Chapter 2
Paradigm for Chaos

One of the major concepts central to the deeper understanding of contemporary physics is the concept of chaos. It would not be an exaggeration to say that chaos is everywhere in physics. The chaotic behavior of physical systems was considered until lately as the result of unknown factors which influence the system. In other words, chaos was supposed to appear in physical systems because of the interactions with other systems. Earlier, as a rule, it was considered that these actions were complicated and uncontrolled and usually random. Random parameters and fields arose in dynamical systems due to these phenomena. As a result, the variables, which describe the dynamical systems, are random. The development of non-linear physics and the discovery of deterministic chaos led to an important change of point of view on both the apparition of chaos in physical systems and the nature of chaos.

Currently, there are many good books and reviews dealing with the theory and applications of deterministic chaos. In most descriptions of deterministic chaos a pragmatic point of view prevails: chaos appears in dynamical systems with trajectories utterly sensible to minor changes in initial conditions. At that, the individual trajectory is, as a rule, so complex that it is practically impossible to distinguish it from a chaotic one. At the same time, this trajectory is completely determined. This point of view, which is sufficient enough for practical work with non-linear dynamical systems, is the one most commonly used. However, the question of the deeper origins of deterministic chaos is rarely discussed. When and why is the behavior of a determined trajectory not only complex and "similar to random," but really random? In other words, can we apply probability laws to it, despite the fact that at the same time, it is quite determined and unique with the same initial conditions? It is clear that answers to these questions are of fundamental importance even if they will not contribute to perfecting techniques of practical calculations of chaos characteristics.

In this chapter we will state the foundations of the algorithmic theory of randomness of Kolmogorov–Martin-Löf which can provide a deeper understanding of the origins of deterministic chaos. The algorithmic theory of randomness does not

Y. Bolotin et al., *Chaos: Concepts, Control and Constructive Use*,
Understanding Complex Systems, DOI 10.1007/978-3-319-42496-5_2

deal directly with dynamical systems. Instead, this theory examines the work of the universal Turing machine. A Turing machine works according to some determined program and a prints finite or infinite sequence of symbols from a finite alphabet (for example, 0 or 1). We can consider the Turing machine an abstract model for a determined dynamical system or as a model of a computer programmed to solve a motion equation for a dynamical system. We can think of the sequence printed by this Turing machine as the trajectory of the dynamical system.

Now we can examine the same question again but this time within the frame of the mathematical theory of Turing machine: supposing that the sequence (trajectory) is complex, would it be in any sense random? The theory of Kolmogorov–Martin-Löf gives the answer to this question. Kolmogorov formalized the concept of complexity when analyzing the length of programs for the Turing machine. He introduced the concept of complexity, now called Kolmogorov's complexity. Kolmogorov's disciple, Martin-Löf, proved the remarkable theorem. Complex sequences, according to Kolmogorov, are random to the extent that they obey all the theorems of the theory of probability with an accuracy up to set of zero measure. This theorem is astonishing, because its proof concerns not only already-known theorems of probability theory, but also theorems which are not yet proven.

Thus, it was strictly proven that the complexity of determined sequences (trajectories), which is understood as the absence of laws, actually turns into true randomness. As a result, the theory of Kolmogorov–Martin-Löf, whose importance is probably not yet fully appreciated in physics, gives a new understanding of the origins of randomness and of deterministic chaos. This is applicable to individual objects without using statistical ensembles.

2.1 Order and Disorder

In order to discuss these concepts, it is natural to start with the most obvious ones. It seems normal that order is simpler than disorder. Let us imagine an experimenter who works with an instrument and who measures the value of some variables. If his instruments record the value $7, 7, 7, 7, 7, \ldots, 7$ multiple times, a rule becomes obvious and simple, under the condition that the experimenter is sure that it would continue in the same way. Other results can also appear, like $7, 2, 7, 2, \ldots, 7, 2$ or $7, 2, 3, 5, 7, 2, 3, 5, \ldots$, so the rule can be seen without any difficulty if the experimenter is sure to repeat the same results as before. However, there are situations when the rule is more complicated and its finding requires efforts which go beyond the scope of these simple observations. The reasoning above suggests that as a definition of ordered behavior or, in this case, ordered sequences of numbers, one can propose a seemingly simple definition. This naive definition means that we can predict all the terms of the sequences using its limited part only. But this definition is not very useful since it is practically impossible to guess the rules of construction for a complex sequence. For instance, if we took the sequence of the first thousand decimal digits belonging to the fractional part of number π, it would seem random. However, when we investigate the simple rule of its construction (a short program for a computer), we can no longer consider this sequence as being random. Actually, if we have a limited part of the sequence, we can imagine an endless number of rules which give different sequences where the limited part of the beginning is present. All this shows that this attempt to define the concept of order is not at all sufficient for understanding the concept of chaos. Therefore some strict mathematical methods are needed to distinguish well-ordered sequences from chaotic.

Let us now consider a different way of introducing these concepts of order and disorder. For the sake of simplicity, we will consider sequences containing only 0 and 1. Then the main question appears: how to distinguish ordered sequences of 0 and 1 from disordered ones. It goes without saying that the origins of these sequences are not important.

The first attempt to define random consequences using the frequency approach was made by Mises [1] who tried to formulate the essential logical principles of the probability theory. To begin with let us examine the Mises's scheme.

Let us suppose an infinite sequence $x_1, x_2, x_3, \ldots$ made of zeros and ones. According to Misses, above all, the necessary condition of randomness is to be fulfilled, i.e., the limit must exist:

$$P = \lim_{N\to\infty} \frac{1}{N} \sum_{i\le N} x_i \,. \tag{2.1}$$

It is clear that this condition is not sufficient, as, for example, the sequence $0, 1, 0, 1, 0, 1 \ldots$ obeys condition (2.1), but can in no way be considered random. Therefore, Mises believed that there is another condition for randomness. Let us choose from the infinite sequence of numbers $1, 2, 3, \ldots$ a subset of numbers and let us designate it as $n_1, n_2, n_3, \ldots, n_k, \ldots,$. Following this choice let us consider variables Mises' second idea was that the initial sequence of variables x_i is random if for the chosen subsequence $x_{n_1}, x_{n_2}, \ldots, x_{n_k}, \ldots$ the limit (2.1) remains the same, i.e.,

$$P = \lim_{M\to\infty} \frac{1}{M} \sum_{k\le M} x_{n_k} \,. \tag{2.2}$$

It is clear that the choice of the subsequence is not arbitrary. For instance, it is impossible to choose all the variables x_i as zeros or ones. That is the reason why the rules of the choice of subsequence are most important in Mises' theory. Mises gave only the general characteristic of these rules and restricted himself to some examples: in particular, prime numbers can be chosen as variable numbers, etc. But Mises could not formulate the mathematical scheme of choice or rules, since the concept of rules and laws of choice were not formulated mathematically in his time. In other words, the concepts of algorithm, recursive functions, and Turing

Machine, which formalized his intuitive ideas of laws and rules of choice were not yet developed.

The next studies of the foundation of the theory of randomness were stopped for a long time because of Kolmogorov's proposition to consider the probability theory as an applied measure theory [2]. The elegance of the axioms of Kolmogorov as well as its great possibilities led to the fact that the main efforts of scientists were concentrated on the development of the probability theory in this direction. The question of foundation was forgotten until the appearance of Kolmogorov's new work where he started to study this problem again. The starting point of his work was his introduction of the new concept of complexity as a measure of chaos. The complexity of finite sequences of 0 and 1, according to Kolmogorov, can be measured by the minimum length of its "description": the rules used to construct this sequence. A good example of sequence description is given in the book "The Good Soldier Švejk" by Jaroslav Hašek [3] when Švejk explains a method to remember the number of the railway engine 4268 which the track supervisor recommends to the engine-driver. "On track no. 16, is engine no. 4268. I know you have a bad memory for figures and if I write any figures down on paper you will lose the paper. Now listen carefully and as you're not good at figures I'll show you that it's very easy to remember any number you like. Look: the engine that you are to take to the depot in Lysa nad Labem is no. 4268. Now pay careful attention. The first figure is four, the second is two, which means that you have to remember 42. That's twice two. That means that in the order of the figures 4 comes first. 4 divided by 2 makes 2 and so again you've got next to each other 4 and 2. Now don't be afraid! What are twice 4? 8, isn't it? Well, then get it into your head that 8 is the last in the series of figures in 4268. And now when you've already got in your head that the first figure is 4, the second 2, and the fourth 8, all that's to be done is to be clever and remember the 6 which comes before 8. And that's frightfully simple. The first figure is 4, the second 2 and 4 and 2 are 6. So now you've got it: the second from the end is 6 and now we shall never forget the order of figures. You now have indelibly fixed in your mind the number 4268. But of course you can also reach the same result by an even simpler method. So he then began to explain to him the simpler method of how to remember the number of the engine 4268. 8 minus 2 is 6. And so now he already knew 68. 6 minus 2 is 4. So now he knew 4 and 68, and only the two had to be inserted, which made $4-2-6-8$. And it isn't too difficult either to do it still another way by means of multiplication and division. In this way the result will be the same."

As we can see, there are plenty of ways to describe a sequence. This is the reason why the main problem consists of how to find a method which would contain all the ways to describe the $0,1$ sequence, from which we need to pick the smallest one. The theory of algorithms by Turing and Post [4, 5] gives the foundations of the formal description of the construction rules for sequences. Their works laid the basis for many mathematical branches, such as mathematical logic, the theory of recursive functions, cybernetics, and the theory of information. Let us consider these works in more detail.

2.2 Algorithms and Turing Machine

In mathematics, an algorithm is the rule which permits one to find a solution to a mathematical problem (if a solution exists), using only regular procedures, without additional heuristic methods. The classical example is the Euclidian algorithm of division. The word algorithm comes from the name of the great Arabian mathematician Mohamed al-Horezmi, whose treaty in Latin begins with the words "Dixit algorizmi" which means "al-Horezmi said." Turing's reflections on the concept of algorithms led him to introduce a new mathematical concept, which is currently called the Turing machine. Nowadays, by definition, the Turing machine is the set:

$$M = \left(\Sigma, S, P, q_0, q_f, a_0\right) , \tag{2.3}$$

where Σ is an external alphabet, with which you can write down the input and output words (sets of letters which are contained in the external alphabet). S is an internal alphabet which describes the internal states of the Turing machine, q_0 is the initial state, q_f is the final state, a_0 is the empty cell, P is the machine program, i.e., the list of commands. As regarding the commands there are three kinds of words:

1.

$$qa \to rb$$

The meaning of this expression is the following. The Turing machine in the state of q and watching the letter a must pass to the state r and write down letter b on the band.

2.

$$qa \to rbR$$

This expression means that the machine in the state of q and watching the letter a must pass in state r, and write down the letter b and move to the right.

3.

$$qa \to rbL$$

This means that the machine in the state of q and watching the letter a must pass in the state of r, and write down the letter b and move to the left.

R, L, and $\to$ are not part of the alphabet Σ and S. By definition the program is a finite sequence of these commands. It is convenient to see the Turing machine as an endless band which is divided into separate cells, on one of which the Turing machine is fixed (see Fig. 2.1). On the band on each cell only one letter can be written from the alphabet Σ. As an example in Fig. 2.2, we present three commands

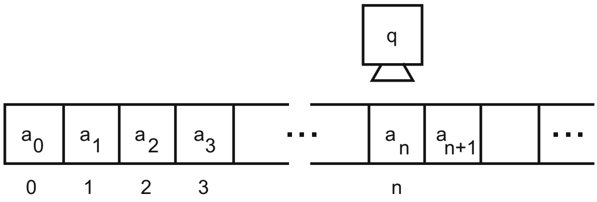

Fig. 2.1 Turing machine. The band of symbols, empty cells, and the reading head of the Turing machine are shown

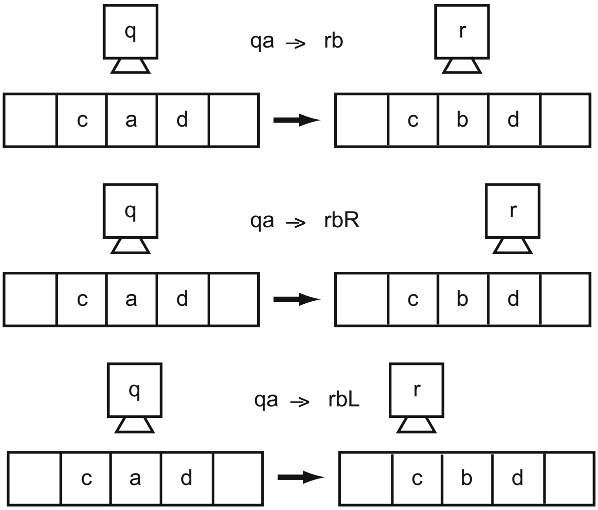

Fig. 2.2 Examples of the execution of commands by the Turing machine

performed by the Turing machine that were described earlier. The state of the Turing machine at any cycle is denoted as $A(q, a)B$. This means, that in the state of q, the machine is fixed on the letter a, on the band on the left of the letter a is the word A and on the right is written the word B.

Let us see now how the Turing machine works. We suppose that there is a machine word or a configuration of a word on the band. The Turing machine's work starts with the initial configuration $(q_0, a_0)B$. After the realization of program P the machine will stop at the final configuration $C(q_f, a_0)D$. The transition between the initial and the final configuration is performed by the command of the P program. Functions, which are calculated by the Turing machine, are called particular recursive functions (the word "particular" refers to the fact that the function is not defined at initial configuration). If the particular recursive function can be defined at any initial configuration it is called recursive. In the case of recursive functions the Turing machine starts to work with any input of integer numbers and always finishes its work giving the value of calculated functions.

The machine can be defined in the initial configuration if it performs its program and stops with the final result $C(q_f, a_0)D$. However, the initial configuration can be undefined for two reasons. First of all, in the process of program realization, the machine can find a configuration to which no commands of this program can be applied. The second reason is that the process of the execution of the program can be endless and the machine might not stop.

The Turing machine that is described gives an algorithm definition in a mathematical sense. It implies that there is an algorithm for a calculation or a process. The machine that is described has modest possibilities. However, it is possible to build a complex Turing machine thanks to the unification of simple ones. At the same time, the calculation possibilities of this machine will grow.

It is important to say that it has been demonstrated that a universal Turing machine can be built. This kind of machine can do whatever any Turing machine M does with an initial configuration. As an input for the universal Turing machine we give the initial configuration and description of Turing machine M. As a result, the initial configuration of treatment will be the same as for the machine M. We have to notice that this machine has impressive capacities. In principle, any modern computer can be coded on a band, and as a result, the universal Turing machine can do anything a modern computer does. The last question that interests us is whether all algorithms, from an intuitive point of view, coincide with the formal definition of the Turing machine. Generally speaking this question is not a mathematical one because there cannot be an algorithm definition in an intuitive sense. Church [6, 7] first answered this question when he proposed a thesis in which he said that every alphabetical, discreet, massive, determined, and closed algorithm can be defined by the Turing machine. We can say that any algorithm in an intuitive sense is given by the Turing machine. We have to emphasize that Church's thesis is not a mathematical assertion, it is more like a statement about energy conservation in physics. However, mathematical experience supports this thesis.

2.3 Complexity and Randomness

Now we have a universal method to describe finite sequences thanks to the Turing machine. Actually, any sequence can be associated with a program P of the Turing machine, thanks to which the Turing machine can write it down. It is clear that for every chosen sequence there are an endless number of programs that can perform it. This is why, according to Kolmogorov, we can define the concept of complexity [8] as related to the Turing machine M. Let us say that the machine M writes down n-value sequences of 0 and 1. By definition, the complexity K_M coincides with the length of the shortest program (in bytes) after the realization of which the machine will write down the given sequence $X = (x_1, x_2, \ldots, x_n)$:

$$K_M = \begin{cases} \min \ell\,(P)\,, & M\,(P) = X, \\ \infty, & M(P) \neq X \end{cases}.$$

This concept of complexity clearly depends on the machine M. However Kolmogorov managed to prove that there is a universal Turing machine for which:

$$K\,(X) \leq K_M\,(X) + C_M\,.$$

The constant C_M does not depend on the sequence and this is why it can be chosen identically for all the sequences. So the complexity K will be minimal and $K(X)$ will be called the complexity of the sequence X according to Kolmogorov. It is possible to prove that there are sequences X (with a n length) for which $K(x) \geq n$. This means that there are no simpler algorithms or ways to describe them than this sequence. Such sequences correspond to our intuitive understanding of random sequences because they do not contain any rules which could simplify them. Nevertheless, for the time being, we do not have any reason to think that probability laws are applicable to these sequences. Thus we have defined the concept of complex finite sequences or a random sequence. In a certain sense this definition can be considered as final. However, it is necessary to extend our definition to infinite sequences. It seems natural to define random infinite sequences of 0 and 1 $X = (x_1, x_2, x_3, \ldots)$ so that

$$K\,(X^n) \geq n + \text{const}\,. \tag{2.4}$$

for any final segment $X^n = (x_1, x_2, x_3, \ldots, x_n)$.

Here the constant depends on the sequence X. However, this definition is not satisfactory. One can prove that sequences for which conditions (2.4) are fulfilled for every n do not exist. We can intuitively understand the reason for these phenomena. Experience shows that in every random sequence, for example, the one you obtain after a coin toss, there are ordered parts of sequence numbers (for example, $1, 1, 1, \ldots, 1$). This means that in every random infinite sequence there is an endless number of ordered segments. Thus complexity behaves like an oscillation function with a growth of n (see Fig. 2.3). In other words, there are many values of n for which:

$$K\,(X^n) < n\,. \tag{2.5}$$

This means that definition (2.4) is not appropriate for random infinite sequences. In order to avoid these difficulties we chose another definition.

The sequence $X = (x_1, x_2, x_3, \ldots)$ is called random according to Kolmogorov if there is a constant C such that for each number n the following condition is satisfied:

$$|K\,(X^n) - n| \leq C\,. \tag{2.6}$$

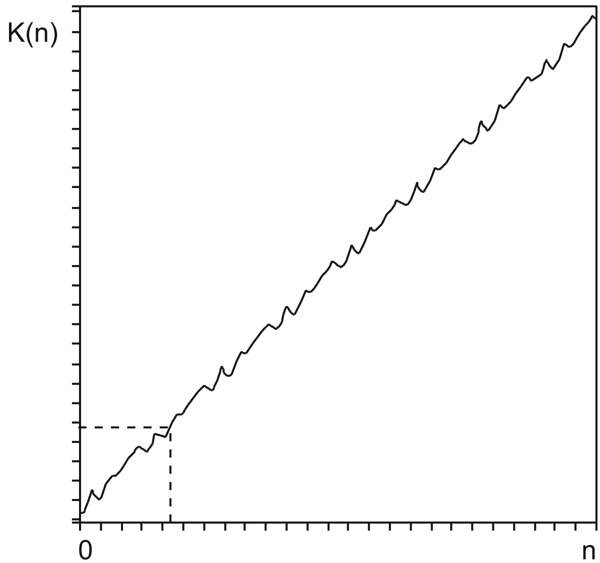

Fig. 2.3 The behavior of complexity with the growth of the number of sequence members. The features of the decrease of complexity on ordered segments of sequences are shown symbolically

From a physical point of view, this means that the decrease of complexity does not go over a certain level and with the growth of n the relative contribution of this decrease to complexity is small.

We will understand random infinite sequences as sequences that fulfill the conditions (2.6). According to Kolmogorov this is a final definition of individual random sequences. Another way to define random sequence is to use monotonous complexity or monotonous entropy (see, for example, [9]) instead of the simple Kolmogorov's complexity. The introduction of monotonous entropy permits us to avoid the difficulty of oscillation in the simple Kolmogorov's complexity.

It seems very natural to think that these sequences are chaotic. However, one question remains: will these sequences be random in the sense that they will obey all theorems of the probability theory? Martin-Löf obtained a positive answer to this question when developing Kolmogorov's ideas [10, 11]. It is not our aim to present the Martin-Löf theory, which is quite complex. However, because of the importance of his results, we are going to explain the main ideas of this theory. Let us consider a set Ω of infinite sequences of 0 and 1. It is clear that the power of set of all sequences has the capacity of a continuum. Let P be a measure for this set. How does the observer exclude all the sequences which have all possible laws from this set? The observer can treat the initial segments of the sequences, find some laws in them as, for example, repetitions, calculate how many bits of regularity he finds, and then exclude these sequences from the admissible set. As a result, the measure of admissible candidates for random sequences will fall. Since these actions are recursive, the observer can charge the Turing machine with these actions. This idea is the base of the "universal test of randomness" or universal sequential test of Martin-Löf. Martin-Löf's test P for randomness is a recursive function F (or the Turing machine) which treats finite sequences of 0 and 1 with a length of m and find how many bytes of regularity there are in it (roughly speaking the common segment). Then for every byte of regularity, recursive function F decreases by at least twice the set of admissible sequences ω and its measure P. The sequences which remain in the admissible set for any length m are considered a sequence which goes through the test P. The union of all P-tests gives a universal sequential test which is a limiting concept. The sifting of the sequences through the Martin-Löf test eliminates all the sequences which have any laws. As the number of regular sequences is much smaller than that of complex ones, as a result we have only complex sequences in the admissible set after sifting. Martin-Löf demonstrated that the complex sequences after Kolmogorov go through the universal test. We have to explain how the random sequences after Martin-Löf satisfy all the theorems of probability theory which can be tested effectively. Let us suppose that we have the sequence ω which does not satisfy one of the theorems of probability. In this case this theorem can be reformulated and added in the new P-test. Now this sequence does not go through the new P-test, i.e., does not go through the universal test and must be rejected as not random.

So it has been proved that random sequences exist. The power of the set of random sequences is continuum. Let us emphasize that now the Martin-Löf random sequences are random in the classical sense. These sequences obey the theorems of the probability theory.

We have managed to establish a relationship between Kolmogorov and Martin-Löf random sequences. These two sets coincide. Now we can affirm that complex sequences or random in the sense of Kolmogorov are random from the probability theory point of view. Let us present another important theorem. Random sequences cannot be calculated with the Turing machine. There are no algorithms for the calculation of random sequences. This is a very important property. In conclusion, let us note that Church formulated the intuitive ideas of Mises using the theory of algorithms [12]. The final theory which developed the frequency approach of the probability of Mises was proposed by Kolmogorov and Loveland [13–15]. Kolmogorov considered the frequency approach very important, as it explains why the abstract probability theory is applicable in the real world, where the number of trials is always finite[4]. Unfortunately the algorithmic notion of randomness according to Mises–Kolmogorov–Loveland does not correspond in full to our intuitive understanding [16]. The random sequence according to Martin-Löf is random according to Mises–Kolmogorov–Loveland. However there are sequences which are random according to Mises–Kolmogorov–Loveland, but are not random according to Martin-Löf.

2.4 Chaos in a Simple Dynamical System

Now let us look at the source of the appearance of chaos in determined systems. Before introducing the motion equation or system evolution we shall discuss its phase space. For the sake of simplicity, let us limit ourselves to one-dimensional models which evolve in the bounded area of the real R^1 line. Then, as a phase space, without the loss of commonness, we can choose the segment [0, 1]. All the points of this segment can be used as values of our system positions during evolution and also as its initial conditions. Now we are reminded of some of simple facts of the number theory. The segment [0, 1] is filled up with real numbers which are separated as rational numbers and irrational ones. The power of the set of rational numbers coincides with the power of integer ones, i.e., a countable set. The power of the set of irrational numbers is a continuum. For the description of numbers in the calculus system with the radix b (integer number) one uses their single valued representability like a series:

$$x = \frac{a_1}{b} + \frac{a_2}{b^2} + \frac{a_3}{b^3} + \cdots = \sum_{i=1}^{\infty} \frac{a_i}{b^i}$$

$$x \in [0, 1] .$$

where $a_i = 0, 1, 2, \ldots, b - 1$. Another expression to write down the numbers in the selected calculus system:

$$x = 0, a_1, a_2, a_3, \ldots$$

is named the b form of number presentation. This form is well known by everyone through the decimal form of writing rational numbers. One often uses binary calculus system in which the radix $b = 2$. For instance, in Babylon, the calculus system with the radix $= 60$ was used because ancient mathematicians did not like fractions. From this point of view, the relatively small number 60 with the big number divisor 12 was a very convenient foundation for the calculus system. As a result, we inherited the division of 60 min in an hour and 60 s in a minute from that Babylonian calculus system. Rational numbers written down in the b form can be easily distinguished from the irrational ones. Actually, any rational number is represented as:

$$\frac{p}{q} = 0, a_1, a_2 \cdots a_m a_1 a_2 \cdots a_m a_1 a_2 \cdots a_m \cdots \tag{2.7}$$

In other words, after m figures in the writing of any rational numbers the e is a block of figures fully repeated periodically. Irrational numbers do not have such periodicity. Obviously it does not help too much to recognize numbers. For example, it is hard to use this fact even to prove the irrationality of $\sqrt{2}$. To do this, one would have to have the infinite recording of this number but that is impossible. That is why the well-known proof of the irrationality of $\sqrt{2}$ obtained by the ancient Greeks was based on another principle.

In some cases the demonstration of rationality or irrationality of some concrete numbers can be very difficult. For example, we currently do not know if the Euler constant $c \approx 0.6$ is a rational or irrational number. By definition the Euler constant is given by the following expression:

$$c = \lim_{n\to\infty} \left(\sum_{i=2}^{n} \frac{1}{i} - \ln n \right) .$$

What is important for us is that rational or irrational numbers are dense everywhere on a segment $[0, 1]$ (see, for example, [17]) and that between two different real numbers there is an infinite number of rational and irrational numbers. Let us note that it is not important if these two numbers are rational or not. The theorem is always true. Hence in the small neighborhood of any number there is an infinite number of rational and irrational numbers. The fact that rational numbers are dense everywhere permits us to have a good approximation of irrational numbers by the rational ones. It has been proven that any irrational number x can be approximated by rational numbers p/q with precision so that:

$$\left| x - \frac{p}{q} \right| < \frac{1}{q^2} .$$

As an example we give an approximation of the irrational number π, $\pi \approx 355/113$. It is easy to verify that:

$$\left| \pi - \frac{355}{113} \right| = 2,66 \cdot 10^{-7} .$$

This means that this fraction coincides with the number π up to the sixth order and corresponds to the inequality that was given before.

Let us transfer some properties of chaotic sequences onto the points of our phase space. It is easy to understand that if we attach 0 with comma (0,) to any sequence, we obtain one-to-one correspondence between the sequences and the point of our phase space. At the same time, the coordinates of our phase space are written down in the calculus system on radix 2. Since the continuum of points exists and there is no algorithm to calculate their coordinates, the coordinates of these points are random.

One might think that these points correspond to all irrational numbers. However, this is not true. For example, the number e is irrational but is not complex according to Kolmogorov, since there is a simple algorithm to calculate it with the expression:

$$e = \lim_{n\to\infty} \left(1 + \frac{1}{n} \right)^n .$$

This example shows that not all the sequences which at first sight are complex or chaotic are such in reality, because they can contain some hidden algorithms which were used for their construction. Now let us consider the non-linear dynamical system in discrete time:

$$x_{n+1} = 2x_n \mod 1\,. \tag{2.8}$$

Here x mod 1 is a fractional part of the number x. If we know the coordinates x_n at the moment n, it is easy to have its coordinates x_{n+1} at the moment $n+1$. If we substitute x_{n+1} in the right part, we find the same way the coordinates of the system at the moment $n+2$. It is hard to imagine a simpler and more determined system. The exact solution of this equation:

$$x_n = 2^n x_0 \mod 1 \tag{2.9}$$

gives a guarantee of the existence and uniqueness of the solution. Thus, the mapping (2.9) has all the features of a strictly determined system. Let us consider one of the trajectories of the system with initial conditions X which belong to the set of "random" points. The dynamic of the system (2.8) means in reality the shift of the comma of one position to the right and the rejection of the integer part of the number at each step. This is why the whole trajectory is actually number x, which is random, hence the trajectory of the motion of the system can be shown to be random.

This example shows that determination is not in contradiction with randomness. Mapping (2.8) is determined and does not contain any random parameters. Moreover, it has an exact solution (2.9). The trajectories with the same initial conditions repeat exactly. However, the behavior of the system, or of the trajectory, is random. In this sense we can speak about a deterministic chaos.

There is a simple way to test it. Let us divide the segment [0, 1] into two segments [0, 1/2] and [1/2, 1]. Now the question is to know in which segment the solution is. The answer to this question depends on whether 0 or 1 is in the corresponding place in the presentation of number x. At the same time, an expert to whom we can present the data about particle position in the first or second segment will not be able to find any difference between this data and the data of a coin toss when 0 or 1 is associated with heads and the other figure is associated with tails. In both cases, he will find that the probability to find the particle in the left segment will be 1/2—the same as for the right segment. In this regard, we can say that the dynamic system (2.8) is the model of coin toss and describes the classical example of the probability process.

Such indeed is the meaning implied when one speaks of continuous phase space as the reason for the chaotic behavior of the system. One might think that this result appeared after a too simple partition of the phase space in cells. However, this is not true. It is possible to partition the segment [0, 1] into more cells and study the transition between them (here we enter into another mathematical branch, the so-called symbolic dynamics [18–20]). The transitions between these cells are described by the Markov process, which are classical examples of probability processes [21, 22]. We can even work with infinitely small cells but if we do not take into account some mathematical difficulties we will not yield anything more than randomness x which has already been proven.

Let us go back to the reasons for chaotic behavior of the trajectory of dynamic systems. First of all it is the continuality of phase space. However, if the chaotic behavior of the system is related to the uncalculated initial data, then why do we observe chaotic behavior in one system but not in others? We can try to answer this question. As a matter of fact if we look attentively at the system (2.8) we can remark that the behavior of this system in time depends more and more on the distant figures of the development of the number x_0. So if we know only a finite number of symbols in the development of the initial data x_0, for example, m, we can describe our system only on a finite interval m. As a result, the dynamical system is sensitive in an exponential way with respect to the uncertainty of the initial data. In such dynamical systems, the potential randomness which is contained in the phase space continuity transforms itself into an actual randomness. At the same time we obtain the first criterion of stochastic behavior of a determined system: stochastic behavior of the trajectory is possible in systems which have an exponential sensitivity to the uncertainty of initial data. This criterion can be presented differently. Chaotic behavior is achieved in non-linear systems with an exponential divergence in neighboring trajectories. We can easily understand this if we consider uncertainty as a module of the difference between the two possible initial conditions. For our system, the distance $|\Delta x_0|$ between trajectories which are close at the initial time will grow with the time as $|\Delta x_0|\, e^{n \ln 2}$.

Another important observation can be made from the study of this simple dynamical system, concerning periodical trajectories or orbits. Let us consider the positions of the periodical trajectories in the phase space. Taking into account the fact that all rational numbers have the form (2.7) it is easy to see that all trajectories with initial conditions x_0, which coincide with the rational numbers, will be periodical. Hence, the periodical orbits are a countable set and are dense everywhere in the phase space. Obviously, the trajectories which were initially close to the periodical orbits will go far away exponentially fast. This is why we call these periodical orbits unstable. Thus, periodical orbits are everywhere dense in the phase space of the dynamical system (2.8). As we will see later, this feature will always be observed in the dynamical systems with chaotic behavior.

Let us pay attention to one important property. If we choose a small neighborhood of initial conditions ω and launch the trajectories out of their neighborhood so, at the moment n, ω will occupy a certain neighborhood ω_n. Our system has the following property, which is easy to test: for any neighborhood ω one can find the time n when we have $\omega_n \cap \omega \neq \varnothing$. Dynamic systems which have this property are called transitive.

Our simplest system which shows the chaotic behavior is transitive. Let us note that the choice of a one-dimensional system is not important. All these properties are exactly applicable in multidimensional systems. For example, a bi-dimensional system $(x_1, \ldots, x_n, \ldots)$ and $(y_1, \ldots, y_n, \ldots)$ can be easily reduced to a one-dimensional case if we write down these sequences in the form of one sequence $(x_1, y_1, x_2, y_2, \ldots, x_n, y_n, \ldots)$. All these qualitative features are similar in the multidimensional case.

References

1. Mises, R.V.: Math. Z. **5**, 52–99 (1919)
2. Kolmogorov, A.N.: Foundations of Probability. Chelsea Publishing Company, New York (1950)
3. Hašek, J.: The Good Soldier Švejk. Penguin Books, New York (1974)
4. Turing, A.M.: Proc. Lond. Math. Soc. (Ser. 2) **42**, 230–265 (1936)
5. Post, E.L.: J. Symb. Log. **1**, 103–105 (1936)
6. Church, A.: Am. J. Math. **58**, 345–363 (1936)
7. Church, A.: Introduction to Mathematical Logic. Princeton University Press, Princeton (1956)
8. Kolmogorov, A.N.: Probl. Inf. Transm. **3**, 3–7 (1969)
9. Zvonkin, A.K., Levin, L.A.: Usp. Math. Nauk **25**(6), 85–127 (1970)
10. Martin-Löf, P.: Inf. Control **9**, 602–619 (1966)
11. Martin-Löf, P.: Algorithms and Random Sequences. Erlangen University Press, Erlangen (1966)
12. Church, A.: Bull. Am. Math. Soc. **46**(N2), 130 (1940)
13. Kolmogorov, A.N.: Sankhya A **25**(N4), 369 (1963)
14. Loveland, D.: Ztrschr. Math. Log. und Grundl. Math. Bd. **12**(4), 279 (1966)
15. Loveland, D.: Trans. Am. Math. Soc. **125**(N3), 497 (1966)
16. Schen, A.: Semiotics and Information, vol. 18, p. 14. VINITI, Moskva (1982) (in Russian)
17. Brannan, D.A.: A First Course in Mathematical Analysis, p. 459. Cambridge University Press, Cambridge (2006)
18. Dieudonné, J.: Foundations of Modern Analysis. Academic, New York/London (1960)
19. Alekseyev, V.M.: Symbolic Dynamics. 11th Summer Mathematical School. Mathematics Institute of the Ukrainian Academy of Science, Kiev (1976) (in Russian)
20. Bowen, R.: Lecture Notes in Mathematics, vol. 470. Springer, Berlin/Heidelberg/New York (1975)
21. Feller, W.: An Introduction to Probability Theory and Its Applications, vol. I. Wiley, New York (1957)
22. Dynkin, E.B.: Markov Processes. Academic, New York (1965)