Chapter 5
The Kinetics of Biological Systems

This section provides an introduction to the basic ideas of systems theory as applied to living organisms which is an important part of theoretical biophysics. Originally, thermodynamics, as a universal theory of energetic basis of all movements in nature, and kinetics, as a theory of time courses of predicted processes, were considered as two separate theoretical approaches. Thermodynamics answers the questions: what are the reasons for, the driving force of, and the direction of a movement, and what, finally, is the equilibrium situation which will be arrived at if the energetic potentials are equilibrated? Kinetics, on the other hand, just studies the course of a given reaction, its time constants and rates, possible stationary states, stabilities, and oscillations.

Nonequilibrium thermodynamics, evaluating the course of irreversible processes, introduced time as a new variable into the framework of thermodynamics. This was the point when kinetics and thermodynamics became more and more linked. *Synergetics* was the new name proposed to characterize this connection (see Haken 2010).

In recent decades, systems theory experienced an enormous upturn. This was caused by the development of powerful computer systems which are tightly connected to new analytical techniques delivering an enormously increased amount of data and structural insights at all levels of biological organization. This development not only requires methods of storage and managing of this data avalanche but also demands the finding of functional connections, leading to a deeper understanding of complex biological functions. This kind of analysis offers the possibility to analyze existing hypotheses, to check them quantitatively, and to subsequently formulate new concepts for experimental research.

In the framework of this book, and according to its intension, we will impart just a basic understanding of some approaches and ideas of systems theory. In the first part, we will introduce some general approaches to describe the complex network of biological interactions, followed by short introductions to particular applications.

A number of profound textbooks and monographs exist that allow us to go deeper into these theories.

R. Glaser, *Biophysics*,
DOI 10.1007/978-3-642-25212-9_5,

5.1 Some General Aspects of Systems Theory

The basic idea of systems analysis is *isomorphism*. This means that quite different processes in nature, technology, or even sociology, can be formally described by similar mathematical approaches. They show similarities in their properties despite profound differences in the particular details of their elements, or their mechanisms of interaction. For biological systems, this means that there are general principles in the understanding of metabolic and epigenetic, as well as neuronal, networks, and even in systems of ecological interactions.

The general problem in calculating the kinetics of biological systems is their enormous complexity, embedded in a hierarchical structure. For example, the metabolic network of the cell, as complex as it is, is also governed by tissue and organ regulation; this again is controlled by the neuro-humoral system of the organism, and so on. In practice, to answer a particular question, limitations are always necessary. This means defining the system level and limiting the conditions as much as possible.

We will follow here the definition of the term "system" as it has already been introduced in Sect. 2.1.3, and use the depiction of the system properties as explained in Sect. 3.1.1. The definitions of stationary states, and the conditions of their stability (Fig. 3.6), which have been discussed from the thermodynamic point of view in Sect. 3.1.4, are also important here.

Some Textbooks in Biological Systems Theory Alberghina and Westerhoff 2005; Alon 2006; Haken 2010; Heinrich and Schuster 1996; Klipp et al. 2009; Koch et al. 2011.

5.1.1 Basic Equations of Kinetic Processes

Usually, a system is characterized by three types of quantities, which are conventionally named constants, parameters, and variables. *Constants* are quantities which are valid for each natural system, like the Faraday constant (F), the molar gas constant (R), or Avogadro's number (A). In contrast, *parameters* are the constants just for the particular case. These, for example, can be the rate constants of the observed reactions (k_n), equilibrium constants (K_n), or the Michaelis constants of enzymatic processes. Unfortunately, these quantities also contain the term "constant" in their name, which indicates that this terminology is not obligatory. *Variables* are time-, and probably even space–dependent, quantities which describe the actual state of the system. In thermodynamics (Sect. 3.1), there is not a clear differentiation between the terms "parameter" and "variable" because this can change, for example when considering the same system under isochoric ($V =$ const), isothermal ($T =$ const), or isobaric ($p =$ const) conditions.

In addition, the term *compartment* must be introduced. In some papers, this expression is simply used for a space which is occupied by particular reactants.

For formal kinetic analysis, however, this term must be generalized. Thus, the usual kinetic equations can formally represent the flux rate of a transport process, in the same way as the rate of a chemical reaction, or even some kinds of proliferation rates of cells or organisms. Therefore, a term is needed which can be formally applied to a phase with spatial borders, as well as to a component of a chemical reaction, or even to a population of cells or organisms.

These requirements are fulfilled by the following definition: *A compartment is an element of a system that, with respect to the matter under investigation, has limits which are defined by a specific process, and that contains no gradients within itself.* The term *matter* is used in a generalized form. It can represent a substance in a reaction system, a cell in a proliferation system, an organism in an ecosystem, etc. *Specific process* means the type of the considered movement. It can be an unidirectional flux, as well as the biochemical reaction rate, the proliferation rate of cells, and the birth or death rate of a species. The absence of gradients inside the compartment means that it is to be considered as homogeneous, or *well mixed.* Otherwise, the compartment must be subdivided into smaller compartments, and a multi-compartment system appears. A compartment model is always defined for a special question and is only valid *with respect to the matter under investigation.* In general, it corresponds to the nodes in a complex graph.

The kinetics of a system, i.e. its time-dependent state can be described by the time derivatives of its variables. In this way, differential equations are obtained. If these equations contain only time derivatives, they are called *ordinary differential equations.* If they additionally include derivatives of the variables by space, i.e. by x-, y-, and z-directions, then *partial differential equations* occur.

The simplest reaction is a decay of a substance in a reaction of the first order, the reaction rate of which is proportional to the concentration. The minus sign indicates that the concentration decreases with time.

$$\frac{\mathrm{d}c}{\mathrm{d}t} = -kc \tag{5.1}$$

transforming this equation into:

$$\frac{\mathrm{d}c}{c} = -k\mathrm{d}t \tag{5.2}$$

and integrating both sides:

$$\int_0^t \frac{1}{c}\mathrm{d}c = \int_0^t -k\mathrm{d}t \tag{5.3}$$

it gives:

$$\ln c = -kt + C \tag{5.4}$$

where C is an integration constant. This equation can be transformed into:

$$c = Ce^{-kt} \tag{5.5}$$

It is possible to transform this *general* solution of Eq. 5.1 into a *particular* one, by defining an initial condition, as, for example: at $t = 0$, $C = c_0$:

$$c = c_0 e^{-kt} \tag{5.6}$$

This equation resembles the kinetics of the radioactive decay, or a simple efflux from a compartment. The time at which the concentration is diminished to 50%, i.e. where $c/c_0 = 0.5$, is the *half-life time* ($t_{1/2}$):

$$t_{1/2} = -\frac{\ln 0.5}{k} = \frac{0.693}{k} \tag{5.7}$$

To calculate more complex reactions, the definition of the reaction rate, or velocity (**v**), is useful. For this simple case, it equals just: $\mathbf{v} = kc$, according to Eq. 5.1. For reactions with more components, the degree of advancement ($\mathrm{d}\xi$) is a useful parameter (see Sect. 3.1.6, Eq. 3.73) which is a measure of the progress of the reaction ($\mathrm{d}\xi = 1/\nu_i \cdot \mathrm{d}n_i$) step by step. Containing the stoichiometric coefficient (ν_i) of the particular reactants, it assigns a step of the whole reaction. This allows the defining of the reaction rate of complex reactions as:

$$\mathrm{v} = \frac{d\xi}{dt} \tag{5.8}$$

In kinetic equations, the stoichiometric coefficients become signed, depending on the direction considered in the model. In the case of a reaction:

$$\mathrm{A} + 2\mathrm{B} \rightarrow 3\mathrm{C}$$

the differential coefficients of the three components can be formulated as:

$$\frac{\mathrm{d}c_A}{\mathrm{d}t} = -\mathrm{v}\ ; \quad \frac{\mathrm{d}c_B}{\mathrm{d}t} = -2\mathrm{v}\ ; \quad \frac{\mathrm{d}c_C}{\mathrm{d}t} = +3\mathrm{v} \tag{5.9}$$

In this case, the reaction rate (**v**) is defined for the reaction, proceeding from left to right. Therefore, the components A and B are diminished if the reaction proceeds in this direction, whereas the component C will be enriched. In biochemical networks with a number of interconnected reactions, the set of stoichiometric coefficients can be listed in matrices, which allow generalized considerations.

The generalized kinetic equation of such a set of reactions is:

$$\frac{\mathrm{d}c_i}{\mathrm{d}t} = \sum_{j=1}^{r} \nu_{ij}\, \mathrm{v}_j \qquad for:\ i = 1, \ldots, m \tag{5.10}$$

This indicates that metabolic networks correspond to systems of many differential equations. Their combination results in differential equations of higher orders, i. e. such as with the higher order of derivatives. Furthermore, reactions between several components, i.e. reactions of a higher order in the terminology of chemistry (unfortunately not identical with the mathematical terminology of the order of differential equations!) mostly lead to nonlinear differential equations, containing the product of concentrations.

To understand some basic properties of the steady state, let us consider a simple system, like that shown in Fig. 3.4b (Sect. 3.1.4), where water with a constant flux J is pumped into a vessel, flowing back depending on its filling level. Let the volume of the vessels correspond to the concentration of a component in a chemical reaction. We will formally characterize it by the symbol c:

$$\frac{dc}{dt} = J - kc \tag{5.11}$$

where J indicates the constant influx, and $-kc$ the volume-dependent outflux of the system (see Fig. 5.1). The integration of this equation leads to:

$$c = \frac{J}{k}(1 - e^{-kt}) \tag{5.12}$$

This is already a particular solution of the Eq. 5.11 for the initial condition: at $t = 0$, $c = 0$, i.e. the case where the vessel at the beginning of the experiment was empty. This is represented by the red line in Fig. 5.1. A general solution just leads to an array of curves, some of which are depicted in Fig. 5.1 in blue.

Figure 5.1 indicates that, independent of the initial condition, all curves finally end at the same point (c_∞), thus illustrating a general property of all linear systems, namely the occurrence of only one single stationary state, which depends not on the initial conditions but only on the system parameters. This property is called

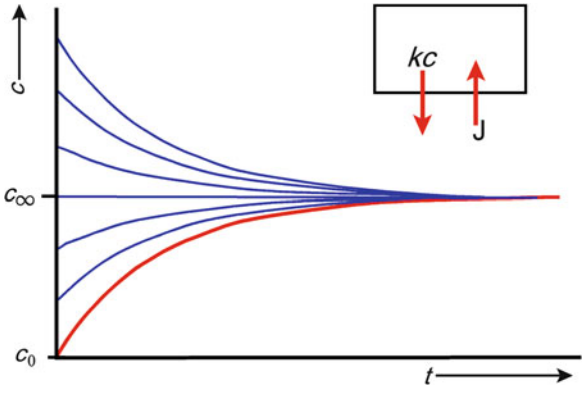

Fig. 5.1 An example demonstrating the property of equifinality. It indicates the kinetics of a single compartment system with a constant influx (J) and a concentration-dependent efflux (kc). Some curves show general solutions of the differential equation (Eq. 5.11 (*blue lines*)). The red line corresponds to the particular solution (Eq. 5.12), starting with the concentration c = 0

equifinality, which means: all paths, independently of the starting point, lead to the same end.

Extrapolating Eq. 5.12 to $t \to \infty$, this steady state of equifinality can be obtained as:

$$c_\infty = \frac{J}{k} \quad \text{for} \quad t \to \infty \tag{5.13}$$

The same result can be obtained using the original differential equation (Eq. 5.11) and considering the steady state, which by definition means: $\mathrm{d}c/\mathrm{d}t = 0$. In this way, even larger sets of differential equations can be transformed into a system of algebraic equations. This makes it possible to calculate the variables $c_{n\infty}$ for the case of the stationary state without further conditions. Therefore, these equilibrium variables only depend on the rate constants (k_{nj}). If all equations are linear, typically only one single solution for each of these values will be obtained.

The property of equifinality occurs only in linear systems. It is easy to understand that solutions of nonlinear algebraic equations obviously provide more than one solution. Some of them may be real, others could just represent mathematical solutions without physical meaning, as, for example, negative concentrations. This multitude of solutions in the case of nonlinear systems corresponds to a system with several stationary states. We will come back to this problem in the next section.

5.1.2 General Features of System Behavior

In general, the behavior of a system can already be predicted by considering the order and the degree of the corresponding differential equations. The *order* of a differential equation corresponds to the order of its highest derivative. The *degree* of an equation is determined by the highest power to which one of its variables, or their derivatives, is raised, provided that the equation has been arranged in such a way that only positive powers of the dependent variables, or their derivatives, occur. The product of two variables has to be considered as a kind of second power. The thermodynamic characteristics of nonlinear processes have already been discussed in Sect. 3.1.4. We will complete these considerations by some kinetic circumstances.

To illustrate this, let us consider a simple mechanical system. Suppose there is a spring hanging from a rigid support (Fig. 5.2). At the lower end of the spring is a pointer which marks the length (l) of the spring. A certain force (F) holds the spring under tension and is in constant equilibrium with the pulling force of the spring according to Young's law ($k_1 l$) (see: Sect. 3.6.3, Eq. 3.226). If the force is suddenly changed (blue line in Fig. 5.2), the spring elongates (red line). Because of a brake which is included in the construction, a fictional force (F_{fr}) occurs, which is taken to be proportional to the rate of displacement ($\mathrm{d}l/\mathrm{d}t$) by the factor k_2. In this case, the action of the force (F) is compensated by the fictional force (F_{fr}) as well as by the

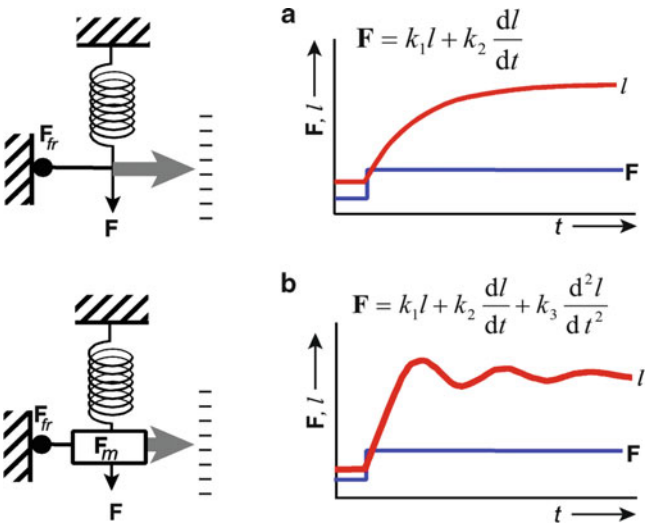

Fig. 5.2 Mechanical constructions illustrating the kinetic behavior of systems of first (**a**), and second (**b**) order as response (*red lines*) to a rectangular function of the extending force F (*blue lines*)

extension force of the spring. Thus, this system is described by a linear differential equation of the first order (Fig. 5.2a). Its solution and integration leads to a function $l(t)$ as a simple exponential expression. It shows a simple continuous approach to the equilibrium, corresponding to the empirically expected behavior.

Completing this construction by an additional mass (m), which imparts a certain amount of inertia to the system, a further summand must be included in the function (Fig. 5.2b). An inertial force (F_m) occurs, which is proportional (k_3) to the acceleration, i.e. to the second time derivative of the length (l), according to Newton's law. In this way, a differential equation of the second order appears. It is easy to imagine that this second order system will show oscillations of the function $l(t)$ as it settles down, similar to pairs of scales. These oscillations can be made to slow down at different rates by the use of various damping devices, i.e. changing the relationship between the parameters k_2 and k_3.

A further complication of the behavior of this system would occur at increasing force (F). In this case, the friction becomes a nonlinear function of the elongation velocity (dl/dt) (see also Sect. 3.1.3, Fig. 3.3).

Non-linear systems can exhibit some further kinds of behavior. Figure 5.3 shows in principle how a given variable c in a system can change when it is displaced by an amount Δc from its stationary value. When the system is stable, it returns to its original state. This can occur without oscillations (black line) at a high degree of damping, i.e. a low time constant of reaction. But increasing the time constant above an aperiodic limiting condition causes the oscillations to die away asymptotically (red line). With periodic limiting conditions, non-damped oscillations are

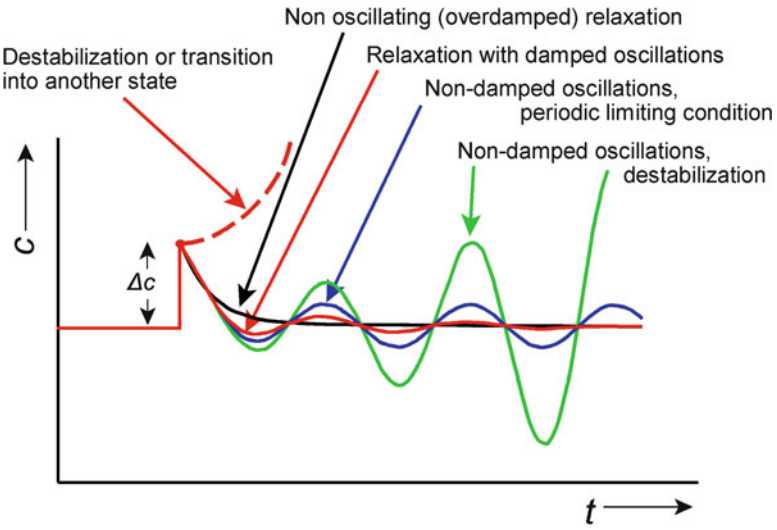

Fig. 5.3 Possible kinds of system behavior, illustrated as a time course of the variable $c(t)$ of a system after a particular displacement Δc

obtained (blue line). Disturbances of the system may also lead to instabilities, possibly leading to another stationary state (dashed red line, and green line).

In the case of more time-dependent variables, two methods are possible to illustrate the system graphically. If, for example, a system contains two variables, c_1 and c_2, it is possible to characterize it by plotting both curves, $c_1(t)$, and $c_2(t)$, separately (Fig. 5.4, left part). At any moment (t), a value of the variable c_1 corresponds to a particular value of c_2. The state of the system, therefore, can also be represented by a graph, where c_2 is plotted against c_1. At the time t_B, for example, the state of the system corresponds to the point B on the righthand graph. Analyzing the functions $c_1(t)$ and $c_2(t)$ on the left plot at the moments t_A, t_B, t_C, t_D, t_E . . ., it is possible to construct a *trajectory* following the points A, B, C, D, E. . . in the graph on the righthand side of Fig. 5.4. This represents a particular change in the system. When there are only two variables, the trajectory is just a line in the plane. The arrow of the trajectory (green line) indicates its time course, showing where the system proceeds.

In systems with n variables, an n-dimensional space is required to depict the behavior of the whole system. In fact, it is a vector space, which means that at each point in it, a signpost is located, indicating the direction of the further direction of system development.

Let us firstly characterize the general kind of system behavior in the case of two variables (n_1, n_2), i.e. in a two-dimensional plane (Fig. 5.5). Later, we will demonstrate this for two particular cases: for a biochemical reaction model (Sect. 5.2.1, Fig. 5.12), and for the depiction of population dynamics (Sect. 5.2.3, Fig. 5.16). Here, we will just discuss some principles of system behavior in a more general way.

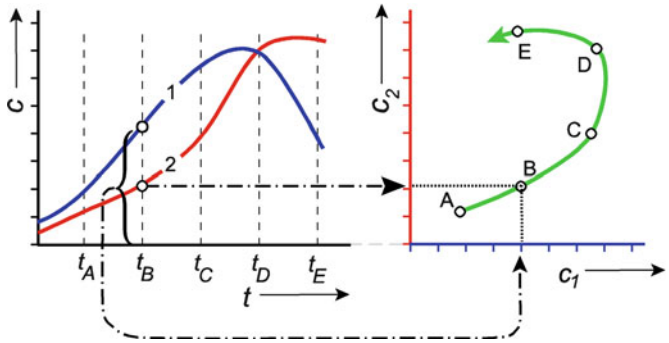

Fig. 5.4 Plot of the time dependent functions of two variables $c_1(t)$ (*blue*), and $c_2(t)$ (*red*) on the left, and the depiction of the trajectory (*green*) of these variables $c_2(c_1)$ on the right (the *dash-dotted arrow lines* illustrate the method of construction of the trajectory)

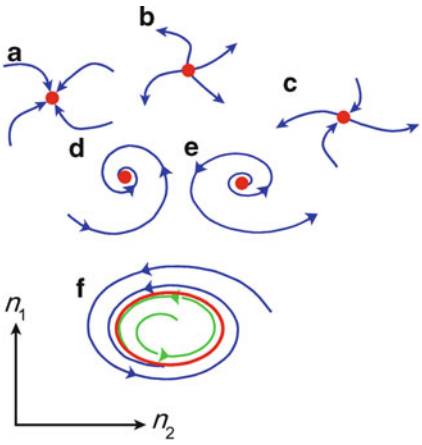

Fig. 5.5 Trajectories in a two-dimensional space to characterize various states of the system: (**a**) stable node, (**b**) unstable node, (**c**) saddle point, (**d**) stable focus, (**e**) unstable focus, (**f**) stable limit cycle

The curves in Fig. 5.5 were obtained in the same way as the trajectory in the righthand part of Fig. 5.4. They are the kinetic completion of the energetic schemes in Figs. 3.5 and 3.6, and represent numerical solutions of differential equations, characterizing the kinetic behavior of complete systems.

At first, these presentations show whether the systems indicate a kind of *stability* as an answer to a small perturbation. The curves A, D, and F show reactions of stable systems. The arrows demonstrate a return of the systems to their origin after the perturbation. In contrast to this, the curves B, C, E show that even a small deviation from the starting point produces a further shift, thus a destabilization of the system (see also Fig. 3.5 in Sect. 3.1.4). The curve F represents a system with stable permanent oscillations. In this case, the stability arrives at a *limit circle* which is a sign of undamped oscillations. Metastable states (see Fig. 3.5), which are not plotted in Fig. 5.5, would be characterized by an arrow, leading from one point to another.

In addition, these figures demonstrate different ways of arriving at the stable state. Case A indicates an aperiodic approximation, as shown in Fig. 5.3. The approach by damped oscillations in Fig. 5.3 is represented in case D of Fig. 5.5, and the undamped oscillation corresponds to case E in Fig. 5.5. The distortion of the circle into an ellipse (case F) simply depends on the chosen coordinates. In the case of oscillations superimposed by several frequencies, this ellipse will be further deformed.

A further important property of system behavior can be demonstrated more easily in a three-dimensional system (Fig. 5.6). Because of causal connections between these three parameters, the only possible properties of the system form a *control surface*. This can be best illustrated by the physicochemical properties of water. Take z to be the viscosity or the structure of water, which depends on pressure (x) and on temperature (y). When the temperature falls (trajectory A in Fig. 5.6), the system will reach a *bifurcation point* at 0°C at normal pressure (point P in the green projection). It can either freeze, i.e. move in the lower surface to the point C, or remain as a super-cooled liquid in the upper surface and move to B. The supercooled liquid, however, will spontaneously freeze (trajectory B–C) if there is any disturbance of the system (alteration of pressure, vibration, etc.).

In 1969, the French mathematician Rene Thom developed a comprehensive topological theory of dynamic systems which primarily enables the occurrence of multi-stationary states to be systematized. He called a sudden change from one state to another a *catastrophe*. The case of Fig. 5.6 was classified as a *cusp catastrophe*. This system is characterized by the conditions that, in most cases, there is only one point in the z-coordinates that corresponds to a point in the x,y-plane. At the folded region of the control surface, however, three possible values of z occur. At some points in the x,y-coordinates, the system can therefore obtain three different states.

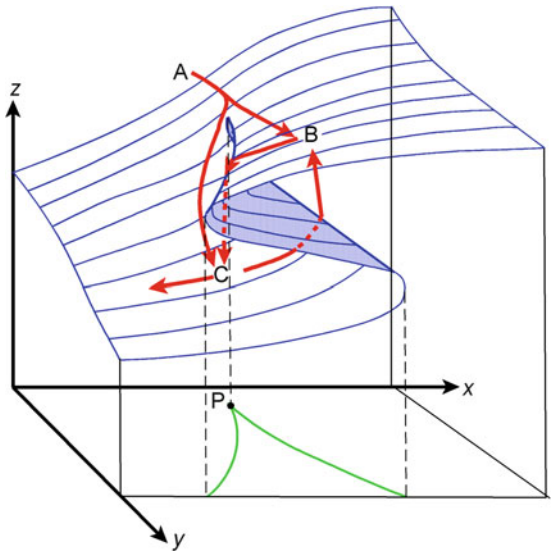

Fig. 5.6 A control surface (*blue area*) of a system with three variables: *x, y, z,* as an example of a cusp catastrophe. (**A**) Trajectory with bifurcation point (P); (**B**, **C**) trajectories with a phase jump ("catastrophe") near the borderline of stability; shaded area: region of unstable states

The points on the middle plane (the shaded area in Fig. 5.6), however, indicate unstable states. This can be compared with the thermodynamic picture of a metastable state (Fig. 3.5), i.e. an unstable state between the two stable stationary states.

With respect to the isomorphism of the systems of different natures, these kinds of cusp catastrophe can frequently be observed in various concrete examples. In some kinds, not only the catastrophe B, C (Fig. 5.6) is possible but also the opposite movement (not observed in the case of water).

In Sect. 5.2.1. we will characterize the behavior of a biochemical reaction chain with such a function (Fig. 5.12b). Furthermore, the electrical potential of a nerve cell was explained in this way. Even examples in the behavioral biology can be found, as, for example, the reaction of a dog, which as a result of irritation can either run away (i.e. the upper surface), or attack (lower surface). The running away in panic can suddenly change to an attack or vice versa.

In recent decades, a type of system behavior was investigated which is called *chaotic*. This means an obviously irregular and non-predictable behavior, which nevertheless is the result of the interaction of deterministic processes. Nonlinear systems with several stationary states can jump from one state to another. Sometimes, oscillations occur, but these oscillations may degenerate into chaotic movements. The turbulent flow, which we discussed in Sect. 3.7, may be taken as an example of a chaotic process. In medicine, the disturbance of cardiac cycle is analyzed by approaches of the chaos theory.

Further Reading

Heinrich and Schuster 1996; Thom 1969; Zeeman 1976.

5.1.3 The Graph-Theory as an Topological Approach to Describe Complex Network Systems

Biological systems are characterized by an enormous number of interactions on each level of organization. A large number of subcellular and neuronal machines are known, such as metabolic circuits and signaling networks containing hundreds to thousands of different types of components with particular kinds of interactions. The description of all of these processes by particular differential equations leads to a system which is not only impossible to solve analytically but which in addition contains an enormous number of unknown constants, so that even numerical solutions will become quite ambiguous.

To overcome this problem, some generalized approaches of topological mathematics are used, in the end based on *graph theory* which had already been introduced in 1736 by the German mathematician Leonhard Euler. He applied this method at that time to an actual problem of his students. They tried to find a walk that would cross each of the seven bridges of the river Pregel in Königsberg exactly once and return to the starting point. Graph theory first looked like a futile bauble, but nowadays it appears as an important instrument to investigate complex networks of interaction.

A *graph* is a mathematical structure consisting of a number of discrete nodes, linked by a network of connections. It differentiates between *arcs* (or *interactions*) and *edges* (or *links*) (see Fig. 5.7). Arcs, graphically represented by arrows, indicate ordered pairs of nodes. In biochemical networks, they demonstrate, for example, direct interactions between proteins, transcription factors, gene regulation effects, or enzyme substrate interactions. Some of these interactions can occur in a reversible way, in the terminology of chemistry (see Sect. 3.1.4), while others are unidirectional. Edges are represented graphically by single lines. They stand for unordered pairs of nodes, for links indicating a certain influence, but without a clear direction.

A graph, the nodes of which are connected by arcs, is called a *directed graph* (or *digraph*). Cell signaling pathways are commonly represented by mixed graphs. In this case, activation or inhibition processes are represented by arcs, whereas physical protein–protein interactions without a clear-cut directionality are classified as edges. In some cases, processes of self-inhibition or even self-activation are possible. Simple graphs like that in Fig. 5.7, without intersecting edges, allow a planar depiction. In the case of complex networks, nonplanar representations are necessary.

The number of interaction arrows of a particular node is defined as its *degree* (or *size*). This is the most fundamental measure of a network. In the case of reversible reactions, it can be separated into *in-degrees* and *out-degrees*. For a directed graph with m edges and n nodes, an average degree is defined by the number of m/n. Furthermore, a *connection density* of a graph can be calculated, which is the proportion of the actual number of edges in relation to the number of possible ones. This is the simplest estimator of the physical cost, namely of the required energy or other resources.

The degrees of all the nodes show a function of *degree distribution*. In random graphs, each pair of nodes has an equal probability of connection. Therefore, all connections are distributed according to a symmetrical Gaussian function. Most graphs describing real-world networks, however, significantly deviate from the

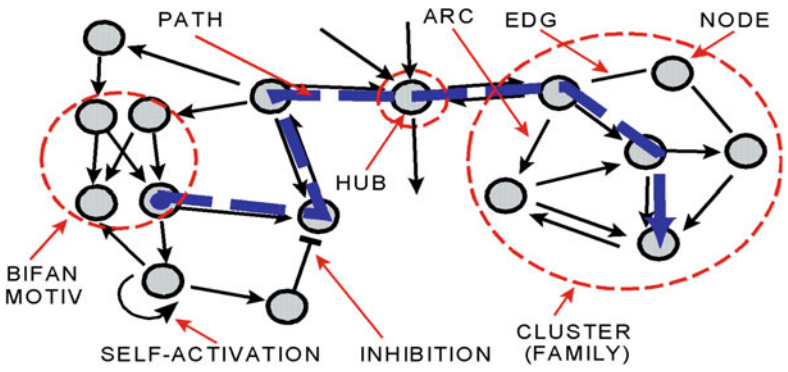

Fig. 5.7 Illustrations of some elements and properties of a graph

simple random-graph model. The analysis of them in some cases shows that the probability function of the number of nodes with a particular degree follows a power function with a negative exponent which often ranges between -2 and -3. This very gradual decay of the degree distribution implies that the network lacks a characteristic scale. These kinds of graphs are therefore called *scale-free*.

In contrast to this, in biological systems, most graphs occur with high levels of local clustering among nodes even forming particular modules as *families* (or *cliques*). This type of network is called a *small-world* graph. Evidence for small-world properties has been found in most genetic, metabolic, and neural networks. These networks therefore deviate from randomness and reflect a kind of specific functionality. Some families representing small circuits in the complex graph show structural similarities, and collectively are termed *network motifs*. In fact, these are subgraphs that are over-represented in the network. Small motives, for example, are particular feedback loops. In gene regulatory networks, *bifan motifs* occur, which are comprised of two upstream regulators both regulating the same two downstream effectors (see Fig. 5.7).

The degree of clustering of a network can be quantified by the *clustering coefficient*, as a proportion of the number of connections that exist between the nearest neighbors of a node in relation to the maximum number of possible connections. The clustering is associated with high local efficiency of information transfer and robustness.

A further parameter in functional graphs is the *path length*. This is defined as the minimum number of edges that must be traversed to go from one node to another. The efficiency of a function is inversely related to its path length. In some cases, it is also possible to estimate the topological distances between nodes of a graph by the functional characteristics. Central nodes, i.e. those with a high number of the shortest paths crossing them, are called *hubs*. Many complex networks can therefore be divided in a number of modules which frequently indicate a hierarchical structure. Usually, each module contains several densely interconnected nodes, whereas there are relatively few connections between nodes of different modules.

A particular limitation of graph theory applications in analyzing biological networks is their dynamics. The nodes and links of biological networks change with time. The abstraction of developing and self-modifying biological networks by graphs with permanent structures sometimes masks temporal aspects of information flow. Modeling the dynamics of these networks to develop more quantitative hypotheses provides a closer insight into the system's behavior.

In summary, the graph analysis of biological networks has been useful to obtain an overview of the organization, and of different types of reactions, and it suggests a number of topological patterns of behavior. Furthermore, some frequently obtained clusters in metabolic systems have already been analyzed in detail, and can be introduced as a more or less clarified module in the network. These approaches allow conclusions on the dynamical behavior of a system, including its stability, noise filtering, modularity, redundancy, and robustness to failure, as well as variations of kinetic rates and concentrations.

Further Reading

Bullmore and Sporns 2009; Lima-Mendez and Van Helden 2009; Ma'ayan 2009.

5.1.4 Systems with Feedback Regulation

In 1948, Norbert Wiener published his book: "Cybernetics – or Control and Communication in the Animal and the Machine." In it, he recommended applying experiences with engineering control systems to understand various processes of physiological and neuronal regulations. Meanwhile, systems theory uses various other possibilities for mathematical modeling of biological processes, but for some considerations, especially in various feedback systems, these cybernetic approaches nevertheless seem to be a useful addition. Generalized feedback systems are mostly applied to understand neuromuscular systems and sensory processes in neurophysiology. Especially in the borderline between physiology and techniques, namely in various branches of medical techniques, these systems engineering tools are common. In some cases, these engineering control system approaches are applied even in processes of cellular regulation.

The basic elements of a feedback control system, and their terminology in the language of engineers, are depicted in Fig. 5.8. In its simplest form, it consists of two subsystems, the *controlled system* and the *controller*. Both of them convert one or more input signals into output signals corresponding to their *input–output law*, which in the simplest case is governed by a system of linear differential equations.

The controller may influence the controlled system in the sense of stimulation, or inhibition, producing in this way a positive, or a negative, feedback. The feedback can work by just triggering, increasing, or inhibiting a process, or it may be part of a sensitive control system. In this case, it adjusts a *controlled variable* $y(t)$ to a given standard which is specified by a *command variable* $z(t)$, in spite of a supplementary effect of a *disturbance variable* $x(t)$. If the command variable remains constant, then the network is called a *retaining control circuit*. However, if it is a function of time, which means that the system has an additional control over it, this is a *follow-up control circuit*. The command variable, the controlled variable, and the disturbance variable may differ from each other in their physical character, both in quality and quantity. The regulation of blood pressure (a controlled variable), for example, on the one hand can be controlled through a sequence of nerve impulses acting as the manipulating variable, while on the other hand, it could also be regulated and controlled by the concentration of a hormone in the blood. Hence, the control signals can be classified as either discrete (sequence of nerve impulses) or analog (hormone concentration).

The difference between the command variable and the controlled variable is called *control deviation* $[\Delta y(t)]$. Depending on the type of the controller, the *manipulated variable* $z(t)$ can be calculated in two ways, corresponding to two types of controllers:

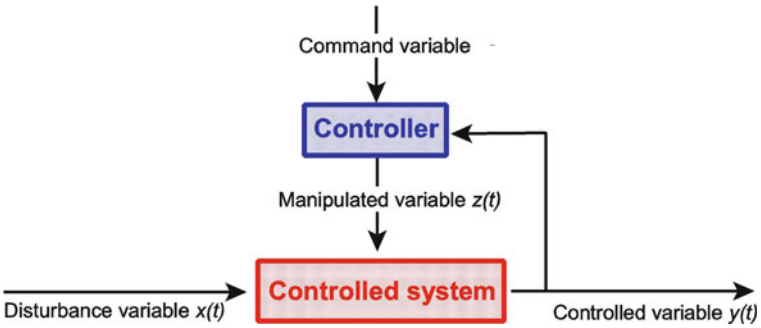

Fig. 5.8 Terminology of the elements of an automatic control system

– In the case of proportional controller (*P-controller*):

$$z(t) = k_p \, \Delta y(t) \tag{5.14}$$

– In the case of integral controller (*I-controller*):

$$z(t) = k_i \int \Delta y(t) \, \mathrm{d}t \tag{5.15}$$

The P-controller, therefore modified the variable [$z(t)$] proportional to the control deviation [$\Delta y(t)$]. This type of controller is not able to completely compensate an induced disturbance of the system. As illustrated in Fig. 5.9, there always remains a proportional deviation (P), because, if $\Delta y = 0$, then according to Eq. 5.14, z = 0.

This problem does not arise with the I-controller. In this case, the modified variable is proportional to the time integral of the control deviation. If the system is brought back to its starting positions, then the parameter z in Eq. 5.15 becomes constant, which means that an optimal compensation for the disturbance variable is achieved.

In biological systems, P- and I-controllers usually exist in combination with each other or are combined with a differential controller (*D-controller*).

For a D-controller, the following equation holds:

$$z(t) = k_D \frac{\mathrm{d}[\Delta y(t)]}{\mathrm{d}t} \tag{5.16}$$

This type of controller does not respond to the control deviation but to the rate at which it is changing.

Neither the controller nor the controlled system respond without inertia, i.e. with a certain time delay. In Eqs. 5.14, 5.15, and 5.16, terms have to be included that take into account the time dependence of the adjustments. This will result in differential

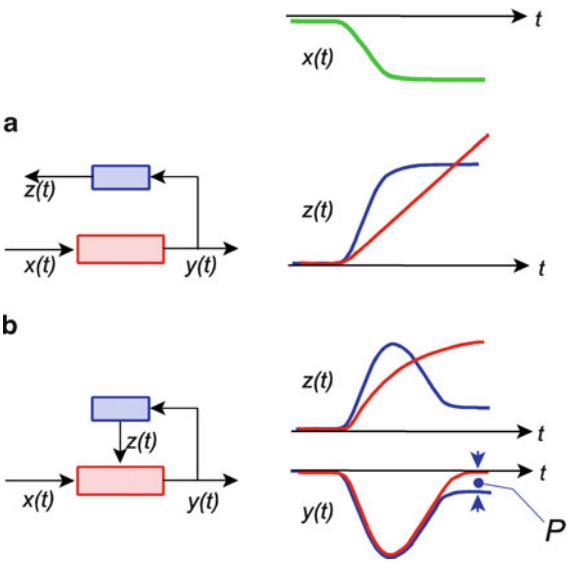

Fig. 5.9 Characteristic responses of an open (**a**), and a closed (**b**) control loop with P-controller (*blue lines*), and I-controller (*red lines*) on a disturbance variable $x(t)$ (*green line*). P proportional deviation

equations of the nth order. The system therefore behaves like those demonstrated as mechanical examples in Fig. 5.2. This means that the curves representing the functions $y(t)$ and $z(t)$ in Fig. 5.9 are in fact more complex than shown. Curves like those illustrated in Fig. 5.3 will more likely result. At least there are no multi-stationary states to be expected with linear systems. The amplification depends on the factors k_P, k_I and k_D in Eqs. 5.14, 5.15, and 5.16, whereas the attenuation correlates with the time constants of the rates at which the systems make necessary adjustments.

The sooner and stronger a system reacts to a disturbance, the faster it will achieve the required correction. However, because control circuits are systems of higher order, the compensation can overshoot. If a system is deflected by a step-wise function, then the following types of system behavior can occur as the amplification is increased (see also Fig. 5.3):

- At a low degree of amplification, up to the aperiodic limiting condition, the control variable approaches the target correction value asymptotically without an overshoot.
- As the amplification increases, exceeding the aperiodic limiting condition, the system starts to oscillate. It reaches the target correction value faster, but overshoots it and is eventually adjusted by damped oscillations. As a result, it may take longer to achieve a stable value than in the first case.
- A further increase in the amplification leads to the periodic limiting condition whereby undamped oscillations around the correct parameter occur.
- If the periodic limiting condition is exceeded because of a further increase in the amplification, then the system becomes unstable. The oscillations get bigger and bigger and cause the system to become increasingly more deflected.

Such instabilities can lead to the destruction which, however, in biological systems is usually limited by the available energy.

These properties illustrate the factors involved in optimizing a control circuit. The fastest possible rate of attaining the target correction value, of course, can to be considered as an optimal condition. This, in fact, cannot be done by simply increasing the amplification. It is best achieved by a combination of different control systems.

Such a kind of complex regulation in the first period of biocybernetic research had already been demonstrated for the case of target-oriented movement of human hand: A person is asked at a given command to immediately move a pointer from position A to position B, whereas the resulting movement is recorded as a function of time (blue line in Fig. 5.10). After a lag period of about 145 ms, the hand starts to move, and reaches the target 500 ms after the signal, without oscillation. This behavior cannot be achieved by a simple I-controller. The aperiodic limiting condition, i.e. the fastest possible reaction of this controller without oscillations, is also shown in the diagram (red line). In fact, the target-oriented movement of the hand is controlled by two processes. First, the distance the hand has to travel is estimated through the eyes, and a rapid movement with a built-in limit is programmed. This takes up the observed lag period of 145 ms. Only during the last phase of the target-oriented movement, do the eyes correct the hand, a process that can be regarded as being performed by an I-controller with a time constant of 47 ms. This model is not only mathematically consistent with the observations but can also be verified by temporarily preventing the subject from seeing what they are doing during the experiment.

Cybernetic approaches to understand physiological reactions, as demonstrated in Fig. 5.10, were chiefly used in the decades after the publication of Norbert Wiener's book. Nowadays, these engineering approaches are directed away from basic biological mechanisms and toward prosthetic applications, and the optimal use of

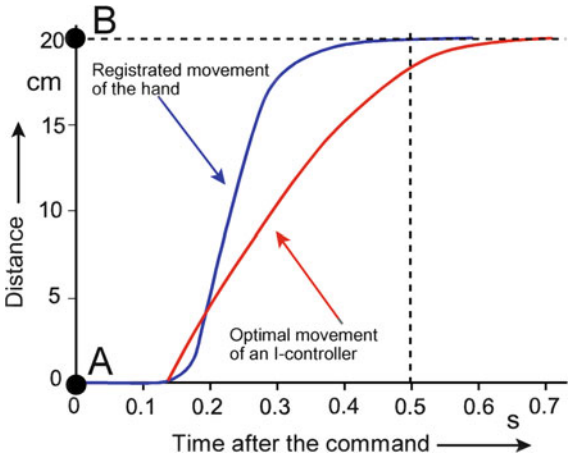

Fig. 5.10 The course of moving the hand (*blue line*) after an order, given at time $t = 0$, to move to a point B with a distance of 20 cm from the starting-point A. The *red line* indicates an optimal reaction of an I-controller with a lag period of 145 ms and an amplification up to the aperiodic limit (Data from Hassenstein 1966)

neural information for their control. Feedback systems in metabolic and neuronal processes are mostly treated by direct mathematical modeling.

Further Reading

Flanders 2011; Rudall 2004.

5.2 Model Approaches to Some Complex Biological Processes

After introducing general approaches of systems analysis, and some general properties of complex systems, this section will give insights into the theoretical treatment of some particular processes. The enormous increase in research in this field, however, only allows to outline these considerations, simply showing their basic ideas and models.

5.2.1 Models of Biochemical Reactions

The first step in investigating biochemical reactions is the construction of the net of interactions as described in Sect. 5.1.3. To concretize these calculations, the particular edges and arcs must be represented by equations. In some cases, the conventional equations of chemical reactions of the first or second order are applicable. Usually, however, biochemical reactions are catalyzed by enzymes. In its simplest form, an enzymatic reaction may be described as follows:

$$\mathrm{S} + \mathrm{E} \rightleftarrows \mathrm{ES} \Rightarrow \mathrm{E} + \mathrm{P}.$$

A substrate S forms a complex ES with the enzyme E. This complex then breaks down to give the product P and release the enzyme E. The energetic advantage of these reactions has already been discussed in Sect. 2.1.5 and illustrated in Fig. 2.5. Here, we will concentrate on the kinetic consequences.

Usually, the rate-limiting step of this sequence of reactions is the break-down of the enzyme–substrate complex. Therefore, the rate of this reaction can be described in the following way:

$$J = k\, c_{ES} \tag{5.17}$$

where k is the rate constant of this break-down reaction. Further assuming that the first step of this reaction, i.e. the formation of the enzyme–substrate complex is always in thermodynamic equilibrium, then by the mass action law the equilibrium constant (K_M) can be evaluated:

$$\mathrm{K}_M = \frac{c_S\, c_E}{c_{ES}} \tag{5.18}$$

The total concentration of the enzyme (c_{E0}) must remain constant.

$$c_{E0} = c_E + c_{ES} \tag{5.19}$$

Combining Eqs. 5.18 and 5.19 one obtains:

$$c_{ES} = \frac{c_S\, c_{E0}}{\mathrm{K}_M + c_s} \tag{5.20}$$

Introducing this into Eq. 5.17, one gets the *Michaelis–Menten equation* which describes the reaction rate (J) of simple enzymatic processes as a function of the substrate concentration (c_s):

$$J = \frac{k\ c_S c_{E0}}{\mathrm{K}_M + c_S} \tag{5.21}$$

K_M, as the dissociation constant of the enzyme–substrate complex, is known as the *Michaelis constant*. It has the units of concentration, and corresponds to the concentration of the substrate when the reaction rate is at half of its maximum value (see Fig. 5.11). It usually lies somewhere between 10^{-2} and 10^{-5} M.

As shown in Fig. 5.11, there are two particular regions in this function where some simplifications are allowed. At high substrate concentrations ($c_S \gg \mathrm{K_M}$), a maximum reaction rate $J_{\max}$ is attained. For this case, Eq. 5.21 leads to:

$$J_{\max} = k\, c_{E0} \tag{5.22}$$

In this case, the flux is independent of substrate concentration (c_s). J_{max} can be introduced as a constant parameter into the general system of equations.

In contrast, at very low substrate concentrations ($c_S \ll \mathrm{K_M}$), Eq. 5.21 becomes:

$$J \approx \frac{k\, c_S c_{E0}}{\mathrm{K}_M} = \frac{J_{\max}}{\mathrm{K}_M} c_S \quad \text{for:} \quad c_S << \mathrm{K}_M \tag{5.23}$$

In this case (dash-dotted line in Fig. 5.11), a linear flux equation similar to Eq. 5.17 can be used. In many particular situations, therefore, simple approaches

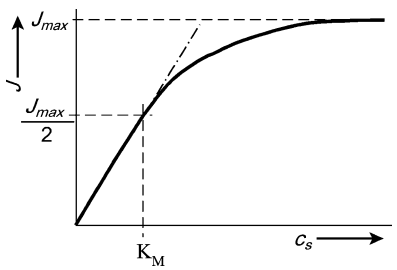

Fig. 5.11 The reaction rate (J) as a function of the substrate concentration (c_S) of an enzymatic reaction with constant enzyme concentration

are possible even in cases of enzymatic reactions. In all other cases, the Michaelis–Menten equation must be directly introduced into the system of flux equations. Obviously, this leads to nonlinear equations.

In fact, enzymatic reactions in metabolic systems are frequently regulated in a more complex way. A number of effectors may bind to the enzyme and interact as inhibitors or activators. Even the product of the enzymatic reaction itself can interact as a positive (product activation) or negative (product inhibition) feedback system. In a similar way, the substrate at higher concentration may occupy a second binding place of the enzyme, forming a ESS-complex that cannot form the product P. As the consequence, the Michaelis−Menten function (Fig. 5.11) with a final reaction rate J_∞ will be modified to a curve like the red line in Fig. 5.12a, which finally arrives at a lower level.

Some consequences of complex enzymatic systems can be demonstrated by a small reaction chain with nonlinear behavior and some stationary states (Fig. 5.12). A product (P) is formed by an enzymatic reaction from a substrate (S). The substrate (S) itself is produced by a first order reaction from a substance A. It can, however, break down with the same kind of kinetics to Z. The time dependence of the concentration of the substrate (c_S) can be formulated as follows:

$$\frac{dc_S}{dt} = J_{AS} - J_{SZ} - J_{SP} \tag{5.24}$$

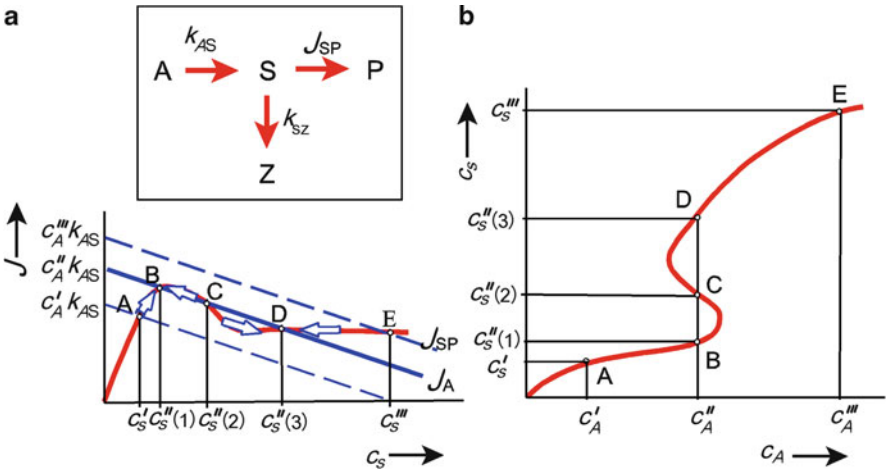

Fig. 5.12 A kinetic model of a simple biochemical reaction chain (insert). (**a**) The reaction rate (J_{SP}) (*red line*) as a function of the substrate concentration (c_S) for the enzymatic reaction S⇒P with substrate inhibition, and the difference of fluxes: $J_A = J_{AS} - J_{SZ}$ (*blue lines*) at various concentrations of the substance A (c_A', c_A'', c_A'''). The intersection points (A, B, C, D, E) of the *blue curves* with the *red* indicate stationary states. The *arrows* are directed towards stabile states. (**b**) The trajectory of the stationary states of the *left picture* as a function $c_S(c_A)$

Let the flux J_{SP}, which denotes the rate of the enzymatic destruction, be dependent on c_S in a kinetics with substrate inhibition as explained before. The function $J_{SP}(c_s)$ is illustrated in Fig. 5.12a. The fluxes J_{AS} and J_{SP} shall be determined by linear approaches according to Eq. 5.1. In this case, Eq. 5.24 can be transformed in the following way:

$$\frac{\mathrm{d}c_S}{\mathrm{d}t} = k_{AS}c_A - k_{SZ}c_S - J_{SP} \tag{5.25}$$

Introducing formally a flux:

$$J_A = J_{AS} - J_{SZ} = k_{AS}c_A - k_{SZ}c_S \tag{5.26}$$

then Eq. 5.25 transforms into:

$$\frac{\mathrm{d}c_S}{\mathrm{d}t} = J_A - J_{SP} \tag{5.27}$$

For the case of a stationary state, the following condition must apply:

$$\frac{\mathrm{d}c_S}{\mathrm{d}t} = 0 \quad \text{and therefore} \quad J_A = J_{SP} \tag{5.28}$$

The function $J_A(c_S)$, as defined by Eq. 5.26, is a straight line with a slope of $-k_{SZ}$, intersecting the ordinate at a point $k_{AS}c_A$ (Fig. 5.12a). The stationary states of the system according to Eq. 5.28 are marked by the points, where the blue lines of the functions J_A are cutting the red curve of the function J_{SP}.

Let us first consider the case $c_A = c_A''$. This straight line intersects with the function J_{SP} at points B, C and D. This means that there are three stationary states for one single value of c_A. However, these points are not equivalent to each other. Only states B and D are stable, C being unstable. This can easily be deduced from the curve.

First, consider the point B: a small increase in the substrate concentration will lead to the situation where function $J_{SP} > J_A$. Reference to the balance equation (Eq. 5.27) will show that this will bring about a reduction in c_S. The system, therefore, is moving towards its initial state. If c_S'' is reduced slightly, then $J_{SP} < J_A$, and this results in an increase in c_S. Small deviations from the stationary value B will thus be corrected. This behavior is indicated by the arrows on the function drawn in Fig. 5.12a. Around the point D, the situation is the same.

Near the point C, the system behaves differently. Small reductions in c_S'' cause J_{SP} to become larger than J_A. The substrate concentration will continue to fall, until point B is reached. An increase in the value of $c_S''(2)$ by a small value will cause the system to slide in the direction of state D which is stable like the state B.

The stationary states in this system can be summarized by a trajectory. In Fig. 5.12b, the function of the substrate concentration (c_s) versus the concentration

of the substance A (c_A) is shown. First, the three stationary values for c_A'' shown in Fig. 5.12a are plotted. A change in c_A causes a parallel shift of the curve J_A in Fig. 5.12a (for example, the dotted lines). The three stationary values occur only within a relatively narrow range of c_A (i.e. $J_A = c_A k_A$). If all the stationary states of c_s for different values of c_A are plotted in the coordinates of Fig. 5.12b, then an S-shaped trajectory is formed. This corresponds to a section through the behavioral surface of the cusp catastrophe, which is shown in Fig. 5.6. This particular example discussed here clearly shows that the points on the shaded area of the surface, corresponding to the point C in Fig. 5.12, are unstable states.

For the curve Fig. 5.12b, the flux J_{SZ} is taken to be constant. Any variation in the rate at which the substrate (S) is used, i.e. a modification of the coefficient k_{SZ}, will change the slope of the straight lines representing the function J_A in Fig. 5.12a (according to Eq. 5.26). It can easily be seen that, if the slope of this curve exceeds a critical value, a situation will develop where, independent of c_A, there will only be a single possible intercept. When compared with the functional surface in Fig. 5.6, this situation corresponds to a section through the bifurcation point P. In this instance, k_{SZ} (or J_{SZ}) would have to be plotted on the y-axis of the coordinate.

The analyses of metabolic networks of course lead to systems of many differential equations. The integrations of them are not only complicated, because most of them are nonlinear, but they are furthermore quite often ambiguous. Many possibly occurring stationary states are probably metastable. By small alterations of the parameters of the system, or by changing the initial conditions, the characters of the stationary states may modify in a non-easily predictable way. Even chaotic situations are possible. These properties not only make calculations difficult but they also sometimes lead to a higher degree of system destabilization.

In Sect. 3.1.4, we have already mentioned the time hierarchy of biological systems. In fact, the rate constants of biologically important reactions are distributed over a broad region of magnitudes. In the context of thermodynamic properties of systems, we have already discussed that, as a consequence in a steady state ($\sigma > 0$) of the whole system, fast reacting equilibria ($\sigma = 0$) of subsystems can be established. Considering this time hierarchy from the standpoint of systems analysis, the following conclusion about the behavior of systems can be drawn from mathematical considerations.

Formally, in a system of many differential equations, simplifications occur because differential quotients of fast processes may approach zero. This is the expression of the above-mentioned case where, in a steady state, as a non-equilibrium stationary state, some parts of the system are in equilibrium. Other processes, being very slow, and taking a shorter time scale, can be considered as non-existing, i.e. the corresponding parameters can be taken as constants. These kinds of simplifications not only help in mathematical modeling they automatically reduce the number of stationary states and stabilize the system itself. It seems that the established time hierarchy of biological reactions is the result of biological optimization in the course of evolution.

Further Reading

Heinrich and Schuster 1996; Klipp et al. 2009; Koch et al. 2011; Palmer 1995.

5.2.2 Pharmacokinetic Models

Optimization of modern therapy requires detailed knowledge of the kinetics of drug distribution in the body, as well as the time course of the induced physiological reactions. The first question concerns the *pharmacokinetics*, the second is the subject of *pharmacodynamics*. The latter includes sets of biochemical and physiological reactions, and concerns the problems of the previous sections in this book.

Pharmacokinetics can be considered as a particular kind of compartment analysis. In this case, the compartments, according to the definition given in Sect. 5.1.1, are represented by various organs and tissues of the body. A pharmacokinetic analysis answers the questions: how much, how frequently, in which way, and how long a definite drug must be administered to guarantee an optimal therapeutic efficiency, i.e. the optimal concentration to have the required effect.

The simplest kind of model calculation concerns the intravenously administered drug. In this case, the *bioavailability* is considered as 100%, and the drug is rapidly and homogeneously distributed throughout the whole blood volume. From here, it will be more or less quickly absorbed by other tissues and organs, metabolized, and finally excreted. As the first approximation, this process can be considered as an outflow from the blood volume as a one-compartment system. This can be described by the basic kinetic flux equation (Eq. 5.1, Sect. 5.1.1) and will follow the time function corresponding to Eq. 5.5.

Applied to the conditions of pharmacokinetics, this is:

$$\frac{\mathrm{d}m}{\mathrm{d}t} = -\mathrm{Cl}\, c \tag{5.29}$$

where c is the concentration of the drug in the blood, which equals m/V, i.e. the total amount of the drug administered (m) per volume (V). k is the rate constant, which is displaced by the *clearance constant Cl* as applied by the volume. This *clearance* is defined as the volume of blood that is totally cleared of its content of the drug per unit time. It has the unit $\mathrm{m}^3\mathrm{s}^{-1}$, or more commonly: 1/h.

The integration of Eq. 5.29 leads to:

$$c = c_0\, e^{-kt} = \frac{m_0}{V}\, e^{-\frac{\mathrm{Cl}}{V}t} \tag{5.30}$$

where m_0 is the total amount of drug applied at time t_0. As explained in Sect. 5.1.1 (Eq. 5.7), a half-life time of the drug can be obtained by:

$$t_{1/2} = \frac{V \ln 2}{\mathrm{Cl}} \tag{5.31}$$

It should be noted that the clearance constant (Cl) is defined as being related to the volume of blood, whereas the volume (V) in this compartment analysis means the total available volume for the distribution of the particular drug. These, in some cases, may differ from each other.

For a particular pharmacokinetic analysis, the distribution of the drug in different parts of the body is of interest. For this, the linear approaches of the compartment analysis have been proved to be useful.

Figure 5.13 demonstrates an example of such a system, which can be regarded as a basic pharmacokinetic model. Schematically, it indicates the case of a single oral dose of a drug which is then taken from the stomach into the blood stream. From here, it is distributed into the tissues of various organs, and at the same time excreted by the kidney.

In this way, the drug distribution can be calculated for various kinds of administrations, considering differences in the time constants of resorption, excretion, and destruction by metabolism. Computer simulations allow one to consider stationary concentrations in particular tissues after repeated administrations of a drug or following a permanent infusion.

Usually, pharmacokinetic studies are based on linear models. In some instances, a concept known as *capacity-limited elimination* is used. This is the excretion of a drug using saturation kinetics. For the direct binding of the drug to a receptor, nonlinear approaches should be used, similar to the Michaelis–Menten equation (Eq. 5.17).

Further Reading

Gabrielsson and Weiner (2006), Rowland and Tozer (2011).

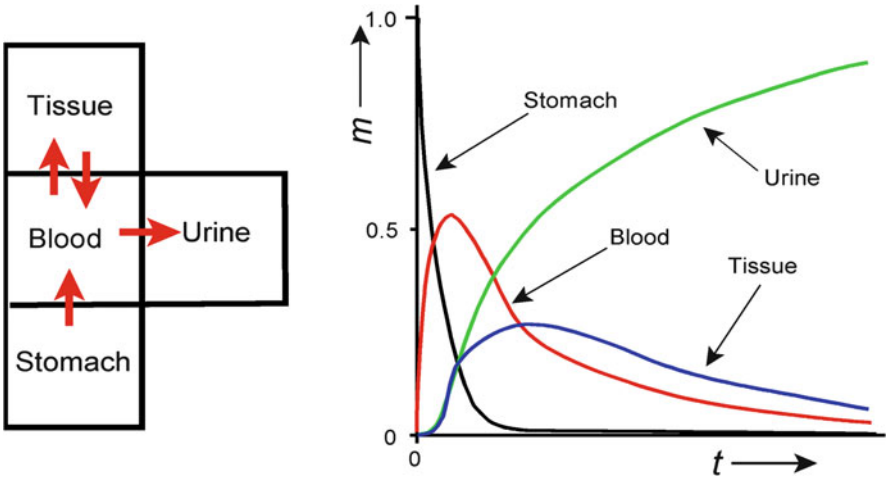

Fig. 5.13 Time course of drug distribution in a basic pharmacokinetic model: At $t = 0$, the drug is taken *per os*, subsequently transferred by the blood stream into the tissue and finally excreted. (The amount of the drug (m) is used as relative unit.) (Data from Knorre 1981)

5.2.3 Models of Propagation and Ecological Interactions

Ecological systems and their interactions are in fact extremely complex. Besides the *abiotic* conditions of the environment, which are mostly time dependent, a large number of *biotic* influences occur, like competition, predation, food webs, migrations, etc. The difficulties of the mathematical treatment of biological systems become evident in ecological models in an extreme way.

Actually, two kinds of models are used to handle this problem. A first approach uses *educational models*, which limit the investigation to one or two essential processes, separated from the complexity of the whole problem. These models are usually restricted to some simple assumptions, but in their favor, they lead to analytically tractable equations. In this way, even reflecting just more or less idealized conditions, this approach is suitable to gain insights into some general properties of the process, and furthermore include the possibility to extend the calculations to more complex models in a next step.

The second kind of approach is to use *practical models* which are based on a larger number of more or less realistic data and assumptions. These models essentially include large numbers of variables and parameters, which to some extent must be evaluated by rough estimations. The corresponding sets of equations are much too complicated for an analytical treatment. Results can be obtained just by numerical simulations. With increasing complexity, these models, however, become more and more ambiguous, and therefore more doubts are indicated whether they in fact describe realistic situations.

Anyway, these models can be formulated mathematically either by deterministic or by stochastic approaches. The *deterministic models* of growth and interactions in biological populations use phenomenological variables, and are finally based on the same kind of mathematical constructions as demonstrated in the previous sections for calculations of chemical reactions or transport processes. The change of the number (n) of individuals, which is described by the differential quotient $\mathrm{d}n/\mathrm{d}t$, can be considered, for example, in first approximations as proportional to their number (kn). These kinds of approaches are possible in considering large populations and averaging the whole process. In contrast to this, the *stochastic approaches* start with the consideration of the probability of the event of the birth of a new subject or of the death of an existing one. In this way, the probability $P(n, t)$ can be evaluated, indicating that the number of the individuals in the population increase or decrease.

In fact, there exist some differences in the results of these two approaches, at least in some particular situations. Let us consider firstly a simple deterministic example. Let the time constant (k) of the population be composed of a coefficient of propagation (k_p), and another for the mortality rate (k_m):

$$k = k_p - k_m \quad \left(k_p,\ k_m > 0\right) \tag{5.32}$$

In this way, the proliferation of a population can be described by the following differential equation:

$$\frac{dn}{dt} = k\,n \tag{5.33}$$

This expression states that both the rate of proliferation as well as the rate of mortality are proportional to the actual number of individuals in the population. Considering that at $t = 0$ the initial number of individuals is n_0, and $k_p > k_m$, i.e.: $k > 0$, then the integration of Eq. 5.33 leads to the following equation:

$$n = n_0\, e^{kt} \tag{5.34}$$

This function shows an exponential growth up to infinity. It was discovered as early as 1798 by the English economist Thomas Robert Malthus. The time course of this function is depicted in Fig. 5.14 (blue line).

In this case, where k_m dominates, i.e. if the birth rate is lower than the rate of mortality, it becomes $k < 0$, and the number of individuals declines exponentially. From the mathematical point of view, the variable (n) in this way can become extremely small, but never zero. In other words, this approach excludes the realistic possibility of a full extinction of the population. In contrast, in the probabilistic models, this situation is included.

Another difference in these two approaches concerns the real-time dependence of biological propagation. In contrast to the deterministic models which assume a continuous process of birth and death, the stochastic approaches may better consider discrete intervals of time between these processes and even a completely non-overlapping of the generations, as happens, for example, in some insect populations.

In fact, educational models usually consider deterministic models, which we will discuss here, in order to demonstrate some general principles of population growth.

The original assumption of Eq. 5.33 supposes that the external milieu is constant and not affected by the organisms themselves. As a result, an infinitely growing Malthus function occurred (Eq. 5.34). This, however, reflects an idealized condition. Only when the number of subjects in the reservoir, i.e. cells in a suspension, is

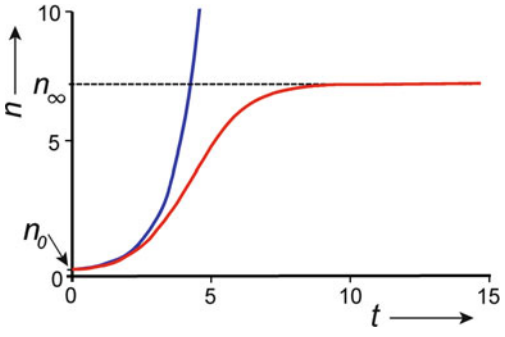

Fig. 5.14 The function $n(t)$ according to the Verhulst−Pearl equation (Eq. 5.37) (*red line*), and the Malthus equation (Eq. 5.34) (*blue line*). Using arbitrary values in the ordinate and abscissa, the parameters are: $n_0 = 0.1$; $k_1 = 1$; $k_2 = 0.14$. In the case of Malthus curve, $k = k_1$

very small, and the time of consideration is sufficiently short, can this condition apply. In real situations, this in fact marks the initial exponential period of growth.

In most cases, a mutual influence of the individuals on one another is to be expected, as, for example, competition for food. This circumstance can be taken into account by modifying Eq. 5.32. It can be better postulated that k_p and k_m are not constant parameters but themselves functions of the population density. So, for example, one can assume that the rate constant of proliferation (k_p) will decrease and the rate constant of mortality (k_m) will increase, when n increases, because of shortage of food. The degree of physiological variability of these parameters, however, will be limited for genetic reasons. The simplest approach is the following linear equation:

$$k = k_1 - k_2 n \quad \text{for: } k_1,\ k_2 \geq 0 \tag{5.35}$$

This takes into account an optimal propagation constant, k_1, determined genetically, in combination with a further parameter, k_2, leading to a diminution of k_1 with an increase of the population density (n). If $n \Rightarrow 0$ then $k_1 \Rightarrow k$. Introducing Eq. 5.26 into Eq. 5.33, one obtains:

$$\frac{\mathrm{d}n}{\mathrm{d}t} = (k_1 - k_2 n)n = k_1 n - k_2 n^2 \tag{5.36}$$

This approach therefore leads to a first order nonlinear differential equation. The integration of this equation for the same initial conditions as above ($t = 0$; $n = n_0$) gives an equation which in mathematics is known as the *logistic function*. In population kinetics, it is called the *Verhulst–Pearl law*:

$$n = \frac{n_0 k_1\ \mathrm{e}^{k_1 t}}{k_1 + k_2 n_0 (\mathrm{e}^{k_1 t} - 1)} \tag{5.37}$$

In contrast to Eq. 5.34, in this function, n does not approach infinity at $t = \infty$ but rather tends towards a saturation value (n_∞). This saturation value can easily be determined from Eq. 5.37 by considering that, with increasing t, the exponential terms increase up to a value where neither the subtraction of 1 nor the addition of k_1 become important. This leads to:

$$\lim_{t \to \infty} n = n_\infty = \frac{k_1}{k_2} \tag{5.38}$$

Furthermore, it is easy to show that, for small values of t, the Eq. 5.37 takes the form of Eq. 5.34. Both functions, for some arbitrarily chosen parameters, are illustrated in Fig. 5.14. An example of experimentally obtained data is shown in Fig. 5.15.

If the system contains individuals of several different species, then the interrelationship between them must be taken into consideration. This can involve competition for the food supply (Fig. 5.15, below) or a predator–prey relationship

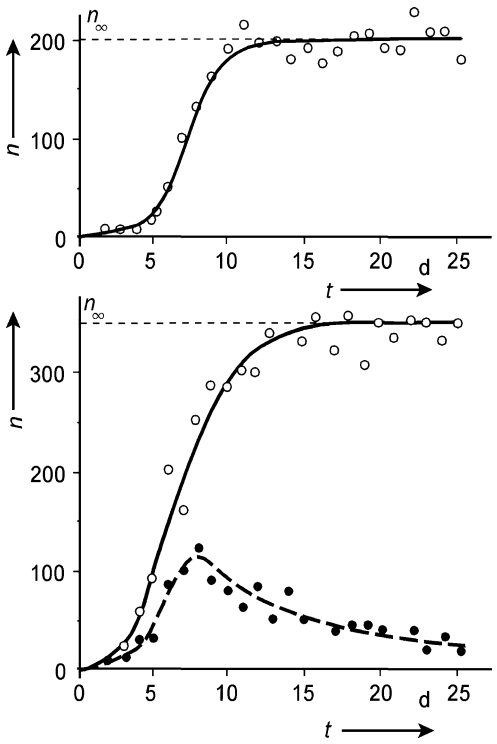

Fig. 5.15 The time course of the growth of a pure population of *Paramecium caudatum* (*full line*) alone (*above*), and mixed together with a population of *Paramecium aurelia* (*dashed line*) (*below*). The ordinates indicate the number of individuals (n) per 0.5 ml medium. The abscissa gives the time in days (d) (Data from Gause 1935)

(Fig. 5.16). Systems like this had already been studied in the 1930s by Vito Volterra and Alfred Lotka.

The following set of equations describes the simplest case of a predator−prey relation between two species:

$$\left.\begin{aligned}\frac{\mathrm{d}n_P}{\mathrm{d}t} &= \left(k_{pP} - k_{mP}n_H\right)n_P \\ \frac{\mathrm{d}n_H}{\mathrm{d}t} &= \left(k_{pH}n_P - k_{mH}\right)n_H\end{aligned}\right\} \tag{5.39}$$

The subscripts p and m have the same meaning as in Eq. 5.32, and the index P in these equations refers to "prey," whereas H means "hunter." For the prey organisms, the mortality ($k_{mP}n_H$) will increase as the number of predators, i.e. hunters (n_H) increases. The greater the number of predators, the shorter will be the life expectancy of the prey organisms. Otherwise, an increasing number of prey organisms means better living conditions for the predators. For this reason, the propagation coefficient of the predators (k_{pH}) is multiplied by the amount of prey organisms (n_P) in the second equation.

This set of Eq. 5.39 represents an extremely simplified situation. The interaction of individuals of the same species, for example, which was already reflected in

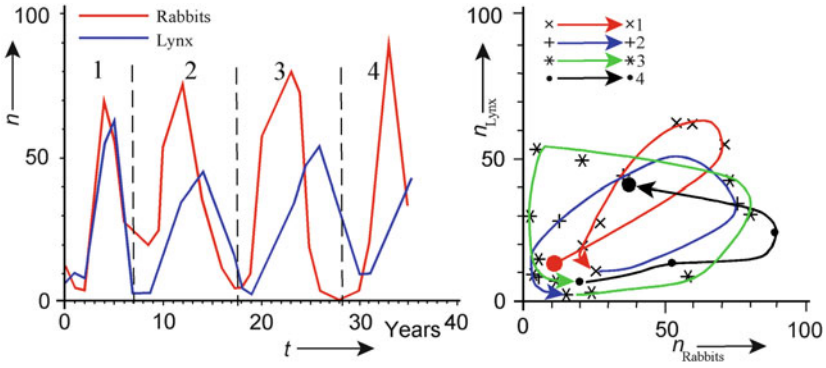

Fig. 5.16 A particular example of population kinetics: the number of pelts obtained from hunted lynx and snow rabbits in Canada after 1900 (numbers in thousands). On the *left*: time course; on the *right*: the same numbers in a parameter plot. The population waves 1. . .4 of the left part of this figure appear as Volterra cycles in the graph on the right, starting at the *red point* and ending at the *back point* (Data from Haken 2010)

Eq. 5.36, has been ignored here. This example, however, demonstrates that even the simplest ecological models already lead to complicated systems of nonlinear equations. The solution of such systems of equations frequently leads to functions indicating oscillations. This corresponds to experiences where ecological systems oscillate even without additional input of external signals, such as daily or annual oscillations of light and temperature.

These oscillations are called *Volterra cycles*. They can either be plotted versus time, or represented as trajectories. In the case of Eq. 5.39, for example, the functions $n_P(t)$ and n_H(t) can be plotted separately or combined as the trajectory $n_P(n_H)$. Again, a system trajectory would appear like those shown in Figs. 5.4, 5.5, 5.6, and 5.12b.

Figure 5.16 demonstrates a concrete example of such oscillations. It uses data from the hunting lists of the Hudson Bay Company in Canada which were kept from 1845 to 1935. Supposing the amounts of hunted animals were proportional to the number of those really existing, one can roughly use this as a representation of the ecological situation. In Fig. 5.16, the data from 1900 to 1935 were used, indicating the time course with typical oscillations (left) and the same data in a parameter plot showing Volterra cycles (right).

In a similar way to the predator−prey system, a kind of symbiotic inter-relationship of two species can be formulated:

$$\left.\begin{aligned}\frac{dn_A}{dt} &= \left(k'_{pA} + k''_{pA}\, n_B - k_{mA}\right) n_A \\ \frac{dn_B}{dt} &= \left(k'_{pB} + k''_{pB}\, n_A - k_{mB}\right) n_B\end{aligned}\right\} \tag{5.40}$$

This set of equations takes into account the fact that when $k'_{pA} > k'_{mA}$, i.e. $k'_{pB} > k'_{mB}$, then one species of this symbiotic community can survive without the other. The symbiotic cooperation is expressed by the product terms $k''_{pA}\, n_B$ and $k''_{pB}\, n_A$.

Mathematical approaches of this kind form the basis for modeling of ecological systems. However, other factors must also be considered such as a large number of trophic levels and complex inter-relationships, time delays caused by processes of developments, annual rhythms, and also ethological and sociological structures.

Further Reading

Ginzburg and Colyvan 2004; Jeffries 1989; Okubo and Levin 2010; Renshaw1990; Turchin 2003.

5.2.4 Models of Growth and Differentiation

The morphology of an organism, its geometric structure and the corresponding morphogenetic mechanisms, have provided time and again subjects for modeling. In 1917, in his famous book "On Growth and Form," D'Arcy Thompson first formulated the mathematical ideas of morphometry. His theory states that different types of body shape of an animal, or parts of it, if projected onto a system of coordinates, can be systematically transformed into another by means of the mathematically defined distortion, i.e. by a sort of coordinate transformation. D'Arcy Thompson demonstrated this method for the shapes of various species of fishes as well as forms of crab carapaces, mammalian skulls, and even leaves of plants and many other structures. In this way, for example, phylogenetic changes can be expressed mathematically. Such approaches belong more to the realms of biometry. They do not include the mechanisms of differentiation and shape formation. The same applies to *allometric models* of growth which are discussed in Sect. 3.8.

A striking phenomenon of living subjects is their ability of self-organization. What mechanisms finally transform the genetic information of an organism into its final structure? In which way a more or less homogeneous egg, or even a single embryonic stem cell, may produce a well-organized animal? What processes control the specific pattern formation in plants and animals?

Experiments for the purpose of understanding processes of regeneration and development range back to the eighteenth century. The German biologists Hans Spemann and Hilde Mangold were crowned by the Nobel prize in 1935 for their experiments on eggs and embryos of amphibians. They supported the idea of an organizing center, later named *Spemann organizer*, which influences the differentiation of cells around a concentration gradient of a certain activator substance.

In the following, a large number of genetic and epigenetic mechanisms were found, controlling growth and differentiation in various plants and animals by systems of activator- and inhibitor molecules. These substances control the

morphogenetic movements of cells, as well as the activation of the genes that are responsible for differentiation. In this context, reference must be made to the discussions on the possible role of electric fields in embryogenesis, leading to electro-diffusion of growth factors (Sect. 3.5.2).

This progress in molecular biology meets the advantage of the theory of nonlinear processes, and of course was strongly supported by the development of powerful computer techniques. As already explained in Sect. 3.1.4, dissipative structures can be established spontaneously in nonlinear systems as a result of superimposition of two or more processes. The Volterra cycles of population dynamics, as described in Sect. 5.23 (Fig. 5.16), like many other oscillations in biological systems can be considered as dissipative structures in time. Evolutionary patterns, on the other hand, are examples of dissipative structures in a spatial "frozen" state. In this way, they are the consistent result of time-dependent concentration gradients.

A first mathematical formulation of this idea was published by Alan Turing in 1952. He demonstrated that spontaneous pattern formation is possible as the result of interaction of two components with different diffusion rates. This is the origin of the *reaction–diffusion model of morphogenesis*, a mathematical formulation of the physiological idea of Spemanns organizer system.

The basic equation for these calculations is the second Fick's law (Eq. 3.32, Sect. 3.3.1) which explains the time-dependent formation of a concentration gradient of a substance as the result of diffusion, starting at the point of its production or activation. In the case where the molecules were supported permanently from a source into the target, this system does not have steady states, but just temporarily stable gradients can be formed. To attain a concentration gradient, as a time-independent steady state, the diffusion term must be completed by a certain reaction of molecule depletion:

$$\frac{\mathrm{d}c}{\mathrm{d}t} = D\frac{\partial^2 c}{\partial x^2} - kc \tag{5.41}$$

This is the simplest approach, because it represents just a one-dimensional system, and furthermore, a linear kinetics of activator depletion where the parameter k represents the time constant of this reaction. In fact, the particular kind of depletion reaction is crucial for the formulation of a realistic model.

Furthermore, a reaction term must be introduced, characterizing the source of the activator. In many models, an autocatalytic kind of reaction is supposed. This, in the simplest way, is a kind of bimolecular reaction which is nonlinear, i.e. determined by the square of concentration. An important aspect was the introduction of an antagonist to the activator, a kind of inhibitor, the formation of which is actually initiated by the activator itself. Finally, it was established that pattern formation is possible only if a locally restricted self-enhancing reaction is coupled with an antagonistic reaction that acts on a longer range.

In the simplest way, two substances, an activator (A) and an inhibitor (I), are considered, which are formed at particular sites and then distributed by diffusion. In this way, complex concentration patterns of both substances are set up, generating a

particular space−time structure. The growth rate, or the mode of differentiation, finally depends on the concentration ratio of these two factors.

Consider again just a simple unidimensional case where the concentration of an activator (c_A) and an inhibitor (c_I) are distributed only in the x-direction. For this case, the following set of equations can be formulated:

$$\left.\begin{aligned}\frac{\partial c_A}{\partial t} &= k_A \frac{c^2{}_A}{c_I} - k'_A\, c_A + D_A \frac{\partial^2 c_A}{\partial x^2} \\ \frac{\partial c_I}{\partial t} &= k_A c_A^2 - k'_I c_I + D_I \frac{\partial^2 c_I}{\partial x^2}\end{aligned}\right\} \tag{5.42}$$

This is a system of partial, nonlinear differential equations: *partial,* because the variables c_A and c_I are differentiated both with respect to time and position, and *nonlinear,* because the square of the concentrations appears as the result of an autocatalytic process.

In particular, the following assumptions are made in this model. The rate of formation of the activating factor A is determined by a process that is autocatalytically promoted by c_A, and furthermore suppressed by the concentration of the inhibitor (c_I). This corresponds to the term: ($k_A \frac{c_A^2}{c_I}$). This quadratic term of activator formation overcomes the decay of the activator which is expressed by the simple linear term ($-k'_A\, c_A$). The diffusion along the x-axis ($D_A \frac{\partial^2 c_A}{\partial x^2}$) again follows the second Fick's law. The rate of inhibitor production also depends on the square of the activator concentration. It differs from that of the activator (A) only by the circumstance that the autocatalytic activation ($k'_A\, c_A$) does not include an additional inhibitor term.

The behavior of such a system depends to a great extent on the choice of the constants. The relation of the diffusion coefficients D_A and D_I is of particular importance. If the diffusion constant for the inhibitor (D_I) is much larger than that for the activator (D_A), then the distribution of the factor I can be regarded as being homogeneous.

This figure shows an extremely simplified case. However, it illustrates the formation of a more or less periodic structure in a one-dimensional system. The question arises: Why there is no self-inhibition of the heterocysts by their own inhibitor production? The answer is the $c_A{}^2$-terms in Eq. 5.42. This nonlinear self-enhancement, together with the linear process of lateral inhibition, makes a uniform distribution unstable and eventually leads to the periodic structure.

The simplified one-dimensional approach is applied to a particular case: the cell differentiation in the blue-green algae *Anabaena*. Under nitrogen-limiting conditions, their filaments at some intervals differentiate vegetative cells into larger non-dividing heterocysts, which are specialized for nitrogen fixation. The distance between two heterocysts in the filament usually is larger than 12–14 cells.

To understand this reaction, it is postulated that the activator of heterocyte-differentiation, which has actually been identified as the transcription factor HetR, which activates itself in an autocatalytic reaction like Eq. 5.42. In the same way, however, it activates the formation of a small peptide PatS as an inhibitor. In contrast to the inhibitor, which diffuses along the filament, the activator does not

leave the cell ($D_A = 0$). The result of this proposal is illustrated schematically in Fig. 5.17.

Another model shows an approach to explain the formation of filamentous branched structures in a two-dimensional system with a similar activator−inhibitor mechanism (Fig. 5.18). In this case, it is assumed that at first the plane contains a homogeneous distribution of a certain trophic factor which allows cell differentiation. If this occurs at a particular point, a trace starts to develop behind a moving local signal. Furthermore, the differentiated cells lower the concentration of the trophic substrate in their vicinity. Only in some cases, where this reaction is not strong enough, do branches of the filament appear.

Figures 5.17 and 5.18 show two simple kinetic models of differentiation processes, just to demonstrate the basic ideas of this reaction–diffusion model. In addition to the initial idea of Turing (1952), an autocatalytic activator−inhibitor mechanism must be included, to attain these results. The molecular basis of such systems, meanwhile, has been found experimentally in many concrete systems. In this way, this model, starting with more or less hypothetical assumptions, had been verified and concretized in many cases.

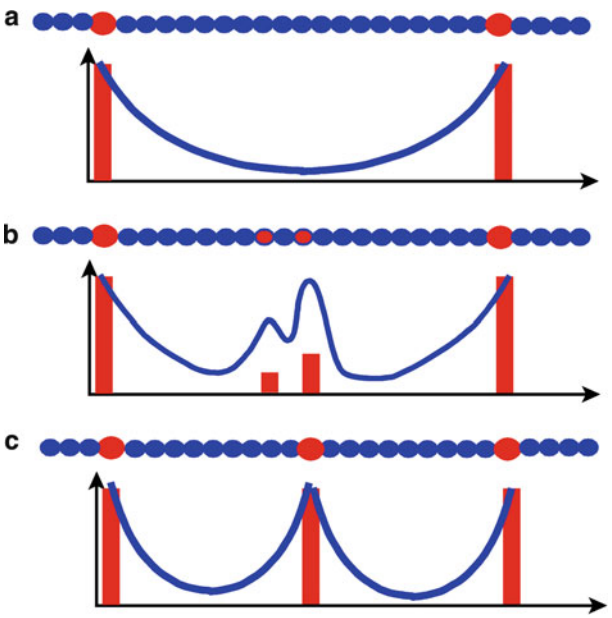

Fig. 5.17 A scheme of heterocyst formation (*red spheres*) in a filament of normal vegetative cells (*blue spheres*) of *Anabaena* as an interplay of an non-diffundable activator (*red bars*) produced in individual cells, and an inhibitor (*blue lines*), which itself is produced catalytically by the activator (corresponding to an equation similar to Eq. 5.42). If the distance between two inhibitor producing heterocysts is large enough (**a**), new heterocysts began to differentiate. They are in competition among each other, because of their own inhibitor production (**b**). Finally, only one of them becomes effective (**c**) (Modified after Meinhardt 2008)

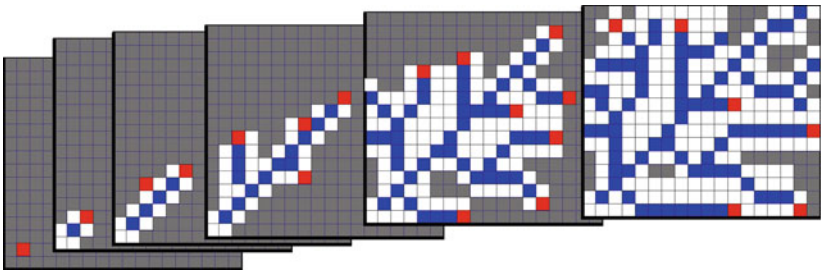

Fig. 5.18 Formation of a branched structure behind moving signals. A homogeneously distributed guiding factor (*gray squares*) stimulates a spontaneous cell differentiation. Starting with a single differentiated cell (*red square*, *left plane*) an elongation occurs (*blue squares*), whereas the differentiated cells remove the trophic substrate in their neighborhood (*white squares*). In some cases, where the lateral inhibition temporary is not strong enough, branches can be formed (After Meinhardt 2008, redrawn)

The one-dimensional approach of Fig. 5.17 in this way was extended to the two-dimensional system in Fig. 5.18. Finally, it is possible to formulate even three-dimensional systems of differentiation. In many cases, however, the three-dimensional orientation is finally based on two-dimensional events, because it takes place in two-dimensional cell sheets. In organs such as three-dimensional structures, firstly a two-dimensional pattern is generated which subsequently folds or form tubes or other epithelial structures.

Starting from more or less simple models, like tentacle formation of the fresh-water polyp hydra, or the regeneration of the flatworm planaria, which are supported by a sufficient experimental basis, similar approaches are meanwhile used to explain differentiation processes on higher organized animals. So, for example, the embryonal orientation of the main body axis and the pattering of limbs in vertebrates, the formation of insect legs, but also many examples of pattern formation in reptiles and mammals, have been investigated. Some of these models are purely hypothetical, while others are based on molecular biological data.

Further Reading

Maini and Othmer 2000; Meinhardt 2008, 2009; Murray 2003; Wartlick et al. 2009.

5.2.5 Models of Origin and Evolution of Life

The interest in the question, how life could have originated and developed, is not new. It is stimulated to a great extent by progress in our knowledge of the early history of the earth, by results in molecular biology, and last but not least by the advances in systems theory in general. In the 1950s, Stanley L. Miller showed experimentally that electric discharges applied to a mixture of CH_4, NH_3, H_2, and water vapor could generate a number of amino acids. These experiments confirmed

earlier speculations but at the same time they provoked a number of new questions: What concrete physicochemical conditions had led to the origin of the first biologically significant molecules? Were these polypeptides first, or instead the nucleic acids? How could the statistically completely improbable amino-acid sequence of functioning proteins occur? What is the origin of the genetic code? The answers to these questions are expected from geophysics, from physical chemistry, and from various theoretical approaches.

The chemical evolution could have happened only after the temperature on the earth surface allowed the existence of liquid water, which was considered to have occurred only about 3.8 billion years ago. On the other hand, the first fossil cells have been assigned an age of 3.5 billion years. The period of chemical evolution, therefore, probably took less than 300 million years. Considering the above-mentioned hen-and-egg problem, it is assumed that probably the polypeptides developed before the nucleic acids. This is based on some arguments: A chemical evolution is possible only with the help of some enzyme-controlled mechanisms, which can be realized even by particular polypeptides. On the other hand, the phosphates, which are essential for the formation of RNA, would have been precipitated by heavier metal ions in the "primordial soup," and therefore not have been easily available. Furthermore, these polypeptides in contrast to nucleic acids indicate a larger resistance to the extremely high UV-irradiation at that time.

The formation of polypeptides seems to be catalyzed by surface reactions of minerals like kaolinite, montmorillonite, hectorite, or other smectites. This was probably promoted by local heating−drying−wetting cycles in this period. In this context, the hypothesis of the *salt-induced peptide formation* (SIPF) *reaction* should be mentioned. It considers that, above a concentration of 3 M, the NaCl can be considered as a condensation reagent, helping amino acids to form oligomers through peptide linkages.

So far, the physicochemical background of recent speculations on the origin of the first bio-organic molecules, concerning the first two questions which are mentioned above, have been considered. The next questions are the subject of various theories of probability, as well as information, game, and network theory. This mainly concerns the problem of the generation of a particular amino-acid sequence transforming a polypeptide with stochastic composition into a functional protein.

Consider a particular enzymatic biopolymer, a rather small one with an information content of only 10^3 bits. As a product of stochastic combination, it would result, on average, once in the course of $2^{1,000} \approx 10^{300}$ chances. This means that, out of 10^{300} resulting molecules, only one could be considered as being able to realize this particular function. To illustrate this abstract situation in a more realistic picture, one should realize that the whole hydrosphere of the earth contains only about 10^{32} molecules. Even considering the whole time of 10^{17} s of the existence of the earth and an extremely fast recombination reaction of 10^{-9} s, the probability of stochastic origin of this enzyme would be minute (see also Sects. 2.1.2 and 2.1.3). Albert Einstein once compared this probability with the chance of a complete dictionary resulting from an explosion in a printing office!

A solution of this problem offers the *quasispecies theory,* proposed by Manfred Eigen (1971). This suggests that, already at the beginning of the chemical evolution, the principles of competition are responsible for the development of particular chains of reactions and patterns of polymerizations. The formation of a functional amino-acid sequence of proteins is therefore not a product of chance but of selection, as in the case of further evolution, according to the principles of Charles Darwin.

These effects of *prebiotic molecular selection* can be considered as processes of catalysis and auto-catalysis. Auto-catalysis can be regarded in this context as a forerunner of self-reproduction.

The number of monomers in a given region is limited, and this will lead to a form of competition between the different kinds of polymerization reactions. However, such a system would rapidly reach an equilibrium state if it were not exposed to a continuous flow of energy and entropy. In fact, it will be disturbed permanently by geophysical effects of heating, cooling, drying, etc. From the thermodynamic point of view, the resulting combinations can be considered as a kind of dissipative structures (see Sect. 3.1.4).

Under these conditions, a selection of polymers with high catalytic activity will take place. A network of catalytic relationships between these individual molecules is built up (Fig. 5.19). If a closed chain occurs in such a network, this will form an auto-catalytic cycle. This is a structure with a good chance for survival. Such cycles sustain themselves, and in doing so, they can modify their own structure. Side chains, which are formed through the poly-catalytic activity of several molecules, are removed and distensible intermediate linkages are eliminated. In this way, the cycle becomes smaller. At the same time, it can be shown that

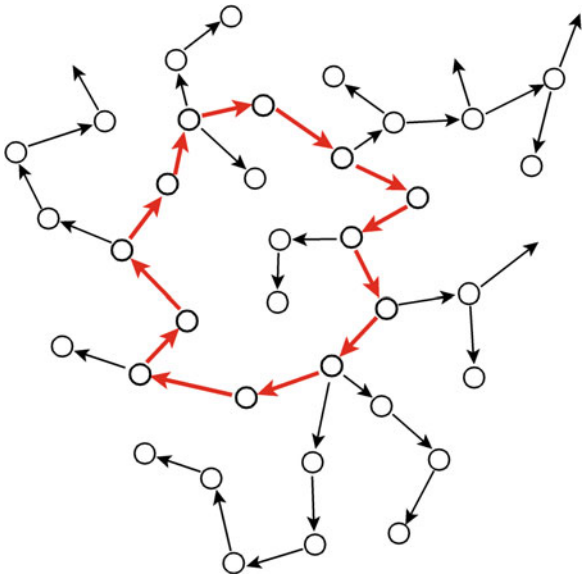

Fig. 5.19 The network of catalytic correlations in a mixture of polymers. The *red arrows* indicate an auto-catalytic hypercycle

heterogeneous cycles, made up of alternating proteins and nucleic acids, possess a selection advantage. Eigen called such structures *hypercycles*.

This concept of evolution, therefore, indicates a dialectic unity of stochastic and deterministic processes. The stochastic component is derived from the stability of bonds, and from the stochastics of molecular contacts in the course of the reaction. At higher levels of development, this stochastic element is realized by the mutation rate or the dependability of self-reproduction.

This theory is backed up by mathematical modeling and computer simulations. It is possible to formulate equations for systems of chemical reactions as well as for processes of reproduction and selection at the highest levels of evolution. With regard to the equations that were given in the previous sections, the following simplified approach can be demonstrated:

$$\frac{\mathrm{d}n_i}{\mathrm{d}t} = \left(k_{pi}q_i - k_{mi} - k'_i\right)n_i + \sum_{j \neq i} k_{m,ji}n_j \tag{5.43}$$

This is the equation for the time course of the number of individuals (n_i) of a species i, which embodies the following postulates: The individuals propagate at a rate that bears a linear relationship to the existing number of individuals and is determined by a time factor k_{pi}. This coefficient is, however, reduced by the factor q_l being a little less than one. In this way, the fact is taken into account that, by mutations, the identical reproduction of the individuals is disturbed. The coefficient governing the mortality (k_{mi}) is the same as in Eq. 5.32. In addition, another coefficient, k'_i, is introduced, expressing the possibility that some individuals will leave the area, i.e. will diffuse away. The summing term at the end of the equation reflects the possibility of the emergence of an individual of the species i from another of the species j by mutation. In contrast to the ecological equations in Sect. 5.2.3, n_i is thus capable of increasing even under the initial conditions, when $t = 0$ and $n_i = 0$. This sum term is a prerequisite for the neogenesis of a species.

This is, of course, a very primitive approach. It presupposes that propagation is based on linear relationships, an assumption having limited validity as has already been discussed in Sect. 5.2.3. The relationships between species have also not been considered. Nevertheless, evolution because of competition between species can be demonstrated on the basis of this approach if the following mass balance is used:

$$\sum_{i-1}^{m} n_i = \text{const}, \qquad \text{therefore: } \sum_{i-1}^{m} \frac{\mathrm{d}n_i}{\mathrm{d}t} = 0. \tag{5.44}$$

The advantage for selection of a species will then result from the relationships of the chosen system constants, one to another. With more complex evolutionary systems, weighting functions will have to be introduced but these must necessarily be of a speculative nature.

Besides the Eigen model, J. F. Crow and M. Kimura in 1970 introduced another quasispecies model of evolution. In contrast to the Eigen model, where the

sequences are subjected to point mutations only during the process of replication (*connected mutation-selection*), in the Crow−Kimura model, mutation and selection are two independent processes (*parallel mutation-selection*). The relationships of these models indicate that the optimum mutation rate in the connected mutation−selection model is lower than that in the parallel mutation−selection model. The Crow−Kimura model therefore gives an adaptive advantage under changing environmental conditions.

A further problem of prebiotic evolution concerns the development of the genetic code which translates the four-letter language of the nucleic bases, A, T, G, and C, into the 20-letter language of the amino acids. In fact, it is highly redundant, with the consequence that all amino acids, except methionine and tryptophan, are encoded by multiple codons. It is suggested that, in a primitive living systems, amino acids and nucleotides coexist and bind non-specifically. This kind of random association of nucleotides nevertheless could catalyze the synthesis of random polypeptides, barely functioning as proto-proteins. The evolution from the *non-coding* to the *coding* state is considered as a result of the interplay of the three conflicting forces: the need for diverse amino acids, for error-tolerance, and for minimal cost of resources. Therefore:

$$(\text{fitness}) = (\text{diversity}) - (\text{error - load}) - (\text{cost})$$

These relationships can be evaluated by mathematical expressions. The result indicates a certain degree of smoothness. In fact, the view of some codes in bacteria, protozoa and mitochondria indicate that the standard genetic code is not strictly universal. This suggests that the code was open, at least for some time, to evolutionary change.

Interestingly, these sorts of calculations lead to the *map coloring problem*. The question is about the minimal number of colors which are required to color an arbitrary map on a surface such that no two bordering countries have the same color. The result is known as the *coloring,* or *chromatic*, *number*.

Further Reading

Ancliff and Park 2010; Dyson 1999; Eigen 1992; Obermayer and Frey 2010; Rode et al. 2007; Tlusty 2010.

5.2.6 Models of Neuronal Processes

The neuronal system, starting from the excitation mechanism of a single neuron up to reactions of neuronal networks in the brain, is extremely complex and multilayered. Correspondingly, the theoretical approaches attempting to explain and model these processes are quite heterogeneous. A number of newly introduced experimental techniques for in vivo, as well as in vitro investigations provide a mass of new information requiring to be functionally connected, and included in model

considerations. This concerns methods for simultaneous recording of multiple single units and optical recordings with voltage and ion-sensitive dyes, as well as large-scale measurements of brain structure and activity with positron emission tomography (PET), magnetoencephalogram (MEG), 2-deoxyglucose (2-DG), and magnetic resonance imaging (MRI). In parallel, the enormous development of computer techniques in the last decades has stimulated the efforts of theoretical modeling.

The development of neuronal and brain models is helpful for a number of questions. In general, it makes the brain system with many interacting components more accessible. Furthermore, it is able to simulate and predict some properties of the neuronal system, which leads to new experiments, and makes it possible to understand and concentrate experimentally obtained data. In some cases, experiments that are difficult or even impossible to perform in living tissue can be simulated by the use of a model. The major research objective of computational neuroscience, however, is to discover the algorithms used in the neuronal systems.

As the result, the number of neuronal models becomes nearly unmanageably large. In the second half of the 1980s, the term *computational neuroscience* was introduced, some years before the coining of the expression *systems biology*. To some extent, this term replaced the *biocybernetics* which in the 1960s had been used to summarize research on many neuronal models.

Nevertheless, the trend to display neuronal functions by simplified analog models is as old as neurobiology itself. Already in 1907, Louis Lapicque used an electronic trigger circuit to model this process. A capacitor is charged up to a certain threshold of voltage, leading to a reaction of instantaneous discharge. This model was completed in the 1930s by Archibald Hill.

After this, various models have been proposed using approaches of Boolean algebra. In the 1940s, McCulloch and Pitts investigated networks consisting of *formal neurons*. These are elements performing basic logical functions, resembling in some instances nerve cells with synaptic connections. In the late 1950s, John von Neumann introduced his theory of self-organization in cellular automata, suggesting the intimate relationship between computers and the brain. This basic model has been modified in subsequent years. So, for example, the refractory period of nerves was included, which means that, for some milliseconds after an input signal, the nerve is not able to accept a further one. These concepts of neuronal networks mostly support the development of new technical systems of artificial intelligence and intelligent computing.

The central problem of these models, however, consists in the requirement of their synchronous reactions. They require a synchronization of all impulses in the network, i.e. the introduction of a certain sequence of operations. Despite the fact that in some particular brain neurons a certain sequence of firing exists, for example in cortical cell columns, there is no evidence of an overall synchronization in the brain. In contrast, it is necessary to take into account the random times of arrival of the synaptic inputs.

In the 1960s, a number of approaches using these digital elements led to models resembling various kinds of behavior and sensory perception. This includes some

systems of feedback regulation as explained in Sect. 5.1.4. As already mentioned, these models at that time were subsumed under the term *biocybernetics*.

In general, neuronal and brain models according to their approaches can be classified as realistic and as simplified ones. *Realistic brain models* start with the entire process of nerve excitation, for example with the Hodgkin−Huxley model as explained in detail in Sect. 3.4.4. In this case, one tries to include as much of molecular and cellular details as are available. The extremely large number of parameters of these models makes it impossible for them to be analyzed analytically. They can be handled only using numerical simulations. Furthermore, the large number of parameters and their variability make the results of such computations sometimes rather ambiguous. The large amount of detail contained in these models can often obscure rather than illuminate the problem. One way of overcoming this dilemma is the more or less rigorous reduction of the Hodgkin−Huxley equations to some essential properties.

As a consequence, this finally leads to the *simplified brain models*. In exchange for analytical tractability, they do abstract from the complexity of individual neurons and the patterns of connectivity. These simplifying models, of course, are essential, but to some extent they are also dangerously seductive because they can become an end in themselves and lose touch with reality.

Some of these models consider *spiking* neurons, reacting simply as digital elements, in contrast to *firing-rates* models, where the code is based on the frequency of the firing rate. Such frequency coded signals in fact occur in most sense organs. In some cases, however, the reaction times are too short to allow long temporal analysis.

The *integrate-and-fire model* (Fig. 5.20) has become widely accepted as one of the canonical models for the study of neural systems. It can be considered as a combination of the earliest electronic trigger-circuit models with the approaches of Boolean algebra. This model considers the state of the neuron as a function of its membrane potential [$\psi(t)$]. This is controlled by stochastic excitatory or inhibitory

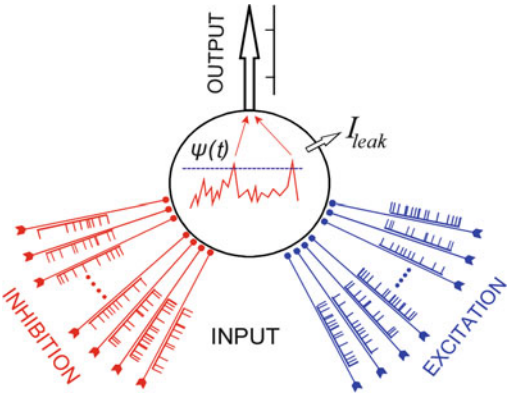

Fig. 5.20 Schematic representation of an integrate-and-fire model. The membrane potential $\psi(t)$ of the neuron is decreased by stochastic inhibitory (*red*) and increased by excitatory (*blue*) synaptic inputs. Furthermore, a leak current (I_{leak}) (*small black arrow*) permanently degrades the membrane potential with a time constant (τ_m). If the membrane potential arrives a critical value (*dashed line*), output spikes (*black*) are generated

synaptic inputs. These inputs are modeled either as injected currents (current synapses) or as a change in the membrane conductance (conductance synapses). Furthermore, due to ion leakage, the membrane potential permanently decays with a characteristic time constant (τ_m). When the membrane potential reaches a definite maximal threshold, the neuron generates an output spike.

With respect to the membrane capacity (C_m), this process can be characterized as follows:

$$C_m \frac{\mathrm{d}\psi(t)}{\mathrm{d}t} = I_{leak}(t) + I_{syn}(t) \tag{5.45}$$

whereas the current (I_{syn}) characterizes the input current from the synapses, and (I_{leak}) stands for the leak current which results from:

$$I_{leak} = -\frac{C_m}{\tau_m}[\psi(t) - \psi_o] \tag{5.46}$$

This includes the time constant (τ_m) which results from product of the capacity and resistance of the membrane as explained in Sect. 3.5.4 (Eq. 3.214). ψ_0 indicates the resting potential of the neuron.

The crucial point of the integrate-and-fire model, and the source of many variations of it, is the formulation of the term (I_{syn}). This is an integral expression which includes the weighted amount of stochastic negative, or positive pulses, depending on the inhibitory or excitatory character of the synapses. In the case of *current synapses*, it is the linear sum of weighted current pulses of various strengths, durations, and polarity. In the case of *conductance synapses*, a nonlinear relationship occurs because the currents, produced by individual pulses, depend on the degree of the corresponding conductance variations, and additionally on the actual membrane potential. The arrival time of the synaptic input pulses is generally modeled as a time-homogeneous stochastic process based on Poisson distribution. Under certain assumptions, it is possible to calculate the output spiking rate of the neuron for this case.

To understand processes coded by the firing rate of pulses, the integrate-and-fire model was extended to inhomogeneous and especially to periodic synaptic input. This also includes the mechanism of stochastic resonance which is important to analyze periodic processes, superimposed by stochastic noise (see Sect. 3.1.5). This, for example, allowed the analyses of auditory neural processing, where neural responses show activity that is phase-locked to periodic acoustical signals.

The integrate-and-fire neuron model appears as a very useful tool to understand a number of neurological observations, how neural systems function, and how they process information. On the one hand, it is sufficiently simple to allow mathematical analysis, while on the other hand, it contains experimentally measurable parameters, such as the firing threshold of the neuron, its resting potential, functional spikes, the response of the membrane potential to stochastic synaptic inputs, etc. Of course, there are a number of problems associated with the simplifying

assumptions. Nevertheless, it represents a balance between conceptual simplicity and biological accuracy.

Beside these models, which at least contain basic neurobiological facts, there are some others which in fact only show some formally analog phenomena. A frequently used example is the *spin-glass model*. This considers the cooperative behavior of molecules in amorphous materials, theoretically treated by approaches of statistical thermodynamics.

Spin-glasses are viscous materials containing particles of different magnetic properties. These include ferromagnetic particles, orienting themselves by mutual influences parallel to each other, as well as anti-ferromagnetic particles, orienting anti-parallel (Fig. 5.21). The analogy between spin-glasses and neuronal networks is the direction of the particle and the state of the neuron as well as their mutual interactions. A parallel or anti-parallel orientation of a particle in the spin-glass would correspond to a "silent" or an "excited" cell in a neural network. The ferromagnetic, i.e. anti-ferromagnetic interaction of the particles, simulates the excitation or inhibitory interaction of the nerve cells.

In this respect, spin-glasses may show a quite complicated pattern of self organization. Of particular interest are the situations where two opposite kinds of interaction come together in one point. In this case, a particle will be influenced at the same time by two forces tending to orientate it in opposite ways. This disturbs the stability of the whole system. This situation is called a *frustrated* state. It can be considered as a sort of metastability and could reflect a kind of memory state of a neuronal network. It is a special kind of phase transition which cannot be described like a catastrophe as depicted in Fig. 5.6, but instead shows a more continuous kind of transition. These spin-glass models, therefore, resemble not only the problem of self-organization but also directly the function of neuronal networks.

The development of new analytical techniques, combined with the sharp increase in computational power, have led to formidable progress in neuroscience in recent decades. This, however, should not mask the fact that the brain is a multilayered functional system, and the path from the action potential of a single neuron up to complex neurological processes such as thinking or even consciousness cannot be modeled by a single mathematical approach. It is the merit of the

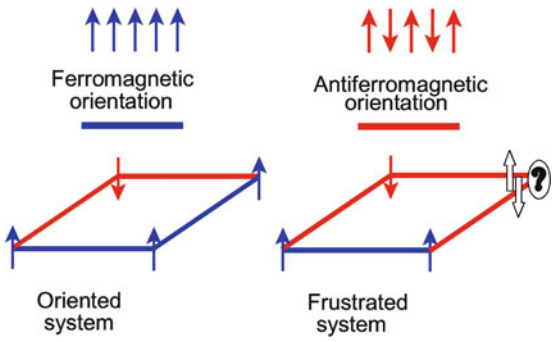

Fig. 5.21 Ferromagnetic and antiferromagnetic interactions in spin-glasses and the occurrence of frustrated systems

research of the last 50 years to realize that there is a principal difference between the data processing in brain versus that in the recent kinds of digital computers.

Further Reading

Banerjee and Chakrabarti 2008; Burkitt 2006; Cessac and Samuelides 2007; Faugeras et al. 2009; Hanrahan 2011; Prettejohn et al. 2011.