Instability, Evolution and Morphogenesis

S. Vakulenko [1], **D. Grigoriev** [2]

[1] North-Western Institut of Printing, S. Petersburg and Institute for Mechanical Engineering Problems of Acad. Sci, S. Petersburg, Russia

[2]IRMAR Université de Rennes 1, Beaulieu, 35042, Rennes, France

Abstract

We introduce new random dynamical systems generalizing neural networks with random sources. We study homeostasis of such system. Namely, following the viability theory, we suppose that there is a domain D in the phase space such that if the system state leaves D, the system will be destroyed.

Under some assumptions, we show that a generic system of such type is, in a sense, unstable under fluctuations. For a system with fixed parameters, the system state leaves D within the time T with a probability $P(T)$ such that $P(T) \to 0$ as $T \to \infty$. However, such systems can survive for large times, i.e., $P(T) > \delta > 0$ for all times, if the system parameters evolve in time.

Some arguments show that if fluctuations are, in a sense, strong, the parameters should be discrete. This allows to connect this evolution problem with theory of complexity and to show that the problem of survival may be very difficult, at least NP-hard.

We consider some morphogenesis problems for genetic networks. We show that these networks are capable to construct any spatio-temporal patterns. As an illustration, the segmentation problem in *Drosophila* is considered and the pattern stability problem is investigated.

Using some recent ideas for NP-complete problems, we formulate, as a hypothesis, "Freedom Principle": if a system has sufficiently many internal parameters to adjust, then the survival is possible, namely, there exists an effective heuristic algorithm of parameter evolution such that $P(T) > \delta > 0$ for all times T.

Contents

1. INTRODUCTION

1.1 Homeostasis problem. Circuits. Circuit evolution

Our start point is a remark from [43], where M. Gromov and A. Carbone formulated the following problem: "Homeostasis of an individual cell cannot be stable for a long time as it would be destroyed by random fluctuations within and out of cell. There is no adequate mathematical formalism to express the intuitively clear idea of replicative stability of dynamical systems" ([43], p.40).

This assertion contains two hypothesis. First, that functioning of biological systems are unstable under random perturbations. Second, these systems can be stabilized by replication (evolution).

The goal of this paper is to formulate mathematically and prove these hypothesis for some classes of systems important in biology, chemistry and other applications. We introduce a measure of homeostasis stability under random perturbations. For some important classes of systems we show that almost all individual systems with *fixed* parameters are unstable, in a sense, for large times T, however, populations of *evolving* systems with changing (from time to time) parameters can be stable even as $T \to \infty$. Our approach to this homeostasis problem uses probabilistic methods and theoretical computer science approaches. The goal is to show mathematically that this approach explains main properties of biological evolution, for example, gene code existence, replication with seldom mutations, existence of great evolution tree etc. (for biological consequences see Conclusion).

The homeostasic concept was proposed by Claude Bernard [12]: "La fixité du milieu intérieur est la condition d'une vie libre et indépendante." (" Constance of the internal environment is the condition for a free and independent life"). This is the underlying principle: homeostasis means supporting of life functions of a system. Biological molecules and chemical mechanisms in the cell are fragile. Thus, in order to support their functioning, main characteristics of the cell (temperature, pressure, pH, reagent concentrations) must be within a narrow domain [3, 11] independently of external medium oscillations. For example, the temperature of a human body must lie within $35 - 42C^0$. Sharp changes in the external medium can destroy the system. Biological, economical and social systems can survive only when their states stay within some precribed domains (we denote these domains by Π).

Basing on these ideas we considered some known and new mathematical models. These models contain a dynamical component and a stochastical part describing a random environment. For such models a natural measure of the stochastical stability can be introduced. This measure is a probability $P_T(\Pi)$ that for $t \in [0, T]$ the system state (that can evolve in time) stays in the domain Π. This measure is well known and studied [86]. For brevity, if the system state stays within Π for $t \in [0, T]$, we say that our system survives on $[0, T]$. Dynamical systems with such admissible domains Π are well studied [5, 6, 7, 8].

Besides this stability measure, in this paper the idea of a "generic" system plays an important role. Two concepts of genericity are known.

Suppose a system depends on parameters $\mathcal{P}$. Following standard ideas [46, 15] of differential topology, we say that a property holds for a generic system if this property holds for an open dense set in the space of possible values of the parameters $\mathcal{P}$. Another approach is to introduce a measure μ on the set of values $\mathcal{P}$. Then "generic" property is valid for all values $\mathcal{P}$ besides, maybe, a set S such that $\mu(S) = 0$. In the other words, this property holds for almost all $\mathcal{P}$ with respect to μ (an interesting discussion on these two concepts one can find in [52]). We shall mainly use the second approach.

For the considered systems we show their instability *if their parameters $\mathcal{P}$ are fixed*. More precisely, we show that the survival probability $P_T(\Pi) \to 0$ as $T \to \infty$ for a generic system. For some important particular class of the systems (genetic circuits), this property holds for any circuits and the probability $P_T(\Pi)$ can be estimated.

The second main idea is that system evolution can stabilize unstable systems. If we consider a set of unstable systems with parameters $\mathcal{P}_i(t)$, which can change from time to time, then the limit of the survival probability $P_T(\Pi)$ as $T \to \infty$ may be positive. Briefly, a fixed system is almost always unstable but an infinite chain of evolving systems may be stable.

As a model of complex biological systems, we consider here circuits (networks). This choice is natural since last decades a large attention is given to problems of global organization, stability and evolution of complex networks such as protein and gene networks, networks of metabolic reactions, neural and economical circuits, Internet etc. (see [54, 55, 44], for an overview [2]).

The simplest mathematical model of such network is a (directed) graph. For example, for a gene network we can associate with this network a graph where a node describes a gene, the i-th node is connected with the j-th one if the i-th gene acts on j-th one. The evolution of such graphs can be considered as an algorithm adding or removing edges and nodes. Stability can be examined in different contexts. For example, we can examine how many edges (or nodes) must be eliminated in order to destroy connectivity of the graph. In biological applications, such an elimination may simulate mutations.

The first theory of graph evolution was developped by Erdos and Rényi [32, 2, 60]. They supposed that, at time moments $0, 1, 2, ...$, one adds to graph a new edge with probability p. This theory leads to a Gaussian distribution $\bar{C}(k)$ of the valency of a node. Recall that the valency of a node is the number of the nodes adjacent to this node. The quantity $\bar{C}(k)$ is the probability that a node has k adjacent nodes [2]. Recently it was found that real networks has another structure, namely, the so-called scale-free structure. Here $\bar{C}(k) \approx const\ k^{-\gamma}$, where the exponent γ lies usually within $(2, 3)$. Such networks have few number of nodes with a great valency, whereas the most of the nodes have a small valency.

Other interesting properties of graphs associated with actual biological, informational and economical systems can be described as follows. The graph diameter is restricted (the diameter is the maximal length of the shortest path connecting two nodes). The diameter defines the speed of dynamical processes in the circuit, thus a small diameter is useful to survive in the random environment. Moreover, studying of biological circuits showed that the averaged valency $\langle C \rangle$ has increased during evolution. Here $\langle C \rangle$ can be computed via $\bar{C}(k)$: $\langle C \rangle = \sum_{0 \le k \le \infty} k\bar{C}(k)$. Another property found experimentally for protein nets is that more connected proteins are more important for organisms: letality correlates with valency [55].

Stability of the free-scale structures is high with respect to a random attack when nodes to eliminate are chosen randomly. However, this stability is weak with respect to a "terroristic attack" (when one eliminates the most connected nodes).

The first evolution algorithm leading to the scale-free organization was proposed by Albert and Barabasi (see [2]). This algorithm uses the idea of a preferential attachment: the probability that a new edge is incident to the $i-$ th node is proportional to the valency of this node. Besides this algorithm, after it was investigated a number of other growth algorithms leading to the scale-free organisation, in particular, an algorithm proposed by [73] generates an hierarchical modular structure observed in methabolic networks.

However, metabolic reaction networks or gene networks cannot be described completely by a simple graph model. They constitute some complex dynamical system, where a scheme of interaction of substrats, ferments or genes can be associated with a graph. A part of the substrats enters this system from an external medium (input) and another part can be considered as an output (products). It is well known that these systems succesfully support an output independent of fluctuating input [61, 3].

It is difficult to propose a mathematical model for metabolic reactions, neural or gene interactions in detail. Circuit models were proposed ([47, 57, 37, 31, 66, 77, 76] among many others, see [80, 14] for an overview) to take into account theoretical ideas and experimental information. Some models [31, 82] use Boolean algebra (so-called boolean switch networks). Gene net models [66, 77] can be considered as a generalization of the famous Hopfield model of attractor neural network [47]. We consider mainly here a simplified version of [66], however, we take into account random fluctuations and evolution of network parameters. This network model is formulated in subsection 1.2, the model for network evolution is described in subsection 1.4. We assume that gene (protein) interaction is a pair one and it is defined by a $m \times m$ matrix $\mathbf{K}$. A directed graph can be associated with this matrix: two nodes are connected if the corresponding entry of $\mathbf{K}$ is not equal 0.

Our evolution model can be considered as a combination of graph evolution approaches described above and dynamical circuit models. It can be described as follows. One has a discrete set Y (finite or countable). This set can be considered as a "genetic code". One also has a map from Y to the set $\mathcal{P}$ of the network parameters (the number of genes m, the interaction matrix $\mathbf{K}$ and other). "Evolution" is a time continuous Markov process with values in Y, which changes y and, therefore, system parameters. Such approach is known in the neural and gene network theory [14], here we extend previous models addmitting that the network may grow unboundedly with time.

Our attention is focused on the following problems for genetic networks and their evolution.

I Stochastic stability of the networks with respect to fluctuations describing an internal noise and environment changes;

II Pattern capacity of the gene networks and patterning algorithms;

III We investigate stable evolution algorithms such that $\lim P_T(\Pi) > 0$ as $T \to \infty$, when a chain of evolving unstable systems has non-zero chances to survive for large times. In this part our goal is to explain, with the help of this stability approach, the main properties of evolution (why systems must make copies and the mutation probability is small, why the genetic code size cannot be bounded during the evolution process, why the evolution tree must be large and the networks should be scale-free, etc).

To study these problems we introduce a concept of *a priori computational complexity* of evolution problems. It allows us to apply some ideas and notions from complexity theory [35, 78, 41, 42, 69]. Indeed, it seems that many evolution problems are, in a certain sense, "complex". Roughly speaking, since "almost all" systems are unstable, to construct a stable system is a "complex" problem. In fact, for simplified models the evolution algorithm must resolve a complicated , maybe, even NP-complete problems (about NP-completeness see books [35, 69], for biological applications, see [72]). We formulate some such problems.

We also consider evolution algorithms for genetic networks. This question is connected with the graph evolution theory pioneered by Erdos and Rényi [32] since circuits can be associated, in a natural way, with directed graphs.

IV Problem of evolution speed: to estimate running time of evolution algorithms allowing to construct an unique structure with a large fitness ("biologically reasonable") among many possible structures. This problem is one of key points of evolution theory, was posed still by Charles Darvin. In fact, it is not obvious how to obtain a complex effectively working organ by an evolution using a local search, based on random mutations and selection. For example, the cell can be considered as "a biological computer" proceeding a complicated feedback [67, 3] and it is unclear how one can construct such a computer (consisting of unstable elements) by evolutionary mechanisms. It is clear only that this problem is very difficult.

"To suppose that the eye, with all its inimitable contrivances for adjusting the focus to different distances, for admitting different amounts of light, and for the correction of spherical and chromatic aberration, could have been formed by natural selection, seems, I freely confess, absurd in the highest degree. When it was first said that the sun stood still and the world turned round, the common sense of mankind declared the doctrine false; but the old saying of Vox populi, vox Dei, as every philosopher knows, cannot be trusted in science. Reason tells me, that if numerous gradations from a perfect and complex eye to one imperfect and simple, each grade being useful to its possessor, can be shown to exist; if further, the eye does vary ever so slightly and the variations be inherited, which is certainly the case; and if any variation or modification in the organ be ever useful to an animal under changing conditions of life, then the difficulty of believing that a perfect and complex eye could have been formed by natural selection, though insuperable by our imagination, can hardly be considered real. How a nerve comes to be sensitive to light, hardly concerns us more than how life itself first originated; but I may remark that, as some of the lowest organisms, in which nerves cannot be detected, are known to be sensitive to light, it does not seem impossible that certain elements in their tissues or sarcode should have become aggregated and developed into nerves endowed with special sensibility to its action." (Ch. Darwin, [24], Chapter VI).

This organ development problem has been considered in many books and papers, for example, see [74] and references in it. We propose a mathematical formalization of this problem. Then an answer can be connected with the following mathematical problem: for some NP-complete problems with a random structure, whether there exist algorithms [21, 35], which solve these problems within a relatively short running time for a certain subclass of instances. Recall that the main difficulty in NP-complete problems is that we need a global exhaustive search, which takes an extremely large running time. In framework of some simplified models we are going to show, by some algorithmic ideas invented recently (see [19] for a review) and results [89, 91, 92, 93, 94], that complex organs can be constructed by genetic circuits step by step, in a gradual manner. Such a conclusion confirms Ch. Darwin ideas (see above).

The main idea is as follows. In Section 4 we show that genetic networks can construct any spatio-temporal patterns [89, 91, 92, 93, 94]. However, it is not sufficient yet to answer to the formidable running time need, in general, for a local search algorithm to find a biologically reasonable structure among many of potentially existing structures. To overcome the key obstacle, we propose to address to recent works on NP-complete problems such as the satisfiability of boolean constraint problem (k-SAT), different graph problems, boolean programming problem (see [1, 19, 4, 25, 18]. It has been shown that, under some conditions, these problems can be effectively resolved.

To explain the main idea, recall that it is a fundamental conjecture of theoretical computing science that there exists no algorithm capable of solving NP-complete problem with inputs of size N in a time bounded by a polynomial of N. Consequently, when we are dealing with such a problem one necesseraly uses algorithms which take exponential times on some inputs.

Let us consider, for example, the k-SAT problem. The task is to satisfy m boolean constraints. Each constraint is a disjunction of k boolean variables $A_{i_1}, A_{i_2}, ...A_{i_k}$ or their negations taken from the list $A_1, ..., A_n$ of n boolean variables. The total search takes 2^n trials whereas the best known algorithms have running time $2^{\alpha n}$, where $\alpha << 1$ is a positive coefficient [23].

However, if we consider a *generic* problem, where constraints are chosen randomly, the situation becomes better. Then the solution set can be described as follows. If the relation $f = m/n$ is sufficiently large, the most of randomly constructed problems have no solutions. In the contrary, if f is small, a solution almost always exists and can be found by relatively simple algorithms (say, greedy ones). Such algorithms, based on heuristic ideas, have a polynomial in n running time, i.e., they are fast. Only in relatively small domain of values f the "generic" k-SAT problem is actually difficult: a solution can exist and, to seek this solution, the known algorithms require an exponentially running time $2^{\alpha n}$. The good situation, where f is large, describes the "free" case, where the alphabet $A_1, ...A_n$ is great enough with respect to the constraint number. The fact that then, at least in a generic situation, a solution can be found by a fast algorithm, can be named the "Freedom Principle".

In many works the genes were considered as boolean variables, a gene can be expressed (turned on) or not expressed (turned off) [3, 57, 80, 14]. We assert here, using these mathematical arguments, that, for biological evolution, the Freedom Principle works successfully since there is a formidable set of genes generating a large set of proteins with different properties. However, to use this richness, the Nature needs genetic circuits.

1.2 Systems under consideration

1.2.1 Network models

Let us consider first the model from [66, 76], which has the following form:

$$\frac{\partial u_i}{\partial t} = R_i \sigma(\sum_{j=1}^{m} K_{ij} u_j + \sum_{l=1}^{p} M_{il} \theta_l(x) - \eta_i) - \lambda_i u_i + d_i \Delta u_i, \tag{1.2.1}$$

where m is the number of genes included in the circuit, $u_i(x, t)$ is the concentration of the i-th protein, λ_i are the protein decay rates, R_i are some positive coefficients and d_i are the protein diffusion coefficients. We consider (1.2.1) in some bounded domain Ω with a boundary $\partial\Omega$.

The real number K_{ij} measures the influence of the j-th gene on the i-th one. The assumption that gene interactions can be expressed by a single real number per pair of genes is a simplification excluding complicated interactions between three, four and more genes. Clearly such interactions are possible, however in this case the problem becomes mathematically much more complicated.

The parameters η_i are activation thresholds and σ is a monotone function satisfying the following assumptions

$$\sigma \in C^{\infty}(\mathbf{R}), \lim_{z\to-\infty} \sigma(z) = 0, \lim_{z\to+\infty} \sigma(z) = 1, \tag{1.2.2}$$

$$|\frac{d\sigma}{dz}| < C\exp(-c|z|), \quad \sigma'(0) = 1. \tag{1.2.3}$$

The well known example is $\sigma(z) = (1 + \tanh z)/2$.

The functions $\theta_l(x)$ can be interpreted as densities of proteins associated with the maternal genes. The matrix M_{il} describes the action of the l-th maternal gene on the i-the gene. To (1.2.1), we should add the boundary and initial conditions. Notice that then eqs. (1.2.1) are well posed: solutions exist, they are a priori bounded, globally defined for all $t > 0$. This system possesses a global finite dimensional attractor. These general facts can be easily proved by standard methods.

This model takes into account the three fundamental processes: (a) decay of gene products (the term $-\lambda_i y_i$); (b) exchange of gene products between cells (the term with Δ) and (c) gene regulation and protein synthesis. Notice that if $d_i = 0$ this model of gene circuit can be considered as a Hopfield's neural network [47] with thresholds depending on x. The Hopfield system is a basic model in the theory of attractor neural networks [47].

Let us fix a function σ satisfying (1.2.2), (1.2.3) and functions θ_i. On the contrary, we consider $m, K_{ij}, M_{il}, \lambda_i, d_i, R_i$ and η_i as parameters $\mathcal{P}$ to be adjusted, $\mathcal{P} = \{m, \mathbf{K}, \mathbf{M}, \eta, \lambda, d, R\}$. Model (1.2.1) allows to use experimental data on gene regulation (see [76, 66, 53]) to fit parameters $\mathcal{P}$. The main method to study (1.2.1) was numerical simulations, for example, works [76, 66, 53] are devoted to the segmentation in *Drosophila*. The [77] analyses complex patterns occurring under a random choice of the coefficients K_{ij}. Numerical results of [76, 66, 53] have elucidated gap-gene interactions in *Drosophila*. Let us recall that, during Drosophila's embryo development, 6 gap genes (Knips, Hunchback, Kruppel, Tll, Cadal, Giant) and pair-rule genes form a periodic pattern formed by some segments (segmentation process). The main maternal gene, involved in this process, is Bicoid (Bc). An interesting experimental fact about this process is that the segmentation process is remarkably stable with respect to mutations (elimination of some genes), fluctuations of bicoid concentration and variations of embryo size L [51].

To investigate this stability with respect to the morphogene concentrations and mutations, we can consider random θ_l or random thresholds η_i. For example, one can set

$$\eta_i = \bar{\eta}_i + \xi_i(t)$$

where ξ_i are random time functions.

Eqs. (1.2.1) are very complex and thus it could be useful to consider simplified versions of these equations. Let us consider a variant of (1.2.1), where diffusion is removed, namely

$$\frac{\partial u_i}{\partial t} = R_i \sigma(\sum_{j=1}^{m} K_{ij} u_j + \sum_{l=1}^{p} M_{il} \eta_l(x) - \xi_i(t)) - \lambda_i u_i, \tag{1.2.4}$$

where x is involved as a parameter through thresholds $\eta_i(x)$, $\xi_i(t)$ being random noises. A similar model was first introduced for biochemical applications in [38]. The initial data are $u_i(x, 0) \equiv s_i(x)$. Another possible model is a dynamical system with discrete time, for example, defined by the following iterative process

$$u_i^{t+1} = r_i \sigma(\sum_{j=1}^{m} K_{ij} u_j^t + \theta_i(x) - \xi_i^t), \tag{1.2.5}$$

where $t = 0, 1, 2, ..., T$, T is an integer, ξ_i^t are random functions of discrete time t. This model can be considered as a discrete time version of (1.2.4), where diffusion and degradation are removed. On the other hand, this system makes a biological sense. In fact, if θ_i are constants then eqs. (1.2.5) reduce to the classical model of the neural network theory well studied in last decades [14, 31, 37, 27]. For patterning problems, this system describes pattern formation by the so-called sequential induction [3]. The signals that organize spatial patterns of an embryo typically act over short distances. We can consider concentrations $u^1(x)$ as a result of the first patterning round, $u^2(x)$ of the second round etc. In principle, for such a process K_{ij} and σ also can depend on t. "It is chiefly through sequential inductions that the body plan of a developing animal, after being first roughed in miniature, becomes elaborated with finer and finer details as development proceeds" ([3], p.1169).

If we assume that σ is the Heaviside step function, i.e. $\sigma(z) = 0$ for $z < 0$ and $\sigma(z) = 1$ for $z > 0$, and that $r_i = 1$, eqs. (1.2.5) give an example of a boolean circuit. Such circuits are well studied and applied to genetic regulation problems [80, 14, 82]. "Like an input-output logic device, an individual gene is thus turned on and off according to the particular combination of proteins bound to its regulatory regions at each stage of development" ([3], p. 1187).

One can also use general boolean circuits with non-pair interactions, for example

$$u_i^{t+1} = \sigma_i(u_{i_1}^t, u_{i_2}^t, ..., u_{i_l}^t, \theta(x)) \tag{1.2.6}$$

where σ_i are boolean functions of boolean arguments, indices i_k and integer l can depend on i, real arguments $\theta_1, ...$, where given functions θ_l describe effects of maternal gene influences. Boolean models for gene nets were introduced in [57].

1.2.2. Polynomial and other systems

Let us consider general systems with random sources

$$\frac{du}{dt} = F(u, \xi), \tag{1.2.7}$$

where $F = (F_1, ..., F_n)$, $u = (u_1, ..., u_n)$ are unknown time functions, $\xi(t) = (\xi_1, ..., \xi_p)$ are random functions. It is difficult to study such general situation and usually one linearizes (1.2.7) with respect to ξ [49]. The well studied classical models are given by (see [49] for detail):

a) stochastical differential equations with white noises

$$du = f(u)dt + \sum_{l=1}^{p} g_l(u)dW_l, \tag{1.2.8}$$

where W_l are the standard Wiener processes;

b) stochastical differential equations with Ornstein-Ulenbeck noises

$$\frac{du}{dt} = f(u) + \sum_{l=1}^{p} \eta_l g_l(u), \tag{1.2.9}$$

where η_l are defined by

$$d\eta_l = -\gamma_l dt + \sigma_l dW_l$$

or more complicated stochastical equations.

Systems with polynomial g_l, f often occur in biological and chemical applications for example in population dynamics. In this case

$$f_i = P_{i0}(u), \quad (g_l)_i = P_{il}(u), \tag{1.2.10}$$

where P_{il} are polynomials of degree r_{ik}:

$$P_{ik} = \sum_{|\alpha| \le r_{ik}} a_{\alpha,ik} u^\alpha. \tag{1.2.11}$$

Here α denotes the multiindex $\alpha = (\alpha_1, \alpha_2, ..., \alpha_m)$, where $\alpha_i \geq 0$, $|\alpha| = \sum_i \alpha_i$ and $u^\alpha = u_1^{\alpha_1} u_2^{\alpha_2} ... u_m^{\alpha_m}$.

Polynomial and circuit systems are related: under some assumptions a polynomial system can be reduced to a circuit (1.2.4) [68]. Such situation occurs if we are dealing, for example, with a simple chemical reaction involving a substrate S, an enzyme X and a product P. This reaction can be described by three differential equations with nonlinearities quadratic in concentrations X, P, S. Supposing that the concentrations S, P are fast variables and excluding equations for X by a standard procedure, one obtains a differential equation for the concentration S with a nonlinearity having sigmoidal form, where σ is the Michaelis -Menten function [68]: $\sigma(X) = 0$ if $X < 0$, $\sigma(X) = C(K + X)^{-1}$, if $X \geq 0$ and where K, C are positive constants.

Let $\mu(a)$ be a gaussian measure on the space of all polynomials, with a strictly positively defined covariation matrix. This measure induces a measure on the set of all polynomial systems, since each system is defined by polynomials P_{ik}. Below the words "almost all" mean that an assertion holds for all poynomials except for a set S of polynomials of measure 0: $\mu(S) = 0$.

New class of systems. We introduce here a new class of systems with analytical nonlinearities, which includes both many circuits and all polynomial systems. Let us suppose that in (1.2.7) functions $F_i(u, \xi)$ are *pfaffian* [58], a definition of the pfaffian functions can be found in subsection 4.3. Pfaffian functions enjoy remarkable properties [58].

1.3 Assumptions to random processes

1.3.1 Assumptions to fluctuations

Let us formulate some assumptions on the random process $\xi_i(t)$. We suppose that this process is a time homogeneous Markov process with values in $\mathbf{R}^p$. Denote by $P(t, x, B)$ the transition probability from x to the set B within time t, $t \geq 0$. Let us recall [39] that such a process is *stochastically continuous* if for each point $x \in \mathbf{R}^p$ and each neighborhood U_x of this point

$$\lim_{t \to 0} P(t, x, U_x) = 1. \tag{1.3.1}$$

The following assumption is very important in coming sections.

Assumption 1.3.1. *Suppose ξ satisfies the above assumptions, i.e., it is a time homogeneous stochastically continuous Markov process. Moreover, let us assume that for each compact*

$K \subset \mathbf{R}^p$, for any subset $B \subset \mathbf{R}^p$ of non-zero measure and any $t > 0$

$$inf_{x \in K} \ P(t, x, B) > 0. \tag{1.3.2}$$

Notice that the standard Wiener and many diffusion processes satisfy this assumption [39]. Condition (1.3.2) can be interpreted as an extremality of random perturbations, they can take large values. One can say that the system is in an *extremal random environment*.

1.3.2 Stochastic stability under random fluctuations

We say that a system (1.2.4) "survives" (supports "homeostasis") if the concentrations u_i lie inside a compact set $\Pi \subset \mathbf{R}^m$ with non-empty interiour in the u -phase space.

This set Π can be called the *admissible domain*. Let us assume, moreover, that $\Pi \subset \mathbf{R}^m$ is a subset of the nonnegative cone $\mathbf{R}^m_+ = \{u_i \geq 0\}$. If, moreover, Π is a subset of the cone

$$Con_{i_1, i_2, ..., i_s} = \{u : u \in \mathbf{R}^m_+, \ u_{i_l} > 0, \ l = 1, 2, ..., s\}, \tag{1.3.4}$$

for the maximal possible s, then we say that $i_1, i_2, ..., i_s$ are *key indices*.

Remark 1. If i is a key index, then u_i must be positive (i-th node must be active). For biological applications, such a node may correspond to a gene important for organism functionning. Let us consider another, non-biological example: let our circuit simulate a country, nodes are cities, then key indices can correspond important administrative centers.

Remark 2. The set Π can depend on time, then the set of key indices also can depend on time.

As a measure of the stochastic stability of system (1.2.4) with the initial state u^0, we consider the probability

$$P(\mathcal{P}, \Pi, u^0, T_1, T_2) = Prob\{\mathbf{u}(t) \in \Pi \text{ for each } \ t \in [T_1, T_2]\}, \tag{1.3.5}$$

where $\mathbf{u} = (u_1, ..., u_m)$. This probability depends on the circuit parameters $\mathcal{P}$ and the homeostasis domain Π. The quantity $P(\mathcal{P}, \Pi, u^0, T_1, T_2)$ can be named the survival probability on the time interval $[T_1, T_2]$. For brevity, we shall denote it sometimes by $P(T_1, T_2)$ or P_T (if $T_1 = 0$), (omitting the dependence on the parameters $\mathcal{P}$, Π).

1.3.3 Complexity

Let us define a complexity for systems (1.2.4). Each interaction matrix $\mathbf{K}$ generates a directed graph with m nodes and at most $m(m-1)/2$ edges. We suppose that i and j -th nodes are connected by a directed edge if the corresponding entry $K_{ij} \neq 0$. We estimate the stochastic stability via the following parameters: the valency V that will be defined below, $K_* = \max_{ij} |K_{ij}|$, $\bar{\theta} = \max |\sum_{l=1}^{p} M_{il}\theta_l|$, $r = \max_i R_i \lambda_i^{-1}$. It is important to take into account the valency since it is well known that biological circuits are far from being completely connected: for each fixed node i we have a valency $V_i < m$: only V_i among the entries K_{ij} are not equal zero. In applications, typically, $V_i << m$ [2]. We denote $V = \min V_i$, where the minimum is taken over all key indices, N_{key} is the number of such indices. So, the parameter V is defined as the minimal connectivity of the key nodes.

Definition 1.3.3 *Complexity $Comp_N$ of network (1.2.4) is the quadruple* $(V, K_*, \bar{\theta}, r)$.

For polynomial systems (1.2.10) complexity is defined by the degrees r_{il} and the multiples α involved in relations (1.2.10), (1.2.11).

It is difficult to define an analog of such complexity for systems with general analytical or smooth nonlinearities, however, it is possible for pfaffian systems. Complexity of a system (1.2.7) with a pfaffian nonlinearity F is complexity of a pfaffian chain that defines F, see subsection 4.3 . The definition of pfaffian chains can be found in subsection 4.3.

1.4. Evolution models

Let us suppose that the parameters $\mathcal{P}$ of systems (1.2.4) can depend on some variable y, which defines circuit internal parameters P by a map $y \to P(y)$. Let us consider the case, where y takes some discrete values $y_i \in \tilde{Y}$, $\tilde{Y}$ is a finite or countable set. If $\tilde{Y}$ is finite, let us denote by $N(\tilde{Y})$ the number of such states, $i = 1, 2, ..., N$.

Suppose that if a trajectory $u(t)$ of system (1.2.4) leaves Π, our circuit homeostasis instantaneously falls. To take into account this assumption, it is convenient to extend formally the set $\tilde{Y}$ adding to $\tilde{Y}$ a marked state $y_0 = \{\infty\}$ (see [39]). Denote $Y = \tilde{Y} \cup \{\infty\}$. The transition probabilities from this marked state to the other states $y_i, i = 1, ..., N$ equal zero. The phase space of our evolution model is $H = \mathbf{R}^m \times Y$.

Suppose that, for fixed y, the time evolution of u is defined by eqs. (1.2.4). Since ξ_l satisfy condition (1.3.1), a solution of (1.2.4) exists, unique, and (1.2.4) defines a Markov process with continuous trajectories $u(t)$.

Suppose, moreover, that the time evolition of y is defined by a continuous jump like homogeneous Markov process with the state set $\tilde{Y}$ [39].

The conjoint evolution of y, u can be defined as follows. In each state $y_i, i \neq 0$, the u -process is defined by (1.2.4) with the parameters $\mathcal{P}(y_i)$ while $u \in \Pi$ and $y = y_i$. If u leaves Π, the process finishes at the absorbing state $y_0 = \{\infty\}$. If the process makes a jump from y_i to $y_j, j \neq 0$ at $t = t_0$, we assume that u is continuous: $u(t_0-) = u(t_0+)$. At last, if this jump involves new nodes we assume that these new nodes are in zero states at $t = t_0+$.

The process is defined by transition probabilities $P(t, B, i \mid u^0, j)$ (the probabilities that the random trajectory $u(t, u_0, y_j)$ of (1.2.4) with the initial data $u^0, y(0) = y_j$ enters for B at the moment t and $y(t) = y_i$). Such processes are well studied [39, 56].

1.5. Patterning problems

Mathematical approaches to pattern formation problem in the developmental biology has started with the seminal work by A. M. Turing [84] devoted to pattern formation from a spatially uniform state. Turing's model is a system of two reaction-diffusion equations. After [84], similar phenomenological models were studied by numerous works (see [63, 64, 68] for the review). Computer simulations based on this mathematical approach give patterns similar to really observed ones [64]. However, there is no direct evidence of Turing's patterning in any developing organism ([100], p.347). The mathematical models are often selected to be mathematically tractable and they do not take into account actual experimental genetic information.

Moreover, within the framework of the Turing-Meinhardt approach some important theoretical questions are left open. For example, whether there exist "universal" mathematical models and patterning algorithms that allow to obtain any, even very complicated, patterns. In fact, a difficulty in using of simple reaction-diffusion models with polynomial or rational nonlinearities is that we have no patterning algorithms. To obtain a given pattern, first we choose a reasonable model (often using intuitive ideas) and later we adjust coefficients or nonlinear terms by numerical experiments (an excellent example of this approach is given by the book of H. Meinhardt on pigmentation in shells [63, 64]). To overcome this algorithmic difficulty we use genetic circuit models. We show that they are capable to generate any spatio-temporal patterns and that there are different algorithms to resolve patterning problems.

Let us formulate now the patterning problem.

Pattern generation problem for gene circuit model.

Let $T_0 > 0$ and $T_0 < T$. Given functions $z_k(x, t), x \in \Omega, t \in [0, T]$, $k = 1, .., s$ and a positive number ϵ, to find the parameters $\mathbf{P}$ such that the solution of system (1.2.4) with initial *conditions $u_j = 0$ satisfies*

$$\sup_{x,t} |z_k(x,t) - u_k(x,t)| < \epsilon, \quad x \in \Omega, \quad t \in [T_0, T]. \tag{1.5.1}$$

The functions z_k can be called *target pattern*.

Pattern generation problem for time discrete gene circuits

Let $T_0 > 0$ and $T_0 < T$, where T_0, T are integers. Given functions $z_k^t(x) \in [0,1], x \in \Omega, t = 0, 1, ..., T,\ k = 1, ..., s$ and a positive ϵ, to find parameters $\mathcal{P}$ such that the functions generated by relations (1.2.5) satisfy

$$\sup_{x,t} |z_k^t(x) - u_k^t(x)| < \epsilon, \quad x \in \Omega, \quad t = T_0, ..., T,\ k = 1, ..., s. \tag{1.5.2}$$

For logic (boolean) networks $z, u_m \in \{0, 1\}$ and we can set $\epsilon = 0$. Then inequalities (1.5.2) transform into

$$z_k^t(x) = u_k^t(x). \quad t = T_0, ..., T. \tag{1.5.3}$$

Let us give a biological interpretation of these problems. Among the variables u_i, we select special variables, say $u_1, ...u_s$. They can be interpretated as structural genes. The cell states depend on the expression of these genes.

All the rest genes $u_{s+1}, ..., u_m$ are "hidden unites", or regulatory genes. They are involved in a biochemical machinery, but they do not act directly on the cell states. Such an approach is consistent with experimental facts (see [100, 3, 99, 17]). Let us consider, for example, the pigmentation process well studied for *Drosophila melanoguster* [99]. This patterning process is controlled by regulatory genes, which control the expression of other genes, and structural genes, which encode enzymes. These enzymes are involved in biochemical pathways used for pigment synthesis. Different regulatory genes control expression of the structural genes in different body regions. The activity of most enzymes is limited to the cells in which they are expressed [99].

Such problem statement reminds classical approaches of neural network theory [97, 98], where, similarly, one distinguishes "input", "output" neurons and "hidden" ones. This helps to resolve approximation and classification problems by multilayered networks. The output genes can change the cell states and, therefore, they can predetermine an output pattern z.

The hidden genes do not influence directly the cell states, they are involved only in an internal cellular gene regulation.

For stationary patterns z (independent of time) the solution of patterning problem follows from the well known results on multilayered neural networks [10, 45, 48, 22]. If $s = m$, i.e., without hidden genes, the pattern problem transforms into the fitting problem posed in [66]. In this case we have experimental data on all gene concentrations and we try to find circuit parameters, which give the best approximation of the data on some time interval.

The patterning and fitting problems admit many solutions; for gap-gene system in *Drosphila* the fitting problem was studied by numerical simulations [66, 76, 53, 71]. Even in this case, where $m = 6, p = 1$, the problem is difficult: numerical simulations take a large computing time.

1.6. Outline of main results

The main results can be described as follows. Notice that all these results are obtained analytically and do not use computer simulations.

I. Instability

For circuit models we show that they are stochastically unstable. This means that for a circuit of a fixed structure the probability P_T to stay in the admissible domain Π within time T converges to 0 as $T \to \infty$. Estimates of P_T yield that the more is the valency of a key node the stabler is the circuit with respect to perturbations in this node. So, homeostasis generated by a fixed circuit will be broken as time tends to infinity. The similar results also hold for many other systems: for example, a generic differential polynomial system is unstable.

To find a stochastically stable system with fixed parameters is, in general, a computationally complicated task. So, parameter adjusting to construct a stable system is a difficult problem.

These results are stated in Section 2, subsection 2.1 is focused on circuits, subsection 2.2 concerns other systems. Instability entails important corrolaries on evolution.

II Evolution. Inside the considered system classes, a generic system with fixed parameters is stochastically unstable but a system with evolving in time parameters can be stable even for large times as $T \to \infty$. To obtain stability, we need non trivial evolution algorithms: for example, for circuits a random growth, following Erdos-Renýi rule, doest not allow to attain a stable evolution. The problem of finding of stable evolution mechanism is, in general, computationally difficult task.

This fact has an important consequence: one can show, under some assumptions, that the evolution parameters should be discrete (biological interpretation: unstable systems should evolve by a genetic code). Moreover, in extremal conditions evolution, to be successful, should have the following properties: evolution makes copies and the mutation probability is small, the genetic code size must increase during the evolution process and the evolution tree should be large.

Moreover, for the considered systems a *complexity* can be introduced. The circuit complexity must increase in evolution process (on average): a circuit with a priori bounded complexity is not capable to survive in an extremal random environment. The evolution mechanism is non-trivial.

These results are stated in Section 3.

III Patterning and hierarchical modular structure of genetic networks

Furthermore, we show that, roughly speaking, any pattern formation process based on a reaction-diffusion model can be performed as well by a genetic network, with a sufficiently large number of the genes. For each reaction-diffusion model one can find an approximating gene network, with the almost same pattern formation capacity. This result implies that classical Turing-Meinhardt's models can be reformulated as gene circuits.

The second main result on patterning asserts that, under natural conditions on maternal genes densities θ_i, the pattern generation problem always has a solution. Moreover, there is a constructive and numerically effective algorithm that allows us to find a circuit generating a given pattern.

We show that the modular hierarchical organisation and sigmoidal interaction are effective tools to form complex hierarchical patterns. Indeed new, more refined patterns, can be obtained by using of previous old ones.

Given a final pattern $z^T(x)$, one can estimate the minimal number of genes in a network that generates this pattern. We give definitions of "complexity" of the circuits and pattern "complexity". The Khovanski theory [58] gives then that there exists a connection between these complexities: it is impossible to obtain a "complex" pattern by a "simple" circuit.

These results can be found in Section 4 (also see works [89, 91, 92, 93, 94]). However, to explain evolution success, we should seek still additional arguments. We should consider the evolution rate problem.

IV Evolution rate problem.

The methods and models for patterning problems allow to advance as well the evolution rate problem. To make this problem more mathematically tractable, we consider simplified boolean versions of genetic circuits. By these boolean models we can study system response on environment changes. The environment state is defined by a boolean input. This network consists of some types of nodes, "regulator" and "structural". We have N regulator and n structural "genes" and also additional nodes which define environment state. The state of each node is a boolean variable, taking the values "1", or "0" (active, passive or expressed, no expressed respectively). The structural genes respond to an organism features: if such a gene is expressed, the organism has the corresponding feature. If there are n structural genes we can obtain 2^n possible boolean patterns.

In this model, evolution is a growth of N, n plus a formation of connections between the two types of the genes. Each connection also is a boolean variable K_{ij} taking the value 0 or K. To obtain a given pattern for each input, it is a boolean programming problem.

Under some simple restrictions on the circuit structure we show that if the "Freedom Principle" holds, i.e., $f = N/n >> 1$, then even for a random circuit there is a simple heuristic algorithm allowing us to find correct connections K_{ij} in linear time $O(N)$.

2. INSTABILITY IN RANDOM ENVIRONMENT

2.1 Instability in circuits

Consider problem (1.2.4). Suppose that condition (1.3.2) holds, $i_1, i_2, ..., i_s$ are key indices. There holds

Theorem 2.1.1 *Stochastic stability P_T of solutions of problem (1.2.4) can be estimated through $Comp_N$, i.e.,*

$$P_T < g(T, V, K_*, \bar{\theta}, r), \tag{2.1.1}$$

where a function g converges to 0 as $T \to \infty$ for all fixed values $V, K_*, \bar{\theta}, r$, (defined in subsection 1.3.3) and g is monotone increasing in V.

Proof. For time discrete networks (1.2.5) an analogous theorem is obtained in [89, 95]. Let us prove first an auxiliary lemma about the processes satisfying assumption 1.3.1.

Suppose $\eta(t)$ is a smooth function with the values in $\mathbf{R}^p$ defined on $[t_1, t_2], t_2 > t_1$. Let V_{η,δ,t_1,t_2} be a tubular neighborhood of the trajectory η:

$$V_{\eta,\delta,t_1,t_2} = \{v : \ \textit{there is } t \in [t_1, t_2] \ \textit{such that} \ \ |v - \eta(t)| < \delta\}. \tag{2.1.2}$$

Lemma 2.2.1 *Under Assumption 1.3.1 one has*

$$P_{T,\delta,\eta} = Prob\{\xi : \ \xi(t) \in V_{\eta,\delta,0,T} \ \textit{for each } t \in [0, T]\} > 0. \tag{2.1.3}$$

To prove this lemma, let us consider the time points $t_0 = 0 < t_1 < t_2 < ... < t_n < T = t_{n+1}$ such that $\max |t_i - t_{i+1}| < \epsilon$ (the maximum over i). If $\epsilon > 0$ is small enough, then $P_{t_1,\delta/2,\eta} > 0$. In fact, since our process is stochastically continuous, relation (1.3.1) implies that with a non-zero probability the values $\eta(t), t \in (t_1, t_2)$ of the process lie in a small neignborhood of $\eta(0)$. Now relation (2.1.3) can be proved by an induction. Suppose $P_{t_i,\kappa,\eta} > 0$ for all $\kappa > 0$. Let us show that $P_{t_{i+1},\delta,\eta} > 0$. Let us take $\kappa = \delta/2$. Let U_x be $\delta/2$ neighborhood of the point $x = \eta(t_i)$. Condition (1.3.1) implies that $P(t, x, U_x) > 0$ for $0 < t < \epsilon$, if ϵ is small enough. Then, since we are dealing with a Markov process, one concludes that $|\xi(t) - \xi(t_i)| < \delta$ for $t \in [t_i, t_{i+1}]$ with a positive probability. The lemma is proved.

The next step is a priori estimate of solutions of (1.2.4). One obtains

$$|u_i| < M_i = \max\{R_i \lambda_i^{-1}, s_i\}. \tag{2.1.4}$$

Now, using this estimate and properties of σ, one has the following inequality

$$\sum_{j=1}^{m} K_{ij} u_j(t) + \theta_i - \xi_i(t) \leq V_i K_* M_i + \bar{\theta} - \xi_i(t), \tag{2.1.5}$$

where K_*, r are defined in subsection 1.3.3. Let us take a sufficiently large $a = a(V, \bar{\theta}, K_*, r)$ and consider a set Ξ_a of trajectories $\xi(t)$ such that

$$\xi_i(t) > a, \quad t \in [T_1, T_2] \tag{2.1.6}$$

for some i. The probability that a trajectory $\in \Xi_a$ is positive, due to Lemma 2.2.1. Consider system (1.2.4) on the interval $[T_1, T_2]$. Let us fix a key index i. By conditions (1.2.2), one finds a priori estimate

$$0 < u_i(t) < M_i \exp(-\lambda_i(t - T_1)) + \delta(a) R_i \lambda_i^{-1}(1 - \exp(-\lambda_i(t - T_1)), \tag{2.1.7}$$

where $\delta(a) > 0$ is a small number depending on a and converging to zero as $a \to \infty$. For sufficiently large $T_2 - T_1$ and a one has then that $u_i(T_2)$ is small enough. Therefore, the system state $\mathbf{u}$ leaves the domain Π at the moment $t = T_2$. This shows that $P_T < 1$ for sufficiently large T and P_T can be estimated through the network complexity by a function g of the parameters $V, K_*, \bar{\theta}, r$ and time interval length T. Let us prove that $P_T \to$ as $T \to \infty$.

The function g is independent of the initial state u^0 since our estimates are uniform in $u^0 \in \Pi$. Thus, $P_{nT} < g(V, K_*, \bar{\theta}, r)^n$. Letting n go to ∞ one concludes that $P_{nT} \to 0$.

2.2 Instability in other systems

One can show that if Π is a compact set and the number of noises $p > 1$, then a generic (in sense of differential topology, see [46]) system (1.2.8) is instable. An elementary proof [95] based on results of [62], which, in turn, hold on fundamental results of R. Thom [46, 83]. The same result is valid for (1.2.9) and even more general. The main idea is that under (1.3.2) we can remove the bounded term f from eqs. (1.2.7) and system (1.2.7) can be reduced to a symmetric differential polydynamical system defined by vector fields g_l, $l = 1, ..., p$. It is a classical objet of geometric control theory (see [62]). The great transversality theorem of R. Thom allows us then to show that, for $p > 1$, a generic polydynamic system is completely controllable. This property shows, in turn, that always there is a trajectory of this polydynamical system starting with any point in the compact Π and leaving Π.

These results can be obtained for polynomial systems, where a generic polynomial system is defined now in another way, by a measure (see subsection 1.2.2).

We consider system (1.2.9) with the right hand sides (1.2.10), where $p > 1$ (at least two independent noises). We assume that η_l are random processes satisfying assumptions of subsection 1.3.1. Notice that if $p = 1$, i.e., one has a single noise, a generic system can be stable. As an example, one can take $m = 1$, $u_1 = u$, $P_0(u) = -u$, $\Pi = [0, b], b > 0$ P_1 is a polynom in u having a root at a point $c, c > b$.

Theorem 2.2.1 *Suppose p > 1. Then, for almost all systems (1.2.6), the stochastic stability P_T satisfies*

$$P_T < \tilde{g}(T), \tag{2.2.1}$$

where a function $\tilde{g}$ converges to 0 as $T \to \infty$.

For proof see Appendix.

One can prove that this theorem does not hold for general pfaffian systems with nonlinearly involved random ξ.

Example: Let us take the pfaffian system

$$\frac{du}{dt} = u - a - (u-a)^3 + \sum_i^p b_i \sigma(\xi_i),$$

where $\sigma(z) = (1+\exp(z))^{-1}$, $a > 0$, $\Pi \subset (r, \infty)$ with $r < a$ and initial state $u^0 \in \Pi$, and b_i are sufficiently small.

In some cases the stability problem can be investigated in detail. Let us consider the following situation. We consider admissible domains Π, which, in a sense, are narrow. ¿From biological point of view, it can be explained by fragility of biological systems. We suppose that there always exists a direction such that acting in this direction can destroy our system. To formalize this idea, we introduce the following class of domains Π.

Definition. *We say that a set $\Pi \subset \mathbf{R}^n$ is δ-narrow at the point x_0, where $\delta > 0$, if there exists a unit vector e such that the ray $x_1 = x_0 + \tau e$, $\tau > \delta$, lies outside Π.*

The supremum over all the points x_0 of the infimum of δ satisfying this definition can be named the width of the set Π. The width determines the maximal radius of inscribed balls.

The δ-narrow at x_0 set can be large in some directions, but it should be sufficiently narrow at least in one direction defined by the vector e.

If Π is δ-narrow at some x_0 with a δ small enough, then analysis of stochastical stability reduces to some complicated polynomial equations. We are going to use the following known results of geometric control theory.

Lemma 2.2.2 (Kalman criterium of controllability).

Consider the linear system

$$\frac{dx}{dt} = Ax + B\xi(t), \quad x(0) = 0, \tag{2.2.2}$$

where $x \in \mathbf{R}^n$, A is a $n \times n$ matrix, B is a vector $\in \mathbf{R}^p$ and $\xi(t), t \in [0,T]$ is a control. Then system (2.2.2) is controllable, i.e., for each x_1 there exists a $\xi(\cdot)$ such that the corresponding trajectory of (2.2.2) attains x_1 if and only if the following condition holds:

$$dim\ Span\{B, AB, A^2B, ..., A^{n-1}B\} = n \tag{2.2.2}$$

Consider now a system (1.2.9) with a polynomial right hand side, defined by (1.2.10), (1.2.11). To simplify situation we suppose that $p = 1$, i.e., we have only one fluctuating parameter $\xi_1 = \xi$. We investigate a stability at an equalibria of a non-perturbed system, i.e., we suppose that for $\xi = 0$ there exists a point x_0 such that $F(x_0, 0) = 0$. We can suppose, without loss of generality, that $x_0 = 0$. Linearizing eq. (2.1) at 0 we obtain the system (2.2.2) with $A = DF(0,0),\ B = \frac{\partial F}{\partial \xi}(0,0)$.

Proposition 2.2.3 *If Π is δ-narrow at $x_0 = 0$ with a sufficiently small δ then polynomial system (1.2.9) is stochastically stable only if*

$$dimSpan\{B, AB, A^2B, ..., A^nB\} < n, \quad A = DF(0,0),\ B = \frac{\partial F}{\partial \xi}(0,0). \tag{2.2.3}$$

Proof is simple, see [95].

This assertion shows that the analysis of stochastical stability of equilibria reduces to solution of the complicated system of polynomial equations:

$$dim\ Span\{B, AB, A^2B, ..., A^nB\} < n, \quad F(x,0) = 0$$

$$A(x) = DF(x,0),\ B(x) = \frac{\partial F}{\partial \xi}(x,0). \tag{2.2.4}$$

In general, this system is overdetermined and one can expect that generically this system has no solutions and thus equilibrium states of (1.2.9) are stochastically unstable.

3. EVOLUTION OF UNSTABLE SYSTEMS

3.1 Circuit evolution

Let $P(y)$ be a mapping that transforms $y \in \tilde{Y}$ into a value $P(y)$ of the circuit parameters.

Theorem 3.1.1 *Assume that the parameters $r, K_*, \bar{\theta}, V$ of network (1.2.4) are fixed and are independent on y whereas the matrix $\mathbf{K}$ and the number of the nodes m depend on $y(t)$. If the set Y is finite then survival probability $P_T \to 0$ as $T \to \infty$.*

The proof of this theorem immediately follows from the estimates of section 2.1, since they are uniform in $r, K_*, \bar{\theta}, V$.

If the state set $\tilde{Y}$ is countable, then it is possible that the stochastical stability does not vanish for large times: $P_T > P_* > 0$ for all $T > 0$. In this case the parameter V satisfies the following asymptotical relation

$$Prob\{V(y(t)) < A\} \to 0 \tag{3.1.1}$$

as $t \to \infty$ for any A. This means that in this case the circuit complexity is unbounded as $t \to \infty$.

This assertion is trivial, if there are no a priori restrictions to averaged valency of circuit and valency growth rates. To see this fact, let us consider to a Markov process with a countable state set defined by a master (Kolmorogov's) equation [39]. Denote by $p_i(t)$ probabilities to be in the state y_i at the moment t, and p_∞, the probability to be into an absorbing set $\{\infty\}$. The master equation has the form

$$\frac{dp_i}{dt} = \sum_{j \neq i, j \neq \infty} w_{ij} p_j - \sum_{j \neq i} w_{ji} p_i - w_{\infty i} p_i, \quad i = 1, 2, ..., N, \tag{3.1.2}$$

$$\frac{dp_\infty}{dt} = \sum_{j \neq \infty} w_{\infty j} p_j \tag{3.1.3}$$

Here w_{ij} are the transition probabilities to go to the state y_i from the state y_j per unit time.

Due to (3.1.3), the function p_∞ is a time non-decreasing function (Lyapunov function). Moreover, if $w_{\infty,i} > \delta > 0$ (that always holds when $N(Y) < \infty$) one has $p_\infty \to 1$ as $t \to \infty$. These facts express "the second law of termodynamics" for such systems. If $w_{\infty i} = 0$ for all i, then the Lyapunov function is the entropy $H = -\sum p_i \log p_i$.

To show a possibility of the stable evolution, consider the case when the complexity C_i of i-th state defined by $C_i = V(y(i))$ is an increasing function of i. Suppose $w_{i+1,i} > \delta > 0$, the rest entries $w_{ij} = 0$. Then, if C_i increases sufficiently fast with i, one has $p_\infty(t) < 1$ as $t \to \infty$.

The number $N(y)$ can be interpreted as a "genom" size. These results yields that " genome size " must increase in time, otherwise the evolution stops. Moreover, the circuit size m also grows, at least on average, because $V(y)$ cannot stay bounded.

To analyze the evolution process in more detail, let us consider the simplest model, where for each i, j either $K_{ij} = K_*$ or $K_{ij} = 0$. Then time changing of the matrix K can be considered as a time evolution of a directed graph associated with K and vice versa, a growth of a directed graph generates an evolution of a network (1.2.4). The graph evolution can be then considered as an algorithm adding edges and nodes (see Introduction).

Following [90], let us compare two evolution algorithms, the Erdos - Rényi one, [32, 60], and the preferential attachment algorithm [2]. We suppose, in both cases, that averaged (over the whole circuit) valency is bounded: $\bar{V} = m^{-1} \sum_i V_i$.

In Erdos -Rényi's algorithm, at time moments $0, 1, 2, ...$, one adds to a graph a new edge with probability p. This leads to a Gaussian distribution for the degree. In the preferential attachment algorithm [2] the probability that a new edge goes to the i-th node is proportional to the valency (connectivity) of this node.

It can be shown that the Erdos-Rényi growth mechanism is always unstable (under some natural assumptions) [90], whereas Albert-Barabasi preferential attachment evolution can be stable. Indeed, in the Erdos-Rényi case almost all the nodes have valency close to the average one, which is bounded, and the probability that the key nodes have a great valency is small. In the preferential attachment case, the key nodes can have a great increasing in time valency if these nodes have had a great valency at the initial moment. For details, see [90].

Notice that the preferential attachment algorithm can loose stability if the set of the key indices depends on time.

This preferential evolution can be illustrated by a country evolution. Looking on the map of a country we can often see that there exist a few of great cities and many of small cities. In development of many countries we can observe such an effect: there are a large administrative center attracting a great part of resources and population, and many of small cities (to some extent, such an example can be given by Russia, where Moscow attracts $5 - 10$ percents of population, more 60 percents financial resources and almost all power, an opposite example is USA).

A more nontrivial problem is to demonstrate that a stable evolution is possible under some natural restrictions. Let us formulate such restrictions. To simplify the statement, let us consider the case of time discrete networks (1.2.5).

R1 The averaged valency of whole network is a priori bounded for all times:

$$\lim_{t\to\infty} m(t)^{-1} \sum_{i}^{m}(t) V_i(t) < K_c, \tag{3.1.4}$$

where K_c is a positive constant, $m(t)$ is the number of nodes involved in the circuit. This assumption is consistent with experimental data [2, 54, 55]. Let us notice that the averaged valency of the key nodes is not bounded, according to Theorem 3.1.1.

R2 The evolution rate is bounded, i.e., at each evolution time step, we add to the graph associated with the circuit at most one edge and at most one node.

R3 The noises ξ^t are independent random processes with discrete time satisfying

$$0 < P(\xi_i^t > a) = \phi(a) < \exp(-\beta a) \tag{3.1.5}$$

for each $a > 0$ and for each fixed t, where β is positive constant independent of t.

Theorem 3.1.2 *There is a growing circuit (1.2.5) satisfying* **R1**, **R2**, **R3** *such that that its stochastical stability does not vanish for large times: $P_T > P_* > 0$ for all $T > 0$. In this case the valency satisfies the following asymptotical relation (3.1.1) as $t \to \infty$ for any A. This means that the circuit complexity (defined as an averaged connectivity of key nodes) is unbounded as $t \to \infty$.*

Proof. Let us consider the boolean circuit (1.2.5), where σ is the step fucntion. Let us set $N_{key} = 1$, $r_i = 1, \bar{\theta} = 1, K_* > 0$. Let us suppose that at the initial moment we have $N = V_0$ nodes and the matrix $\mathbf{K}$ is defined by $K_{1j} = K_*,\ K_{j1} = K_*,\ K_{jj} = 0$, where $j = 1, 2, ..., V_0$. We set $\theta_i(x) = h > 0$.

At the time moment t, where $t = 1, 2, ...$, we can add a node and one edge connecting this new node with our key node. For the new node, we have weights $K_{1j} = K_*, K_{j1} = K_*$. Let us denote by $V(t)$ the valency of the key node at the moment t. The total number of the nodes at the time moment t also is $N(t) = V(t)$. We shall refer the nodes $2, 3, ...N(t)$ usual nodes.

Let us find first an upper estimate of the probability $Q(t)$ that the circuit will be destroyed at a moment t. Suppose that, at the moment t, exactly k of the N usual nodes have values 0. Then the value u_1^{t+1} can become zero at the time moment $t+1$ as a result of the noise action to the key node. To obtain $u_1^{t+1} = 0$, these noise should satisfy the inequality:

$$\xi_1^t > h + K_*(V(t) - k). \tag{3.1.6}$$

On the other hand, at the tim emoment t the i-th usual node is not active only under the inequality

$$\xi_i^{t-1} > h + K_*. \tag{3.1.7}$$

Therefore, due to our hypothesis **R3** and (3.1.6), (3.1.7), the probability $Q(t)$ admits the estimate

$$Q(t) < \sum_{k=2,...,V(t)} C_n^k \exp(-\beta(h + K_*(V(t) - k))) \exp(-\beta(h + K_*)k).$$

By summarizing over k one obtains

$$Q(t) < \exp(-\beta(K_* V(t) + h))(1 + \exp(-\beta h))^V(t) = \rho^V \exp(-\beta h), \tag{3.1.8}$$

where $\rho = \exp(-\beta K_*)(1 + \exp(-\beta h))$.

Suppose $\rho < 1$. By summarizing (3.1.8) over $t = 1, 2, ..., T$ one finds

$$\log P_T > -\beta h + \sum_{t=1}^{T} \log(1 - \rho^V(t)). \tag{3.1.9}$$

Assume now that the valency of the key node $V(t)$ increases at least linearly: $V(t) > \alpha t + V(0)$ where $\alpha > 0$. Then

$$\sum_{t=1}^{T} \log(1 - \rho^V(t)) > -C - 2\sum_{t=1}^{T} \rho^V(t) > -C_1,$$

where C, C_1 are positive constants. This uniform in T estimate finishes the proof.

Remark 1. If the condition **R3** does not hold, this proof is not correct. (although the theorem, maybe, is correct). Indeed, if the noises are correlated, then the probability of destruction of many nodes may be not small.

Remark 2. If $V(0)$ is large, running of this algorithm is similar to the preferential attachment. The preferential attachment can be considered then as a probabilistic variant of the described algorithm.

Remark 3. This algorithm can be interpreted as a greedy algorithm. Let us consider a node without adjacent edges. The algorithm chooses a new edge, adjacent to this node, to increase maximally the node stability, since the stability grows with valency (connectivity).

Remark 4. There is an interesting analogy between this growth algorithm and the well known Hebb rule of neural network theory. Let us consider a neural network, where synaptic weights take only values K_* or 0. For such and more general networks with discrete synaptic weights a Hebb rule is studied in [16]. This rule can be as a Markov evolution of the entries K_{ij}. If the both nodes (neurons), say, i-th and $j - th$ ones, are active at the moment t (i.e., $u_i^t = u_j^t = 1$), and $K_{ij} = 0$, then with a probability p the weight K_{ij} goes to 1. If one neuron is active and another one is not active, and $K_{ij} = 1$ then, with a probability q, the weight K_{ij} returns to 0.

Let us consider now our network organisation with a key node of a large valency $V >> 1$ encercled by a number of usual nodes of valency 1. This single key node is always active while our circuit survives. Under strong noises the usual nodes will be often disfunctioning. One can expect then that the Hebb evolution of such network will be similar to the preferential attachement evolution.

Remark 5. We do not know whether this algorithm is optimal (gives maximal value P_T for large T) or not. Moreover, other stable algorithms are possible. They drastically depend on properties of $\phi(a)$ and on the parameters h, K_*. To explain this dependence, let us consider the case, where $\phi(a)$ has a threshold behaviour: this function is not small for bounded a, say, on $[0, a_0]$ and it is fast decreasing to zero for $a > a_0$. Let us compare the stability of two structures. The first one is described above, it consists of a single key node connected with V usual nodes. Each usual node is connected with the key node, but there are no connections between the usual nodes.

The second structure is formed by a key node and by clusters. Each usual node is involved in a cluster consisting of n usual nodes. Inside this cluster all the nodes are completely interconnected, one has $N_c = n(n-1)/2$ connections in each cluster. Moreover, each cluster contains a marked central node (this node also is an usual one). Only this central node is connected with the key node. Such structure like the network organization is proposed in [73].

Clearly that the second structure can be essentially stabler than the first one. The stability depends on K_*, h, n and a_0. For large h and small K_* the first structure (no clusters) becomes stabler, the second one is better for small positive or negative h and large K_*.

Remark 6. Theorem 3.1.2 holds in a relatively simple situation when the admissible domain Π is fixed. Actually, biological economical and social systems survive in much more complicated situations when Π depends on time. The key problem is to find stable evolution algorithms in this case. Some ideas can be found in Section 5. One can assume that the Hebb evolution (see Remark 4) may be effective in this complicated case.

It is interesting to interprete the growth algorithm from Theorem 3.1.2 in the framework of our analogy with a strongly centralized country (an Empire) development consisting of a number of regions and a bureacratic center. The evolution sense is to conserve the center. The parameter h can be considered as an internal region resource: for greater h only a great noise ξ_i leads to the region disfunctioning. The parameter K_* determines the connection intensity between the center and the regions. The noises can be considered as instability sources in the regions.

We notice that the Empire should be extending. The described algorithm is as follows: the center obtains resources from the all regions giving, in turn, a minimum of resources for each region. The regions are disconnected. The algorithm works successfully under condition $\rho < 1$. This condition holds if the internal resource parameter h and the connection force K_* are both large enough, if $K_* > 0$ is small and h is very large, or if K_* is very large and $h > 0$ is a small. The second situtation, by our opinion, corresponds to a model of development that reminds Russia (great resources, weak connections). Thus, the algorithm, described in Theorem 3.1.2, can be called " russia way". One can suppose that, in an opposite situation (small resources, strong connections), an algorithm leading to cluster formation should be more effective (such an algorithm can be called "european way").

The russian way does not work if the noises ξ_i are correlated. Appearance of correlated noise can be interpreted, for example, as vanishing of resources in a large region or even in the whole country, or a large rebellion. So, a great Empire can be destroyed only by a correlated action, for instance, by an attack simultenously on many regions or by general economical disfunctioning in a number of regions (USSR fall).

3.2 Evolution as a computational problem. Relation to NP-complete problems

Above we have explained that a stable circuit evolution must use special algorithms. Let us consider now some restrictions to the connection graph K taking into account a real structure of biological molecules.

Above we have supposed that during evolution process any two nodes can be connected. This could give an impression that the network evolution is an easy process. However, this evolution cannot be such a simple process. Indeed, biomolecules consist of numerous polymer groups. During a chemical reaction, they loose (or accumalate) only one such group. This explains, in particular, why enzyme reactions proceed in many steps (see [61]). We conclude, therefore, that it is impossible, in general, to connect two arbitrary nodes. An analogous picture can be observed for other graphs corresponding to processes in the Nature and society [2].

To take into account possible restrictions on the matrix K fixed a priori, we can introduce a large graph (V, E), where V is a set of nodes, E is a set of edges. An entry K_{ij} in (1.2.4) could be non-zero only if it is prescribed by E, i.e., if v_i, v_j a priori can be connected $(v_i, v_j \in E)$.

Now an evolution can be formally described as a time change of subgraphs (V, D^t), $D^t \subset E$, where $t = 0, 1, , 2, ...$ and $D^0 \subset D^1 \subset D^2 ...$ and the time depending matrices $\mathbf{K}^t$ such that

$K_{ij}^t = 0$ if $v_i, v_j \notin D^t$.

Let us consider now some problems connected with such an evolution. Let us fix time t. To obtain a chemical reaction that transforms a substrat $s \in V$ to a product $p \in V$, we have to find a simple path in (V, D^t) leading from s to p.

It is clear as well that the length of this way may be large, but a priori majorated by a number L_{max}. Otherwise, the relaxation processes will be very long and such a system could not survive.

Let us recall our main principle, namely, that the system must be stable in stochastical environment. This implies, in particular, that the system should be stable with respect to mutations or random vanishing of some substrats needed for producing the product p. Mutations can eliminate of some nodes or edges (see above beginning of Section 2). So, evolution should form more than one way from different nutrients to products. The more different ways one has, the stabler is our system. Thus, these ideas lead us to the following problem:

Problem 3.2.1 Given a graph $G = (V, E)$, collection of disjoint node pairs $(s_1, \bar{s}_1), ..., (s_k, \bar{s}_k)$. Does G contain J or more mutually pairwise node-disjoint simple paths connecting s_i and $\bar{s}_i$ for each $i = 1, ..., k$?.

This problem is NP-complete (see [35]). The given nodes s_i could correspond to nutrients (substrates), nodes $\bar{s}_i$ could correspond to products, the paths correspond then to some metabolic paths. Suppose that a system, defined by the graph, survives if the environment contains at least one type of nutrients s_i. Environment fluctuations are eliminations of some nutrients.

There are possible different models of such fluctuations and their action on the system. We shall distinguish two cases: hard environments and soft ones.

Example. Suppose each nutrient s_i can vanish independently with a probability r_i. The system will be destroyed *if all possible nutrients are absent*. Then, if k paths have been found, the probability to survive (per unit of time) becomes $f(k) = 1 - r_1 r_2 ... r_k$.

We say that an environment is hard, if the quantity $f(k)$ (the probability to survive per unit time after finding of the k-th pathway) admits, for large k, the following estimate:

$$f(k) < 1 - \delta k^{-\mu}, \qquad (3.2.1)$$

for some $\mu > 0$, where $\delta > 0$. Otherwise, the environment is soft.

Problem 3.2.1 gives rise to a natural hierarchy of the computational problems (to find a path, to find two paths are found, ... to find k paths).

Circuit stability ideas lead to another natural NP-connected problem that can be formulated as follows:

Problem 3.2.2 Given a graph (V, E), positive integers $K \leq |V|$ and $B \leq |E|$, is there a subset $\tilde{E} \subset E$ with $|E'| \leq B$ such that the graph (V, E') is K -connected, i.e. cannot be disconnected by removing fewer than K nodes?

This problem is simple for a fixed K but it is NP-hard for K varying [35].

The stabilization of differential polynomial system also can entail computationally hard problems. Let us consider the situation studied above with a narrow admissible domain. The main idea is based on the following observation: a "generic" symmetric polydynamical systems defined by g, h are not completely controllable. Indeed, the system

$$g(x, y) = h(x, y) = 0 \tag{3.2.2}$$

can, even in a generic case, have a solution x for some y (however, to find such y is a hard problem, see below).

Similarly, for polynomial dynamical systems with $F = F(x, y, \xi)$ we seek for y such that the system

$$dimSpan\{B, AB, A^2B, ..., A^nB\} < n, \quad F(x, y, 0) = 0,$$
$$A(x) = DF(x, 0), \; B(x) = \frac{\partial F}{\partial \xi}(x, y, \xi)|_{\xi=0} \tag{3.2.3}$$

is resolvable. We suppose that coefficients of polynomials involved in relations (3.2.3) lie in $h\mathbf{Z}$, where h is a rational positive number.

The problems (3.2.3) and (3.2.2) are well known in real algebraic geometry and named "elimination of quantifiers" (see [9]). It is a hard problem. Algorithms for these problems were found by D. Grigoriev et al. [41], another method was proposed by M. F. Roy et al. (see book [9]). The known algorithms take an exponential number N_E of steps

$$N_E = (dn)^{O(n^2)}, \tag{3.2.4}$$

where d is the maximal degree of polynomials A, B, F in x and y. Notice that, in general, the problems of quantifier elimination or even of solvability of polynomial systems are NP-hard

[35]. In the next subsection we shall see that appearance of such problems yields important consequences: evolution should have special properties.

To conclude this subsection, let us notice the following. First, a number of practical problems of bioinformatics are quite complicated, see [72]. The second fact is quite fundamental. Some natural evolution problems are not only complicated. They are not decidable, i.e., there are no algorithms to resolve them.

To show it, one can use results proved in [13]. Let us consider iterations of some map, similar to map (1.2.5) (but without random noises). The problem is as follows: whether these maps attain a some domain in the phase space or not?. It is clear this problem like our problem of existence of stable evolution. This problem is not decidable according to [13]. So, one can expect that it is impossible to decide evolution stops or not. We are not capable to foresee the End of the World.

3.3 Why evolution uses genetic code?

Let us consider other evolution models. For example, one can suppose that the parameter y is continuous. Then an evolution model consists of a random dynamical system, admissible domains Π_t and a second random dynamical system governing y-evolution:

$$\frac{du}{dt} = F(u, y, \xi), \quad u(t) \in \Pi_t \tag{3.3.1}$$

$$\frac{dy}{dt} = Y(u, y, \eta), \tag{3.3.2}$$

where u is the system state, y is an evolving parameter, ξ, η are random noises. This model is well studied (under some restrictions to F, η, ξ, which make this problem mathematically tractable) in viability theory, see [5, 6, 7, 8].

There are some reasons why we have chosen the case, where y is discrete. The first, in principle, continuous evolution can be always approximated by a discrete evolution. The classical example is the organism size L. One can assume that this parameter is continuous, because L is controlled by many genes [74].

The second, it is difficult to describe a graph growth (for example, formation of new connections) by eqs. (3.3.1), (3.3.2) with a continuous parameter y.

At last, the third argument (most important) is as follows. To create a biological system effectively functionning in an extremal environment, it is a complicated problem (see above).

Among all possible parameter values, only a small part gives to biologically reasonable ones. Therefore, if, fortunately, evolution has found a sufficiently stable system then it is necessary to retain the corresponding parameter y. Actually, this argument is known in another form ([74], Ch. 2). For such parameter fixation, a discrete version of eq. (3.3.2) is more effective than a continuous one. To illustrate this intuitive assertion, let us compare two variants of (3.3.2) corresponding to a random parameter search. We also take into account a system extinction connected with leaving of Π_t. The first variant describes a discrete diffusion in the state space:

$$\frac{dp_i}{dt} = d_1(p_{i+1} - 2p_i + p_{i-1}) + \omega Q(y_i)p_i = (\mathcal{L}p)_i, \tag{3.3.3}$$

where p_i is the probability to be in the state $y_i \in Y$, $Q(y) \geq 0$ is the extinction intensity, $Y = \mathbf{Z}$ is a countable set, $i = ... - 2, -1, 0, 1, 2, ...$, d is a diffusion coefficient (diffusion in the state space Y), and ω is a large parameter (large values ω correspond to extremal conditions). Let us suppose that there is a state $y = y_0$ such that $Q(y)$ is minimal at this state and that $\omega Q(y_0) = O(1)$, i.e., the value y_0 gives a stable state. Let us assume that, at $y = 0$, the Taylor approximation holds:

$$Q(y) = Q_{y_0} + c(y - y_0)^2 + ... \tag{3.3.4}$$

The time asymptotics of $p_i(t)$ for large times is defined by the eigenfunction Ψ_0 of the linear operator $\mathcal{L}$ with the minimal eigenvalue λ_{min}. To calculate λ_{min}, we use the standard perturbation theory. For large ω by a time rescaling one can obtain the case, where the diffusion operator is small. Standard calculations show then that

$$\lambda_{min} = \omega Q(y_0) + O(d_1) \quad \omega \to \infty. \tag{3.3.5}$$

This means that the mean survival time τ has the order $O(1)$, i.e., P_T has the asymptotic $\exp(-t/\tau)$ (we suppose that the initail state is y_0).

Let us consider now the second case describing a continuous diffusion in the state space:

$$\frac{\partial p(y,t)}{\partial t} = d_2 p_{yy} + \omega Q(y)p, \tag{3.3.6}$$

where $y \in Y = \mathbf{R}$. An estimate based on (3.3.3) shows that

$$\lambda_{min} = \omega Q(y_0) + O(d_2\omega^{1/2}), \quad \omega \to \infty, \tag{3.3.7}$$

where the term $O(\omega)$ is positive. Comparing (3.3.5) and (3.3.6) one observes that the discrete variant is stabler for large ω, at least if $d_2 > cd_1\omega^{-1/2}$, where c is constant. So, in the continuous case the diffusion coefficient d_2 should be very small. But small values d_2 are not admissible because then there appears a new difficulty: the system is not able to attain at the stable state within a relatively short time. Let us set $y_0 = 0$. Let us denote by $p(y, 0)$ the initial probability density for (3.3.6). Suppose this densityis localized at $y = y_1 > 0$ (for example, the support of $p(x, 0)$ lies in $(y_1/2, 3y_1/2)$). Then $p(0, t) > 1/2$ in the time $\tau_{trans} = O(d_2^{-2})$, and if $d_2 = O(\omega^{-1/2})$, one obtains $\tau_{trans} = O(\omega)$. This last relation implies that the transition time has the same order that the survival times for non-stable states. Roughy speaking, if we are in a stable state, we survive but it is impossible to attain this stable state.

In the discrete case (3.3.3) one concludes that values of d_2 such that the transition time is not too large and the stability is high, should be small.

Using analogous arguments, one seeks for an optimal structure of a Markov operator, which define evolution with maximal chances to survive, independently of initial states. Instead of eq. (3.3.3), let us consider a general kinetic equation (3.1.2) with the extinction probabilities $Q_i = w_{\infty i}$ and with a finite number N of the states.

Proposition 3.3.1 *Suppose that among $Q_i > 0$ there is a Q_{i_0} such that $Q_{i_0} < \delta$, the rest $Q_i > c_1 > 0$. If for all sufficiently small δ the survival probability $P_t(i)$ of the state y_i satisfies*

$$P_t(i) > C\exp(-c_0\delta t), \quad t > 0, c_0 > 0, \tag{3.3.8}$$

where c_0 is a constant independent of δ, then the following estimate holds:

$$\sum_{j=1}^{N} w_{ji} < C\delta, \quad C > 0, \tag{3.3.9}$$

where C is a constant independent of δ.

Moreover, for each i there is such a $j \neq i$ that $w_{ij} > 0$.

The proof uses the standard perturbation theory. Let us decompoze the linear kinetic operator $\mathcal{L}$ into two parts. The first part defines transitions $(\mathcal{W}p)_i = \sum_{j\neq i, j\neq\infty} w_{ij}p_j - \sum_{j\neq i} w_{ji}p_i$, the second one defines extinction $\mathcal{Q}p_i = -Q_ip_i$. The first part can be considered as a perturbation, while the second operator is a diagonal operator with eigenfunctions $\Psi_i^{(l)} = \delta_{li}$ and the corresponding eigenvalues $\lambda_l = -Q_l$. The perturbation $\tilde{\lambda}$ of λ_{i_0} under $\mathcal{W}$ is defined by $\mathcal{W}$ is $\tilde{\lambda} = \sum_{i=1}^{N} w_{ii_0}$. This proves the upper estimate (3.3.9). The second assertion trivially holds because otherwise it is impossible to attain the stable state i_0 from another state i and thus estimate (3.3.8) is not fulfilled.

Biological interpretation: since the stable state i_0 can be aribitrary, depending on the random environment, this elementary proposition yields that evolution almost always makes copies, the mutation probabilities are small, however, they should be positive.

So, one can assume, that many years ago, when there were yet no complicated evolution algorithms, and primitive biosystems lived in extremal conditions, they used a genetic code to survive. Fragile systems without such a code have vanished under fluctuations and chaos. Some fragile systems could survive using first a primitive random search and after more sophisticated evolution algorithms were found. Notice that a concrete mathematical model of genom evolution is proposed in seminal work [30].

To conclude this section, let us notice that from our arguments an interesting consequence follows: organism death, probably, should be genetically programmed (this idea is well known for biologists, see [29]). Replication cannot work without death. In fact, since resources are bounded, old organisms should be destroyed.

3.4 Evolution for countable state sets: branching processes, algorithms, NP-hardness and evolution properties

We suppose here, that at each time moment t the state y may proceed to new states $y'_1, ..., y'_{n(t,y)}$. The number n of new potentially possible states is finite but it may depend on the moment of time and the previous states. One can think about an evolution "tree" growing with time. During tree extending, some states can vanish.

We state the following problem: how to estimate the size of the evolution tree providing a stable evolution, when the survival probability limit P_T stays greater than a positive constant? (i.e., the limit relation $P_T \to 0$ as $T \to \infty$ does not hold). Our goal is to explain increasing of evolution tree and genetic code with time growing. The main idea is to connect this problem with the theory of algorithmic complexity. In fact, we have just seen that to create a system making a stable homeostasis, is a complex problem.

To formalize more the problem, one assumes the following. Let us suppose that each state y is defined by a code C_y. To simplify, we consider the problem with discrete time: $t = 0, \tau, 2\tau, ...$, where τ is a time step. At each instant of time, we transform this code to another code.

We suppose moreover that the survival problem has some "a priori computational complexity" $Comp_{apriori} = Comp_a$. Let us observe that there exists a tradeoff between a memory Mem needed to perform an algorithm and the number of steps N_{time} of this algorithm. For certain compututional problems there was obtained the estimate (first it was obtained in [42], see book [78] for a review):

$$Mem \cdot N_{time} \geq Comp_a. \tag{3.4.1}$$

We can illustrate this fundamental relation by an example. Let us consider the following famous salesman problem, which is NP-complete. Let n cities be located in a country. Distances between cities are given. The problem is to find a tour running all n cities (each city once) and having the minimal total length. Here the algorithm of the exaustive search has an exponential time cost $N_{time} = O(ne)^n$ but it uses the memory $O(n)$. On the other hand, if we use a memory 2^n, we can solve the salesman problem in $O(n)$ steps (see [69]).

Remark: It is important to note that if $P \neq NP$ and, for a NP-complete problem the N_{time} depends polynomially on the input size $|C|$, and $Comp_a$ is not polynomial in $|C|$, then the memory size should be non-polynomial in $|C|$.

Furthermore, we suppose that the evolution solves a chain of computational problems to survive. Namely, we deal with problems $Pr_1, \ldots Pr_k, \ldots$ of increasing a priori complexities $Comp_a(1),\ Comp_a(2), \ldots,\ Comp_a(k), \ldots$.

Let us formulate an important assumption.

Assumption 3.4.1 *At the moment t, all states can be destroyed simultaneously by the random environment within time interval $[t, t+1]$ with the probability $Q(y)$ independent of t (thus we suppose that the random processes are homogeneous in time).*

Example. Let us turn to problem 3.2.1. Recall that given nodes s_i could correspond to nutrients (substrates), given nodes $\bar{s}_i$ could correspond to products, the unknown paths correspond then to metabolic paths. in this case the problem Pr_k is to find k mutually disjoint paths from s_i to $\bar{s}_i$, i.e., from substrats to products. Under assumptions from subection 3.2.1, one has $Q(y) = r_1 r_2 ... r_k$.

We suppose, moreover, that if the corresponding state y is a solution of the problem Pr_k, then the probability $Q(y)$ satisfies

$$Q(y) > 1 - f(k) > 0, \tag{3.4.2}$$

where $f(k) > 0$ is a function of the integer argument k. This means that, at each step, there is an uniform lower bound for the destruction probability (depending on the step number).

By solving a sequence of the computational evolutionary problems the population increases the survival probability. The chances to survive depend on the evolution algorithm speed and on the environment properties.

Let us introduce the quantity $S_{ev}(k)$, which is the number of the states (the nodes of the evolution tree) with *pairwise different* codes obtained to this moment, when k-th computational problem is resolved. Notice that only different codes are essential for evolution, it follows from Assumption 7.1. Moreover, let us observe the inequality $S_{ev}(k) \leq \max |C(y)|$, where the maximum is taken over all states at the k-th step. Therefore, if the tree is large, the code length also is large.

Proposition 3.4.2 *Suppose that the evolution is stable, i.e., $P_T > p_\infty > 0$ as $T \to \infty$. Assume that the evolution solves a sequence of computational problems (as described above) such that their a priori complexities $Comp_a(k)$ increase faster in k than any polynomial $k^{O(1)}$ and that for these problems the estimate (3.4.1) holds. Assume that the population is in an hard environment, i.e.*

$$f(k) < 1 - \delta k^{-\mu}, \quad \delta, \mu > 0. \tag{3.4.3}$$

Then, if $P \neq NP$, the evolution tree size $S_{ev}(k)$ tends to ∞ as $k \to \infty$.

The proof see in [95].

4. CIRCUITS and MORPHOGENESIS

4.1 Generation of complicated patterns

Let us consider pattern generation problem for time discrete gene circuits (1.2.5). To simplify the statement, let us set $s = 1$.

Theorem 4.1.1 *Suppose $T_0 > 2$ and that there exist continuous functions $\phi_l(\theta)$, $l = 1, ..., d$ defined on $\mathbf{R}^m$ such that $x_l = \phi_l(\theta_1(x), ..., \theta_m(x))$ for each $x \in \Omega \subset \mathbf{R}^d$. Then the pattern generation problem has a solution.*

Remark 1. The assumption of the theorem implies that at least d functions θ_i are nontrivial: $\theta_i \neq const$. In the one-dimensional case $d = 1$, to satisfy this assumption, it is sufficient to suppose that at least one function θ_i is strictly monotone. Moreover, under the condition of the theorem, any function $f(x_1, ..., x_d)$ can be presented as a function of $\theta = (\theta_1, ..., \theta_m)$. Indeed, $f(x_1, ..., x_d) = f(\phi_1(\theta), ..., \phi_d(\theta)) = \tilde{f}(\theta)$.

Remark 2. We also observe that the assumption on θ_i is necessary to approximate any sequences $z^t(x)$. In fact, chain (1.2.5) can generate only such sequences z^t, where each $z^t(x)$

depend on x through $\theta(x) = (\theta_1(x), ..., \theta_m(x))$. This means that for each z^t there exists a function $G^t(\theta)$ such that $z^t(x) = G^t(\theta)$. If our assumption does not hold, the trivial target sequence $z^t = x_k$ cannot be approximated by iterations (1.2.5). Consequently, the assumption of the theorem is sufficient and necessary to resolve the pattern generation problem for any outputs z^t.

This theorem can be considered as a generalization of the well known results for multilayered neural networks (multilayered perceptrons) [97, 98, 10, 45, 48, 22] and time recurrent networks [34, 88, 87]. In fact, removing t from (1.2.4) and introducing the input $v = u^t$ and the output $w = u^{t+1}$, we transform (1.2.4) to a multilayered network. The multilayered perceptrons are capable to generate complicated patterns and resolve classification problems [97, 98, 10, 45, 48, 22]. If we remove x from (1.2.4) we obtain a time recurrent network. It is well known that these networks can generate all possible time trajectories [34]. Also they are capable to generate all possible structurally stable attractors (up to a topological orbital equivalency) [88, 87]. Roughly speaking, they can induce all time depending patterns. Theorem 4.1.1 is a development of these previous results.

This theorem can be obtained by the following lemma.

Superposition Lemma 4.2.1 *Consider a family consisting of p circuits (1.2.5) generating functions $u^t_{i,s}$, where $t = 0, ..., T_1$, $s = 1, ..., p$ and $i = 1, 2, ..., m_s$ (here the index s marks the functions generated by s-th circuit, m_s is the number of the genes involved in s-th circuit). Denote by $\mathbf{u}^t$ the vector with the components $u^t_{1,1}, u^t_{2,1}, ..., u^t_{m_1,1}, u^t_{1,2}, ..., u^t_{m_2,2}, ..., u^t_{1,p}, ...u^t_{m_p,p}$.*

Suppose that $z^t(x) = F(\mathbf{u}^t(x))$, where F is a continuous function of N variables defined on N -dimensional cube $Q_N = [0, 1]^N$ and $N = \sum_{s=1}^p m_s$ is the complete number of functions involved in the circuits. (This means that the target pattern can be expressed through the patterns generated by our family). Then for any $\epsilon > 0$, there exists a circuit (1.2.5) satisfying (1.5.2) with $T_0 = 2$ and $T = T_1 + 2$.

The main idea of the proof of this lemma is based on the biological fact: the gene networks have modular hierarchical structure and are organized in blocks [44, 73]. As a mathematical basis, we use as well the following known approximation result: for each $\kappa > 0$ there exist a number m and coefficients A_{kjs}, b_k, η_k such that

$$|\sigma^{-1}(F(\mathbf{u})) - \sum_{k=1}^{m} b_k \sigma(\sum_{s=1}^{p}\sum_{j=1}^{m_s} A_{kjs} u_{j,s} - \eta_k)| < \kappa, \quad \mathbf{u} \in Q^N, \tag{4.1.1}$$

for detail see [94]. The theorem easily follows from Lemma 4.2, see [94].

This proof gives, moreover, an algorithm to resolve the pattern generation problem. Namely, the key step of the proof (approximation (4.1.1)) can be realized by a constructive procedure (see [87]). An explicit estimate of the gene number m can be obtained under some supplementary assumptions on F from this lemma and on z^t from Theorem 4.1.1. Namely, we suppose that the functions $F(\mathbf{u})$ and $z^t(x)$ are Lipshitzian, with the Lipshitz constants $Lip(F)$ and $Lip(z^t)$.

Analogues of these results can be obtained for time continuous patterning problem (1.5.1), see [92, 93]. Lemma 4.2.1 confirms the famous law "moprhogenesis repeats evolution" [74, 3, 100]: new, more refined patterns, can be always obtained by old patterns by a sequential induction.

Notice that the obtained algorithm, based on a superposition, is not unique. There are many variants to resolve the patterning problem. Even for the gap-gene fitness problem, when we have experimental data on all the gene concentrations, numerical results show that there are a number of solutions. Finding of these solutions needs a global search and takes a formidable processing time. The first works have used simulated annealing [66, 76], afterwards asymptotic approaches were developped in order to diminish processing time [71]. Moreover, to handle experimental data correctly, we need a priori hypothesis on the net structure. It is well known that such inverse problems for gene nets are very difficult, although different approaches are proposed [57, 101, 71, 70]. Furtermore, we should take into account a fundamental robustness of circuit with respect to variations in maternal gene concentrations, mutations embryo sizes. The output pattern should be proportional to the embryo length [51, 50]. There are possible different reaction - diffusion models to explain this stability observed in experiments [50], however we think that they are still far from biological reality. The problem is far from to be resolved: it is a topic for coming investigations.

4.2 Approximation of reaction-diffusion systems by gene networks

We consider, for simplicity, the case of two component reaction-diffusion systems

$$\frac{\partial u}{\partial t} = d_1 \Delta u + f(u, v), \quad \frac{\partial v}{\partial t} = d_2 \Delta v + g(u, v). \tag{4.2.1}$$

The phenomenological approach based on eqs. (4.2.1) gives excellent results for some pattern formation problems, see [64, 63, 68].

In these equations, u and v are unknown functions of the space variables $x = (x_1, x_2, x_3)$

defined in a bounded domain Ω. System (4.2.1) must be complemented by standard initial and boundary conditions. The general multi-component case can be studied in a similar way. Assume solutions of (4.2.1) remain globally bounded, i.e., for some positive constants C_i we have the estimate

$$|u(x,t)| < C_1, \quad |v(x,t)| < C_2, \tag{4.2.2}$$

for all $t > 0$, if it holds for $t = 0$. Let us define the domain D_{C_1,C_2} as follows:

$$D_{C_1,C_2} = \{(u,v) : \quad 0 \le u < C_1,\ 0 \le v < C_2\}. \tag{4.2.3}$$

We suppose that initial condition belongs to D_{C_1,C_2} for each x.

One can show that, for a given reaction-diffusion system we can always find an "ϵ- equivalent" circuit (1.2.1). Namely, for this equivalent circuit there exists a smooth map $b(y)$: $(y_1, y_2, \ldots, y_m) \to (u, v)$ transforming the gene concentrations to the reagent concentrations and such that time evolution of u, v is defined by a new reaction -diffusion system with nonlinearities $\Phi_1(u,v), \Phi_2(u,v)$, ϵ- close to nonlinearities $f(u,v), g(u,v)$. Therefore, one can say that reaction -diffusion systems can be realized as gene circuits.

To construct this circuit, we use the same modular approach (subsection 4.1), and algebraic tools from [87]. Let us consider a system (1.2.1) having a special modular structure. Namely, we assume that there exist two kinds of the genes. We denote these groups of the genes by y and z, where vector $y(x,t)$ contains m_1 components and $z(x,t)$ contains m_2 components. Naturally, $m = m_1 + m_2$. We consider a system (1.2.1) of the special form

$$\frac{\partial y_i}{\partial t} = \sigma(\mathbf{K}_i^{yy} y + \mathbf{K}_i^{yz} z - \theta_i) + d_1 \Delta y_i, \tag{4.2.4a}$$

$$\frac{\partial z_i}{\partial t} = \sigma(\mathbf{K}_i^{zy} y + \mathbf{K}_i^{zz} z - \bar{\theta}_i) + d_2 \Delta z_i. \qquad 4.2.4b$$

Here we use notation $\mathbf{K}_i y = \sum_{j=1}^{m} K_{ij} y_j$ and matrices $\mathbf{K}^{yy}$, $\mathbf{K}^{zz}$, $\mathbf{K}^{zy}$ and $\mathbf{K}^{yz}$ describe interactions between different groups of the genes.

In general, these interactions are not symmetric, i.e., $\mathbf{K}^{yz}$ is not equal to the transpose of $\mathbf{K}^{zy}$. The coefficients d_1 and d_2 coincide with the diffusion coefficients in equations (4.2.1). We choose the entries of the matrices $\mathbf{K}^{yy}$, $\mathbf{K}^{zz}$, $\mathbf{K}^{zy}$ and $\mathbf{K}^{yz}$ as follows:

$$K_{ij}^{yy} = a_i b_j, \quad K_{ij}^{yz} = \gamma_i \bar{b}_j, \quad K_{ij}^{zy} = \bar{\gamma}_i b_j, \quad K_{ij}^{zz} = \bar{a}_i \bar{b}_j$$

where $a_i, \bar{a}_i, \gamma_i, \bar{\gamma}_i, b_i, \bar{b}_i$ are unknown coefficients.

Let us define "collective variables"

$$u = \sum_{i=1}^{m_1} b_i y_i, \quad v = \sum_{i=1}^{m_2} \bar{b}_i z_i.$$

After some calculations (see [93]) we obtain

$$\frac{\partial u}{\partial t} = d_1 \Delta u + \Phi_1(u, v), \quad \frac{\partial v}{\partial t} = d_2 \Delta v + \Phi_2(u, v),$$

where

$$\Phi_1(u, v) = \sum_{i=1}^{m_1} b_i \sigma(a_i u + \gamma_i v - \theta_i), \quad \Phi_2(u, v) = \sum_{i=1}^{m_2} \bar{b}_i \sigma(\bar{a}_i v + \bar{\gamma}_i u - \bar{\theta}_i).$$

The well known approximation results of the neural network theory [10, 48, 45, 34] yield that for any $\epsilon > 0$ there exist numbers m_1, m_2, vectors $a, b, \bar{a}, \bar{b}, \gamma, \bar{\gamma}$ and $\theta, \bar{\theta}$ such that

$$|\Phi_1(u, v) - f(u, v)| < \epsilon, \quad |\Phi_2(u, v) - g(u, v)| < \epsilon$$

for all u, v from some bounded domain. This proves the following result [92, 93]:

Proposition 4.2.1

Consider equations (4.2.1) whose solutions remain in a domain D_{C_1,C_2}.

Then, if functions f, g are continuous, for any $\epsilon > 0$, there exist such a system (4.2.1) with a sufficiently large number m and coefficients $r = (r_1, r_2, \ldots, r_m)$ and $s = (s_1, s_2, \ldots, s_m)$ such that the functions

$$u = ry = \sum_{i=1}^{m} r_i y_i, \quad v = sy = \sum_{i=1}^{m} s_i y_i$$

satisfy the system

$$u_t = d_1 \Delta u + \tilde{f}(u, v), \quad v_t = d_2 \Delta v + \tilde{g}(u, v),$$

where

$$|f(u, v) - \tilde{f}(u, v)| < \epsilon, \quad |g(u, v) - \tilde{g}(u, v)| < \epsilon$$

for $(u, v) \in D_{C_1,C_2}$.

Therefore, any reaction-diffusion patterning processes on a bounded time interval $[0, T]$ can be performed as well by genetic networks. In other words, the pattern capacity of the gene circuits on bounded time intervals are not less than the pattern capacity of reaction-diffusion systems.

4.3 Complexity of pattern and complexity of network

In this subsection we consider the following problem. Suppose we observe some sequence of patterns $z^t(x), \ x \in \Omega, \ t \in [0, T]$. We would like to estimate the number of the genes required to create this sequence.

To resolve this problem we can use different characteristics of pattern complexity. In this paper we employ the following three quantities: $C_1(z^t(\cdot), c)$, $C_2(z^t(\cdot), c_1, c_2)$, $E(z^t(\cdot))$. They are functions of the discrete time t.

The quantity C_1 is the number of the connected components of the set

$$D_{c,t} = \{x : \ z^t(x) = c\}. \tag{4.3.1}$$

To define C_2, let us consider a set $D_{c_1,c_2,t}$ depending on two parameters c_1, c_2 and t. Namely, let us define

$$D_{c_1,c_2,t} = \{x : \ c_1 \leq z^t(x) \leq c_2\}. \tag{4.3.2}$$

Then C_2 is the number of the connected components of this set.

Both complexity measures are discrete, whereas E is a continuous quantity defined by

$$E(t) = \int_\Omega |\nabla z^t|^2 dx. \tag{4.3.3}$$

Let us discuss now the biological sense of C_1, C_2 and E and relations between them.

Organisms consist of cells and these cells can be in different states. Following classical ideas [67, 3] we assume that different cell states appear as a result of expression of different genes. We consider here the case of one gene. Let u_m be such a gene.

Then we can study structures consisting of two kinds of cells: modified and the usual ones. If u_m is expressed at x then we have here a modified cell at x, otherwise the cell remains in a usual state.

Following the usual threshold approach we suppose that the gene u_m is expressed if $u_m > c$ and it is not expressed in the opposite case ($u_m \leq c$). In this case we obtain, as a natural measure of complexity, the quantity C_1.

The measure C_2 admits a similar interpretation. Here we assume that u_m is expressed if $u_m > c_2$ and it is not expressed if $u_m < c_1$. In the case $c_1 < u_m < c_2$ we deal with an intermediate (transient) state.

Thus both measures C_1 and C_2 relate to the number of transitions between the cells of different types.

Notice that using Sard' theorem [46] we can choose c, c_1, c_2 in definitions (4.3.1) and (4.3.2) such that at least locally the boundaries of the connected components will be smooth submanifolds of Ω of the codimension 1. In particular, if Ω is an interval, these components will be isolated points.

Example. For a periodical in x function $z^t(x)$ ("layered structure") C_1=C_2= number of layers (for appropriate c, c_1, c_2).

The third measure, the quantity E, can be interpreted as a mean value of "oscillations" of z.

The results for C_1 and C_2 are quite different. To estimate m through C_1 we use so-called Pfaffian chains [58], under some additional assumptions on σ. It allows us to obtain rough estimates of C_1 by Khovanski's results. Estimates of C_2 and E can be derived in a simpler way and appear to be essentially better.

Up to now, nobody knows whether the Khovanskii bounds can be improved. The key difference between estimates of C_1 on the one hand, and C_2, E on the other is that the estimates of C_2 and E depend, in particular, on the diameter $diam(\Omega)$ of domain Ω whereas the ones of C_1 are independent of this diameter.

An estimate of m via C_1

Let us introduce the key notion of a Pfaffian chain [58], [40].

Definition. A Pfaffian chain of the length r and degree $d \geq 1$ is a sequence of real analytic functions $f_1(x), f_2(x), ...f_T(x)$ in $\mathbf{R}^n$ with the following property: every f_j, $1 \leq j \leq T$ satisfies a Pfaffian equation

$$\frac{\partial f_j}{\partial x_k} = g_{kj}(x, f_1(x), ..., f_j(x)), \tag{4.3.4}$$

where g_{kj} are polynomials of degrees $\leq d$. Then T is called the length and d the degree of the Pfaffian chain.

Pfaffian functions are well studied. They enjoy the following properties: the sum and the product of two Pfaffian functions f_1 and f_2 of lengths r_i and degrees d_i are again Pffafian functions of length $r_1 + r_2$ and degree $d_1 + d_2$ for both the sum and the product. Superpositions of Pfaffian functions also are Pfaffian (see [40] for details).

Consider some elementary examples. The exponent $\exp(ax)$, $x \in \mathbf{R}$ is a Pfaffian function of length 1 and degree 2. More generally, any real analytic function $f(z),\ z \in \mathbf{R}$ satisfying an equation

$$\frac{df}{dz} = P(z, f) \tag{4.3.5}$$

is a Pfaffian of degree $degP$. We observe thus that many classical sigmoidal functions are Pfaffian. For example, $f = (1 + \exp(z))^{-1}$ satisfies (4.3.5) with $P = f^2 - f$. Superposition $\sigma(\exp(ax))$ also is a Pfaffian, etc.

Let us show first that chain (1.2.5) is a Pfaffian chain if $\sigma(z) = 1 + \exp(-az)^{-1}$, where $a > 0$. Let us introduce complexity of chain (1.2.5) as the tuple of integers

$$Comp = \{\ m,\ T,\ r_\theta,\ d_\theta,\ degP\}, \tag{4.3.6}$$

where r_θ is the sum of the lengths of Pfaffian chains for θ_i, d_θ is the maximum of the degrees of Pfaffian chains determining θ_i, $degP$ is the degree of the polynomial from (4.3.5) that defines σ.

Using induction, let us consider now the functions u_i^1. By differentiating, one has

$$\frac{\partial u_i^1}{\partial x_l} = \sigma'(\mu_i \theta_i - \eta_i)\mu_i \frac{\partial \theta_i(x)}{\partial x_l}.$$

Consequently by (4.3.5) one obtains

$$\frac{\partial u_i^1}{\partial x_l} = P(\mu_i \theta_i - \eta_i)\mu_i P_{i,l}(x, v_1^i, v_2^i, ..., \theta_i), \tag{4.3.6}$$

where $P_{j,l}$ are appropriate polynomials, v_k^j are functions of chains determining θ_j. Thus, u_i^1 and θ_j form a chain of the degree $d_\theta + degP$ and the length $r_\theta + m$. Repeating these calculations, we conclude that $u_i^t, u_i^{t-1}, ...\theta_i$ form a chain of the degree $d_\theta + tdegP$ and the length $r_t = r_\theta + tm$.

Now the complexity of the pattern $u_m^T(x)$ can be estimated applying known results ([58], see also [40], Proposition A4).

Theorem 4.3.1. *The number C_1 of the connected components of the pattern $u_m^T(x)$ generated by (1.2.5) can be bounded from above by*

$$C_1 < 2^{(r_\theta + Tm)^2}(d_\theta + TdegP)^{O(r_\theta + Tm + n)}. \tag{4.3.7}$$

Thus given C_1 we can bound from below $R = r_\theta + Tm$ roughly as $(\log_2 C_1)^{1/2}$, provided that $\log(degP), \log(d_\theta), n^{1/2}$ are less than $r_\theta + Tm$. The quantity R can be interpreted as a "complexity" of gene circuit (1.2.5).

This estimate does not look optimal but in general case up to now there exist no methods that could improve it.

5 CAN EVOLUTION CONSTRUCT A "BIOLOGICAL COMPUTER" ?

Following classical concepts [3, 67] we consider here organisms as "biological computers", which should give a correct response to different environment challenges. How an evolution based on local search and mutations can construct such a computer? The key question: Is there a stable evolution algorithm generating step by step such a computer? This problem is much more complicated than patterning problem studied in the previous section although there exists a connection. We are not capable yet to resolve this problem, but we state here some ideas that allow us to expect to a positive answer.

To simplify the problem, in this section we consider boolean circuits. We suppose that they involve structural genes $z_1, z_2, ...z_n$ and regulatory genes $u_1, u_2, ...u_N$. In real systems there are many different regulator elements (enchancers, transcription factors etc). Recall the fundamental experimental fact: organism complexity is generated mainly by regulatory elements. In fact, only three genoms are decoded: the genom of the worm *C. elegans*, of the fly *Drosophila melanoguster* and *Homo sapiens*. The first contains 19000 genes, the second contains 14000 ones and the third has 30000 genes. Main genes making body design in *C. elegans* and *Drosophila* are similar. The fly is more complex with respect to the worm because the fly genom contains more regulatory elements. The simplest well studied patterning processes are pigmentation and segmentation in *Drosophila*. The segmentation is controlled by both maternal (zygotical) genes (Bicoid, Nanos) and regulatory genes (gap-genes, pair-rule genes, segment-polarity genes). Details of their interaction are unknown, but it it well known that this interaction is redundant and some genes repress a part of other ones, and activate another part. The same redundancy can be observed in pigmentation. "For instance, the *yellow* locus contains five independent *cis*-regulatory elements (enchancers) that control its expression in the developing body, wings, bristles, laval mouthparts and denticle belts. There is also evidence for a simialr modular organization of *cis*-regulatory elements of other enzyme genes" [99]. Because each enchancer element is independent, random modification of an enchancer do not affect the regulatory functions. It helps gene expression to evolve independently in each body part [99]. *Drosophila melanoguster* development involves, besides gap, pair-rule and segment -polarity genes, homeotic selector genes and many others. These gens are organized in modules (see above and [3, 44].

We consider two simplest models of such redundant regulations. In the both cases one can aplly some contemporary ideas of the algorithm theory.

The first model is a particular case of (1.2.5)

$$z_i = \sigma(\tilde{r}_{1i} + r_{1i}u_{i_1} + \tilde{r}_{2i} + r_{2i}u_{i_2} + ... + \tilde{r}_{ki} + r_{ki}u_{i_k} - h_i) \tag{5.1.1}$$

where σ is the step function, $h_i \in (0, k)$, the indices $i_1, i_2, ...i_k$ are chosen randomly for each i. We suppose that parameters r_{li} are random variables, taking values 1 or -1 and $\tilde{r}_{il}$ to be defined by the condition: if $r_{li} = 1$ then $\tilde{r}_{il} = 0$ and if $r_{li} = -1$ then $\tilde{r}_{il} = 1$. We consider the time following stationary assignment problem: are there values $u_1, ..., u_N$ such that

$$z_l(u) = 1.$$

It is easy to see that this problem reduces to the random k-SAT problem: to satisfy n randomly constructed disjunctions

$$z_l = v_{i_1(l)} \; Or \; v_{i_2(l)} \; Or ... \; Or \; v_{i_k(l)}, \quad i = 1, 2, ..., n \tag{5.1.2}$$

where each v_{i_r} is either u_{i_r}, or the negation $\bar{u}_{i_r}$ of u_{i_r}. The k-SAT problem has been a central for theoretical computer science since S. Cook established that it is NP-complete in 1971 [20].

Of course, as a model for gene network, this circuit is too simplified, however, all our arguments are valid for general networks (1.2.6) if they possess a redundancy parameter. We say that a family of boolean functions $\sigma_i(u^t_{i_1}, u_{i_2}, ..., u_{i_s})$ has the redundancy k if for each i boolean function σ_i contains k-disjunctions $v_{j_1} \; Or \; v_{j_2} \; or... \; Or \; v_{j_k}$, where each v_{i_r} is either u_{i_r}, or the negation $\bar{u}_{i_r}$ of u_{i_r}.

Notice that problem (5.1.2) is complicated, namely, it is NP-hard [35, 69], there are 2^N possible inputs. This problem admits such biological intepretation: we suppose that our "organism" survives only if all n structural genes are correctly assigned. Structural genes predetermine "organism" features. If $z_l = 1$ then one can say that an "organism" possesses l -th feature. To form a complex organ, the organism should form simultaneuosly many features, and, moreover, this formation process should be proceeded step by step (mathematically, this means that (5.1.2) should be resolved by a local search algorithm, for example, a greedy algorithm).

In general case problem (5.1.2) cannot be resolved in a polynomial in n time: the best known algorithms use running time $2^{cn}, 0 < c < 1$ [23]. This means that this problem is absolutely hopeless for a darwinian biological evolution based on gradient descent methods (greedy ones). Such algorithms use some function of input u trying to increase this function at each step. In biology, this function is the so-called fitness introduced by R. Fisher [33, 81]. The hardness of (5.1.2) means, that for large n, say $n > 40$, in general there are no fitness that could resolve (5.1.2) in a "biologically reasonble" time.

Nonetheless, one can show that problem is not so hopeless. Suppose that $N < (2^{-k}/k)n$. Then there is an algorithm, which resolves (5.1.2) with a probability, close to 1 as $n \to \infty$ (for rigorious mathematical results [1, 18], for a review see [19]. First such algorithms were invented by sophisticated methods of theoretical physics, see works [4, 25, 26], where these methods have applied for different NP-hard problems). These mathematical ideas yield such a

Biological corollary. *There is a relation between the number of regulatory genes and of structural genes. The more redundancy, the smaller may be the proportion α of regulatory genes with respect to structural ones. If α is too small, evolution stops.*

Notice that redundancy plays an important role in evolution [96].

This assertion can be called **Freedom Principle**: *the number of regulatory elements should be sufficiently large with respect to the number of structural elements*. The relation α between these numbers depends on the regulation mechanism, mathematically, on the boolean function. In more complicated situations, the critical level α may be higher.

To illustrate it, let us consider a model describing a boolean gene circuit which responds to external medium changes. This model is similar to systems considered in [14, 27] and it has the following form

$$y_k = \sigma(\sum_{j=1}^{n} w_{kj} s_j^{in}), \quad k = 1, ..., N \tag{5.1.3}$$

$$s^{out} = \sigma(\sum_{k=1}^{N} W_k y_k). \tag{5.1.4}$$

We assume that σ is the step function and $s^{out}, s_j^{in}, y_k \in \{0, 1\}$. This circuit contains three kinds of nodes. The nodes s_j^{in} contain information about external environment, y_k are regulatory ones, s^{out} is an output state. Let us suppose that w_{kj} and W_k take values $0, 1, -1$. We assume that there exists a graph V, E describing connections between nodes. Each input node s_i^{in} is connected with a random set B_i of y -nodes consisting of m nodes, $|B_i| = m$.

The problem is to find a correct correspondence between inputs and the output. Suppose that l-th type of environment (where $l = 1, 2, ..., L$) activate s_i^{in} in such a way: $s_i^{in} = \delta_{il}$. We set $w_{kj} = 1$. For each l the system should give a correct output, 1 or -1. This leads to the following problem of boolean programming: to find W_k satisfying the inequalities:

$$\sum_{k=1}^{N} W_k y_{kl} > (<) 0, \quad l = 1, 2, ..., L \tag{5.1.5}$$

where the sign $>$ or $<$ depends on l and any sequence of the signs is possible, y_{kl} are boolean coefficients describing a reaction to l-th input defined by (5.1.3). In general, such a problem is NP-complete [35], however, in our case the freedom principle works successfully. If $\alpha = N/m$ is large enough, a simple heuristic algorithm resolves the problem (with a probability close to 1). To see it, let us notice that problem (5.1.5) can be rewritten (under above assumptions) by

$$\sum_{k \in B_l} W_k > (<) 0, \quad l = 1, 2, ...L. \tag{5.1.6}$$

This chain of the inequalities can be satisfied step by step in such a way. If for $l = 1$ one has $>$, let us set $W_k = 1$ for all $k \in B_1$ otherwise one takes $W_k = -1$ for $k \in B_1$.

At $l + 1$ step, we set $W_k = 1$ for $>$ in (5.1.6) and $W_k = -1$ for $<$ for all $k \in B_{l+1}$ such that W_k is not yet defined at the previous steps. If the value W_k is defined earlier, we do not change this value.

Correctness of the algorithm follows from the next elementary Lemma.

Lemma 5.1.1. *Let us consider the set $I_N = \{1, 2, ..., N\}$. Let $B_{m'}$ be a subset of I_N consisting of $1 \leq m' < N$ elements. Let us consider the probability $P(m, m', k)$ that a random subset R_m consisting of $1 \leq m < N$ elements has an intersection with $B_{m'}$ having at least k elements. Then this probability satisfies the estimate*

$$P(m, m', k) < (const \frac{mm'}{kN})^k. \tag{5.1.7}$$

By applying this estimate inductively one obtains that the algorithm works with probability P^*

$$P^* > (1 - (\frac{cm}{N})^{m/2})(1 - (\frac{2cm}{N})^{m/2})...(1 - (\frac{Lcm}{N})^{m/2}).$$

Since $\log(1 - x) > x$, one finds that

$$\log P^* > L(\frac{cmL}{N})^{m/2}.$$

So, for large N and m, L such that the parameter mL/N is small enough this algorithm works with a probability close to 1.

Notice that this model is similar to multilayered perceptron models [97, 98], it is a completely discrete version of multilayered perceptron, where synaptic weights (connection forces) take discrete values. Such neural models are well studied (see [16]). In the time reccurent case, they can generate a complicated dynamics (an analytical proof see [88]) and, therefore, they can simulate any Turing machine [13, 59].

The proposed algorithm realizes a training process for such multilayered circuit. This algorithm is essentially simpler than the usual backtraining procedure [97, 98]. Moreover, this algorithm is fast and proceeds step by step, where the need step number can be estimated by linearly $O(mL)$. The last fact is very important: this means that, using Freedom principle, evolution can create complicated organs step by step (see the citation from Charles Darwin in Introduction). Taking into account that evolution uses a great tree (see subsection 3.4), the process can be accelarated.

However, it is necessary to note that such local search algorithms can be succesful in real biological situation only when there is a fitness having a biological meaning and growing with each step (or, at least on average, during some steps). This hypothesis is not mathematical assertion, it is a biological one that can be verified only by experiments. The assertion is not evident, because the first untuitive impression that, in many cases, only a completely constructed organ (or subsystem) can help to survive. As an example one can consider the famous Krebs cycle that plays the key role in organism energetics [3]. This cycle consists of many chemical reactions. It is unclear why such a cycle can be obtained by the discussed scheme of step by step evolution: a reaction chain, that still non compeletely constructed, does not function. A natural hypothesis is that, possibly, such noncomplete chains could function in transient extincted organisms. Some arguments show that the hypothesis of such kind could be correct [74], however, this problem is far from to be resolved.

CONCLUSION

In this paper we have developped a mathematical approach to evolution of complex systems. This approach seems opposite to R. Thom's structural stability ideas [83, 36]. Indeed, the fundamental concept of structural stability meets serious difficulties (see, for example, a discussion in [79]). A general dynamic system is structurally stable under very restrictive conditions [28, 79, 75, 52]. Paradoxally, by R. Thom's results one can show that many systems are unstable under strong fluctuations of their parameters (multiplicative noises) [62, 95].

Notice that the instability ideas were known in biology under the name "Red Queen hypothesis" and was proposed by Van Valen [85] in 1973. They are confirmed by experimental data on species survival. On average, the species survival times are bounded by some million years. The Red Queen hypothesis asserts that species extinction probabilities are positive.

Following [85, 43] we develop here the idea that biological systems are fundamentally unstable under fluctuations, however, evolution can stabilize them.

The large classes of stochastically unstable systems have been found, in particular, circuits are unstable. Independently of an initial state, the state of an unstable system leaves an "admissible domain", where this system is viable.

We have introduced here a definition of stable evolution algorithms. Stable evolution algorithms give, for a set of unstable evolving circuits, non-zero chances to stay stable even under strong multiplicative noises (earlier or later, noises destroy each individual circuit, but a chain of the circuits can live eternally).

Mathematical tools based on network theory, graph theory, control theory, probabilistic approaches and theoretical computer science allow us to understand, at least for classical circuit models, main properties of stable algorithms and even to describe some stable algorithms. There is a connection between stable algorithms and such classical models of graph and network evolution as the preferential attachment, the Hebb rule. In some situations, the stable evolution leads to free-scale structures, in other cases to cluster formation.

It seems that the developped approach is in a good accordance with fundamental experimental data. Namely, it explains, at least to some extent, 1) existence of genetic code; 2) a growth of this code (maybe, non-monotone); 3) existence of a great evolution tree; 4) a bounded average time of species life; 5) death is a genetically programmed phenomen.

In the framework of this approach the most interesting and complicated problem is as follows. Since seminal work [67], biologists consider the cell as a complicated mechanism performing a sophisticated feedback. This feedback based on gene switches helps to survive in the response of environment challenges [3]. Are there stable evolution algorithms allowing to construct "biological computer" , consisting of unstable elements but working in stable manner, within an evolutionary reasonable time period?

To advance this problem we show first that gene networks, organized in modular hierarchical structures, can construct any complicated spatio-temporal patterns; one can estimate theoretically the gene number need for this process and connect a circuit complexity with a pattern complexity. In the last section we have considered a toy model of "biological computer".

The obtained results allows us to expect that an evolution based on random mutations and selection actually can construct effective gene switches within a bounded time period, nonetheless, the question on existence of a stable evolution leading to such gene circuits is open.

To conclude, let us notice a remarkable and fundamental fact. Although each fragile organism should die, and, probably, this process is genetically programmed [29], but the problem on the evolution end is, probably, not decidable. This means that it is impossible, due to fundamental reasons, to foresee the End of the World.

Acknowledgements

The authors thank Ovidiu Radulescu, John Reinitz, Vitaly Gurskii, Stephane Genieys, Thierry Dumont, Nikolai Petrov for fruitful discussions and a help. The first author is grateful to Vitaly Volpert for his hospitality, discussions and a help in University Claude Bernard Lyon 1 (1998- 2003).

6. Appendix

The proof of Theorem 2.2 is based on the following lemma.

Lemma 6.2 *Suppose* Π *is a compact set. Let us consider the system of polynomial equations*

$$P_{ik}(u) = 0. \tag{6.1.1}$$

Assume that the number of equations $N(P)$ in (6.1.1) is more than the number of variables m. Then the probability that this system has a solution $u_* \in \Pi$ equals 0.

The formal proof of this obvious assertion is as follows. Let us introduced an auxiliary function of the variables u and the coefficients a_α

$$\phi_\epsilon(a, u) = \exp(-\epsilon^{-2} S(u)) \tag{6.1.2}$$

depending on a parameter ϵ, where

$$S(u) = \sum_{ik} P_{ik}(u)^2.$$

Let us consider now the integral

$$I_\epsilon = \int_\Pi \int \phi_\epsilon(u,a)d\mu(a)d^m u. \tag{6.1.3}$$

Since μ is an exponentially descreasing function as $|u| \to \infty$, we can change the order of integration in (6.1.3).

Suppose there is a solution u_* of system (6.1.2). Then

$$\phi_\epsilon > \delta > 0 \tag{6.1.4}$$

in a ball of radius ϵ^{1-r} centered at u_*, where $r \in (0,1]$ and δ is uniform in ϵ as $\epsilon \to 0$. Thus, by integrating first over u and then over a, one obtains that in this case

$$I_\epsilon > \delta\epsilon^{(1-r)m} \int d\mu(a) > c\epsilon^m. \tag{6.1.5}$$

Let us find an upper estimate of this integral. Integrating first over all coefficients $a_{000...0,i}$ at the term , one sees that

$$\int \phi_\epsilon d\mu(a) < C\epsilon^{N(P)}.$$

Therefore, due to compactness of Π, one has

$$I_\epsilon < C\epsilon^N. \tag{6.1.6}$$

For $N > m$ estimates (6.1.6) and (6.1.5) lead to a contradiction as $\epsilon \to 0$. The lemma is proved.

Notice that in our case $N(P) = pm$. Thus since $p > 1$ we can use the lemma. Therefore, for each initial state u^0 there are indices i, k_0 such that $P_{ik_0}(u) \neq 0$. Let us choose functions $\eta_k(t)$ defined on $[0, T/a]$, where a is a positive parameter, such that $\eta_{k_0} = a$, $\eta_k \equiv 0$ for $k \neq k_0$. Let us define a small $\delta(a)$ tubular neighborhood $\mathcal{W}$ of $\eta(t)$. One can show now that, if a is large enough and $\delta(a)$ is sufficiently small, and if $\xi(t) \in \mathcal{W}$, the solution $u(t)$ leave the compact Π (see [95] for more detail). Now the proof can be finished By Lemma 2.2.1.

References

[1] Achlioptas D. and Moore C., (2002) The asymptotic order of the random k-SAT threshold. IEE Comput.Soc. Press. Los Alamitos, CA.

[2] Albert R. and Barabási A. L., (2002) Statistical mechanics of complex networks, Rev. Modern Physics, **74**, 47-97

[3] Alberts B., Bray D., Lewis J., Raff M., Roberts K., Walter P., Molecular Biology of the Cell, 4nd ed. (Garland Publishing, Inc., New York, 2002.

[4] Altarelli F., Monasson R., Zamponi F. (2007) Can rare SAT formulae be easily recognized? On the efficiency of message-passing algorithms for K-SAT at large clause-to-variable ratios. . J. Phys. A: Math. Theor. **40**, 867-886

[5] Aubin J.-P., Bayen A., Bonneuil N. and Saint-Pierre P. (2005) Viability, Control and Games: Regulation of complex evolutionary systems under uncertainty and viability constraints, Springer-Verlag.

[6] Aubin J.-P. (2000) Mutational and morphological analysis: tools for shape regulation and morphogenesis, Birkhauser.

[7] Aubin J.-P. (1997) Dynamic economic theory: a viability approach, Springer-Verlag.

[8] Aubin J.-P. (1996) Neural networks and qualitative physics: a viability approach, Cambridge University Press.

[9] S. Basu, R. Pollack, M. F. Roy, Algorithms in real algebraic geometry, Berlin-Heidelberg, New-York, Springer, 2003.

[10] Barron A., (1993) Universal Approximation Bounds for superpositions of a sigmoidal functions, IEEE Trans. on Inf. Theory, **39**, 930-945.

[11] Begon M., Harper J.L., Townsend C. R., Ecology, Vol.2, Blackwall Scientific Punlications, Oxford London Edinburgh, 1986.

[12] Bernard C., An Introduction to the Study of Experimental Medicine, Macmillan and Co., Ltd., 1927.

[13] Blondel V., Bournez O., Koiran P. and Tsitsiklis J. (2001) The stability of saturated linear dynamical systems is undecidable. Journal of Computer and System Sciences, **62**, 442-462

[14] Bornholdt S., (2001) Modeling Genetic Networks and their Evolution: A Complex Dynamical Systems Perspective, Biol. Chem. **382**, 1289-1299.

[15] Brocker T., Lander K., Diffential Germs and Catastrophes, Cambridge Uiniversity Press, 1975.

[16] Brunel N., Carusi F. and Fusi S.(1998), Slow stochastic Hebbian learning of classes in recurrent neural networks, Network: Computation in Neural Systems, **9**, 123-152

[17] Carroll S. B. et al., From DNA to Diversity: Moleculer Genetics and the Evolution of Animal Design, Blackwell Science, 2001.

[18] Chao M-T. and Franco J. J., (1990) Probabilistic analysis of a generalisation of the unit-clause literal selection heuristics for k-satisfiability problem. Inform. Sci., **51 (3)**, 289-314

[19] Cocco S., Monasson R., Montanari A., Semerjan S. Approximate analysis of search algorithms with "physical methods", (2004) Chapter for "Phase transitions and Algorithmic complexity" edited by G. Istrate, C. Moore, A. Percus

[20] Cook S., (1971) The complexity of theorem proving prcoedures. In Proceeding of the Third Annual ACM Symposium on Theory of Computing, 151-155

[21] Cormen T. H., Leiserson C. E., Rivest R. L., Introduction to Algorithms, MIT Press, New York etc. 1990.

[22] Cybenko G., (1989) Approximation by superpositions of a sigmoidal function, Mathematics of Control, Signals and System, **2**, 303-314.

[23] Dantsin E., Hirsch E. A., Ivanov S., Vsemirnov M., (2003) Algorithms for SAT and Upper Bounds on Their Complexity, Journal of Mathe. Sciences **118 (2)** 4948 -4962

[24] Darwin C. R., On the Origin of Species, John Murray. London, 1859.

[25] Deroulers C., Monasson R., (2006) Criticality and Universality in the Unit-Propagation Search Rule. Eur. Phys. Journal, **B49**, 339

[26] Deroulers C., Monasson R. (2004) Critical behaviour of combinatorial search algorithms, and the unitary-propagation universality class. Europhys. Lett. **68**, 153

[27] Derrida B., Pomeau Y., (1986) Random networks of automata: a simple annealed approximation, Europhys. Lett. **1**, 45-49

[28] *Dynamical Systems with Hyperbolic Behaviour*, D. V. Anosov (ed). (Dynamical Systems 9). Encyclopedia of Mathematical Sciences, Vol. 66. Translated from Russian. Springer V., Berlin, Heidelberg, New-York, 1995.

V

[29] Dilman V. M. (1994) Aging, Rate of Aging and Cancer. A Search for Preventive Treatment Annals of the New York Academy of Sciences **719 (1)**, 454-455. doi:10.1111/j.1749-6632.1994.tb56849.x

[30] Eigen M. and Schuster P., The Hypercycle – A Principle of Natural Self-Organization, Springer, 1979.

[31] Edwards R., Siegelmann T.H., Aziza K. and Glass L., (2001) Symbolic dynamics and computation in model gene networks, Chaos, **11** , 160-169.

[32] Erdos P., A. Rényi A., (1960) On the evolution of random graphs, Publ. Math. Inst. Hungarian Academy of Sciences, **5**, 17-61

[33] Fisher, R. A., The genetical theory of natural selection. Oxford, Clarendon Press, 1930.

[34] Funahashi K. and Nakamura Y., (1993) Approximation of Dynamical Systems by Continuous Time Recurrent Neural Networks, Neural Networks, **6**, 801-806.

[35] Garey M. H., Jonnson D. S., Computers and Intractability, A Guide to the Theory of NP-completeness, W. H. Freeman and Company. New York, 1979.

[36] Gilmore R., Catastrophe Theory for scientists and engineers, John Wiley and Sons, New -York etc. 1981.

[37] Glass L. and Kauffman S., (1973) The logical analysis of continuous, nonlinear biochemical control networks, J. Theor. Biology, **34**, 103-129.

[38] Glass L., (1975) Combinatorial and topological methods in nonlinear chemical kinetics. J. Chem. Phys. **63 (4)** 1325 -1335

[39] Gichman I. I. and Scorochod A. B., Introduction to Theory of Random Processes (in Russian, Nauka, Moscow, 1977).

[40] Grigoriev D. and Vorobjov N., (1996) Complexity Lower Bounds for Computation Trees with Elementary Transcendental Functions Gates, Theoret.Comput.Sci., **157**, 185-214.

[41] Grigoriev D., (1988) Complexity of deciding Tarski algebra, J. Sym. Computations, **5**, 65-108.

[42] Grigoriev D., (1980) Application of separability and independence notions for proving lower bounds of circuit complexity, J.Soviet Math., **14**, 1450-1456

[43] Gromov M. and Carbone A.(2001), Mathematical slices of molecular biology, Preprint IHES/M/01/03.

[44] Hartwell L. H., Hopfield J. J., Leibler S. and Murray A. W., (1999) From molecular to modular cell biology, Nature(London), **402**, C47-C52.

[45] Hecht-Nielsen R., (1987) Kolmogorov's mapping neural network existence theorem, IEEE International Conference on Neural Networks, San Diego: SOS Printing, 11-14.

[46] Hirsch M. M., Differential topology, Springer-Verlag, New-York, Heidelberg, Berlin. 1976.

[47] Hopfield J. J., (1982) Neural networks and physical systems with emergent collective computational abilities, Proc. of Natl. Acad. USA, **79** 2554-2558.

[48] Hornik K., Stinchcombe M. and White H., (1989) Multilayer feedforward networks are universal approximators, Neural Networks, **2**, 359-366.

[49] Horsthemke W., Lefever R, Noise-induced Transitions, Springer-Verlag, Berlin etc. 1984.

[50] Howard M., ten Wolde P. R. (2005) Finding the center reliably: robust patterns of developmental gene expression, arHiv: q-bio.TO/0511011 v1

[51] B. Hougmandzadech B., Wieschaus E., and Leibler S. (2002) Nature **415**, 798- 802.

[52] Ilyashenko Yu, Weigu Li, Nonlocal bifurcations, AMS, 1999.

[53] Jaeger J, Blagov M, Kosman D, Kozlov K.N, Manu, Myasnikova E, Surkova S, Vanario-Alonso CE, Samsonova M, Sharp D.H and Reinitz J, (2004) Dynamical analysis of regulatory interactions in the gap gene system of Drosophila melanogaster. Genetics **167**, 1721-1737.

[54] Jeong H., B. Tombor B., Albert R., Otvai Z. N., and Barabási A.L., (2000) The large-scale organisation of metabolic networks, Nature (London), **407**, 651-654.

[55] Jeong H., Mason S. P., Barabási A.L. and Otvai Z. N., (2000) Lethality and centrality in protein networks, Nature (London), **411**, 41-42.

[56] Karlin S., A first course in stochastic processes, Academic Press, New York, London. 1968.

[57] Kauffman S. A., (1969) Metabolic stability and epigenesis in randomly constructed genetic nets. J. Theor. Biology, 22, pp. 437-467

[58] Khovanskii A., Fewnomials, Translations of Mathem. Monographs, Amer. Math. Soc., **88**, 1991.

[59] Koiran P. and Moore C. Closed-form analytic maps in one and two dimensions can simulate Turing machines. (1999) DIMACS Technical Report 96-25. Theoretical Computer Science, **210(1)** 217-223

[60] Kolchin V. F., Random graphs (in Russian). FizMatLit. Moscow. 2004.

[61] Lehninger A. L., Nelson D. L. and Cox M. M., Principles of Biochemistry, 2nd. ed. Worth, New York, 1993).

[62] Lobry C., Une proprieté generique des couples de champs de vecteurs, (1972) Chechoslovak Mathematical Journal, **22 (97)** 230-237.

[63] Meinhardt H., Models of Biological Pattern Formation, Academic Press, 1982.

[64] Meinhardt H., The Algorithmic Beauty of Sea Shells, 2nd Enlarged Edition, Springer, 1998.

[65] Mendoza L. and Alvarez-Buylla E. R., (1998) Dynamics of Genetic Regulatory Networks for *Arabodopsis thaliana* Flower Morphogenesis, J. Theor. Biol., **193**, 307-319.

[66] Mjolness E., Sharp D. H. and Reinitz J., (1991) A connectionist Model of Development, J. Theor. Biol. **152**, 429-453.

[67] Monod J., Jacob F., (1961) General conclusions: teleonomic mechanisms in cellular metabolism, growth and differentiation, Cold Spring Harb. Symp. quant. Biol. 26, 389-401

[68] Murray J. D., Mathematical Biology, Springer, New York, Berlin, Heidelberg, 1993.

[69] Papadimitriou C. H., Steglitz K., Combinatorial optimization, Algorithms and Complexity, Prentice Hall, Inc. Englewood Cliffs. New Jersey, 1982

[70] Perkins T. J., Hallette M., Glass L., (2006) Dynamical properties of model gene networks and implications for the inverse problem, Biosystems **84** 115 - 123

[71] Perkins T. J., Jaeger J., Reinitz J., Glass L., (2006) Reverse Engineering the Gap Gene Network of Drosophila Melanoguster, PloS Computational Biology, **2**, 0417 -428.

[72] Proceeding of WCB05 Workshop on Constraint Based Methods for Bioinformatics, Eds. R. Backofen and A. Dovier, 2005. Spain.

[73] Ravasz E., Somera A. L., Mongru D. A., Oltvai Z. N., Barabasi A-L., (2002) Hierarchical organisation of Modularity in Metabolic Networks, Science, **297** 1551- 1555

[74] Ridley M., Evolution, 2nd ed. Blokwell Scientific Publications Ltd, Oxford, 1996

[75] Ruelle D., Elements of differentiable dynamics and bifurcation theory, Acad. Press, Boston, 1989.

[76] Reinitz J. and Sharp D. H., (1995) Mechanism of formation of eve stripes Mechanisms of Development, **49**, 133-158.

[77] Salazar-Ciudad I., Garcia-Fernadez J. and Solé R.V., (2000), Gene Networks Capable of Pattern Formation: ¿From Induction to Reaction-Diffusion, J. Theor Biology, **205** 587-603.

[78] Savage J. E., Models of Computations. Exploring the Power of Computing, Addison-Wesley, 1997.

[79] Smale S., Mathematics of Time, Springer, New - York, 1980.

[80] Smolen P., Baxter D., (2000) Byrne J. H., Mathematical modelling of gene networks, Review in Neuron, **25**, 247-292.

[81] Sviregev, Ju. M. and Pasekov V. P.: Foundations of theoretical genetics. Moskow, Nauka 1982

[82] Thieffry D. and Thomas R., Dynamical behaviour of biological regulatory networks, II.Immunity control in bacteriophage lambda, Bull. Math. Biology, **57**, 277-295

[83] Thom R., Stabilité structurelle et morphogénèse, New York, Benjamin, 1972.

[84] Turing A. M., (1952) The chemical basis of morphogenesis, Phil. Trans. Roy. Sco. B, **237** 37-72

[85] Van Valen L., (1973) A new evolutionary law, Evolutionary Theory, **1**, 1-30

[86] Ventsel A. D. and Freidlin M. I., Random Perturbations of Dynamic Systems, Springer, New York, 1984.

[87] Vakulenko S. A., (2000) Dissipative systems generating any structurally stable chaos, Advances in Diff. Equations, **5**, 1139-1178

[88] Vakoulenko, S. Complexité dynamique de reseaux de Hopfield, (2002) C. R. Acad. Sci. Paris Sér. I Math., **t.335**.

[89] Vakulenko S. and Grigoriev D., Complexity of patterns generated by genetic circuits and Pfaffian functions, Preprint IHES, 2003.

[90] Vakulenko S., Grigoriev D., (2005) Stable growth of complex systems, Proceeding of Fifth Workshop on Simulation, 705- 709.

[91] Vakulenko S., Grigoriev D., (2003) Complexity of gene circuits, Pfaffian functions and the morphogenesis problem, C. R. Acad. Sci, Ser I. **337** 721-724

[92] Vakulenko S. and Genieys S., (2003) Pattern programming by genetic networks, Patterns and Waves, Collection of papers, eds. A. Abramian, S. Vakulenko and V. Volpert, S. Petersburg, 346-366.

[93] Vakulenko S. and Genieys S., (2005) Patterning by genetic networks, Mathematical Methods in Applied Sciences, **29**, 173 -190.

[94] Vakulenko S., Grigoriev D., (2006) Algorithms and complexity in biological pattern formation problems, Annales of Pure and Applied Logic **141**, 421 -428

[95] Vakulenko S., Grigoriev D., (2005) Evolution in random environment and structural stability, Zapiski seminarov POMI RAN, **325**, 28 -60

[96] Wagner A., (1999) Redundant gene functions and natural selection, J. Evol. Biol. **12**, 1 -16

[97] Nguyen D., Widrow B. The truck backer-upper: An example of self-learning in neural networks. In W.T. Miller, R. Sutton, and P. Werbos, Eds. Neural Networks for Robotics and Control. Cambridge, MA: M.I.T. Press, 1990.

[98] Widrow B., Lehr M.A. (1990) 30 years of adaptive neural networks: perceptron, madaline, and backpropagation, Proceedings of the IEEE, **78**, No. 9, 1415-1442.

[99] Witcopp P. J., Carroll S. B. and Kopp A., (2003) Evolution in black and white: genetic control of pigment patterns in *Drosophila*, TRENDS in Genetics, **19**, No9., 495-504

[100] Wolpert L., R. Beddington R., Jessell T., Lawrence P, Meyerowitz E. and Smith J., Principles of Development Oxford University Press, 2002.

[101] Yeing M. K. S., Tegner J., Collins J. J., (2002) Reverse engineering gene networks using singular vale decomposition and robust regression, Proc. Natl Acad. USA, **99** 6163- 6168