ELSEVIER

Biophysical Chemistry 85 (2000) 101–123

Biophysical Chemistry

www.elsevier.nl/locate/bpc

Natural selection: a phase transition?

Manfred Eigen*

Max-Planck-Institut für Biophysikalische Chemie, Am Fassberg 11, D-37077 Göttingen, Germany

Abstract

Information has two aspects: a quantity to be called 'extent' and a quality which may be termed 'content' since it deals with meaning. The latter originates via selective self-organization, which can be described also in quantitative physical terms. A prerequisite is the reproducibility of the informational substrate forming the basis of selection. This paper focuses on selection being the analogue of a physical phase transition. In Section 1 the criteria for phase transitions are formulated. Section 2 introduces the concept of information space and describes information as selected points or regions in this space. In Section 3 selection is analyzed in terms of the criteria for phase transitions, and in Section 4 the concept is confronted with experimental data. The conclusion is reached that information content is generated via selection, which can be described as a phase transition in information space.

Keywords: Phase transitions; Information space; Coding theory; Error threshold; RNA and DNA replication

1. How to define a phase transition?

'Phase transition' is a typical subject of physical chemistry, 'natural selection' *the* central theme of biology. I chose this title as a tribute to a biophysical chemist, colleague and friend — coeditor of this journal for many years — on the occasion of his 70th birthday. Gerhard Schwarz — I am tempted to say — was my first scholar (and I use this word having in mind both its meanings). More correctly, he was a student of Karl Friedrich Bonhoeffer, one of the early promoters of biophysical chemistry, who asked me to supervise his thesis work. This was back in 1954 and it started a fruitful cooperation which in the late 1950s and 1960s induced the study of cooperative conformation changes in biopolymers and their kinetics [1–3].

Cooperativity is one of the characteristics of phase transitions. Let us, for a moment, stay with biopolymers. The transformation of a helical structure — such as the α-helix of a polypeptide or the DNA double helix — into a randomly coiled form appears to be of a cooperative nature. 'Cooperation' means that the individual monomers in the polymeric chain do not change

* Corresponding author. Tel.: +49-551-201-1437; fax: +49-551-201-1435.

0301-4622/00/$ - see front matter
PII: S0301-4622(00)00122-8

their conformation independently from one another. They 'correspond' with their neighbors. In other words, they are more likely to change if the neighbor has already changed, or they 'dislike' it and behave anti-cooperatively. Pörschke [4], studied rates and stabilities of base pairing with oligonucleotides of various lengths and found that the stability of a basepair next to an already existing one is by a factor $\sigma^{-1} \approx 10^4$ larger than that of a single (isolated) pair.

In 1925 Ising [5] tried to understand the cooperative behavior of ferro-magnets by considering a simplified model and deriving its statistics. The 'Ising model' assumed a linear chain of elementary magnets, the moments of which are in either of two states ('up' or 'down'), influenced by the state of their nearest neighbors. Ising concluded (correctly) that the linear chain may undergo cooperative changes, but no true phase transitions, i.e. 'all or none' changes of the whole chain. He further extrapolated (incorrectly) that nearest neighbor models for any dimension > 1 (that are more realistic for ferromagnets) would show no true phase transitions either.

The one-dimensional Ising model had a late comeback with the helix-coil transition theories of linear biopolymers [6–8]. Kinetic analysis was done at our laboratory at Göttingen by Crothers [9], Lumry [10], and Schwarz [11] and revealed the validity of the linear Ising model. From my Harvard lectures [12], given in 1965, I recovered the shape of the transition curves obtained (Fig. 1). The essential result is that for sufficiently long (linear) chains there is a finite correlation length of cooperative transformation, essentially given by the reciprocal square root of the cooperativity parameter σ. This parameter describes how less likely the nucleation than the continuation of the transition proceeds. For a value of σ of 10^{-4} the 'melting curve' behaves as if blocks with an average length of hundred monomers simultaneously change from one state to the other. For the later discussion we should keep in mind that in the case of linear models true phase transitions cannot be expected, even for indefinitely long chains. They would require an unlimited correlation length, i.e. an unlimited σ^{-1}, which physically doesn't make much sense.

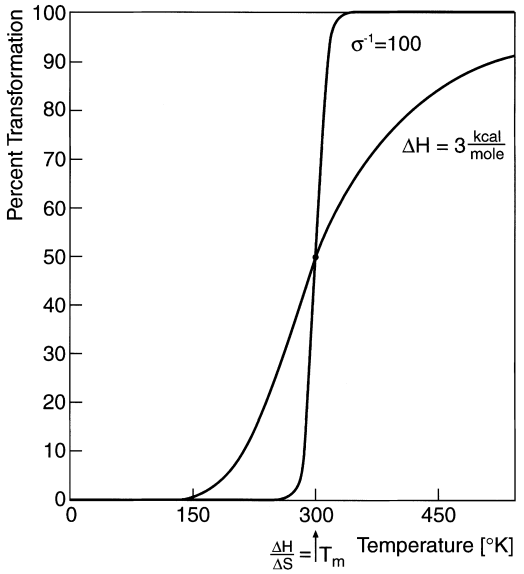

Fig. 1. Schematic representation of cooperative as compared to non-cooperative change. If in a linear polymer the subunits would change independently of one another with a reaction enthalpy ΔH of 3 kcal/mol the transition curve (centered at $\Delta H/\Delta S = T_m = 300$°K) would extend over a fairly large temperature range of nearly 300°K (10–90%). A nearest neighbor interaction described by a linear Ising-model would steepen the curve. The cooperativity factor σ indicates how much more difficult it is to nucleate than to propagate the change, based on nearest neighbor interactions. The original curve steepens by a factor $\sigma^{-1/2}$, in the example shown equivalent to a simultaneous change of 10 subunits, each involving a ΔH of 3 kcal/mol. For helix-coil transitions σ factors down to 10^{-4} have been observed.

We now have a second criterion for true phase transitions: cooperativity as such is not sufficient, we need in addition, an unlimited correlation length. In 1944 Onsager [13] showed that a two-dimensional Ising model indeed may have these properties and hence show true phase transitions. Mathematically, Onsager's exact solution was a 'tour de force', unprecedented in condensed matter physics. It dealt with a magnet at zero external field. So far no one was able to improve Onsager's solution as to include higher dimensions or finite external fields. Instead attempts to come closer to reality have been made by suitable approximations such as mean field theories. The most successful approach to the problem was the Landau theory of phase transition. It solved many prob-

lems — except one — the critical phenomena, previously designated as phase transitions of second order. It led to a better approximation than the semiempirical classical models did, but it yields a critical exponent who differs clearly from well established experimental data. Let me specify this point with the help of an example well known to the chemists: the transition gas ↔ liquid.

As early as 1881 van der Waals [14] came up with a two-parameter equation that reproduces remarkably well the phase transition between liquid and gas. In his doctoral thesis (1873) he had dealt already with the subject 'On the Continuity of the Liquid and Gaseous State', knowing that the Ideal Gas law could be derived from kinetic theory under the assumptions of point-like molecules and an absence of intermolecular forces of attraction. The two parameters he introduced were meant to take care of those assumptions: the volume V was corrected by subtraction of a constant hard sphere volume b, and the (external) pressure by adding an 'internal' pressure term a/V^2. This latter term can be explained by the assumption that the internal energy, due to attractive forces among molecules, is inversely proportional to the volume. At constant temperature the pressure p equals $-(\partial F/\partial V)_T$ yielding a contribution to pressure by internal interactions which is inversely proportional to V^2. Fig. 2 shows the pressure–volume relation according to the van der Waals equation. In this Figure two regions are emphasized, namely the one designated to a coexistence of two phases (light shading) and the one forbidden by stability criteria for homogenous phases, i.e. $(\partial V/\partial p)_T > 0$ (darker shaded). Both regions are peaked by the critical point, and it is on this, which I shall focus further on.

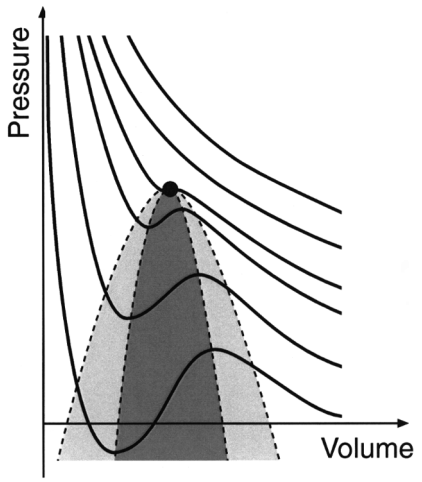

Fig. 2. Graphic representation of van der Waals isotherms: $(p + (a/V^2))(V - b) = RT$. Left of the shaded area the system is in the liquid state, right of the area it is a gas. The black dot indicates the critical point. The isotherm, which passes this point, refers to the critical temperature. At pressures above critical pressure, gas and liquid are undistinguishable. The weakly shaded area denotes the coexistence range of liquid and gas. The deeply shaded part is a region of instability in which $(\partial V/\partial p)_r$ is positive.

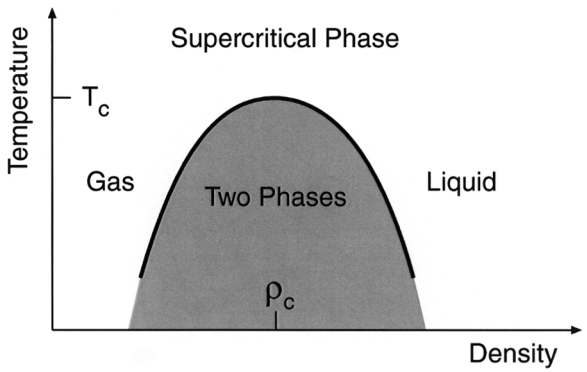

Fig. 3. Temperature–density diagram around the critical point. The critical curve $|\rho_c - \rho|$ is related to $(T_c - T)$ in a power law fashion as shown in the text with a critical exponent near 0.33... but being non-rational as substantiated by the error limits.

First, let me say that it is not my intention to reproduce in this paper standard knowledge from textbooks. For that purpose I refer to the excellent text produced by Berry, Rice and Ross [15]. What I want to demonstrate below is a new kind of phase transition that applies to a totally different non-material system in a peculiar space. In order to do so I have to digress a little further into the physics of phase transitions.

The critical point to which I referred above appears in the temperature (T) density (ρ) diagram shown in Fig. 3, or in analytical representation as a power law:

$$(\rho_{\text{liquid}} - \rho_{\text{gas}}) \propto (T_c - T)^{\beta} \tag{1}$$

with β being universally near 0.33. The error limits in the most accurate detections suggest that β — although close to one-third — is not a rational number. In particular it differs clearly from 0.5, the power that came out of the Landau theory [16]. The universality behind the correct power law is demonstrated by the fact that the same critical coefficient is being found for entirely different kinds of critical phase transitions, for instance in the behavior of ferromagnets or antiferromagnets at their critical temperatures ($T_c \sim$ Curie temperature for ferro — or Néel-temperature for antiferromagnets).

So what is wrong with the Landau theory? Nothing is wrong! Wherever mean field approaches apply, this theory is almost unbeatable. In particular, it reflects universality and in many cases yields an excellent representation of experimental data. Moreover, it tells us where it is prone to fail. Being a 'mean field' approach it ignores fluctuations. All microscopic phenomena include fluctuations, but these fluctuations usually cancel out if macroscopic dimensions are involved. This is not the case for critical phenomena. We all know the phenomenon of critical opalescence, which results from fluctuations of density and hence of the refractive index at all scales of length. The same is true for magnetic phenomena near the Curie temperature, or for superfluidity phenomena in ^4H_e [16]. Fluctuations at all length scales require an iterative reconsideration of the partition function at all length scales. This is accomplished by the renormalization group, introduced by Wison [17].

Renormalisation was originally introduced in relativistic quantum field theory in order to remove infinities that occur in higher terms of the solutions obtained by perturbation theory. Wilson did his thesis work with Murray Gell-Mann and Francis Low who both came up with a renormalization scheme for perturbation theory. Wilson's version of the renormalization group was based on quantum field theory but then was given a more general physical interpretation. His colleagues at Cornell, the physical chemists Ben Widom and Michael Fisher, whose influence Wilson [17] specially acknowledges in his Nobel lecture, had prepared the ground by the notion of scaling laws, that generally apply to processes in the neighborhood of a critical point. According to a classification by Ehrenfest, critical phenomena were called second order phase transitions. This classification assigned an nth order to a discontinuity of the nth derivative of free energy at the point of phase transition. According to the modern view this definition fails except for the first order process, which involves the generation of a latent heat. The failure of Ehrenfest's definition is the failure of mean field assumptions. 'Second order phase transitions' involve non-analytic behavior rather than true discontinuities. A λ-point in the temperature dependence of heat capacity is a cusp rather than a true divergence. It was Wilson's method that showed the universality of so different phenomena as ferromagnetism near Curie temperature and liquid–gas transitions at the critical point yielding the correct (irrational) number for the critical exponent, to which mean field theory assigned a rational, such as $1/2$ in Landau theory. A profound discussion of these quite recent advances can be found in Goldenfeld's [16] 'Lectures on Phase Transitions and the Renormalization Group'. Let me summarize the results of this (purely descriptive) introductory part by three statements:

1. Phase transitions in physical space involve cooperative interactions.
2. The correlation length in the thermodynamic limit ($N \to \infty$) is infinite.
3. Critical phenomena involve fluctuations at all length scales. Renormalization leads to the correct (experimentally observed) critical coefficients.

2. How to generate 'information'?

Next we consider an arrangement of symbols, we call 'information'. Those arrangements commonly require material carriers. Genetic information, for instance, is contained in the base sequences of nucleic acid molecules. However, the material aspect here is of quite minor importance. Information as such is a property that

transcends its material form of appearance. The type of order characterizing information is reflected in the structure of a particular space, which we have called 'sequence' or 'information space'. The concept goes back to Hamming [18] who used it first for binary sequences in a 'geo-

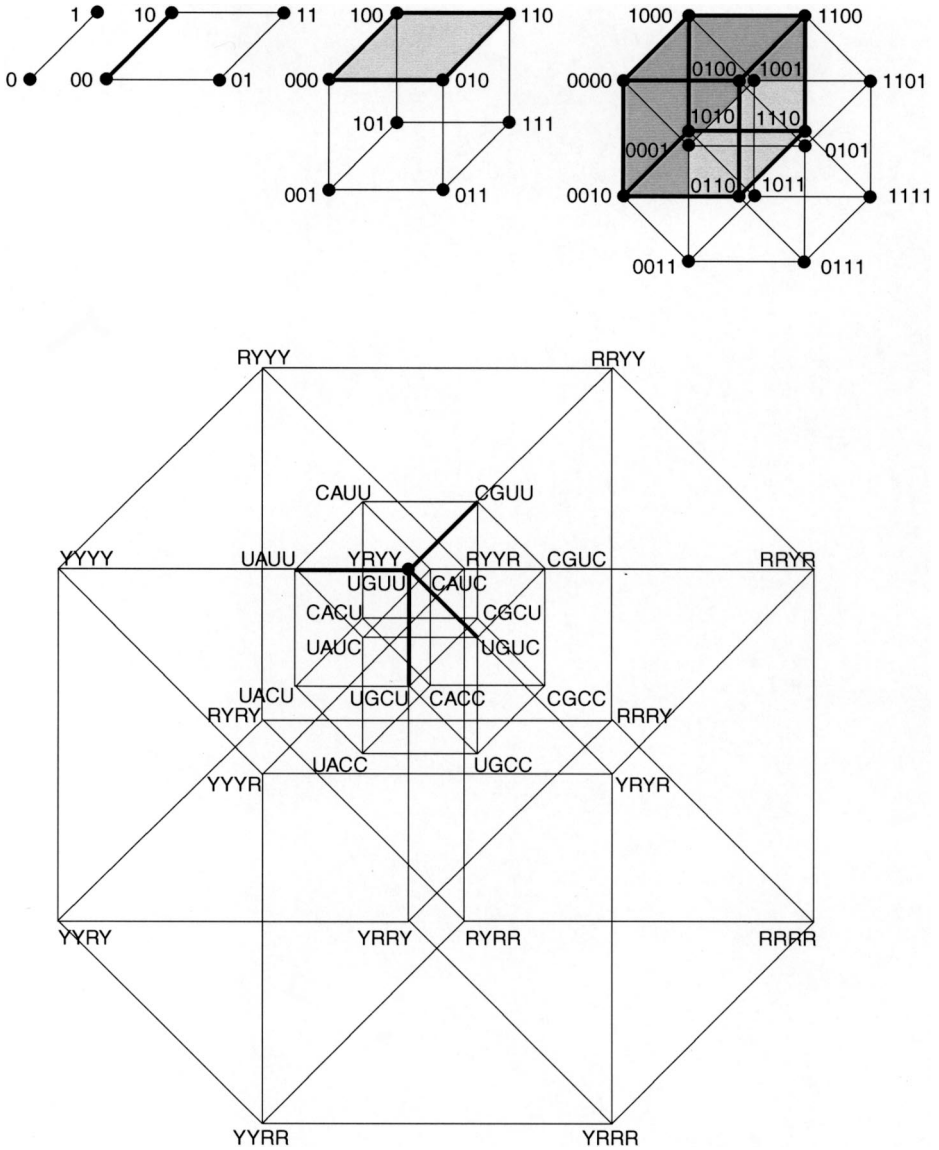

Fig. 4. Iterative construction of sequence space. Starting with one position (occupied by 0 or 1) of a binary sequence, positions are iteratively added resulting in a doubling of the preceding diagram in which corresponding points are connected by lines. The binary sequence space is a hypercube, each corner representing one of the 2^{ν} possible binary sequences of length ν (ν is the number of binary symbols — 0 or 1 — that comprise the total sequence). The lines connect nearest neighbors, i.e. mutants that differ in only one position. For nucleic acids the quaternary sequence space can be constructed by assigning first only purines (R) and pyrimidines (Y) yielding a binary hypercube. Each point represents one sequence in R and Y notations. Assigning for any sequence the two purines (A or G) and pyrimidines [U(T) or C] yields at each point of the binary R, Y-hypercube another hypercube as subspace. This is shown in the lower diagram for the sequence YRYY.

metric approach' to his theory of error-correcting codes. My paper on 'self-reproducing macromolecules' [19] is also based on the same concept, then adapted to quaternary sequences (such as RNA and DNA). The construction of this space is explained in Fig. 4.

The idea is that any possible symbol arrangement — in Hamming's case: any of the 2^{ν} binary sequences of length ν — is uniquely represented by a point (or a cell), and that all points are arranged in a way that kinship relations are correctly reflected by distances. A Hamming distance 'd' means that d positions in two arbitrary sequences are occupied by different (binary) symbols. The iterative construction of the space up to dimension four, starting with one single position, is demonstrated in the upper part of Fig. 4. Proceeding from any dimension to the next higher just means doubling of the former diagram and connecting equivalent points. The final result is a hypercube of dimension ν, each vertex of which represents one of the 2^{ν} possible sequences. The Hamming distance between two sequences is the minimal sum of edges that connect their two representative points. Hamming has shown that it is a legitimate distance function, which satisfies the three conditions:

1. that a distance from a point to itself is zero;
2. that a distance between two points is positive and convertible (i.e. AB = BA); and
3. that the triangle inequality holds which says AB + BC ≥ AC. Otherwise the Hamming metric is non-Euclidean.

The lower part of Fig. 4 shows how to extend the concept to quaternary sequences. R and Y are the two base classes of nucleic acids: purines and pyrimidines, which again have the binary choice of being R = A or G, Y = U or C (in RNA) or Y = T or C (in DNA).

Hamming [18] has used the distance concept for constructing an economical error-detecting and -correcting code. Given the billions of operations modern computers do within short times one needs equipment that is highly reliable. How to detect an error in a string of symbols that is sent through a noisy transmission line, is elegantly solved by Hammings theory. The easiest, however, quite costly way would be to send every message more than twice and decide by a majority vote. However, there are safer ways, which, in addition, are much less costly in computer time. They are based on parity checks, which means that additional positions are appended to the message positions and used to satisfy certain parity relations that are properly designed in order to fulfill the job.

Focus on a single vertex of the hypercube. Then all vertices with a given Hamming distance represent the surface of a 'sphere'. As in Euclidean space the surface of a sphere is defined by all points having the same distance from a center. Hence the volume of a Hamming sphere with radius 1 are all ν points with Hamming distance '1' plus the center point. Radius 2 then will include $1+\nu+\frac{\nu(\nu-1)}{2}$ vertices, and so on with the binomial series where the sum of all terms always yields the 'volume' and the last term the surface of the sphere. The procedure then is to fill the whole volume of space (2^{ν} vertices) with such spheres and use only the center of the spheres as message points. The ratio of the total volume and the sphere volume yields the maximum number of spheres that can be accommodated. It is a 'maximum number', that can be reached only when all the sphere volumes precisely match the total volume. This works ideally only in particular cases; otherwise the message spheres don't fill precisely the total volume. A perfect fit, for instance, is obtained for $\nu = 7$ and message sphere radius = 1, as depicted in Fig. 5. We can express this in mathematical terms. If ν is the total number of sequence positions, of which m positions are used as message points while $k = \nu - m$ positions are used for parity checks, we have for spheres with radius one as an example:

$$\frac{\text{total volume}}{\text{sphere volume}} = \frac{2^{\nu}}{1+\nu} \geq 2^{m} = \text{maximum number of spheres} \tag{2}$$

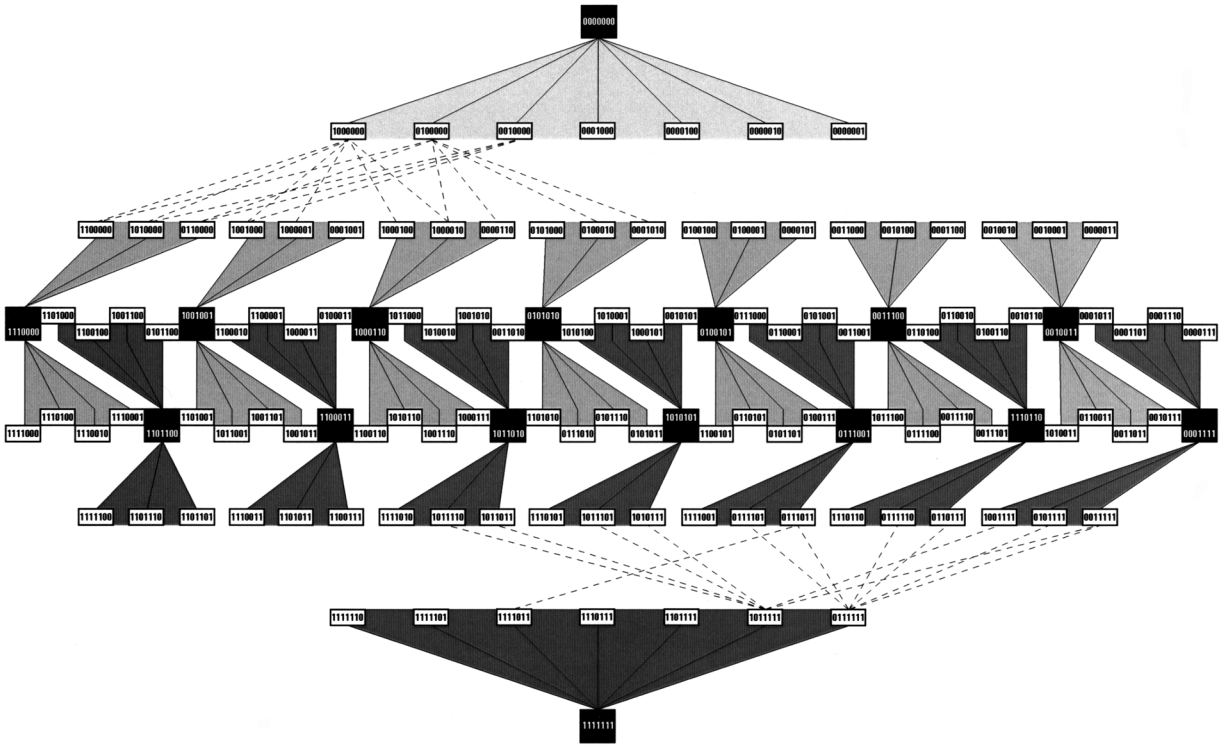

Fig. 5. Ideal partition of a Hamming space into densely packed non-overlapping spheres. The number of binary symbols is seven, so there are $2^7 = 128$ vertices of the hypercube. The radius of each sphere is 1. Hence, one may form $128/8 = 2^4 = 16$ non-overlapping spheres. The center points of the spheres are blackened. The Hamming distance between the center points of two neighboring spheres is 3. Hence, there are 16 message points, each comprising seven neighboring points allowing for localization of one mistake. Hence, in a sequence comprising seven digits, four are used for the message, and three for parity checks. This ratio is much more favorable for long sequences because of the exponential relations. A sequence space for $\nu = 90$ positions (2^{90} sequences) can be ideally partitioned by spheres with radius 2, having a volume of $1 + \nu + \frac{\nu(\nu-1)}{2} = 4096 = 2^{12}$ points, allowing for $2^{90}/2^{12} = 2^{78}$ possible messages. In other words: 78 of the 90 positions can be used for messages, while only 12 positions are necessary for each parity checks. Here the Hamming distance between sphere centers is 5 and two errors in each message could be identified.

or, since $\nu = m + k$:

$$2^{m+k} \geq 2^m\ (1 + \nu), \text{ or } 2^k \geq 1 + \nu \qquad (3)$$

For spheres with larger radii, the $1 + \nu$ is to be replaced by the corresponding sum of binomial coefficients, up to an order defined by the radius of the message sphere. Another example of perfect fit is $\nu = 90$ and radius 2. $1 + \nu + \frac{\nu(\nu-1)}{2} = 4096 = 2^{12}$. Hence, $k = 12$ and $m = 78$.

Note that the idea behind this scheme is to separate and isolate the message points, i.e. those points that carry a message. Minimum separation is at a Hamming distance 1, but that would use all vertices for messages. Hamming distance 2 would allow for spheres with radius 1, which share their surface points with those of neighboring spheres. This allows for single error detection but not for its localization and subsequent correction. Parity checks in this case just tell whether or not the message contains one single error. Only a Hamming distance 3 allows spheres of radius 1 not to overlap. Here localization and correction of the error is possible. This procedure can be continued in order to yield multiple error correction. For instance, double error correction requires a Hamming distance of 5 so that non-overlapping spheres of radius 2 can be formed (cf. the

above example). Of course, one can alternatively combine error detection and correction. For instance, a Hamming distance of 4 would allow for one of two alternatives: triple error detection or single error correction plus double error detection.

How economic this method of error correction is may be seen from an example. Assume a sequence length $\nu = 1000$. In order to correct for errors, which may occur in (up to) 20 positions, we have to isolate the message points by surrounding them with spheres with a radius of 20. The mutual Hamming distance of neighboring spheres in this case is 41. A sphere of radius 20 includes $\mu = \sum_{m=1}^{20} \binom{\nu}{m} \approx 10^{41.5} \approx 2^{138}$ points. Hence, we use only approximately 14% of the message symbols for error correction via parity checks and have approximately 86% of the symbols free for encoding messages. The message capacity then involves some $2^{860} \approx 10^{258}$ possible symbol combinations. I should say that an assumption of 20 possible errors (at least in the context discussed further below) is a very permissive one.[1] If the error expectation value per symbol is 10^{-3} (yielding for 10^3 symbols about one mistake), the probability of getting 20 errors in one sequence is as low as nearly 10^{-19}.

With these considerations we are (still) in the midst of classical information theory [20], which emphasizes the 'extent' of information, leaving 'content', i.e. meaning or value of information entirely unspecified. Classical information theory focuses on those properties, which are important for reliable transmission of messages, which involve channel efficiency and noise protection. In this connection Shannon's [21] concept of entropy, an analogue of Boltzmann's or Gibbs' entropy of statistical mechanics, is of importance. Entropy is a property of information space where individual points play a role only in as far as they are assigned different probabilities p_i of being populated. If all points had equal probabilities, Shannon's entropy were simply the logarithm of the volume of information space (which is 2^ν, i.e. the sum of all points). If the dual logarithm of 2 is used as the unit, entropy (often called 'information', better: 'extent of information') would be just ν bits. Non-uniform probabilities p_i take account of general constraints of language. Entropy H then reads:

$$H = \sum_i p_i \log\left(\frac{1}{p_i}\right) \quad \text{with } p_i \geq 0$$

and

$$\sum_i p_i = 1 \tag{4}$$

It is important to stress that the probabilities p_i in Shannon's expression are constrained by the structure of the language rather then by the particulars of the message sent. Due to the complex structure of languages it is not easy to put these constraints into quantitative terms. Fig. 6 represents the situation. Accordingly, we could interpret the extent of the measure of information as 'what we do not know, after we have taken into account everything we do know'. Given the length of the message it is an amount determined by the number of possible (statistically weighed) alternative symbol combinations which does not specify any particular message. For telecommunication it is — as Jaynes [22] said very aptly — 'the information of the telephone company which designs the coding device and the transmission line'. As an average quantity it does not apply to the peculiarities of real messages. We still could call it a logarithmic measure of a (statistically weighed) volume of information space, just as in Gibbsian terminology thermal entropy is a logarithm of phase space volume [23].

I am stressing this point in order to bring out now the different aspects of 'extent' and 'content' of information. The information contained in any message sent must have originated in some brain,

[1] If we call q the fidelity and hence $(1-q)$ the error rate per symbol a sequence of ν symbols then has the fidelity $Q = q^\nu \approx e^{-\nu(1-q)}$ or an error expectation of $1 - Q = 1 - e^{-\nu(1-q)}$. For the above values the expectation value for one error in the sequence is $1 - e^{-1} = 0.632$. The probability for k errors in a sequence of length ν is $\binom{\nu}{k} q^{\nu-k}(1-q)^k$, *yielding the above value*.

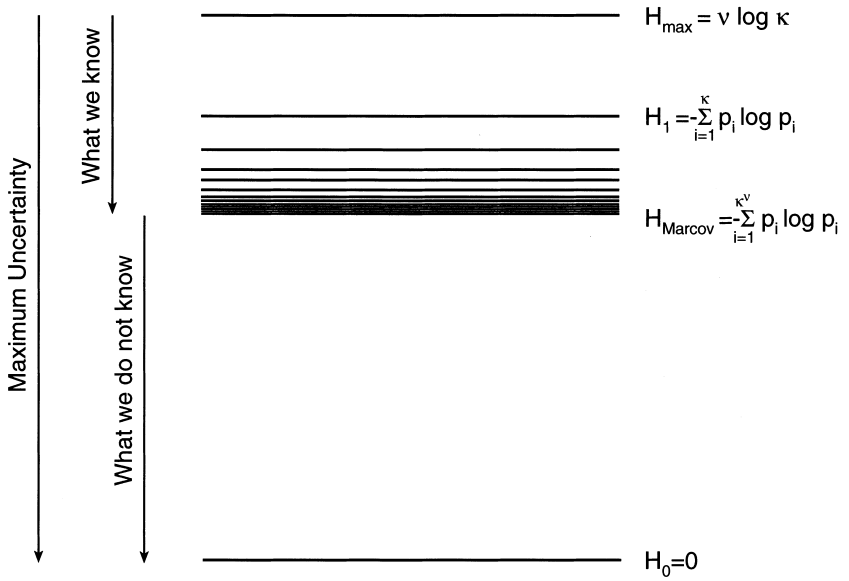

Fig. 6. Schematic diagram representing the meaning of Shannon's entropy in its meaning as 'extent' of information. The number of symbols in the message is ν, while κ denotes the number of symbol classes (e.g. $\kappa = 2$ for binary alphabets). In H_1, p_i is the probability (or frequency) of appearance of a symbol of class i; the sum is taken over all κ classes of symbols. In H_{Marcov}, p_i denotes the probability for the appearance of a certain symbol sequence i. Here the sum has to be taken over all κ^ν possible symbol sequences of length ν. H_{Marcov} takes into account all possible long-range symbol interactions. In a protein chain a given symbol might be in interaction with any other symbol via folding of the chain.

but long before brains were in existence, i.e. back to 4 000 000 000 years ago, biological information has been formed in what became the genes of living organisms. In order to generate the information that is laid down in a message (genotype), its meaning (phenotype) must somehow be evaluated and it is our general understanding that this occurs via some self-organizing process. Biologists refer to Darwin, which in modern language could be formulated as, genetic information came about by evolution based on 'natural selection'. If this process started at the molecular level (and still is continued at this level in a highly sophisticated form), we might well ask: what does 'natural selection' mean in terms of physics?

I am not going to redescribe all the theoretical and experimental work, done at various laboratories during the past 30 years that has established the validity of an extension of the Darwinian thesis to the molecular level which now reads:

1. genetic information comes about by evolution;
2. evolution is based on natural selection; and
3. natural selection is a consequence of self- or complementary, mutagenic replication.

In the picture of information space (Fig. 4) this means: evolution of any gene must have started in some region of the space representing sequences that were somewhat better adapted to a certain function than others simultaneously present and, therefore, became selected. Adaptation was obviously quite weak under any condition at which such a process could have started. Therefore, if there is mutagenic replication plus selection the system will proceed steadily along a trajectory in sequence space until it arrives at a point of optimal adaptation. This sounds reasonable, but the logical sequence from (i) to (iii) doesn't yet seem plausible. So let's invert it: the complexity of sequence space is exponential with respect to sequence length, exceeding soon the capacity of any realistic scenario. Therefore, any given sequence can be populated reproducibly only if it is replicated with a rate larger than its decomposi-

tion rate. (Initially this is quite a tough condition.) If copying by replication were entirely correct there would be selection of the most efficiently multiplying sequence, but no further evolution. The process would get stuck at a point in sequence space that represents the locally best-adapted sequence. Given the complexity of any sequence of reasonable length (i.e. representing genetic information) a 'locally' best sequence generally must be far inferior to the 'globally' best sequence. Only variation (by point mutation, deletion, insertion or new combination) can free the system from such a trap and start a trajectory from the 'locally' to the 'globally' best point.

So far this all sounds very descriptive, but before starting a more quantitative approach two points should be clear:

1. Reproducibility is a necessary prerequisite of selection in view of the complexity of sequences. Note that a gene length of 100 (like a tRNA molecule) or of a 1000 (like many 'true' genes) bases means $4^{100} \approx 10^{60}$ or $4^{1000} \approx 10^{600}$ possible alternative sequences, the lower of both values already exceeding the capacity of our universe. Clearly, those alternatives can't, and don't need to occur during the evolutionary process, but this is so only thanks to an inherent reproductivity. ('Inherent' means that all members of the class can behave in this way.) Absence of reproduction, for instance: stochastic de novo production, would fill the sequence space like an ideal gas with exceedingly low density.
2. Reproduction is the basis of selection. It requires a system far from equilibrium where autocatalytic formation and first order decomposition are not linked in a reversible fashion. 'Selection'(c.f. [2]) differs from a mere favoring of some state, which could be materialized also in equilibrized systems, if the state is distinguished by a minimal free energy. 'Selection' requires complete disappearance of 'less fit' competitors. On the other hand, such an ideal selection of a single state, which I would compare with an 'ideal crystal', would not allow for evolution because the system gets stuck at the 'fittest' local state.

The three prerequisites for evolutionary behavior are:

1. inherent replication as the basis of selection;
2. mutational variation as the basis of evolutionary adaptation; and
3. metabolism for keeping the system stable far from equilibrium.

This essentially confirms the Darwinian notion, now extended to molecules. The detailed treatment, however, uncovers facts that cannot be reached by verbal argumentation and partly also remained obscured in the neo-Darwinian approaches of population theory.

A system of phenomenological equations that takes care of the above prerequisites was proposed in 1971 [19]. It has the form:

$$\dot{\vec{\mathbf{x}}}(t) = \left\{(W_{ik}) - \delta_{ik}\bar{E}(t)\right\}\vec{\mathbf{x}}(t) \qquad (5)$$

where $\vec{\mathbf{x}}$ is a vector the components of which are the relative concentrations introduced above,

[2] Consider the rate equation of an autocatalytic process far from equilibrium:

$$\dot{n}_i(t) = (A_i - D_i)n_i(t) \equiv E_i n_i(t) \geq 0$$

where $n_i(t)$ is the time derivative of a particle number (per volume) $n_i(t)$, A_i the rate parameter of autocatalytic 'amplification', and D_i that for first order 'decomposition' while E_i denotes (positive) 'excess' production. Suppose that N different competitors are present. Introduction of relative concentrations

$$x_i(t) = n_i(t) / \sum_{k=1}^{N} n_k(t) \quad \text{with} \quad \sum_{k=1}^{N} x_k = 1$$

yields

$$\dot{x}_i(t) = \{E_i - \bar{E}(t)\}x_i(t)$$

which describes selection, because the asymptotic solution for $x \to \infty$ reads:

$$x_{i=\max} \to 1;\ x_{k\neq i} \to 0 \quad \text{and} \quad \bar{E}(t) \to E_{\max}.$$

In other words, only those compounds grow (exponentially) for which $E_i > (t)$, all others decay (exponentially). By this very fact $\bar{E}(t)$ steadily grows until it reaches $E_{\max}$. At this point the corresponding relative concentration reaches one, while all others have decayed to zero.

(W_{ik}) a matrix of rate coefficients, $\bar{E}(t)$ the average overall production as defined in footnote 2 and δ_{ik} the Dirac δ (δ_{ik} equals zero for $i \neq k$ and one for $i = k$). For a binary sequence of length ν the vector $\vec{\mathbf{x}}(t)$ has 2^ν components and (W_{ik}) is a $2^\nu \times 2^\nu$ matrix. (For quaternary sequences such as RNA or DNA 2^ν is to be replaced by 4^ν.) This matrix can be written as a product of an error matrix (Q_{ik}) times a vector whose components are the rate coefficients to be assigned to each of the 2^ν vector components of $\vec{\mathbf{x}}(t)$:

$$(W_{ik}) = (Q_{ik} A_k - \delta_{ik} D_k) \tag{6}$$

The product, i.e. the matrix (W_{ik}) is intrinsically non-symmetric because the rates of replication differ individually and it is this non-uniformity which eventually causes selection. The diagonal coefficients $W_{ii} = A_i Q_{ii} - D_i$ correspond to the term in footnote 2 describing precise replication. Q_{ii} is the diagonal term of the error matrix expressing the fraction of error-free replication processes as introduced in footnote 1. For an average symbol fidelity $\bar{q}$ (having values between 0 and 1) it equals $\bar{q}^\nu$. The non-diagonal terms $W_{ik} = Q_{ik} A_k$ certainly are non-symmetric because of $A_i \neq A_k$. However, in the case of uniform fidelities q the matrix (Q_{ik}) is symmetrical: $Q_{ik} = Q_{ki}$, because of the binomial form

$$Q_{ik} = \binom{\nu}{d_{ik}} q^{\nu - d_{ik}} (1-q)^{d_{ik}}.$$

In other words: in a binomial distribution the error frequency depends only on the Hamming distance d_{ik} between state i and k. In this case (W_{ik}) is the product of a symmetric matrix with a vector causing all eigenvalues to be real. On the other hand, the assumption of an average $\bar{q}$, though being an useful approximation, does not really involve symmetry of the error matrix. The fidelity $\bar{q}$ is a geometric mean and involves the product $\prod_{i=1}^{\nu} q_i$ for all positions of the sequence. A hot spot, i.e. a higher error probability at a certain position as a consequence of its microenvironment, by no means needs be symmetric with respect to the reverse process.

Eq. (5) has a quasi-linear appearance, but it is basically non-linear. The rate coefficient A contains at least the concentrations of the monomeric substrates. Moreover, in enzymic catalysis it involves the enzyme concentration and those of compounds that regulate enzymic activity. We have formulated these expressions analytically and, in particular, Biebricher [24] in our laboratory has tested them experimentally in RNA-model systems. In Eq. (5) all these concentrations are assumed to be buffered at constant values, which easily can be effected in experiments — except for the template of the reaction which at the same time is the reaction product that grows 'autocatalytically', which under many conditions means 'exponentially'. A reaction behavior as described by Eq. (5) requires an exchange of material with the environment which may be regulated in a way that either reaction fluxes or forces are kept constant [19]. As a result, even the simplified quasi-linear form is not linear since it contains the variable term $\bar{E}(t)$.

Originally, solutions of Eq. (5) in explicit form were obtained by second order perturbation theory. It was Mark Kac — at the occasion of a lecture I gave at Rockefeller University — who thought that an exact solution of the equations could be given. These appeared soon afterwards in two independent papers by Thompson and McBride [25] and by Jones, Enns and Rangnekar [26]. The trick is to introduce the variable

$$z_i(t) = x_i(t) \exp\left\{ -\int_0^t \bar{E}(\tau) \mathrm{d}\tau \right\} \tag{8}$$

yielding a system of linear differential equations that can be solved. After transforming back to the variables $x_i(t)$ the solutions can be expressed in a 'normal mode' form introducing the variables $y_i(t)$, which contain contributions from all variables $x_i(t)$:

$$\dot{y}_i(t) = \left(\lambda_i - \bar{\lambda}(t)\right) y_i(t) \tag{9}$$

with the λ_i being the eigenvalues obtained from the solutions of the linear equations and $\bar{\lambda}(t)$ their average which is equal to $\bar{E}(t)$ [27]. Now —

in the same way as was explained in footnote 2 — we see that the system describes 'natural selection'. However, the target of selection is no longer a single sequence 'species', but rather a clan of

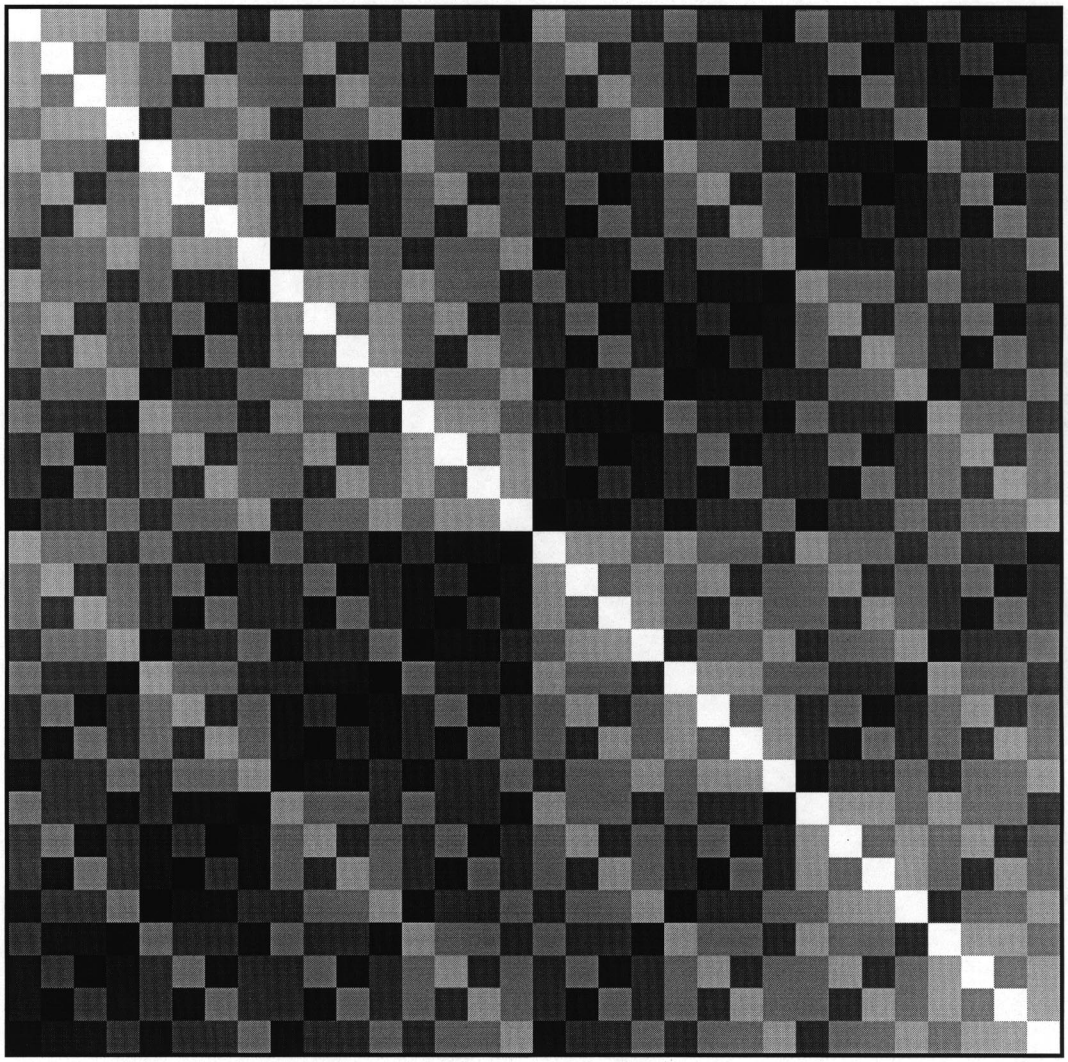

Fig. 7. Schematic representation of the error matrix (Q_{ik}), built up according to the same iterative principle as was applied to sequence space in Fig. 4. Starting in the upper left corner there are four boxes for a single position being either 0 or 1. The two (white) diagonal boxes represent error free copying $0 \rightarrow 0$ or $1 \rightarrow 1$. The non-diagonal boxes refer to errors $1 \rightarrow 0$ and $0 \rightarrow 1$. For two positions we have to square the first diagram yielding 16 boxes. Again the four diagonal boxes refer to $00 \rightarrow 00$; $10 \rightarrow 10$; $01 \rightarrow 01$ and $11 \rightarrow 11$. For each additional position the former diagram has to be squared, which discloses itself clearly in the picture referring to a sequence with five positions. The diagonal terms refer to the $2^5 = 32$ sequences that reproduce themselves correctly. The blackening deepens with increasing number of errors. The black counter diagonal refers to complete conversion of all symbols, such as $00000 \rightarrow 11111$. If this iterative scheme is extended to the total sequence each diagonal term represents one of the 2^ν complete sequences of length ν, reproducing itself correctly, just as any of the 2^ν vertices of the hypercube after complete iteration represents one of the 2^ν complete sequences. All non-diagonal terms refer to errors. The rows refer to the replicas produced, the columns to the producing templates. The term in ith row and kth column Q_{ik} refers to a sequence k producing sequence i through erroneous copying of k.

'species', ordered by kinship relations. These clans behave as if they were single species in the sense of footnote 2; therefore Peter Schuster and I have termed them 'quasi-species'. However, selection now is somewhat different: the target is not one individual 'species' (which might be a molecular one), but rather a particular kind of distribution among all species, which as such outgrows alternative distributions. The selected distribution, which is called the quasi-species, has a 'center of gravity', resembling the 'wild-type'. It may, but need not be, the dominating master species, defined by W_{mm} (m for max.). It may as well be the center of several degenerate or neutral master species which as such needs not be characterized by a maximum W_{mm}.

The simple form of Eq. (9) involves a wealth of peculiar and sometimes entirely unexpected solutions. We need perturbation theory to arrive at any such individual solution. Just one general remark: as in particle physics, those perturbation solutions may contain singularities. Common terms in these solutions, describing the population number of mutants relative to the wild type, are of the form $\frac{W_{ii}}{W_{mm} - W_{ii}}$, or products of those terms, which would become infinite for a neutral mutant, i.e. if $W_{ii} = W_{mm}$. McCaskill [28] has shown that — as in particle physics — this problem can be solved by renormalization.

However, what I wanted to address in this paper is another problem, more closely linked to the title of this paper. Rumschitzki [29] has shown that the original Eq. (5) can be explicitly solved in many cases if it is arranged in the same way as we introduced sequence space in Fig. 4. One starts with one position and adds successively all other positions, thereby finally yielding the rate equation $\dot{x}_i(t)$ for any of the 2^{ν} (or 4^{ν}) possible states. The error matrix for binary sequences then has the recursive form depicted in Fig. 7.

Let's start with the four squares in the upper left corner. They represent the four possible changes of one position. The two diagonal squares represent faithful reproduction of either 0 or 1. The two non-diagonal boxes mean the mutations $1 \rightarrow 0$ (upper) and $0 \rightarrow 1$ (lower). Now add a second position. Similar to the construction of sequence space the diagram now gets squared. There are now four states: 00, 10, 01 and 11 which may mutate in 12 different ways, i.e. by eight possible one-error and four two-error mutations. In Fig. 7 the Hamming distance involved in the mutation is indicated by the darkness of the square, which here reaches up to a Hamming distance 5, but is to be iterated as to include all 2^{ν} possible sequences. Faithful (error-free) replications are represented by the (2^{ν}) diagonal squares. The symmetry of the error matrix in the case of uniform fidelities q is neatly reflected by the graph. A deeper mathematical foundation (in terms of the hyper-octahedral group) is given by Dress and Rumschitzki [30].

The picture of the quasi-species now evolving is that of a (partly filled) Hamming-sphere in sequence space that, given a certain population size, has a relatively large radius by the fact that certain (neutral) mutants still appear with finite population numbers (≥ 1 individual) near its periphery. Of course, this is as little a true sphere as any living being can be represented by an 'ideal sphere'. 'Sphere' means here an extension characterized by some 'radius' which refers to an 'evaluated' quasi-species, which includes many (in fact most) states that are not populated at all, but which shows finite population numbers for neutral (or nearly neutral) mutants at fairly large Hamming distances from the wildtype. The center of the sphere is defined by the 'most central' neutral in this distribution.

Here I am coming back to my remarks about Hammings theory of coding at the beginning of this section. The recursive construction of the rate equations, following the same principle as the construction of sequence space in Fig. 4, provides a dynamic fitness landscape that guides the process of selection. It determines the population structure of the sequence space, i.e. it assigns population numbers to the vertices of the hypercube. One might ask whether the three parameters $\bar{q}$, A and D are sufficient to describe any sophisticated task of evolutionary adaptation. The answer is that these parameters themselves may be sophisticated functions, specified by the particular need of adaptation. Here they appear as phenomenological parameters, which may ex-

press any kind of selection pressure, and in this form are widely applicable.

We see that the quasi-species structure is a general consequence of mutagenic replication. The resulting properties make the quasi-species a system which is — so to speak — inverse to Hamming's coding model. First, in Hamming's spheres (in the ideal case) all points have a function. The center point carries the message and the surrounding points are used for the various parity checks. By contrast, a quasi-species in order to cover a large region leaves most points in its 'sphere' unoccupied. Second, the purpose of the coding spheres is to conserve the message, while the structure of the quasi-species should aid the process of evolution, i.e. to change rather than to conserve. Third, Hamming's spheres have the purpose of separation, there should be no 'kissing points' [31] among the spheres. Quasi-species should penetrate one another and should thereby build up many 'kissing points' among which neutral bridges may be found. Ideally, at any level of adaptation neutral nets should span across the whole space (which is only possible in multi-dimensional spaces). Evolution then is a combination of random walk along neutral nets with quasi-species formation of increasing fitness values.

3. Does information originate by phase transitions?

'Information' now includes both extent and content. According to the results of Section 2 extent is related to the 'constrained' volume of the information space and most suitably expressed by a logarithmic term which converts products of probabilities into additive quantities. Being expressed in the form of an entropy it refers to an extremum principle that guides equilibration within a certain distribution. According to Jaynes [22], maximum uncertainty of a probability distribution is to be accounted for by a principle of maximum entropy [32].

Content, on the other hand, refers to a particular point or region in information space, specified by dynamic parameters, which cause the system to respond in an optimal functional manner. Again, the optimal response involves an extremum principle that was identified to be the basis of 'selection'. Equilibration and selection dynamically are two inverse processes, one reaching to all states that possibly may be populated, the other narrowing down to one or few particular states that optimally match a certain situation. Does this 'condensation' resemble a true phase transition?

Let us look at some example. The decisive variables for selection are obviously the 'superiority' of the wild type based on maximum efficiency in producing offspring and the mean symbol copying fidelity $\bar{q}$ or its complementary symbol error rate $(1-\bar{q})$. Schuster and Swetina [33] have looked at such an example considering the complete set of rate equations. The system consists of a binary master sequence of length 50, which reproduces 10 times more efficiently than its (degenerate) mutants of equal length. Fig. 8 shows the results for different error rates $(1-\bar{q})$: The relative population variable $\bar{\mathbf{x}}_d$ (where d refers to Hamming distance with respect to the master sequence), is plotted vs. the error rate $(1-\bar{q})$. Glancing at this picture clearly creates the impression that we are dealing with some kind of phase transition, however, we have to analyze it in more detail.

At zero error rate $\bar{x}_0$ equals to 1, i.e. 'survival of the fittest'. In fact, this would be true for any smaller selective advantage than given by $W_{oo} = 10$ W_{kk} $(k \neq 0)$. With increasing error rate the master population is reduced in favor of the successively occurring mutant populations. The model case of degenerate mutants (having equal reproduction and decomposition rates) allows us to treat error classes by sum expressions; it was chosen for lucid representation, but otherwise is unrealistic. Fitness landscapes are entirely unsymmetrical, sometimes monotonous, and sometimes rugged including neutral or nearly neutral peaks. At an error rate $(1-\bar{q}) = 1/\nu$ $(= 0.02)$ the Poissonian distribution of errors discloses itself (unperturbed by the monotonous fitness landscape) by approximately one-third $(1/e)$ of master copies $(d = 0)$, another third of one-error copies $(d = 1)$ and the residual third shared by all other error

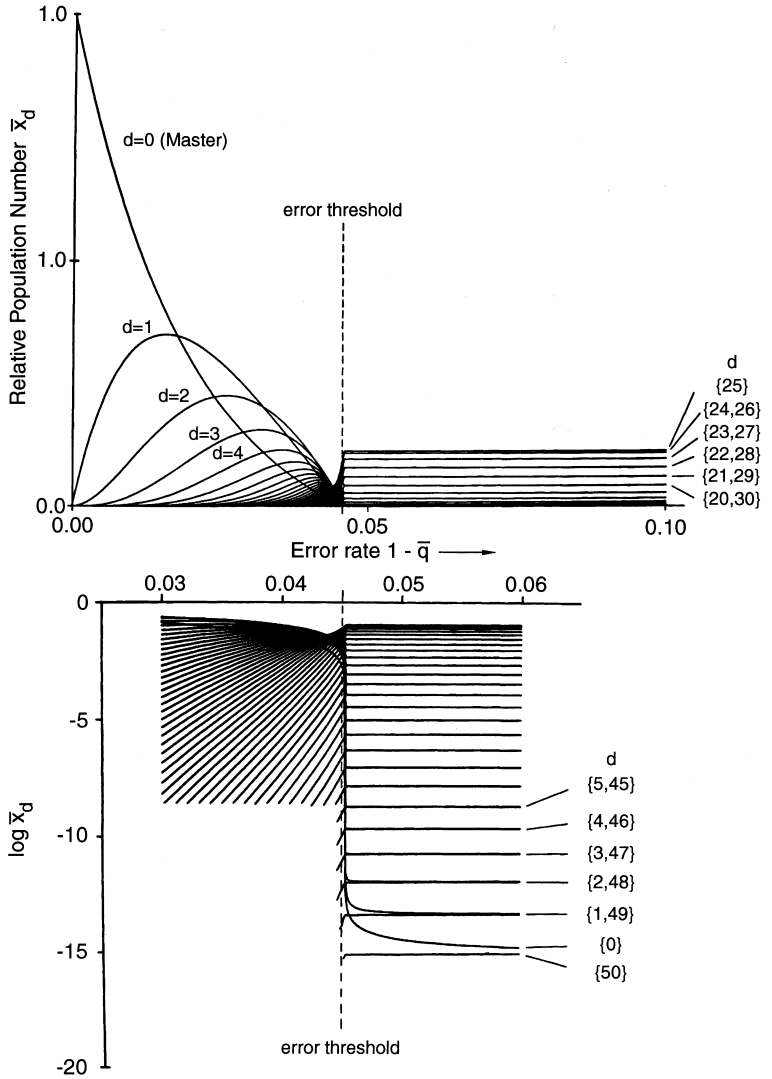

Fig. 8. Quasi-species as a function of average single-digit error rate $(1-\bar{q})$ for chain length $\nu = 50$. Transition is very sharp at this chain length already. We show a linear plot (upper) and, in order to demonstrate sharpness even more clearly, a logarithmic plot (lower) of relative concentrations around the critical point $q = q_{\min}$. (For further explanation cf. text.)

copies, weighed by their binomial coefficients (i.e. $1/2$, $1/6$.....).

All curves eventually converge to a threshold value of error rate (cf. below) of $\frac{\ln 10}{50} \approx 0.046$. The numbers refer to the selective superiority of the master, i.e. its 10-fold replication rate, and to the length ($\nu = 50$) of the sequences. Although individual master copies are already quite rare near the error threshold, the consensus sequence of the whole mutant distribution remains that of the master up to close to the error threshold [34,35]. Its definition requires sufficiently large population numbers because, despite conservation of the average (1st moment) the variance (2nd moment) increases steadily toward the threshold. The main and very drastic change of the master and mutant population occurs exactly at the threshold value, as is seen best by the

logarithmic plot. The relative population number of the master drops abruptly from fractions of 1% to $2^{-50} \approx 10^{-15}$. Above the threshold all mutants appear with an individual frequency of 2^{-50}. There is only one copy of each the master and its antipode with Hamming distance 50, but there is a maximum of individual 25-error copies amounting to approximately 10% of the total set. All information, which was conserved pretty well in the quasi-species distribution up to the threshold, 'evaporates' leaving nothing to reverberate once the threshold is passed.

My terminology suggests that we are dealing with a true phase transition, but we are not dealing with matter which 'evaporates' or 'condenses', it is information, a property that is associated with a material carrier, but has otherwise little to do with the material character of the carrier, in fact as little as the content of a novel is an attribute of printer's ink. In order to prove the point we have to analyze quantitatively the process using the criteria for phase transitions, derived in Section 1 of this paper. These refer to: (i) cooperativity; (ii) correlation length; and (iii) large scale fluctuations.

3.1. Cooperativity

The sequence, as such, is treated as a cooperative unit. Formation and decomposition of the sequence are described by rate parameters which refer to them as individual units. The analogue of cooperativity which might extend through the whole population is of kinetic nature. The phase transition that is supposed to take place at the error threshold is a kinetic process in population space. Information is a property of distribution in information space rather than of distribution of matter in geometrical space. Autocatalytic reproduction certainly could be seen in analogy to nearest neighbor interactions due to physical forces. What are the consequences of this kind of kinetic cooperativity?

Let us rewrite the vectorial Eq. (5) for a single component:

$$\dot{x}_m(t) = \big(W_{mm} - \bar{E}(t)\big)x_m(t) + \sum_{k \neq m} W_{mk}x_k(t) \tag{10}$$

where W_{mk} is defined by Eq. (6). Let us, in particular, assume that the index 'm' refers to the master species which in Fig. 8 was denoted by the index o (Hamming distance zero). The average $\bar{E}_{(t)}$ contains a major term $E_m x_m$ which becomes dominant towards completion of selection. With

$$\sum_{k \neq m} x_k = 1 - x_m$$

we obtain:

$$\bar{E}(t) = \bar{E}_{k \neq m} + x_m\big(E_m - \bar{E}_{k \neq m}\big) \tag{11}$$

Likewise, the last term describing a 'backflow' from mutants 'k' producing 'm' by erroneous copying can be expressed as an average:

$$\sum_{k \neq m} W_{mk}x_k = \bar{W}_{mk}(1 - x_m) \tag{12}$$

In the case of the example in Fig. 8 the terms in a given error class defined by Hamming distance d are equal but different from terms in other error classes. Eq. (10) then can be written as:

$$\dot{x}_m(t) = \bar{W}_{mk} + \big(W_{mm} - \bar{E}_{k \neq m} - \bar{W}_{mk}\big)x_m(t) - \big(E_m - \bar{E}_{k \neq m}\big)x_m^2(t) \tag{13}$$

showing its true non-linear form expressed by the non-negligible square term and the x_m-dependence of $\bar{W}_{mk}$. The backflow term $\bar{W}_{mk}$ is exceedingly small as compared to all other terms, involving the (small) error rate $(1 - \bar{q})$ for individual one-error mutants and kth powers of this term for k-error mutants. The term approaches zero for large ν-values, but nevertheless will turn out to be important for the discussion of phase transitions.

In this paper we do not consider the time courses, but rather ask what happens for $t \to \infty$ when $\dot{x}(t) \to 0$. We consider now the selection of

the master sequence 'm' and its mutant 'clan'. Putting Eq. (13) to zero yields a quadratic equation with the solution [27]:

$$\bar{x}_m = 1/2\left\{a - b + \sqrt{(a-b)^2 + 4b}\right\} \tag{14}$$

with

$$a = \frac{W_{mm} - \overline{E}_{k\neq m}}{E_m - \overline{E}_{k\neq m}} \quad \text{and} \quad b = \frac{\overline{W}_{mk}}{E_m - \overline{E}_{k\neq m}} \tag{15}$$

The b-term has been found to be exceedingly small, reaching 0 for $\nu \to \infty$. If it could be neglected altogether we would be left with

$$\bar{x}_m = a = \frac{\sigma_m Q_{mm} - 1}{\sigma_m - 1} \tag{16}$$

Here I introduced the terms σ_m and Q_{mm} with the definition

$$\sigma_m = \frac{A_m}{\overline{E}_{k\neq m} + D_m}$$
$$\text{(note that } W_{mm} = A_m Q_{mm} - D_m) \tag{17}$$

while Q_{mm} was known to us from the error matrix:

$$Q_{mm} = \bar{q}^{\nu} \approx e^{-\nu(1-\bar{q})} \tag{18}$$

σ_m is called the superiority of the master m, in the above example $\sigma_m = 10$. As is obvious from Eq. (16) $\sigma_m Q_{mm}$ must be larger than 1 for selection, yielding the important error threshold relation:

$$(1-\bar{q}) \leq \frac{\ln \sigma_m}{\nu} \tag{19}$$

At the error threshold $\sigma_m Q_{mm}$ becomes one and since $\bar{x}_m$ cannot become negative we have an abrupt change for $\bar{x}_m = 0$. But how small is zero?

If 'a' becomes 0 we cannot neglect anymore 'b'. In fact Eq. (14) never truly becomes 0 except for $\nu \to \infty$.

The $b \to 0$ approximation at $a = 0$ reads:

$$\bar{x}_m = \frac{b}{2}\left(\sqrt{1 + \frac{4}{b}} - 1\right) \tag{20}$$

which approaches 0 with $\sqrt{b}$ for $\nu \to \infty$ or $(1-\bar{q}) \to 0$ and, as Eq. (14) shows, even for $a < 0$, that x_m always remains positive.

However, b is not a constant, it rather changes drastically at the transition where x_m decays to (correspondingly) small values. Does this change involve an instability which may explain the 'phase transition'?

3.2. Correlation length

We remember from Section 1 that a linear Ising model as applied to helix-coil transitions in linear biopolymers does not yield an infinite correlation length in the thermodynamic limit. The correlation length there is related to the square root of the cooperativity parameter σ^{-1} which for sufficiently long polymers is independent of length and never gets infinite. The situation is different in the case of the above model where transformation curves can become indefinitely sharp. Cooperativity builds up with increasing chain length.

In Fig. 9 two examples of low chain length, $\nu = 5$ and $\nu = 10$, are presented, again simulations by Schuster et al. [33]. The two pictures speak for themselves. While for $\nu = 5$ the quasi-species is not well defined showing no real 'error catastrophe', for $\nu = 10$ the phase separation clearly emerges. If we compare with the curves in Fig. 8 for $\nu = 50$ we realize how quickly cooperativity and phase transition-like behavior build up. For chain lengths of the order of magnitude of 1000 (i.e. genes) the master population variable $\bar{x}_m$ at the transition point may change by factors reaching almost $2^{1000} \approx 10^{300}$.

In order to understand this behavior we have to go back to the rate equations. What is plotted in Figs. 8 and 9 is the stationary outcome of selection for $\dot{x}_i(t) = 0$ as a function of error rate. The sharpness of the curve especially in the range $(1-\bar{q}) = \ln \sigma_m/\nu$ suggests some instability. The

term, linear in x_m, becomes 0 if $W_{mm} = \overline{E}_{k \neq m} + \overline{W}_{mk}$. Then x_m for $t \to \infty$ is exactly equal to

$$\left(\frac{\overline{W}_{mk}}{E_m - \overline{E}_{k \neq m}} \right)^{1/2}.$$

However, $\overline{W}_{mk}$ is not a constant; it has dropped already to a quite small value. If W_{mm} as a function of $1-q$ decreases further towards the error threshold: $W_{mm} = \overline{E}_{k \neq m}$, $\overline{W}_{mk}$ undergoes a dramatic change (approaching 0 for $\nu \to \infty$). Hence, it is the change of $\overline{W}_{km}$ which is mainly responsible for the phase transition-like behavior, which is also suggested by the normal mode form of Eq. (9). The 'correlation length' becomes truly infinite only in the 'thermodynamic limit' $\nu \to \infty$; but realistically the transition is perfectly sharp at lengths between 100 and 1000. This situation is analogous to that of phase transitions in macroscopic matter. Note that here we are dealing with a deterministic model, which still has to be complemented by stochastic treatment.

Is there a physical model for this non-material type of phase transition? Fig. 9 suggests a close analogy, namely to the behavior of ferro-magnets or antiferromagnets around Curie- or Neél-Temperature, respectively. Both diagrams in Fig. 9 show a variation of error rate $(1-\bar{q})$ from 0 to 1. Error rate zero means that each symbol is precisely copied as such. Error rate one in a binary

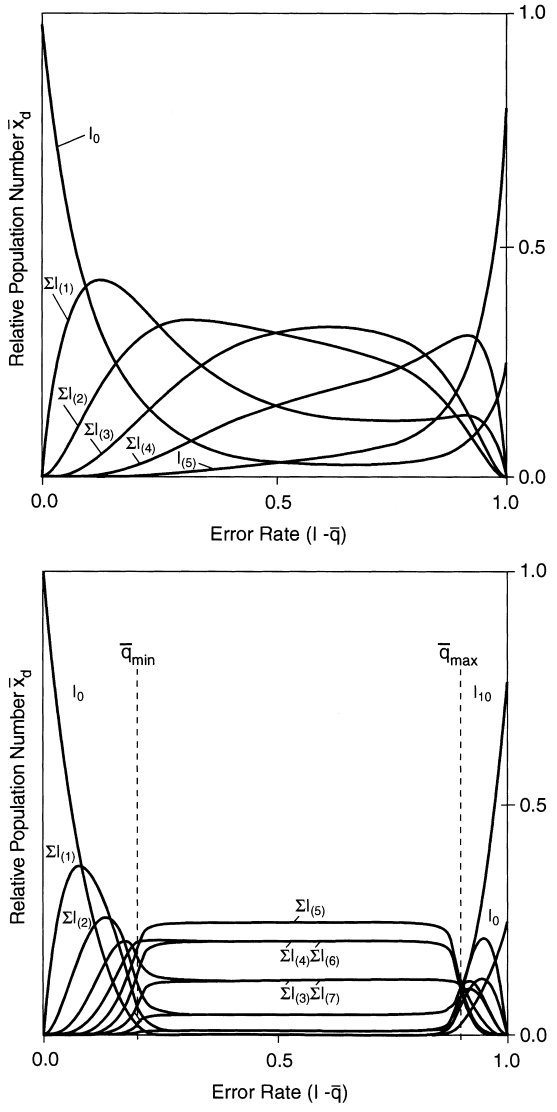

Fig. 9.

Fig. 9. Upper: Quasi-species as a function of single-digit error rate $(1-\bar{q})$ for chain $\nu = 5$. There are altogether $2^5 = 32$ sequences (named $I_0 \ldots I_{31}$). We plot relative stationary concentrations of master sequence $(\bar{x}_0)$, sum of relative stationary concentrations of *all* one-error mutants $(\bar{x}_1)$, of all two-error mutants $(\bar{x}_2)$, etc. Note that we have only one five-error mutant $I_{(5)} = I_{31}$ in this particular example. We observe selection of the master sequence at $(1-\bar{q}) = 0$. Then relative concentration of the master sequence decreases with increasing $(1-\bar{q})$. At $(1-\bar{q}) = 0.5$ all sequences are present in equal concentrations. Hence, sums of concentrations of two- and three-error mutants are largest — they have statistical weight of 10 — those of the one- and four-error mutants are half as large — they have statistical weight of 5 — and finally master sequence I_0 and its complementary sequence, the five-error mutant I_{31}, are present in relative concentration of 1/32 only. At $(1-\bar{q}) = 1$ we have selection of a complementary 'master pair', which consists of I_0 and I_{31}. Thus, we have direct replication with errors in the range $0 < (1-\bar{q}) < 0.5$ and complementary replication with errors in the range $0.5 < (1-\bar{q}) < 1$. Rate constants chosen as $A_0 = 10[t^{-1}]$ and $A_k = 1[t^{-1}]$ for all mutants $k \neq 0$. Here we denote arbitrary reciprocal time unit by $[t^{-1}]$. All degradation rate constants were put equal: $D_0 = D_1 = D_2 = \cdots = D_{31} = 0$. Lower: Quasi-species as a function of single-digit error rate $(1-\bar{q})$ for chain length $\nu = 10$. Computations were performed in complete analogy to those shown in the upper figure. Note that the range of 'random replication' has increased substantially compared to the case $\nu = 5$. We observe fairly sharp transitions between direct and random replication at critical value $\bar{q} = \bar{q}_{\min}$ and between random and complementary replication at $\bar{q} = \bar{q}_{\max}$.

system, however, means that each symbol is precisely copied into its complement. This is the way nucleic acids work, i.e. by complementary reproduction. Maximum uncertainty prevails at $\bar{q} = (1 - \bar{q}) = 0.5$. Accordingly we have two critical error thresholds, one for self- and one for complementary copying.

Now consider magnetic crystals where the lattice points have a binary choice: 'spin up' or 'spin down'. Some materials have a lower free energy when spins line up in the same direction, we call them ferro-magnets, others prefer to the anticonfiguration of neighbored spins, they are called antiferromagnets. If we call 'spin up' 0 and 'spin down' 1, we have a cooperative binary model. As was mentioned in Section 1 a three-dimensional Ising-model has not yet been worked out up to date, but in 1944 Lars Onsager came up with an exact solution for the two-dimensional model in the absence of external fields, showing that such a model, contrary to Ising's conclusion, can show true phase transitions.

The selection model discussed above corresponds to a two-dimensional Ising model. In the spin model we have horizontal and vertical lines of spins, interacting with their nearest neighbors. In the molecular selection models we have as horizontal lines the sequences behaving as units that are characterized by a particular W_{mm}-value for their (effective) rate of reproduction. The vertical line represents the sequence of generations with their 'interaction' brought about by auto- or cross-catalytic reproduction. The analogy has been worked out independently by Leuthäusser [36,37] in her thesis at Göttingen and by Demetrius [38,39] at his visit to the laboratory of Peter Schuster. It is discussed in detail in the quasi-species paper [27], to which I refer for further discussion.

One of the results of this striking analogy is the assignment of variables that correspond to one another. For instance the quantity $\ln\frac{(1-q)}{q}$ is analogous to reciprocal temperature. Both $(1 - \bar{q})$-values — close to 0 and close to 1 — mean T-values close to 0, while $(1 - \bar{q}) \approx \bar{q} \approx 0.5$ (i.e. $\ln 1 = 0$) means high temperatures. Similarly, the logarithm of the dominant eigenvalue corresponds to free energy. Since according to Ehrenfest the derivatives of free energy define the order of the phase transition (not precisely correct for continuous critical phase transitions, cf. Goldenfeld [16]), we might even think of selection being a phase transition of first order, since differentiation of λ with respect to $(1 - \bar{q})$ yields finite values. However, perhaps, we overexert the analogies among those models at the present stage. Similarly, any calculation of rate constants from sequence, in a similar way as magnetization is related to spin sequences, is far beyond our power. The selection model is of a phenomenological nature and at best comparable to a deterministic mean field approach.

3.3. Large scale fluctuations

The third criterion of continuous critical phase transitions is the appearance of fluctuations at all scales of lengths. The length scale in our model is the Hamming distance. Fluctuations to be expected would refer to populations of mutants characterized by their mutual Hamming distances. In order to calculate fluctuations of population numbers (and hence fluctuations of rates) we would require a stochastic theory of the model which exists only for limiting cases. What would we expect? We must direct our attention not only to errors which occur on a stochastic basis, but also to the other important parameter controlling selection, i.e. the superiority parameter σ_m describing the relative efficiency of the master species (or its ensemble analogue). The idea behind it is that microscopic fluctuations can build up to a macroscopic scale in the absence of a process that responds to the amplitude of the fluctuation in a counteracting way. One might conjecture that in an autocatalytic network fluctuations are automatically amplified which, however, under flux control at steady state conditions is not the case. The negative quadratic term in Eq. (13), which responds to the sign of fluctuation, counteracts the build-up of fluctuations, and is quite effective when σ_m is large. However, when σ_m approaches 1, in other words if the total distribution contains many neutral or nearly neutral mutants there is no immediate fluctuation

control. Hence, large scale fluctuations may be expected near the borderline where selection disappears because of $\sigma_m \to 1$. I shall return to such a case in Section 4. If the neutrals can isolate themselves against the rest of less efficient competitors the fluctuations will be more or less restricted to the 'exclusive club' of neutrals. This is the basis of Kimura's [40] theory of 'neutral drift'. Kimura's model is based on stochastic theory using either the Fokker–Planck equation or Kolmogorov's forward and backward equations. Appearance of neutrals at a macroscopic scale is independent of the population size which uncovers interesting clues on the evolution of higher species beyond Darwinian selection.

4. Experimental evidence of the error threshold

Theorizing in biology is of value only if it leads to predictions that can be tested experimentally. The model of the quasi-species has been subjected to various tests with such a success that the term quasi-species has almost become a common place in virology, though sometimes not quite with the precise meaning originally given to it. The nature of the quasi-species had been independently found by Weissmann and coworkers [41,42] in their experiments with phage Q_β who found in cloning experiments that the wild-type, identified by the consensus sequence, as an individual represents only a minor fraction of the total population. Domingo and Holland [43] extended the work to foot and mouth disease virus and demonstrated the lability of the wild type at the error threshold. Loeb [44] tested the hypothesis that viruses might disintegrate if 'driven' across the error threshold. The feasibility of this method as a tool in treating virus diseases [45] has not yet been fully explored. More systematic studies of

Table 1
Error rates and genome sizes of RNA viruses as compared to autonomous organisms[a]

Virus	Genome size: ν (no. of nt or bp)	Error rate: $(1-q)$ (per replication round and per nt)	Error rate $\nu(1-q)$ (per replication round and per genome)
RNA			
Bacteriophage Q_β	4200	3×10^{-4}	1.3
Polio-1 virus	7400	3×10^{-5}	0.2
Vesicular stomatitis virus	11 000	1×10^{-4}	1.1
Foot and mouth disease virus	8400	1×10^{-4}	0.8
Influenza-A virus	14 000	6×10^{-5}	0.8
Sendai virus	15 000	3×10^{-5}	0.5
HIV-1 (AIDS virus)	10 000	1×10^{-4}	1.0
Avian myeloblastosis virus	7000	5×10^{-5}	0.4
DNA			
Bacteriophage M13	6400	7×10^{-7}	4.6×10^{-3}
Bacteriophage λ	48 500	8×10^{-8}	3.8×10^{-3}
Bacteriophage T4	166 000	2×10^{-8}	3.3×10^{-3}
E. coli	4.7 000 000	7×10^{-10}	3.3×10^{-3}
Yeast (*Saccharomyces cerevisiae*)	13.8 000 000	3×10^{-10}	3.8×10^{-3}
Neurospora crassa	41.9 000 000	1×10^{-10}	4.2×10^{-3}
Human	3 000 000 000	$\sim 10^{-12}$	$\sim 3 \times 10^{-3}$

[a] The Table is reproduced from Eigen [47] where references for individual data are given. Systematic study of mutation rates during the past 20 years is reported in papers by Kunkel and Drake.

quasi-species structures have been carried out by Christof Biebricher and are described in a number of papers (for a review cf. [46]).

What I wanted to focus upon in this section is the error threshold relation as such, as it was specified by Eq. (19). Table 1 comprises some data from literature referring specifically to the error rate $(1-\bar{q})$. The authors who provided the data have emphasized the precautions required to arrive at reliable results. Fidelity of replication may be quite variable at different positions of a sequence. The average symbol copying 'quality' $\bar{q}$ that appears in measured values is a geometric mean taken over all positions of a sequence studied. The data in Table 1 reveal two major results:

1. For RNA-viruses the values of $(1-\bar{q})$ are close to the reciprocal sequence length, such that $(1-\bar{q})\nu \approx 1$. This means that σ_m is clearly larger than 1, mainly varying between 1.2 ($\ln \sigma_m \approx 0.18$) and 4 ($\ln 4 \approx 1.4$).
2. For DNA the products $\nu(1-\bar{q})$ are lower by 2–3 orders of magnitude. This might mean a σ_m-value close to 1 (i.e. for $\sigma_m = 1+\varepsilon$ and $\varepsilon \ll 1$: $\ln \sigma_m = \varepsilon$). There are certainly arguments that suggest for longer DNA-sequences a larger number of neutrals and hence σ_m-values closer to 1, but in addition there is a principal difference resulting from the particular replication mechanisms.

The most straight forward mechanism would be that for replication of a single-stranded palindromic RNA as a template. It does not involve error correction because a single strand cannot tell the enzyme what is right or wrong. Since the template is complementary to the product of replication the simplest case is that of a palindromic strand which yields 'vectorially' identical plus and minus strands (RNA-replication proceeds from the 3′ to the 5′ end of the template). Here the error threshold relation should hold in its simple form of Eq. (19).

A slight (formally not distinguishable) modification is required if the strands are not palindromic and replication for plus and minus strand differ. It changes the interpretation of σ_m, which now involves a (geometric) mean of the rate constants for plus and minus strand replication, usually leaving $Q_{ii} = \bar{q}^{\nu}$ unchanged. If replication rates have been adapted to maximal values the optimum is for plus and minus strand to have (approx.) equal rate constants. However, we know cases, e.g. an in-vitro evolutionary experiment we have carried out at our laboratory [47,48], where the selection pressure requires dissymmetric amplification of both strands. In the experiment referred to the dissymmetry in rate parameters reached a factor of 10^2.

Other cases, such as DNA-dependent RNA polymerization, reverse transcription, and also in vitro amplification such as effected by 'polymerization chain reaction' (PCR) or 'self-sustained sequence replication' (3SR) can be formally reduced to the above cases. Reidys, Forst and Schuster [49] considered another, more specific case applying to selective neutrality where random drift is present among neutral genotypes, that encode an optimal phenotype. It leads to a modification of the error threshold relation where Eq. (19) represents only one limiting case. It shows, again, the importance of fluctuations in 'neutral networks', requiring attention for a number of practical applications.

The most interesting case to be discussed with respect to Table 1 is that of double-stranded DNA replication. Unlike RNA replication, where the template strand remains unchanged, DNA synthesis is 'semiconservative'. This means that the original double strand, serving as the template, disappears in the process in favor of two new strands, one containing the parental plus strand, the other the parental minus strand. Since the polymerase always proceeds from the 3′ to the 5′ ends of the unwound template strand, the process is dissymmetric with respect to both the plus and the minus strand. The parental plus strand along which the polymerase can move continuously from 3′ to 5′ towards the (moving) replication fork is called the 'leading' strand. (I shall use for its daughter double strand the index '+'.) The other strand, where the polymerase has to move away from the (moving) replication fork, is called the lagging strand. Its daughter strand (index '−') can only be synthesized in the form of relatively short fragments, called Okasaki frag-

ments. These fragments, which vary in lengths from few hundreds (eukaryotic DNA) to thousands (prokaryotic DNA) of bases, have to be ligated, which means another source of errors.

Hence, we have the situation that DNA replication produces two daughter double strands with different error rates while the parental double strand disappears. DNA replication involves several mechanisms of error correction, but this will not remove a principal dissymmetry of the final products [50]. We can account for this dissymmetry by replacing $Q_{mm} = e^{-\nu(1-\bar{q})}$ in Eq. (19) by $(Q_+ + Q_- - 1)$. If the dissymmetry produces a Q_+ value close to 1 it will be Q_-, which is responsible for the error threshold. If the symmetry would prevail, the sum $Q_+ + Q_-$ must still remain larger than 1, which means $Q_+ \approx Q_- > 0.5$, i.e. a much more restrictive condition than for single strand replication where for smaller Q-values survival still can be guaranteed if σ_m is sufficiently large. (Note that for $(1-\bar{q}) = 1/\nu$, Q_{mm} becomes e^{-1}.)

The result is that DNA replication indeed requires higher fidelity than RNA replication. How much of the 2–3 orders of magnitude difference in error rates, revealed by Table 1, are due to the dissymmetry of replication of the leading and lagging strand, and how much of it is caused by the closeness of σ_m to 1, has to be established. It is important in experiments that yield the error rate to ensure that overall, plus or minus error rates are clearly distinguished.

5. Concluding remarks: what is information?

What we use to call biophysics often is just physics of biological material or the dynamics of the material processes involved. Since biology for a good part is 'tinkering' with gadgets it is important to have good physical models that describe those processes. The purpose of this paper is of a different nature, namely to ask what kind of physics characterizes biology itself, and: can we describe it, that is to say 'something we do not find in inanimate matter, however complex', in terms of physics. I am asking for a 'physics of biology'. 'Information' is a term typical of biology, appearing at quite different levels of molecular organization. One aspect of information is entropy, but even at a very early level, 'information' includes something, that is related to the original meaning of the word; meaning and value for achieving a certain goal which in evolution only in retrospect becomes defined. Information originated in properties of matter but soon transcended the material basis. What I try to show is that there is a physics associated with those processes even after their material basis has become irrelevant, a physics that takes place in new spaces. Do these new terms — such as information space, value landscape, replicative motion, Darwinian selection — have real physical significance?

Physicists are used to abstract spaces. The fact that we do not have sensory organs which perceive those spaces via experience does not diminish their relevance of reflecting physical reality. If we realize that information is the organizing principle of life at its various levels, such as genetic organization, cellular differentiation with its morphogenetic consequences, or neural circuitry — we have to develop a physics of information that covers this general aspect. We are surprised that this physics functions along similar lines as the physics of energy and matter. Here we have to be willing to think in new categories. Selection is not a phase transition that takes place abruptly in conventional time and space, and yet selection in its increasingly complex physical context has brought about our brains. We find general requirements that have to be fulfilled, we find analogies to the known laws of physics — despite the fact that the 'substrate' is neither energy nor matter, but rather 'information'. There is self-organization at work, as if there were physical forces that move the matterless substrate we call information. Of course, since there are material carriers of information that have to 'move', we can always build physical models that perform in the supposed way. However, what they achieve is pretty independent of the material realization of the model. In molecular genetics it is the replicative power of DNA and RNA that enslaved the enormous capacity of protein function (I leave open who 'enslaved' whom). In living organisms there is the exploitation of the molecular mecha-

nisms of somatic and germ cells which shifted reproduction to the cellular and organismic level. In our brains it is the reproduction of firing patterns in the network of nerve cells that constitutes the ability of reflecting and creating information. These processes, albeit quite different from molecular replication, are also based on reproductive generation of the circuit activity. Yes, the material realization of the processes differs, but they are guided by common fundamental principles.

There is an important point, which is even more relevant in biology than in other branches of physics. While theory needs abstraction, abstraction cannot separate itself from experiencing reality. Our brain is an adaptive organ. It can construct theories, but only on the basis of experience. There must be an unlimited feedback between theory and experiment. In the complex reality of biology: 'pure thinking would be poor thinking'!

References

[1] M. Eigen, G. Schwarz, J. Colloid. Sci. 12 (1957) 181.
[2] G. Schwarz, On the kinetics of the helix-coil transition of polypeptides in solution, M. Mol. Biol. 11 (1965) 64.
[3] B. Gruenewald, C.U. Nicola, A. Lustig, G. Schwarz, H. Klump, Biophys. Chem. 9 (1979) 137.
[4] D. Pörschke J. Mol. Biol. 53 (1970) 123, Dissertation (Göttingen-Braunschweig) 1968.
[5] E. Ising, Z. Physik 31 (1925) 253.
[6] B.H. Zimm, J.K. Bragg, J. Chem. Phys. 31 (1959) 526.
[7] S. Lifson, A. Roig, J. Chem. Phys. 34 (1961) 1963.
[8] G. Schwarz, Biopolymers 6 (1968) 873.
[9] D. Crothers, Ber. Bunsenges. Phys. Chem. 68 (1964) 840.
[10] R. Lumry, R. Legare, W.G. Miller, Biopolymers 2 (1964) 489.
[11] G. Schwarz, Ber. Bunsenges. Phys. Chem. 68 (1964) 843.
[12] M. Eigen, Harvard Lectures (1965) (distributed notes).
[13] L. Onsager, Phys. Rev. 65 (1944) 117.
[14] J.D. van der Waals (1881), cf. Les Prix Nobel (1910)
[15] S. Barry., S. Rice, J. Ross, Physical Chemistry, John Wiley & Sons, New York, 1980.
[16] N. Goldenfeld, Lectures on Phase Transitions and the Renormalization Group, Perseus Books, Reading Mass, 1992.
[17] K.G. Wilson, Phys. Rev. B4 (1971) 3174–3184.
[18] R.W. Hamming, Coding and Information Theory, Prentice Hall, Englewook Cliffs, N.J, 1980.
[19] M. Eigen, Naturwissenschaften 58 (1971) 465.
[20] L. Brillouin, Science and Information Theory, Academic Press, New York, 1962.
[21] C.E. Shannon, W. Weaver, The Mathematical Theory of Communication, U. of Illinois Press, Urbana, III USA, 1949.
[22] E.T. Jaynes, Papers on Probability Theory and Statistical Mechanics, D. Reidel Publishing Co., , Dordrecht, Holland, 1983.
[23] J.W. Gibbs, Elementary Principles in Statistical Mechanics, Yale University Press, New Haven, Conn. USA, 1902.
[24] C.K. Biebricher, M. Eigen, W.C. Gardiner, Jr., Biologically Inspired Physics, Plenum Press, New York, 1991.
[25] C.J. Thompson, L.M. McBride, Math. Biosci. 21 (1974) 127.
[26] B.L. Jones, R.H. Enns, S.S. Rangnekar, Bull. Math. Biol. 38 (1975) 15.
[27] M. Eigen, J.S. McCaskill, P. Schuster, Adv. Chem. Phys. 75 (1989) 149.
[28] J.S. McCaskill, J. Chem. Phys. 80 (1984) 5194.
[29] D.S. Rumschitzki, J. Math. Biol. 24 (1987) 667.
[30] A.W.M. Dress, D.S. Rumschitzki, Acta Appl. Math. 11 (1988) 103.
[31] J.H. Conway, N.J.A. Sloane, Sphere Packings; Lattices and Groups, Springer Verlag, New York and Heidelberg, 1988.
[32] J.N. Kapur, H.K. Kesavan, Entropy Optimization Principles with Applications, Academic Press, Inc., Boston, 1992.
[33] J. Swetina, P. Schuster, Biophys. Chem. 16 (1982) 329.
[34] M. Eigen, R. Winkler-Oswatitsch, A. Dress, Proc. Natl. Acad. Sci. USA 85 (1988) 5913.
[35] K. Nieselt-Struwe, J. Biophys. Chem. 66 (1997) 111.
[36] I. Leuthäusser, J. Chem. Phys. 84 (1986), 1884, Dissertation (Göttingen-Braunschweig) 1987.
[37] I. Leuthäusser, J. Stat. Phys. 48 (1987) 343.
[38] L. Demetrius, J. Stat. Phys. 30 (1983) 709.
[39] L. Demetrius, P. Schuster, K. Sigmund, Bull. Math. Biol. 47 (1985) 239.
[40] M. Kimura, The Neutral Theory of Molecular Evolution, Cambridge University Press, Cambridge U.K, 1983.
[41] E. Batschelet, E. Domingo, C. Weissmann, Gene 1 (1976) 27.
[42] E. Domingo, D. Sabo, T. Taniguche, C. Weissmann, Cell 13 (1978) 735.
[43] E. Domingo, J.J. Holland, P. Ahlquist, RNA Genetics, Vol III. DRR. Press, Boca Raton, 1988.
[44] L.A. Loeb, J.M. Essigmann, F. Kazazi, Proc. Natl. Acad. Sci. USA 96 (1999) 1492.
[45] M. Eigen, Medicinal research revies 13 (4) (1993) 385.
[46] N. Rohde, H. Daum, C.K. Biebricher, J. Mol. Biol. 249 (1995) 754.
[47] M. Eigen, Gene 135 (1993) 37.
[48] G. Strunk, J. Biophys. Chem. 66 (1997), 193, Dissertation (Göttingen-Braunschweig) 1992
[49] C. Reidys, C.V. Forst, P. Schuster, J. Math. Biol. (2000), in press.
[50] M. Furusawa, K.N. Wada, H. Doi, S.I. Tanaka, Y. Wada, Proc. Natl. Acad. Sci. USA 90 (1993) 11934.