The Dynamic World View in Action

4.1 Causality, Teleology and About the Scope and Limits of the Dynamical Paradigm

The dynamical system is a mathematical concept motivated first by Newtonian mechanics. The state of the system is generally denoted by a point in an appropriately defined geometrical space. A dynamical system operates in time. Typically, we take the time set $T$ to be the real line $\mathbb{R}$ (a continuous-time system) or the set of integers $\mathbb{Z}$ (a discrete-time system). We then formalize an autonomous system as a ordered pair $(Q, g)$, where $Q$ is the state space, and $g : T \times Q \to Q$ is a function that assigns to each initial state $x_0 \in Q$ the state $x = g(t, x_0)$, in which the system will be after a time interval $t$ if it started in state $x_0$. A fundamental property of $g$, then, is the validity of the identity

$$g(t + s, x_0) \equiv g(s, g(t, x_0))$$

for all states $x$, and times $t, s$. Loosely speaking $g$ is a fixed rule which governs the motion of the system.

The behavior of a dynamic system may also depend on the time course of the input applied. Systems that take into account the effect of input belong to the family of non-autonomous systems.

Dynamical systems theory became a predominant paradigm, and materialize the philosophical concept of causality by mathematical tools. This concept tells that causes imply effects, consequently the present state determines the future. If the fixed rule acting on the actual state is deterministic, there is only “one” future. In case of probabilistic rules, different futures could be predicted with certain probabilities.
For continuous time continuous state deterministic system the equation is given by
\[ \dot{x}(t) = f(x(t), k) + I(t); \quad x(0) = x_0. \] (4.2)
It describes the temporal behavior of the system variable \( x \) driven by a “forcing function” \( f \), influenced by the parameter \( k \), starting from the initial condition \( x_0 \), taking into account that the system is subject of the time dependent external perturbation (\( I(t) \)) arriving from the environment. The evolution of the system depends on the additive effect of the external perturbation and the internal development.

Robert Rosen, a controversial hero of theoretical biology and of complex systems research suggested [439] that at least three of the four Aristotelian notions of causality can be associated to mathematical objects in equation (4.2):

| Aristotelean categories | mathematical objects   |
|-------------------------|------------------------|
| material cause          | initial values         |
| formal cause            | parameters             |
| efficient cause         | driving force          |
| final cause             | —                      |

“The causality of natural processes may be interpreted as implying that the conditions in a body at time \( t \) are determined by the past history of the body, and that no aspect of its future behavior need to be known in order to determine all of them” – wrote Truesdell and Noll in 1965 in their rigorous encyclopedia on the nonlinear mechanics of continua.

4.1.1 Causal Versus Teleological Description

While differential equations (and dynamical systems) describe motions locally, for certain systems there is an equivalent, global description by using integral principles for the whole motion. Specifically, a quantity is defined, which should have an extreme (i.e., maximum or minimum) value for the whole motion from the beginning to the end. The concept of optimal processes reached physics at least four hundred years ago.

Pierre de Fermat (1601–1665), a French lawyer and mathematician (probably more famous for the latter profession) posed and solved a well-defined physical (specifically geometrical optical) problem by mathematical methods.
How does a light ray move from an initial point to a given endpoint through an optical medium, where the optical density may vary from point to point? There is a set of possible solutions, and Fermat’s “principle of least time” gives an answer. While there are infinite number of possible trajectories which connect the two fixed points, there is a selection criterion to choose one of them. The answer suggests that the time of the traveling (and nothing else, e.g. not the geometrical distance between the two points) should be minimal.

The “principle of least action” (most likely formulated by Pierre-Louis Moreau de Maupertuis (1698–1759)) is the mechanical analogue of Fermat’s principle. There is a mechanical quantity called “action” which should be minimum for the path taken by a particle among all possible trajectories which are compatible with the conservation of the energy. Maupertuis principle is also an optimality principle. About a hundred years later Hamilton’s principle, the most effective optimality principle in physics, has been formulated. The principle of William Rowan Hamilton (1805–1865) states that for mechanical systems there is a quantity, called the Lagrange function, and the integral of this function will be stationary for the actual mechanical process. This function is defined as the difference between the kinetic and potential energy, $L := E_{\text{kin}} - E_{\text{pot}}$. According to the Hamilton’s principle:

$$\delta \int_{t_0}^{t_1} L \, dt = 0. \quad (4.3)$$

The quantity $\int_{t_0}^{t_1} L \, dt$ is the action integral. For the classical mechanical systems the local and global (“teleological”) descriptions are equivalent. The mathematics behind the extremal principles of mechanics (and a few other branches of physics) is based on the mathematical disciplines called variational calculus. The Lagrange function satisfies the equation

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} - \frac{\partial L}{\partial x_i} = 0. \quad (4.4)$$

For the harmonic oscillator $E_{\text{kin}} := 1/2m\dot{x}^2$ and $E_{\text{pot}} := 1/2kx^2$, and the action integral is $J = 1/2 \int_0^T (m\dot{x}^2 - kx^2) \, dt$, and $\delta J(x) = 0$ is fulfilled when $m\ddot{x} + kx = 0$. This is exactly the Newton equation for the harmonic oscillator.

The equation for the pendulum could be also rewritten in the form of Lagrange equation.

The fact that many (but far from all) dynamical equations of natural process based on mechanisms localized to a single time point are mathematically equivalent to integral description referring to a time interval implies that for
this (restricted, but important) class the “causal” description is equivalent with the “teleological” one. The variational principle is not only an equivalent formulation of the local one in the Newtonian mechanics, but has been used in relativity theory and quantum physics, as well.

4.1.2 Causality, Networks, Emergent Novelty

According to the mechanistic worldview, Science, Technology and Metaphysics seemed to be unified by the Newtonian principles. The motion of mechanical machines as well as celestial bodies were thought to be determined by the (same) Laws of Nature. The clockwork world view of Kepler, Galileo and Newton, characterized by causality, determinism, continuity and reversibility, promised to reduce all kinds of dynamic phenomena to mechanical motions. At the end of the 18th century chemistry and medicine started to challenge the view that material nature is nothing but inert mass and motion. As we know, the invention of the steam engine contributed very much to the disorganization of the cyclic reversible mechanistic world concept and to the birth of the theory of irreversibility.

One of the supporting pillar of the Newtonian world view – paradigm, if you like – is the strict principle of causality, which can be stated as follows: “Every event is caused by some other event” ([83], pp. XXII). The concept of the linear causality, which separates ‘cause’ and ‘effect’ by a simple temporal sequence, has been found to be appropriate for describing simple systems, and could not be considered as a universal concept. Circular causality relies on the suggestion that in a feedback loop there is no meaning in separating cause and effect, since they are mixed together. “Circular causality is not just a subcategory of causality, but a concept that supersedes the traditional notions of cause and effect. Hence, these traditional notions no longer apply” ([455], p. 129). The term “network causality” means the ability to take into account interactions among structural loops. The models of linear causal systems are appropriate subsets of single level dynamic systems.

We assume here that the models of these generalized causal phenomena can be given within the framework of dynamic system theory. In a really excellent, thought provoking book [263] George Kampis identified material causation as a creative agent which causes self-modifications of systems. Dynamical system theory can try to tackle the emergent complex structures and creative processes.
“Material causation is just a word. It is clear . . . that what it means is that we do not know anything about the causes that determine things by an invisible and unapproachable ‘actor’. The actions of this actor bring forth something new. It has to be taken very seriously that molecules, thoughts, artefacts, and other qualities never existed in the Universe before they were first produced by a material causation. Therefore, irreducible material causation is creation per se: free construction of new existence, with new properties, i.e., that in no conceivable form pre-exist either physically or logically. Before they are already there, absolutely no hint can be gained about their possibility, about their properties, about how they come into being, what they will look like and what will happen to them next . . .”

Kampis, G: Self-modifying systems in biology and cognitive science. pp. 257–258.

We will not go further into the metaphysical issues of causality and dynamics. Our real concern in this chapter is that by using the concepts of dynamical systems spatiotemporal patterns observed in natural and socioeconomical systems are supposed to be subject of causal explanations. A set of model frameworks and models will be shown to illustrate how structural conditions, interactions among qualities induce predictable or unpredictable temporal, spatiotemporal patterns.

4.2 Chemical Kinetics: A Prototype of Nonlinear Science

The intention of the theory of chemical kinetics is to describe the interactions among the components (species) of a chemical system [164]. The state of a chemical system can generally be characterized by a finite-dimensional vector. The dimension of the vector is the number of interacting qualities (i.e., components); the value of the vector describes the quantities of the qualities. A chemical reaction is traditionally conceived as a process during which chemical components are transformed into other chemical components. Stoichiometry investigates the static, algebraic relation of the network of reactions.

Rules of reaction kinetics govern the velocity of the composition change. The most often used rule to prescribe the dynamics of the system is the “mass action law”.

The velocity of the reaction

$$aA + bB \xrightarrow{k} cC + dD$$

is given by this law as

$$r(t) = -\frac{1}{a} \frac{dc_A}{dt} = -\frac{1}{b} \frac{dc_B}{dt} = \frac{1}{c} \frac{dc_C}{dt} = \frac{1}{d} \frac{dc_D}{dt} = k(c_A(t))^a(c_B(t))^b,$$

where the scalar $k$ is the rate constant characterizing the velocity of the process.

The kinetic behavior of a chemical reaction is traditionally described by a system of (generally) nonlinear differential equations:

$$\dot{c}(t) = f(c(t); k); \quad c(0) = c_0,$$

where $c$, an $m$-dimensional concentration vector ($m$ is the number of the components), is the state of the system, and the $f$ function, which determines the temporal evolution of the system is determined by the stoichiometry, $k$ is the vector of the parameters, i.e., of the rate constants, and $c_0$ is the initial value vector of the components.

However, not all kinds of differential equations, not even the special class of ODEs with polynomial right-hand side, can be considered as reaction kinetic equations. Trivially, the term $-kc_2(t)c_3(t)$ cannot occur in a rate equations for referring to the velocity of $c_1$, since the quantity of a component cannot be reduced in a reaction in which the components in question does not take place. Putting it another way, the negative cross effect is excluded.$^1$

The deterministic models of classical kinetics are appropriate only when the system can be considered as macroscopic, and even then the deviations from the “average values” remain negligible. A number of situations can be given, when the fluctuations are relevant:

- The size of the chemical system is small. In this case the state space is discrete, the continuous approximation is very bad. Discrete state space,

$^1$ Please note that equations (3.34) and (3.37) are not kinetic equations, since they contain a term with negative cross-effect. Even more, orthogonal transformations cannot transform these equations into kinetic ones [516].
but deterministic models are also out of question, since fluctuations cannot be neglected even in the “zeroth” approximation, because they are not superimposed upon the phenomenon, but they represent the phenomenon itself.

- The system operates near instability point of a deterministic model. In this case small fluctuations may be amplified and produce observable, even macroscopic effects.

- Fluctuations\(^2\) can be a source of information. The fluctuation-dissipation theorem connect the spontaneous fluctuations around an equilibrium state, and the dissipative process leading to equilibrium. Using this theorem applied to chemical kinetics, rate constants can be calculated from equilibrium fluctuations.

A continuous time discrete state space stochastic model is defined to describe chemical fluctuation phenomena. The concentration, (a continuous variable) should be transformed into a function of the number of components (discrete variable): \(c(t) \rightarrow V^{-1} n(t)\), where \(V\) is the volume of the system, \(c(t)\) is a concentration vector and \(n(t)\) is the number of components at a fixed time \(t\).

Introducing a stochastic description let \(\xi\) be a stochastic vector process, the dimension of which is equal to the dimension of the concentration vector. The state of a system is described by the \(P_n(t) = P(\xi(\omega) = n)\) probability distribution function. The temporal evolution of the distribution is determined by the assumption that the process is considered as a Markovian jump process. Markovian character of a stochastic process means that the change of state does not depend explicitly on the past values of the process, so the process does not have memory. the future depends only on the current time step.

The structure of a stochastic model is more complicated than of its deterministic counterpart. The equation for the temporal evolution of the absolute

\(^2\) The Scottish botanist Robert Brown discovered the existence of fluctuations when he studied microscopic living phenomena. However, the physical nature of the motion, which was named after its discoverer, was not known for a long time. As Darwin wrote in 1876: “I called on him [Brown] two or three time before the voyage of the Beagle (1831), and on the occasion he asked me to look through a microscope and describe what I saw. This I did, and believe now that it was the marvelous currents of protoplasm in some vegetable cell. I then asked him what I had seen; but he answered me: This is my little secret.” The theory of Brownian motion was given by Einstein (1905) [148] and Smoluchowski (1906) [545], who calculated the temporal change of the expectation of the square displacement of the Brownian particle, and the connection between the mobility of the particle and the – macroscopic – diffusion constant.
distribution function (called the master equation) is derived from taking into account two types of elementary reactions. The effect of the first class of reaction is that the state $j$ is available from $l$ ($l$ can denote different possible states). The second class describes all of the possible transitions from state $j$. Therefore we can write:

$$\dot{P}_j(t) = \sum_i \text{transition to } j \text{ from } i - \text{transition from } j \text{ to } i.$$  

### Stochastic Chemical Reaction: An Example

Here is a simple example to show the importance of fluctuations. Let us consider the reaction ([164], Sect. 5.6.1):

$$
\begin{align*}
A + X & \xrightarrow{\lambda'} 2X \\
X & \xrightarrow{\mu} 0.
\end{align*}
\tag{4.8}
$$

where $A$ is the external and $X$ is the internal component, and 0 denotes a so-called zero complex. This reaction can be associated with a simple birth-and-death process. The deterministic model is the following:

$$dx(t)/dt = (\lambda - \mu)x(t); \quad x(0) = x_0, \tag{4.9}$$

where $\lambda = \lambda'[A]$. The solution is

$$x(t) = x_0 \exp(\lambda - \mu)t. \tag{4.10}$$

If $\lambda > \mu$, i.e., the birth rate constant is greater than the death rate constant, $x$ is exponentially increasing function of the time. For the case of $\lambda = \mu$

$$x(t) = x_0. \tag{4.11}$$

The stochastic model of the reaction is

$$dP_k(t)/dt = -k(\lambda + \mu)P_k(t) + \lambda(k-1)P_{k-1}(t) + \mu(k+1)P_{k+1}(t) \tag{4.12}$$

$$P_k(0) = \delta_{kx_0}; k = 1, 2, \ldots N. \tag{4.13}$$
There are two consequences of the model:

(1) the expectation coincides with the process coming from the deterministic theory, i.e.,

$$E[\xi(t)] = x_0 e^{\lambda - \mu} t,$$

which in case of $\lambda = \mu$ reduces to the form

$$E[\xi(t)] = x_0.$$  \hspace{1cm} (4.14)

(2) the variance of the process is

$$D^2[\xi(t)] = (\lambda + \mu)t.$$ \hspace{1cm} (4.15)

For the case of $\lambda = \mu$

$$D^2[\xi(t)] = 2D\lambda t,$$

i.e., progressing with time, larger and larger fluctuations around the expectation occur) Fig. 4.1.

![Graph showing the expectation and variance of the process over time.](image)

**Fig. 4.1.** Amplifications of fluctuations might imply instability. While the expectation is constant, the variance increases in time.

It is quite obvious that in this situation it is very important to take the fluctuations into considerations. Such kinds of formal reactions are used to describe the chain reactions in nuclear reactors. In this context it is clear that the fluctuations have to be limited since their increase could imply undesirable instability phenomena.
4.2.1 On the Structure – Dynamics Relationship for Chemical Reactions

A complex chemical reaction (i.e., a mechanism) is a set of elementary reactions. Stoichiometry describes the static, algebraic relationship among the chemical components. Since the kinetic differential equations of the chemical reactions have some special structure, it is possible to derive relationships between the structure of the reactions and its dynamics without solving the system of differential equations itself. This is the spirit of the “qualitative theory of differential equations”. The question was at least to exclude chemical reactions to show “exotic” behavior, which in chemical context means periodicity, chaos or multistationarity. Horn [241] and Feinberg [172, 171] gave negative criteria for the existence of exotic behavior. Here one of their results is described by using their notation.

Zero Deficiency Theorem

Let the chemical components or species of a mechanism be $A(1), \ldots, A(M)$, the complexes

$$C(n) = \sum_{m=1}^{M} y^m(n) A(m) \quad (n = 1, 2, \ldots, N).$$  \hspace{1cm} (4.16)

Therefore the complex vectors of stoichiometric coefficients are

$$y(n) = (y^1(n), \ldots, y^M(n))^T \quad (n = 1, 2, \ldots, N),$$  \hspace{1cm} (4.17)

$T$ denotes the transposed vector.

The elementary reactions of the mechanism are

$$C(i) \to C(j) \quad (i \neq j; i, j = 1, 2, \ldots, N).$$  \hspace{1cm} (4.18)

Let $L$ be the number of the connected subgraphs of the directed graph formed by the complexes as vertices and reactions as edges, i.e., $L$ is the number of linkage classes.

The reaction is weakly reversible, if the transitive closure of the reaction determined by the above defined directed graph is a symmetric relation.

Let $s$ be the dimension of the stoichiometric space $S$, where

$$S := \text{span}\{y(j) - y(i); C(i) \to C(j)\}$$  \hspace{1cm} (4.19)
The deficiency of the mechanism is

\[ \delta := N - L - s. \]  

(4.20)

According to one of the assertions of the zero deficiency theorem if a chemical system with \( \delta = 0 \) is weakly reversible, then for mass action kinetics with any choice of positive rate constants the existence, uniqueness and asymptotic stability of positive equilibrium point follows, i.e., the exotic behavior of these systems is excluded.

**Examples**

Let’s see two, slightly different mechanisms, this is mechanism A:

\[
\begin{array}{c}
A(1) + A(2) \\
\downarrow \quad \downarrow \\
2A(4) \\
\uparrow \quad \uparrow \\
A(1) \\
\end{array}
\quad\begin{array}{c}
A(1) \\
\quad \rightarrow \\
A(2) \quad \rightarrow \quad A(3) \\
\end{array}
\]

In this mechanism there are four components, five complexes, and the number of linkage classes is two. This is mechanism B:

\[
\begin{array}{c}
A(1) + A(2) \\
\downarrow \quad \downarrow \\
2A(3) \\
\uparrow \quad \uparrow \\
A(1) \\
\end{array}
\quad\begin{array}{c}
A(1) \\
\quad \rightarrow \\
A(2) \quad \rightarrow \quad A(3) \\
\end{array}
\]

Complex \( 2A(3) \) replaces complex \( 2A(4) \) of mechanism A. Both mechanisms are weakly reversible, \( N = 5 \) and \( L = 2 \).

However, for mechanism A, \( s = 3 \), and for mechanism B, \( s = 2 \). Therefore the deficiency \( \delta \) is zero for mechanism A, and one for B. The zero deficiency theorem ensures certain stability properties of mechanism A, which is not guaranteed for mechanism B.

### 4.2.2 Chemical Kinetics as a Metalanguage

The structure and models describing chemical reactions are almost trivial. Chemical kinetics generally takes into consideration binary, and rarely, ternary interactions among the molecules. It is an extensively used procedure to decompose (not only chemical) complex phenomena into binary, or perhaps
ternary interactions. Therefore the formal theory of chemical kinetics can be extended to describe transformation phenomena in other-than-molecule populations.

Chemical kinetics is able to describe competition, cooperation and selection among the constituents. A number of biomathematical models at different hierarchical levels, such as prebiotic chemical [146, 147], genetic, population dynamic and evolutionary models lead to the same type of differential equations. The replicator equations [461] have the form

$$\dot{x}_i(t) = x_i(t)\left( f_i(x(t)) - \sum_{j=1}^{n} x_j(t)f_j(x(t)) \right). \quad (4.21)$$

The replicator equation and its variations are important mathematical models of evolution, and we shall return to them in Sect. 4.5.

### 4.2.3 Spatiotemporal Patterns in Chemistry and Biology

A variety of spatial structures such as chemical fronts or waves, periodic precipitates and stable spatial patterns have been subject of studies, both experimental and theoretical, for decades. It is widely accepted that mechanisms responsible for patterns and order in a reaction-diffusion system may also play a fundamental role in understanding certain aspects of biological pattern formation. Though it is evident that chemical reactions and diffusion processes are everywhere and always present in biological systems, one question seems to be difficult to answer completely: are the pattern forming reaction-diffusion systems the real basis of biological morphogenesis or do they offer some analogy to obtain insights into the mechanism of pattern formation? How are patterns of sea-shell forms and of animal coasts (zebras, giraffes, lions etc.) formed?

Since chemical reactions and diffusion processes are the “consequences” of the interactions among chemical constituents, chemical patterns may be interpreted at the molecular level. In the first example, stationary spatial chemical patterns (i.e., the so-called Turing structures) are reviewed. Then we follow with a general theory of biological pattern formation applied among others to the regeneration of Hydra’s head after its removal. Finally a very important phenomenon, somitogenesis, a challenging problem of biological morphogenesis is discussed.
Turing Structures

Alan Turing (1912–1954) From T. machines to T. structures. Computer scientists remember Alan Turing for his fundamental contribution around 1936 to theory of computability. Cognitive scientists celebrate his paper of 1950 “Computing machinery and intelligence”. For the biologists (I guess, mostly for mathematical biologists), Turing’s main achievement is the 1952 Royal Society paper “On the chemical basis of morphogenesis”. And of course, he contributed very much to break the secret codes of the Germans during the war. Probably the different fields he worked can be interconnected: that algorithms corresponding to local mechanism of interacting components produce ordered structures.

Turing wanted to show the possibility of the emergence of spatially inhomogeneous (but temporally stationary) stable structures starting from the perturbation of (spatially) homogeneous structures. He constructed a model which he thought to be a reaction diffusion system in which there exists a spatially stable temporally homogeneous stationary state which loses its stability as a result of inhomogeneous perturbations.

Turing’s example was

\[
\begin{align*}
\dot{x} &= 5x - 6y + 1 + D_x \Delta x, \\
\dot{y} &= 6x - 7y + 1 + D_y \Delta y.
\end{align*}
\]

(4.22)

(4.23)

This model, where \( D_x \) and \( D_y \) are diffusion constants, and \( \Delta \) is the second spatial derivation (Laplacian) operator, was able to produce the formation of stable spatial structures and influenced the way of thinking on chemical morphogenesis, though main stream biology neglected it for decades. The importance of this paper, and the relevance of diffusion-driven instability has been recognized by theoreticians. It influenced people in all the three schools (catastrophe theory, dissipative structures, synergetics) very much. It is obviously an abstract model, such numbers as 5, and 6, are not realistic in terms of elementary mechanisms. Furthermore, and more importantly, a term \(-6y\) on the right hand side of the first equation, i.e describing the change of \( x \), materializes “negative cross effect”. Equation (4.23) is not a system of differential equations of chemical kinetic models.

It was shown a few decade later [502, 503] that the presence of cross-inhibition (i.e., \((\partial f_i/\partial x_j)(c) < 0\)) is a necessary condition of Turing instability.
This result implies that the presence of higher than first order reactions is a necessary condition of Turing instability.

However, no well defined experiments had been made until the late eighties. All the systems exhibiting spatial patterns either contained convection or surface effects, thus the origin of pattern formation has never been pure Turing instability. It was putting the CIMA reaction into a carefully designed ‘gel ring reactor’ [384] (in Patrick DeKepper’s group in Bordeaux) [92] which is generally considered to have produced the long-sought-for result first: the emergence of stationary patterns as a result of diffusion driven instability. Lengyel and Epstein [308] were able to illuminate an important condition of generation of Turing structures: it should be a great difference between the diffusion constants of the various species, and it is provided by the starch indicator present in the CIMA system. Detailed analysis showed that the same dynamical system may lead to different spatial patterns, such as hexagonal blobs and stripes, see Fig. 4.2.

![Fig. 4.2.](image)

3 For a thorough mathematical analysis of the CIMA system, see [383].
A Generative Principle of Biological Pattern Formation: The Gierer–Meinhardt Model Framework

A family of models was developed by Hans Meinhardt and Alfred Gierer based on the interaction of two types of (partially hypothetical) morphogen molecules (activators and inhibitors) and the assumption (which was already assumed by Max Delbrück, whom the reader already knows well from previous chapters). It was assumed that the interaction for different initial arrangements and parameter values may lead to different patterns [349].

Suppose that there are two types of diffusible chemicals, an activator \((a)\) and an inhibitor \((h)\), that are produced at the same place in the animal. Both depend on the coordinate \(x\) and time \(t\). The rate of change of \(a\) is:

\[
\frac{\partial a}{\partial t} = \rho + ka^2/h - \mu a + Da \frac{\partial^2 a}{\partial x^2},
\]

(4.24)

where \(\rho\) is the production rate, \(ka^2/h\) expresses that the generation of \(a\) is an autocatalytic process which is hindered by the inhibitor \(h\), \(\mu\) is a decay constant and \(Da\) is diffusion constant.

The inhibitor has a decay in time and it can diffuse as well, but its generation is triggered by the activator:

\[
\frac{\partial h}{\partial t} = ca^2 - \nu h + Da \frac{\partial^2 h}{\partial x^2}.
\]

(4.25)

Of course, the details might be subject of modifications, the important thing is that the system should be locally unstable and globally stable.

Models and Reality. A Case Study: Somitogenesis

Somitogenesis is the segmentation process in vertebrates and cephalochordates which produces a periodic pattern along the head-tail axis of the embryo. Somites are formed by the successive segmentation of the presomitic mesoderm so that the first somite is generated at the most cranial end of the embryo and segmentation propagates caudally, to the direction of the tail (Fig. 4.3a, b). Somites are then divided by a fissure into anterior (A, front) and posterior (P, back) halves that differ in their gene expression and differentiation (Fig. 4.3c). Further differentiation of somitic cells leads then to the formation of bones, musculature and connective tissue of the skin.

\(^4\) Fishlike animals having a flexible, rod-shaped body instead of true spinal column present in vertebrates. Cephalochordates are probably the closest living relatives of the vertebrates.
Fig. 4.3. Somitogenesis in chicken embryo. Formation of somites starts at the cranial side of the embryo (a) and proceeds in caudal direction (b). (c) For the formation of vertebrae, cells of the anterior part of the somite migrate in cranial while cells of the posterior part migrate in caudal direction. Adapted from [401].

The above mentioned mechanism of pattern generation in somitogenesis is a strongly investigated but still an unresolved problem in developmental biology. Two of the existing models will be reviewed in the following paragraphs.

**Reaction–Diffusion Model**

Meinhardt’s somitogenesis model is derived from his former reaction–diffusion models of morphogenesis [349]. These models are able to produce structures with spatial periodicity that are reminiscent of the periodic anterior-posterior half-somite pattern in somitogenesis. We introduce such a reaction-diffusion model first.

Suppose that there are two types of diffusible chemicals, an activator ($a$) and an inhibitor ($h$), that are produced at the same place in the animal. Both depend on the coordinate $x$ and time $t$. The rate of change of $a$ is:

$$\frac{\partial a}{\partial t} = \rho + k\frac{a^2}{h} - \mu a + D_a \frac{\partial^2 a}{\partial x^2},$$

where $\rho$ is the production rate, $k\frac{a^2}{h}$ expresses that the generation of $a$ in an autocatalytic process which is hindered by the inhibitor $h$, $\mu$ is decay constant and $D_a$ is diffusion constant.

The inhibitor has a decay in time and it can diffuse as well, but its generation is triggered by the activator:

$$\frac{\partial h}{\partial t} = ca^2 - \nu h + D_a \frac{\partial^2 h}{\partial x^2}.$$
4.2 On the Structure – Dynamics Relationship for Chemical Reactions

Figure 4.4 shows the periodic structure that this interplay between activator and inhibitor leads to.

![Figure 4.4](image)

**Fig. 4.4.** Spatial structure generated as the result of the interplay between activator and inhibitor. Adapted from [228].

In Meinhardt’s model for somitogenesis a similar periodic pattern is generated first which models the generation of anterior-posterior (A-P) half-somites. Somites are formed only after this by the separation of subsequent A-P pairs (so that one somite will consist of one A-P pair). This is in agreement with experimental observations in the so called bithorax (Drosophila) mutant where it was found that the segmental specification of the insect can be disturbed without changes in the A-P pattern. This suggests that the formation of A-P pattern is the primary and somite formation is the secondary event indeed.

According to Meinhardt’s model [347, 348, 346], a presomitic cell can be in two different states: either gene A or gene P is active which leads to the production of substance A or P, respectively. In the former case the cell will become an anterior while in the latter case a posterior somitic cell. The state of a cell can be influenced by the neighboring cells: if they produce substance A, P production will be reinforced and vice verse. The mechanism of this mutual reinforcement can be imagined so that the substances diffuse to the neighboring cells where they function as transcription factors: A activates the gene of P whereas P triggers transcription from the gene of A. Under these conditions small random initial fluctuation in the level of A or P gene activation would be sufficient to generate the periodic A-P pattern (Fig. 4.5a).

However, the resulting pattern is somewhat irregular and the emergence of A and P stripes occurs parallel along the whole axis of the embryo instead of being subsequently generated from cranial to caudal direction. In order to get around these problems a further feature of the cells is introduced, the
Fig. 4.5. Reaction-diffusion model of segmentation. (a) A small random initial difference in the A and P gene activation in the cells automatically leads to the formation of a more or less regular A-P pattern. (b) Regularity of the pattern can be increased by a slight modification of the model. At the beginning all the cells, except for the most anterior ones, switch from state A to P because of the assumed morphogen concentration gradient. Cells near to the A-P border are mutually stabilized while cells far from it switch back to state A. This process continues until a highly regular A-P pattern is generated throughout the whole axis of the embryo. Adapted from [348].

ability to oscillate: if a cell is surrounded mainly by cells of the same state, it switches to the other state as explained above. If the neighboring cells switch their states as well, the process can be repeated and cells will oscillate between states A and P.

According to the improved model, at the beginning of the segmentation process all presomitic cells are in state A and there is a morphogen concentration gradient in the embryo increasing from cranial to caudal direction. Cells exposed to a morphogen concentration above a certain (low) threshold switch from state A to P by which the first A-P border is born (Fig. 4.5b). A and P cells near the border stabilize each other while all the farther P cells switch back to state A, because of their oscillating tendency, generating a new P-A border. This A-P stripe forming process continues until it reaches the caudal end of the embryo. By this mechanism a caudally propagating and highly regular A-P pattern can be generated (Fig. 4.5b).
The next step in somitogenesis is the generation of sequential pattern of somites: although somites are similar, they also differ from each other, so a mechanism is needed that distinguishes between subsequent somites. The borders of somites has to be exactly at the P-A borders (each somite containing one A-P pair), that is the mechanism has to ensure the precise superposition of the periodic (A-P) and the sequential pattern. These features can be easily explained by Meinhardt’s model: it is a property of the model that the number of A-P oscillations a cell has made correlates with its position along the anterio-posterior axis (Fig. 4.5b). Assuming that every switching from state P to A activates a new gene, different sets of genes will be active in different AP pairs, which can cause the differences between somites. By this mechanism somite borders would automatically coincide with P-A borders.

**Cell Cycle Model**

In the late 1980s observations from single heat shock experiments suggested that cell cycle correlates to somite segmentation [415]: heat shock to chick embryos resulted in anomalies separated by constant distances of six to seven somites, the number of somites that develop during 9 hours which is the duration of the cell cycle. As presomitic cells leave the Hensen’s node (where they are derived from) and settle along the anterio-posterior axis of the embryo strictly in the order in which they were formed (Fig. 4.6), the observation was explained by that heat shock affects an oscillatory process (the cell cycle) within the presomitic cells. This view is supported by further observations as well: similar periodic anomalies can be caused by drugs inhibiting cell cycle progression [415], and there is some degree of cell cycle synchrony between cells that are destined to segment together to form a somite [490].

Based on these experiments Stern et al developed the “cell cycle model” [490]: P1 and P2 denote two time points in the cell cycle, about 90 min apart (Fig. 4.6). Cells that reach P2 are assumed to start to produce a signal to which cells, whose cell cycle is between P1 and P2, will respond by increasing their adhesion to each other (after this they become unable to signal). As mentioned above, cells along the embryonic axis are located according to their time of formation so neighboring cells are in similar cell cycle state. Therefore cells between P1 and P2 are close to each other and due to the adhesion molecules they produced they will aggregate and start to form a somite.

Collier et al. [102] proposed a mathematical formulation of this model: the system is described by two state variables, the concentration of the adhesion molecule \( g(x, t) \) and the signal molecule \( s(x, t) \). The model equations are:

\[
\frac{\partial g}{\partial t} = \frac{(g + \mu s)^2}{\gamma + \rho g^2} \theta(x, t) - \eta g,
\]
Cells cycle model of somitogenesis. Cells, whose cell cycle is between $P_1$ and $P_2$, aggregate to form a somite. Adapted from [102].

\[
\frac{\partial s}{\partial t} = \frac{\kappa}{\epsilon + g} \chi(x, t) - \lambda s + D \frac{\partial^2 s}{\partial x^2},
\]

where

\[
\theta(x, t) = H(ct - x + x_1),
\]

\[
\chi(x, t) = H(ct - x + x_2).
\]

$H$ is the Heaviside function, $x_i$ is the position of $P_i$ on the x axis when $t = 0$, for $i = 1, 2$, thus $g$ and $s$ production can only occur in cells that passed $P_1$ and $P_2$ state, respectively. The $(g + \mu s)^2/(\gamma + \rho g^2)$ term in equation X expresses...
that production of $g$ is autocatalytic, it is enhanced by $s$ and saturates for large $g$ whereas $-\eta g$ represents linear degradation. The production of $s$ is inhibited by $g (\kappa/(\epsilon + g))$, its concentration also decays linearly ($\lambda s$) and it diffuses along the axis (last term).

Figure 4.7 shows the numerical solutions of the equations. At the position of $P2$, a peak in $s$ occurs because cells start to produce the signal molecule $s$ here (Fig. 4.7b). $s$ diffuses rapidly so there is a decreasing level of $s$ in the neighboring positions too. $s$ triggers the $g$ production in cells between $P1$ and $P2$ which in turn inhibits $s$ production in these cells. Due to this mechanism a wavefront of $g$ is propagating down the axis in jumps which ensures the aggregation of groups of neighboring cells which in turn leads to somite formation.

![Figure 4.7. Numerical solution of the cell cycle model. Peaks of signal molecule production (b) trigger an abrupt increase in the production of the adhesion molecule (a) which leads to cell aggregation and somite formation. Adapted from [102].](image)

**Chemical Waves**

Zaikin and Zhabotinsky [578] found not only the spontaneous appearance of periodic temporal and spatial patterns in an initially homogeneous chemical systems, but wave phenomena were also demonstrated. Arthur Winfree (1942–2002) [564] discovered the existence of spiral waves experimentally and basically explained them mathematically, based on the partial differential equations of the reaction-diffusion systems. Figure 4.8 shows a typical spiral wave emerging in chemical systems. Spiral waves in an other excitable media, actually in cardiac muscle, was studied even by founders of the cybernetics Wiener and Rosenbluth in 1946, [559] Wave propagation in chemical systems now seems to be well-controllable [277]. Interestingly, it was shown [364] that
by the aid of in excitable media logical gates are emulate, which may serve,
as element of information processing.

Fig. 4.8. The emergence of spiral waves in chemical medium is a prototype of
nonlinear spatiotemporal phenomena.

4.3 Systems Biology: The Half Admitted Renaissance
of Cybernetics and Systems Theory

4.3.1 Life itself

Systems biology is an emergent movement to combine system-level description
with microscopic details. It might be interpreted as the renaissance of cyber-
etics and of system theory, materialized in the works of Robert Rosen (1934–
1998) [438, 440]. In an excellent review Olaf Wolkenhauer [567] explained how
the concepts of systems theory, and of cybernetics were applied by Rosen to
biology, and how his ideas returned now under the name of systems biology.

Rosen gave a formalism, which connected *phenotype* (i.e., what we can
observe directly about an organism) and *genotype* (the genetic makeup). In
particular, phenotype is interpreted as being “caused” by genotype. He also
argued that to understand biological phenotype, in addition to the Newtonian
paradigm, the organizational principles should be uncovered. He realized that
a crucial property of living systems, that while they are thermodynamically
open systems, organizationally they should be closed. To put it in another
way, all components, which are subject of degradation due to ordinary wear
and tear, should be repaired or resynthesized within the cell. Rosen gave
4.3 Systems Biology

a mathematical framework to show how it is possible to do. The original
treatment use a branch of mathematics called category theory and will not
repeated here. We restrict ourselves here to discuss briefly the question why
cells might be considered as self-referential systems.

4.3.2 Cells As Self-Referential Systems

Robert Rosen, by analyzing machines and organisms noticed that a main dif-
ference is that organisms not only make and reproduce themselves but are
also able to repair themselves. Rosen gave the formal framework of what he
called metabolism-repair or (M,R)-system. Recently, Athel Cornish-Bowden,
a British biochemist in Marseilles (yes, we live in global world, don’t we?) and
his coworkers reanalyzed Rosen’s results [312, 112], and attempted to transfer
to the community of biochemists. What is the message to be transferred? Tra-
ditionally, in cell biology, enzymes are considered as proteins which catalyzes
the metabolic conversion of substrates to proteins. These (and any other) pro-
teins, however, have also finite life-times, so they should be resynthesized. But
this synthesis also needs enzymes, so there is a infinite regress. Rosen’s central
result was to show that logically it is possible to avoid this infinite regress,
and the essence of cell is the existence of organizational closure, see Fig. 4.9.
The operation of the cell is controlled internally.

About Rosen:
His work is almost totally unknown to biologists, but it is essential
for placing in a broader context the idea of understanding the parts
of a system in terms of the whole. He tried to analyze metabolism in
terms of what he called metabolism-repair systems, or (M,R)-systems.
“Repair” was an unfortunate choice of term, and what he meant by
it was not repair but resynthesis. In other words, his (M,R)-systems
were an attempt to give mathematical expression to the ideas . . . where
enzymes are explicitly considered as products of metabolism.
From [312].

I belong to that camp, whose members believe that Rosen was way ahead
of his time. However, he offered a purely functional theory, and did not try to
give any structural realization of the (M,R) system. Therefore biologists less
prone to abstract thinking could not prove or falsify his hypothesis. In any
case, some people in the systems biology community now give credit Rosen’s
pioneer work.
Biological Complexity (To Be) Explained

Biological complexity, as we have already seen, has different roots. The fundamental components of living beings, i.e., cells are organizationally closed, as Rosen discussed. Biological systems, however, are open from the perspective of material, energetic and information flow.

Robert Rosen [441] recalls the physicist’s approach, which denies that the mind can be the object (or subject) of legitimate scientific study, since it cannot be identified with objective reality. Rosen’s analysis points out that this kind of objectivity is narrowly understood and based on mechanistic notions. He also remarks that biologists adopt a more narrow concept of objectivity: it should be independent not only from perceptive agents, but also from the environment: to explain wholes from parts, that is “objective”, but parts in terms of wholes, that is not. To put it another way: closed causal loops are forbidden in the “objective” world. Rosen’s conclusion is that the world of systems determined by linear (and only linear) causal relationships belongs to the class of ”simple systems” or mechanisms. The alternative is not a “sub-
jective” world, immune to science, but a world of complex systems, i.e., one which contains closed causal loops.

Leaving the clear and well-organized world of linear causal systems we find ourselves in the jungle of the second order cybernetics (Sect. 2.2.2). Biological systems contain their own descriptions, and therefore they need special methods or “at least” special language. It is rather obvious that, despite the methodological success of the analytic sciences, the marvelous complexity of life cannot be explained completely in terms of physics. I think, the main methodological problem what we should confront is to understand what dynamical systems can and cannot do. Whether or not the framework of dynamical systems is sufficiently rich to tackle novelty generations, emergence of complexity. I have to tell the kind Linear Reader that the intention of the whole book is to show how well dynamical systems can be applied to understand the emergence of complexity, but its limit also will be analyzed.

4.3.3 The Old–New Systems Biology

As opposed (better yet, complementary) to molecular biology, the systems biological approach emphasizes the integration of components (mostly proteins and genes) by dynamical models. Slightly modifying Kitano’s approach [283] I believe that systems level understanding requires the integration of five basic properties.

Architecture. The structure (i.e., units and relations among these units) of the system from network of gene interactions via cellular networks to the modular architecture of the brain are the basis of any system level investigations.

Dynamics. Spatio-temporal patterns (i.e., concentrations of biochemical components, cellular activity, global dynamical activities such as measured by electroencephalogram, EEG) characterize a dynamical system. To describe these patterns dynamical systems theory offers a conceptual and mathematical framework. Bifurcation analysis and sensitivity analysis reveal the qualitative and quantitative changes in the behavior of the system.

---

5 Such kinds of languages were offered by Maturana and Varela [331] speaking about autopoiesis. Autopoiesis is term to describe the basic properties of living systems, as their ability to self-created.

6 While nowadays systems biology generally not incorporates the brain, we don’t see any reason to exclude it [162]. In Sect. 8.5.2 the application of the systems biological perspective to neuropharmacology will be briefly presented.
Function. This is the role that units (from proteins via genes, cells and cellular networks) play to the functioning of a system (e.g., our body and mind).

Control. There are internal control mechanisms which maintain the function of the system, while external control (such as chemical, electrical or mechanical perturbation) of an impaired system may help to recover its function.

Design. There are strategies to modify the system architecture and dynamics to get a desired behavior at functional level. A desired function may be related to some “optimal temporal pattern”.

I see systems biology as a relation of these five properties as Fig. 4.10 shows.

Fig. 4.10. Interdependence of the key properties of biological systems. Architecture is defined in a general sense. Structure is supplemented with parameters and initial conditions to specify a model. The role of control is to shift the system dynamical state to a desired one.

Systems biology adopts a set of different modeling techniques from deterministic and mostly stochastic chemical kinetics to more phenomenological description of both genetic and biochemical networks [68, 135]. For the phenomenological description the prototypical example is the random Boolean network suggested by Stuart Kauffman. While the original idea was suggested in 1969 [269], the model framework has been subject of extensive studies from the mid-eighties, when Kauffman moved to the Santa Fe Institute (and in any case, computers with large capacity appeared).
4.3.4 Random Boolean Networks: Model Framework and Applications for Genetic Networks

The Structure of the Model

Kaufman introduced the so-called “NK model” in the context of genetics and evolution [269, 270]. Its generalization became a quite generic dynamical system, which is able to show regular and irregular behaviors, and was applied not only in biology, but also in many disciplines from computer science to social dynamics. There is an excellent review of the model framework[8]. The model is called NK, since there are $N$ elements, each of them interacting with $K$ others. The model is Boolean, since the $N$ variables are Boolean, i.e., they take values 0 and 1. Formally it is a discrete time binary state model, the dynamics is defined by the updating rule

$$\sigma_i(t+1) = f_i(\sigma_{j1(i)}(t), \sigma_{j2(i)}(t), ..., \sigma_{jK(i)}(t)),$$  \hspace{1cm} (4.26)

Where $\sigma_i(t)$ is the value of an element.

To specify the dynamics we should know the number of elements which influence all the $i^{th}$ variables, (connectivity and linkages, respectively) and the updating rule $f$. While in the generalized model framework, the number of influencing variables may be different for each node, The Kaufman model assumes that the number of connections are the same, $K$. For variable connectivities it is possible to define the average connectivity of the system $<K> = 1/N \sum_{i=1}^{N} K_i$.

The dynamic behavior of the system strongly depends on the structure of the coupling. It is possible to assume some topology among the variables, say, they form a hyper-cubic lattice, and one may assume that the state of the elements depend on their somehow defined neighbors. Alternatively, the dependence may be chosen from a uniform distribution, so each element has equivalent chance to influence the others. This random uniform Boolean network is called as the Kaufman net. (One may define coupling, which is intermediate between the lattice-like and the purely random, i.e., “small world” networks, see Sect-7.4.1).

The number of coupling functions is $2^{2K}$, so for $K = 3$ there are 256 functions. The set of functions can be classified into groups. E.g., there is a set of canalizing functions, in which the value of the functions is determined by one of its argument only. Say, the value of the $f(\sigma_1, \sigma_2, \sigma_3)$ might be 1, for $\sigma_1 = 0$, and 0 for $\sigma_1 = 1$, independently of the values of $\sigma_2$ and $\sigma_3$. 

Another possibility is to use weighted \( f_i \)s, i.e., the functions are weighted with probabilities \( p \) and \( 1 - p \).

### Characterization of the Dynamic Behavior

**Information flow and phases**

One characteristic property of the dynamics is its sensitivity to initial state. Assume two initial states:

\[
S(t) := [\sigma_1(0), \sigma_2(0), ...\sigma_N(0)]
\quad (4.27)
\]

\[
S^+(t) := [\sigma_1^+(0), \sigma_2^+(0), ...\sigma_N^+(0)].
\quad (4.28)
\]

A time-dependent measure for the (of course time-dependent) distance is defined as

\[
D(t) := \sum_{i=1}^{N} (\sigma_i(t) - \sigma_i^+(t))^2.
\quad (4.29)
\]

Let’s assume that the distance between the two initial state is small. If the information flow is localized, the distance remain relatively small (the system is said to be in *frozen* state). In certain situations \( D(t) \) can diverge for large times, so the information can be transferred to the whole system (the system is said in *chaotic phase*).

Simple assumptions lead to

\[
D(t) = D(0) \exp[t \ln(0.5K)].
\quad (4.30)
\]

The distance is measured by the Hamming distance.\(^7\)

The analysis of this equation implies that three different phases occur by varying the value of the connectivity \( K \):

- **frozen**: if \( K < 2 \) (so \( K = 0 \) or \( K = 1 \)), the Hamming distance is exponentially decaying function of time.
- **chaotic**: if \( K > 2 \), the Hamming distance increases exponentially with time.
- **critical**: \( K = 2 \), the system’s behavior critically depends on fluctuations.

\(^7\) Hamming distance between two strings of equal length is the number of positions for which the corresponding symbols are different.
This critical behavior became famous (and as it often happens with too successful notions, a little bit infamous), as the “edge of chaos”. But first let’s speak about another characterization of the dynamics.

**Cycles**

The state space is finite, and contains $2^N$ configurations. Starting the system from any initial condition there is a cyclic attractor, where the system tends. The different cyclic attractors are characterized by their length and the size of the basin of attraction. The different phases might have very different cycles. The frozen phase has relatively short cycles, and short transients. The chaotic phase contains long cycles and long transients. The cycle length grows as a power of the size. In the critical phase these dependencies are algebraic, and it was found that specifically it is proportional to the square root of $N$. This result seemed to support the applicability of the whole model framework, since such kinds of behavior was found in biological experiments, what we will review soon. It might be the case, that the dependency is linear, as it was suggested several years ago [62], and the situation is under dispute.

The structure of the nodes is very important for the dynamics of Random Boolean Networks. The descendants of a node are the nodes that it affects, while the ancestors of a node are those that affect it. To have cycle attractors, i.e., of period greater than one, there should be at least one node that will be its own ancestor. A circuit of auto-activating nodes is called a linkage loop, and when there is no feedback, linkage trees are formed. Note that loops spread activation through trees, but not vice versa. The relevant elements of a network are those nodes that form linkage loops, and do not have constant functions, for these cause instabilities in the network, which might or not propagate. Note that as there are more connections in a network (higher $K$), the probability of having loops increases. Therefore, finding less stable dynamics for high values of $K$ is natural. From [86].

**Edge of Chaos?**

The critical phase separates the frozen (ordered) and the chaotic (disordered) phase. The verbal hypothesis is that complex systems evolve that way that their parameters will be in the ordered phase, close to the border of the chaotic phase.
It seems to be plausible, in general terms, that somehow, complexity lies between the regular determinism and randomness/chaos. Specific calculations [355] with cellular automata (a method, offered by John von Neumann, see Sect. 5.5) suggested some cautiousness to avoid over-interpretation of the notion. In any case, Kaufman managed to show, that the model framework was very useful to get insight about the operation of genetic networks. Every model framework may be subject of criticism (and actually, they have been), including the random Boolean $NK$ model. No doubt that the model framework is beautiful, illuminating, and helps to understand many aspects of genetic networks. OK, not all, but it never was expected.

$NK$ Modeling of Genetic Networks

The development of cells within the developing embryo into different cell types happens in a process regulated by certain DNA sequences, i.e. “regulator genes”. These genes produce (“express”) proteins that regulate other genes. Interaction of genes can be modeled by networks. Genes are the nodes, and two genes are connected by an edge if the product protein of one gene influences the expression of the other one. While it was obvious that the genetic network is not inherently random, due to the fact the linkages were complicated, and not well-known, the random network approximation was reasonable.

In the model the state of a cell is characterized

- by the state of its genes denoted by the vector $S(t) := \sigma_1, \sigma_2, \ldots, \sigma_N$.
- the state of each gene is binary, i.e., a gene is either “on”, or “off”.
- each gene is connected to the same number $K$ other genes, (and canalizing functions are used in some cases).
- the linkages among the genes are selected randomly.

The dynamics of the gene network in the $NK$ model is specified by the following assumptions:

- The updating rule, or evolution function $f_i$, related to $\sigma_i$ gets value 1 with probability $p$, and value 0 with probability $1 - p$.
- The network updates is synchronous, the state of each gene is calculated in the same time.
Under these assumptions, as it was mentioned, cycles, i.e., periodic attractors appear, but the properties are very different in the frozen and chaotic phase. The biological interpretation is that the attractors represent the resulting cell type. Biological plausibility exert constraints on the properties of the attractors. The length of an attractor, (and the transient leading to it) must be not too long, since a cell should reach it stable state during a reasonable time. Since the chaotic phase is characterized by cycles, which show exponential increase with system size, genetic networks are not supposed to be in the chaotic phase.

While in the frozen phase the cycles are short, they have an other property, which does not support their biological relevance. The number of relevant elements in the frozen phase is close to be zero. Therefore the system is very resistant for point mutations (i.e., when the value of a single gene is subject to change). A real genetic network, which is able to evolve, should have some degree of sensitivity to mutations.

Kaufman’s suggestion, that biologically realistic genetic networks should operate at the edge of chaos, i.e., the systems evolve that way that they are at or close to the critical phase. In this phase the size-dependence of the number and length of attractors is linear, it ensures a relatively quick convergence to stable, and not very long periodic attractors.

One family of experimental data supported the view, that genetic networks perform at the edge of chaos: the number of different cell types is more or less proportional to the square root of its DNA content. Disputes about the specific form of the size dependence are weakening this argument. An other problem is that in real genetic network the connectivity is much larger than two. To keep the system at the critical phase weighted or canalized, updating rules should be used.

**Concluding Remarks**

The random Boolean network model framework has been very popular, since it was intuitively well understandable, biologically plausible. It was able to produce rich dynamic behavior, and specifically offered a concept, which reached the “popular culture”. The view, that “interesting” things happen between the regions of order and randomness is appealing. Intellectually it grew from the John von Neumann’s cellular automata concept. Many extensions of the model framework exist. Among others, instead of assuming constant $K$ for all genes, the input distribution may be taken from some distribution and the network might be so-called scale-free. The concept of “scale-free” will be explained later in Sect. 7.4.1. (Of course, many mathematical results obtained in the last 20+ years are restricted to constant connectivity). Kaufman sees clearly the scope
and limits of the original framework, and participates in developing it. For his own analysis, see [271].

4.4 Population Dynamic and Epidemic Models: Biological and Social

4.4.1 Connectivity, Stability, Diversity

The fundamental question to be answered is how does the stability of an ecological system change if there is a change in the size and/or connectivity of the network of interacting elements, and/or in the strength of the interactions.

The May–Wigner Theorem

While ecologists believed for a while that diversity and stronger interactions among species enhance stability, Robert May [332, 333] proved for certain system sizes connectivity cannot exceed a threshold to ensure stability. The story goes back to cybernetics. Ashby [28] has found and Gardner [193] is said to have found that the probability of stability exponentially decreases with system size (i.e., with diversity). The same authors [194] reported the above mentioned threshold effect.

May [332, 333] published theoretical results, for randomly assembled deterministic systems by using linear stability analysis. The mathematics is a corollary of Wigner’s theorem [560] on the eigenvalues of certain random matrices. May’s argument follows.

Let $B_n$ be a random $n \times n$ matrix with connectivity $c_n$, i.e., $B_n$ has $c_n n^2$ non-zero elements. The non-zero elements are chosen independently from a fixed symmetric distribution with mean 0 and variance $\alpha_n^2$. The interaction matrix is defined as $A_n := B_n - I_n$, and the trivial solution of the differential equations $\dot{x} = A_n x$ is for large $n$ almost surely stable if $\alpha_n^2 n c_n < 1$ and almost surely unstable if $\alpha_n^2 n c_n > 1$. (The number $\alpha_n$ measures the range of the distribution, or it controls the magnitude of the elements of the matrix $A_n$, i.e., it characterizes the strength of connections).

The problem of connecting the structure and dynamics of ecological systems has been discussed as the diversity-stability debate [338, 337].
May’s results suggested that both diversity, and too strong connections tend to destabilize the equilibrium state of a community. The state of a community, in a more general ecological sense can be characterized by different measures. *Richness* is the total number of species, while *diversity* takes into account, somehow, both richness (the *dimension* of a vector), and the quantity of each species (the *value* of each vector component). *Complexity* of a community is determined by the richness, level of connectedness (number of non-zero elements in the interaction matrix), and the strength of the interactions.

**Measures of Ecological Stability**

The stability of complex ecosystems cannot be restricted to the stability of the equilibrium point [337], and several measures of stability can be defined. There are two classes of stability measures, the first is related to the usual stability concepts of dynamical systems, and the second measures the system’s ability to preserve its function after disturbance (“resilience”, “resistance”).

1. Measures of system dynamics
   - Equilibrium stability: external perturbation implying deviation from the equilibrium state decays and the system returns to the vicinity if the state.
   - General stability: the quantities of the each species vary under limits, the lower limit is far from zero, so extinctions is not a threat.
   - Variability: The variance of the quantities of species is the usual experimental measure of stability. Larger general stability implies smaller variation.

2. Measures of resilience and resistance
   - Equilibrium resilience: is proportional with the inverse of the time necessary for the system to return to its equilibrium state.
   - General resilience: is proportional with the inverse of time necessary for the system to reach its original (not necessarily equilibrium) state.
   - Resistance: e.g the ability of the community to survive invasion of new species.
Weak Interactions Enhance Stability

Fig. 4.11. Elementary food-web structures. C, P and R denote consumers, top predator and resource densities. $\Omega_{ij}$ denotes a parameter expressing the preference of a species $i$ to consume species $j$. (a) Simple food chain; (b) exploitative competition (multiple intermediate consumers); (c) apparent competition (top predator feeding on two intermediate consumers); (d) intraguild competition (i.e., the killing and eating of species that use similar resources and are thus potential competitors); (e) omnivorous predator (“eats everything”); (f) food chain with external input. Based on McCann et al. [338].

McCann et al. [338] used a variation of the Rosenzweig-MacArthur model [444], where (as opposed to the assumption of the Lotka-Volterra model) predators’ kill rate will approach an upper bound as the density of prey increases. Here $V$ stands for the victim component (may be $R$ or $C$), and $P$ is
the predator ($P$ or $C$) in Fig. 4.11.

\[
\dot{V} = rV - \frac{kPV}{\chi + V} - \delta V^2 - kPV
\]

\[
\dot{P} = \frac{\beta kPV}{\chi + V} - mP.
\]  

Here $r$ is the victim species’ Malthusian rate of increase, $k$ the prey-predator kill rate, $\beta$ a conversion factor from victims to baby predators, and $m$ the predator’s death rate, $\delta$ expresses the effects of intraspecific competition, $\chi$ is the density of prey at which the predators’ kill rate reaches half its maximum.

Model studies showed the crucial dependence of the dynamic behavior on the strength of interactions. Many weak,\(^8\) i.e., non-zero, and not too high, (it is difficult to give much more quantitative statements) in addition of a few strong connections tend to stabilize ecological communities. Obviously there is not too much experimental data about interaction strength, but based on the small amount available, and the model calculations McCann states:

\[
\text{It seems ... that the weak interactions may be the glue that binds natural communities together. [338]}
\]

**Further Reading**

For a review on modeling food webs, see [136]. They evaluated the available food web data, discussed several classes of models. Static models assign links between species according to some simple rule. The second class uses general kinetic equations for interacting populations. In the third model new species may arrive through “invasion”. The third model framework may lead to large stable network.

\(^8\)Somewhat (somewhat!) analogously a very important concept about social networks [214, 215], Granovetter’s “The Strength of Weak Ties...” argues that not only strong connections among people (“close friends”) but weak connections (“acquaintances”) also have an indispensable structural role in social organization. Weak ties play the role of the glue between densely connected friendship networks.
4.4.2 The Epidemic Propagation of Infections and Ideas

The Basic Kermack–McKendrick Model

In an excellent, very clearly written book Joshua Epstein [153] demonstrated that a set of sociodynamical phenomena could be interpreted by using nonlinear dynamical models elaborated in biological context. Epidemics is a fascinating topic in his books, since he uses the analogy of “revolution as epidemics”, and models of describing and controlling infectious disease for modeling the propagation of ideas between population of people.

In the simplest situation there are two populations, $I(t)$ denotes the number of already infected (either with biological objects capable of transmit infection or with revolutionary ideas to be transferred to others) individuals, and $S(t)$ is the number of susceptible individuals. It is assumed that who is infected is also infective.

What is now the simplest assumption? The encounter between an infective and a susceptible may imply the transformation of a susceptible to an infective one. Chemical kinetics here also serves as a metalanguage to encode the process:

$$ I + S \xrightarrow{r} 2I, $$

where $r$ is the effectivity of the encounter, i.e., the infection rate: the larger this rate the more contagious the infection. By adopting (i) the mass action kinetic assumption and (ii) “prefect mixing” (i.e., the space is homogeneous), the kinetic equation is:

$$ \dot{S} = -rSI, \quad (4.33a) $$
$$ \dot{I} = rSI. \quad (4.33b) $$

Perfect mixing in social applications should be interpreted, as the lack of the existence of any (spatial or social) organization among the participants.

If the total population is constant (no birth and death, no immigration from and to the community), i.e., $I(t) + S(t) = N$, then the temporal change of the infected population is described by the equation:

$$ \dot{I} = rI(N - I). \quad (4.34) $$

This is the logistic differential equation. The equilibrium solution $I_{eq} = 0$ is unstable. Instability here means that its slightest perturbation implies that $I(t)$ tends to $N$, the whole population will be infected.
The oversimplified model just presented consists of two subpopulations and it is called as the $SI$ model. The classical strategy to defend the susceptibles against infection is to remove infectives, so to apply $SIR$ models. Kermack and McKendrick defined a model, by adding a first order removal to (4.33) and so with the structure $S \rightarrow I \rightarrow R$:

$$
\dot{S} = -rSI \quad (4.35) \\
\dot{I} = rSI - \gamma I \quad (4.36) \\
\dot{R} = \gamma I, \quad (4.37)
$$

where $\gamma$ is the removal rate. (Of course, this is the simplest assumption, the removal is proportional with the actual number of the infective ones). The relative removal rate is defined as $\rho := \gamma/r$. A plausible initial condition is that $S(0) > 0$, $I(0) > 0$ and $R(0) = 0$, nobody is in a quarantine before the epidemic breaks out.

The kinetic condition of the outbreak of epidemic is that the infectious population increases, i.e., $\dot{I} > 0$, i.e., $S(t) > \rho$. The number of susceptibles should exceed a threshold to have an epidemic outbreak. Since epidemics is threshold phenomena, one implication is that there is no necessity for full vaccination of the population to avoid the outbreak.\(^9\)

**Modeling More Realistic Epidemics**

In the original model (4.37) the assumption has been made that the contact rate does not depend on the size of the population. For a set of sexually transmitted diseases it seemed to be a more realistic assumption that the contact rate is a non-increasing function of the population size. Another extension is to take into account the change of the total population size. A simple assumption is that a fraction of $f$ of the infective class recover, and the remaining fraction $1 - f$ die of disease. The extended model based on these assumptions is written [49] as

$$
\dot{S} = -r(N)SI, \quad (4.38) \\
\dot{I} = r(N)SI - \gamma I, \quad (4.39) \\
\dot{N} = -(1 - f)\gamma I, \quad (4.40)
$$

More realistic models take into account age-dependence, delay, spatial heterogeneity, etc. In Sect. 7.4.3 instead of assuming “perfect mixing” among the

\(^9\) Well, I leave to social scientists to analyze which subpopulation of a society will be immunized and which will not...
participant, some (network) organization will be assumed. In the first case
the probability of the encounter of any two members of the population is the
same. For organized communities some people may encounter with a subset
of people more frequently than with others. Actually this organization makes
a difference for the spreading of epidemics, as you will see.

Recently the investigation of the effects of infective immigrants to epidemic
dynamics seems to be important, in context of HIV, severe acute respiratory
syndrome (SARS) and avian influenza and measles (say [464]. From the per-
spective of public health policy the possibility of providing access to health
care for populations of illegal immigrants should be seriously assessed. While
deterministic models proved to be efficient, stochastic models are more faithful
[15].

4.4.3 Modeling Social Epidemics

Epstein’s witty perspective is to use models of (biological) epidemics to in-
terpret social revolutions [153]. To do so, the variables, parameters, and each
term in the equations should be reinterpreted. This reinterpretation requires
to assume that everybody, who is not infected is susceptible (so there are no
neutral, inert individuals), and if the contact between a susceptible and an in-
fected has an effect, always the susceptible changes her attitude. Removal can
be identified with imprisonment. Since in the equation (4.37) there are two
parameters, the revolutionary and counterrevolutionary tactics is to modify
these parameters.

The reduction of the $r$, i.e the contact rate between infectives and sus-
ceptibles can be obtained, say, by restricting the right of assembly, and the
increase of the $\gamma$ removal rate can be attained by increasing the rate of impris-
onment. Ambush help to reduce $\gamma$, and the underground literature “Samizdat”
and Radio Free Europe (a certain geographically and demographically defined
(challenged?) subset of the Readers remember very well) increased $r$.

Supplementing the basic model with the birth of susceptible population
with a rate $\mu > 0$, the following equation is obtained:

\[
\dot{S} = -rSI + \mu S, \quad (4.41)
\]
\[
\dot{I} = rSI - \gamma I. \quad (4.42)
\]

\[10\] The author decided not to discuss in this book what is revolutionary idea and
what is not.
The system of equations is identical to the Lotka–Volterra model. The susceptibles play the role of the preys, and the infectives do that of the predators. The interaction between the populations can be viewed as predator-prey interaction in a political ecological field.

In all these models it was implicitly assumed that the population is completely “mixed”, and the result is that the outbreak of the epidemics is a threshold phenomenon. In Sect. 7.4.3 we shall discuss that assuming certain topological structure of the individuals, the propagation is not threshold. Propagation of computer viruses on network belong to this category. Consequently, the slightest infection of the Internet may propagate through the whole network.

4.5 Evolutionary Dynamics

The replicator equation (originally offered by Taylor and Jonker in 1978 [505]) was already mentioned, in the context of using, as a metalanguage:

\[
\dot{x}_i(t) = x_i(t) \left( f_i(x(t)) - \sum_{j=1}^{n} x_j(t) f_j(x(t)) \right).
\] (4.43)

In the context of evolution, \(x_i\) is identified with the relative abundance of a genetic sequence. This equation expresses that the selection is frequency-dependent, (or abundance-dependent); the fitness of a single individual depends on the abundance of the other sequences. \(f_i(x(t))\) is the fitness of type \(i\), and \(\sum_{j=1}^{n} x_j(t) f_j(x(t))\) is the average fitness, also denoted by

\[
\Phi := \sum_{j=1}^{n} x_j(t) f_j(x(t)).
\] (4.44)

This equation is both broader and narrower than the quasi-species equation offered to describe molecular evolution by Manfred Eigen and Peter Schuster [146, 147]:

\[
\dot{x}_i(t) = \sum_{j=1}^{n} x_j(t) f_j q_{ji} - x_i(t) \Phi.
\] (4.45)

Equation (4.45) is narrower than the replicator equation (4.43), since it does not contain frequency selection. However, there is a factor to take into account the probability of mutations in the replication; \(q_{ji}\) is the probability
of that sequence $i$ will be copied as sequence $j$. The trivial unification of the two equations is the replicator-mutation equation:

$$
\dot{x}_i(t) = \sum_{j=1}^n x_j(t) f_j(x(t)) q_{ji} - x_i(t) \Phi.
$$

The trivial unification of the two equations is the replicator-mutation equation:

$$
\dot{x}_i(t) = \sum_{j=1}^n x_j(t) f_j(x(t)) q_{ji} - x_i(t) \Phi.
$$

Further Remarks

The relationship among the different equations were analyzed by Martin Nowak and his coworkers [394, 385].

The replicator equation has stochastic version, too [181]. Such kinds of concepts, such as stability, survival, extinction have some different meanings in stochastic context. There is another school of modeling evolutionary systems, called adaptive dynamics. It connects phenomena with different time-scales, specifically the more rapid population and ecological dynamics and the slow evolutionary changes. It is a promising field, and one fundamental paper is [198].

We shall return to evolutionary dynamics in Chap. 9, related to digital evolution and evolutionary game theory.

4.6 Dynamic Models of War and Love

4.6.1 Lanchaster’s Combat Model and Its Variations

The Basic Model

R.W. Lanchester, a British engineer derived and analyzed a mathematical model of warfare (e.g., [153]). For a combat between two sides, the state of the system is a two-dimensional vector, the number of soldiers on both sides. The interaction between the two sides implies the reduction of the number of soldiers, i.e., the attrition of forces. The simplest possible assumption is that there is a linear relationship between the attrition of a population of soldiers, and their number of the other side, the $k_b > 0$, $k_r > 0$ rate constants are
the firing efficiencies. By denoting the number of “red” and “blue” combatants with $R$ and $B$, the Lanchester equation is written as:

\[
\frac{dR}{dt} = -k_B B, \tag{4.47}
\]
\[
\frac{dB}{dt} = -k_R R, \tag{4.48}
\]

with the initial values $R(0) = R_0$ and $B(0) = B_0$. Instead of solving the equation, but eliminating its time-dependence, and adopting some stalemate assumption, the Lanchester “square law” is derived:

\[
B_0 = \sqrt{\frac{k_R}{k_B} R_0}. \tag{4.49}
\]

The interpretation of the square law has a dramatic military implications: if your initial adversary twice as numerous as that of your enemy, it is not sufficient for your enemy to double her firing efficiency. To attain stalemate she must be four time as effective. Under this assumption numbers count more than effectiveness. Advantage of numbers strongly influenced the Pentagon strategy, maybe even today.

**Density-Dependence**

The assumption of the Lanchester equation is that the battle field is “spatially homogeneous”, i.e., completely mixed. Also, it assumes that the increase of the numbers of combatants has benefits only. But a soldier is not only predator, but also a prey. Increasing numbers implies increasing the risk of being incapacitated. The *density dependent* Lanchester equation

\[
\frac{dR}{dt} = (-k_B B)R, \tag{4.50}
\]
\[
\frac{dB}{dt} = (-k_R R)B, \tag{4.51}
\]

takes into account both the predation benefit (in parentheses) and the prey cost (proportional with its own number). Under this assumption the stalemate condition is $k_B B_0 = k_R R_0$. Here linear increase in technological efficiency is sufficient to compensate the increase of adversary.
Ambush and Asymmetry

It is reasonable to assume some asymmetry between the troops: one side is able to ambush, the other is not. The model is defined as:

\[
\frac{dB}{dt} = -k_B R, \quad (4.52)
\]

\[
\frac{dR}{dt} = -k_R RB. \quad (4.53)
\]

The stalemate condition is

\[b(B_0^2 - B(t)^2) = r(R_0 - R(t)).\]

Assuming equal firing effectiveness \( r = b \), \( B_0 \) can stalemate Red force with number \( B_0^2 \). That was the case in the Battle of Thermopylae.

Reinforcement

A further extension of the basic model is to incorporate some flow of reinforcements. It is reasonable to set some saturation to this flow, so the equation for the logistic population growth is added to the (density dependent) model:

\[
\frac{dR}{dt} = -k_B RB + \alpha R \left(1 - \frac{R}{K}\right), \quad (4.54)
\]

\[
\frac{dB}{dt} = -k_R BR + \beta B \left(1 - \frac{B}{L}\right), \quad (4.55)
\]

where \( \alpha, \beta, K \) and \( L \) are positive constants. This equations is a generalized Lotka-Volterra equation, a famous model of competition of “the struggle for existence”, given by Georgyi Frantsevitch Gause (1910-1986), a biologist from Moscow. There are four basic solutions of (4.55), and they have different interpretations, as Fig. 4.12 shows. Two cases implements the “principle of competitive exclusion”. In these cases one side wins, since the other population becomes extinct. In the third case there is a stable fixed point (technically a so-called node), with finite coordinates. This scenario implements “coexistence” by using ecological terminology, but in this “combat dynamical” model it means “permanent war”. The fourth case is an unstable (so-called saddle) equilibrium. Any small perturbation implies the win of the one or the other side.
4.6 Dynamic Models of War and Love

Fig. 4.12. Phase plane analysis of the Gause model. The upper two plots implement the “principle of competitive exclusion”. The left lower figure visualizes the “permanent war” situation (the equilibrium point is stable focus). The lower right plot shows an unstable saddle point.

4.6.2 Is Love Different from War?

“Love is what we do.” There are several models to use dynamical models to describe the temporal changes in love. The models are semi-serious, and help motivate students to learn in thinking with dynamical models, but they give an insight of possible scenarios. Based on some previous studies Julien Sprott gave a series of models to describe the dynamics of romantic relationships, say, between Romeo and Juliet [480].

The state of the system is characterized by a two-dimensional vector $(R(t), J(t))$. $R(t)$ is Romeo’s love (or hate if negative) for Juliet at time $t$ and $J(t)$ is Juliet’s love for Romeo. A simple assumption to model the interaction is that their change of their emotional state depends on their own and the other’s state, and the dependence is linear and additive. These assumptions lead to the linear model:

\[
\frac{dR}{dt} = aR + bJ, \quad (4.56)
\]

\[
\frac{dJ}{dt} = cR + dJ. \quad (4.57)
\]
The constants $a, b, c, d$ determine the romantic style of the relationship. If $a > 0$, than Romeo’s feeling is subject of self-amplification (positive feedback), $b > 0$ means that his feelings is encouraged by Juliet’s feeling.

| From emotional attitude to love dynamics |
|-----------------------------------------|
| 1. Eager beaver: $a > 0$, $b > 0$ (Romeo is encouraged by his own feelings as well as Juliet’s.) |
| 2. Narcissistic nerd: $a > 0$, $b < 0$ (Romeo wants more of what he feels but retreats from Juliet’s feelings.) |
| 3. Cautious (or secure) lover: $a < 0$, $b > 0$ (Romeo retreats from his own feelings but is encouraged by Juliet’s.) |
| 4. Hermit: $a < 0$, $b < 0$ (Romeo retreats from his own feelings as well as Juliet’s.) |
| 5. out of ouch with one own’s feeling: $a = 0$ or $d = 0$ |
| 6. oblivious to the other feeling: $b = 0$ or $c = 0$ |

The terminology goes back to Strogatz and his students [491], Chap. 5.3.

There are $6 \times 6 = 36$ possible dynamics, but the $R/J$ symmetry reduces this number to 21. Specifically, the case of two nerds (i.e., $b < 0$ and $c < 0$) with “out of ouch with their owns feelings” is equivalent with the original Lanchester equation. Of course, it not surprising that occasionally love models and war models coincide. The case of $b > 0$ and $c > 0$, i.e mutual activation, also may be interpreted in combat dynamic context.

The model framework has obvious difficulties. It is not trivial to assign numbers to our love state. Furthermore, more realistic parameters should be time-dependent, etc. Still, it is worth to take a look to some possible outcomes of this simple model of love affair dynamics.

Equation (4.57) has a single equilibrium at $R = J = 0$. The nature of the dynamics around an equilibrium is determined by the eigenvalues of the coefficient matrix $A := \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. 
The eigenvalues are the solutions of the characteristic equation (3.18) in this case are

\[
\lambda = \frac{a + d}{2} \pm \frac{1}{2} \sqrt{(a + d)^2 - 4(ad - bc)}. \tag{4.58}
\]

As is well known, nature of the equilibrium points depend on whether the eigenvalues are real, complex conjugate, or purely complex, and we get three classes (focus, node, saddle point), as Fig. 4.13 shows.

**Fig. 4.13.** Dynamics in the vicinity of an equilibrium point. From http://sprott.physics.wisc.edu/pubs/paper277.htm.

The stability depends on the sign of the real parts of the eigenvalues. Some specific examples for the love dynamics follow.

*Fire and Ice*

This situation is defined, if the two lovers are opposites, so \( b = -c \) and \( a = -d \), and the eigenvalues are \( \lambda = \pm \sqrt{a^2 - b^2} \). Since the dynamics depends on \( a \) and \( b \), there are two possibilities:

- \( ab > 0 \): *eager beaver plus hermit*
- \( ab < 0 \): *narcissistic nerd plus cautious lover*

There are two cases, again. If \( a > b \), i.e., the responsiveness to one’s own state is larger than to the other’s, both \( \lambda \) is real, one is positive, and one is negative. The dynamics can end up in any of the four quadrant, so all four combination of love and hate may emerge depending on the initial values. If \( a < b \), i.e., the self-responsiveness is smaller that to responsiveness the other’s state), the equilibrium point is a *center*, as it was explained in Sect. 3.5.1. The dynamics is oscillatory.
**Other Possibilities**

The *peas in a pod* situations is characterized by the interaction of two equivalently romantic characters: \( b = c \), and \( a = d \). If \(|a| > |b|\), the singularity point (i.e., the 0, 0) is a stable node, the relationship will end with mutual apathy (not the worst case scenario). In the *Romeo the robot* scenario, Romeo’s feeling is unchanged \((dR/dt = 0)\). Juliet feeling is determined by the sign of Romeo’s (constant, so constant sign) feeling and her own romantic style. The equilibrium point is \( J_{eq} = -cR/d \). What if Romeo loves Juliet, i.e., \( R > 0? \) Juliet will love him back if \( cd < 0 \). This can be implemented if she either is a cautious lover or narcissistic nerd. But the singularity point is stable only if she is cautious \((d < 0)\). Narcissism \((d > 0)\) will lead either unbound love or hate. Her feeling never goes to zero, or and will not show oscillation between love and hate.

Further most exciting possibilities nonlinear models, love triangles (*ménage à trois*) are also analyzed by Sprott. But even simple linear models produced dynamics with sufficiently interesting insights.

### 4.7 Social Dynamics: Some Examples

#### 4.7.1 Segregation dynamics

Thomas Schelling published a very influential paper in 1971 with the title “Dynamic Models of Segregation” [457]. This is a paradigmatic paper to demonstrate how local rules (micromotives in Schelling’s terminology) imply global ordered social structures (macrobehaviors). Technically it is a *cellular automata* model, space can be one- or two-dimensional. Each player (or agent or turtle, the terminology may differ) sits on a grid point surrounded by others. (You may imagine that the points are people living in a house encircled by their neighbors). The players may have a predominant parameter. This parameter in the USA is a visible color, and it is correlated to the race of the players, and the model describes *racial segregation*.

The model was elaborated in the context of residential segregation, but of course the phenomena cannot be restricted to it, as Beverly Daniel Tatum’s book, “Why Are All The Black Kids Sitting Together in the Cafeteria?” explains [504].

---

11 Of course, other than racial segregation exists, such as gender segregation in the labor market, distribution of different nationalities in an international meeting, etc.
The state of each player is now characterized by the degree of satisfaction. Here this degree depends on the number of neighbors having the same parameter only. The dynamics of the system is determined by a local rule, which tend to eliminate the individual dissatisfaction by reallocating the player into a suitable place. The “degree of satisfaction” may be player-dependent. Some players may feel themselves unsatisfied by having a single different neighbor, others are more “tolerant” to “race inhomogeneity”. Obviously, it would be easy to explain segregation, if the individual players would be inherently racists. What the simulation results suggested is that slight preference to live “among their owns” imply global segregation, and the formation of ghettos.

In an extensive analysis Pancs and Vriend studied both 1D and 2D model [398], with a supplement [399], and Schelling’s basics results proved to be robust. Four types of utility functions were defined. (This is what theoretical social models can do: to define classes of idealized behavior types and analyze the different outcomes due to different structural conditions). See Fig. 4.14.

The functions define the degree of satisfaction in the function of unlike neighbors. All the figures have a special value at 50% (“perfect integration”). There is no other percentage what the player better prefer than to live in a perfect integration, at least in some sense. The flat utility reflects Schelling’s original perspective. This player is indifferent for having unlike neighbors (or not to have any) while they does not exceeded the 50% tolerance threshold.
For this player, complete segregation is as acceptable as perfect integration, and unacceptable to live in any minority.

The ‘p50’ utility function has a clear preference for perfect integration, and the larger the better for having unlike neighbors at the 50% point, where there is a cut-off. The nature of this function reflects that this player also has aversion being in minority.

The ‘p100’ function is symmetric with maximum at 50%. Her preference is to live in perfect integration, and any deviation from it is symmetrically (un)acceptable.

The spiked utility function defined the case when any deviation from the perfect integration is highly non-preferred.

A simulation result with 4000 players of each type on a $100 \times 100$ board is presented in Fig. 4.15.

Fig. 4.15. Simulation results from [398]. The left board is the control set randomly. The middle board illustrates that if the utility functions is set to be flat, the separation is close to be complete. The “p50” utility function implies the formation of patches, stripes. Occasionally embedded islands occur.

The flat utility function implied almost full separation, while ‘p50’ implied the appearance of patches, islands. Occasionally islands in the islands also emerge.

Schelling’s segregation became popular, since it attacked a socially and intellectually challenging problem by simple conceptual tools, produced easily interpretable, really interesting, somewhat paradoxical results. The emerging macroscopic structure is not designed, it is not in accordance (but occasionally also not against) to the intention of the players. The principle of “Think globally, act locally!” is violated.
4.7.2 Opinion Dynamics

Interactions of people in a group (in extreme case this may be the whole society) imply changes in their opinion about different issues and may lead to consensus, fragmentation or polarization. Consensus means that all players share the same opinion, fragmentation occurs when several opinions emerge, polarization is a special case of fragmentation for having two parties of people with two different opinions.

Mathematical models help to understand the dynamics of the interactions among players. A family of models were set and investigated by Hegselmann and Krause [234]. There are \( n \) number of players. The opinion of a player at a certain time point \( t \) is a real number, \( x_i(t) \) and the complete state of the model is described by a continuous opinion vector. There is a simpler situation, if there are two possible opinions the state space is binary. The \( n \)-dimensional \( \mathbf{x}(t) = (x_1(t), \ldots, x_n(t)) \) giving the state of the system is called the opinion profile.

The change of the state, i.e., the opinion dynamics, is governed by the local interactions among people through their opinions, and the \( \mathbf{A} \) matrix specifies this influence; an element \( a_{ij} \) denotes the influence of \( i \) on \( j \). For simplicity \( \sum_{j=1}^{n} a_{ij} = 1 \). In the following we assume that \( a_{ij} \geq 0 \), this expresses that a player is never influenced negatively by another player (but of course the influence may be 0). The time may be discrete. The simplest dynamic model (the “classical model”) is based on the assumption that player \( i \) adjusts her opinion by taking a weighted average of all other \( j \) players’ opinion, and the influence matrix is constant:

\[
\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t). \tag{4.59}
\]

Assuming that the weights change with time or opinion, a quasi-linear (basically nonlinear) model can be defined:

\[
\mathbf{x}(t+1) = \mathbf{A}(t, x(t))\mathbf{x}(t). \tag{4.60}
\]

The main questions to be answered are related to the dynamic change of opinions starting from some initial opinion profile \( x(0) \). A specific problem is to find out the conditions, which imply that the group of players approach a consensus \( c \), i.e., \( \lim_{x \to \infty} x_i(t) = c \), for all players \( i \). (Structural conditions leading to fragmentation and polarization are also analyzed).

There is another variation of the classical model (called the Friedkin and Johnsen model). While the assumption behind the classical model is that the only driving force of the dynamics is the influence, the Friedkin and Johnsen model takes into account the nonfidelity of the player to an original idea. The
degree of fidelity is denoted by $g_i$, and the $1 - g_i$ is the measure of a player’s susceptibility for social influence. The model is defined as

\[ x(t + 1) = Gx(0) + (I - G)Ax(t). \]  

(4.61)

Here $G$ is a diagonal matrix with elements $g_i$, $0 \leq g_i \leq 1$ in the diagonal and $I$ is the identity matrix.

There are some analytical results for certain models of opinion dynamics. These results connect structural conditions of the models (such as the structure of the influence matrix), and the qualitative dynamic behavior of the system.

1. There are sufficient and necessary conditions for the model (4.59). If for any two players $i, j$ there is a third one $k$, with positive weights $a_{ik} > 0$ and $a_{jk} > 0$ then the consensus property holds for all initial value. (Of course, the value of the consensus depends on the initial value.)

2. For the Friedkin-Johnsen model there is a sufficient condition to converge a fixed vector of fragmented opinions. The existence of at least one positive degree $g_i$ under some technical condition implies that any initial profile will converge to a stationary opinion pattern. This pattern has the consensus property if and only if there were preexisting consensus among all players with positive degree.

Hegselmann and Krause [234] extended these models by introducing the concept of bounded confidence. This concept expresses the assumption that players with too different opinion don’t influence each other. The opinions of two players should be closer than a given $\epsilon$ confidence bound. Only the opinion of those other players will modify the opinion of player $i$, who are within this confidence interval.

The model is a special case of the nonlinear model (4.60), (for time-independent influence matrix). One (maybe too simplifying) assumption is that the confidence level is the same for all players: $\epsilon_i = \epsilon$ for all $i$. Obviously, $\epsilon$ is a parameter of the model, and as always, there is the question: “How robust are the results of changing the parameters?”

The bounded confidence (BC) model is defined as:

\[ x_i(t + 1) = |I(i, x(t))|^{-1} \sum_{j \in I(i, x(t))} x_j(t), \]  

(4.62)

where $|I|$ is the cardinality of the set $I$, and the set of players whose opinions are taken into account is: $I(i, x) = \{1 \leq j \leq n; |x_i - x_j| \leq \epsilon_i\}$, and $\epsilon_i > 0$ is the confidence level of player $i$. 
Analytical results were obtained for specific cases. \( \epsilon_i \) may be uniform for all players, so \( \epsilon_i = \epsilon \) for all \( i \). A confidence interval maybe symmetric or asymmetric. Symmetric confidence interval \( [-\epsilon, +\epsilon] \) here means that a player equivalently tolerant for a “less” and a “more” opinion. Asymmetric confidence interval is denoted by \( [-\epsilon_l, +\epsilon_r] \), where the indices \( l \) and \( r \) denotes “left” and “right”. \( \epsilon_l \) and \( \epsilon_r \) are positive, but they may be different.

An asymmetric confidence interval may imply one sided split. The term means that while one player \( (i, \text{if } \epsilon_l < \epsilon_r) \) takes into account the opinion of \( j \), but not vice verse. In an opinion profile \( x \) there is a split between players \( i \) and \( j \), if \( |x_i - x_j| > \epsilon \). The dynamical model has some characteristic properties:

1. The dynamics is “order preserving”, i.e., if \( x_i(t) \leq x_j(t) \) for all \( i \leq j \) then \( x_i(t+1) \leq x_j(t+1) \) for all \( i, j \).

2. The dynamics is split preserving. If there is a split between two players, this property is persistent.

The general take home message is that these models help to get an insight about structural conditions of different qualitative outcomes, such as consensus, polarization and fragmentation. Of course, the assumptions of the models may vary, other effects (say repulsion) could be included. Propagation of extreme views are not very different for social epidemics. Specifically, a model resembling for epidemic spread was introduced to analyze the interaction among four subpopulations, such as general, susceptible, excited, and fanatic ones [483]. In a certain parameter regime the excited and fanatic sub-populations survive, in another one die out. The model, (as usually model do) give hint about the conditions of theses different qualitative outcomes.

4.8 Nonlinear Dynamics in Economics: Some Examples

4.8.1 Business Cycles

Kaldor Model

The Goodwin model, an application and extension of the Lotka-Volterra model was briefly mentioned in Sect. 3.5.1. Nicolas Kaldor (1908–1986) gave a mechanism for the generation of temporal oscillatory dynamics in income \( (y) \) and capital \( (k) \). Kaldor assumed nonlinear dependence of investment \( (I) \) and saving \( (S) \) on income.
Assuming time-independent investment and saving the dynamics is given as:

\[
\dot{y}(t) = a[I(k(t), y(t)) - S(k(t), y(t))],
\]
\[
\dot{k}(t) = I(k(t), y(t)) - \delta k(t).
\]

Equation 4.63a describes income dynamics. If the rate parameter \( a > 0 \), the direction of change of the income depends on the sign of the difference between investment and saving. Equation 4.63b specifies that the increase in the capital is equal to the investment, and is reduced by depreciation controlled by the parameter \( \delta \). Chang and Smyth [95] proved (actually for the case of \( \delta = 0 \)) the conditions of the existence of limit cycle. Kaldor made assumptions for the form of the functions \( I \) and \( S \). Investment, savings, and their difference, \( dy \), as functions of income, \( y \), for capital at \( k = k_e \), with \( a = 4 \) was shown in the left side of Fig. 4.16. With some other restrictions, the model may lead to limit cycle, as the right-hand side of Fig. 4.16 shows.

**The Kaldor–Kalecki Model**

In a version of the model another mechanism to induce periodicity, i.e., some delay in investment, is also incorporated. The modified model is a time-delay differential equation:

\[
\dot{y}(t) = a[I(k(t), y(t)) - S(k(t), y(t))],
\]
\[
\dot{k}(t) = I(k(t), y(t - \tau)) - \delta k(t)
\]

where \( \tau \) is the time delay. It was shown [299] that \( \tau \) is bifurcation parameter, and there is a Hopf bifurcation mechanism for the onset of limit cycle behavior.
4.8.2 Controlling Chaos in Economic Models

Controlling of some economic process might be very important. While traditional economics used linear models, nonlinear dynamics helps to understand the mechanism of chaos generation and control in micro- and macroeconomics, too. The use of a chaos control technique (discussed in Sect. 3.6.4) in a microeconometric model was demonstrated [240] Chaos occurs in a simple microeconometric model of two competing firms.

Two firms X and Y competing on the same market of goods. The firms perform active investment strategies, i.e., their temporary investments depend on their relative position on the market. The strategies can be asymmetric:

- Firm X invests more when it has an advantage over the firm Y;
- Firm Y invests more if it is in a disadvantageous position compared to the firm X.

A discrete-time dynamical model was defined for the temporal change of the sales $x_n, y_n$ of the two firms as:

$$
x_{n+1} = F^x(x_n, y_n) = (1 - \alpha)x_n + \frac{a}{1 + e^{xexp[-c(x_n - y_n)]}}, \quad (4.65a)$$

$$
y_{n+1} = F^y(x_n, y_n) = (1 - \beta)y_n + \frac{b}{1 + e^{xexp[-c(x_n - y_n)]}}. \quad (4.65b)$$

The parameters $0 < \alpha < 1$ and $0 < \beta < 1$ control the velocity of sales reduction without investment, while $a$ and $b$ denote the investment effectiveness. This is the Behrens–Feichtinger model, which also well describes the control of the drug market Sect. 4.9. The analysis showed that there is a parameter region, where chaotic behavior occurs.

Chaos suppression might be advantageous, since the unpredictable dynamics will be at least predictable. Different control strategies could be applied:

1. Change in the forcing function. A time-delay feedback is given to the right-hands site of the governing equation: To ensure periodic (say, period-1, or period-2 behavior), the equation for sales dynamics can be written as

$$
y_{n+1} = F^y(x_n, y_n) + F_K, \quad (4.66)$$

where the form of the $F_K$ control force is by an additive term.

$$
F_K = K_{yy}(y_n - y_{n-m}), \quad (4.67)$$
where \( m = 1 \) or \( m = 2 \) for period-1 or period-2 orbits, and \( K_{yy} \) is a control constant.

2. Change in the parameters. A time-delay feedback term is added to a system parameter. It is installed by

\[
y_{n+1} = (1 - \beta)y_n + \frac{b + K_{yy}(y_n - y_{n-m})}{1 + e^{x(-c(x_n - y_n))}}.
\]

Time-delay feedback is a plausible control mechanism in this context. The policy of firms depend on the difference between actual and past values of sales. There are realistic mechanisms (strategies adopted) to suppress chaos.

### 4.9 Drug Market: Controlling Chaos

A set of models were established to study the dynamics of the number of addicts \( A(t) \) and sellers \( D(t) \) in a city’s illegal drug market [53].

The assumptions of the model describe the increase and decrease of the size of the two populations. The decrease term is described by a first-order decay term. Increase of both population size depend on their relative numbers. There is a positive feedback effect, since relative large number of users will increase the number of sellers (since the market offers good profit opportunities). The increase in the number of new users depends, however, mostly on the number of the actual users, since their habit propagates through their social network. The number of sellers influence the number of new users only indirectly. The structure of the model is shown in Fig. 4.17.

![Fig. 4.17. Flow diagram of the drug market model. Adapted from [53].](image)
The discrete-time dynamical model equations are specified, as

\[
\begin{align*}
A_{t+1} &= (1 - \alpha)A_t + af(A_t, D_t), \quad A_0 = A(0), \quad (4.69a) \\
D_{t+1} &= (1 - \beta)D_t + bf(A_t, D_t), \quad D_0 = D(0), \quad (4.69b)
\end{align*}
\]

\(a, b > 0, \) and \(\alpha, \beta \in [0, 1].\) The positive and negative feedback effects should be expressed in the form of the function \(f.\) A rather general form for the function is

\[
f(A, D) = \frac{1}{1 + e^{-c(A-D)}}, \quad c > 0. \quad (4.70)
\]

Statistical data suggest that the outflow rate of the dealers (due to death, law enforcement, policy) is larger than the outflow rate of the addicts, i.e., \(\beta > \alpha.\) It was shown that for a reasonable region of parameters, chaos may occur, as the bifurcation diagram Fig. 4.18 shows.

Since a chaotic drug market is interpreted as unpredictable, there is a natural question whether it is possible to control the system to stable cycles or even (low level) equilibrium.

![Bifurcation diagram of the model of the drug market. Chaos may occur in a certain parameter region. Adapted from [53].](image)