})$ ) itself. Note that as the Little-endian is capable of performing measurements at the smallest of scales, the time a system reaches a state that belongs to some minimal automaton is not 0, i.e., the “wave function collapse” is not momentary, it takes some minimal time intervals (e.g., some Planck time). Note that the collapse of wave functions as a finite-time process is discussed in the literature; see, e.g., [60].

To start with, we need to restate some results from Section 5.4 in terms of prefix codes since in what follows, we use some basic properties of the codes which may be found, e.g., in the book [61].

**Definition 13** (Prefix code). *A nonempty set  $\mathbf{C}$  of finite nonempty words over a finite alphabet  $\mathcal{A}$  that consists of  $p > 1$  letters is called a prefix code if each word from  $\mathbf{C}$  is a prefix of no other word from  $\mathbf{C}$ .*

Let words from the nonempty set  $\mathbf{G}$  of finite nonempty words over  $\mathcal{A}$  be ordered with respect to a nondecreasing order of their lengths, and let  $\ell_i$  be the length of the  $i$ -th word (so  $\ell_1 \leq \ell_2 \leq \dots$ ). The set  $\mathbf{G}$  is a prefix code if and only if the following Kraft–McMillan inequality holds:

$$\sum_{i=1}^{\infty} p^{-\ell_i} \leq 1. \tag{20}$$

**Note 13.** *From the proof of Theorem 17, it follows that the branches of the state transition diagram constitute a prefix code since each word which corresponds to a branch of length  $k$  reaches some minimal affine subautomaton exactly at the  $k$ -th step, thus, the word cannot be a prefix of any other word which corresponds to another branch. Note that words begin from the root of the tree, and the root is the initial state in the state-transition diagram. From the construction, it follows that the Kraft–MacMillan inequality for that code is equality. However, by using Corollary 1 rather than Lemma 2 in the proof, one constructs a prefix code such that  $\sum_{i=1}^{\infty} p^{-\ell_i} = q \leq 1$  for any given  $0 < q \leq 1$ . In this case, the rest infinite paths of the complete  $p$ -adic tree that lead to no minimal finite affine subautomaton constitute a set of Haar measure  $1 - q$ . The automaton having such a state transition diagram will reach minimal subautomata which are finite and affine with probability  $0 < q \leq 1$  rather than exactly 1. In that case, to automaton  $\mathfrak{A}$ , there corresponds a sharp wave function of the form (17) such that  $\sum_{[\mathfrak{S}_{a,b}] \in \text{Spec}(\mathfrak{A})} q_{[\mathfrak{S}_{a,b}]} = q$  which therefore is normalisable. For not to overload the exposition, in what follows we mostly deal with the case when  $q = 1$ , i.e., with ultimately affine automata  $\mathfrak{A}$ , cf. Section 5.4.*

Let  $X$  be a random variable on the prefix code  $\mathbf{G} = (w_i)_{i=1}^{\infty}$ ; we denote via  $q_i = \text{Prob}(X = w_i)$  the probability that  $X$  is equal to the word  $w_i$ . By definition [61], the entropy  $\mathbf{H}(X)$  of the random variable  $X$  is  $\mathbf{H}(X) = -\sum_{i=1}^{\infty} q_i \log_p q_i$ , whereas the mean length of the codeword is  $\mathbf{E}(X) = \sum_{i=1}^{\infty} q_i \ell_i$ .

There exists a prefix code such that  $\ell_i = \lceil \log_p(1/q_i) \rceil$  for which the right-hand side inequality in (21) below holds (that right-hand side inequality is not true in general). The left-hand side inequality in (21) below holds whenever  $\mathbf{H}(X) < +\infty$  and  $\mathbf{E}(X) < +\infty$ , becoming an equality if and only if  $q_i = p^{-\ell_i}$  [61] [Theorem 4.3].

$$\mathbf{H}(X) \leq \mathbf{E}(X) < \mathbf{H}(X) + 1, \tag{21}$$

The time which a (both sharp and fuzzy) wave function takes to collapse can be expressed via the length of a word which reaches a state from some minimal affine subau-

tomaton since the length of the word is the order of time expressed in the respective base; see the explanations in Section 5.3. This is why *in what follows, we deal with the lengths of the words rather than with time itself*. Note that when dealing with the lengths of the words, we may restrict considerations to the words over the alphabet  $\{0, 1, \dots, p - 1\}$  where  $p$  is a prime since fuzzy wave functions are constructed by using words over the alphabet  $\{0, 1\}$ ; see Section 5.4. The only difference between sharp and fuzzy wave function constructions for  $p = 2$  are the numerical values that are assigned to words by both the observers: The Little-endian assigns numbers to words by reading the words as the base-2 expansions of numbers whereas the Big-endian reads these words as  $(1 + \tau)$ -representations of numbers for  $0 < \tau \ll 1$ . We stress that in what follows, “mean time of collapse” is synonymous with “mean word length” although the *actual* mean time of collapse *measured by the both observers* is different due to the inevitable nonzero measurement error. For instance, to the word of length  $k$  whose prefix is  $k - 1$  zeros and whose suffix is 1, the Little-endian assigns the value  $2^k$ , whereas the Big-endian assigns the value  $(1 + \tau)^k$ , which for small  $\tau$  and not too large  $k$ , is indistinguishable for this observer from 1 due to the measurement error. To put it in other words, the Little-endian’s measurements of time elapsed are much more accurate than are the Big-endian’s; the time within which the wave function collapses is large for the Little-endian, whereas that time is zero for the Big-endian up to the measurement accuracy of his equipment; although both the clocks the observers use are 2-adic, according to Theorem 18, the Big-endian can observe digits in the windows that are to the left of the  $(N - 1)$ -th window at the face of the clock for  $N$  large, whereas Little-endian observes digits to the left of the lowest order position, i.e., from the rightmost window (cf., Figure 3). Nevertheless, we are going to show that “time-energy” uncertainty in terms of the length of words in the state transition diagrams of automata holds for both observers.

Let  $\mathfrak{A}$  be an ultimately affine automaton, cf. Section 5.4. Define the *automaton entropy* as

$$H_{\mathfrak{A}} = - \sum_{[\mathfrak{S}_{a,b}] \in \text{Spec}(\mathfrak{A})} q_{[\mathfrak{S}_{a,b}]} \log_p q_{[\mathfrak{S}_{a,b}]}$$

For every  $[\mathfrak{S}_{a,b}] \in \text{Spec}(\mathfrak{A})$ , the probability  $q_{[\mathfrak{S}_{a,b}]}$  is equal to the sum of all  $p^{-\Lambda(w)}$ , where  $\Lambda(w)$  is the length of a finite word  $w$  that reaches some state that belongs to some subautomaton from  $[\mathfrak{S}_{a,b}]$  exactly at the  $\Lambda(w)$ -th step; the summation is over all these words. Let  $C[\mathfrak{S}_{a,b}]$  be a code whose codewords are all these words  $w$ ; then,  $\sum_{(w \in [\mathfrak{S}_{a,b}])} p^{-\Lambda(w)} = q_{[\mathfrak{S}_{a,b}]}$ . Note that  $\sum_{[\mathfrak{S}_{a,b}] \in \text{Spec}(\mathfrak{A})} \sum_{(w \in [\mathfrak{S}_{a,b}])} p^{-\Lambda(w)} = \sum_{[\mathfrak{S}_{a,b}] \in \text{Spec}(\mathfrak{A})} q_{[\mathfrak{S}_{a,b}]} = 1$ , the codes  $C[\mathfrak{S}_{a,b}]$  are disjointed for different  $[\mathfrak{S}_{a,b}] \in \text{Spec}(\mathfrak{A})$ , and the union of all these codes for all  $[\mathfrak{S}_{a,b}] \in \text{Spec}(\mathfrak{A})$  is a prefix code  $C(\mathfrak{A})$  such that  $\sum_{w \in C(\mathfrak{A})} p^{-\Lambda(w)} = 1$ . According to the above convention, the mean time  $T(\mathfrak{A})$  of wave function collapse is the mean length of a codeword of the code  $C(\mathfrak{A})$ :

$$T(\mathfrak{A}) = \sum_{w \in C(\mathfrak{A})} \Lambda(w) p^{-\Lambda(w)}$$

The inequality (21) implies that

$$T(\mathfrak{A}) \geq H_T = - \sum_{w \in C(\mathfrak{A})} p^{-\Lambda(w)} \log_p(p^{-\Lambda(w)}) = \sum_{w \in C(\mathfrak{A})} \Lambda(w) p^{-\Lambda(w)} = T(\mathfrak{A}),$$

where  $H_T$  is the entropy of the code  $C(\mathfrak{A})$ . Therefore, *the mean time of collapse of the automaton wave function is equal to the entropy of the code  $C(\mathfrak{A})$* .

For every  $n \in \mathbb{N}$ , let  $T_n$  be a set of all codewords of length  $n$  from the code  $C(\mathfrak{A})$ . If  $T_n \neq \emptyset$ , then  $T_n$  is a prefix code. All of these codes are disjointed, and their union is  $C(\mathfrak{A})$ . Therefore,  $P_N = \sum_{n=1}^N \sum_{w \in T_n} p^{-n}$  is the probability that the wave function collapses for a time not greater than  $N$ . Let  $H_{\text{time} \leq N}(\mathfrak{A})$  be the entropy of the prefix code  $T(N) = \bigcup_{n=1}^N T_n$ ; that is,  $H_{\text{time} \leq N}(\mathfrak{A}) = - \sum_{w \in T(N)} p^{-\Lambda(w)} \log_p(p^{-\Lambda(w)}) = \sum_{w \in T(N)} \Lambda(w) p^{-\Lambda(w)}$ . As the probability assigned to  $w \in C(\mathfrak{A})$  is  $p^{-\Lambda(w)}$  and as  $\sum_{w \in C(\mathfrak{A})} p^{-\Lambda(w)} = 1$ , then for  $N$  not

less than the length of the shortest word from  $C(\mathfrak{A})$ , it holds that  $H_{\text{time} \leq N}(\mathfrak{A}) + H_{\mathfrak{A}} > 0$ . Moreover, as  $H_{\text{time} \leq +\infty}(\mathfrak{A}) = \sum_{w \in C(\mathfrak{A})} \Lambda(w) p^{-\Lambda(w)} \geq \sum_{w \in C(\mathfrak{A})} p^{-\Lambda(w)} = 1$ , then

$$H_{\text{time} \leq +\infty}(\mathfrak{A}) + H_{\mathfrak{A}} \geq 1.$$

If to minimal affine subautomata there are ascribed “energy levels” (e.g., if in the subautomata functions  $z \mapsto az + b$ , the coefficients  $a$  are different and  $b = 0$ ) these inequalities may be judged as time-energy uncertainty relation since *if an observer measures the time which a wave function takes to collapse, he does not know for sure to which of the states the wave function has collapsed; on the other hand, if he knows to which of the states the wave function has collapsed, he does not know for sure how much time the collapse has taken.*

In a general case, these inequalities cannot be sharpened. Since  $\bigcup_{n=1}^{\infty} T_n = C(\mathfrak{A})$ , then  $C(\mathfrak{A})$  can be split arbitrarily into the disjointed union of sets  $D_1, D_2, \dots$ , and as each of  $D_j$  is itself a prefix code, there is an automaton  $\mathfrak{D}$  such that  $C[\mathfrak{S}_{a,b}] = D_j, (j = 1, 2, \dots)$ . Indeed, the entropy  $H_{\text{time} \leq +\infty}(\mathfrak{A})$  is determined by the code  $C(\mathfrak{A})$  only, whereas  $H_{\mathfrak{A}}$  is determined completely by the partition of the code  $C(\mathfrak{A})$  into arbitrary nonempty subsets and by the “assigning of limit plots” to each of the subsets.

The codeword lengths in  $C(\mathfrak{A})$  can be arbitrary as well.

**Theorem 20** (On maximal prefix codes [62]). *For every non-decreasing map  $\ell: \mathbb{N} \rightarrow \mathbb{N}$  such that  $\sum_{n=1}^{\infty} p^{-\ell(n)} = d \leq 1$  there exists a (maximal) prefix code  $C = \{w_n: n \in \mathbb{N}\}$  such that  $\Lambda(w_n) = \ell(n)$ , for all  $n \in \mathbb{N}$ .*

That is, one can take any such code  $C$  for  $d = 1$ , split all its codewords into a partition  $\mathbf{P}(C)$  of nonempty subsets, assign to every subset  $S \in \mathbf{P}(C)$  a limit plot of a finite automaton  $\mathfrak{S}_{a,b}(S)$  arbitrarily, and construct a respective automaton  $\mathfrak{A}$  so that  $C(\mathfrak{A}) = C$  and all finite paths in every  $S$  lead to the  $\mathfrak{S}_{a,b}(S)$ .

We have that  $H_{\mathfrak{A}} \in [0, +\infty], H_{\text{time} \leq +\infty}(\mathfrak{A}) \in [1, +\infty], T(\mathfrak{A}) \in [1, +\infty]$  (as  $T(\mathfrak{A}) = H_T = H_{\text{time} \leq +\infty}(\mathfrak{A})$ ), and nothing more definite can be said in the general case. It is possible that  $H_{\text{time} \leq +\infty}(\mathfrak{A}) = 1$ . For instance, let  $p = 2$ , and let  $C(\mathfrak{A}) = \{1, 01, 001, 0001, \dots\}$ . Then, the entropy  $H_{\mathfrak{A}}$  may be equal to 1 if different limit plots are assigned to different balls  $\mathbf{B}_{2^{-n-1}}(2^n)$ . The entropy  $H_{\mathfrak{A}}$  may be zero if the limit plots that are assigned to all these balls are equal one to another. One may split the set of all these balls into a partition of pairwise disjointed nonempty subsets and assign to each ball a limit plot so that to all balls from a subset, the same limit plot is assigned, but to balls that belong to different subsets, one assigns different limit plots. In all these cases,  $H_{\text{time} \leq +\infty}(\mathfrak{A}) = 1$  (as the entropy is equal to  $T(\mathfrak{A})$ ), but the entropies  $H_{\mathfrak{A}}$  are different.

Finally, consider *generating series*  $f_{\mathfrak{A}}(x) = \sum_{w \in C(\mathfrak{A})} x^{\Lambda(w)} = \sum_{n=1}^{\infty} t_n x^n$ , where  $t_n$  is the number of all words of length  $n$  in the prefix code  $C(\mathfrak{A})$ . As  $f_{\mathfrak{A}}(1/p) = 1$ , then for the radius  $R_{\mathfrak{A}}$  of convergence of the series, it holds that  $R_{\mathfrak{A}} \geq 1/p$ , with  $f_{\mathfrak{A}}(1/p) = 1$ . Hence, the function  $f_{\mathfrak{A}}(x)$  is differentiable at all points from  $(-R_{\mathfrak{A}}, R_{\mathfrak{A}})$ , but if  $R_{\mathfrak{A}} = 1/p$ , then the derivative  $f'_{\mathfrak{A}}(x)$  may not exist at  $x = 1/p$  or may go to  $+\infty$ . However,  $f'_{\mathfrak{A}}(1/p) = p \cdot T(\mathfrak{A})$ , i.e., the derivative  $f'_{\mathfrak{A}}(1/p)$  determines the entropy  $H_{\text{time} \leq +\infty}(\mathfrak{A})$ .

### 6. Discussion

In the paper, a number of mathematical statements are rigorously proven which, as a whole, advocate that answers to the questions as whether Nature at the smallest of scales is discrete or continuous, random and chaotic, or deterministic and predictable, solely depend on the free choice of metric, real or  $p$ -adic, with respect to which numerical experimental data are processed. The core idea is that rational  $p$ -adic integers, i.e., irreducible fractions whose denominators are coprime to  $p$ , are indistinguishable by measurement from real numbers due to the inevitable nonzero measurement error. The paper is motivated by the ideas of I. Volovich on  $p$ -adic mathematical physics, cf., [6], by G. 't Hooft's cellular automaton interpretation of quantum mechanics, cf., [7], and (to some extent) by recent papers on superdeterminism by J. Hance, S. Hossenfelder, and T. Palmer, [3,5]. As a whole, the paper

is information-theoretic by nature, so the results of the paper concerning causality, wave functions, entropic time-energy uncertainty relation, etc., which are rigorously deduced in the paper, may be considered as a contribution to J. Wheeler's *it from bit* doctrine, cf., [14].

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Data Availability Statement:** Additional data can be found at <https://www.researchgate.net/profile/Vladimir-Anashin/research>.

**Conflicts of Interest:** This author declares no conflict of interest.

## References

- Anashin, V. Toward the (non-cellular) automata interpretation of quantum mechanics: Volovich postulates as a roadmap. *Int. J. Mod. Phys. A* **2022**, *37*, 2243003. [CrossRef]
- Hagar, A. *Discrete or Continuous?* Cambridge Univ. Press: Cambridge, UK, 2014.
- Hance, J.R.; Hossenfelder, S.; Palmer, T.N. Supermeasured: Violating Bell statistical independence. *Found. Phys.* **2022**, *52*. [CrossRef]
- Conway, J.H.; Kochen, S. The strong free will theorem. *Not. Am. Math. Soc.* **2009**, *56*, 226–232.
- Hossenfelder, S.; Palmer, T. Rethinking superdeterminism. *Front. Phys.* **2020**, *8*, 139. [CrossRef]
- Vladimirov, V.S.; Volovich, I.V.; Zelenov, E.I. *p-adic Analysis and Mathematical Physics*; World Scientific: Singapore, 1994.
- Hooft, G. The Cellular Automaton Interpretation of Quantum Mechanics. In *Fundamental Theories of Physics*; Springer: Berlin/Heidelberg, Germany, 2016; Volume 185.
- Cook, M. Universality in elementary cellular automata. *Complex Syst.* **2004**, *15*, 1–40.
- Rendell, P. Turing Machine Universality of the Game of Life. In *Emergence, Complexity and Computation*; Springer: Berlin/Heidelberg, Germany, 2016; Volume 18.
- Anashin, V.S. Quantization causes waves: Smooth finitely computable functions are affine. *p-Adic Numbers Ultrametric Anal. Appl.* **2015**, *7*, 169–227. [CrossRef]
- Lerner, E. The uniform distribution of sequences generated by iterated polynomials. *p-Adic Numbers Ultrametric Anal. Appl.* **2019**, *11*, 280–298. [CrossRef]
- Rényi, A. Representation for real numbers and their ergodic properties. *Acta Math. Acad. Sci. Hung.* **1957**, *8*, 477–493. [CrossRef]
- Parry, W. On the  $\beta$ -expansions of real numbers. *Acta Math. Acad. Sci. Hung.* **1960**, *11*, 401–416. [CrossRef]
- Wheeler, J.A. Information, physics, quantum: The search for links. In *Complexity, Entropy, and the Physics of Information*; Zurek, W.H., Ed.; Addison-Wesley Pub Co.: Redwood City, CA, USA, 1990; pp. 309–336.
- Schikhof, W.H. *Ultrametric Calculus*; Cambridge Univ. Press: Cambridge, UK, 1984.
- Grätzer, G. *Lattice Theory: Foundation*; Birkhäuser: Basel, Switzerland, 2011.
- Goldin, D.; Smolka, S.A.; Wegner, P. (Eds.) Chapter A Theory of System Interaction: Components, Interfaces, and Services. *Interactive Computation*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 41–96.
- Kalman, R.E.; Falb, P.L.; Arbib, M.A. *Topics in Mathematical System Theory*; McGraw-Hill: New York, NY, USA, 1969.
- Salomaa, A. *Theory of Automata*; Pergamon Press: Oxford, UK, 1969.
- Mahler, K. *p-adic Numbers and Their Functions*, 2nd ed.; Cambridge Univ. Press: Cambridge, UK, 1981.
- Lothaire, M. *Algebraic Combinatorics on Words*; Cambridge Univ. Press: Cambridge, UK, 2002.
- Grigorchuk, R.I.; Nekrashevich, V.V.; Sushchanskii, V.I. Automata, dynamical systems, and groups. *Proc. Steklov Institute Math.* **2000**, *231*, 128–203.
- Anashin, V.; Khrennikov, A. Applied Algebraic Dynamics. In *de Gruyter Expositions in Mathematics*; Walter de Gruyter GmbH & Co.: Berlin, Germany, 2009; Volume 49.
- Anashin, V. The  $p$ -adic theory of automata functions. In *Advances in Non-Archimedean Analysis and Applications*; STEAM-H: Science, Technology, Engineering, Agriculture, Mathematics & Health; Springer Nature: Berlin/Heidelberg, Germany, 2021; pp. 9–113.
- Hasselblatt, B.; Katok, A. (Eds.) *Handbook of Dynamical Systems*; Elsevier Science B.V.: Amsterdam, The Netherlands, 2002; Volume 1A.
- Downarowicz, T. *Entropy in Dynamical Systems*; Cambridge Univ. Press: Cambridge, UK, 2011.
- Tao, R. *Finite Automata and Application to Cryptography*; Tsinghua Univ. Press: Beijing, China; Springer: Berlin/Heidelberg, Germany, 2008.
- Durand, F.; Paccout, F. Minimal polynomial dynamics on the set of 3-adic integers. *Bull. Lond. Math. Soc.* **2009**, *41*, 302–314. [CrossRef]
- Hall, R.R. On pseudo-polynomials. *Arch. Math.* **1971**, *18*, 71–77. [CrossRef]
- De Bruijn, N.G. Some classes of integer-valued functions. *Nederl. Akad. Wetensch. Proc. Ser. A* **1955**, *58*, 363–367. [CrossRef]

31. Delaygue, É.; Rivoal, T. On primary pseudo-polynomials (around Ruzsa’s conjecture). *Int. J. Number Theory* **2022**, *18*, 1613–1636. [CrossRef]
32. Anashin, V.S. Uniformly distributed sequences of  $p$ -adic integers. *Math. Notes* **1994**, *55*, 109–133. [CrossRef]
33. Chlodovsky, M.I. Une remarque sur la représentation des fonctions continues par des polynômes à coefficients entiers. *Math. Sb.* **1925**, 472–474.
34. Ferguson, L.B.O. *Approximation by Polynomials with Integral Coefficients*; Amer. Math. Soc.: Providence, RI, USA, 1980.
35. Kolyada, S.F. Li-Yorke sensitivity and other concepts of chaos. *Ukr. Math. J.* **2004**, *56*, 1242–1257. [CrossRef]
36. Devaney, R.L. *An Introduction to Chaotic Dynamical Systems*; 2nd ed.; Addison-Wesley Publ. Co.: Boston, MA, USA, 1989.
37. Banks, J.; Brooks, J.; Cairns, G.; Davis, G.; Stacy, P. On Devaney’s definition of chaos. *Amer. Math. Month.* **1992**, *99*, 332–334. [CrossRef]
38. Assaf, D., IV; Gadbois, S. Definition of chaos. *Am. Math. Mon.* **1992**, *99*, 865. [CrossRef]
39. Knudsen, C. Chaos without nonperiodicity. *Amer. Math. Month.* **1994**, *101*, 563–565. [CrossRef]
40. Kolesov, A.Y.; Rozov, N.K. On the definition of chaos. *Russ. Math. Surv.* **2009**, *64*, 701–744. [CrossRef]
41. Downarowicz, T.; Lacroix, Y. Measure-theoretic chaos. *Ergod. Th. Dynam. Sys.* **2012**, *34*, 110–131. [CrossRef]
42. Halmos, P.R. *Lectures on Ergodic Theory*; Pub. Math. Soc. Japan. Kenkyusha: Tokyo, Japan, 1956.
43. Anashin, V. Ergodic transformations in the space of  $p$ -adic integers. In  *$p$ -adic Mathematical Physics, Proceedings of the 2-nd Int’l Conference, Belgrade, Serbia and Montenegro, 15–21 September 2005*; Khrennikov, A.Y., Rakić, Z., Volovich, I.V., Eds.; American Institute of Physics: Melville, NY, USA, 2006; Volume 826, pp. 3–24. [CrossRef]
44. Anashin, V.S. Uniformly distributed sequences of  $p$ -adic integers, II. *Discrete Math. Appl.* **2002**, *12*, 527–590. [CrossRef]
45. Green, J.W. Functions which assume rational values at rational points. *Duke Math. J.* **1939**, *5*, 164–171. [CrossRef]
46. Makhholm, H. Functions That Take Rationals to Rationals. Available online: <https://math.stackexchange.com/questions/167620/> (accessed on 8 May 2023).
47. Griffith, D.J. *Introduction to Quantum Mechanics*, 2nd ed.; Pearson Prentice Hall: Hoboken, NJ, USA, 2005.
48. Oppenheim, J.; Reznik, B.; Unruh, W.G. Temporal ordering in quantum mechanics. *J. Phys. A Math. Gen.* **2002**, *35*, 7641–7652. [CrossRef]
49. Anashin, V. The non-Archimedean theory of discrete systems. *Math. Comp. Sci.* **2012**, *6*, 375–393. [CrossRef]
50. Aliprantis, C.D.; Burkinshaw, O. *Principles of Real Analysis*, 3rd ed.; Academic Press, Inc.: Cambridge, MA, USA, 1998.
51. Mansurov, V. *Knot Theory*; Chapman & Hall: London, UK; CRC: Boca Raton, FL, USA, 2004.
52. Dubrovin, B.A.; Fomenko, A.T.; Novikov, S.P. *Modern Geometry—Methods and Applications*; Springer: Berlin/Heidelberg, Germany, 1985; Volume II.
53. Mishchenko, A.; Fomenko, A. *A Course of Differential Geometry and Topology*; Mir: Moscow, Russia, 1988.
54. Hasselblatt, B.; Katok, A. *A First Course in Dynamics*; Cambridge Univ. Press: Cambridge, UK, 2003.
55. Hieronymi, P.; Walsberg, E. A tetrachotomy for expansions of the real ordered additive group. *Sel. Math.* **2021**, *27*. [CrossRef]
56. Ushakov, N.G. *Selected Topics in Characteristic Functions*; DeGruyter: Berlin, Germany, 1999.
57. Weaver, G. König’s infinity lemma and Beth’s tree theorem. *Hist. Philos. Log.* **2017**, *38*, 48–56. [CrossRef]
58. Erdős, P.; Joó, I.; Komornik, V. Characterization of the unique expansions  $1 = \sum_{i=1}^{\infty} q^{-n_i}$  and related problems. *Bull. Soc. Math. Fr.* **1990**, *118*, 377–390. [CrossRef]
59. Coles et al. Entropic uncertainty relations and their applications. *Rev. Mod. Phys.* **2017**, *89*, 015002-1–015002-58. [CrossRef]
60. Ignatiev, A.Y. How fast is the wave function collapse? *J. Phys. Conf. Ser.* **2013**, *410*, 012153. [CrossRef]
61. Golomb, S.W.; Peile, R.E.; Scholtz, R.A. *Basic Concepts in Information Theory and Coding*; Springer: Berlin/Heidelberg, Germany, 1994.
62. Staiger, L. On maximal prefix codes. *Bull. EATCS* **2007**, 205–207.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.