POLYNUCLEOTIDE EVOLUTION AND BRANCHING PROCESSES

Lloyd Demetrius*
Peter Schuster**
Karl Sigmund***

August 1984
CP-84-36

*Max-Planck-Institut für Biophysikalische Chemie,
Göttingen, BRD
**Institut für Theoretische Chemie und Strahlenchemie
der Universität Wien, A-1090 Wien, Austria
***Institut für Mathematik der Universität Wien,
A-1090 Wien, Austria

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
2361 Laxenburg, Austria

PREFACE

In this paper the authors are concerned with the upper bound on the length of genomes imposed by the error rate (the frequency of inaccurate replication) of nucleotides. This limiting law was originally derived from a deterministic chemical model by Eigen. The authors make a connection between these deterministic chemical equations and the theory of multitype branching processes in order to study certain properties of the Eigen kinetic equations and to generalize the error threshold criteria. This generalization is based on a criterion for the extinction of branching processes.

This research was carried out as part of the Feasibility Study on the Dynamics of Macrosystems in the System and Decision Sciences Program.

ANDRZEJ WIERZBICKI
Chairman
System and Decision Sciences

ABSTRACT

The theory of multitype branching processes is applied to the kinetics of polynucleotide replication. The results obtained are compared with the solutions of the deterministic differential equations from conventional chemical kinetics.

POLYNUCLEOTIDE EVOLUTION AND BRANCHING PROCESSES

Lloyd Demetrius, Peter Schuster and Karl Sigmund

1. INTRODUCTION

Allometric relations, which set limits to the growth of organisms based on certain physical laws, are very common in nature. For example, the height of trees is restricted by the strength of wood and the capacity for water transport of the trunk, the size of insects is restricted by the rate of oxygen transport through capillary diffusion, and the body weight of vertebrates is limited by the carrying capacity of the skeleton. We are concerned here with a closely related limitation at the molecular level: an upper bound on the length of genomes imposed by the error rate (the frequency of inaccurate replication) of nucleotides. This limiting law was derived from a deterministic chemical kinetic model (Eigen, 1971; Eigen and Schuster, 1979) and is based on the relation between the rate of production of accurate replicas of molecules and the mean total productivity. This paper describes a connection between the deterministic chemical equations and the theory of multitype branching processes. We study this connection, in particular the matrix of mean values of the branching process, in order to look at certain properties of the Eigen kinetic equation and to generalize the error threshold criteria*. This generalization is based on a criterion for the extinction of branching processes.

Eigen (1971) postulated a formal, phenomenological kinetic equation

$$\frac{dx_i}{dt} = \dot{x}_i = \sum_j w_{ij} x_j - x_i \varphi , \qquad i,j = 1,\ldots,m \tag{1}$$

to describe the evolution of a population of replicating units under the idealized experimental conditions of a dialysis reactor (see, e.g., Küppers, 1979). We shall call these units *types* and represent them by $I_1,\ldots,I_m$. We use x_i to denote

*The expectation of a replication error must always remain below a sharply defined threshold (the *error threshold*) if the information accumulated in the evolutionary process is not to be lost.

the relative concentration of type I_i:

$$[I_i] = c_i\,, \quad x_i = \frac{c_i}{\sum_j c_j}\,, \quad \sum_j x_j = 1 \quad .$$

All concentrations x_j are positive and hence the physically accessible domain of variables is restricted to a unit simplex

$$S_m = \left\{ x \in \mathbb{R}_m : 0 \leq x_i \leq 1 \;\forall i = 1,\ldots,m\,,\; \sum_i x_i = 1 \right\} \quad .$$

The flow term φ is given by

$$\varphi = \sum_{rs} w_{rs} x_s \quad .$$

The elements w_{ij} are constants which will be discussed in detail later in the paper. We shall simply note here that they are constructed from rate constants and mutation frequencies in accordance with the replication mechanism. The deterministic equation (1) has been subjected to rigorous mathematical analysis (Thompson and McBride, 1974; Jones et al., 1976; Swetina and Schuster, 1982; McCaskill, 1984a; Eigen et al., 1984), and its solutions have been obtained in terms of the eigenvalues and eigenvectors of matrix $W = \{w_{ij}\}$.

The deterministic equation (1) has a number of serious drawbacks when applied to realistic experimental systems. In the case where the replicating units are polynucleotides, the case we are basically interested in here, there are three sources of stochasticity which are of particular importance:

1. *Finite population size.* The number of potential types, i.e., the number of possible, different polynucleotide sequences, is extremely large (4^ν for a length containing ν bases). Thus, the number of potential types is far greater than the total number of molecules available in any experimental set-up or in nature. Only a tiny fraction of these sequences can possibly be present at any time t. Thus, the population size truncates the existing mutant distribution and introduces a stochastic element into the dynamics of replicating ensembles of polynucleotides. In the case of high replication fidelity, this truncation affects the many types of molecules which are present in only small numbers in the stationary mutant distribution. In cases of low replication fidelity, i.e., in systems which replicate with accuracies below the error threshold, the deterministic description given by

equation (1) fails completely. Indeed, the deterministic solution predicts that all types are present in equal amounts. This is impossible since we really cannot have less than one molecule of a given type. What we should expect, therefore, is a steadily changing population of polynucleotide sequences, with some dying out while others appear through mutations. No stationary distribution of mutants can ever exist in the real world (Swetina and Schuster, 1982).

2. *Kinetic degeneracy.* Conventional deterministic equations are unable to handle cases of kinetic degeneracy, i.e., situations in which two or more types have identical kinetic rate constants. In this case the relative concentrations of such molecules are determined by random drift (Schuster and Sigmund, 1984a,b).

3. *Complex dynamics.* Sensitive dependence on the initial conditions can give rise to a third source of stochasticity in autocatalytic systems. Although this kind of stochastic dynamics (often described as chaotic behavior) arises in some complicated networks of replication processes, it does not occur with equation (1), and hence will not be discussed any further here.

The formal description of chemical reactions by stochastic processes has a long tradition (see, for example, the review by McQuarrie, 1967). More recently, new analytical techniques for the study of chemical master equations have become available and this fact has revived interest in stochastic approaches to biochemical reaction systems. Equation (1) is essentially multi-dimensional, and this makes any analysis of the corresponding master equation particularly difficult (Ebeling and Feistel, 1977). Some attempts to study the master equation of an ensemble of replicating polynucleotides under rather radical simplifying assumptions have been made by Jones and Leung (1981), Heinrich and Sonntag (1981), and Schuster and Sigmund (1984a). Inagaki (1982) reported a study of replication with random mutations using a Langevin-type equation. So far the only error threshold relation derived from an underlying stochastic model was obtained in a very recent study by McCaskill (1984b), which however makes several drastic approximations.

By contrast, the approach described here adopts a less general but very powerful method: the theory of branching processes. This theory, which was originally developed to deal with the extinction of family names, has been applied to a great variety of physical and biological problems since the forties.

The mathematical background can be found, for example, in Harris (1961), Athreya and Ney (1978) and Jagers (1975); the main properties are summarized in Section 3. We shall apply this concept to polynucleotide replication and relate branching processes to the deterministic equation (1). In particular, we shall analyze the "freezing in" of fluctuations which makes the results of the deterministic model so reliable. Section 4 is concerned with the probability of extinction, and the error threshold relation is derived in a stochastic context. The original experimental set-up (Eigen, 1971) can then be broadened considerably in the light of these results. It has already been shown that the results hold for most evolving systems (Eigen and Schuster, 1979). We are now in a position to extend the theoretical predictions even further, to systems with a discontinuously changing environment. Sequential sampling or, more generally, any sequence of alternating phases of growth and sampling is amenable to a similar analysis, thus validating the threshold relation for conditions close to those under which molecular evolution occurs in nature. The paper concludes with a discussion of "complexity", which we interpret in a different way to the concept of algorithmic complexity used, for example, by Ebeling and Jimenez-Montano (1980). The notion of complexity discussed here is an entropy-based invariant which describes the frequency with which there are mutations back to the "wild type" (see later). The complexity parameter, like the extinction parameter on which our threshold criterion is based, is a function of the mean value matrix of the branching process.

2. POLYNUCLEOTIDE REPLICATION AS A MULTITYPE BRANCHING PROCESS

If we try to describe polynucleotide replication in terms of elementary step kinetics we obtain an exceedingly complex reaction network (Biebricher et al., 1983). However, in many cases we can dispense with most of the details as long as we retain certain important steps in a simplified reaction mechanism (see, for example, Gassner and Schuster, 1982). The basic features of selection and evolutionary optimization can be derived from a crude dynamical model which represents the whole polymerization process as one single reaction step. In this simplified model it is only necessary to distinguish between faithful replication and mutation.

2.1 Discrete-time branching processes

2.1.1 *Transition probabilities*

Consider a population consisting of m types of polynucleotides, $I_1,\ldots,I_m$. Each polymer of type I_i can generate polymers of the same type ($I_i \rightarrow 2I_i$) by faithful replication or polymers of different types ($I_i \rightarrow I_i + I_j$) by false replication, i.e., mutation. Molecule replications are assumed to be homogeneous in time and mutually independent. We shall first assume that the molecules exist in discrete generations. In every generation, each polymer of type I_i produces r_1 polymers of type I_1, r_2 polymers of type I_2 and so on up to r_m polymers of type I_m, with probability $P_i(r_1,\ldots,r_m)$.

Let $Z_i(n)$ denote the total number of polymers of type I_i in generation n, where the vector $\mathbf{Z}(n) = \{Z_1(n),\ldots,Z_m(n)\}$ is a random variable.

In order to illustrate the transition law for the stochastic process we shall take $\mathbf{Z}(0) = \mathbf{e}_i$, where $\mathbf{e}_i = (0,\ldots,1,\ldots,0)$ is the unit vector in the direction of type I_i. This implies that the population consists of a single polymer of type I_i at time $n = 0$. In this case the probability generating function of $\mathbf{Z}(1)$

$$F^{(1)}(\mathbf{s}) = F^{(1)}(s_1,\ldots,s_m) = \sum_{Z_1,\ldots,Z_m=0} P^{(1)}(Z_1,\ldots,Z_m)\, s_1^{Z_1} \cdots s_m^{Z_m} \quad , \tag{2}$$

where

$$P^{(n)}(Z_1,\ldots,Z_m) = \text{Prob}\,\{Z_1(n) = Z_1,\ldots,Z_m(n) = Z_m\} \quad , \tag{3}$$

is of the simple form

$$F^{(1)}(\mathbf{s}) = f_i(\mathbf{s}) = \sum_{r_1,\ldots,r_m=0} P_i(r_1,\ldots,r_m)\, s_1^{r_1} \cdots s_m^{r_m} \quad . \tag{4}$$

We may make the following generalization: if $\mathbf{Z}(n) = (Z_1,\ldots,Z_m)$ represents the distribution of polymers in generation n, then $\mathbf{Z}(n+1)$ is the sum of $Z_1 + \cdots + Z_m$ independent random vectors of which a number Z_1 have generating function f_1, Z_2 have generating function f_2, and so on. We may thus dispense with the explicit formula which is rather lengthy and not very informative.

2.1.2 *The mean value matrix*

For reasons which are physically obvious*, we assume that first moments

$$m_{ij} = \mathrm{E}\{Z_i(1) \mid \mathbf{Z}(0) = \mathbf{e}_j\} \tag{5}$$

exist for all i and j. Thus, m_{ij} is the mean number of polymers of type I_i derived from a polymer of type I_j within one generation. In terms of generating functions we have

$$m_{ij} = \left[\frac{\partial f_j}{\partial s_i}\right]_{s_1 = \cdots = s_m = 1}, \qquad i,j = 1,\ldots,m \quad . \tag{6}$$

We are clearly dealing with non-negative first moments $m_{ij} \geq 0$. Unless otherwise stated, we shall assume that the matrix $M = \{m_{ij}\}$ is positively regular, i.e., there exists an $n > 0$ such that M^n has strictly positive elements. This implies that M is irreducible: each type I_i can be derived from every other type I_j by a series of mutations. (Mutation models which consider only point mutations, such as that analyzed by Swetina and Schuster (1982), are generally based on non-zero probabilities of mutation over a sufficiently large number of generations. In more sophisticated models which include deletions and insertions it might be advantageous to have disjoint sets of types.)

According to the Perron–Frobenius theorem (see, e.g., Karlin, 1974), the matrix M has a unique eigenvalue $\lambda > 0$ which is dominant in the sense that $|\mu| < \lambda$ for every other eigenvalue μ of M. The eigenvalue λ is non-degenerate (or simple): there exist right and left eigenvectors, denoted by $\mathbf{u}$ and $\mathbf{v}$, respectively, where $u_i > 0$ and $v_i > 0$ for all $i = 1,\ldots,m$, such that

$$M\mathbf{u} = \lambda\mathbf{u} \quad \text{and} \quad \mathbf{v}M = \lambda\mathbf{v} \quad . \tag{7}$$

Both eigenvectors are normalized in a special but very useful manner:

$$(\mathbf{u}, \mathbf{v}) = 1 \quad \text{and} \quad \sum_i v_i = 1 \quad . \tag{8}$$

The matrix $T = \{t_{ij} = v_i u_j\}$ is idempotent, $T^2 = T$, and in addition we have

$$TM = MT = \lambda T \quad \text{and} \quad \lim_{n \to \infty} \lambda^{-n} M^n = T \quad . \tag{9}$$

*In real systems we always deal with finite populations in finite time and in this case the expectations do not diverge.

No other eigenvalue μ of M is associated with an eigenvector whose components are all strictly positive.

2.1.3 *Probabilities of extinction*

A population is said to become extinct if $\mathbf{Z}(n) = 0$ for some $n > 0$. Let q_i denote the probability of this event given the initial condition $\mathbf{Z}(0) = \mathbf{e}_i$:

$$q_i = \text{Prob}\{\exists n \text{ such that } \mathbf{Z}(n) = \mathbf{0} \mid \mathbf{Z}(0) = \mathbf{e}_i\} \quad . \tag{10}$$

The vector $\mathbf{q} = (q_1,\ldots,q_m)$ is given by the smallest non-negative solution of the equation

$$\mathbf{f}(\mathbf{q}) = \mathbf{q} \quad , \tag{11}$$

where $\mathbf{f}(\mathbf{s}) = \{f_1(\mathbf{s}),\ldots,f_m(\mathbf{s})\}$ and the $f_i(\mathbf{s})$ are given by (4).

Conditions for extinction can be formulated in terms of the dominant eigenvalue λ of M:

(i) if $\lambda \leq 1$ then $q_i = 1$ for all i and extinction is certain,

(ii) if $\lambda > 1$ then $q_i < 1$ for all i and there is a positive probability of survival to infinite time.

2.1.4 *Asymptotic frequencies*

The frequency of type I_i in generation n is a random variable defined by

$$X_i(n) = \frac{Z_i(n)}{Z_1(n) + \cdots + Z_m(n)} \tag{12}$$

provided that the denominator is non-vanishing, i.e., that the system does not become extinct.

If $\lambda > 1$, there exists a random vector $\mathbf{W} = (W_1,\ldots,W_m)$ and a scalar random variable w such that with probability 1

$$\lim_{n\to\infty} \lambda^{-n}\, \mathbf{Z}(n) = \mathbf{W} \tag{13a}$$

and

$$\mathbf{W} = w\,\mathbf{u} \quad , \tag{13b}$$

where $\mathbf{u}$ is the right eigenvector of M from (7). It follows that

$$\lim_{n\to\infty} X_i(n) = \frac{u_i}{u_1 + \cdots + u_m} \tag{14}$$

holds almost everywhere provided that the population does not become extinct.

Equation (14) asserts that the random variable $X_i(n)$ representing the frequency of type I_i converges almost surely to a constant (provided that $w \neq 0$). This asymptotic behavior of the random vector $\mathbf{X}(n)$ is in sharp contrast to that of the population distribution $\mathbf{Z}(n)$ and the total population size $Z(n) = \sum_i Z_i(n)$. Because of the autocatalytic nature of the replication process, $\mathbf{Z}(n)$ may experience large fluctuations in the initial phases which persist and even accumulate in subsequent generations (see, e.g., Schuster, 1983). In the later stages of the stochastic process the system either becomes extinct or grows very large (with probability 1). In the latter case the law of large numbers implies that fluctuations in relative concentrations will be small.

The behavior of the random variable w can be described completely by results obtained by Kesten and Stigum (1966). We have either

(i) $w = 0$ with probability 1 (15)

(which is always the case if $\lambda \leq 1$), or

(ii) $\mathrm{E}\{w \mid \mathbf{Z}(0) = \mathbf{e}_i\} = v_i$, (16)

where v_i is the i-th component of the left eigenvector $\mathbf{v}$ of M (see equation 7). A necessary and sufficient condition for (16) to hold is

$$\mathrm{E}\{Z_j(1) \log Z_j(1) \mid \mathbf{Z}(0) = \mathbf{e}_i\} < \infty \text{ for } 1 \leq i, j \leq m \quad . \tag{17}$$

This condition of finite population size clearly holds for all real populations. If the population initially consists of a single polymer of type I_i, the distribution of w displays a jump of magnitude q_i at zero and has continuous density on the set of positive numbers.

2.2 Continuous-time branching processes

2.2.1 *Transition probabilities*

The assumption of discrete generations generally applies to populations with external (and sometimes also internal) clocks which prevent the mixing of generations. Such conditions are often found in nature, e.g., in populations whose breeding periods are fixed by seasonal requirements. In chemical systems there are usually no such regulators. Indeed, if we start polynucleotide replication in an initially synchronized population the synchronization is lost within a few rounds of replication. Continuous-time multitype branching Markov processes offer an accurate description of polynucleotide replication, but one which is technically quite complicated. The basic results are similar to those obtained in the discrete case, however, and are summarized briefly below.

In the continuous-time model we suppose that, independently of the other polymers, a polynucleotide of type I_i persists for some random period of time (which has an exponential distribution and mean μ_i^{-1}) and then generates copies by replication and mutation according to a distribution whose generating function is $f_i(\mathbf{s})$. This is the case if it is assumed that in a time interval of length Δt, up to probability $o(\Delta t)$, the polynucleotide must experience one of the following:

(i) no change

(ii) it "dies off", or

(iii) it survives and produces a copy of type I_j ($j = 1,\ldots,m$).

The time-homogeneous probabilities of events (ii) and (iii) are proportional to Δt, up to some $o(\Delta t)$. As before, we let $Z_i(t)$ denote the total number of polynucleotides of type I_i at time t, and the random vector $\mathbf{Z}(t) = (Z_1(t),\ldots,Z_m(t))$ denote the distribution of types.

2.2.2 *The mean value matrix*

For physical reasons we assume once again that all the first moments

$$m_{ij}(t) = \mathrm{E}\{Z_i(t) \mid \mathbf{Z}(0) = \mathbf{e}_j\}$$

are finite, for all $t \geq 0$. The mean value matrix $M(t)$ satisfies the semigroup property

$$M(t+u) = M(t)M(u), \quad t,u \geq 0 \tag{18}$$

and the continuity property

$$\lim_{t \to +0} M(t) = Id \quad , \tag{19}$$

where Id is the identity matrix. Conditions (18) and (19) imply that there exists a matrix A such that

$$M(t) = e^{At} \quad , \tag{20}$$

for all $t \geq 0$. A is called the *infinitesimal generator*. The elements of A are given by $a_{ij} = \mu_i c_{ij}$, where $c_{ij} = b_{ij} - \delta_{ij}$ (δ_{ij} is the Kronecker delta) and

$$b_{ij} = \left[\frac{\partial f_i}{\partial s_j} \right]_{s_1 = \cdots = s_m = 1} \quad . \tag{21}$$

Again we assume that each type can give rise to all of the others. It follows that $m_{ij}(t) > 0$ for $t > 0$, and hence that A is essentially positive, i.e., $a_{ij} > 0$ for all $i \neq j$. The Perron–Frobenius theory then implies that there exists a unique real eigenvalue λ of A which is dominant in the sense that it is larger than the real parts of all other eigenvalues. The eigenvalue λ is simple (non-degenerate) and has positive right and left eigenvectors $\mathbf{u}$ and $\mathbf{v}$, which we again normalize such that $\sum v_i = \sum u_i v_i = 1$. The dominant eigenvalue of $M(t)$ is $e^{\lambda t}$, with $\mathbf{u}$ and $\mathbf{v}$ as associated eigenvectors. Taking $t_{ij} = v_i u_j$ we again have

$$\lim_{t \to \infty} e^{-\lambda t} \{M(t)\} = T \quad . \tag{22}$$

2.2.3 *Asymptotic behavior*

As in the discrete case, the extinction conditions are given in terms of λ. If, as before, $\mathbf{q} = (q_1, \ldots, q_m)$ denotes the extinction probabilities, then $\mathbf{q}$ is the unique solution of

$$\mathbf{g}(\mathbf{s}) = \mathbf{0} \quad , \tag{23}$$

where

$$\mathbf{g}(\mathbf{s}) = (g_1(\mathbf{s}), \ldots, g_m(\mathbf{s}))$$

and

$$g_i(\mathbf{s}) = \mu_i (f_i(\mathbf{s}) - s_i) \quad . \tag{24}$$

We have

(i) if $\lambda \leq 0$ then $q_i = 1$ for all i;

(ii) if $\lambda > 0$ then $q_i < 1$ for all i.

Furthermore, we once again obtain

$$X_i(t) = \frac{Z_i(t)}{Z_1(t) + \cdots + Z_m(t)} \to \frac{u_i}{u_1 + \cdots + u_m} \tag{25}$$

provided that the process does not lead to extinction.

3. EXPECTATIONS AND EIGEN'S SELECTION EQUATION

3.1 Eigen's selection equation

In Eigen (1971) and Eigen and Schuster (1979) the evolution of polynucleotides in a dialysis reactor was modeled by a differential equation of the form

$$\dot{x}_i = \sum_j w_{ij} x_j - x_i \left[\sum_{rs} w_{rs} x_s \right] \tag{26}$$

or, in vector notation (with $W = (w_{ij})$, $\mathbf{x} = (x_1,\ldots,x_m)$ and $1 = (1,\ldots,1)$),

$$\dot{\mathbf{x}} = W\mathbf{x} - (1 \cdot W\mathbf{x})\mathbf{x} \tag{27}$$

on the unit simplex

$$S_m = \left\{ \mathbf{x} \in \mathrm{R}^m : x_i \geq 0,\ \sum x_j = 1 \right\} \quad . \tag{28}$$

Here the x_i are the concentrations of polynucleotides of type I_i $(i = 1,\ldots,m)$. The coefficients w_{ij} satisfy $w_{jj} = A_j Q_{jj} - D_j$ and $w_{ij} = A_j Q_{ij}$ $(i \neq j)$, where $A_j > 0$ is the total rate constant for polynucleotide synthesis on template I_j, D_j is the decay rate and Q_{ij}, the "quality factor", gives the probability that a copy of a molecule of type I_j will be of type I_i (for $j \neq i$ this is a mutation rate). We shall first assume that W is positively regular. The term

$$\varphi = 1 \cdot W\mathbf{x} = \sum_r (W\mathbf{x})_r = \sum_{rs} w_{rs} x_s \tag{29}$$

is interpreted as an externally controlled "dilution flow" which keeps the total concentration $\sum x_i$ constant (without loss of generality, equal to 1). The

parameter φ may be viewed as the "average productivity" of the molecular population. It is easy to check that S_m is invariant under (26): if $\mathbf{x}(0) \in S_m$ then $\mathbf{x}(t) \in S_m$ for all $t \geq 0$.

Equation (26), then, was introduced as a phenomenological equation describing the kinetics of self-reproducing molecules in a dialysis reactor under the constraint of constant total population. The aim of this section is to relate this equation to multitype branching processes.

3.2 Preliminary remarks

We begin with a few simple remarks.

1. Let

$$\dot{\mathbf{y}} = W\mathbf{y} \tag{30}$$

be a linear differential equation with (essentially) positively regular W. If $\mathbf{y}(0) \in \mathrm{R}_+^m$, then

$$\mathbf{x}(t) = \frac{1}{\sum y_j(t)}\,\mathbf{y}(t) \tag{31}$$

is well defined and in S_m for all $t \geq 0$. $\mathbf{x}(t)$ is also a solution of (26).

2. It is also possible to obtain (26) from (30) by setting

$$\Psi(t) = \int_0^t \varphi(u)\mathrm{d}u \tag{32}$$

and

$$y(t) = \mathbf{x}(t)\,\mathrm{e}^{\Psi(t)} \tag{33}$$

(see Jones et al., 1976; Thompson and McBride, 1974).

3. The nonlinear equation (26) is therefore easy to solve. Any equilibrium of (26) must satisfy

$$W\mathbf{x} = \varphi\mathbf{x} \tag{34}$$

and therefore be a right eigenvector of W. There is only one such eigenvector in S_m, which is denoted by $\mathbf{u}$: the corresponding eigenvalue Φ is the dominant eigenvalue of W. From the correspondence between (30) and (26) it follows that all orbits of (26) in the state space S_m converge to $\mathbf{u}$.

4. We shall use a canonical method to link the difference equation

$$\dot{\mathbf{v}} = F(\mathbf{v}) \tag{35}$$

with the differential equation

$$\mathbf{v}' = F(\mathbf{v}) - \mathbf{v} \quad . \tag{36}$$

Such an extension to continuous time cannot always be justified, of course. But if the generation length is 1 (say) then (35), or $\mathbf{v}' - \mathbf{v} = F(\mathbf{v}) - \mathbf{v}$, implies that

$$\mathbf{v}(1) - \mathbf{v}(0) = F(\mathbf{v}(0)) - \mathbf{v}(0) \quad .$$

If the generations are not distinct, but blend into each other, then the increase $\mathbf{v}(1/n) - \mathbf{v}(0)$ in time $1/n$ is approximately $(1/n)(F(\mathbf{v}(0)) - \mathbf{v}(0))$, or

$$\frac{\mathbf{v}(\Delta t) - \mathbf{v}(0)}{\Delta t} = F(\mathbf{v}(0)) - \mathbf{v}(0)$$

which, in the limit, implies (36).

3.3 Multitype branching and the selection equation

The relationship between branching processes and the selection equation is summarized in Figure 1. If we start with a discrete multitype branching process $\mathbf{Z}(n)$, then the values of the expectation $\mathbf{Y}(n)$ satisfy $\mathbf{Y}(n) = M^n\mathbf{Y}(0)$, where M is the mean value matrix described in Section 2.1.2. Thus $\mathbf{Y}(n)$ may be obtained by iteration from the difference equation* $\mathbf{y}' = M\mathbf{y}$. This equation can be transformed into the selection equation in two ways:

(i) by first passing to continuous time, i.e., to the differential equation $\dot{\mathbf{y}} = V\mathbf{y}$ (with $V = M - Id$), and then normalizing, as in (26), which leads to

$$\dot{\mathbf{x}} = V\mathbf{x} - \mathbf{x}(1 \cdot V\mathbf{x}) \tag{37}$$

(ii) by first normalizing the difference equation, thereby obtaining

$$\mathbf{x}' = \frac{1}{1 \cdot M\mathbf{x}} M\mathbf{x} \tag{38}$$

on S_m, and then passing to continuous time, which yields

*This difference equation is similar to the discrete-time model given by Demetrius (1983a).

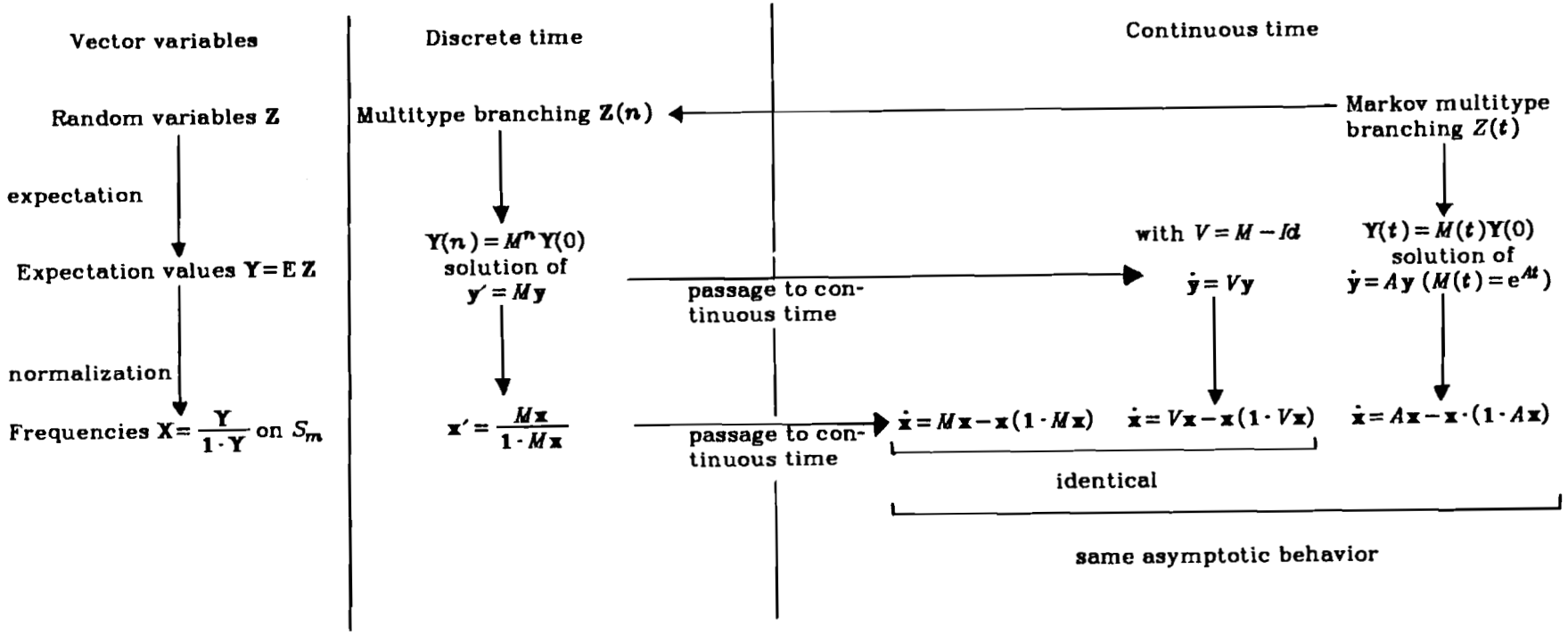

Figure 1. The relationship between branching processes and the selection equation.

$$\dot{\mathbf{x}} = (M\mathbf{x} - \mathbf{x}(1 \cdot M\mathbf{x}))\ \frac{1}{1 \cdot M\mathbf{x}} \quad . \tag{39}$$

Multiplying the right-hand side of (39) by the factor $1 \cdot M\mathbf{x}$ (which does not depend on i and is always strictly positive on S_m) corresponds to a change in velocity. The orbits of (39) are the same as those of

$$\dot{\mathbf{x}} = M\mathbf{x} - \mathbf{x}(1 \cdot M\mathbf{x}) \quad . \tag{40}$$

Since $V = M - Id$, equations (39) and (40) are identical on S_m. Both are of the same form as Eigen's selection equation.

The previous discussion took a discrete process as the starting point. If, however, we begin with a continuous Markovian multitype branching process $\mathbf{Z}(t)$, $(t \geq 0)$, we can either reduce it (by discretization) to the discrete branching process $\mathbf{Z}(n)$, or else obtain $\mathbf{Y}(t) = M(t)\mathbf{Y}(0)$ for the expectation values $\mathbf{Y}(t) = \mathrm{E}\{\mathbf{Z}(t)\}$ (where $M(t)$ is again the mean value matrix: $M(1) = M$). $\mathbf{Y}(t)$ is then the solution of the linear differential equation

$$\dot{\mathbf{y}} = A\mathbf{y} \quad , \tag{41}$$

where

$$A = \lim_{t \to 0} \frac{M(t) - Id}{t}$$

is the infinitesimal generator of the semigroup $M(t)$, and $M(t) = \mathrm{e}^{At}$. Normalization yields

$$\dot{\mathbf{x}} = A\mathbf{x} - \mathbf{x}(1 \cdot A\mathbf{x}) \tag{42}$$

on S_m. This equation generally has different dynamics to (40), but the asymptotic behavior is the same. Indeed, A and $m = \mathrm{e}^A$ have the same eigenvectors. Thus $\mathbf{u}$ is the global attractor for both (40) and (42).

3.4 The reliability of the deterministic equation

We have seen that there are three simple ways of getting from branching processes to an essentially unique version of Eigen's selection equation. There remains the question of whether such a reduction from a stochastic to a deterministic system is of any practical use. The first impression may be that it brings no obvious advantages. Indeed, going from the random variables to their

expectations can be quite misleading, because the variances grow so rapidly. This can be verified most easily for the single-type branching process. If m and σ^2 are the mean and variance, respectively, of the number of descendants of a single individual in the first generation, then the corresponding mean and variance in the n-th generation grow in the supercritical case ($m > 1$) according to

$$m^n \quad \text{and} \quad \sigma^2 \frac{m^n(m^n-1)}{m(m-1)} \quad , \tag{43}$$

so that the ratio of the dispersion (i.e., the root of the variance) to the mean converges to a positive constant. Thus the "window" of probable values of the random variable is rather large. (For a critical process, the situation is even worse: the mean remains constant and the variance increases to infinity.) The situation is similar in the multitype case. Here the variance and correlation formulae are rather complicated (see Harris, 1961, for the discrete and Athreya and Ney, 1972, for the continuous case) but the result is, once again, that the second moments grow so fast that the averages tell us virtually nothing. Normalization changes this, however. The transition from expectations to relative frequencies cancels the fluctuations. More precisely, if the process does not go to extinction, then the relative frequencies of the random variables

$$X_i = \frac{Z_i}{Z_1 + \cdots + Z_m}$$

converge almost surely to the values u_i ($i = 1,\ldots,m$). These are also the limits of the relative frequencies of the expectations

$$x_i = \frac{y_i}{y_1 + \cdots + y_m} \quad .$$

In this sense, the deterministic selection equation yields a description of the stochastic evolution process which is more reliable than that given by the dynamics of the non-normalized means. The qualitative aspects of the selection equation represent the "variance free" part of the deterministic approach.

3.5 "Freezing in" of fluctuations

We should stress here that the initial fluctuations in a supercritical branching process are "frozen in". In order to clarify what we mean by this, let us compare two symmetric random walks on a finite set $\{0,1,\ldots,N\}$ of integers,

where one of the random walks has absorbing boundaries and the other reflecting boundaries. In both cases the mean remains constant, and the variance converges to some positive value. In the absorbing case, however, the initial fluctuations play a decisive role. Sooner or later, the walk reaches a boundary and from then on remains "frozen in". In the reflecting case, on the other hand, the initial fluctuation will be "forgotten" after a sufficiently long time has elapsed. With this example in mind, we say that the initial fluctuations in a stochastic process X_n are frozen in if for every $\varepsilon > 0$ and for all $n \geq N$ we have

$$\text{Prob}\left\{\frac{\text{Var}\{X_n\}}{\text{E}\{X_n\}} < \varepsilon \mid X_N\right\} > 1 - \varepsilon \quad ,$$

provided that N is sufficiently large. In this sense the "deterministic" model (i.e., the sequence $\text{E}\{X_n\}$) is fairly reliable if we wait sufficiently long before starting observations, because by this time the fluctuations will have subsided.

It is easy to check that the above statement also holds for supercritical branching processes. It is only necessary to note that for given (large) $k > 0$, with probability $1 - \varepsilon$, either $X_N = 0$ or $X_N > k$ if N is sufficiently large. If $X_N = 0$, then $\text{Var}\{X_n\} = 0$ for all $n \geq N$; if $X_N > k$, then from (43) we have

$$\frac{\text{Var}\{X_n\}}{\text{E}\{X_n\}} < \frac{\sigma}{k} \cdot \frac{1}{m(m-1)} \quad ,$$

which is smaller than ε provided that k is sufficiently large.

4. THE ERROR THRESHOLD

The parameter λ, the dominant eigenvalue of the mean value matrix M, plays a crucial role in the branching process in that extinction is certain iff $\lambda < 1$. This relation provides both an interpretation and a generalization of Eigen's error threshold relation, as we shall presently see.

4.1 Single-type branching

Let us first consider a single type of macromolecule. In each generation, a polynucleotide yields σ copies before it "dies" by hydrolysis. Here σ is an integer-valued random variable with probability distribution

$$\text{P}\{\sigma = m\} = q_m; \qquad m = 0,1,2,... \tag{44}$$

and expectation

$$\bar{\sigma} = \sum_{m=0}^{\infty} m q_m \quad . \tag{45}$$

We shall assume that the polymer is a chain consisting of ν nucleotides and that there is a fixed probability p of a single nucleotide being copied correctly*. The assumption of some constant, single-component accuracy of replication p which is independent of the molecule type and its position in the sequence is, of course, an oversimplification. However, this assumption may be justified on physical grounds – for details see Eigen and Schuster (1979) and Schuster (1981).

In this case the probability that a given copy is exact is p^ν, and the probability that X, the number of correct copies in one generation, is equal to some integer k, is

$$P\{X = k\} = \binom{m}{k} p^{\nu k} (1 - p^\nu)^{m-k} \quad ,$$

where m is the total number of copies and $m \geq k$. The mean number of correct copies is therefore

$$\sum_{k=0}^{\infty} k P\{X = k\} = \sum_{k=0}^{\infty} \sum_{m \geq k} q_m \binom{m}{k} k p^{\nu k} (1 - p^\nu)^{m-k} =$$

$$= \sum_{m=0}^{\infty} q_m \sum_{k=0}^{m} k \binom{m}{k} p^{\nu k} (1 - p^\nu)^{m-k} =$$

$$= p^\nu \sum_{k=0}^{m} m q_m = p^\nu \bar{\sigma} \quad .$$

From the relation $\lambda \leq 1$ extinction is certain iff

$$p^\nu \bar{\sigma} \leq 1 \tag{46}$$

and hence there is a strictly positive probability of indefinite survival iff

$$\nu \leq \frac{\log \bar{\sigma}}{-\log p} \sim \frac{\log \bar{\sigma}}{1-p} \quad . \tag{47}$$

*In order to avoid confusion with the notation used elsewhere in this paper, we have chosen the letter "p" to represent single-digit accuracy of replication. In our previous publications we have generally used q for this quantity (Eigen and Schuster, 1979).

where the approximation holds when $1-p$ (the probability of an inaccurate copy) is small. For fixed $\bar{\sigma} > 1$, this means that the maximum length of a polynucleotide is inversely proportional to the probability of a replication error in one of its components. If this length is exceeded, long-run survival is impossible.

The probability of extinction q is the smallest positive solution of

$$\Phi(s) = s \quad , \tag{48}$$

where $\Phi(s)$ is the probability generating function for the random number X of correct copies, i.e.,

$$\Phi(s) = \sum_{k=0}^{\infty} P\{X = k\} s^k \quad .$$

Under the previous assumptions,

$$\Phi(s) = \sum_{k=0}^{\infty} \sum_{m \geq k} q_m \binom{m}{k} p^{\nu k} (1-p^{\nu})^{m-k} s^k$$

$$= \sum_{m=0}^{\infty} q_m \sum_{k \leq m} \binom{m}{k} (sp^{\nu})^k (1-p^{\nu})^{m-k}$$

i.e.,

$$\Phi(s) = \sum_{m=0}^{\infty} q_m [1 - p^{\nu}(1-s)]^m \quad . \tag{49}$$

4.2 Single-type branching – a variant case

In order to link this theory with current experimental work on polynucleotide replication, it is useful to introduce a slight modification. It should be recalled that the lifetime of a polynucleotide is not a well-defined constant but rather a random variable which, to a first approximation, has an exponential distribution. On the other hand, the replication time is fairly well defined, at least under appropriate boundary conditions. It is therefore convenient to view this time, rather than the actual lifetime, as the length of a generation.

Let us assume, then, that in unit time, the molecule either survives (with probability w) and produces a copy (which is accurate with probability p^{ν}), or is hydrolysed (with probability $1-w$). The survival probability w is constant if

there is no "aging" under the experimental conditions. A given molecule thus yields 0, 1 or 2 molecules of the same type after one unit of time with probabilities $1-w$, $w(1-p^\nu)$ and wp^ν, respectively. The mean is $w(1+p^\nu)$. We therefore have a non-zero probability of survival to infinite time iff equation (47) is satisfied, where the constant $\bar{\sigma}$ now denotes $w/1-w$. The probability of extinction is easily computed from (48):

$$q = \min\left[1, \frac{1}{\bar{\sigma}p^\nu}\right] \quad . \tag{50}$$

4.3 Multitype branching

So far we have considered only one type of molecule. However, the same results also hold in a multitype situation if the possibility of "back mutations" is excluded, i.e., if mutations from I_j $(j \neq 1)$ to I_1 can be neglected.

In the general case, i.e., allowing all types of mutation to occur, it can be difficult to estimate the dominant eigenvalue λ. We refer the reader to Thompson and McBride (1974) and Eigen and Schuster (1979) for some useful inequalities.

The $2^\nu \times 2^\nu$ matrix M introduced by Swetina and Schuster (1982) provides an interesting example. In a somewhat simplified version of the replication problem only two classes of components (or digits), say 0 and 1, are considered. The polymer is thus a sequence of ν such digits and, in general, we are dealing with 2^ν different sequences. We shall assume that I_j has replication rate A_j, and that the mutation rate from I_j to I_i depends only on the Hamming distance k between the two sequences – this is the minimum number of single-digit mutations needed to transform I_j into I_i. The elements of the matrix M can then be expressed by

$$m_{ij} = A_j p^k (1-p)^{\nu-k} \quad . \tag{51}$$

If $A_j = A$ $(j = 2,3,\ldots,2^\nu)$, if $A_1 \gg A$ and if $(1-p)^2$ can be neglected, then second-order perturbation theory (see Eigen and Schuster, 1979; Thompson and McBride, 1974) yields

$$\lambda \sim m_{11} + \sum_{k=2}^{2^\nu} \frac{m_{1k}m_{k1}}{m_{11}-m_{kk}} = A_1 p^\nu + \frac{1}{(A-1)p^\nu} \sum_{k=1}^{2\nu} \binom{\nu}{k} p^{2k}(1-p)^{2\nu-2k} \sim A_1 p^\nu \quad .$$

Again, the condition for positive survival probability reduces to (47).

4.4 Complementary replication

The basic mechanism of polynucleotide replication does not lead directly to copies of the templates. From the pairing rules ($G \leftrightarrow C$ and $A \leftrightarrow U$ or $A \leftrightarrow T$) for the individual nucleotides, it is clear that *complementary copies* act as intermediates. This mechanism is common in the replication of viral RNA and has been studied in great detail using kinetic methods (Biebricher et al., 1983). The two complementary polynucleotide sequences are usually called *plus strands* and *minus strands* (I^+ and I^-, respectively). We can now apply our previous theory with some slight modifications. Let $I_1^+,\ldots,I_m^+$ and $I_1^-,\ldots,I_m^-$ be the different types of plus and minus strands in the reactor. The matrix of mean values M is then a $2m \times 2m$ matrix of the form

$$M = \begin{bmatrix} 0 & U \\ V & 0 \end{bmatrix} .$$

It is easy to check that the non-vanishing eigenvalues of M are just the square roots of the non-vanishing eigenvalues of UV (or VU). Thus, the dominant eigenvalue of M is the square root of the dominant eigenvalue of UV. In particular, if replication is error-free, i.e., if

$$U = \begin{bmatrix} \lambda_1^+ & & 0 \\ & \ddots & \\ 0 & & \lambda_m^+ \end{bmatrix}, \quad V = \begin{bmatrix} \lambda_1^- & & 0 \\ & \ddots & \\ 0 & & \lambda_m^- \end{bmatrix}$$

are diagonal matrices, the dominant eigenvalue of M is

$$\max \{\sqrt{\lambda_i^+ \lambda_i^-} : i = 1,\ldots,m\} .$$

From equation (1), the deterministic rate equations for the concentrations x_i^+ and x_i^- of I_i^+ and I_i^- are given by

$$\left.\begin{aligned} \dot{x}_i^+ &= \lambda_i^+ x_i^- - x_i^+ \varphi \\ \dot{x}_i^- &= \lambda_i^- x_i^+ - x_i^- \varphi \end{aligned}\right\} \quad i = 1,\ldots,n \tag{52}$$

where $\varphi = \sum (\lambda_i^+ x_i^- + \lambda_i^- x_i^+)$.

This set of equations was first analyzed by Eigen (1971). It can easily be checked that

$$\left[\frac{x_i^+}{x_i^-}\right]^{\cdot} = \frac{\lambda_i^+ (x_i^-)^2 - \lambda_i^- (x_i^+)^2}{(x_i^-)^2} = \lambda_i^+ \left[\left[\frac{x_i^+}{x_i^-}\right]^2 - \frac{\lambda_i^-}{\lambda_i^+}\right]$$

and hence that

$$\frac{x_i^+}{x_i^-} \to \left[\frac{\lambda_i^-}{\lambda_i^+}\right]^{\frac{1}{2}} \quad . \tag{53}$$

In the limiting case $x_i^- = x_i^+ \cdot (\lambda_i^+ / \lambda_i^-)^{1/2}$ and setting $\bar{x}_i = x_i^+ + x_i^-$ we have

$$\dot{\bar{x}}_i = \bar{x}_i \sqrt{\lambda_i^+ \lambda_i^-} - \bar{x}_i \left(\sum \bar{x}_j \sqrt{\lambda_j^+ \lambda_j^-}\right) \quad ,$$

which is just the Eigen selection equation for direct, error-free copying, and leads to the extinction of all pairs I_i^+, I_i^- for which $\sqrt{\lambda_i^+ \lambda_i^-}$ is not maximal.

4.5 The deterministic error threshold

Equation (47) is very similar to the error threshold relation

$$\nu < \frac{\log \sigma_1}{-\log p} \tag{54}$$

derived for the deterministic model by Eigen (1971), Eigen and Schuster (1979) and Swetina and Schuster (1982). In this case, the molecular species I_1 is assumed to be the master sequence (which means by definition that $m_{11} > m_{ii}$ for all $i \neq 1$) and the parameter σ_1 is its *superiority*, which is defined by

$$\sigma_1 = \frac{A_1}{A_1 + \bar{E}_{-1}} \quad . \tag{55}$$

Here $\bar{E}_{-1}$ is the mean excess productivity (number produced minus number hydrolysed) of molecules other than the master sequence, i.e.,

$$\bar{E}_{-1} = \left[\sum_{i=2}^{m} x_i\right]^{-1} \left[\sum_{i=2}^{m} (A_i - D_i) x_i\right] \quad . \tag{56}$$

It is instructive to compare the derivation of (47) with that of (54). Inequality (54) is a consequence of

$$Q_{11} > \sigma_1^{-1} \quad , \tag{57}$$

where Q_{11}, the rate of accurate replication of molecules of type I_1, is again p^ν. Equation (57) is derived from

$$w_{11} > \bar{E}_{-1} \tag{58}$$

and $w_{11} = A_1 Q_{11} - D_1$.

This last inequality states that the value function, i.e., the rate of production of *accurate* copies of the master sequence I_1, is higher than the average production rate of *all* copies (accurate and inaccurate) of all other molecules.

This need not always be true. Let us consider an almost trivial but, nevertheless, illustrative example. Setting $A_1 = 3$, $A_2 = A_3 = 4$, $Q_{11} = 1$, $Q_{22} = Q_{33} = Q_{23} = Q_{32} = 1/2$, and all other Q_{ij} values and the degradation rate constants D_i equal to 0, we obtain

$$M = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{bmatrix} \quad .$$

This leads to $w_{11} = 3$ and $\bar{E}_{-1} = 4$. If the zero terms are replaced by (more realistic) small, non-vanishing terms, (58) is still violated. However, relation (58) clearly holds in physically meaningful situations when the mutation terms are small. In particular, in the limiting case where all the mutation rates vanish (and all the Q_{ii} are 1), relation (58) is an obvious consequence of $w_{11} > w_{ii}$ for $i \neq 1$, i.e., of the assumption that I_1 is the master species. It should, however, be noted that $\bar{E}_{-1}$ (and σ_1) are generally functions of x (see Swetina and Schuster, 1982).

It is natural to evaluate w_{11} and $\bar{E}_{-1}$ at the equilibrium state $\mathbf{u}$. However, in the case with no mutations we have $u_i = 0$ for $i = 2,\ldots,m$, so that E_{-1} is not properly defined. In this case we consider $\lim_{t \to +\infty} \bar{E}_{-1}$, which exists and is equal to the second largest diagonal term. Thus relation (58) also holds under these conditions.

The deterministic error threshold is based on the assumption that, under selection, the rate of production of accurate copies of a molecular species

becomes equal to the mean total productivity of all other species. The master species will always replicate with a fidelity above the error threshold provided that the mutation terms of all other species are sufficiently small. The stochastic error threshold is based on the probability of extinction. Thus, the requirement to operate above the stochastic threshold is always a stronger condition than the corresponding requirement in the deterministic case.

5. COMPLEXITY

In this section we shall introduce another parameter called the *complexity*, which, like the dominant eigenvalue, is a function of the matrix of mean values M.

5.1 The parameter H

We assume once again that the process is positively regular and write

$$p_{ij} = \frac{m_{ij} u_j}{\lambda} \quad . \tag{59}$$

The matrix $P = (p_{ij})$ is Markovian. Let $\pi = (\pi_i)$ denote the stationary distribution of the Markov chain. The *complexity* of the branching process is then defined by

$$H = -\sum_{ij} \pi_i p_{ij} \log p_{ij} \quad . \tag{60}$$

There is a simple relation between H and the dominant eigenvalue λ, which is given by

$$\log \lambda = H - \Psi \quad , \tag{61}$$

where

$$\Psi = -\sum_{ij} \pi_i p_{ij} \log \left[\frac{m_{ij} u_j}{u_i} \right] \quad . \tag{62}$$

The entropy-like parameter H defined by (60) represents the frequency of mutations back to the "wild type". The positive regularity of the process ensures that H is strictly positive.

5.2 An illustrative example

As in the case of the dominant eigenvalue λ, it is generally difficult to compute H exactly if we allow all types of mutations to occur. However, an explicit expression can be obtained from the Swetina–Schuster matrix (47) with $A_1 = A > 1$ and $A_j = 1$ for $j = 2,\ldots,2^\nu$. H measures the degree to which correct and erroneous digits are incorporated in the polynucleotide, and assumes its maximum value when $p = 1/2$. In this case

$$\lambda = \frac{A - 1 + 2^\nu}{2^\nu} \quad . \tag{63}$$

The corresponding stochastic matrix P has all rows equal: they are given by the vector

$$\left[\frac{A\left(\frac{1}{2}\right)^\nu}{\lambda} , \frac{\left(\frac{1}{2}\right)^\nu}{\lambda} , \frac{\left(\frac{1}{2}\right)^\nu}{\lambda} \ldots\ldots , \frac{\left(\frac{1}{2}\right)^\nu}{\lambda} \right] \quad . \tag{64}$$

Using (63), (64) and (60), we have

$$H = -\left[\frac{A}{A - 1 + 2^\nu} \right] \log A + \log (A - 1 + 2^\nu) \quad .$$

We note that λ decreases with sequence length ν, while H increases with ν.

5.3 Genealogies

We define a *genealogy* as a sequence $(I_{k_n})_n$, $k_n \in \{1,\ldots,m\}$ such that I_{k_n} is a direct copy (accurate or inaccurate) of $I_{k_n - 1}$, for $n = 1,2,\ldots$. Demetrius (1983b) has used the Shannon–McMillan theorem on entropy (see, e.g., Billingsley, 1965) to show that the set of all genealogies generated by a given individual after a sufficiently long time n falls into two classes: a class S_1 in which each genealogy occurs with a high probability, and a class S_2 in which each genealogy occurs with an arbitrarily low frequency. The elements of S_1 are called *typical genealogies*.

Let $N_0(n)$ denote the number of genealogies generated up to time n, and $N_1(n)$ the number of typical genealogies. It is known (cf. Demetrius, 1983a,b) that

$$N_0(n) \sim \lambda^n \tag{65}$$

$$N_1(n) \sim e^{Hn} \quad . \tag{66}$$

Tuljapurkar (1982) used the notion of Kullback distance to show, in the context of the Leslie model of age distributions, that H yields a measure of the rate at which a population converges to its stable distribution. Thus the complexity H yields biologically useful information which is not contained in λ.

6. CONCLUSIONS

The theory of multitype branching processes has been shown to provide an appropriate basis for the description of replication in biophysics. The main results derived from the deterministic differential equations of conventional chemical kinetics are valid, on the average, for the corresponding stochastic processes. This is basically a consequence of the important principle by which the initial fluctuations are "frozen in". After a transition period the supercritical multitype branching process either leads to extinction or the total population size becomes very large. In the former case there are no fluctuations, while in the latter the law of large numbers becomes applicable.

Both stochastic and deterministic treatments of replication with errors yield error threshold relations which state that the maximum lengths of faithfully replicated sequences are roughly inversely proportional to the single component error rate. The stochastic threshold turns out to be a stronger condition than the deterministic relation.

ACKNOWLEDGEMENTS

We acknowledge fruitful discussions with Prof. Manfred Eigen and Dr. John McCaskill. Financial support of P. Schuster by the Austrian Fonds zur Förderung der wissenschaftlichen Forschung Project No. 4506 and 5286 is also gratefully acknowledged.

REFERENCES

Athreya, K.B. and P.E. Ney (1972). *Branching Processes*. Springer-Verlag, Berlin.

Biebricher, C.K., M. Eigen and W.C. Gardiner, Jr. (1983). Kinetics of RNA replication. *Biochemistry* 22, 2544–2559.

Billingsley, P. (1965). *Ergodic Theory and Information*. J. Wiley & Sons, New York.

Demetrius, L. (1983a). Statistical mechanics and population biology. *J. Statist. Phys.* 30, 729–773.

Demetrius, L. (1983b). Selection and evolution in macromolecular systems. *J. Theor. Biol.* 103, 619–643.

Ebeling, W. and R. Feistel (1977). Stochastic theory of molecular replication processes with selection character. *Ann. Phys.* 34, 81–90.

Ebeling, W. and M.A. Jimenez-Montana (1980). *Math. Biosciences* 52, 53–71.

Eigen, M. (1971). Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften* 58, 465–526.

Eigen, M. and P. Schuster (1979). *The Hypercycle. A Principle of Natural Self-Organization*. Springer-Verlag, Berlin.

Eigen, M., J.S. McCaskill and P. Schuster (1984). Physics of Evolutionary Optimization. Forthcoming.

Gassner, B. and P. Schuster (1982). Model studies on RNA replication I. The quasiequilibrium assumption and the analysis of a simplified mechanism. *Mh. Chem.* 113, 237–263.

Harris, T.E. (1961) *The Theory of Branching Processes*. Springer-Verlag, Berlin.

Heinrich, R. and T. Sonntag (1981). Analysis of the selection equations for a multivariable population model. Deterministic and stochastic solutions and discussion of the approach for populations of self-reproducing biochemical networks. *J. Theor. Biol.* 93, 325–361.

Inagaki, H. (1982) Selection under random mutations in stochastic Eigen model. *Bull. Math. Biol.* 44, 17–28.

Jagers, P. (1975). *Branching Processes with Biological Applications*. J. Wiley & Sons, London.

Jones, B.L., R.H. Enns and S.S. Rangnekar (1976). On the theory of selection of coupled macromolecular systems. *Bull. Math. Biol.* 38, 15–28.

Jones, B.L. and H.K. Leung (1981). Stochastic analysis of a nonlinear model for selection of biological macromolecules. *Bull. Math. Biol.* 43, 665–680.

Karlin, S. (1974). *A First Course in Stochastic Processes*, 2nd edn. Academic Press, New York.

Kesten, H. and B.P. Stigum (1966). A limit theorem for multidimensional Galton–Watson processes. *Ann. Math. Stat.* 37, 1211–1223.

Küppers, B.O. (1979). Towards an experimental analysis of molecular self-organization and precellular darwinian evolution. *Naturwissenschaften* 66, 228–243.

McCaskill, J. (1984a). A localization threshold for macromolecular quasispecies from continuously distributed replication rates. Submitted to *J. Chem. Phys.*

McCaskill, J. (1984b). A stochastic theory of macromolecular evolution. Submitted to *Biological Cybernetics.*

McQuarrie, D.A. (1967). Stochastic approach to chemical kinetics. *J. Appl. Prob.* 4, 413–478.

Schuster, P. (1981). Prebiotic evolution. In H. Gutfreund (Ed.), *Biochemical Evolution.* Cambridge University Press, Cambridge, pp. 15–87.

Schuster, P. (1983). Selection and evolution in molecular systems – a combined approach by stochastic and deterministic chemical kinetics. In B. Gomez, S.M. Moore, A.M. Rodriguez-Vargas and A. Rueda (Eds.), *Stochastic Processes Applied to Physics and Other Related Fields.* World Scientific Publ. Co., Singapore, pp. 134–159.

Schuster, P. and K. Sigmund (1984a). Random selection and the neutral theory – sources of stochasticity in replication. In P. Schuster (Ed.), *Stochastic Phenomena and Chaotic Behaviour in Complex Systems.* Springer-Verlag, Berlin, pp. 186–207.

Schuster, P. and K. Sigmund (1984b). Random selection – a simple model based on linear birth and death processes. *Bull. Math. Biol.* 46, 11–17.

Swetina, J. and P. Schuster (1982). Self-replication with errors – a model for polynucleotide replication. *Biophys. Chem.* 16, 329–345.

Thompson, C.J. and J.L. McBride (1974). On Eigen's theory of the self-organization of matter and the evolution of biological macromolecules. *Math. Biosciences* 21, 127–142.

Tuljapurkar, S.D. (1982). Why use population entropy? It determines the rate of convergence. *J. Math. Biology* 13, 235–337.