We note that instead of the requirement $M_{jk} = 1$ as in case of a single final attractor for an incident message, M_{jk} can be generalized to

$$0 < M_{jk} \leq 1 \ . \tag{1.13}$$

The form (1.8), left hand side, immediately allows us to write down the formulas for several systems that are coupled one after the other. For instance in the two step process we immediately obtain

$$p_j = \sum_k L_{jk}^{(1)} p'_k = \sum_{kk'} L_{jk}^{(1)} L_{kk'}^{(2)} p''_{k'} \tag{1.14}$$

where one can convince oneself very easily that $\sum_j p_j = 1$ provided $\sum_k p'_k = 1$ and $\sum_j L_{jk} = 1$. The individual steps read

$$\sum_j p_j = \sum_j \sum_{kk'} L_{jk}^{(1)} L_{kk'}^{(2)} p''_{k'} = \sum_{kk'} \underbrace{\left(\sum_j L_{jk}^{(1)} \right)}_{=1} L_{kk'}^{(2)} p''_{k'} \tag{1.15}$$

$$= \sum_{k'} \underbrace{\sum_k L_{kk'}^{(2)}}_{=1} p''_{k'} = 1 \ . \tag{1.16}$$

We may define

$$L'_{jk'} = \sum_k L_{jk}^{(1)} L_{kk'}^{(2)} \ . \tag{1.17}$$

Because the L's are positive we find

$$L'_{jk} \geq 0 \tag{1.18}$$

and because of the normalization properties (in case of no information deficiency)

$$\sum_j L'_{jk'} = \sum_k \sum_j L_{jk}^{(1)} L_{kk'}^{(2)}$$

$$= \sum_k L_{kk'}^{(2)} = 1 \tag{1.19}$$

we readily obtain

$$L'_{jk} \leq 1 \tag{1.20}$$

so that L'_{jk} obeys the inequality

$$0 \leq L'_{jk} \leq 1 \ . \tag{1.21}$$

We mention that the recursion from p'' or still higher order $p^{(n)}$ to p may depend on the paths.

Our above approach not only introduces the new concept of relative importance of a message but it also provides us with an algorithm to determine p_j which has some conceptual and practical consequences. With a given task or ensemble of tasks, this algorithm allows us to select the message to be sent, namely the one with the biggest p_j. If there are several p_j of the same size it does not matter which message is sent. From the conceptual point of view we may then decide whether a dynamical system annihilates, conserves or generates information. To this end we make use of the concept of information in the sense of conventional information theory. But instead of the information content due to the relative frequency of symbols we use the relative importance within a set of messages, i.e. we introduce the quantities

$$S^{(0)} = -\sum_j p_j \ln p_j \tag{1.22}$$

$$S^{(1)} = -\sum_k p'_k \ln p'_k \;, \tag{1.23}$$

where p_j and p'_k have been defined above in the text. If $\sum_k p'_k = 1$, as is always assumed here, and $\sum_j p_j < 1$, an information deficiency is present. In the case $\sum_j p_j = 1$ we shall speak of annihilation of information if

$$S^{(1)} < S^{(0)} \tag{1.24}$$

of conservation of information if

$$S^{(1)} = S^{(0)} \tag{1.25}$$

and of generation of information if

$$S^{(1)} > S^{(0)} \;. \tag{1.26}$$

The meaning of this definition quickly becomes clear when we treat special cases. If, for instance, two messages lead to the same attractor there is a redundancy in the system and the information content (in the traditional technical sense of the word) becomes smaller. It is reduced from

$$S^{(0)} = -K[\tfrac{1}{2}\ln(\tfrac{1}{2}) + \tfrac{1}{2}\ln(\tfrac{1}{2})] = K \ln 2 \tag{1.27}$$

to

$$S^{(1)} = -K \cdot 1 \cdot \ln 1 = 0 \;. \tag{1.28}$$

In the case of a one-to-one mapping of p_j onto p'_k we find the transfer of $\{p_j\}$ into the same set $\{p'_k\}$, except maybe for the permutation of indices, i.e. for different numbering of the states. In such a case (1.25) clearly holds. Finally, in the case (1.26), the p_j, where one $p_j = 1$ and all others $= 0$, are transferred into e.g. $p' = p'' = \frac{1}{2}$ and all others are equal to 0. Then $S^{(0)} = -K \cdot 1 \cdot \ln 1 = 0$ is enlarged to

$$S^{(1)} = -K[\tfrac{1}{2}\ln(\tfrac{1}{2}) + \tfrac{1}{2}\ln(\tfrac{1}{2})] = K \ln 2 \;. \tag{1.29}$$

Of course these examples are not meant to prove the definitions (1.24–26) but rather to illustrate their meaning.

Our approach based on synergetics has some further nice features. Semantics has become the problem of studying the response (attractors) of the dynamic system. The system may be error-correcting (or may supplement partial information). If the incident message does not set the initial state $\boldsymbol{q}$ *on the attractor* (i.e. *not correctly*), it may set the initial state $\boldsymbol{q}$ within the *basin of the attractor* i.e. on the slope of the hill surrounding the bottom of the specific valley which represents the attractor (fixed point). In this way the system pulls the state vector into the attractor corresponding to that basin, i.e. into the *correct* state. It will be an interesting problem to determine the minimum number of bits required to realize a given attractor (or to realize a given value of "relative importance").

Within our present scheme, the learning process of a system can also be modeled. A system can be "sensitized" or "desensitized" with respect to messages j e.g. by letting more or fewer parameters react to specific messages.

In the above treatment we have assumed that the value of the messages is measured with respect to the *same* initial state of the receiver. In the next step of our considerations we may assume that messages apply to a receiver in *another* initial state which has been set for instance by a previous message.

In such a way we obtain an interference of messages and the relative importance of a message depends on the messages previously delivered to the receiver. In the general case, the relative importance of a message will depend in a non-commuting way on the sequence of the messages. In this way the receiver is transformed by messages again and again and clearly the relative importance of messages will become a function of time.

Another remark might be useful, particularly in relation to synergetic processes. A synergetic system not only needs to be a dynamical system showing e.g. limit cycle or chaotic behavior, but it might also be one in which irreversible processes leading for instance from a disorganized liquid state into a structured solid state occur.

Let us conclude this part with a comment on pattern recognition which will be eludicated from various points of view in this book. Pattern recognition can be considered as a processing of incoming messages by a receiver, e.g. the brain or a machine. It is therefore an interesting task to discuss pattern recognition using the ideas just outlined. I suggest that pattern recognition, at least in general, is a multistep process in which the receiver takes an active part. In the first step, the pattern is received at a global level where, in general, several attractors can be reached. Then, the sensory system is requested to focus its attention on the exploration of additional features so that a finer set of attractors can be selected. To be more explicit: In the first step for instance the global shape of the contour lines of an object are determined e.g. close to a circle, rectangular, etc. Then, in the case of a circle, there are several attractors: apple, face, wheel, tree. Then the receiver asks back for further details, e.g. color, vertical lines (nose?), etc. In this way the process can be continued.

Note that our interpretation of pattern recognition differs from the "traditional" approach to which we shall come in Chap. 12. There the pattern is first decomposed into its "primitives" or "features". Here we start from the *global pattern* (contour line) and then proceed to more and more details.

This approach offers us an explanation (or at least a hint) as to why, in human pattern recognition, even interrupted contour lines are supplemented such that a continuous line is "seen".

1.6.3 Self-Creation of Meaning

As was mentioned previously, synergetics may be considered as a theory of the emergence of new qualities at a macroscopic level. By means of a suitable interpretation of the results of synergetics, we may thus study the *emergence of meaning* as the emergence of a new quality of a system, or in other words the *self-creation of meaning*. In order to study how this happens we want to compare a physical system, namely the laser, with several model systems of biology. Let us start with some general remarks on the role of information in biological systems.

One of the most striking features of any biological system is the enormous degree of coordination among its individual parts. In a cell, thousands of metabolic processes may go on at the same time in a well-regulated fashion. In animals, millions to billions of neurons and muscle cells cooperate to bring about well-coordinated locomotion, heartbeat, breathing or blood flow. Recognition is a highly cooperative process, and so are speech and thought in humans. Quite clearly, all these well-coordinated, coherent processes become possible only through the exchange of information, which must be produced, transmitted, received, processed, transformed into new forms of information, communicated between different parts of the system and at the same time, as we shall see, between different hierarchical levels. We are thus led to the conclusion that information is a crucial element of the very existence of life.

The concept of information is a rather subtle one and it will be the goal of this section to further elucidate some of its aspects. As we shall see, information is linked not only with channel capacity or with orders given from a central controller to individual parts of a system – it can acquire also the role of a "medium" to whose existence the individual parts of a system contribute and from which they obtain specific information on how to behave in a coherent, cooperative fashion. At this level, semantics may come in.

Let us first have a look at physics. In closed systems the second law of thermodynamics tells us that structures decay and systems become more and more homogeneous, at least on a macroscopic level. At the microscopic level complete chaos may occur. For these reasons information cannot be generated by systems in thermal equilibrium; in closed systems thermal equilibrium is eventually reached. But a system in thermal equilibrium cannot even *store* information. Let us consider a typical example, namely a book. At first sight, it seems to be in thermal equilibrium, and indeed we can measure its temperature. But in spite of that, it has not reached its final state of complete thermal equilibrium. In the course of time, the printer's ink in the individual letters will diffuse away until a homogeneous state is reached.

This simple example teaches us that any memory consisting of a closed system is out of thermal equilibrium and it is always necessary to ask *how long* the

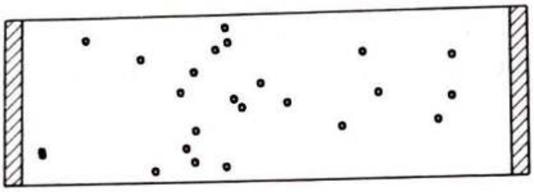

Fig. 1.28. Laser active atoms embedded in a crystal of a laser setup

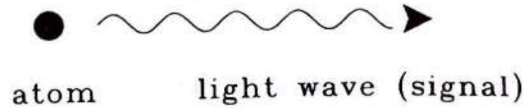

Fig. 1.29. An excited atom emits a light wave (signal)

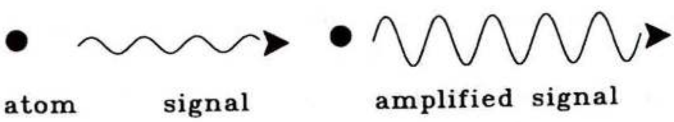

Fig. 1.30. When the light wave hits an excited atom it may cause the atom to amplify the original light wave

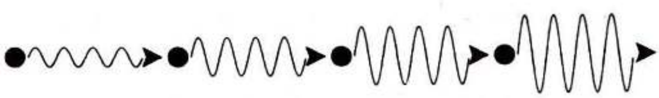

Fig. 1.31. A cascade of amplifying processes

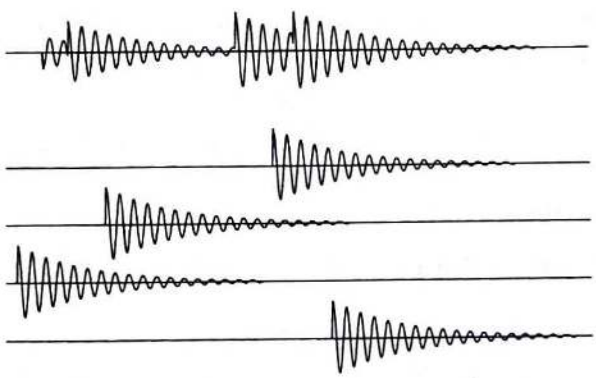

Fig. 1.32. The incoherent superposition of amplified light waves produces a still rather irregular light emission

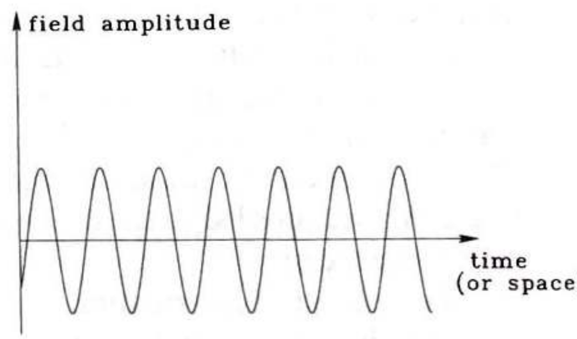

Fig. 1.33. In the laser the field amplitude is represented by a sinusoidal wave with a practically stable amplitude and only small phase fluctuations

information can be stored in each specific case. Let us therefore consider open systems which are kept far from thermal equilibrium by an influx of energy and/or matter into the system. As was mentioned before, in open systems, even in the inanimate world, specific spatial or temporal structures can be generated in a self-organized fashion. Examples are provided by the laser which produces coherent light, by fluids which can form specific spatial or temporal patterns, or by chemical reactions which can show continuous oscillations, or spatial spirals, or concentric waves. Even at this level we can speak to some extent of creation or storage of information. On the other hand, we can hardly attribute words like relevance, purpose or meaning to these processes.

Let us discuss the laser in some more detail because it allows us to introduce a terminology which is also useful for biological and other systems. In the laser a number of atoms are embedded, for instance, in a crystal such as ruby (Fig.1.28). After excitation from the outside, these atoms may emit individual light wave trains

(Fig.1.29). Thus, each atom emits a signal, i.e. it creates information which is carried by the light field. In the laser cavity the emitted wave trains may hit another excited atom and cause it to amplify the original wave (Fig.1.30). In this way, the information serves the purpose of enhancing the signal (Fig.1.31). Because the individual excited atoms may emit light waves indepently of each other and these may then be amplified by other excited atoms, a superposition of uncorrelated, though amplified wave trains results and a quite irregular pattern is observed (Fig.1.32).

But when the signal reaches a sufficiently high amplitude, an entirely new process starts. The atoms begin to oscillate coherently and the field itself becomes coherent, i.e. it is no longer composed of individual uncorrelated wave tracks but has become a practically infinitely long sinusoidal wave (Fig.1.33).

We have here a typical example of self-organization where the temporal structure of the coherent wave emerges without interference from the outside. Order is established. The detailed mathematical theory shows that the emerging coherent light wave serves as order parameter which forces the atoms to oscillate coherently, or in other words it enslaves the atoms (Fig.1.34). Note that we are dealing here with circular causality: On the one hand the order parameter enslaves the atoms, but on the other hand it is itself generated by the joint action of the atoms (Fig.1.35).

From the viewpoint of information, the order parameter serves a double role: it informs the atoms how to behave, and in addition, it informs the observer about the macroscopic ordered state of the system. While an enormous amount of information is needed to describe the states of the individual atoms, once the ordered state is established, only a single quantity, namely the phase of the total light field is necessary, i.e. we have an enormous compression of information. We may call the

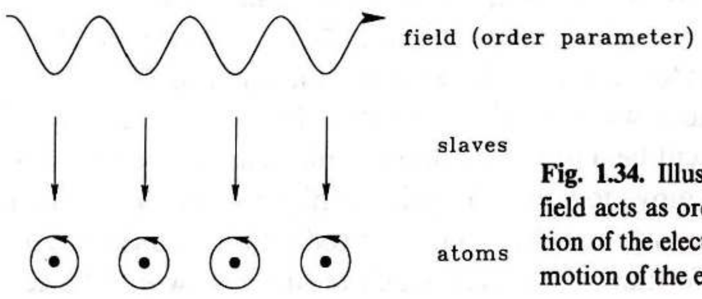

Fig. 1.34. Illustration of the slaving principle. The field acts as order parameter and prescribes the motion of the electrons in the atoms. In other words the motion of the electrons becomes slave to the field

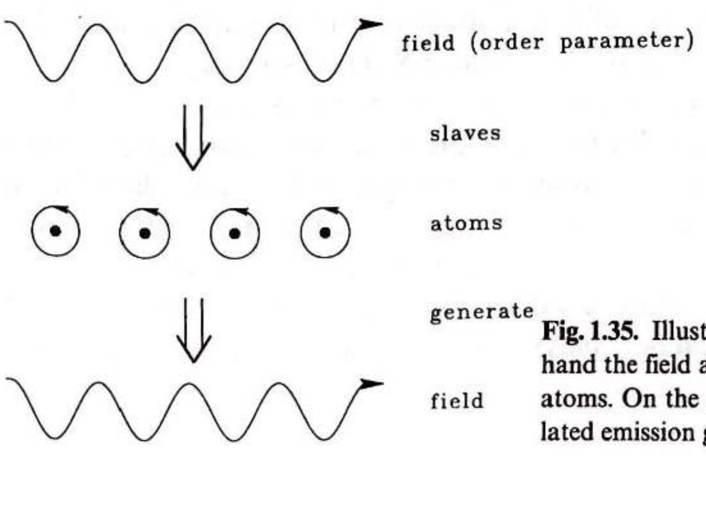

Fig. 1.35. Illustration of circular causality. On the one hand the field acting as order parameter enslaves the atoms. On the other hand the atoms by their stimulated emission generate the field

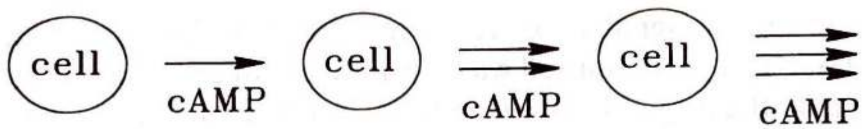

Fig. 1.36. Illustration of the amplification of the number of cAMP molecules in the cells of slime mold

Fig. 1.37. Schematic illustration of the concentration wave of cAMP in slime mold formation

order parameter an "informator". Over the past years, it has been shown that these concepts apply to a large number of quite different physical, chemical and biological systems.

To elucidate the role of information exchange at the level we are presently considering, let us take the example of slime mold (*dictiostelium discoideum*). Usually its cells live individually on a substrate but when food becomes scarce they assemble at a particular point. The mechanism of this kind of self-assembly is as follows:

The individual cells start to emit the substance, cyclic Adenosinemonophosphate (cAMP); thus they send out a signal or a message, i.e. information. Once cAMP molecules hit other cells, these are induced to increase their production in much the same way as the laser atoms amplify the incoming signal (Fig.1.36). Quite clearly, the elements themselves are not aware of the meaning of the information but through the interplay between emission, amplification and diffusion of the cAMP molecules, a spiral pattern of concentration of cAMP is formed, i.e. information at a higher level is generated (Fig.1.37). Because this information is produced by the cooperativity of the system, we may call it *synergetic information*. The spiral waves form some kind of gradient field (the informator) which can be measured by the individual cells which then move towards the point of highest concentration in the field. Clearly we can distinguish here between the production of information, the information carrier and information receiver which in our case would be cell, cAMP and cell, respectively. However at the next level, we observe that a new meaning has arisen, namely the established pattern of a molecular concentration serves the purpose of guiding the cells to the center of their assembly.

Basically the same idea holds for the concept of positional information. Here, it is assumed that the individual cell within a tissue receives its information from a chemical field which has been established by the production and diffusion of chemicals. In general, two kinds of molecules are assumed, namely activator and inhibitor molecules. Where activator molecules have a high concentration it is assumed that specific genes can be switched on which then cause the differentiation of a cell. In this way the chemical field plays the role of the informator. A particular model system has been hydra.

It is useful to recall what we have established so far. Quite evidently, there is a hierarchy of informational levels. At the lowest level, the individual parts can emit information which hits other parts of the system. Such an information transfer can take place between specific pairs of elements or the information can be transferred by a general carrier. An example for the first case are nerve fibers each connecting two neurons; examples of the second case are provided by hormones released to the blood, or by pheromones released into air.

Although, in all these cases the exchange of information may initially occur at random, a competition or cooperation between different kinds of signals sets in, and eventually a new collective state is reached which differs qualitatively from the disordered or uncorrelated state present before. Thus, a new state is described by an order parameter or a set of order parameters or equivalently by one or several informators. The states of the individual parts are determined by means of the slaving principle. But one may describe this process in another way, namely that a specific consensus was reached among the individual parts of the system or that self-organization has happened. At the same time information compression takes place. The information appears manifest at a macroscopic level and, in many cases, increases the reliability and/or efficiency of the system, or serves other purposes as mentioned above.

This new collective level becomes observable to the outer world and by establishing this relationship or context a new semantic level is reached. By the way, the context may be established with the outer world but equally well within the same system. Here then words like useful, useless, or relevance can be applied. This is quite evident from the example of the laser where the cooperative state reaches a high efficiency. In the analogous case of a biological system, such behavior is then useful for the whole system. Beyond instability points the system can acquire different possible states and it needs additional information on which state to choose. One possibility is that this information is provided genetically, or by constraints established by other parts of the system. But often in such a case of degeneracy, the surroundings play an important role, or in other words, it is the context which judges the value of the kind of state to be established. In the opinion of the author it is here that information in the biological sense starts. Through instability a collective state is formed but it acquires its meaning only with respect to the surroundings and, in a way, with respect to its value for the survival of the whole system.

These remarks also apply to the genetic code, though its very origin is not yet too well clarified. One may speculate that at first fluctuations occur which create some biological macromolecule with specific properties. The most important of these is that it can multiply in an autocatalytic fashion. The value of the information conveyed by this molecule to its phenotype is then judged by the environment to which other molecules with their phenotypes may also belong. By the interplay of mutation and selection new types of molecules and their corresponding phenotypes are then generated and in this way we observe the creation of new information. But whether this information is useful or not can be checked upon only by the interaction of the particular species with its environment.

In the considerations above we described the first steps of the formation of ordered or structured collective states. But in contrast to the physical systems mentioned above, such as lasers, fluid dynamics, or chemical reactions, a new feature appears in biology, namely a solidification. For instance, when the genes of a cell are switched on by activator molecules, the cell differentiates into a specific cell which now is no longer modifiable or can no longer be transformed back into the original cell. In a way dynamical processes may lead to solid structures like bones or organs. In a similar way, information is laid down in a rigid manner in DNA, i.e. in the genetic code. It appears that lower animals are constructed more or less by the rules given by the genetic code with a rather rigid "wiring" of their nervous systems.

On the other hand in higher animals, in addition to rigid wiring of the nervous system a good deal of self-organization appears. The interaction of the system with its environment, together with the genetic information laid down in the system leads to the formation of new information. Through the continuous testing of the new information stored and created in the brain by the environment, new contexts are established and thus a new kind of semantics occur. But we may also expect that "solidification" occurs at various hierarchical levels of semantic information and serves for making the system more reliable, and to store information (memory). While the concept of Hebb's synapse, one which is strengthened by its use, may be a correct concept, the building up of semantics requires a high degree of cooperativity within the system and a repeated interaction with the outside world. In this respect, semantic information is not a static property, but rather a process in which contexts and relevance are checked, reinforced or dismissed again and again. By the way, I believe that consciousness is not a static state, but a process in which information is continuously transferred between various parts of the brain and repeatedly processed there.

At this point a word on pattern recognition may be in order. Lower animals immediately react to stimuli such as light flashes and only few criteria are needed, such as theshold of intensity, in order to respond to a signal. In higher animals, however, the incoming information will certainly be compared with stored information. However, our picture of how this comparison is done is changing slightly.

Quite often it is assumed that the incoming pattern is compared with templates. However, the storage of a template would require quite a large amount of information. Therefore, one might imagine in the sense of synergetics, that only specific characteristic features are stored in the form of order parameters which may then be called upon to generate a detailed picture. In this sense then, pattern recognition becomes an active process in which new patterns are formed in a self-organized fashion by the brain which, using certain hypotheses, checks them repeatedly against the incoming patterns. For instance, it is well known that when people look at faces, they focus their attention on specific parts like eyes, or nose, or mouth and look at them again and again.

Let us finally discuss a point which applies specifically to humans. In contrast to animals, human beings can transfer information not only by the genetic code, but also by teaching which in the world of animals takes place only in a very limited way. So a good deal of our culture is based on this new way of transferring information from one generation to the next. But here an enourmous difficulty arises because of the tremendous amount of knowledge which has been accumulated by humanity. Therefore, quite in the spirit of synergetics, it will be important to find unifying ideas and principles to cope with this large amount of information.

In addition our approach provides us with a picture rather different from those conventionally drawn from biological systems. There, it is assumed that there exists one single command center, say in the brain which then organizes all the behavior. The model that we are strongly supporting calls rather for processes of self-organization, and more recently we were able to prove this hypothesis by our quantitative theory of specific experiments on the correlation of hand movements and their changes. In these experiments, performed by S. Kelso, test persons were asked to oscillate their fingers in parallel. At an increased oscillation frequency an involuntary change to an antiparallel oscillation occurred. The way in which this transition occurs can be represented in all its details by the assumption of self-organization of the behavior of neurons and muscles.

This is certainly an extreme case and in general the information production and transfer in biological systems must be considered in two ways: the one is the conventional one in which specific motor programs serve for specific actions, whereas other phenomena occur in an entirely self-organized fashion. We may hypothesize that self-organization in information processing in biological systems plays a widespread and major role. This is borne out by the great flexibility of biological systems and their adaptability and plasticity.

In my opinion, the study of information in biological systems is also of interest to modern society whose proper functioning relies on the adequate production, transfer, and processing of information. Perhaps the most important aspect which has emerged is that of circular causality which results in a collective state which in sociology may represent a social climate, a general public opinion, a democracy or a dictatorship.

1.6.4 How Much Information Do We Need to Maintain an Ordered State?

Let us consider our standard example, namely the laser. Let us assume that there are atoms in the laser each having two levels. The total number of atoms in the lower state will be denoted by N_1, the number of atoms in the upper state by N_2. We have the relation

$$N_1 + N_2 = N \ . \tag{1.30}$$

In the sense of quantum mechanics we may relate the occupation numbers N_1 and N_2 to the occupation probability

$$p_j = \frac{N_j}{N} \ , \qquad j = 1, 2 \tag{1.31}$$

for a single atom. Thus, the information per atom is given by

$$i = -p_1 \ln p_1 - p_2 \ln p_2 \tag{1.32}$$

and then for all atoms by

$$I = -N(p_1 \ln p_1 + p_2 \ln p_2) \ . \tag{1.33}$$

As we know, an exited atom may emit a photon either by spontaneous emission, or if other photons are already present, by so-called stimulated emission. We may identify a single photon with a symbol that carries an element of information. As we know, photons can escape through the mirrors. Therefore, we may ask the question of what production rate of photons is necessary in order to maintain a coherent state?

According to laser theory we must not only introduce the number of photons n as a variable, but in addition the inversion which is defined as the difference between the occupation numbers of the upper and lower state:

$$D = N_2 - N_1 \ . \tag{1.34}$$

According to the theory the production rate of photons is given by the equation

$$\frac{dn}{dt} = WDn - 2\kappa n \ . \tag{1.35}$$

The first term on the right-hand side describes the production rate of photons, where W is a rate constant for this production, whereas the second term describes the escape of photons through the mirrors so that the actual production is diminished. As is shown in laser theory, (1.35) describes the production of coherent photons; the production of incoherent photons is neglected. Because of the laser process the inversion also changes in time. Its rate of change is given by the equation

$$\frac{dD}{dt} = \frac{D_0 - D}{T} - 2WDn \ . \tag{1.36}$$

Here D_0 is the inversion produced by the pump process and relaxation processes which do not give rise to laser light emission. T is the time in which any deviation of the inversion relaxes towards the inversion D_0. The last term in (1.36) stems from the laser process in which photons are produced. In general, the decay constant κ is much smaller than the rate constant $1/T$. This allows us to apply the so-called adiabatic approximation in which we may write

$$\frac{dD}{dt} \approx 0 \ . \tag{1.37}$$

Using (1.37) in (1.36) we can immediately solve (1.36) for D thus obtaining

$$D = \frac{D_0}{1 + 2TWn} \ . \tag{1.38}$$

When the laser is not too far above the onset of laser action, we may expand the denominator as a power series in the photon number n so that in the leading approximation we obtain

$$D \approx d_0 - 2D_0 TWn \ . \tag{1.39}$$

Inserting this result into the equation for the production rate of photons (1.35), we readily obtain

$$\frac{dn}{dt} = (WD_0 - 2\kappa)n - 2TW^2 D_0 n^2 \ . \tag{1.40}$$

While the second term in (1.40) will always lead to a decrease in the production rate of the photons, the first term will give rise to a positive production rate provided the inequality

$$WD_0 - 2\kappa > 0 \tag{1.41}$$

holds. Equation (1.41) is identical with the laser condition. Thus (1.41) guarantees a positive net production rate of photons, or in other words, a *positive net production rate of signals*. This is necessary for the maintainance of a nonzero flux of photons and thus for the ordered state of the laser. According to (1.41) this can be established only if the inversion D_0 which is achieved by pumping is sufficiently high. From

$$n = \frac{WD_0 - 2\kappa}{2TW^2 D_0} \tag{1.42}$$

we may deduce that the condition (1.41) guarantees a non-vanishing number of photons which must be present in the laser at all times.

Let us now study the behavior of the information (1.32) or (1.33) when we increase the pump rate or in other words the inversion D_0. To this end we insert (1.42) in (1.39) and obtain

$$D \approx \frac{2\kappa}{W} = \text{const.} \ ! \qquad \text{for } n \geq 0 \text{ or} \tag{1.43}$$

$$D = D_0 \tag{1.44}$$

for $n = 0$. In other words when we start from a low pump rate, D_0 is small and no photons are present. Then D increases at the same rate as D_0. But once laser action sets in, the inversion D becomes a constant according to (1.43) and shows no further increase. All the additional energy fed into the laser is transformed into coherent photons. Using (1.34) and (1.30) we have the relations

$$N_1 = \tfrac{1}{2}(N - D) \tag{1.45}$$

$$N_2 = \tfrac{1}{2}(N + D) \tag{1.46}$$

which can then be transformed according to (1.31) into the occupation probabilities

$$p_1 = \frac{1}{2}\left(1 - \frac{D}{N}\right) \tag{1.47}$$

$$p_2 = \frac{1}{2}\left(1 + \frac{D}{N}\right) . \tag{1.48}$$

Inserting (1.47) and (1.48) into (1.32) we obtain

$$i = -\frac{1}{2}\left(1 - \frac{D}{N}\right)\ln\frac{1}{2}\left(1 - \frac{D}{N}\right) - \frac{1}{2}\left(1 + \frac{D}{N}\right)\ln\frac{1}{2}\left(1 + \frac{D}{N}\right) . \tag{1.49}$$

In the following we shall use the parameter γ defined by

$$\gamma = \frac{D}{N} . \tag{1.50}$$

Because D takes values in the range

$$D: -N, \dots, +N , \tag{1.51}$$

γ must lie in the range

$$\gamma: -1 \dots +1 . \tag{1.52}$$

The behavior of i as a function of γ is plotted in Fig.1.38. The result (1.49) jointly with (1.43) and (1.44) allows us to study the change in the information of an individual atom when we increase the pump rate D_0. According to Fig. 1.39 the information first rises, goes through a maximum and then saturates. Our present approach is not capable of dealing with the information contained in the light field because so far we have not considered any fluctuations, i.e. any probability distribution over the photon numbers n. One of the main objectives of our book will it be to study the information as a function of pump strength not only for the atoms but also for the photons.

Indeed we shall see that a surprising result is obtained, namely that the interesting information close to the point where laser action starts is contained in the photons rather than in the atoms.

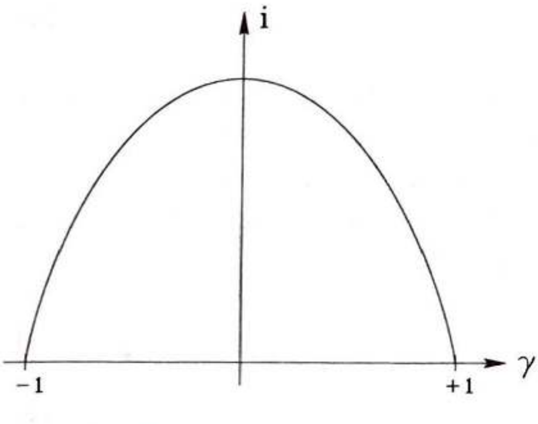

Fig. 1.38. The information of an atom versus the parameter γ

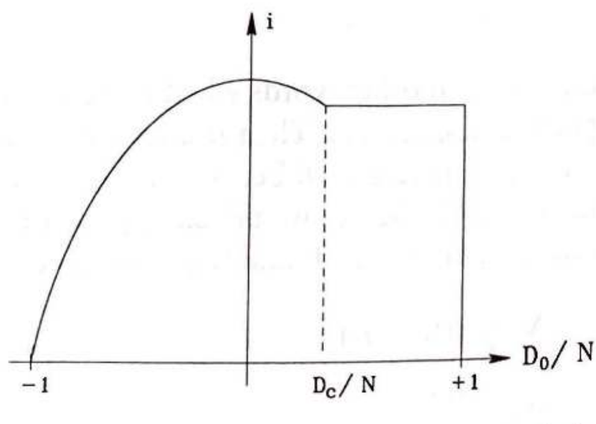

Fig. 1.39. The behavior of information of a single atom in the laser

1.7 The Second Foundation of Synergetics

After having discussed the qualitative aspects of information and self-organization in the previous sections, we now wish to come to the hard core of our approach which will then be followed up in the remainder of this book. Let us briefly recall what we have been doing in the field of synergetics so far. There, we started from the microscopic or mesoscopic level at which we formulated equations. Then, by using the concepts of instability, order parameters, and slaving, which can be cast into a rigorous mathematical form, we could show the emergence of structures and, concomitantly, of new qualities at a macroscopic level.

In a way parallels can be drawn between the latter approach and that of statistical mechanics. We wish now to develop an approach which can be put in analogy with that of thermodynamics. Namely, we wish to treat complex systems by means of macroscopically observed quantities. Then we shall try to guess the microscopic structure of the processes which give rise to the macroscopic structure or the macroscopic behavior. The vehicle we shall use for this purpose is the maximum entropy principle, or the maximum information entropy principle which was developed quite generally by Jaynes.

We shall give a detailed presentation of this principle in Chap. 3. Here it will suffice to summarize the basic idea. We start from macro-observables which may fluctuate and whose mean values are known. We distinguish the macro-variables by an index k and denote their mean values by f_k. We wish then to make a guess at the probability distribution p_j of the system over states labeled by the index j. This is achieved under the maximization of the information

$$i = -\sum_j p_j \ln p_j \tag{1.53}$$

under the constraint that

$$\sum_j p_j f_j^{(k)} = f_k \ . \tag{1.54}$$

Evidently, $f_j^{(k)}$ is the contribution of state j to the macro-variable labeled by k. Furthermore we require

$$\sum_j p_j = 1 \ , \tag{1.55}$$

i.e. that the probability distribution p_j is normalized to unity. As was shown by Jaynes and as will be demonstrated in Chap. 4, this principle allows us to derive the basic formulas of thermodynamics in a very short and elegant fashion. For this derivation the constraints refer to the conserved quantities of a closed system, i.e. energy, particle numbers etc. The crux of the problem of extending this maximum entropy principle to systems far from thermal equilibrium or even to non-physical systems lies in the adequate choice of constraints.

As we shall see, the constraints which have been used so far, of energy conservation or even of regulated energy fluxes into the system, are inadequate to treat open systems and especially to treat the transition from a structureless to a structured state as occurs in non-equilibrium phase transitions. The maximum entropy principle has been criticized occasionally because the choice of the constraints seems to introduce a certain subjectivity in that the constraints are said to be chosen arbitrarily at the will of the observer rather than by objective criteria.

This criticism has been debated by Jaynes in detail, but I should like to add here another point of view. Scientific progress relies on a general consensus being reached within the scientific community; results are made objective by general agreement. One might call this attitude "relative objectivism". This is actually the most which can be said about any physical theory because in the natural sciences a theory can never be verified but only falsified, a point quite correctly made by Popper. Thus, what we have to adopt is a learning process based on the correct choice of adequate constraints. This is in fact what has happened in thermodynamics where by now we all know that the adequate constraints are the conservation laws.

In the field of non-equilibrium phase transitions, or more generally speaking, of open systems, we wish to make the first steps by showing what these constraints are. Indeed, when we confine our analysis to non-equilibrium phase transitions, we find complete agreement between the macroscopic approach by the maximum (information) entropy principle and the results derived from a microscopic theory for all cases where the microscopic distribution functions are known. Therefore, I am sure that a consensus can be found here too.

There is another aspect important from the mathematical point of view. Namely, when we prescribe specific constraints which are given experimental mean values, the maximum (information) entropy principle will always provide us with distribution functions which reproduce these mean values. In this sense we are dealing here with a tautology. Then, however, in the next step we may infer new mean values by means of the probability distribution and then predictions are made which can be checked experimentally. If these predictions are not fulfilled, we may choose these new experimental data as additional constraints which then give rise to altered probability distribution functions. In this way an infinite process has been started. But in spite of this cautioning remark, we may find a consensus on the proper choice of a limited set of constraints, provided we confine our analysis to specific classes of phenomena.

One such class is, as mentioned, closed (thermodynamical) systems with their appropriate constraints. Another class consists of non-equilibrium phase transitions which will be treated here. As we shall see, this class comprises numerous phenomena in various fields, such as the emergence of spatial patterns, of new types of information and even of oscillatory phenomena. The appropriate choice of constraints for processes leading to deterministic chaos remains, at least partly, a task for the future.

As we shall see, the main new insight which we are gaining by our approach into the constraints is the following: In a first step one may guess that the adequate constraints must include the macroscopic variables, or in other words, the order parameters. But it is known that in non-equilibrium phase transitions critical fluctuations of the order parameters occur, i.e. that their fluctuations become macroscopic variables. Indeed, it will turn out that the inclusion of the fluctuations of the order parameters is the crucial step in finding adequate constraints for this class of phenomena.

Using the results of the microscopic theory as a guide, we are then able to do much more; namely, we can do without the order parameters from the outset. Instead our approach will start from correlation functions, i.e. moments of observed variables from which we may then reconstruct the order parameters and the enslaved modes. Incidentally, we can also construct the macroscopic pattern, or in other words we may automatize the recognition of the evolving patterns which are produced in a non-equilibrium phase transition.

In conclusion, let us return to the discussion of the relation between the analytical (or microscopic) approach and the holistic (or macroscopic) approach, and let us make a further point in favor of a macroscopic approach. Quite often, even the subsystems become very complicated so that it is difficult or even impossible to formulate the microscopic or mesoscopic equations explicitly. When we go to the extreme case, namely the human brain, the subsystems are, for instance, the nerve cells (neurons) which are themselves complicated systems. A nerve cell contains its soma, an axom and up to 80 thousand dendrites by which the cell is connected with other nerve cells. In the human brain there are about 10 billion nerve cells. It is proposed in synergetics that despite this enormous complexity, a number of behavioral patterns can be treated by means of the order parameter concept, where the order parameter equations are now established in a phenomenological manner.

Recently, we were able to find a paradigm, namely the coordination of hand movements and especially involuntary changes between hand movements. Though the relevant subsystems are quite numerous and consist of nerve cells, muscle cells and other tissue, the behavior can be represented by a single order parameter. We shall describe these experiments in this book and elucidate our general approach by this example. Further examples will be taken from laser physics and fluid dynamics. Looking at the numerous examples treated from the microscopic point of view in synergetics, it is not difficult to find many more applications of our new macroscopic approach. In addition, numerous examples, especially in biology, can be found where only the macroscopic approach is applicable.

We conclude with the remark that we shall use the expression "maximum information principle" exchangeably with "maximum entropy principle". But, as will transpire, the term information is the more appropriate to the situation in non-equilibrium systems.

2. From the Microscopic to the Macroscopic World...

2.1 Levels of Description

In this chapter I present the basic concepts and methods which have been used in synergetics to study self-organization by means of a microscopic approach. Readers familiar with this approach may skip this chapter and proceed directly to Chap. 3. On the other hand those readers who are unfamiliar with these concepts and methods and who wish to penetrate more deeply into them are advised to read my books "Synergetics. An Introduction" and "Advanced Synergetics" where all these concepts are explained in great detail.

When we deal with a system, we first have to identify the variables or quantities by which we wish to describe the system. Such a description can be done at various levels which are, however, interconnected with each other. Let us discuss the example of a fluid (Fig.2.1). At the microscopic level, the fluid can be described as being composed of individual molecules. Thus, for a complete description of the fluid, we have to deal with the positions and velocities of the individual molecules. However, for many purposes it is sufficient to deal with the fluid at a mesoscopic level. Here, we start from volume elements which are still small compared to the total size of the fluid, but which are so large that we may safely speak of densities, velocity fields, or local temperatures. Finally, our concern will be the macroscopic level at which we wish to study the formation of structures, or of patterns, for instance a hexagonal pattern in a fluid which is heated from below. A similar subdivision of levels can also be made with respect to biological systems. But as the reader will recognize quickly, we have here a much greater arbitrariness in choosing our levels.

Let us consider the example of a biological tissue. At the microscopic level we may speak of biomolecules or equally well of organelles and so forth. At the mesoscopic level we may speak of cells and finally at the macroscopic level we may be concerned with whole tissues or organs formed by these cells. Quite clearly, for the transition from the mesoscopic to the macroscopic level we must neglect many detailed features of the cells and their constituents. Instead we must pick out those features which are relevant for the formation of organs. Therefore, at this mesoscopic level we are already dealing with an enormous compression of information.

In this chapter we shall mainly be concerned with the transition from the mesoscopic to the macroscopic level, though in a number of cases a direct transition from the microscopic to the macroscopic level can also be performed. Within

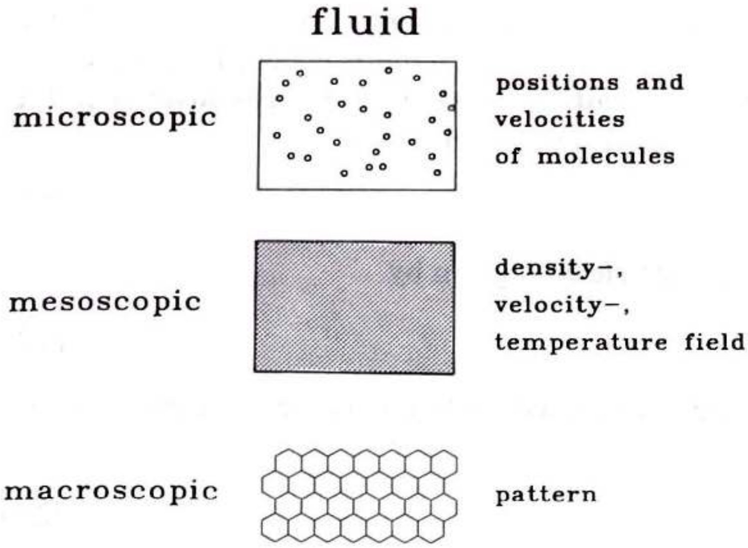

Fig. 2.1. Illustration of the microscopic, mesoscopic and macroscopic approach by means of the example of a fluid

Table 2.1. Examples of mesoscopic variables

Field of Study	Variables
chemical reactions, solidification	densities of molecules in different phases
fluids	velocity fields
flames	temperature fields
plasmas	electric and magnetic fields
lasers, parametric oscillators	atomic polarization, inversion
solid state physics, Gunn oscillator, filamentation	densities of electrons and holes
morphogenesis	densities of cells in tissues
pre-biotic evolution	numbers of biomolecules
population dynamics	numbers of animals
neuronal nets	firing rates of neurons
locomotion	elongation and contraction of muscles
economy	monetary flows
sociology	number of people with specific attitudes
synergetic computers	activation of elements

physics, the mesoscopic level can be reached by means of statistical mechanics where certain relevant variables are then introduced and treated. In most cases, however, such as in chemistry and biology we shall use phenomenological equations for the corresponding variables.

Let us give the reader an impression of the variety of problems to be treated by means of listing a number of examples of mesoscopic variables (Table 2.1).

2.2 Langevin Equations

A quite typical example for the description at the mesoscopic level is provided by the Brownian motion of a particle, say a dust particle, which is immersed in a fluid. Its motion is described by the Langevin equation where the variable q is then to be identified with the velocity of the particle. The microscopic motion of all the molecules of the liquid has two effects. On the one hand it leads to a damping of the velocity and on the other hand it leads to random impulses delivered to the particle under consideration. The Langevin equation reads

$$\dot{q} = K(q) + F(t) \tag{2.1}$$

where in the case of the Brownian particle K is given by

$$K(q) = -\gamma q \ . \tag{2.2}$$

A form which we are quite often concerned with in synergetics is given by the non-linear expression

$$K(q) = \alpha q - \beta q^3 \ . \tag{2.3}$$

The fluctuating forces F are characterized by the properties

$$\langle F(t) \rangle = 0 \qquad \text{and} \tag{2.4}$$

$$\langle F(t) F(t') \rangle = Q \delta(t - t') \tag{2.5}$$

where the average is taken over the stochastic process. When we deal with several variables $q_1, \dots, q_N$ which are lumped together into a state vector $\boldsymbol{q}$, the Langevin equation reads

$$\dot{\boldsymbol{q}} = \boldsymbol{K}(\boldsymbol{q}) + \boldsymbol{F}(t) \tag{2.6}$$

and the fluctuating forces are assumed to possess the properties

$$\langle F_j(t) \rangle = 0 \tag{2.7}$$

$$\langle F_j(t) F_{j'}(t') \rangle = Q_j \delta_{jj'} \delta(t - t') \ . \tag{2.8}$$

Note that $\boldsymbol{q}$ may be a vector in a high dimensional space thus representing a very complicated system. In a number of cases the fluctuating forces are themselves dependent on the state variable $\boldsymbol{q}$. In such a case a number of specific problems arise which were solved by the Îto or Stratonovich calculus. In the *Îto calculus* the Langevin equation (2.1) must be replaced by the following equation

$$dq(t) = K(q(t))\, dt + g(q(t))\, dw(t) \ . \tag{2.9}$$

Here K and g are in general non-linear functions of q whereas dw describes a stochastic process where we make the assumptions

$$\langle dw \rangle = 0 \tag{2.10}$$

$$\langle dw(t)\, dw(t') \rangle = \delta(t - t')\, dt \ . \tag{2.11}$$

In the Îto formulas it is assumed that $q(t)$ and dw which occur in the last term of (2.9) are statistically uncorrelated. When we deal with a multidimensional state