ISSN 0005–1179, Automation and Remote Control, 2007, Vol. 68, No. 4, p. 1. © Pleiades Publishing, Ltd., 2007.
Original Russian Text © V.S. Kozyakin, N.A. Kuznetsov, 2007, published in Informatsionnye Protsessy, 2007, No. 4, pp. 327–363.

INFORMATION PROCESSES

Feasibility of Numerical Modelling: Information Aspect[1]

V. S. Kozyakin, N. A. Kuznetsov

Institute for Information Transmission Problems, Moscow, Russia

Received December 27, 2007

Abstract—In the paper some problems related to computer studies of continuous objects are considered. Examples are presented which demonstrate that even in trivial situations results of computer modelling can differ drastically from properties of original continuous objects. This observation is aggravated by the fact that often the situation cannot be improved by raise of accuracy of approximating of considered continuous objects. The mathematical models are offered which allow to a certain extent to explain originating phenomena and also to use and/or parry their negative aftereffects.

PACS number 02.30.Yy
DOI: 10.1134/S0005117906040171

1. INTRODUCTION

The traditional approach to numerical methods is based on the fundamental hypothesis that *any required accuracy of solution of a problem and adequacy of discrete model to its continuous analogue can be achieved by means of appropriate precision of computer arithmetic*, which is often supplemented by the following important remark: *insufficient accuracy of computer arithmetic can be compensated by extending computer memory and time of calculation.*

These hypotheses are strongly supported by many rigorous theoretical methods of mathematics and theoretical physics and many heuristic and empirical tools. The collection of examples which support these hypotheses is too extensive to be detailed. Accordingly, the following scheme of computer modelling became now typical:

- selection of a continuous model;
- computer realization of the selected model;
- interpretation of the numerical results, their application in practice and for the further investigation.

Nevertheless, recently computer modelling is often applied just for the analysis of those problems *where the reliable theoretical methods of understanding the relationships between properties of the continuous objects and their discretized counterparts are not applicable and, at the same time, empirically justified tools are not yet adequate.* For example, for investigation of systems with chaotic behavior (calculation of phase portraits of the Lorentz attractor and other chaotic systems, various plots with fractal-like structures), for providing the so-called "computer-aided" proofs and so on. In such situations there only remains as an article of faith that results of computer modelling are rather robust with respect to accuracy of calculation and increasing time of calculation, that is in the validity of the hypothesis formulated above. Unfortunately *such hopes are sometimes deceptive.* More than that it seems that the situations when the hypothesis is wrong are typical

[1] This work was supported by the Russian Foundation for Basic Research, projects nos. 06-01-00256, 06-01-72552-NCNIL-a.

in some areas. It is paradoxically enough, but the development of the computers technology and research methods did not reduce but, on the contrary, enforced the acuteness of the following principal question:

TO WHICH DEGREE WE CAN RELY UPON THE RESULTS OF COMPUTER MODELLING?

There arise daily some new unexpected and shocking examples, which demonstrate that even in seemingly trivial situations the results of computer modelling can be rather far from the theoretical predictions. Even that the cause of the discrepancy can be pinned down in each particular case, this situation indicates that "something is wrong in the Danish kingdom". In our view, the root of the troubles is that majority of the users of computers are bounded by the ideology of *continuous mathematics*. It seems that, as it was in physics, to grasp the meaning of the relations between continuous mathematical objects and their computer analogous, the sort of ideology of *quantum* mathematics is necessary, when the object will be considered as a *continuous-discrete* object, not only as a continuous one or only as a discrete one.

The concepts of proximity and continuity are the corner-stones of continuous mathematics. Usually these notions are used rather formal in computer modelling. For example, a computer realization is considered to be close to its continuous original, if it had been created through discretization on a fine enough lattice. The main target of a computer experiment — to get an *information* about the modelling object — is often lost of sight. Meanwhile, one should keep in mind the following point: while a transition from one continuous model to another is often feasible without any leakage of information (homeomorphic change of variables, etc.),

THE TRANSITION FROM A CONTINUOUS TO A DISCRETE MODEL IS USUALLY IMPOSSIBLE WITHOUT LOOSING OF INFORMATION.

A trivial example: discretization of an invertible linear system on a uniform lattice is as a rule a non-invertible mapping. Another example: the main information characteristic of a dynamical system, its entropy [56], measures the exponential rate of increasing of the quantity of distinct trajectories against length of trajectories of the system under consideration. But for any single-valued spatial discretized model of the system only a finite set of infinite trajectories is possible and the definition of entropy is getting meaningless. Here the conflict is apparent and the necessity of indirect methods to estimate entropy of continuous system in terms of its discretizations is evident. Note, that also some powerful methods have been created, the problem is rather difficult up to now. Often conflict is more disguised, but not less dangerous. We believe, that the first question of computer experiment should be the following

WHAT KIND OF INFORMATION ABOUT UNDERLYING CONTINUOUS SYSTEM IS PROBABLY/INEVITABLY LOST IN A CHOSEN SCHEME OF COMPUTER REALIZATION?

The second fundamental question is connected with such notions of continuous mathematics as robustness and structural stability. In continuous modelling it is actually the same as to understand how robust are properties of the object with respect to continuous, smooth etc. disturbances. As soon as we realized that the main problem in computer modelling is the problem about information, than we should ask the question:

CAN WE GUARANTEE INFORMATION ROBUSTNESS OF DISCRETIZATION SCHEMES?

It seems likely, that detailed analysis of these two questions and other related problems will be one of the strategic direction in the development of the exact sciences. To give the detailed

description of the situation in this field is hopeless and even more hopeless is to forecast its future development. Some topics related to the problem of discretization of continuous mappings are discussed in [7,41] where one can also find an extensive bibliography. We are going only to consider some typical troubles, which arise in this area, and to discuss some perspective, in our opinion, general ideas on a few examples. The chosen examples are connected with the very fact of embedding a continuous system in finite computer arithmetic rather than with inevitable calculating errors.

The outline of the paper is as follows. In Introduction we have tried to explain necessity of the "cautious attitude" to results of computer simulation of continuous systems.

In the second part of the paper (Sections 2.1, 2.2 and 2.3) three situations are discussed in which computer modelling of simple continuous models results in unexpected effects. They are the collapsing effects in discretization of dynamical systems, influence of discretizations on behavior of plane rotations and aliasing effects during computer visualization of simple two-dimensional images. These demonstrate also that arising effects reflect, maybe in a distorted way, properties of original continuous models. We have chosen these examples by the following two reasons. At first, they may by easily described and are rather spectacular, secondly, they are a good polygon for development and check of the common methods of qualitative understanding and the quantitative investigation of the discretization phenomena.

In Sections 3 and 4 the phenomenological models allowing precisely enough to compute statistical characteristics of the collapsing effect, so as lengths of cycles and transients of discretizations of dynamic systems are offered. At the heart of these models the theory of random maps with one absorbing centre lays (Sections 3.1 and 4.1). As one of approaches to compensate negative aftereffects of the collapsing effect various strategies based on application of the theory of stochastic interval matrices are considered (Section 3.3).

Section 4.3 is devoted to discussion of the common principles to which discretization of a continuous dynamic system should fulfil in order to its behavior reflect behavior of the original continuous system. Here concepts of consistency of a discretization, and also the concept of shadowing and mutual shadowing for a class of mappings are considered (Section 4.4). Application of the principle of mutual shadowing is illustrated on the example of the so-called semihyperbolic Lipschitzian maps (Section 4.5).

Section 5 is devoted to the models which explain, to a certain extent, effects arising during discretization of the plain rotations considered in Section 2.2. Here the problem of invertibility of discretizations of the linear maps of the plane is discussed (Section 5.1). It is noted, that generally discretizations of the linear maps of the plane are not invertible, and namely this is one of reasons of information losses under discretization of multidimensional systems. To describe the frequency properties of discretizations of the linear maps and to get estimates of "gaps" not filled by the trajectories of discretized planar rotation maps, the approach based on the analysis of the quasiperiodic frequency measurable sets, properties of the so-called "quantizers" of the linear maps and ergodic properties of discretizations of the linear maps is developed (Sections 5.2, 5.3 and 5.4).

In Section 6 we undertake an attempt to develop general approaches to estimate information losses at a discretization of stochastic processes.

The paper is completed by Section 7 in which the general structure of moirés is investigated (Section 7.1) and the problem how irregular may be the picture generated by a given function and how to characterize numerically this irregularity is considered (Section 7.2). To evaluate the extent of this regularity, the concept of aliasing dimension is introduced and some estimates of aliasing dimension are obtained.

Last years a substantial proportion of all computer experiments is carrying out in networks when a lot of processors work asynchronously. This leads to a superposition of discretization effects and desynchronization effects. These effects are very important and recently to their analysis the

increasing attention is paid. We did not discuss these effects partly because of restricted volume of this text, and partly because these question were carefully considered in recent journal papers and monographs, see bibliography in [4, 44].

Emphasize once more, that our purpose was not to sum recognized methods in the considered field, but to discuss rather new phenomena and approaches to their description, application and/or suppressing. Naturally we will base mainly on works which have been carried out during last 15 years in our small non-formal international team. The driving force of this team are Ph. Diamond (Australia), P Kloeden (Germany), V. Kozyakin (Russia), A. Pokrovskii (Russia, Ireland), A. Vladimirov (Russia), I. Vladimirov (Russia, Australia) and others. We would like also to thank M. Blank for discussions of discretizations of rotations of the plane (see Section 2.2), A. Vladimirov, who provided the first version of the Section 2.3, and M. Suzuki (Australia, Japan) for many numerical experiments In preparation of this paper we used surveys [48, 50] supplementing them by recent results.

One of incentive motives of writing this work was also the fact that overwhelming majority of publications on the given subjects has been published in foreign magazines and until recently in Russian-speaking magazines we did not meet similar works.

2. EXAMPLES

In this Section we give some typical examples in which the discordance between properties of continuous objects and their discretizations are demonstrated. These examples are just illustrative. They are simplified to the extent which makes possible, on the one hand, to proceed the thorough theoretical analysis of underlying continuous mappings and, on the other hand, to carry out the necessary experimental calculations, with a view of matching of the gained results.

Due to finiteness of internal computer arithmetic the natural computer representations of dynamical systems can be interpreted as mappings of finite sets into itself, that is by directed graphs. Nevertheless, it should be emphasized that the classical graph theory methods are insufficient to answer most interesting theoretically and important practically questions which arise in analysis of computer representations of continuous dynamical systems. This is explained by the fact that properties of procedures of a discretization of continuous systems are tightly linked to geometry and topology of (continuous) phase space of a prototype system which is difficult enough to express in terms of classical graph theory. Besides, as a rule, an individual discretizations appear to be not very informative due to their high sensitivity to concrete choice of a procedure of discretization (to a disposition of nodes of a lattice, its size, etc.). At the same time the averaged (or statistical) characteristics of corresponding discretizations with respect to "sizable" enough families of discretizations are turned out to be rather robust and reflect properties of a prototype system (though often and by not clear in some cases reasons in a distorted form). Therefore often the object of the computer modelling should be not an individual discretization of some continuous system but a family of discretizations generated by this system.

2.1. Collapsing effects in discretization of dynamical systems

We commence with the Feigenbaum–Cvitanovich mapping

$$F^{(\gamma)}(x) = 1 - |2x-1|^{1/\gamma}, \qquad x \in [0,1], \quad \gamma \le 1, \tag{1}$$

playing an important role in theoretical physics. When $\gamma = 1/2$ this mapping is called also the *logistic mapping* (see Fig. 1), when $\gamma = 1$ this mapping is called the "baker's mapping" (see Fig. 2).

The "baker's mapping" has the feature that for it the variability of the continuous system's behavior can "degenerate" during discretization into trivial and not interesting behavior. So, this

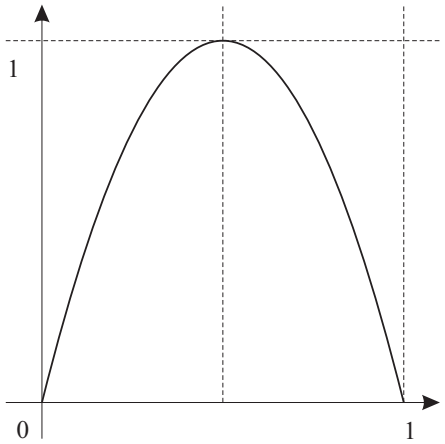

Fig. 1. Plot of the logistic mapping.

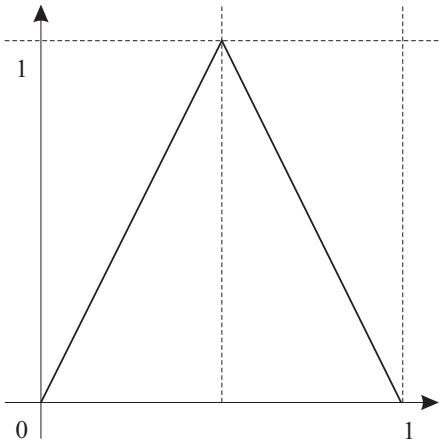

Fig. 2. Plot of the "baker's mapping".

mapping has the cycles of all periods, the Lebesgue measure is invariant and ergodic with respect to $F \equiv F^{(1)}$, it is mixing etc. Consider an ideal computer model of this mapping: its restriction F_ν on the ν-digit binary lattice

$$\mathbb{L}_{2^\nu} = \left\{0, \frac{1}{2^\nu}, \frac{2}{2^\nu}, \frac{3}{2^\nu}, \ldots, \frac{2^\nu - 1}{2^\nu}\right\}.$$

This restriction is asymptotically trivial: $F_\nu^m \equiv 0$ for $m \geq \nu$ (here $F^m(\cdot)$ denotes an operator power of the mapping $F(\cdot)$). Naturally, F_ν has only zero cycle; only measure concentrated in zero is invariant for F_ν etc. In this case the cause of crash is clear: every next iteration "kills" the last not–zero digit, i.e., "kills" information.

In general case the mapping $F^{(\gamma)}(x)$ so as the "baker's mapping" has unique absolutely continuous invariant measure $\mu_{F^{(\gamma)}}$. Therefore for almost all initial values $x_0 \in [0,1]$ the sequence

$$x_0, F^{(\gamma)}(x_0), \ldots, \left(F^{(\gamma)}\right)^n(x_0), \ldots \tag{2}$$

is spreaded in the natural sense over the whole interval $[0,1]$ with the density μ_G.

Denote by $\mathbb{L}_\nu$ the uniform $1/\nu$–lattice on the interval $[0,1]$ and let

$$[x]_\nu = \frac{\text{ROUND}(\nu x)}{\nu}, \qquad x \in [0,1],$$

where $\text{ROUND}(\cdot)$ is the standard roundoff operator. Denote by $F_\nu^{(\gamma)} : \mathbb{L}_\nu \mapsto \mathbb{L}_\nu$ the $\mathbb{L}_\nu$-*discretizations* [66] of $F^{(\gamma)}$ defined by

$$F_\nu^{(\gamma)}(\xi) = [F^{(\gamma)}(\xi)]_\nu, \qquad \xi \in \mathbb{L}_\nu.$$

The point $\xi \in \mathbb{L}_\nu$ is called ν-*collapsing* for the mapping f if $f_\nu^n(\xi) = 0$ for some (and therefore for all subsequent) ν. Denote by $Y(\nu; f)$ the set of ν-collapsing points. Denote by $P(\nu; f)$ the proportion of collapsing points of the lattice $\mathbb{L}_\nu$:

$$P(\nu; f) = \frac{\#(Y(\nu; f))}{\nu}. \tag{3}$$

What kind of properties it is natural to expect from the discretization $F_\nu^{(\gamma)}(\xi)$ of the mapping (1)? As is known, the mapping $F^{(\gamma)}(x)$ has cycles of arbitrary large periods, it is mixing, it

has unique absolutely continuous invariant measure and so on. Hence, the typical sequence $\{x_n\}$ satisfying

$$x_{n+1} = F^{(\gamma)}(x_n), \qquad n = 0, 1, \ldots, \tag{4}$$

is distributed with some density μ_γ over the whole interval $[0, 1]$. Then this sequence sometimes will "come" close to the zero point spending, nevertheless, the main amount of time rather far from it. Therefore it looks natural to suppose that the main bulk of trajectories of a sufficiently fine discretization of the mapping (4) would behave similar. That is, the proportion $\mathcal{P}(F_\nu^{(\gamma)})$ of collapsing elements of the mapping $F_\nu^{(\gamma)}$ *should be small* for large ν. On the other hand, for a typical trajectory $\{x_n\}$ of the mapping (4) there exists an index n such that $|x_n| < \frac{1}{2\nu}$. At the same time, all the points $x \in [0, 1]$ satisfying the inequality $|x| < \frac{1}{2\nu}$ are "stuck" with the zero fixed point under discretization on the lattice $\mathbb{L}_\nu$. Therefore it seems reasonable to expect that the main bulk of the points of the lattice $\mathbb{L}_\nu$ should be collapsing for the mapping $F_\nu^{(\gamma)}$, that is the proportion $\mathcal{P}(F_\nu^{(\gamma)})$ *should be close to* 1.

The two considered possibilities are incompatible and there arises a question which one of the two heuristics above is the reliable one. Surprisingly, the experimental calculations demonstrate that the proportions $\mathcal{P}(F_\nu^{(\gamma)})$ behave quite different for different γ and ν. For $\frac{1}{2} \le \gamma < 1$ under appropriate conditions can happen either the first or the second possibility whereas for $\gamma < \frac{1}{2}$ *both the first and the second possibilities are wrong!* In the latter case the mean value $p(\gamma)$ for the proportion of collapsing elements of the discretized mappings $F_\nu^{(\gamma)}$ (with respect to ν) varies from 0 at $\gamma \to \frac{1}{2}$ up to 1 at $\gamma \to 0$. For instance, $p(\frac{1}{3}) \approx 0.6$. Hence, the properties of a typical trajectory (4) of discretized system can be quite different from the predictions of both heuristics above and this conclusion remains unchanged for arbitrary fine lattices $\mathbb{L}_\nu$, or for floating point discretizations.

Below $\widetilde{P}(\nu; f)$ denotes the proportion of collapsing elements of the mapping f in a simple sampling consisting of 100 elements $\xi \in \mathbb{L}_\nu$. The value $\widetilde{P}(\nu; f)$ is statistically close to $P(\nu; f)$. Figure. 3 graphs $\widetilde{P}(\nu+n; F^{(\gamma)})$ for $\gamma = 1/3$, $\nu = 2^{27}$, $n = 1, 2, \ldots, 500$. This graph is rather typical. It is easy to notice that it represents a function with an irregular character and rather large mean value. Thus the properties of the typical sequence of iterations of the discretized system are quite different from those of a typical sequence (2). Perhaps we have chosen too rough lattice? Consider instead of the lattice $\mathbb{L}_{2^{27}}$ the lattice $\mathbb{L}_{2^{44}}$. How the corresponding graph will change? And here is the first surprise: seemingly the graph did not change at all, see Figure 4.

Let us forget about the artificial discretization of the mapping $F^{(\gamma)}$ and let us calculate on the computer the sequence (2) for random x_0 using the standard floating point operations. The result will support the observation above. Very often, in more than 90% experiments depending of the specific program used, after a few thousands of iterations the sequence (2) collapses to zero. In other rare cases the elements of the sequence are not zero for all n.

So, a conflict between behavior of underlying continuous system and behavior of its discretizations can not be explained by the insufficiently small step of discretization or a choice of specific discretization procedure. It seems, the root of this conflict is *in the very fact of discretization*. To resolve this conflict we should develop methods for detailed qualitative and quantitative investigation of combinatorics of discretizations of dynamical systems.

2.2. "Saturn rings": discretizations of rotations of the plane

Consider the problem of the numerical modelling of the phase portraits of some mapping. Choose as an example one of the most simple and well investigated mappings, a rotation of the plane on

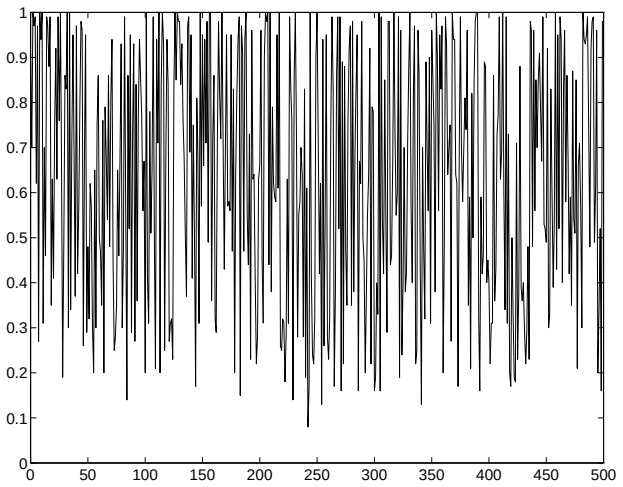

Fig. 3. Proportion of collapsing elements $\tilde{P}(\nu + n; F^{(\gamma)})$ for $\nu = 2^{27}$, $1 \le n \le 500$ for the mapping $F^{(\gamma)}$ with $\gamma = 1/3$.

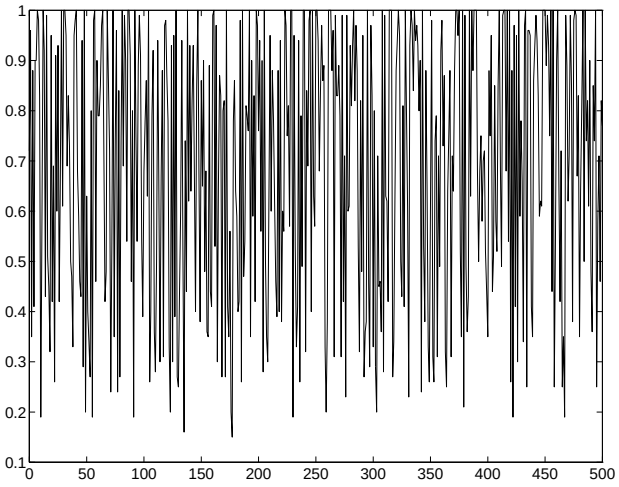

Fig. 4. Proportion of collapsing elements $\tilde{P}(\nu + n; F^{(\gamma)})$ for $\nu = 2^{44}$, $1 \le n \le 500$ for the mapping $F^{(\gamma)}(x)$ with $\gamma = 1/3$.

an angle ϑ around some point (x_0, y_0):

$$\begin{aligned} x &\mapsto x_0 + (x - x_0)\cos\vartheta - (y - y_0)\sin\vartheta, \\ y &\mapsto y_0 + (x - x_0)\sin\vartheta + (y - y_0)\cos\vartheta. \end{aligned} \tag{5}$$

This mapping is, on one hand, rather "bad" for numerical investigation because it is not structurally stable in a reasonable sense. On the other hand, this is a typical example of the mappings with an invariant measure, which arise, for example, in conservative mechanics. The phase portrait of the mapping (5) is quite simple — all its trajectories are situated on the concentric circumferences with their centers in the point (x_0, y_0), filling in the corresponding circles dense for ϑ such that ϑ/π is irrational and creating the finite set of points otherwise.

For the purpose of numeric modelling, consider the mapping (5) on the square lattice $\mathbb{L}_{1,1}$ with a step $h = 1$ and let us try to construct the trajectories with the initial points, which are equidistant from the point (x_0, y_0) with the step Δ.

In particular, consider two procedures of the numeric calculation for the modelling of the mapping (5):

$$\begin{aligned} x &\mapsto \textsc{trunc}(x_0 + (x - x_0)\cos\vartheta - (y - y_0)\sin\vartheta), \\ y &\mapsto \textsc{trunc}(y_0 + (x - x_0)\sin\vartheta + (y - y_0)\cos\vartheta) \end{aligned} \tag{6}$$

and

$$\begin{aligned} x &\mapsto \textsc{round}(x_0 + (x - x_0)\cos\vartheta - (y - y_0)\sin\vartheta), \\ y &\mapsto \textsc{round}(y_0 + (x - x_0)\sin\vartheta + (y - y_0)\cos\vartheta), \end{aligned} \tag{7}$$

where $\textsc{trunc}(x)$ is the procedure of cutting off the fractional part of the number x, and $\textsc{round}(x)$ is the procedure of rounding the number x to the nearest integer.

Intuitive apprehension of computer modeling as "feasible" is reflected, in our mind, in the situation when *in a neighborhood of any real trajectory of the mapping (5) there exists at least one computed trajectory (trajectory of the mappings (6) or (7)) and vise versa — when in a neighborhood of any computed trajectory can be found at least one true trajectory of the mapping (5).*

Figure 5 graphs the results of numerics by formula (6) for $\vartheta = 17°$, $x_0 = y_0 = 150$, $\Delta = 3$. It turned out that the trajectory started from rather distant from (x_0, y_0) points tend to some

bounded region, which is not depending on the starting point. "Inside" this region the picture is reminiscent the "Saturn rings" with typical "densifyings" of trajectories and "gaps" between them. Stress, that the density of observed cyclic trajectories is turned out to be approximately 30% of theoretical which evidences that some trajectories are sticked together.

Figure 6 graphs the results of numerics by formula (7) for $\vartheta = 7^\circ$, $x_0 = y_0 = 100$, $\Delta = 6$. Outwardly, the result is turned out to be more consistent with theoretical predictions. At the same time, the resulting picture once again is similar to "Saturn rings" — some trajectories for large number of iterations "float away" from their theoretical positions and as a result we can see "gaps" between them. Another trajectories became very "cloudy" (up to 5% of the values of their radii). At last, very often we can see "transposition" of trajectories — some trajectories starting from points, more distant from (x_0, y_0), tend to cycles lying nearer to (x_0, y_0) than the trajectories starting from points, nearer to (x_0, y_0).

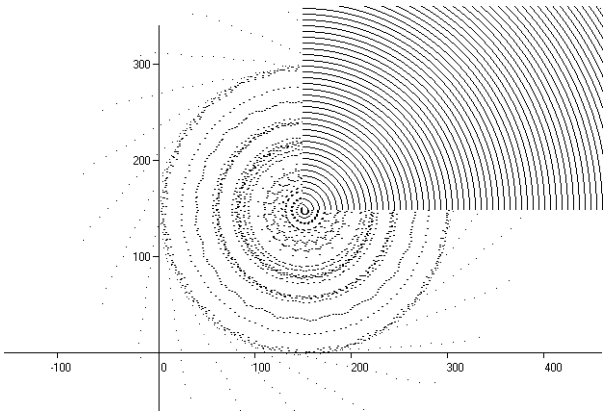

Fig. 5. Phase portrait of the plain rotation mapping on integer lattice around the point $(150, 150)$ under truncating of errors of computed coordinates; the angle of rotation $\vartheta = 17^\circ$, the increment of radii of initial points $\Delta = 3$. The upper-right quadrant contains the fragment of ideal phase portrait for $\Delta = 12$.

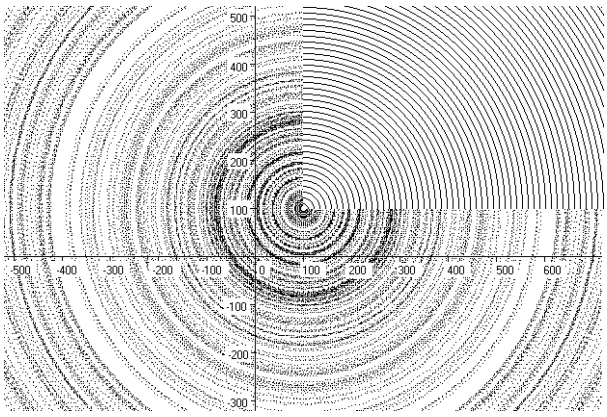

Fig. 6. Phase portrait of the plain rotation mapping on integer lattice around the point $(100, 100)$ under rounding of errors of computed coordinates; the angle of rotation $\vartheta = 7^\circ$, the increment of radii of initial points $\Delta = 6$. The upper-right quadrant contains the fragment of ideal phase portrait for $\Delta = 24$.

Thus, the results are both typical and unsatisfactory in their extreme sensitivity to the mode of rounding (more general, to the scheme of modeling). If we didn't been know in advance the true phase portrait of the system, we would possible interpret the computed results as indication to the complicated structure of the phase portrait of the system under investigation. This leads to the question:

WHICH CONDITIONS GUARANTEE THE SIMILARITY OF PHASE PORTRAITS OF UNDERLAYING CONTINUOUS DYNAMICAL SYSTEM AND ITS SPATIAL DISCRETIZATION?

Note that forming of bounded region of attraction under modeling by formulas (6) are easy to explain (see Theorem 7 below); less clear theoretically is the behavior of trajectories inside this region. It is unclear also how will behave the trajectories computed by formulas (7) — will they be bounded, will they tend to zero, can they escape to infinity etc.

2.3. Moiré structures: artifacts in discretization of oscillating images

Consider the problem of numerical investigation of a smooth function $f = f(x, y)$, $x, y \in \Re^2$ with scalar values.[1] Suppose we would like to represent the behavior of the function f on the black-white digit monitor, which may be considered as a part of the integer lattice $\mathbb{L}_{1,1}$. The natural way to do this is the following: first to draw the level lines of the function $f(x, y)$, say, lines $L_i(f) = \{(x, y) : f(x, y) = i + 1/2\}$, and then to color one by one the gaps between the lines into white and black. In other words we pain alternatively in black and white the Lebesgue sets

$$S_i(f) = \{(x, y) : i - 1/2 < f(x, y) \le i + 1/2\}, \qquad i = 0, \pm 1, \pm 2, \dots \ .$$

The resulting "zebra" represents the behavior of the function f; for example, we can recognize its points of local extremum, but we cannot distinct maximums of the function from its minimums.

It seems natural that for more detailed investigation of the function f it suffices to make the level lines more frequent, i.e. to consider the sets $S_i(\mu f) = \{(x, y) : i - 1/2 < \mu f(x, y) \le i + 1/2\}$ and to tend μ to infinity. But the discrete structure of the screen, which is essentially the rectangular lattice, changes its qualitative portrait drastically just for rather moderate values of the parameter μ: new parasitic regular and irregular structures arise. Figures 7, 8 and 9 represent the copy of screens for the described procedure of the coloring of the level sets of the function

$$\mu f(x, y) = x^2 y - y^4 \qquad -1 \le x, y \le 1 \tag{8}$$

On Figure 7 for $\mu = 10$ everything is in order — we see the real level lines (more precisely the Lebesgue sets) of the function f. But even for $\mu = 300$, on Figure 8, the main fragment of the picture reminds the sequences of the phase portraits of some conics. For $\mu = 288000$ (Figure 9) we can see only the "Persian carpet", which seemingly does not have any relation to the initial function.

Fig. 7. Computer screen with bicolored level sets of the function $10(x^2y - y^4)$ on the lattice 300×200 dots.

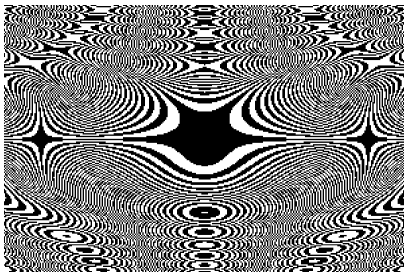

Fig. 8. Computer screen with bicolored level sets of the function $300(x^2y - y^4)$ on the lattice 300×200 dots.

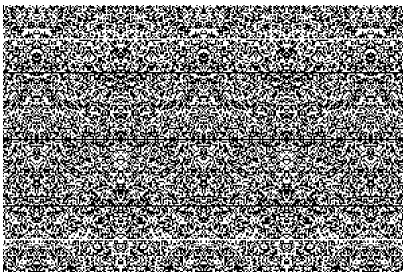

Fig. 9. Computer screen with bicolored level sets of the function $288000(x^2y - y^4)$ on the lattice 300×200 dots.

This sort of artifacts are called *moirés* in the computer graphics. It is important (both for reconstruction not distorted by the discretization image, as well as for reconstruction the lattice of discretization itself) to create the system of the mathematical notions that can adequately describe the asymptotic behavior of the moirés when the pixel is decreasing.

3. PHENOMENOLOGICAL MODEL OF COLLAPSING EFFECTS

It seems important to propose a simple phenomenological model which explains both qualitatively and quantitatively effects mentioned in Subsection 2.1. Such model should allow straightforward theoretical analysis and so provide a means of predicting the severity of such collapsing

[1] We restrict our considerations only by functions of two variables for a convenient graphic representation of effects.

effects in numerical simulation of chaotic dynamical systems. A special model proposed in [27] will be discussed in this subsection.

3.1. Random mappings with single attracting center and equal transient probabilities

The main technical tools below are random mappings. First we consider the so-called random mappings with a single attracting centre. Let us recall the definition. Let $\Delta > 0$. Define a random mapping $T_{\Delta,\kappa}$ of the set $(0, 1, 2, \ldots, \kappa)$ into itself by the formula

$$P(T_{\Delta,\kappa}(i) = j) = \begin{cases} \frac{\Delta}{\kappa+\Delta}, & j = 0, \\ \frac{1}{\Delta+\kappa}, & j \neq 0 \end{cases}$$

and the demand that the image of an element i is chosen independently of those of other elements of $(0, 1, 2, \ldots, \kappa)$.

Define the *collapsing component* $Z_{\Delta,\kappa}$ of the mapping $T_{\Delta,\kappa}$ as a random subset of $(0, 1, \ldots, \kappa)$ defined by

$$Z_{\Delta,\kappa} = \{i \in (0, 1, 2, \ldots, \kappa) : T^n_{\Delta,\kappa}(i) = 0 \text{ for some } n\}.$$

Introduce the random variable

$$Q_{\Delta,\kappa} = \frac{\#(Z_{\Delta,\kappa})}{\kappa + 1}.$$

That is, $Q_{\kappa,\Delta}$ is the proportion of elements of $(0, 1, \ldots, \kappa)$ belonging to the collapsing component of the mapping $T_{\Delta,\kappa}$.

Let $f : [0, 1] \mapsto [0, 1]$ be a function which has a unique absolutely continuous measure μ_f. Suppose that the density of this measure is positive and only the end-points of the interval $[0, 1]$ are singular for the density, moreover, that for a given $l \geq 2$ there exist limits

$$\alpha_0 = \lim_{s\to 0} \mu([0, s])s^{1/l}, \qquad \alpha_1 = \lim_{s\to 1} \mu([1 - s, 1])s^{1/l}.$$

These assumptions hold for the Feigenbaum-Cvitanovich function (1) and many others, see [58] and references therein.

Now, we are able to formulate the hypothesis linking the proportion of collapsing points of the arbitrary mapping $f : [0, 1] \mapsto [0, 1]$ and the property of the collapsing component of the random mapping $T_{\Delta,\kappa}$ as follows:

Hypothesis 1. *There exist positive constants $a = a(f)$, $b = b(f)$ with the following property. For large ν and $1 \ll n \ll \nu$, the statistical characteristics of the sequence*

$$\mathbf{P}(\nu, n; f) = P(\nu; f), P(\nu + 1; f), \ldots, P(\nu + n; f),$$

defined by (3) are similar to those of the random variable $Q_{\Delta(\nu;f),\kappa(\nu;f)}$. The parameters $\Delta(\nu, f)$ and $\kappa(\nu, f)$ are determined by

$$\Delta(\nu; f) = a(f)\frac{\sqrt{\nu}}{\ln \nu}, \qquad \kappa(\nu; f) = \left[b(f)\frac{\nu}{\ln(\nu)}\right],$$

if $l = 2$, and by

$$\Delta(\nu; f) = \frac{a(f)}{\nu^{1/l}}, \qquad \kappa(\nu; f) = \left[b(f)\frac{\nu}{\nu^{2/l}}\right],$$

if $l > 2$.

The justification for this hypothesis "on the physical level" see in [27].

The properties of random mapping with a single attracting centre were studies in details in [10]. In particular the following assertion is valid.

Lemma 1. *In the case $l = 2$ the distributions of random variables $Q_{\Delta(\nu;f),\kappa(\nu;f)}\ln(\nu)$, $\nu = 1, 2, \ldots$, converge to the function*

$$d_c(x) = erfc(1/\sqrt{2cx}), \tag{9}$$

where $c = a(f)^2/b(f)$ and $erfc(y)$ is the complimentary error function [57]

$$erfc(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2}\, dt, \qquad x \geq 0.$$

In the case $l > 2$ the distributions of the random variable $Q_{\Delta(\nu;f),\kappa(\nu;f)}$, $\nu = 1, 2, \ldots$, converge to the function

$$d_c^*(x) = erfc\left(\sqrt{\frac{c}{2}\left(\frac{1}{x} - 1\right)}\right), \qquad 0 < x < 1, \tag{10}$$

where again $c = a(f)^2/b(f)$.

Lemma 2. *Denote by $q(\nu; f)$ the probability of the event*

$$Q_{\Delta(\nu;f),\kappa(\nu;f)} = \kappa(\nu; f) + 1.$$

i.e. the probability that the collapsing component is the whole set $(0, 1, \ldots, \kappa)$. Then

$$q(\nu; f) = \frac{\Delta(\nu; f)}{\kappa(\nu; f) + 1} \approx \frac{a(f)}{b(f)} \frac{1}{\nu^{1/l}}$$

for $l \geq 2$.

3.2. Comparison with experimental data

Statistical analysis of results of experimental calculations is in a good agreement with Hypothesis 1 and theoretical predictions which follow from Lemmas 1 and 2, see [40]. Let us discuss a few examples.

Let $\mathbf{S}$ be a finite set of non-negative values and define the *distribution function* of the set $\mathbf{S}$

$$D(a, \mathbf{S}) = \frac{\#(\{s \in \mathbf{S} : s \leq a\})}{\#(\mathbf{S})}, \qquad a \geq 0.$$

Hypothesis 1 and Lemma 1 imply that $D(a; \ln(\nu)\widetilde{\mathbf{P}}(\nu, n; F^{(\gamma)}))$ should be close, for reasonably large ν and $1 \ll n \ll \nu$, to the function (9) with appropriate c if $l = 2$ and to the function (10) with appropriate c if $l > 2$.

Figure 10 shows $D(a; \ln(\nu)\widetilde{\mathbf{P}}(\nu, n; F^{(\gamma)}))$ for $\gamma = 1/3$, $\nu = 2^{44}$, $n = 1, 2, \ldots, 500$, and the distribution function with density $d_1^*(x)$ (the smooth line). Here,

$$\widetilde{\mathbf{P}}(\nu, n; F^{(\gamma)}) = \widetilde{P}(\nu; F^{(\gamma)}), \widetilde{P}(\nu + 1; F^{(\gamma)}), \ldots, \widetilde{P}(\nu + n; F^{(\gamma)}),$$

and the $\widetilde{P}(\nu + n; F^{(\gamma)})$ were calculated as described in Subsection 2.1.

Analogously, Figure 11 graphs the distribution of the set

$$\widetilde{P}(2^{44}; F)\ln(2^{44}),\ \widetilde{P}(2^{44} + 1; F)\ln(2^{44}),\ \ldots\ , \widetilde{P}(2^{44} + 356; F)\ln(2^{44}) \tag{11}$$

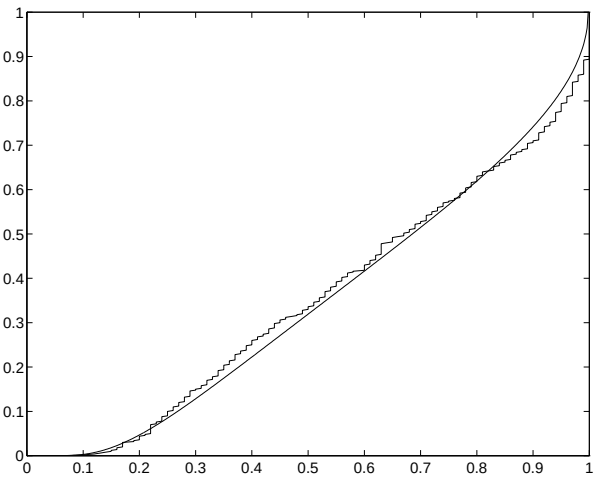

Fig. 10. The graph of the distribution function $D(a; \ln(2^{44})\widetilde{\mathbf{P}}(2^{44}, 500; F^{(\gamma)}))$ for the mapping $F^{(\gamma)}(x)$ with $\gamma = 1/3$ (step function) against the graph of $d_1^*(x)$ (smooth function).

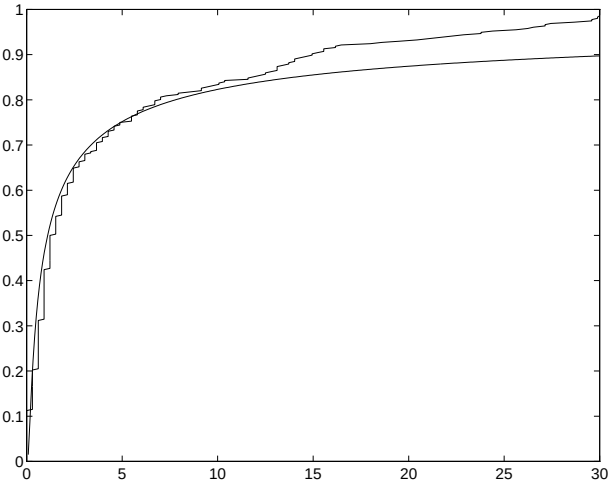

Fig. 11. The graph of the distribution function $D(a; \ln(2^{44})\widetilde{\mathbf{P}}(2^{44}, 356; F))$ for the logistic mapping $F^{(1/2)}(x) = 4x(1-x)$ (step function) against the graph of $d_2(x)$ (smooth function).

for the logistic mapping $F(x) = 4x(1-x)$ (which has a quadratic singularity) and the function (9) for $c = 2$. (It was proposed at first to consider the set

$$\widetilde{P}(2^{44}; F)\ln(2^{44}),\ \widetilde{P}(2^{44}+1; F)\ln(2^{44}),\ \ldots\ , \widetilde{P}(2^{44}+500; F)\ln(2^{44})$$

instead of (11). Unfortunately the computer experiment on Cray was terminated after 24 hours of cpt–time, by the fault of operator who didn't inform his colleague about high priority of this experiment). Since each $\widetilde{P}(\nu + m; F) \le 1$, $m = 0, 1, \ldots, n$, $\ln \nu = \ln(2^{44}) \approx 30$,

$$D_{27} := D(a; (\ln \nu)\widetilde{\mathbf{P}}(\nu, n; F))$$

is a truncated distribution as shown, with domain $[0, 30.0]$. On the other hand, $d_2(x)$ is a distribution with infinite domain $[0, \infty)$. Obviously, the truncated tail of D_{27} is inflated by the tail of d_2 in $[30.0, \infty)$. When this is taken into account, the two distributions are in good agreement.

Say that the lattice $\mathbb{L}_\nu$ is *absolutely collapsing* for f if iterations of each point $\xi \in \mathbb{L}_\nu$ under mapping f are eventually zero. Some general results about absolutely collapsing discretizations were discussed in [24]. Hypothesis 1 and Lemma 2 imply that the quantity $w(N; f)$ of absolutely collapsing lattices in the sequence $\mathbb{L}_1$, $\mathbb{L}_2$, …, $\mathbb{L}_N$ should be of order $(2a(f)/b(f))N^{(l-1)/l}$. In particular, $w(N; f)^{l/(l-1)}$ should be approximately linear in N. Figures 12 and 13 demonstrate the corresponding graphs for the cases $f = F$ and $f = F^{(\gamma)}$.

Many simulations of this type were performed, on grids of various sizes and with different non-degenerated unimodal functions. All of them unambiguously justified the validity of Hypothesis 1.

3.3. Interval stochastic matrices in preventing collapsing effects

What can be done to avoid such collapsing effects described in Subsection 2.1? There are some general strategies, including the following:

S1. To use special "anti–collapsing" discretizations.

S2. To use discretizations which include random perturbations; this is usually done by replacing f_ν by a Markov chain on the lattice $\mathbb{L}$.

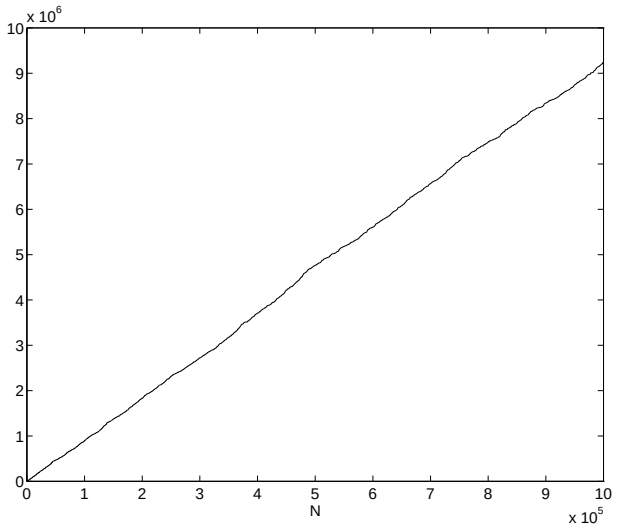

Fig. 12. Squared accumulated number, $w(N;F)^2$, of absolutely collapsing points for the logistic mapping $F^{(1/2)}(x) = 4x(1-x)$.

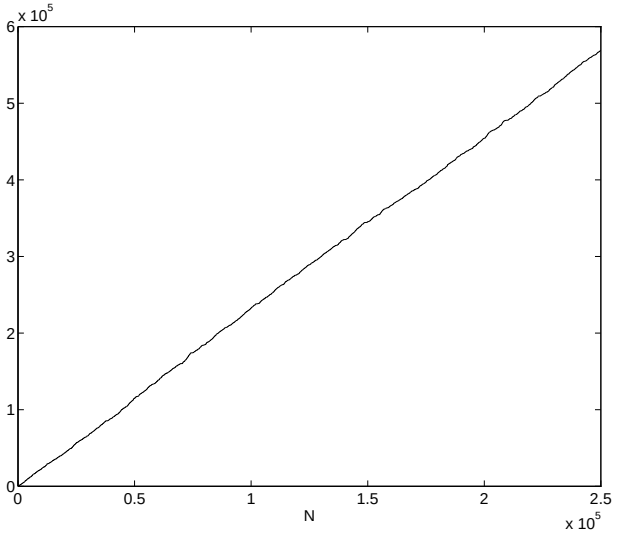

Fig. 13. Powered accumulated number, $w(N;F^{(\gamma)})^{3/2}$, of absolutely collapsing points for the for the mapping $F^{(\gamma)}(x)$ with $\gamma = 1/3$.

S3. To use multi–valued discretizations.

S4. To sweep the problem under the carpet.

All these strategies are quite reasonable though the fundamental theoretical questions concerning each of them are still open.

The issues concerning the first strategy are quite clear:

to give recommendations, at least general, for choosing a discretization which guarantees that collapsing effects are avoided.

Point out one of possible ways to resolve this problem. But firstly, let us present some terms and definitions. Let Ω be a compact metric space endowed with metric ρ. Denote by $\mathbf{Sep}(y,X) = \inf_{x\in X}\rho(y,x)$ the *separation* of an element $y \in \Omega$ from the set $X \subseteq \Omega$, and denote by $\mathbf{Sep}(Y,X)$, $Y, X \subseteq \Omega$, the *Hausdorff separation*[2] of Y from X:

$$\mathbf{Sep}(Y,X) = \sup_{y\in Y}\mathbf{Sep}(y,X).$$

The same symbol **Sep** will be used for the Hausdorff separation between sets belonging to the Cartesian product $\Omega \times \ldots \times \Omega$ with the metric

$$\rho_N((x_0,x_1,\ldots x_N),(y_0,y_1,\ldots y_N)) = \max_{0\le n\le N}\rho(x_n,y_n).$$

At last, by $\mathbf{Gr}(f)$ denote *the graph* of the map f.

Let now $\mathbb{T}^d$ be the standard d–dimensional torus, let $f : \mathbb{T}^d \mapsto \mathbb{T}^d$ be a mapping with invariant Lebesgue measure and let $\mathbb{L}$ be a lattice on $\mathbb{T}^d$ induced by the standard uniform $1/\nu$ lattice on the cube.

[2] The notion of Hausdorff separation should not be confused with the notion of Hausdorff metric which is $\chi(X,Y) = \max\{\mathbf{Sep}(X,Y), \mathbf{Sep}(Y,X)\}$.

Theorem 1 (see [21]). *For every $\varepsilon > 0$ there exists a permutation π of the lattice $\mathbb{L}_\nu$ satisfying* $\mathbf{Sep}(\mathbf{Gr}(\pi), \mathbf{Gr}(f)) \le \varepsilon$.

This theorem seems to be similar to Lax result from [52]. At the same time it is somewhat surprising because in contrast to Lax theorem the mapping f is not even assumed to be injective. It can be regarded as an example of how to choose a good approximation. Clearly, mappings such as permutations avoid collapsing effects. It is important that there are rapid algorithms to find such permutations corresponding to f constructively. They can be implemented as simple computer programs.

Concerning the first strategy, there are some deep theoretical results (see for instance [9]), mainly connected with using the Stetter discretization [66], but these are not always applicable to the situation under consideration.

Strategies **S2** and **S3** are quite natural from the point of view of information theory. Indeed, as it was told in Introduction, entropy of single–valued deterministic discretization is, formally speaking zero. To restore the entropy we should add a certain source of information artificially. The evident ways to do it are to introduce a random (Strategy **S2**) or multi–valued (Strategy **S3**) noise. There are a number of deep theoretical results associated with the second strategy, especially general theorems of Kifer [39] and Blank [6], but, again, the main question is open:

WHAT IS AN APPROPRIATE LEVEL OF RANDOMNESS OR "MULTI–VALUEDNESS" IN STRATEGIES S2 OR S3?

If the stochastic component in the second strategy is too large then the dynamics of the model will differ markedly from those of the original system, but if it is not strong enough, then collapsing effects will be present.

With the strategy **S4** the following theoretical question is connected:

WHEN AND WHY THE STRATEGY OF IGNORING THE PROBLEM IS JUSTIFIED?

That is, when are collapsing effects unlikely (although still possible by virtue of Theorem 5 from [24]).

The approach which will be described below is useful when using either the second or the third strategy and would possibly also be useful in a rigorous analysis of the first and the fourth strategy. What is especially important is that it demonstrates that all these strategies are related.

Let $\mathcal{M}_d$ denote the totality of all real square $d \times d$ matrices $A = (a_{ij})$ with nonnegative entries $a_{ij} \ge 0$. Vectors $v \in \mathbb{R}^d$ will be treated as columns and matrix multiplication Av will be on the left. The set $\mathcal{M}_d$ has a natural partial order given by

$$A \le B \;\Leftrightarrow\; a_{ij} \le b_{ij}\ ,\ i,j = 1,\dots,d\ ,$$

where $A = (a_{ij})$ and $B = (b_{ij})$ are matrices from $\mathcal{M}_d$. Recall that a matrix $C = (c_{ij})$ is a *stochastic matrix* if

$$\sum_{i=1}^{d} c_{ij} = 1, \qquad j = 1,\dots,d\ .$$

The class of all stochastic matrices in $\mathcal{M}_d$ will be denoted by $\mathcal{S}_d$.

Let $\mathcal{M}_d^-$ be the set of all matrices $A = (a_{ij}) \in \mathcal{M}_d$ satisfying

$$\sum_{i=1}^{d} a_{ij} \;\le\; 1, \qquad j = 1,\dots,d$$

and $\mathcal{M}_d^+$ the set of all matrices $B = (b_{ij}) \in \mathcal{M}_d$ satisfying

$$\sum_{i=1}^{d} b_{ij} \ \geq \ 1, \qquad j = 1, \ldots, d \ .$$

For any two matrices $A \in \mathcal{M}_d^-$ and $B \in \mathcal{M}_d^+$ such that $A \leq B$, let $\widehat{AB}$ denote the set of all stochastic matrices between A and B, that is

$$\widehat{AB} = \{C \in \mathcal{S}_d : A \leq C \leq B\} \ .$$

The set $\widehat{AB}$ will be called *the interval stochastic matrix* with boundaries A and B.

Finally, if σ_d is the standard simplex in $\mathbb{R}^d$, that is,

$$\sigma_d = \left\{ (p_1, \ldots, p_d) \in \mathbb{R}^d \ : \ p_i \geq 0, \ i = 1, \ldots, d \quad \text{and} \quad \sum_{i=1}^{d} p_i = 1 \right\} \ ,$$

then for any vector $p \in \sigma_d$ and for any interval stochastic matrix $\widehat{AB}$ define

$$\widehat{AB}\, p = \{Cp : C \in \widehat{AB}\}.$$

Our principal result is an explicit representation of the set $\widehat{AB}\, p$. Let $\mathcal{I}_d$ be the class of all subsets of $\{1, 2, \ldots, d\}$. For any $j \in \{1, \ldots, d\}$ and $I \in \mathcal{I}_d$ define

$$H_j(I, \widehat{AB}) = \min \left\{ \sum_{i \in I} b_{ij}, \ 1 - \sum_{i \notin I} a_{ij} \right\}.$$

Theorem 2. *The set $\widehat{AB}\, p$ is precisely the set of all vectors $q \in \sigma_d$ satisfying*

$$\sum_{j=1}^{d} p_j H_j(I, \widehat{AB}) \ \geq \ \sum_{i \in I} q_i, \qquad \forall \ I \in \mathcal{I}_d \ .$$

An important corollary to this theorem can be formulated as follows. A vector $p \in \sigma_d$ is said to be *semi–invariant* for the interval stochastic matrix $\widehat{AB}$ if

$$\sum_{j=1}^{d} p_j H_j(I, \widehat{AB}) \ \geq \ \sum_{i \in I} p_i, \qquad \forall \ I \in \mathcal{I}_d \ .$$

For any stochastic matrix C, let $\mathrm{Fix}(C)$ denote the set of all vectors $x \in \sigma_d$ such that $Cx = x$.

Corollary 1. *The set $\bigcup\{\mathrm{Fix}(C) : \ C \in \widehat{AB}\}$ is precisely the set of all semi–invariant vectors of the interval stochastic matrix $\widehat{AB}$.*

Now we will show how Theorem 2 works in analysis of spatial discretizations.

Example 1 (**Application to Strategy S2**). Let $\mathcal{P}$ denote the totality of Borel probability measures on Ω. The *Prokhorov metric* ρ_P on $\mathcal{P}$ can be defined by

$$\rho_P(\mu_1, \mu_2) = \inf\{\varepsilon > 0 : \mu_1(\mathcal{O}_\varepsilon(S)) \leq \mu_2(S) - \varepsilon, \ \forall S \in \mathcal{B}\}, \tag{12}$$

where $\mathcal{O}_\varepsilon(S)$ is the ε–neighborhood of S. This metric is a standard in many branches of probability theory (see, for instance, [32]). By virtue of Theorem 2 *for any invariant measure μ of f there exists a Markov chain on $\mathbb{L}$ with a stationary measure μ_* satisfying*

$$\rho_P(\mu_*, \mu) \le h(\mathbb{L}), \tag{13}$$

where h is defined by

$$h(\mathbb{L}) = \sup_x \inf_\xi \{\rho(\xi, x) : \xi \in \mathbb{L},\ x \in \Omega\},$$

and with transient probabilities $p(\xi, \eta)$ satisfying

$$P(\xi, \eta) = 0 \quad \text{for} \quad \mathbf{Sep}((\xi, \eta), \mathbf{Gr}(f)) > h(\mathbb{L}). \tag{14}$$

Inequalities (13), (14) provide an estimate of the minimal level of stochasticity which is sufficient to suppress collapsing effects. This level is surprisingly low. In an obvious sense it coincides with the spatial step of the discretization under consideration. Certainly, it does not provide a full answer to the question about randomization level, but it is a step in this direction. We mention also that analogous assertions were useful in analyzing some algorithms to find interesting invariant measures [25].

Example 2 (**Reduction to a linear programming problem**). From Theorem 2 it follows that *for any invariant measure μ of f there exists a measure μ_* on $\mathbb{L}$ satisfying* (13) *and*

$$\mu_*(\mathbb{L}_*) \le \mu_*(\Phi_*^{-1}\mathbb{L}) \quad \text{for} \quad \mathbb{L}_* \subseteq \mathbb{L}.$$

This is interesting because often an invariant measure with certain extremal properties is of interest. Above relations reduce the numerical search of such a measure to a linear programming problem.

Example 3 (**Application to Strategy S3**). Choose a fixed multi–valued discretization Φ_f of the mapping f satisfying

$$\mathbf{Gr}(\Phi_f) = \{(\xi, \eta) : \rho((\xi, \eta), \mathbf{Gr}(f) \le h(\mathbb{L})\},$$

Then the points from the graph of Φ_f are points from the $h(\mathbb{L})$–tube around the graph of f. In a sense it is a reasonable minimal multi–valued realization of f. From Theorem 2 it follows that *for any invariant measure μ of f there exists an invariant measure μ_* of Φ_f satisfying the inequality* (13). This proposition estimates the minimal level of "multi–valuedness" sufficient to suppress collapsing effects.

The results which have been discussed in this Subsection are based on the papers [22, 26]. Generally speaking, troubles in avoiding collapsing effects are similar to those in solving of incorrect problems. In particular collapsing effect could be to a certain extend avoided if the computer experiment will be stopped rather early. This question is discussed in particular in [9]. New methods of estimating transition modes are necessary to understand constructively, when it is reasonable to stop computational process. Possibly here will be useful the scheme proposed in [45].

4. PHENOMENOLOGICAL MODEL OF STATISTICS OF CYCLE LENGTHS AND FIRST RECURRENCE TIMES FOR DISCRETIZATIONS OF DYNAMICAL SYSTEMS

Unfortunately, for the Feigenbaum–Cvitanovich mapping (so as for other systems with complicated behavior) the straightforward theoretical analysis of combinatorics of their discretizations is extremely difficult. For this reason we used a compromise approach based on rigorous analysis of

phenomenological models for statistics of cycle lengths, first recurrence times and other combinatorial characteristics of discretizations of dynamical systems. Naturally, this approach is not new. In particular, a phenomenological model [31, 35, 53], based on the theory of completely random mappings [8], has been most successful in the situations when a "typical" discretization of the continuous system does not have a strong algebraic structure and the underlying continuous system has a stochastic attractor for which the Hausdorff dimension coincides with its correlation dimension (see details in [28, 35]). The condition concerning the dimension of attractor is rather restrictive and does not hold even for the simplest one-dimensional systems generated by the mappings from the family (1) with $\gamma < \frac{1}{2}$, for which the Hausdorff dimension of the stochastic attractor $[0, 1]$ is equal to 1, whereas the correlation dimension is equal to 2γ.

The first successful models for systems with different Hausdorff and correlation dimensions of stochastic attractors were suggested in [27, 28]. These models were based on the theory of random mappings [33, 42], in particular, random mappings with a single attracting centre [65] and was refined in [29]. While these models gave better results than those in [33, 42], they are still not quite satisfactory quantitatively. Below a more appropriate class of models is presented and analyzed.

4.1. Random mappings with single absorbing center and different transient probabilities

In this Section we discuss the model based on the notion of a random mapping with single absorbing center whose properties are recalled below.

Let N be a natural number, $\alpha > 0$ and $\gamma \in (0, 1/2)$. Consider the random mapping $T^{(\gamma)}_{\alpha,N}$ of the set $X(N) = \{0, \dots, N\}$ into itself, which is defined by the following conditions. Define $q_0(\alpha, \gamma) = \alpha^\gamma$ and $q_i(\alpha, \gamma) = (\alpha + i)^\gamma - (\alpha + i - 1)^\gamma$, $i = 1, 2, \dots, N$. Suppose that the point 0 is fixed for each realization $\widehat{T}$ of the random mapping $T^{(\gamma)}_{\alpha,N}$ and that the probability of the realization $\widehat{T}$ is equal to

$$\prod_{i=1}^{N} (\alpha + N)^{-\gamma} q_{\widehat{T}(i)}(\alpha, \gamma).$$

In other words, in the construction of a realization $\widehat{T}$ the images $\widehat{T}(i)$ of the points $i \in X(N)$ are chosen independently (i)and equiprobably with the probability of the event $T^{(\gamma)}_{\alpha,N}(i) = j$ to be proportional to $q_j(\alpha, \gamma)$. This definition is a natural analog of that for random mappings with a single attracting centre [65] (see also [8, 10]) and thus $T^{(\gamma)}_{\alpha,N}$ is natural to call *the random mapping with single absorbing center and different transient probabilities.*

Each realization $\widehat{T}$ of the random mapping $T^{(\gamma)}_{\alpha,N}$ is a deterministic mapping. Thus for each $i \in X(N)$ the trajectory $\mathbf{Tr}(i, \widehat{T})$, that is the sequence $i_0, i_1, \dots, i_n, \dots$ which satisfies the equalities $i_0 = i$ and $i_n = \widehat{T}(i_{n-1})$, $n = 1, 2, \dots$, is uniquely defined. For each such trajectory $\mathbf{Tr}(i, \widehat{T})$ the first recurrence time $Q(i, \widehat{T})$ is defined to be the first n after which the trajectory is cyclic with the minimal period, say $C(i, \widehat{T})$.

Let $\#(X)$ denote the number of elements in a finite set X and let

$$\mathcal{Q}(x, \widehat{T}) = \frac{1}{N}\#[\{i : Q(i, \widehat{T}) < xN^\gamma\}], \quad \mathcal{C}(x, \widehat{T}) = \frac{1}{N}\#[\{i : C(i, \widehat{T}) < xN^\gamma\}], \qquad x \geq 0 \tag{15}$$

denote the scaled distribution functions of the first recurrence moments and of the minimal periods for the totality of trajectories of the mapping $\widehat{T}$. Also denote by $\mathcal{P}(\widehat{T})$ the proportion of collapsing elements of the mapping $\widehat{T}$, i.e. the proportion of those i for which $C(i, \widehat{T}) = 1$. Note that for fixed α, γ, N the quantities (15) are random functions; so let $Q^{(\gamma)}(x; \alpha, N)$, $C^{(\gamma)}(x; \alpha, N)$, $x \geq 0$, denote the corresponding mathematical expectations. Finally denote for $x \in [0, 1]$ the distribution function of the random variable $\mathcal{P}(\widehat{T})$ by $P^{(\gamma)}(x; \alpha, N)$.

Theorem 3 (see [46, 49]). *For each $\alpha, \beta > 0$ and $0 < \gamma < 1/2$ there are valid the limit equalities*

$$\lim_{N\to\infty} Q^{(\gamma)}(x;\alpha,\beta N) = 1 - F^{(\gamma)}(\beta x;\alpha), \qquad \lim_{N\to\infty} C^{(\gamma)}(x;\alpha,\beta N) = H^{(\gamma)}(\beta x;\alpha),$$

$$\lim_{N\to\infty} \int_0^1 (1 - P^{(\gamma)}(x;\alpha,N))\,dx = \alpha^\gamma \int_0^\infty F^{(\gamma)}(x,\alpha)\,dx,$$

where

$$F^{(\gamma)}(x;\alpha) = e^{-\alpha^\gamma x} \prod_{i=1}^{\infty} \frac{1 + q_i(\alpha,\gamma)x}{e^{q_i(\alpha,\gamma)x}},$$

$$H^{(\gamma)}(x;\alpha) = 1 - F^{(\gamma)}(x;\alpha) + \alpha^\gamma \int_x^\infty F^{(\gamma)}(y;\alpha)\,dy + x\int_x^\infty F^{(\gamma)}(y;\alpha) \sum_{i=1}^{\infty} \frac{q_i(\alpha,\gamma)^2}{1 + q_i(\alpha,\gamma)y}\,dy.$$

Note that only one particular case of random mappings with a single absorbing centre and with transient probabilities generated by discretizations of the function $q(\alpha,\gamma;t) = (t+\alpha)^{\gamma-1}$ has been considered. The choice of this particular random mapping was conditioned by the heuristic reason that the mapping $F^{(\gamma)}$ has the absolute continuous invariant Sinai–Ruelle–Bowen measure with the singularity of power $\gamma - 1$ at the points $t = 0, 1$.

Proof of Theorem 3 is based on investigation of asymptotics of the elementary symmetric functions on probability distributions of a special form. Let the function $p(k,n)$ id defined for $k = 1, 2, \ldots, n$ by the equality

$$p(k,n) = \int_{\frac{k-1}{n}}^{\frac{k}{n}} q(t)\,dt,$$

where

$$q(t) = \alpha\beta(t)t^{\gamma-1} \geq 0, \qquad \int_0^1 q(t)\,dt = 1, \qquad \beta(0) = 1, \quad \alpha > 0, \quad 0 < \gamma \leq 1.$$

Theorem 4 (see [49]). *The following relations are valid for $n \to \infty$*

$$\pi(m,n) \to e^{-\frac{1}{2}\left(\int_0^1 q^2(t)\,dt\right)x^2} \qquad \text{for} \qquad \frac{m}{\sqrt{n}} \to x, \quad \frac{1}{2} < \gamma \leq 1;$$

$$\pi(m,n) \to e^{-\frac{1}{2}\alpha^2 x^2} \qquad \text{for} \qquad \frac{m}{\sqrt{n(\log n)^{-1}}} \to x, \quad \gamma = \frac{1}{2};$$

$$\pi(m,n) \to \mathcal{L}(\gamma,\alpha x) \qquad \text{for} \qquad \frac{m}{n^\gamma} \to x, \quad 0 < \gamma < \frac{1}{2}$$

where $\pi(m,n) = m!\sigma_m(p(1,n), p(2,n), \ldots, p(n,n))$, $\sigma_m(x_1, x_2, \ldots, x_n)$ is the mth elementary symmetric function and $\mathcal{L}(\gamma,z)$ is the entire analytic function defined by the infinite product

$$\mathcal{L}(\gamma,z) = \prod_{k=1}^{\infty} \left\{ \left[1 + \frac{k^\gamma - (k-1)^\gamma}{\gamma} z\right] e^{-\frac{k^\gamma - (k-1)^\gamma}{\gamma} z} \right\}.$$

Remark that for $0 < \gamma < \frac{1}{2}$ and for the values $|x| < \frac{\gamma}{\alpha}$ the following more convenient representation

$$\pi(m,n) \to e^{-\sum_{r=2}^{\infty} \frac{(-1)^r}{r} \left\{ \sum_{k=1}^{\infty} \left(\frac{k^\gamma - (k-1)^\gamma}{\gamma} \right)^r \right\} (\alpha x)^r} \quad \text{for} \quad \frac{m}{n^\gamma} \to x\ .$$

is also valid.

4.2. Principle of Correspondence

Let us come back to the analysis of discretizations of continuous dynamical systems from the family (1). Every ν-discretization $F_\nu^{(\gamma)}$ maps a finite lattice $\mathbb{L}_\nu = \{0, 1/\nu, \dots, (\nu-1)/\nu, 1\}$ into itself, so for each $\xi \in \mathbb{L}_\nu$ the first recurrence moment $Q(\xi, F_\nu^{(\gamma)})$, and the corresponding minimal period of a cyclic part of the trajectory $Q(\xi, F_\nu^{(\gamma)})$, are well defined, as are the distribution functions $\mathcal{Q}(x; F_\nu^{(\gamma)})$, $\mathcal{C}(x; F_\nu^{(\gamma)})$ and the number $\mathcal{P}(F_\nu^{(\gamma)})$.

As it was mentioned above, a theoretical study of the sequences $\{\mathcal{Q}(x, F_\nu^{(\gamma)})\}_{n=1}^\infty$, $\{\mathcal{C}(x, F_\nu^{(\gamma)})\}_{n=1}^\infty$ and $\{\mathcal{P}(F_\nu^{(\gamma)})\}_{n=1}^\infty$ is more complicated than that of the sequences $\{\mathcal{Q}(x; \widehat{T}_\nu)\}_{n=1}^\infty$, $\{\mathcal{C}(x; \widehat{T}_\nu)\}_{n=1}^\infty$ and $\{\mathcal{P}(\widehat{T}_\nu)\}_{n=1}^\infty$. Nevertheless, this difficulty can be overcome by means of the *principle of correspondence*, to be formulated below. This principle is not a rigorous mathematical theorem, but admits an heuristic explanation analogously to the reasoning in [27, Section 2], which in turn is not dissimilar that in [35, 53]. Numerical experiments demonstrate that this principle hold with rather high accuracy.

For each $x \geq 0$ introduce the functions

$$\mathbf{q}^{(\gamma)}(x; N, M) = \frac{1}{M} \sum_{\nu=N+1}^{N+M} \mathcal{Q}(x; F_\nu^{(\gamma)}), \qquad \mathbf{c}^{(\gamma)}(x; N, M) = \frac{1}{M} \sum_{\nu=N+1}^{N+M} \mathcal{C}(x; F_\nu^{(\gamma)}).$$

which are the distribution functions $\mathcal{Q}(x; F_\nu^{(\gamma)})$ and $\mathcal{C}(x; F_\nu^{(\gamma)})$ averaged on M lattices $\mathbb{L}_{N+1}$, $\mathbb{L}_{N+2}$, $\dots$, $\mathbb{L}_{N+M}$. Denote also the distribution function of the set $\{\mathcal{P}(F_\nu^{(\gamma)}) : N < \nu \leq N+M\}$ by $\mathbf{p}^{(\gamma)}(x; N, M)$, $x \in [0, 1]$. A family of positive integers $\mathcal{N}$ is said to be *dense* if

$$\lim_{n\to\infty} \frac{\#[\{m \in \mathcal{N} : m \leq n\}]}{n} = 1.$$

Principle of Correspondence. *There exist constants $\alpha(\gamma)$, $\beta(\gamma) > 0$ and a dense set of integers $\mathcal{N}(M)$ depending on a positive integer M such that the functions $\mathbf{q}^{(\gamma)}(x; N, M)$, $\mathbf{c}^{(\gamma)}(x; N, M)$ and $\mathbf{p}^{(\gamma)}(x; N, M)$ are close in the Levy metric [34] to the corresponding functions $Q^{(\gamma)}(x; \alpha, \beta N)$, $C^{(\gamma)}(x; \alpha, \beta N)$, and $P^{(\gamma)}(x; \alpha, N)$ for all sufficiently large $N \in \mathcal{N}(M)$.*

The Principle of Correspondence for the sequences $\mathbf{q}$ and $\mathbf{c}$, together with Theorem 3 suggest that for randomly chosen $1 \ll M \ll N$ the functions $\mathbf{q}^{(\gamma)}(x; M, N)$ and $\mathbf{c}^{(\gamma)}(x; M, N)$ should be similar to the functions $1 - F^{(\gamma)}(\beta x; \alpha)$ and $H^{(\gamma)}(\beta x; \alpha)$. This assertion can be tested numerically through simulation. Furthermore the Principle of Correspondence for the sequences $\mathcal{P}$ together with Theorem 3 suggest also that for $1 \ll M \ll N$ the mean value

$$\mu = \frac{1}{M} \sum_{m=N+1}^{N+M} \mathcal{P}(F_m^{(\gamma)}) = \int_0^1 (1 - \mathbf{p}^{(\gamma)}(x; M, N))\, dx$$

should be close to the value $\mu^{(\gamma)} = \alpha^\gamma \int_0^\infty F^{(\gamma)}(x, \alpha(\gamma))\, dx$. Again the experimental result $\mu \approx 0,675$ appeared to be indeed quite close to the theoretical prediction $\mu^{(\gamma)} \approx 0,678$ which was calculated for $\alpha = 0,3$.

Similar experiments were carried out also for other values of γ, such as $\gamma = 2/5, 2/7$ etc., and also for other values of the parameters M, N. All of these experiments supported the Principle of Correspondence.

The Principle of Correspondence formulated above for three concrete combinatorial characteristics of the spatial discretizations of continuous dynamical systems is also applicable to the

investigation of other combinatorial characteristics, such as statistics of absolutely collapsing discretizations [27, p. 566], basins of attractions etc. *We don't know any natural combinatorial characteristics of discretizations of the mappings (1) for which the Principle of Correspondence is not valid.*

The Principle of Correspondence can be used without change for other quasi-chaotic systems which have an invariant stochastic attractor for which the corresponding invariant measure has a singularity on a preimage of a fixed point of the mapping. When the last condition does not hold, the Principle of Correspondence should be modified slightly: in the definition of random mapping $T^{(\gamma)}_{\alpha,N}$ the requirement that 0 is a fixed point there must be omitted and, correspondingly, the expression for transitive probabilities should be changed. Nevertheless, the problem of *formal justification of the Principle of Correspondence* is still open [17, 19, 46, 47].

4.3. Consistent discretizations and shadowing

Suppose that a mapping $f: \Omega \mapsto \Omega$, on a compact metric space Ω endowed with metric ρ, generates a discrete dynamical system. As was mentioned in Section 2.1, its set of all trajectories can differ dramatically from that of any single–valued discretization φ, even very fine [11]. Such effects are an inevitable consequence of discretization in the sense that *there always exists some discretization which collapses a given system f onto a given f–invariant set*, in particular onto a fixed point or cycle [23, 24]. At the same time, as it was mentioned it Subsection 3.3 we can eliminate such degenerate, collapsing behavior if instead single–valued discretizations we will deal with stochastic or multi–valued perturbations of the original system f. In this Subsection we consider in more details the question about using multi–valued discretization to represent the flow of trajectories of the underlying system. The choice of an appropriate model system φ introduces a conundrum which frequently arises in the theory of ill–posed problems. If the perturbation that is introduced is too large, then the behavior of the system φ, while not degenerate can differ markedly from f. On the other hand, if the perturbation is not strong enough, collapsing effects will not be avoided. Consequently, questions about the robustness of systems to various levels of stochastic or multi–valued perturbation are very important. To make a reasonable choice, idea of *consistence of multi–valued discretization* is often useful.

Below, a dynamical system is generated by a Borel mapping $f: \Omega \mapsto \Omega$. Suppose that $\mathbb{L}$ is a finite subset (lattice, grid) of Ω and consider a map $\varphi: \mathbb{L} \mapsto 2^{\mathbb{L}}$ with the graph $\mathbf{Gr}(\varphi) \subseteq \mathbb{L}\times\mathbb{L} \subseteq \Omega\times\Omega$ as a discretization of the system f. An estimate of the accuracy of such a discretization φ is provided by the two quantities

$$d(\varphi, f) = \mathbf{Sep}(\mathbf{Gr}(\varphi), \mathbf{Gr}(f)) \qquad \text{and} \qquad d(f, \varphi) = \mathbf{Sep}(\mathbf{Gr}(f), \mathbf{Gr}(\varphi)).$$

Note that d is not a metric.

The mapping $\varphi: \mathbb{L} \mapsto 2^{\mathbb{L}}$ will be called an *α–consistent discretization of f* if at least one of the following two conditions hold:

C1. There exists a cover $X(\xi)$, $\xi \in \mathbb{L}$ of Ω with $X(\xi) \subseteq \overline{\mathcal{O}_\alpha(\xi)}$, $\xi \in \mathbb{L}$ and

$$f(X(\xi)) \subseteq X(\varphi(\xi)).$$

C2. There exist nonempty sets $\Xi(x)$, $x \in \Omega$ of $\mathbb{L}$ with $\Xi(x) \subseteq \overline{\mathcal{O}_\alpha(x)}$ and

$$\varphi(\Xi(x)) \supseteq \Xi(f(x)).$$

Consider some simple examples of α–consistent discretizations.

Example 4. For any $\xi \in \mathbb{L}$ denote by $X(\xi)$ the subset of Ω defined by

$$X(\xi) = \left\{ x : \rho(x,\xi) = \min_{\eta\in\mathbb{L}} \rho(x,\eta) \right\},$$

and for any subset $\Omega_* \subseteq \Omega$ denote

$$\Xi(\Omega_*) = \left\{ \xi : X(\xi) \bigcap \Omega_* \neq \emptyset \right\}.$$

Then the multi–valued mapping $\varphi : \mathbb{L} \mapsto 2^{\mathbb{L}}$ defined by $\varphi(\xi) = \Xi(f(X(\xi))$ for $\xi \in \mathbb{L}$ will be an $h(\mathbb{L})$–consistent discretization of f with $d(\varphi, f) \leq h(\mathbb{L})$. In this example the condition C1 is fulfilled.

Example 5. Denote $\mathcal{A} = \overline{\mathcal{O}_{h(\mathbb{L})}(\mathbf{Gr}(f))} \bigcap (\mathbb{L}\times\mathbb{L})$ and consider the multi–valued mapping $\varphi : \mathbb{L} \mapsto 2^{\mathbb{L}}$ with the graph $\mathcal{A}$. Then this map is a $h(\mathbb{L})$–consistent discretization of f with $d(\varphi, f) \leq h(\mathbb{L})$. In this example the condition C1 holds. with the same cover as in the previous example.

Example 6. Let $\psi : \mathbb{L} \mapsto 2^{\mathbb{L}}$ be a given mapping with $d(f, \psi) \leq \alpha$. Then the multi–valued mapping $\varphi(\xi) = \overline{\mathcal{O}_{2\alpha}(\psi(\xi))} \bigcap \mathbb{L}$ will be an α–consistent discretization of f with $d(\varphi, f) \leq 3\alpha$. In this example the condition C2 is fulfilled with $\Xi(x) = \overline{\mathcal{O}_{\alpha}(x)} \bigcap \mathbb{L}$.

Denote by $\mathrm{C}(\varphi, f)$ the greatest lower bound of those α for which φ is an α–consistent discretization of f. Consider a mapping $\varphi : \mathbb{L} \mapsto 2^{\mathbb{L}}$ with the graph $\mathbf{Gr}(\varphi) \subseteq \mathbb{L}\times\mathbb{L} \subseteq \Omega\times\Omega$ as a discretization of the system f.

Recall definitions of trajectories and pseudo–trajectories for single– and multi–valued mappings. A sequence $\mathbf{y} = y_0, y_1, \ldots, y_N$ is called *a trajectory* of a single–valued mapping f if $y_{n+1} = f(y_n)$ for $n = 0, 1, \ldots, N-1$; it is called *a trajectory* of a multi–valued mapping φ if $y_{n+1} \in \varphi(y_n)$ for $n = 0, 1, \ldots, N-1$. A sequence $y_0, y_1, \ldots, y_N$ is called *a γ–pseudo–trajectory* of a single–valued mapping f if $\rho(y_{n+1}, f(y_n)) \leq \gamma$ for $n = 0, 1, \ldots, N-1$; it is called *a γ–pseudo–trajectory* of a multi–valued mapping φ if $\mathbf{Sep}(y_{n+1}, \varphi(y_n)) \leq \gamma$ for $n = 0, 1, \ldots, N-1$. Denote by $\mathbf{Tr}(f)$ the totality of trajectories of f and denote by $\mathbf{Tr}(\varphi)$ the totality of trajectories of φ. Sometimes it is necessary to distinguish more accurately to which set belong elements of a trajectory, or to which γ corresponds a given pseudo–trajectory. In this case we will denote by $\mathbf{Tr}(f, S, \gamma)$ the totality of finite or infinite γ–pseudo–trajectories belonging entirely to a subset $S \subseteq \Omega$. Since a true trajectory can be regarded as a 0–pseudo–trajectory, the set of all finite or infinite trajectories of f which belong entirely to S will be denoted by $\mathbf{Tr}(f, S, 0)$. Since a trajectory is also a γ–pseudo–trajectory for any $\gamma > 0$, $\mathbf{Tr}(f, S, 0) \subset \mathbf{Tr}(f, S, \gamma)$. The inclusion is strict because not every pseudo–trajectory is a trajectory.

Pseudo–trajectories arise naturally due to the presence of roundoff error in computer calculations of trajectories, though accumulated roundoff error can rapidly destroy any meaningful connection between a computed pseudo-trajectory and an original trajectory. The concept of shadowing introduced below provides an alternative, more practical form of comparison of trajectories and pseudo–trajectories.

Pseudo-trajectories often arise as a result of roundoff errors during computation of real trajectories. Moreover, accumulation of roundoff errors may result in quick loss of any relation between computed pseudo–trajectory and original trajectory. The notion of shadowing given below provides an alternative and more practical way to compare trajectories with pseudo–trajectories.

The system f is said to be *shadowing with positive parameters α and β* on a closed subset K of Ω if for any given finite γ–pseudo–trajectory $\mathbf{y} \subseteq K$ with $0 \leq \gamma \leq \beta$ there exists a trajectory $\mathbf{x} = \{x_n\} \in \mathbf{Tr}(f)$ such that

$$\rho(x_n, y_n) \leq \alpha\gamma$$

for all n for which $\mathbf{y}$ is defined.

Many systems with chaotic behavior are shadowing for appropriate α, β. From this it is often concluded that the behavior of a computed system reflects that of the original system, at least over finite time intervals, in the sense that there will always be some true trajectory near any observed, pseudo–trajectory. The classical examples of shadowing system are hyperbolic diffeomorphisms.

Theorem 5. *(a) For any dynamical system f the relation*

$$\mathbf{Sep}(\mathbf{Tr}(f), \mathbf{Tr}(\varphi)) \le C(f, \varphi)$$

holds.

(b) If the system f is shadowing on Ω with constants α and β then for any φ with $2d(f, \varphi) \le \beta$ the following relation holds

$$\mathbf{Sep}(\mathbf{Tr}(\varphi), \mathbf{Tr}(f)) \le 2\alpha d(f, \varphi) + d(\varphi, f).$$

This Theorem shows that the flow of the original system f is closely represented by the flow of the multi–valued discretization φ provided that φ is consistent, the system f is shadowing, and the graphs $\mathbf{Gr}(f), \mathbf{Gr}(\varphi)$ are close in the Hausdorff metric.

From Theorem 5 a number of useful corollaries follow; let us mention one of them. A point $\xi \in \mathbb{L}$ is called *cyclic* for $\varphi : \mathbb{L} \mapsto 2^{\mathbb{L}}$ if there exists a natural number p such that $\xi \in \varphi^p(\xi)$. Recall that an infinite trajectory $\mathbf{x} = x_0, x_1, \dots$ of f is called *recurrent* if for any $\varepsilon > 0$ there exists a natural N such that for any natural M, $\mathbf{Sep}(\mathbf{x}, \mathbf{x}_*) \le \varepsilon$ where $\mathbf{x}_* = x_M, \dots, x_{M+N-1}$. In other words, a trajectory is recurrent if it is approximated with a given accuracy by each sufficiently long subtrajectory.

Corollary 2. *(a) Let a trajectory $\mathbf{x} = x_0, x_1, \dots$ be recurrent for f and suppose that φ is a discretization. Then for any $\alpha > C(\varphi, f)$ there exists a periodic trajectory $\xi_1, \xi_2, \dots$ of φ satisfying*

$$\rho(x_n, \xi_n) \le \alpha, \qquad n = 0, 1, \dots \ .$$

(b) If the system f is shadowing on Ω with constants α, β and $\xi_0, \xi_1, \dots$ is a periodic trajectory of φ then there exists a recurrent trajectory $\mathbf{x}$ of f with

$$\rho(x_n, \xi_n) \le 2\alpha d(f, \varphi) + d(\varphi, f), \qquad n = 0, 1, \dots \ .$$

4.4. Concept of bi–shadowing

Now discuss an important point concerned the interpretation of shadowing property in numerical modelling. Theorem 5 is sufficient for comparison of properties of original system and its discretization would f described the dynamic of the system under consideration *precisely*. However, it is *not the case* in many practical situation. Often f itself should be considered as an approximation/idealization of a certain unknown in explicit form dynamical system g. Information about the system g can be usually summed as some qualitative properties of the mapping g, for instance continuity, and some quantitative estimates, concerning the proximity between f and g pretty often these are estimates of the uniform distance. In this situation it would be better to use instead statements like Theorem 5 assertion about similarity between the flows of a discretization φ of an ideal system f and a possible flow of a real, unknown explicitly, system g.

Here is more natural to use instead of the concept of shadowing the one of *bi–shadowing*. Recall once more that the shadowing is often interpreted as the confirmation of the fact that the behavior

of a computed system reflects that of the original system, at least over finite time intervals, in the sense that there will always be some true trajectory near any observed, pseudo–trajectory.

The inverse question as to whether every true trajectory can be approximated by some pseudo–trajectory is of no less practical importance. While any γ-pseudo-trajectory is possible in principle, only those belonging to some particular class $\mathcal{T}$ occur in practice. These might be generated by a discretization method being applied or arise from specific processes associated with computer arithmetic. Typically, only general characteristics of such pseudo–trajectories will be known rather than a complete definition of $\mathcal{T}$ itself. The problem of *inverse shadowing* with respect to such a class $\mathcal{T}$ is to determine whether every true trajectory of a given system f can be approximated by some pseudo–trajectories from $\mathcal{T}$. A discussion of direct and indirect shadowing can be found in [61, Appendix C].

The class of pseudo–trajectories $\mathcal{T}$ plays a somewhat different role in the two forms of shadowing. In the classical Shadowing Lemma, $\mathcal{T}$ consists of *all conceivable* pseudo–trajectories of f and is thus as large as possible. On the other hand, inverse shadowing should be compatible with more restricted classes $\mathcal{T}$, such as *the trajectories of a continuous mapping φ that is sufficiently C^0 close to f*. It will be shown that the two forms of shadowing with the respect to such classes are usually both present (see also [12–15]).

Define the distance between two mappings $f, g : \Omega \mapsto \Omega$ by

$$\rho_\infty(\varphi, f) = \sup_{x\in\Omega} \rho(\varphi(x), f(x)).$$

A dynamical system $f : \Omega \mapsto \Omega$ is said to be *bi–shadowing with positive parameters α and β on a subset K of Ω* if for any given finite pseudo-trajectory $\mathbf{y} = \{y_n\} \in \mathbf{Tr}(f, K, \gamma)$ with $0 \le \gamma \le \beta$ and any mapping $\varphi : \Omega \mapsto \Omega$ satisfying

$$\gamma + \rho_\infty(\varphi, f) \le \beta,$$

there exists a trajectory $\mathbf{x} = \{x_n\} \in \mathbf{Tr}(\varphi, \Omega, 0)$ such that

$$\rho(x_n, y_n) \le \alpha(\gamma + \rho_\infty(\varphi, f))$$

for all n for which $\mathbf{y}$ is defined.

Now, we can reformulate Theorem 5 in the following form:

Corollary 3. *If the system f is bi–shadowing on Ω with constants α, β then for any φ with $2d(f, \varphi) + \rho_\infty(\varphi, f) \le \beta$ the estimate*

$$\mathbf{Sep}(\mathbf{Tr}(\varphi), \mathbf{Tr}(g)) \le \alpha(2d(f, \varphi) + \rho_\infty(\varphi, f)) + \max\{d(\varphi, f), C(\varphi, f)\}$$

holds.

4.5. Semi–hyperbolic Lipschitz mappings

Applications of Theorem 3 above and of many other assertions of such kind require constructive theorems about bi–shadowing. Shadowing results (cf. [5, 15, 60, 62]) typically establish only direct shadowing and involve rather stringent assumptions such as that the dynamical system is generated by a hyperbolic diffeomorphism. Hyperbolicity, however, imposes far more structure on a dynamical system than is required for shadowing, for example the continuity of the splitting and the existence of invariant stable and unstable manifolds. In addition, as we have seen above there are many important and interesting dynamical systems which lack either the smoothness or the invertibility. Many useful properties of hyperbolic diffeomorphisms are retained by the semi–hyperbolic mappings

that were introduced in [15] for local diffeomorphisms (see also Anosov [3] and Sataev [63] where related concepts are discussed) and extended to Lipschitz mappings in [16].

Below is shown that semi–hyperbolicity of Lipschitz mappings implies bi–shadowing. A four-tuple $\mathbf{s} = (\lambda_s, \lambda_u, \mu_s, \mu_u)$ of nonnegative real numbers is called a *split* if

$$\lambda_s < 1 < \lambda_u \tag{16}$$

and

$$(1 - \lambda_s)(\lambda_u - 1) > \mu_s \mu_u. \tag{17}$$

Clearly, for any given λ_s, λ_u satisfying (16) the four-tuple $\mathbf{s}$ is a split if the product $\mu_s \mu_u$ is small enough.

Below Ω is a compact subset of $\mathbb{R}^d$ with non–empty interior. Let $\mathbf{s} = (\lambda_s, \lambda_u, \mu_s, \mu_u)$ be a split and K a compact subset of interior of Ω. A Lipschitz mapping $f : \Omega \mapsto \Omega$ is said to be $\mathbf{s}$*–semi-hyperbolic on the set* K if there exist positive real numbers k, δ such that for each $x \in K$ there exists a splitting (decomposition)

$$\mathbb{R}^d = E^s_x \oplus E^u_x,$$

with corresponding projectors P^s_x and P^u_x satisfying the following properties:

SH0. $\dim(E^s_x) = \dim(E^s_{f(x)})$ if $x, f(x) \in K$.
SH1. $\sup_{x \in K}\{|P^s_x|, |P^u_x|\} \le k$.
SH2. The inclusion

$$x + u + v \in \Omega$$

and the inequalities

$$\begin{aligned}
|P^s_{f(x)}\,(f(x+u+v) - f(x+\widetilde{u}+v))\,| &\le \lambda_s |u - \widetilde{u}|,\\
|P^s_{f(x)}\,(f(x+u+v) - f(x+u+\widetilde{v}))\,| &\le \mu_s |v - \widetilde{v}|,\\
|P^u_{f(x)}\,(f(x+u+v) - f(x+\widetilde{u}+v))\,| &\le \mu_u |u - \widetilde{u}|,\\
|P^u_{f(x)}\,(f(x+u+v) - f(x+u+\widetilde{v}))\,| &\ge \lambda_u |v - \widetilde{v}|
\end{aligned}$$

hold for all $x \in K$ with $f(x) \in K$ and all $u, \widetilde{u} \in E^s_x, v, \widetilde{v} \in E^u_x$ such that $|u|, |\widetilde{u}|, |v|, |\widetilde{v}| \le \delta$.

Note that continuity in x of the splitting subspaces E^s_x, E^u_x or of the projectors P^s_x, P^u_x is not assumed here, nor is invariance of the splitting subspaces, as is the case in the definition of hyperbolicity (of a diffeomorphism). The smooth $(\mathbf{s}, k)$–semi-hyperbolic mappings, as they were introduced in [15] are, clearly, $\mathbf{s}$–semi-hyperbolic for any split $\mathbf{s}_\varepsilon = (\lambda_s + \varepsilon, \lambda_u - \varepsilon, \mu_s + \varepsilon, \mu_u + \varepsilon)$, $\varepsilon > 0$. A hyperbolic (with respect to the Euclidean metric) diffeomorphism on K is semi-hyperbolic in the sense of the definition in [15] and hence in the above sense as a Lipschitz mapping for an appropriate triple $(\mathbf{s}, k, \delta)$.

The main result of this Section is that semi–hyperbolicity is sufficient to ensure bi–shadowing of a dynamical system generated by a Lipschitz mapping with respect to perturbed systems generated by continuous mappings.

Theorem 6. *Let $f : \Omega \mapsto \Omega$ be a Lipschitz mapping which is $\mathbf{s}$–semi–hyperbolic on a compact subset K of Ω with constants k, δ. Then it is bi–shadowing on K with parameters*

$$\alpha(\mathbf{s}, k) = k\,\frac{\lambda_u - \lambda_s + \mu_s + \mu_u}{(1 - \lambda_s)\,(\lambda_u - 1) - \mu_s \mu_u}$$

and

$$\beta(\mathbf{s}, k, \delta) = \delta k^{-1}\,\frac{(1 - \lambda_s)(\lambda_u - 1) - \mu_s \mu_u}{\max\{\lambda_u - 1 + \mu_s, 1 - \lambda_s + \mu_u\}}.$$

with respect to continuous mappings $\varphi : \Omega \mapsto \Omega$.

Other properties of semi–hyperbolic mappings see in [1, 2, 12–16, 18]

5. DISCRETIZATION OF PLANAR LINEAR SYSTEMS WITH INVARIANT MEASURE

An impression can appear that the difficulties of investigating discretized systems are intrinsic for systems with chaotic behavior only. However, that is not the case. Consider, for example, the problem of numerical modeling of the dynamics of a *linear mapping describing the rotation of the plane by an angle* ϑ:

$$\begin{array}{rcl} x & \mapsto & x_0 + (x - x_0)\cos\vartheta - (y - y_0)\sin\vartheta, \\ y & \mapsto & y_0 + (x - x_0)\sin\vartheta + (y - y_0)\cos\vartheta. \end{array} \tag{18}$$

When treating the results of the numerical modeling, we again encounter paradoxical results.

As can be seen from 5, the phase portrait of the mapping (6) differs drastically from that of the rotation mapping (18). The basic distinction consists in that the mapping (6) is dissipative. But this is precisely the property of the mapping (6) that can be justified theoretically. Nevertheless, the proof turns out to be not as easy as it might seem, and this illustrates how poorly we understand the geometry of discretized (even linear) mappings for the time being!

Theorem 7 (see [20]). *If* $\vartheta = 0, \pm\pi, \pm\frac{\pi}{2}$ *then every trajectory of the mapping (6) is eventually periodic with periods* $1, 2, 4$ *respectively.*

If $(x_0, y_0) = 0$ *and* $\vartheta \neq 0, \pm\pi, \pm\frac{\pi}{2}$ *then every trajectory of the mapping (6) eventually gets into the zero point.*

If $(x_0, y_0) \neq 0$, $(x_0, y_0) \in \mathbb{L}_{1,1}$ *and* $\vartheta \neq 0, \pm\pi, \pm\frac{\pi}{2}$ *then there exists a number* $r_0 > 0$ *such that every trajectory of the mapping (6) eventually gets into the disk of the radius*r_0 *centered at the point* (x_0, y_0).

The results of modeling by formulae (7) (see Figure 6) look closer to behavior of trajectories of the continuous rotation mapping T_ϑ, although qualitative distinctions can be seen in this case as well. The obtained pattern turns out to resemble the "rings of Saturn" with characteristic zones of condensation and rarefaction. To justify the obtained results formally and, in particular, to answer the questions:

- WHAT ARE THE SPATIAL DENSITIES OF CYCLES OF THE MAPPING (7) OR OF POINTS WITH EMPTY PREIMAGES?
- WHETHER TRAJECTORIES OF THE MAPPING (7) ARE ALL BOUNDED OR AMONG THEM THERE ARE THOSE ONES DIVERGING TO INFINITY?

etc., turned out to be substantially more difficult than in case of the mapping (6).

Investigation of the problem of invertibility of discretized plane rotations is fulfilled in the next Section.

5.1. Problem of invertibility of the roundoff discretization of linear mappings

Let as usually $\mathbb{Z}$ be the set of integers and $\mathbb{L}_{1,1} := \mathbb{Z} \times \mathbb{Z}$ be the lattice of points with integer coordinates in $\mathbb{R}^2$. Denote by $\text{ROUND}(x)$ the *roundoff operator to the nearest integer* on $\mathbb{R}$ defined by the equality

$$\text{ROUND}(x) = i \in \mathbb{Z} : \quad x - \frac{1}{2} \leq i < x + \frac{1}{2}, \qquad x \in \mathbb{R},$$

and analogously define the coordinate-wise roundoff operator on the plane $\mathbb{R}^2$:

$$\text{ROUND}(x) = (\text{ROUND}(x_1), \text{ROUND}(x_2)) \in \mathbb{L}_{1,1}, \qquad x = (x_1, x_2) \in \mathbb{R}^2.$$

The standard scalar product in $\mathbb{R}^2$ will be denoted below by $\langle\cdot,\cdot\rangle$.

Consider the linear mapping $A:\mathbb{R}^2\to\mathbb{R}^2$ and let $\mathcal{A}:\mathbb{L}_{1,1}\to\mathbb{L}_{1,1}$ be its *roundoff discretization* defined as

$$\mathcal{A}(x):=\textsc{round}(Ax),\qquad x\in\mathbb{L}_{1,1}. \tag{19}$$

In virtue of (19) the mapping $\mathcal{A}$ acts from $\mathbb{R}^2$ to $\mathbb{L}_{1,1}$; it can be also uniquely defined as such a mapping $\mathcal{A}:\mathbb{L}_{1,1}\to\mathbb{L}_{1,1}$ which satisfies the inclusion

$$\mathcal{A}(x)-Ax\in\Pi:=\left[-\frac{1}{2},\frac{1}{2}\right)\times\left[-\frac{1}{2},\frac{1}{2}\right),\qquad z\in\mathbb{L}_{1,1}. \tag{20}$$

The roundoff discretization is a particular case of the so-called Stetter or general discretizations. Recall [66] that the mapping $\mathcal{A}:\mathbb{L}_{1,1}\to\mathbb{L}_{1,1}$ is a *Stetter or general* discretization of the mapping Ax on $\mathbb{L}_{1,1}$ if it satisfies

$$\mathcal{A}(x)-Ax\in\overline{\Pi}:=\left[-\frac{1}{2},\frac{1}{2}\right]\times\left[-\frac{1}{2},\frac{1}{2}\right],\qquad z\in\mathbb{L}_{1,1}. \tag{21}$$

Remark that in contrast to (20) the mapping $\mathcal{A}$ is determined by (21) ambiguously.

In what follows we will be interested mainly in studying the properties of roundoff discretizations although we will not avoid also investigating the properties of general discretizations. One of the first problem arising during consideration of roundoff discretizations, the answer to which is not quite obvious, is that about their invertibility:

ARE THERE SUCH LINEAR MAPPINGS A WHOSE ROUNDOFF DISCRETIZATIONS $\mathcal{A}$ ARE INVERTIBLE ON $\mathbb{L}_{1,1}$?

Examples given below show that the answer to the above question is positive.

Example 7. Let

$$A=\begin{bmatrix} m & n\\ p & q\end{bmatrix},\qquad m,n,p,q\in\mathbb{Z},$$

where $mq-np=\pm1$. Then the mapping A acts on the lattice $\mathbb{L}_{1,1}$, i.e. $\mathcal{A}\equiv A$, and is invertible on it.

Example 8 (see [55]). Let

$$A=\begin{bmatrix}\gamma & -1\\ 1 & 0\end{bmatrix},\qquad \gamma\in\mathbb{R},\quad \gamma>0.$$

Then the roundoff discretization $\mathcal{A}$ of the mapping A is invertible on the lattice $\mathbb{L}_{1,1}$, and the mappings $\mathcal{A}$, $\mathcal{A}^{-1}$ are of the form:

$$\mathcal{A}(x)=\begin{bmatrix}\textsc{round}(\gamma x_1)-x_2\\ x_1\end{bmatrix},\quad \mathcal{A}^{-1}(x)=\begin{bmatrix}x_2\\ -x_1+\textsc{round}(\gamma x_2)\end{bmatrix},\qquad x=\begin{bmatrix}x_1\\ x_2\end{bmatrix}\in\mathbb{L}_{1,1}. \tag{22}$$

So as in the previous case the assertion of Example 8 is proved by direct verification. Remark that Example 8 is not so evident as it seems on the first glance. So, the representation (22) for $\mathcal{A}$ is not valid for any negative rational γ of the form $-\frac{m}{2n}$; this fact directly follows from the following relation

$$\textsc{round}(-x)\not\equiv-\textsc{round}(x),\qquad x\in\mathbb{R}^2. \tag{23}$$

The same relation (23) implies also that $\mathcal{A}^{-1}(x)\not\equiv\textsc{round}(A^{-1}x)$ for any negative rational γ of the form $-\frac{m}{2n}$.

Example 9. Let

$$A = \begin{bmatrix} \frac{4}{5} & -\frac{3}{5} \\ \frac{3}{5} & \frac{4}{5} \end{bmatrix}.$$

Then the roundoff discretization $\mathcal{A}$ of the mapping A is invertible on the lattice $\mathbb{L}_{1,1}$ and $\mathcal{A}^{-1}(x) = \text{ROUND}(A^{-1}x)$.

Example 10. Let

$$A = \begin{bmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{bmatrix},$$

and the angle ϑ is such that the values of $\cos\vartheta$ and $\sin\vartheta$ are irrational. Then the roundoff discretization $\mathcal{A}$ of the mapping A is not invertible on the lattice $\mathbb{L}_{1,1}$.

In connection with the examples presented above there arise a question about characterization of linear mappings A possessing an invertible roundoff discretization $\mathcal{A}$. One should keep in mind that it is not clear a'priori whether the mapping invertible to $\mathcal{A}$ is the roundoff discretization of some linear mapping of $\mathbb{R}^2$ into itself. If the answer is positive then we can pose further question: whether it is possible to choose the corresponding mapping from the class of the roundoff discretizations or from the broader class of general (Stetter) discretizations? So, the main question now can be formulated more specific:

ARE THERE SUCH LINEAR MAPPINGS A WHOSE ROUNDOFF DISCRETIZATIONS $\mathcal{A}$ ARE INVERTIBLE ON $\mathbb{L}_{1,1}$ IN THE CLASS OF ROUNDOFF DISCRETIZATIONS OF LINEAR MAPPINGS, I.E. FOR WHICH THERE EXISTS SUCH A LINEAR MAPPING B WHOSE ROUNDOFF DISCRETIZATION $\mathcal{B}$ SATISFIES $\mathcal{A} \circ \mathcal{B} \equiv I$?

A more general form of the same question is as follows:

ARE THERE SUCH LINEAR MAPPINGS A WHOSE ROUNDOFF DISCRETIZATIONS $\mathcal{A}$ ARE INVERTIBLE ON $\mathbb{L}_{1,1}$ IN THE CLASS OF GENERAL STETTER DISCRETIZATIONS, I.E. FOR WHICH THERE EXISTS SUCH A LINEAR MAPPING B WHOSE STETTER DISCRETIZATION $\mathcal{B}$ SATISFIES $\mathcal{A} \circ \mathcal{B} \equiv I$?

The following Lemma demonstrates that the class of linear mappings possessing invertible discretizations is quite restrictive.

Lemma 3. *Let $\mathcal{A}$ and $\mathcal{B}$ be general (Stetter) discretizations of linear mappings A and B, respectively, and let $\mathcal{A} \circ \mathcal{B} = I$ on $\mathbb{L}_{1,1}$. Then $AB = I$ and $\det A = \det B = 1$.*

Lemmata 4, 5 and Theorem 8 below even more restrict the class of linear mappings discretizations of which are invertible (in the class of discretizations of linear mappings).

Derive some auxiliary relations. Let A, B be linear mappings in $\mathbb{R}^2$ and $\mathcal{A}, \mathcal{B}$ be their general (Stetter) discretizations. Let also

$$(\mathcal{A} \circ \mathcal{B})(x) \equiv x, \qquad x \in \mathbb{L}_{1,1}. \tag{24}$$

Then by Lemma 3 $B = A^{-1}$ and one may write down the chain of equalities:

$$\begin{aligned} x &= (\mathcal{A} \circ \mathcal{B})(x) = \mathcal{A}(\mathcal{B}(x)) = A\mathcal{B}(x) + \alpha(\mathcal{B}(x)) = \\ &= AA^{-1}x + A\beta(x) + \alpha(\mathcal{B}(x)) = x + A\beta(x) + \alpha(\mathcal{B}(x)), \end{aligned}$$

where

$$\alpha(x) = A(x) - \mathcal{A}(x), \qquad \beta(x) = B(x) - \mathcal{B}(x).$$

Then

$$A\beta(x) + \alpha(\mathcal{B}(x)) = 0$$

or

$$\beta(x) = -A^{-1}\alpha(\mathcal{B}(x)), \qquad x \in \mathbb{L}_{1,1}. \tag{25}$$

Since (24) is equivalent to

$$(\mathcal{B} \circ \mathcal{A})(x) \equiv x, \qquad x \in \mathbb{L}_{1,1}, \tag{26}$$

then the latter relation implies:

$$\alpha(x) = -A\beta(\mathcal{A}(x)), \qquad x \in \mathbb{L}_{1,1}. \tag{27}$$

Define now two sets:

$$\mathbb{A} = \left\{ z \in \mathbb{R}^2 : \ z = \alpha(x), x \in \mathbb{L}_{1,1} \right\}, \qquad \mathbb{B} = \left\{ z \in \mathbb{R}^2 : \ z = \beta(x), x \in \mathbb{L}_{1,1} \right\}.$$

Then from the definitions of the mappings $\alpha(\cdot)$ and $\beta(\cdot)$ it follows that

$$\mathbb{A}, \mathbb{B} \subseteq \overline{\Pi}. \tag{28}$$

On the other hand, relations (27) and (25) imply

$$\mathbb{A} \subseteq -A\mathbb{B}, \qquad \mathbb{B} \subseteq -A^{-1}\mathbb{A},$$

from which due to invertibility of the mapping A (see Lemma 3)

$$\mathbb{A} = -A\mathbb{B}. \tag{29}$$

Finally, from (28) and (29) we get

$$\mathbb{A} \subseteq -\overline{\Pi} \cap A\overline{\Pi}, \qquad \mathbb{B} \subseteq -\overline{\Pi} \cap A^{-1}\overline{\Pi}. \tag{30}$$

Lemma 4. *If the mapping A has an invertible discretization and*

$$A \neq \pm I, \qquad A \neq \pm \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

then the polygons $\overline{\Pi} \cap A\overline{\Pi}$ and $\overline{\Pi} \cap A^{-1}\overline{\Pi}$ are proper subsets of the closed square $\overline{\Pi}$, and therefore $\mathbb{A}$ and $\mathbb{B}$ are not dense in $\overline{\Pi}$ (see Figure 14).

The next Lemma specifies the conditions under which the set $\mathbb{A}$ (or $\mathbb{B}$) is dense in Π, and thus by Lemma 4 the corresponding discretization of the mapping A is not invertible.

Lemma 5. *Let $\mathcal{A}$ be the roundoff discretization of the mapping A. Then the set $\mathbb{A}$ is dense in Π (or in $\overline{\Pi}$) if and only if the following Vladimirov Condition (see, e.g., Theorem 9 below and also [68, p. 28]) holds:*

(A1) *the matrix $A \in \mathbb{R}^{2\times 2}$ is such that the columns of the matrix $\left[I, A^{\mathrm{T}}\right] \in \mathbb{R}^{2\times 4}$ are rationally independent.*

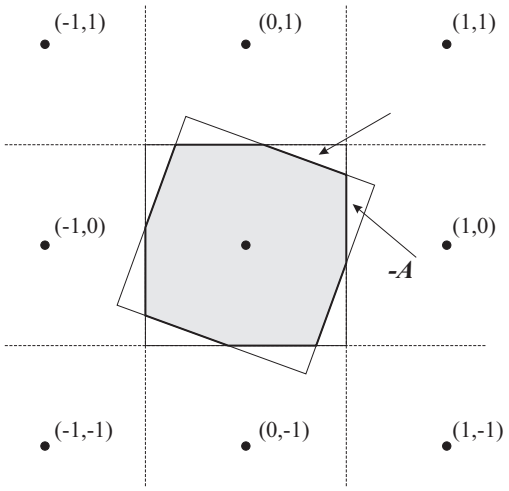

Fig. 14. Disposition of the sets Π and $A\Pi$ for a linear mapping A with an invertible discretization.

Lemmata 4 and 5 imply the following

Theorem 8. *Let $\mathcal{A}$ be the roundoff discretization of the linear mapping A, and $\mathcal{B}$ be a general (Stetter) discretization of the linear mapping B such that $\mathcal{A} \circ \mathcal{B} = I$. Then* $\det A = 1$ *and there are nonzero vectors $z, w \in \mathbb{L}_{1,1}$ such that $A^T z = w$.*

Remark that the statement of Example 10 immediately follows from the above Theorem.

So, results of this Section show that discretizations of linear mappings may be invertible in exceptional cases. This means that in the course of numerical modelling of dynamics of invertible continuous processes *we inevitable will face the loss of information.* How big this loss of information and to which extent this loss of information may influence the results of modelling is another problem to partial study of which the next Sections are devoted.

5.2. Algebra of quasiperiodic frequency measurable sets

Answering the questions related to the density of cycles of the mapping (7) or the points which have no preimages requires developing and justifying probability theoretical constructions for measures that are *not countably additive*, i.e. for those ones not obeying the now classical Kolmogorov's axiomatics of the probability theory. Since the approach proposed in [68, 69] is applicable for different situations connected with the analysis of discretizations of continuous mappings, we will describe it in this Section in more details.

One of the basic properties of the continuous rotation mapping T_θ to preserve the Lebesgue measure is lost when we pass to its discretization $\text{ROUND}(T_\theta(x))$ whatever "reasonable" and "natural" definition of a measure on the lattice $\mathbb{L}_{1,1} \subset \mathbb{R}^2$ is used. At the same time, it is clear that the measure-preserving property plays a key role in the study of "frequency" properties of many continuous dynamical systems. That is why one of the first questions which arise is *whether it is possible to define a measure on the lattice $\mathbb{L}_{1,1}$ in such a manner that the measure be preserved by the mapping* $\text{ROUND}(T_\theta(x))$? Only in the case of a positive answer to this question one may hope to obtain substantive assertions on the properties of discretized continuous mappings.

The answer to the question turned out to be positive, and we shall give a brief description of the construction of an appropriate measure. For a quadruple of integers a_1, a_2 and $l_1, l_2 > 0$ denote by $\mathcal{R}(a_1, a_2, l_1, l_2)$ the rectangle $([a_1, a_1 + l_1) \times [a_2, a_2 + l_2)) \cap \mathbb{L}_{1,1}$. A set $A \subset \mathbb{L}_{1,1}$ is called *frequency*

measurable if there exists a number $\mathcal{F}(A)$, $0 \le \mathcal{F}(A) \le 1$ (*a frequency* of the set A) such that

$$\sup_{l_1,l_2\ge N,\, a_1,a_2} \left| \frac{\#\left(A \bigcap \mathcal{R}(a_1,a_2,l_1,l_2)\right)}{l_1\, l_2} - \mathcal{F}(A) \right| \to 0 \quad \text{for} \quad N \to +\infty.$$

Example 11. Obvious examples of frequency measurable sets are given by the lattice $\mathbb{L}_{1,1}$ for which $\mathcal{F}\left(\mathbb{L}_{1,1}\right) = 1$, and by any finite subset $A \subset \mathbb{L}_{1,1}$ of the lattice; for any such set $\mathcal{F}(A) = 0$.

The totality $\mathcal{S}$ of frequency measurable sets is closed under unification of a finite number of pairwise disjoint sets and complementation of a set to the lattice $\mathbb{L}_{1,1}$, but is not an algebra since this totality is not closed under intersection of sets (and thus, the totality $\mathcal{S}$ is not a σ–algebra). In addition, the functional $\mathcal{F} : \mathcal{S} \to [0;1]$ is a finitely additive probability measure and hence the triplet $(\mathbb{L}_{1,1}, \mathcal{S}, \mathcal{F})$ can be regarded as a *probability space which does not obey the axiomatics of Kolmogorov* [43, 64].

Now, give an example of a less trivial and more essential to our goals class of frequency measurable sets.

Theorem 9. *Let $\mathcal{L} \in \mathbb{R}^{\infty\times 2}$ be a matrix consisting of countable number of rows and two columns whose every $(m \times 2)$–submatrix Λ is such that*

(A1) *the columns of the matrix $\left[I_2, \Lambda^{\mathrm{T}}\right]$ are rationally independent,*

with I_2 standing for the identity matrix of the second order.

Then for any Jordan measurable set[3] *$G \subset [0;1)^m$ and for any $(m\times 2)$–submatrix Λ of the matrix $\mathcal{L}$, the set*

$$Q_m(G,\Lambda) \equiv \{x \in \mathbb{L}_{1,1} :\ \Lambda x \in G + \mathbb{Z}^m\} \tag{31}$$

is frequency measurable and its frequency $\mathcal{F}\left(Q_m(G,\Lambda)\right)$ coincides with the m–dimensional Lebesgue measure $\operatorname{mes} G$ *of the set G. In addition, the totality $\mathcal{Q}(\mathcal{L})$ of all the sets of kind (31) is an algebra (entirely consisting of frequency measurable sets).*

Note that the matrices Λ satisfying Condition (A1) form a set of full Lebesgue measure and, hence, are typical in the sense.

5.3. Rules of interpretation

From the property that the frequency $\mathcal{F}$ of any finite subset of the lattice equals zero, it follows that the functional $\mathcal{F}$ is substantially finitely additive. From here, it also can be seen that the frequency $\mathcal{F}$ is concentrated on infinite subsets of the lattice. These rather unusual circumstances aggravate perception of frequency measurable sets with frequencies different from 0 and 1. That is why, when carrying out a computer experiment, one needs rules for interpreting assertions like this: *an event $A \subseteq \mathbb{L}_{1,1}$ holds with a frequency or probability p, $0 \le p \le 1$, with respect to the probability measure $\mathcal{F}$.* An appropriate interpretation should be consistent with the definition of the frequency $\mathcal{F}$ and one of its simplest version is as follows.

For sequences of integers $l_1^{(n)}, l_2^{(n)} > 0$ and $a_1^{(n)}, a_2^{(n)}$ with $l_1^{(n)}$ and $l_2^{(n)}$ unboundedly increasing, denote by $p(A;\mathcal{R}^{(n)})$ the proportion of those points x of the rectangle

$$\mathcal{R}^{(n)} = [a_1^{(n)}, a_1^{(n)} + l_1^{(n)}) \times [a_2^{(n)}, a_2^{(n)} + l_2^{(n)}) \bigcap \mathbb{L}_{1,1}$$

for which the inclusion $x \in A$ is valid. Then

[3] Recall that a set $G \subset \mathbb{R}^n$ is called *Jordan measurable* if it is bounded and its boundary ∂G satisfies $\operatorname{mes}_n \partial G = 0$, where $\operatorname{mes}_n(\,\cdot\,)$ is the n–dimensional Lebesgue measure of a set.

an event $A \subseteq \mathbb{L}_{1,1}$ holds with a probability p with respect to the probability measure $\mathcal{F}$ if and only if the limiting relation $\lim_{n\to\infty} p(A;\mathcal{R}^{(n)}) = p$ holds independently of the choice of the sequences $a_1^{(n)}, a_2^{(n)}$.

When calculating the proportions $p(A;\mathcal{R}^{(n)})$, the sets $\mathcal{R}^{(n)}$ are not allowed to be chosen arbitrarily (satisfying the only condition that their cardinalities unboundedly grow). Nevertheless, there may be a certain arbitrariness in choosing the sets $\mathcal{R}^{(n)}$ for calculating the proportions $p(A;\mathcal{R}^{(n)})$. For example, one can take, as the sets $\mathcal{R}^{(n)}$, the unions of finite numbers of pairwise disjoint rectangles with jointly increasing lengths of sides. Some other ways of choosing the sets $\mathcal{R}^{(n)}$ are also allowed.

5.4. Quantizers and ergodic properties of discretized linear systems

The sets belonging to some algebra $\mathcal{Q}(\mathcal{L})$ have a fractal-wise structure, although they do not formally possess the self-similarity property and that is why are not fractals. These sets will be called *$\mathcal{L}$–quasiperiodic*. The quasiperiodic sets naturally arise in considering a sufficiently wide class of discretizations of linear mappings, quantized linear systems. Adduce appropriate definitions.

A mapping $Q : \mathbb{R}^2 \to \mathbb{L}_{1,1}$ is called a *quantizer* if it commutates with the additive group of translations of the lattice $\mathbb{L}_{1,1}$ (i.e. $Q(x+z) = Q(x) + z$ for all $x \in \mathbb{R}^2$, $z \in \mathbb{L}_{1,1}$), and the full preimage $Q^{-1}(0) \subset \mathbb{R}^2$ of the zero vector is a Jordan measurable set. The superposition $Q \circ T : \mathbb{L}_{1,1} \to \mathbb{L}_{1,1}$ of a nonsingular linear mapping T with a quantizer Q will be called *a quantized linear system*.

The most interesting quantizer is the roundoff operator ROUND to the nearest point of the lattice $\mathbb{L}_{1,1}$ which generates a quantized linear system ROUND $\circ\, T_\theta$. Note that the operator of truncation TRUNC is not a quantizer and the mapping (6) is not a quantized linear system. It is precisely by this reason that there is a difference in behavior of the discretized mappings (6) and (7).

The following theorem shows that the frequency functional $\mathcal{F}$, when being considered on an appropriately chosen algebra of frequency measurable quasiperiodic sets, is the finitely additive probability measure invariant for a quantized linear system, and thus a quantized linear system possesses strong ergodic properties.

Theorem 10. *Let the matrix $T \in \mathbb{R}^{2\times 2}$ of the quantized linear system $Q \circ T$ be such that the rows of the infinite dimensional matrix $\mathcal{T}^+ \equiv \operatorname{blockcol}_{k\geq 0}\left(T^k\right) \in \mathbb{R}^{\infty\times 2}$ are rationally independent. Then the operator $Q \circ T : \mathbb{L}_{1,1} \to \mathbb{L}_{1,1}$ is measurable with respect to the algebra of $\mathcal{T}^+$–quasiperiodic sets, preserves the frequency $\mathcal{F}$ on this algebra and possesses a mixing property, i.e.*

$$(Q \circ T)^{-1}(A) \in \mathcal{Q}(\mathcal{T}^+), \qquad \mathcal{F}\left((Q \circ T)^{-1}(A)\right) = \mathcal{F}(A)$$

and

$$\lim_{k\to+\infty} \mathcal{F}\left((Q \circ T)^{-k}(A) \bigcap B\right) = \mathcal{F}(A)\,\mathcal{F}(B)$$

for all $A, B \in \mathcal{Q}(\mathcal{T}^+)$.

The collection of facts concerning the quantized linear systems is not exhausted by the assertions of Theorem 10 only. By appropriately choosing algebras of frequency measurable quasiperiodic sets (depending on the quantizer and Q matrix T), one can also establish some other properties of discretizations of linear systems. For instance, it turns out that quantization errors $E_k \equiv (I - Q) \circ T \circ (Q \circ T)^{k-1} : \mathbb{L}_{1,1} \to Q^{-1}(0)$ are *mutually independent, uniformly distributed on the*

set $Q^{-1}(0)$ and measurable with respect to the algebra of $\mathcal{T}^+$–quasiperiodic sets which allows to compute frequency characteristics of the Nth deviation

$$\Delta^{(N)}(x) \stackrel{\text{def}}{=} T^N x - (Q \circ T)^N(x) = \sum_{k=1}^{N} T^{N-k} E_k(x).$$

of trajectories of the quantized linear system $Q \circ T$ from those of the original linear system T. One can also compute the frequency of the points of the lattice $\mathbb{L}_{1,1}$ with full preimages consisting of a given number of points, etc.

Now, discuss some corollaries from the above results for the quantized linear system with the roundoff quantizer $Q = \text{ROUND}$ and the rotation matrix $T = T_\theta$ with the angle $\theta \in (0; \pi/2)$. An exhaustive answer to the above asked question — is there a possibility of defining a measure on the lattice $\mathbb{L}_{1,1}$ which is preserved by the mapping $\text{ROUND} \circ T$? — was given by Theorem 10. A rather satisfactory answer to another question — what are the spatial densities of cycles of the mapping (7) or points with empty preimages? — will be given by Theorem 11. Denote by an $\sigma_n(\mathcal{R}, \theta)$ the relative proportion of those points of a rectangle $\mathcal{R}$ whose n-th full preimages with respect to the mapping $\text{ROUND}(T_\theta x)$ are empty. Then the following theorem is valid.

Theorem 11. *Let the rotation angle $\theta \in (0; \pi/2)$ be such that the rows of all the nonnegative powers of the matrix T_θ are rationally independent. Then, uniformly with respect to a_1 and a_2*

$$\lim_{\min\{l_1,l_2\}\to\infty} \sigma_1(\mathcal{R}(a_1, a_2, l_1, l_2), \theta) = \sigma_1(\theta) \equiv \left(\sqrt{2}\,\cos\left(\theta - \frac{\pi}{4}\right) - 1\right)^2.$$

From this theorem, it follows that, for sufficiently large rectangular fragments of the lattice $\mathbb{L}_{1,1} \subset \mathbb{R}^2$ the relative proportion of the points with empty full preimages is approximately equal to $1 - \sigma_1(\theta)$. The upper estimates for the relative proportion of the points belonging to cycles of the mapping $\text{ROUND}(T_\theta x)$ are given by the numbers $1 - \sigma_n(\theta)$, $n = 2, 3, \ldots$, which can be computed with the aid of explicit formulae.

The assertion of Theorem 11 seems sufficiently natural and, if we ignore a concrete shape of the formula, trivial in essence. Nevertheless, this impression is not completely true. The nontriviality of Theorem 11 is testified by the fact that *the theorem is no longer valid if, when calculating the relative proportions $\sigma_1(\mathcal{R}, \theta)$, arbitrary finite sets $\mathcal{R}$ with growing cardinalities are used* instead of the rectangles $\mathcal{R} = \mathcal{R}(a_1, a_2, l_1, l_2)$. In a, sense, Theorem 11 resembles the method of computing divergent integrals by means of taking their principal values.

As to the above asked question - whether the trajectories of the mapping (7) are all bounded, or among them there are those ones diverging to infinity? - we succeeded in obtaining a partial answer to the question. It turns out that *for the overwhelming, in the sense of the measure $\mathcal{F}$, majority of trajectories of the discretized mapping (7), their deviation from corresponding trajectories of the ideal system in N steps of evolution grows like $\sqrt{N}$ with N increasing.* This result follows from the theorem below.

Theorem 12. *Let the rotation angle θ be such that the rows of all the nonnegative powers of the matrix T_θ are rationally independent. Then the Euclidean deviation of trajectories of the quantized linear system $\text{ROUND} \circ T_\theta$ from those of the original linear system T_θ during N steps of evolution is characterized by the following asymptotical distributions: for any $\alpha \geq 0$ there exist the limits*

$$\lim_{N\to+\infty} \mathcal{F}\left\{x \in \mathbb{L}_{1,1} :\ \sqrt{\frac{12}{N}}\left|T_\theta^k(x) - (\text{ROUND} \circ T_\theta)^k(x)\right| > \alpha\right\} = \exp\left(-\alpha^2/2\right),$$

$$\lim_{N\to+\infty} \mathcal{F}\left\{x\in\mathbb{L}_{1,1}:\ \max_{1\le k\le N}\sqrt{\frac{12}{N}}\left|T_\theta^k(x)-(\text{ROUND}\circ T_\theta)^k(x)\right|>\alpha\right\}=\tau(\alpha),$$

where $\tau(\alpha)$ stands for the probability of the random event that the maximal Euclidean deviation of the two-dimensional standard Wiener process from the zero point on the time interval $[0;1]$ exceeds the level α.

Note that the set Θ of all the values θ satisfying the assumption of Theorem 12 has full Lebesgue measure, and moreover, the set$(0;\pi/2)\setminus\Theta$ formed by "pathological" rotation angles is countable. Also, note that for any $\theta\in\Theta$ the number θ/π is irrational.

The first of the limiting relations from Theorem 12 can informally be rephrased as follows: if, for a sufficiently large and fixed N, we calculate the empirical tail-distribution function

$$D_{\mathcal{R}}^{(N)}(\alpha)\equiv\frac{1}{\#(\mathcal{R})}\#\left\{x\in\mathcal{R}:\ \sqrt{\frac{12}{N}}\left|T_\theta^k(x)-(\text{ROUND}\circ T_\theta)^k(x)\right|>\alpha\right\}$$

of the aforementioned normed deviations for a large rectangle $\mathcal{R}$ on the lattice $\mathbb{L}_{1,1}$, then this distribution will be close to the theoretically predicted function $\exp(-\alpha^2/2)$ uniformly in $\alpha\ge 0$. This relation admits an experimental verification which for $N=10$ and the rectangle $\mathcal{R}(-25,-25,50,50)$ is in good agreement with the theoretical prediction.

6. ASYMPTOTICS OF INFORMATION LOSS UNDER DISCRETIZATION OF RANDOM PROCESSES

Physical measurements with the aid of digital equipment and computer simulations of dynamical systems involve a phase space which is the finite set of machine arithmetic and hence, are accompanied by distortion of functional properties of original continual objects under spatial discretization. Development of information theoretical techniques for estimating the consequences of modeling continual phenomena by their computer counterparts seems important in this circle of problems [48, 50]. In this Section the asymptotic analysis of losses of the information is carried out at an analogue-digital conversion of random data in arithmetics with the fixed point. In the presentation of this part we follow [30, 70].

6.1. Lattice discretization

Denote by $R:\mathbb{R}^n\to\mathbb{Z}^n$ the *roundoff operator* which maps a vector $u=(u_k)_{1\le k\le n}$ to the nearest node $R:\mathbb{R}^n\to\mathbb{Z}^n$ of the n-dimensional integer lattice $\mathbb{Z}^n$, where $\lfloor\cdot\rfloor$ is the floor integer part of a number. Clearly, R commutes with the additive group of translations of $\mathbb{Z}^n$,

$$R(u+z)=R(u)+z\quad\text{for all } u\in\mathbb{R}^n,\ z\in\mathbb{Z}^n,$$

and the preimage of the zero vector under the mapping is the half-open cube

$$V=R^{-1}(0)=[-1/2,1/2)^n. \tag{32}$$

For any $\varepsilon>0$, define the mapping $R_\varepsilon=\varepsilon\,R\circ\varepsilon^{-1}:\mathbb{R}^n\to\varepsilon\mathbb{Z}^n$ which can be interpreted as a model for the computer discretization procedure in fixed-point arithmetic with accuracy ε.

6.2. Entropy of discretized random vector

Let $\xi:\Omega\to\mathbb{R}^n$ be a random vector on a probability space $(\Omega,\mathcal{F},P)$. Suppose that its probability distribution is absolutely continuous with density $p:\mathbb{R}^n\to\mathbb{R}_+$ with respect to the n-dimensional

Lebesgue measure. Then the ε-discretization $\xi_\varepsilon = R_\varepsilon(\xi)$ of the vector is distributed with probabilities

$$p_\varepsilon(x) = \varepsilon^n \mathbf{E} p(x + \varepsilon\theta), \quad x \in \varepsilon\mathbb{Z}^n, \tag{33}$$

where θ is the random vector uniformly distributed over the cube (32) and $\mathbf{E}(\,\cdot\,)$ is the mathematical expectation. The entropy of the vector ξ_ε

$$H(\xi_\varepsilon) = -\sum_{x\in\varepsilon\mathbb{Z}^n} p_\varepsilon(x)\ln p_\varepsilon(x) \tag{34}$$

coincides with the information $I(\xi_\varepsilon;\xi)$ in this vector about ξ. From (33) and (34) it follows that

$$H(\xi_\varepsilon) = -n\ln\varepsilon - \varepsilon^n \sum_{x\in\varepsilon\mathbb{Z}^n} f_\varepsilon(x), \tag{35}$$

where

$$f_\varepsilon(x) = \mathbf{E} p(x+\varepsilon\theta)\ln \mathbf{E} p(x+\varepsilon\theta). \tag{36}$$

By the Jensen inequality and convexity of the function $u\ln u$ in $u \geq 0$, from (36) it follows that $f_\varepsilon(x) \leq \mathbf{E}(p(x+\varepsilon\theta)\ln p(x+\varepsilon\theta))$, and then by (35)

$$H(\xi_\varepsilon) \geq -n\ln\varepsilon - \varepsilon^n \sum_{x\in\varepsilon\mathbb{Z}^n} \mathbf{E}(p(x+\varepsilon\theta)\ln p(x+\varepsilon\theta)) = -n\ln\varepsilon + h(\xi), \tag{37}$$

where

$$h(\xi) = -\int\limits_{\mathbb{R}^n} p(x)\ln p(x)dx \tag{38}$$

is the differential entropy [56] of the vector ξ. To state the theorem below which gives an asymptotic estimate for the proximity of the inequality (37) to equality, some auxiliary notions are needed.

A function $f : \mathbb{R}^n \to \mathbb{R}$ is called *supersummable* if

$$\int_{\mathbb{R}^n} U_\varepsilon(f)(x)dx < +\infty$$

for some $\varepsilon > 0$, where the function $U_\varepsilon(f) : \mathbb{R}^n \to \mathbb{R}_+$ is defined by

$$U_\varepsilon(f)(x) = \sup_{y\in V} |f(x+\varepsilon y)|.$$

Note that the supersummability of a function $f : \mathbb{R}^n \to \mathbb{R}$ is a stronger property than its Lebesgue summability. For example, the function $f : \mathbb{R} \to \mathbb{R}_+$ of the form

$$f(x) = \sum_{r\geq 0} g\left(2^r\left(|x| - \sum_{k=1}^{r}\frac{1}{k}\right)\right),$$

where $g(x) = \max(0,\, 1-2|x|)$, is integrable by Lebesgue but not supersummable. Indeed, $\int_{\mathbb{R}} f(x)dx = 3/2$, although $U_{1/r}(f)(x) = 1$ for any $r \in \mathbb{N}$, $|x| > \sum_{k=1}^{r}\frac{1}{k}$, and therefore $\int_{\mathbb{R}} U_\varepsilon(f)(x)dx = +\infty$ for all $\varepsilon > 0$.

The role of the supersummability of a function $f : \mathbb{R}^n \to \mathbb{R}$ is clarified by the inequality

$$\varepsilon^n \sum_{x\in\varepsilon\mathbb{Z}^n} |f(x)| \leq \int_{\mathbb{R}^n} U_\varepsilon(f)(x)\,dx\ ,$$

due to which supersummability implies absolute summability of the above series for all sufficiently small ε.

For any $N \in \mathbb{Z}_+$, denote by F_N the class of N times continuously differentiable functions $f : \mathbb{R}^n \to \mathbb{R}$, Lebesgue summable together with all their partial derivatives up to order N and such that the N-th order partial derivatives are all supersummable. In particular, F_0 is the class of continuous supersummable functions.

Theorem 13. *Let the probability density function p of the random vector ξ satisfies the conditions:*

(a) *p is strictly positive and four times continuously differentiable everywhere in $\mathbb{R}^n$;*
(b) *$p \ln p \in F_4$;*
(c) *$(1+\ln p)\triangle p \in F_2$ where $\triangle = \sum_{k=1}^n \frac{\partial^2}{\partial x_k^2}$ is the Laplacian;*
(d) *the functions $U_\varepsilon(1+\ln p)U_\varepsilon\left(\frac{\partial^4 p}{\partial x_i \partial x_j \partial x_k \partial x_l}\right)$ and $U_\varepsilon\left(\frac{1}{p}\right)U_\varepsilon\left(\left(\frac{\partial^2 p}{\partial x_i \partial x_j}\right)^2\right)$ are Lebesgue summable for some $\varepsilon > 0$ and any $1 \le i,j,k,l \le n$.*

Then the entropy (34) is asymptotically related with the differential entropy (38) by

$$H(\xi_\varepsilon) = -n\ln\varepsilon + h(\xi) + \mathbf{D}(\xi)\,\varepsilon^2 + O(\varepsilon^4) \qquad \varepsilon \to +0,$$

where

$$\mathbf{D}(\xi) = \frac{1}{24}\int\limits_{\mathbb{R}^n} \frac{|\nabla p(x)|^2}{p(x)}dx = \frac{1}{24}\,\mathbf{E}\,|\nabla \ln p(\xi)|^2 \tag{39}$$

and $\nabla = \frac{\partial}{\partial x}$ denotes the gradient operator.

The quantity $\mathbf{D}(\xi)$ is further on called the *entropy shifting factor*. For example, suppose that the random vector ξ is Gaussian distributed with mathematical expectation μ and nonsingular covariance matrix Σ. That is, its probability density function is given by

$$p(x) = (2\pi)^{-n/2}(\det\Sigma)^{-1/2}\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right) \tag{40}$$

where $(\,\cdot\,)^T$ denotes the transpose. Any partial derivative $\frac{\partial^r p(x)}{\partial x_{i_1}\ldots\partial x_{i_r}}$ of the function p is the product of $p(x)$ and some algebraic polynomial in the coordinates of x. Then the conditions (a)–(d) of Theorem 13 are satisfied. Furthermore, (40) implies that $\nabla \ln p(x) = \Sigma^{-1}(x-\mu)$ and so the entropy shifting factor (39) is

$$\mathbf{D}(\xi) = \frac{1}{24}\,\mathbf{E}\left|\Sigma^{-1}(\xi-\mu)\right|^2 = \frac{1}{24}\,\mathbf{Tr}\Sigma^{-1},$$

and does not depend on μ (here, $\mathbf{Tr}(\,\cdot\,)$ denotes the trace of a matrix).

6.3. Information in discretized random vectors

Let ξ and η be two random vectors of dimensions n_1 and n_2, respectively. Suppose that the probability distribution of the n-dimensional vector

$$\zeta = \left[\begin{array}{c}\xi\\ \eta\end{array}\right] \tag{41}$$

is absolutely continuous with density $p : \mathbb{R}^n \to \mathbb{R}_+$ where $n = n_1 + n_2$. Then the marginal probability density functions $p_1 : \mathbb{R}^{n_1} \to \mathbb{R}_+$ and $p_2 : \mathbb{R}^{n_2} \to \mathbb{R}_+$ of the vectors ξ and η are given by

$$p_1(x) = \int\limits_{\mathbb{R}^{n_2}} p(x,y)dy, \quad p_2(y) = \int\limits_{\mathbb{R}^{n_1}} p(x,y)dx. \tag{42}$$

For any $\varepsilon > 0$, denote by ξ_ε and η_ε the discretizations of the vectors ξ and η onto the ε-grids $\varepsilon\mathbb{Z}^{n_1}$ and $\varepsilon\mathbb{Z}^{n_2}$ so that

$$\zeta_\varepsilon = R_\varepsilon(\zeta) = \left[\begin{array}{c}\xi_\varepsilon\\ \eta_\varepsilon\end{array}\right].$$

Then the information

$$I(\xi_\varepsilon;\eta_\varepsilon) = H(\xi_\varepsilon) + H(\eta_\varepsilon) - H(\zeta_\varepsilon)$$

in ξ_ε about η_ε does not exceed the information

$$I(\xi;\eta) = h(\xi) + h(\eta) - h(\zeta)$$

in ξ about η. Therefore, the non-negative quantity

$$L_\varepsilon(\xi,\eta) = I(\xi;\eta) - I(\xi_\varepsilon;\eta_\varepsilon) \tag{43}$$

measures the loss of the mutual information between ξ and η under the discretization of these vectors onto the corresponding ε-grids.

Theorem 14. *Let the probability density functions p and (42) of the vector (41) and its subvectors ξ and η satisfy the conditions of Theorem 13. Then the loss of information (43) is asymptotically described by*

$$L_\varepsilon(\xi,\eta) = \mathbf{L}(\xi,\eta)\,\varepsilon^2 + O(\varepsilon^4) \quad \text{for } \varepsilon \to +0,$$

where

$$\begin{aligned} \mathbf{L}(\xi,\eta) &= \mathbf{D}(\zeta) - \mathbf{D}(\xi) - \mathbf{D}(\eta) \\ &= \frac{1}{24}\,\mathbf{E}\left(|\nabla_{x,y}\ln p(\zeta)|^2 - |\nabla_x \ln p_1(\xi)|^2 - |\nabla_y \ln p_2(\eta)|^2\right) \end{aligned} \tag{44}$$

with $\nabla_x = \frac{\partial}{\partial x}$, $\nabla_y = \frac{\partial}{\partial y}$ and $\nabla_{x,y} = \begin{bmatrix} \nabla_x \\ \nabla_y \end{bmatrix}$ denoting the gradient operators.

The non-negative quantity $\mathbf{L}(\xi,\eta)$ is further on called the *information loss factor*. For example, suppose that the vector (41) is Gaussian distributed with zero mean and nonsingular covariance matrix

$$\Sigma = \begin{bmatrix} \Phi & \Pi \\ \Pi^T & \Psi \end{bmatrix} \tag{45}$$

where

$$\Phi = \mathbf{E}(\xi\xi^T), \quad \Psi = \mathbf{E}(\eta\eta^T), \quad \Pi = \mathbf{E}(\xi\eta^T).$$

Since Σ is positive definite, so are the covariance matrices Φ and Ψ, and the matrix

$$G = \Phi^{-1/2}\Pi\Psi^{-1/2}$$

is contracting which implies the nonsingularity of the matrices $I_{n_1} - GG^T$ and $I_{n_2} - G^TG$.

Theorem 15. *For the subvectors ξ and η of the Gaussian vector (41) with nonsingular covariance matrix (45), the information loss factor (44) is computed as*

$$\mathbf{L}(\xi,\eta) = \frac{1}{24}\left(\mathbf{Tr}\left(\Phi^{-1}GH_2G^T\right) + \mathbf{Tr}\left(\Psi^{-1}G^TH_1G\right)\right), \tag{46}$$

where

$$H_1 = \left(I_{n_1} - GG^T\right)^{-1}, \quad H_2 = \left(I_{n_2} - G^TG\right)^{-1}. \tag{47}$$

In particular, if ξ and η are scalar Gaussian random variables with variances $\Phi, \Psi > 0$ and correlation coefficient $G \in (-1,1)$, the formulas (46)–(47) yield

$$\mathbf{L}(\xi,\eta) = \frac{1}{24}\left(\frac{1}{\Phi} + \frac{1}{\Psi}\right)\frac{G^2}{1-G^2}.$$

6.4. Entropy and information for discretized Gaussian sequences

Let $Z = (\zeta_k)_{-\infty<k<+\infty}$ be a n-dimensional stationary Gaussian sequence with zero mean and covariance function $\sigma_k = \mathbf{E}\left(\zeta_k \zeta_0^T\right)$. For any $N \in \mathbb{N}$, the initial fragment

$$Z_N = (\zeta_k)_{1\le k\le N} \tag{48}$$

of the sequence is a Gaussian distributed vector with block Toeplitz covariance matrix

$$\Sigma_N = \operatorname*{block}_{1\le j,k\le N} (\sigma_{j-k}).$$

Suppose that the matrix spectral density

$$S(\omega) = \sum_{k=-\infty}^{+\infty} \sigma_k \, \mathrm{e}^{ik\omega}$$

of the sequence Z is continuous and positive definite for any $\omega \in [-\pi, \pi]$. Then by the well-known property of Toeplitz forms [36]

$$\lim_{N\to+\infty} \left(\frac{1}{N}\mathbf{Tr}\Sigma_N^{-1}\right) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathbf{Tr}(S(\omega))^{-1} d\omega.$$

Hence, there exists the average value of the entropy shifting factor $\mathbf{D}(Z_N) = \frac{1}{24}\mathbf{Tr}\Sigma_N^{-1}$ of the fragment (48) per unit of time,

$$\mathcal{D}(Z) = \lim_{N\to+\infty} \left(\frac{1}{N}\,\mathbf{D}(Z_N)\right) = \frac{1}{48\pi}\int_{-\pi}^{\pi} \mathbf{Tr}(S(\omega))^{-1}\, d\omega. \tag{49}$$

Let the Gaussian sequence Z be described by the state-space equations

$$s_{k+1} = As_k + Bv_k, \tag{50}$$
$$\zeta_k = Cs_k + Dv_k, \tag{51}$$

where A, B, C, D are constant matrices of appropriate dimensions, with A asymptotically stable (its spectral radius $\rho(A) < 1$) and DD^T nonsingular; $(v_k)_{-\infty<k<+\infty}$ is the sequence of mutually independent m-dimensional Gaussian distributed vectors with zero mean and identity covariance matrix. Associate with (50)–(51) the matrix algebraic Riccati equation

$$Q = AQA^T + BB^T - \Lambda\Theta\Lambda^T, \tag{52}$$
$$\Lambda = (BQC^T + BD^T)\,\Theta^{-1}, \tag{53}$$
$$\Theta = CQC^T + DD^T, \tag{54}$$

A solution Q of this equation is called *stabilizing* if the matrix is symmetric and positive semidefinite, and $\rho(A - \Lambda C) < 1$. As is well-known [51], such a solution of the Riccati equation (52)–(54) is unique.

Theorem 16. *For the stationary Gaussian sequence Z governed by (50)–(51), the average entropy shifting factor (49) is computed as*

$$\mathcal{D}(Z) = \frac{1}{24}\,\mathbf{Tr}\left(\Theta^{-1}(CMC^T + I_n)\right), \tag{55}$$

where the matrix M is the solution of the Lyapunov matrix algebraic equation

$$M = (A - \Lambda C)M(A - \Lambda C)^T + \Lambda\Lambda^T, \tag{56}$$

and the matrices Λ and Θ are determined by the admissible solution Q of the Riccati equation (52)–(54).

Note that the quantity $\mathcal{D}(Z)$ defined by the equations (52)–(56) is a function F of the matrices A, B, C, D from (50)–(51). Hence, if the elements of the sequence Z are partitioned into subvectors ξ_k and η_k of dimensions n_1 and n_2 so that

$$\zeta_k = \left[\begin{array}{c} \xi_k \\ \eta_k \end{array} \right],$$

then the average entropy shifting factors of the stationary Gaussian sequences $X = (\xi_k)_{-\infty<k<+\infty}$ and $Y = (\eta_k)_{-\infty<k<+\infty}$ take the form

$$\mathcal{D}(X) = F(A, B, C_1, D_1), \quad \mathcal{D}(Y) = F(A, B, C_2, D_2)$$

where C_k and D_k are the blocks of the matrices C and D consisting of n_k rows respectively,

$$C = \left[\begin{array}{c} C_1 \\ C_2 \end{array} \right], \quad D = \left[\begin{array}{c} D_1 \\ D_2 \end{array} \right].$$

This allows to compute the average information loss factor of the sequences X and Y per unit of time as

$$\begin{aligned} \mathcal{L}(X,Y) &= \mathcal{D}(Z) - \mathcal{D}(X) - \mathcal{D}(Y) = \\ &= F(A, B, C, D) - F(A, B, C_1, D_1) - F(A, B, C_2, D_2). \end{aligned}$$

7. ANALYSIS OF MOIRE

Results of this Section were mainly published in [38].

7.1. General structure of moirés

Below μ is a big parameter. For any positive integer n denote by $L(n; f, \mu)$ the Lebesgue set

$$L(n; f, \mu) = \{(x, y) \in \mathcal{Z} : n - 1/2 \le \mu f(x, y) \le n + 1/2\}.$$

Let $\mathbb{Z}$ be the set of integers and $q \in \mathbb{Z}$ be a fixed prime number. Denote

$$L_q(n, f, \mu) = \bigcup_{k\in\mathbb{Z}} L(qk + n, f, \mu), \qquad n = 0, 1, \ldots, q - 1.$$

and let $\mathbb{L}_{h,\alpha}$ be a rectangular lattice in the plane with steps $h_x = \alpha h$ and $h_y = h/\alpha$ in x and y directions; the number α is considered below as given. Introduce the sets

$$\mathcal{L}_q(n; f, h, \alpha, \mu) = L_q(n; f, \mu) \bigcap \mathbb{L}_{h,\alpha}.$$

Associate with each element (x_0, y_0) of the lattice $\mathbb{L}_{h,\alpha}$ the rectangle

$$Q(x_0, y_0) = \{(x, y) : |x - x_0| < \alpha h/2, |y - y_0| < h/(2\alpha)\}.$$

Let now $\mathcal{M}$ be a subset of the lattice $\mathbb{L}_{h,\alpha}$. Denote by $Q(\mathcal{M})$ the union of rectangles $Q(x_0, y_0)$ for all points (x_0, y_0) from the set $\mathcal{M}$. Let $\chi \in (0, 1]$. The set $M \subseteq \Re^2$ with a finite and positive Lebesgue measure is said to be χ–*representable* with the set $\mathcal{M} \subseteq \mathbb{L}_{h,\alpha}$, if the inequality $\operatorname{mes}(Q(\mathcal{M}))/\operatorname{mes}(M) < 1 + \chi$ is valid. If μ is fixed, then the sets $\mathcal{L}_q(n, f, h, \alpha, \mu)$ represent the corresponding Lebesgue sets $L_q(n; f, \mu)$ as better as smaller is h. But χ not necessarily tends to 1, if decreasing of h is accompanied with rather fast increasing of μ (below usually $\mu h = \sigma$, where σ is fixed). We will say that the sets, $\mathcal{L}_q(n; f, h, \alpha, \sigma/h)$ *asymptotically represent an auxiliary function* $G(x, y)$ *in a domain* Ω, if for each sufficiently small h there exist $\xi = \xi(h)$, $\chi = \chi(h)$, such that $\mathcal{L}_q(n; f, h, \alpha, \sigma/h) \bigcap \Omega$ χ–represents the set $L_q(n; G(x, y) + \xi(h), \sigma/h)) \bigcap \Omega$ for each $n = 0, 1, \ldots, q-1$, where $\chi \to 1$ as $h \to 0$. Denote by $C_x(i, f)$ and $C_y(j, f)$ curves, which are defined at all integer i and j by the equalities $f'(x, y) = iq/(\alpha\sigma)$, $f'(x, y) = jq\alpha/\sigma$. Finally, denote by $D(f, \sigma)$ the totality of points $d(i, j; f, \sigma)$ where curves $C_x(i, f)$ intersect curves $C_y(j, f)$ transversally and bring to correspond to each such a point $d = (x_0, y_0) \in D(f, \sigma)$ the function

$$G_d(x, y) = f(x, y) - f'(x_0, y_0)(x - x_0) - f'(x_0, y_0)(y - y_0).$$

Theorem 17. *For each $d \in D(f)$ there exist $r > 0$, such that the sets $\mathcal{L}_q(n, f, h, \sigma/h)$ asymptotically represent the function $G_d(x, y)$ in the ring $|(x - x_0)^2 + (y - y_0)^2| < r^2$.*

Level sets of the auxiliary function G_d are those that are perceived by an observer as moirés; points from the set $D(f, \sigma)$ coincides with the centres of moirés. If in a certain domain the Gessian of the function $f(x, y)$ is not degenerated and changes rather slowly, then the system of curves $C_x(i, f)$ and $C_y(j, f)$ is sufficiently regular and the relations

$$d(i + 1, j; f, \sigma) - d(i, j; f, \sigma) \sim (q/(\alpha\sigma))(f''(d(i, j; f, \sigma)))^{-1}(1, 0),$$

$$d(i, j + 1; f, \sigma) - d(i, j; f, \sigma) \sim (q\alpha/\sigma)(f''(d(i, j; f, \sigma)))^{-1}(0, 1).$$

hold. The matrix $f''(d(i, j; f, \sigma))$ so as the matrix $(f''(d(i, j; f, \sigma)))^{-1}$ are symmetric; therefore two last relations allow to estimate the parameter α by formula $\alpha \sim (d_x/d_y)^{1/2}$, using the observable vectors

$$d_x = d(i + 1, j; f, \sigma) - d(i, j; f, \sigma), \qquad d_y = d(i, j + 1; f, \sigma) - d(i, j; f, \sigma).$$

If, on the contrary, the function $f(x, y)$ is unknown, but the system of curves $C_x(i, f)$, $C_y(j, f)$ is seen rather clearly, then the above relations can be used to estimate entries of the matrix $(\sigma f''(d(i, j; f, \sigma))^{-1}$.

7.2. Aliasing dimension

One of main general questions which appear in analysis of moirés (see Sections 2.3 and Figure 9) is the following:

HOW IRREGULAR MAY BE THE PICTURE GENERATED BY A GIVEN FUNCTION $f(\cdot)$ AND HOW TO DESCRIBE THIS IRREGULARITY, AND HOW TO CONNECT REGULARITY OR IRREGULARITY OF THE IMAGE WITH THE COMPLEXITY OF ITS GENERATING FUNCTION?

Below some principal results in this direction will be discussed, see details in [37]. The paper provides an insight to these problems. For the sake of simplicity the analysis is provided only for one-dimensional case (functions f of single scalar argument). The notion of aliasing dimension is introduced which naturally corresponds to the complexity of the visual images. The place of aliasing dimension in the family of other well-known complexity characteristics is discussed. Bilateral bounds for aliasing dimension polynomials $f(\cdot)$ are also provided.

Given a number x, denote by $[x]$ the integer part of x (i.e., the maximal integer number which is less than or equal to x). Denote also by $\{x\} = x - [x]$ the fractional part of x. Let $\mathcal{P}_k$ be the set of all k-th degree polynomials of one variable.

A binary sequence (u_i), where $i \in \mathcal{Z}$, is a *PG-sequence of k-th degree* (polynomial generated) if it may be represented as

$$u_i = [2\{f(i)\}].$$

for some $f \in \mathcal{P}_k$. The class of all such sequences (u_i) is denoted by $\mathcal{B}_k$. To describe the *diversity* of the class $\mathcal{B}_k$ we have to define a measure of such diversity. The natural way to do it is to determine what finite subsequences (words) can be extracted from the sequences of this class and what is the number of these words. The next problem is what is the maximal N such that *any* binary N-word may be found as a subsequence of some element of $\mathcal{B}_k$?

Note that all finite N-subsequences

$$(f(i), f(i+1), \ldots, f(i+N-1)), qquadN > k,$$

considered as points in $\mathbb{R}^N$ belong to some $(k+1)$-dimensional subspace P_k^N. For example, if $k = 1$ then

$$f(i+1) - f(i) \equiv f(i+2) - f(i+1).$$

However, this is not the case for N-subwords from $(u_i) \in \mathcal{B}_k$.

Let $u = (u_i)$, $i \in \mathbb{Z}$ be a binary sequence. Denote by $w_N(u)$ the set of all N-subwords of u and denote by $S_u(N)$ (subword complexity of u) the number of elements in $w_N(u)$. Denote by $B(k, N)$ the set of all N-words contained in all PG-sequences of k-th degree, i.e.

$$B(k, N) = \bigcup_{u \in \mathcal{B}_k} w_N(u).$$

Let $S(k, N)$ be the number of elements in $B(k, N)$. Note that $S(k, N) \le 2^N$.

Consider the largest integer N such that $S(k, N) = 2^N$. This N is called *aliasing dimension* of the class $\mathcal{P}_k$ and is denoted by $D(k)$.

The notion of aliasing dimension is to some extent similar to the notion of Vapnik-Chervonenkis dimension (VC-dimension) [67], it has much in common with ideas of Sturmian sequences (see, for example, [59]), linear recurring sequences [54], and symbolic dynamics [56]. Note especially the similarity with the fragmentary complexity as it was introduced in [17].

Theorem 18. *The aliasing dimension of the class $\mathcal{P}_k$ satisfy the estimates*

$$k + 1 \le D(k) \le 8k^2 \log k \tag{57}$$

It is not easy to calculate exactly aliasing dimension even for small k. Let for the beginning $k = 1$. Then the class $\mathcal{P}_1$ consists of linear polynomials.

Theorem 19. *The aliasing dimension $D(1)$ of the class $\mathcal{P}_1$ is equal to four.*

Let $k = 2$. Then the class $\mathcal{P}_2$ consists of quadratic polynomials. We could not calculate it so an attempt was made to establish a lower bound for $D(2)$ by numerical experiments. In these experiments for each N "random" polynomials $ax^2 + bx + c$ binary N-words were generated until all possible 2^N combinations of N-words are exhausted. The largest N found with this method was 13: any binary word of length 13 may be generated by the appropriate quadratic polynomial. However, the computer search failed to generate 32 "unused" binary words of length 14. Other modifications of the program (using integer arithmetic with enlarged precision) produced exactly the same set of unused words. For example, the program failed to generate the binary word "00001100011110" using quadratic polynomials. Thus there are strong reasons to believe that $D(2) = 13$ though the strict proof of that is yet unavailable.

REFERENCES

1. Al-Nayef, A., Diamond, P., Kloeden, P., Kozyakin, V., and Pokrovskii, A., *Bi-Shadowing and Delay Equations*, Dynamics & Stability of Systems, 1996, vol. 10, pp. 121–129.
2. Al-Nayef, A., Diamond, P., Kloeden, P., and Pokrovskii, A., *Expansivity of Nonsmooth Functional Differential Equations*, J. Math. Anal. Applns., 1997, vol. 208, pp. 453–461.
3. Anosov, D., *Geodesic Flows on Closed RIemannian Manifolds with Negative Curvature*, Proc. Steklov Inst. Math., 1967, vol. 90, pp. 1–235, in Russian.
4. Asarin, E.A., Kozyakin, V.S., Krasnosel'skii, M.A., and Kuznetsov, N.A., *Analysis of the Stability of Asynchronous Discrete Systems*, Moscow: Nauka, 1992, in Russian.
5. Beyn, M.J., *On the Numerical Approximation of Phase Portraits Near Stationary Points*, SIAM J. Numer. Anal., 1987, vol. 24, no. 5, pp. 1095–1113.
6. Blank, M., *Small Perturbations of Chaotic Dynamical Systems*, Russian Math. Surveys, 1989, vol. 44, no. 6, pp. 1–33.
7. Blank, M., *Discretness and Continuity in Problems of Chaotic Dynamics*, no. 161 in Translation of Mathematical Monographs, Providence: American Mathematical Society, 1997.
8. Bollobas, B., *Random Graphs*, London: Academic Press, 1985.
9. Boyarsky, A. and Gora, P., *Why Computers Like Lebesgue Measure*, Comp. Math. Applns, 1988, vol. 16, pp. 321–329.
10. Burtin, Y.D., *On a Simple Formula for Random Mappings and its Applications*, J. Appl. Prob., 1980, vol. 17, pp. 403–414.
11. Diamond, P. and Kloeden, P., *Spatial Discretization of Mappings*, Comp. Math. Applns., 1993, vol. 25, pp. 85–94.
12. Diamond, P., Kloeden, P., Kozyakin, V., Krasnosel'skii, M., and Pokrovskii, A., *Periodic Trajectories of Nonsmooth Perturbations of Systems with Chaotic Behavior*, Automat. Remote Control, 1995, vol. 56, no. 5, pp. 637–643.
13. Diamond, P., Kloeden, P., Kozyakin, V., Krasnosel'skii, M., and Pokrovskii, A., *Roughness of Trajectories of Dynamical Systems with Respect to Hysteresis Disturbances*, Soviet Math. Dokl., 1995, vol. 52, no. 1, pp. 138–140.
14. Diamond, P., Kloeden, P., Kozyakin, V., Krasnosel'skii, M., and Pokrovskii, A., *Robustness of Dynamical Systems to a Class of Nonsmooth Perturbations*, Nonlinear Analysis, TMA, 1996, vol. 26, no. 2, pp. 351–361.
15. Diamond, P., Kloeden, P., Kozyakin, V., and Pokrovskii, A., *Computer Robustness of Semi-Hyperbolic Mappings*, Random and Computational Dynamics, 1995, vol. 3, no. 1–2, pp. 57–70.
16. Diamond, P., Kloeden, P., Kozyakin, V., and Pokrovskii, A., *Expansivity of Semi-Hyperbolic Lipschitz Mappings*, Bulletin of the Australian Mathematical Society, 1995, vol. 51, no. 2, pp. 301–308.
17. Diamond, P., Kloeden, P., Kozyakin, V., and Pokrovskii, A., *On the Fragmentary Complexity of Symbolic Sequences*, Theoretical Computer Science, 1995, vol. 148, no. 1, pp. 1–17.
18. Diamond, P., Kloeden, P., Kozyakin, V., and Pokrovskii, A., *Semihyperbolic Mappings*, J. Nonlinear Science, 1995, vol. 5, no. 5, pp. 419–431.
19. Diamond, P., Kloeden, P., Kozyakin, V., and Pokrovskii, A., *A Model for Roundoff and Collapse in Computation of Chaotic Dynamical Systems*, Math. Comput. Simulation, 1997, vol. 44, no. 2, pp. 163–185.
20. Diamond, P., Kloeden, P., Kozyakin, V., and Pokrovskii, A., *Boundedness and Dissipativity of Truncated Rotations on Uniform Planar Lattices*, Math. Nachr., 1998, vol. 191, pp. 59–81.
21. Diamond, P., Kloeden, P., and Pokrovskii, A., *Numerical Modelling of Toroidal Dynamical Systems with Invariant Lebesgue Measure*, Journal of Complexity, 1993, vol. 7, pp. 415–422.

22. Diamond, P., Kloeden, P., and Pokrovskii, A., *Interval Stochastic Matrices and Simulation of Chaotic Dynamics*, in *Chaotic Numerics*, P. Kloeden and K.J. Palmer, eds., Contemporary Mathematics, pp. 203–216, Providence: American Mathematical Society, 1994.

23. Diamond, P., Kloeden, P., and Pokrovskii, A., *Multivalued Spatial Discretization of Dynamical Systems*, in *Proceedings of the Centre for Mathematics and its Applications of the Australian National University*, G. Martin and H. Thompson, eds., vol. 33, 1994 pp. 61–70.

24. Diamond, P., Kloeden, P., and Pokrovskii, A., *Weakly Chain Recurrent Points and Spatial Discretizations of Dynamical Systems*, Random & Computational Dynamics, 1994, vol. 2, no. 1, pp. 97–110.

25. Diamond, P., Kloeden, P., and Pokrovskii, A., *Analysis of an Algorithm for Computing Invariant Measures*, Nonlin. Anal. TMA, 1995, vol. 24, no. 3, pp. 323–336.

26. Diamond, P., Kloeden, P., and Pokrovskii, A., *Interval Stochastic Matrices, a Combinatorical Lemma and the Computation of Invariant Measures of Dynamical Systems*, Journal on Dynamics and Differential Equations, 1995, vol. 7, no. 2, pp. 341–364.

27. Diamond, P., Kloeden, P., Pokrovskii, A., and Vladimirov, A., *Collapsing Effects in Numerical Simulation of a Class of Chaotic Dynamical Systems and Random Mappings with a Single Attracting Centre*, Physica D, 1995, vol. 86, pp. 559–571.

28. Diamond, P., Kloeden, P.E., Klemm, A.D., and Pokrovskii, A., *Basin of Attraction of Cycles of Space Discretizations of Dynamical Systems with SRB Invariant Measure*, Journal of Statistical Physics, 1996, vol. 84, no. 3/4, pp. 713–733.

29. Diamond, P. and Pokrovskii, A., *Statistical Laws for Computational Collapse of Discretized Chaotic Mappings*, Journal of Bifurcation and Chaos in Applied Sciences and Engineering, 1996, vol. 6, no. 12A, pp. 2389–2399.

30. Diamond, P. and Vladimirov, I., *Higher-Order Terms in Asymptotic Expansion for Information Loss in Quantized Random Processes*, Circuits, Systems, and Signal Processing, 2001, vol. 20, no. 6, pp. 677–693.

31. Erber, T. and Gavelek, D., *The Iterative Evolution of Complex Systems*, Physica A, 1991, vol. 177, pp. 394–400.

32. Ethier, S. and Kurtz, T., *Markov Processes: Characterization and Convergence*, New York: Wiley, 1986.

33. Gerchbakh, I.B., *Epidemic Processes on a Random Graph: Some Prelimenary Results*, J. Appl. Prob., 1977, vol. 14, pp. 427–438.

34. Gnedenko, B.V. and Kolmogoroff, A.N., *Limit Distributions for Sums of Independent Random Variables*, Moscow–Leningrad: Gostekhizdat, 1949, in Russian.

35. Gregobi, C., Ott, E., and Yorke, J.A., *Roundoff–Induced Periodicity and the Correlation Dimension of Chaotic Attractors*, Phys. Rev. A, 1988, vol. 34, pp. 3688–3692.

36. Grenander, U., *Toeplitz Forms and Their Applications*, 2nd ed., Chelsea Pub Co, 1984.

37. Izmailov, R. and Vladimirov, A., *Dimension of Aliasing Structures*, Computers & Graphics, 1993, vol. 17, no. 5, pp. 539–547.

38. Izmailov, R.N. and Pokrovskii, A.V., *Asymptotic Classification of Aliasing Structures*, Journal of Applied Mathematics and Stochastic Analysis, 1992, vol. 5, no. 3, pp. 193–204.

39. Kifer, Y., *Random Perturbations of Dynamical Systems*, Boston: Birkhäuser, 1988.

40. Klemm, A. and Pokrovskii, A.V., *Random Mappings with a Single Absorbing Center and Combinatorics of Discretizations of the Logistic Mapping*, J. of Appl. Math. and Stoch. Anal., 1999, vol. 12, no. 3, pp. 205–221.

41. Kloeden, P.E. and Palmer, K.J., Eds., *Chaotic Numerics*, *Contemporary Mathematics*, vol. 172, Providence, Rhode Island: American mathematical Society, 1993, an International Workshop on the Approximation and Computation of Complicated Dynamical Behavior, July 12–16, 1993, Deakin University, Geelong, Australia.

42. Kolchin, V.F. and Khokhlov, V.I., *On a Number of Cycles in a Random Non-Equiprobable Graph*, Discrete Mathematics, 1990, vol. 2, no. 3, pp. 137–145, in Russian.

43. Kolmogoroff, A., *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Berlin–Göttingen–Heidelberg: Springer–Verlag, 1933.

44. Kozyakin, V., *A Short Introduction to Asynchronous Systems*, in *Proceedings of the Sixth International Conference on Difference Equations. Augsburg, Germany 2001: New Progress in Difference Equations*, B. Aulbach, S. Elaydi, and G. Ladas, eds., Boca Raton: CRC Press, 2004 pp. 153–166.

45. Kozyakin, V.S., Kuznetsov, N.A., and Pokrovskii, A.V., *Transients in Quasi–Controllable Systems. Overshooting, Stability and Instability*, in *Preprints of the 12th World IFAC Congress, Sydney, Australia, 18–23 July, 1993*, 1993 pp. 465–468.

46. Kozyakin, V.S., Kuznetsov, N.A., and Pokrovskii, A.V., *A New Model for Statistics of the First Recurrent Moments and Cycle Lengths in Discretizations of Dynamical Systems*, CADSEM Report 96–006, Deakin University, Geelong, Australia, 1996.

47. Kozyakin, V.S., Kuznetsov, N.A., and Pokrovskii, A.V., *A Phenomenological Model of the Statistics of Cycle Lengths and the First Recurrence Moments in Discretizations of Dynamic Systems*, Soviet Math. Dokl., 1996, vol. 54, no. 1, pp. 524–527.

48. Kozyakin, V.S., Kuznetsov, N.A., Pokrovskii, A.V., and Vladimirov, I.G., *Some Problems in Analysis of Discretizations of Continuous Dynamical Systems*, Nonlinear Analysis, TMA, 1997, vol. 30, no. 2, pp. 767–778.

49. Kozyakin, V.S. and Pokrovskii, A.V., *The Asymptotic Behavior of Elementary Symmetric Functions on a Probability Distribution*, Journal of Appl. Math. and Stoch. Anal., 2001, vol. 14, no. 3, pp. 237–248.

50. Kuznetsov, N. and Kloeden, P., *The Problem of Information Stability in Computer Studies of Continuous Systems*, Math. Comput. Simulation, 1993, vol. 43, no. 2, pp. 143–158.

51. Lancaster, P. and Rodman, L., *Algebraic Riccati Equations*, New York: Oxford University Press, 1995.

52. Lax, P., *Approximation of Measure Preserving Transformations*, Communications on Pure and Applied Mathematics, 1971, vol. 24, pp. 133–135.

53. Levy, Y.E., *Some Remarks about Computer Studies of Dynamical Systems*, Physics Letters, 1982, vol. A88, pp. 1–3.

54. Lidl, R. and Niederreiter, H., *Introduction to Finite Fields and their Applications*, Cambridge: Cambridge University Press, 1986.

55. Lowenstein, J., Hatjispyros, S., and Vivaldi, F., *Quasi-Periodicity, Global Stability and Scaling in a Model of Hamiltonian Round-off*, Chaos, 1997, vol. 7, no. 1, pp. 49–66.

56. Martin, N.F.G. and England, J.W., *Mathematical Theory of Entropy*, *Encyclopedia of Mathematics and its Applications*, vol. 12, London: Addison-Wesley, 1981.

57. The MathWorks, Inc., Natick, *MATLAB. Reference Guide*, 1992.

58. Melo, W. and van Strien, S., *One-Dimensional Dynamics*, Berlin: Springer–Verlag, 1993.

59. Morse, M. and Hedlund, G.A., *Symbolic Dynamics II: Sturmian Trajectories*, American Journal of Mathematics, 1940, vol. 62, pp. 1–42.

60. Palis, J. and de Melo, W., *Geometric Theory of Dynamical Systems. An Introduction*, Heidelberg: Springer–Verlag, 1982.

61. Pilyugin, S.Y., *The Space of Dynamical Systems with the C^0–Topology*, *Lecture Notes in Mathematics*, vol. 1571, Heidelberg: Springer–Verlag, 1994.

62. Ruelle, D., *Elements of Differentiable Dynamics and Bifurcations Theory*, Boston: Academic Press, Inc., 1989.

63. Sataev, E.A., *Invariant Measures for Hyperbolic Maps with Singularities*, Russian Math. Surveys, 1992, vol. 47, no. 1, pp. 191–251.

64. Shiryaev, A.N., *Probability*, Moscow: Nauka, 1989, in Russian.

65. Stepanov, V.E., *Random Mappings with a Single Attracting Centre*, Theory Prob. Appl., 1971, vol. 16, pp. 155–161.

66. Stetter, H.J., *Analysis of Discretization Methods for Ordinary Differential Equations*, Berlin: Springer–Verlag, 1976.

67. Vapnik, V.N. and Chervonenkis, A.J., *On Uniform Convergence of Relative Frequencies of Events to their Probabilities*, Theory of Probability and its Applications, 1971, vol. 16, pp. 264–280.

68. Vladimirov, I., *Quantized Linear Systems on Integer Lattices: Frequency-based Approach. Part I.*, CAD-SEM Report 96–032, Deakin University, Geelong, Australia, 1996.

69. Vladimirov, I., *Quantized Linear Systems on Integer Lattices: Frequency-based Approach. Part II.*, CAD-SEM Report 96–033, Deakin University, Geelong, Australia, 1996.

70. Vladimirov, I.G., *Asymptotics of Information Loss under Discretization of Random Processes*, Doklady Mathematics, 1999, vol. 60, no. 2, pp. 300–303.