As in Section 4.1.1, the Law of Large Numbers suggests the approximations

$$\frac{B(t)}{Z_0} \approx \overline{b}(t) = \frac{b(t)}{Z_0} \qquad \frac{Z_t}{Z_0} \approx \overline{m}(t) = \frac{m(t)}{Z_0} \ , \tag{4.24}$$

where $\overline{b}$ satisfies Equation (4.21), with Z_0 replaced by 1, and $\overline{m}$ is given by Equation (4.22), with Z_0 replaced by 1 and b replaced by $\overline{b}$. This approximation is, indeed, valid in the sense that $B(t)/Z_0$ converges to $\overline{b}(t)$ and Z_t/Z_0 converges to $\overline{m}(t)$ as Z_0 goes to infinity.

Limit theorems of this type can be found in Kurtz (1983) and Solomon (1987). The limiting model is essentially the standard continuous-time demographic model (e.g., Keyfitz 1977).

4.2 Discrete-Time Dynamical Systems as Population Models

F.C. Klebaner

Deterministic discrete-time processes are characterized by a recurrence equation $x_{n+1} = f(x_n)$ that specifies the relation between the value of a state at time n, x_n, and its value one time step later, x_{n+1}. Such processes are used as models for changes in population size or density. An example is the Ricker model introduced in Section 1.4. The (asymptotic) dynamics of these models is described in many textbooks on biological systems (e.g., Case 2000). We thus include only a short survey here, to connect the field with branching processes and to acquaint those unfamiliar with the area with some basic concepts. Others may well skip Section 4.2.1 and proceed to Section 4.2.2, in which we relate these models to branching processes.

4.2.1 Dynamics of deterministic models

Recall the formulation in Section 4.1 of the Ricker model

$$z_{n+1} = m z_n e^{-b z_n} \ , \tag{4.25}$$

where z_n denotes population size. It can be reformulated in terms of population density by the transformation $x_n = z_n/K$, where K is the area occupied by the population, and rescaling the parameter b. This yields a model of the form

$$x_{n+1} = x_n R(x_n) \ , \tag{4.26}$$

which is the general form of density-dependent models that we consider here. Note that in this section we use "density" in a very loose way. If K is an area, x_n is, indeed, a density in the strict sense, but in the following we also consider cases in which K represents a different constant, such as carrying capacity in the logistic model (see below). In either case, we refer to x_n as a density.

The function $R(x)$ is the individual reproduction function. This function typically depends on the availability of resources and, therefore, it is reasonable to assume that, as population density increases, $R(x)$ tends to zero (the reader may verify that in the Ricker model this is indeed so). This is called negative density dependence.

Klebaner FC (2005). Discrete-Time Dynamical Systems as Population Models. In: Branching Processes: Variation, Growth, and Extinction of Populations, Haccou P, Jagers P & Vatutin VA, pp. 88–94. Cambridge University Press. © International Institute for Applied Systems Analysis

In the Ricker model, and the other examples given in this section, $R(x)$ tends to a positive constant as x approaches zero, which signifies that at low population densities there is no resource limitation and the population grows at a constant rate independent of its density. The branching process equivalent of this situation is the Galton–Watson process. There are also models, however, in which a negative density dependence occurs at low densities too. This so-called "Allee effect" (Allee 1931) can, for instance, result from the difficulty of finding suitable mates when population size is small.

In ecological models a limitation on population size is often assumed, usually called a *carrying capacity* (of the environment for the population in question). When the population is far from its carrying capacity it reproduces at a constant rate, and the reproduction rate declines when the population approaches its carrying capacity and resources become exhausted. In this case, we can let K represent the carrying capacity. The simplest such dependence is linear, with x denoting the density $0 \leq x \leq 1$,

$$R(x) = r(1 - x) \ . \tag{4.27}$$

This is how the logistic model for population dynamics is obtained,

$$x_{n+1} = r x_n (1 - x_n) \ . \tag{4.28}$$

It is the most studied scheme from a mathematical perspective (see Thompson and Stuart 1986), and represents a prototype of simple models that exhibit complex behavior (May 1976).

Another famous density-dependent population approach uses the relation

$$x_{n+1} = r x_n / (1 + a x_n)^b \ , \tag{4.29}$$

which was applied by Hassell *et al.* (1976) to compare population dynamics of 28 species of insects (see also Smith 1974).

Iterating the recurrence equation

$$x_{n+1} = f(x_n) = f(f(x_{n-1})) \tag{4.30}$$

and so on, we obtain that the density in the nth generation is given by the nth iterate of the function f evaluated at the initial population density x_0,

$$x_n = f^{(n)}(x_0) \ , \tag{4.31}$$

where $f^{(n)}$ means f taken n times.

Behavior of the iterates of functions such as the logistic or Ricker is studied within non-linear discrete-time dynamics and chaos theory. Typically, the functions have a shape parameter (r in the logistic model) and, depending on the value of this parameter, the iterates of f converge to a fixed point (a stable fixed point), or they oscillate between a finite number of points (convergence to a stable cycle), or they exhibit chaotic behavior (which means that their positions for large n are described by a distribution function, rather than a limited set of predetermined points).

A point x is called a *fixed point* for f if $f(x) = x$. Clearly, for functions of the form $f(x) = xR(x)$ the point $x = 0$ is a fixed point, but there may be others. What happens if the population density is very low, near to zero, but positive? Development in this case depends on the stability of the point zero. A fixed point x^* is called *stable* or *attracting* if for all the initial points x_0 near x^*, the iterates $f^{(n)}(x_0)$ converge to x^* as $n = 1, 2, \ldots$ increases.

Alternatively, if there is an interval that includes x^* such that for some n the iterate x_n is outside this neighborhood, the fixed point x^* is called *unstable* or *repelling*. A sufficient condition for the stability of a fixed point x^* is that $|f'(x^*)| < 1$, and for it to be unstable the condition is $|f'(x^*)| > 1$. In the case $|f'(x^*)| = 1$, x^* may be attracting or repelling. For a necessary and sufficient condition for a fixed point to be attracting, see, for example, Theorem 2.2.1 in Sharkovskii *et al.* (1993).

If 0 is a repelling fixed point, f has another fixed point x^* that may also be either attracting or repelling. If it happens to be attracting, the long-term iterates converge to it, and if it is repelling, a cycling behavior occurs. For example, a cycle of period 2 means that there are two points, x_1^* and x_2^*, such that $f(x_1^*) = x_2^*$ and $f(x_2^*) = x_1^*$. This cycle is attracting if for large even values of n the iterates x_n are in the vicinity of a point x_1^* and if for large odd values of n the iterates x_n are in the vicinity of a point x_2^*. We can also describe a cycle of period 2 by means of fixed points of the twice-iterated function, $f^{(2)}$. If f has a two-cycle $\{x_1^*, x_2^*\}$, then $f^{(2)}$ has two fixed points x_1^* and x_2^*. The cycle is stable or attracting if these fixed points are stable. For a cycle to be attracting it is enough that

$$|f^{(2)\prime}(x_1^*)| < 1 . \tag{4.32}$$

Using the chain rule of differentiation we find that $f^{(2)\prime}(x_1^*) = f'(f(x_1^*))f'(x_1^*) = f'(x_2^*)f'(x_1^*)$. Thus, a sufficient condition for a cycle to be attracting is given by

$$|f'(x_1^*)f'(x_2^*)| < 1 . \tag{4.33}$$

Of course, a cycle of period d and its stability are defined similarly.

A large class of dynamical systems has asymptotically periodic trajectories: a dynamical system is called *simple* if each of its trajectories is periodic or asymptotically periodic. Moreover, there is a class of simple dynamical systems in which the stable cycle is unique and trajectories (x_n) are attracted to it for almost all initial points x_0 (*ibid*).

Example 4.1 We examine the behavior of the iterates in the logistic model $x_{n+1} = rx_n(1 - x_n)$. The function $f(x) = rx(1 - x)$ has a single fixed point 0 if $r \leq 1$, which is attracting, and $x_n \to 0$ as $n \to \infty$ for any x_0. For $1 < r \leq 3$, the fixed point zero becomes repelling and another fixed point appears, $x^* = 1 - 1/r$. This point is attracting if $r < 3$, as $|f'(x^*)| < 1$, $(f'(x^*) = 2 - r)$. If $r = 3$, x^* is still attracting, although $|f'(2/3)| = 1$. If $r \leq 3$, then $x_n \to x^*$ as $n \to \infty$ for any $x_0 \neq 0, 1$. When $r > 3$, then $x^* = 1 - 1/r$ is repelling as $|f'(x^*)| > 1$. For values of r in the range $3 < r < 1 + \sqrt{6} \approx 3.449$, f has a stable cycle of period 2. These points are determined as roots of $f^{(2)}(x) = x$. For all x_0 outside an (actually finite) exceptional set, x_n converges to this cycle.

When r increases further, $1 + \sqrt{6} \leq r$, the stable two-cycle becomes unstable and a stable cycle of period 4 is created. For all points x_0 outside an exceptional set (which contains the fixed points, cycles, and their pre-images), $f^{(n)}(x_0)$ converges to this four-cycle. This phenomenon (appearance of stable cycles of higher powers of two instead of unstable ones) is known as period doubling bifurcation, and continues until r reaches some value $r_c \approx 3.569\ldots$, at which point no stable trajectories of longer periodicity exist, and the system displays no simple dynamics.

For any $r < r_c$ there is a stable cycle of period 2^k (k depends on r), and $f^{(n)}(x_0)$ converges to this cycle for all x_0 except those that go to repelling cycles of periods 2^i, $i = 0, 1, \ldots, k-1$. The value r_c is known as the value for the onset of chaos. For $r > r_c$ there are infinitely many cycles, all of which may be repelling. For certain values (*periodic windows*) of the parameter the system admits attracting cycles of periods not restricted to the powers of 2.

When $r = 4$, the long-term behavior of $f^{(n)}(x_0)$ is described by the probability distribution

$$\frac{1}{\pi\sqrt{x(1-x)}}\,, \tag{4.34}$$

that is, for large n the probability of finding $f^{(n)}(x_0)$ in an interval $[a, b]$ is given by

$$\int_a^b \frac{dx}{\pi\sqrt{x(1-x)}}\,. \tag{4.35}$$

◇ ◇ ◇

Remark. For the Ricker model, $f(x) = xe^{r-x}$, a bifurcation to a cycle of period 2 occurs at $r = 2$, then further from a 2-cycle to a 4-cycle at 2, 2.526, etc., and the value of onset of chaos is $r_c \approx 2.692$.

◇ ◇ ◇

4.2.2 Density dependent branching processes and dynamical systems

Deterministic models of the form $x_{n+1} = f(x_n)$ are macroscopic, they give a rule according to which the whole population evolves. Branching processes, however, are microscopic, built upon the individual behavior of population members and determined by their offspring distribution.

As we have seen, though, the macroscopic models allow feedback in the form of the effects of population density on growth, and may exhibit periodic behavior, which does not appear in branching processes. The bridge between the two approaches is provided by branching processes with a similar feedback (i.e., population-size- and density-dependent branching processes).

In branching processes dependent on population size, the distribution of offspring numbers depends on the size of the population z, and in density-dependent branching processes the distribution of offspring depends on the population density, or concentration z/K, where K is a parameter such as (but not necessarily) the carrying capacity. If this parameter is fixed, there is no difference between the two types of models, but when it becomes large, density-dependent models may simplify and allow approximations. Indeed, for large values of K they reduce to deterministic dynamical systems plus a small noise. The dynamical system part represents the deterministic approximation and the noise admits a Gaussian approximation.

This Gaussian approximation is different from the diffusion approximation of branching processes, treated in the preceding sections. In diffusion approximations the time is scaled to arrive at a continuous-time process, but in this case time remains discrete and the limiting process is not continuous, but jumps as the corresponding dynamical system.

As indicated in Section 2.6, density-dependent branching processes are defined in the same way as the classic Galton–Watson process, except that offspring distributions are allowed to depend on population density. This can be written as

$$Z_{n+1} = \sum_{j=1}^{Z_n} \xi_{j,n}(Z_n/K) , \tag{4.36}$$

to indicate that the distribution of $\xi_{j,n}(Z_n/K)$ is dependent on Z_n/K. The offspring numbers themselves are independent, with the common distribution being that of $\xi(x)$ if $Z_n/K = x$.

Example 4.2 Consider branching processes that occur in polymerase chain reactions (PCRs). PCR is a stepwise procedure in molecular biology whereby in each step DNA molecules either remain or are replaced by two copies. It is further described in Section 7.5.

The reaction can be modeled as a (single-type) Galton–Watson branching process, in which each individual has one or two offspring in the next generation. The probability of the latter event is usually termed the *efficiency* in the present connection. It is natural from the experimental setup that the efficiency of the reaction should decrease.

Under classic, so-called Michaelis–Menten kinetics, largely valid for enzymatic reactions, it follows that the probability of successful copying (two offspring) is given by

$$p(z) = \frac{K}{K+z} , \tag{4.37}$$

where z is the number of molecules and K is the Michaelis–Menten constant of the reaction. Initially, the efficiency is close to 1, since K is large compared to z.

The result is a density-dependent binary splitting Galton–Watson process, whereby the alternative to splitting is remaining into the next generation (experiment cycle), or equivalently giving birth to one offspring. The offspring number $\xi(z)$ takes values 1 and 2 with probabilities $1 - p(z)$ and $p(z)$, respectively.

◇ ◇ ◇

As shown in Section 4.2.1, it is more convenient to consider the density process $x_n^K = Z_n/K$, which evolves according to

$$x_{n+1}^K = \frac{1}{K} \sum_{j=1}^{Kx_n^K} \xi_{j,n}(x_n^K) . \tag{4.38}$$

We show that this model is a stochastic analog of the deterministic model $x_{n+1} = f(x_n)$ with a suitable function f. Indeed, denote by $R(x) = \mathbb{E}[\xi(x)]$ the mean offspring number when the population density is x. By subtracting and adding it within the sum, we have

$$x_{n+1}^K = \frac{1}{K}\sum_{j=1}^{Kx_n^K} R(x_n^K) + \frac{1}{K}\sum_{j=1}^{Kx_n^K}[\xi_{j,n}(x_n^K) - R(x_n^K)]\,, \tag{4.39}$$

or

$$x_{n+1}^K = x_n^K R(x_n^K) + \eta_n^K = f(x_n^K) + \eta_n^K\,, \tag{4.40}$$

where $f(x) = xR(x)$ and

$$\eta_n^K = \frac{1}{K}\sum_{j=1}^{Kx_n^K}[\xi_{j,n}(x_n^K) - R(x_n^K)] \tag{4.41}$$

is random. This random term is small for large values of K, essentially by the Law of Large Numbers. Let x_n be the nth iterate of $f(x) = xR(x)$, starting from $x_0 = Z_0/K$. The results below state that, for large values of the carrying capacity, the process x_n^K is approximated by the deterministic sequence x_n, with the difference $x_n^K - x_n$ being approximately normal with mean zero and variance of order $1/K$.

In the next two theorems we assume that the function $f(x) = xR(x)$ has a continuous derivative and that the variances $\sigma^2(x)$ of offsprings $\xi(x)$ are bounded by some constant, say $\sigma^2(x) \leq C$.

Theorem 4.1 (Consistency theorem) *For any fixed n, as $K \to \infty$, $x_n^K \to x_n$ in probability.*

The proof uses induction on n and Chebyshev's inequality,

$$\mathbb{E}[(\eta_n^K)^2] = \mathbb{E}[\mathbb{E}[(\eta_n^K)^2|x_n^K]] = \frac{1}{K}\mathbb{E}[\sigma^2(x_n^K)] \leq C/K \to 0\,. \tag{4.42}$$

Theorem 4.2 (Fluctuation theorem) *Assume, in addition, that the third absolute moments $\mathbb{E}[(\xi(x) - R(x))^3]$ are bounded. Then, for any fixed n, as $K \to \infty$, $(x_n^K - x_n)\sqrt{K}$ converges in distribution to a normal random variable $N(0, D_n^2)$, where $D_0 = 0$ and D_n is defined by the recurrence relation*

$$D_{n+1}^2 = x_n\sigma^2(x_n) + [f'(x_n)D_n]^2\,. \tag{4.43}$$

The proof of this result follows by induction on n and an analysis of characteristic functions, and can be found in Klebaner (1993) under less stringent assumptions. In Klebaner and Nerman (1994) these results are established not only for a single fixed n, but also for any collection of times $n_1, n_2, \ldots, n_k$ (which corresponds to the convergence of processes, or the functional version of the limit theorem). Watkins (2000) generalized these results to multi-type processes (structured populations) and referred to them as consistency and fluctuation theorems in biology.

Example 4.2 (*continued*) We can express the probability of successful division in PCR as a function of the so-called dimensionless reduced concentration $x = z/K$ (Schnell and Mendoza 1997a),

$$p(x) = \frac{1}{1+x} . \tag{4.44}$$

Since the expected number of offspring per individual is

$$m(x) = 1 - p(x) + 2p(x) = 1 + 1/(1+x) , \tag{4.45}$$

we obtain that the population density process has the representation

$$x_{n+1}^K = f(x_n^K) + \eta_n^K , \tag{4.46}$$

with $f(x) = xm(x) = x/(1+x)$. The reproduction variance can be checked easily to satisfy

$$\sigma^2(x) = p(x)(1 - p(x)) = x/(1+x)^2 , \tag{4.47}$$

obviously bounded. Thus, for large values of K, the sequence $x_n^K = Z_n/K$ can be approximated by the deterministic sequence x_n obtained as the nth iterate of f starting at x_0. This is further developed in Jagers and Klebaner (2003).

◇ ◇ ◇

4.3 Branching Processes and Structured Population Dynamics

M. Gyllenberg and P. Jagers

4.3.1 Introduction

Most of the classic deterministic population models developed by Lotka, Volterra, and others during the “golden age” of theoretical ecology in the 1930s are concerned with a single homogeneous population or with the interaction between several homogeneous populations. In particular, these models are based on the assumption that, at least on average, all individuals in a population behave identically with respect to reproduction, survival, exploitation of resources, competition, and other processes of importance for the dynamics of the population. They are deterministic versions of single-type Markov branching processes, or modifications of such processes.

In reality, matters are more complex. The reproductive behavior is age specific and depends on nutrition, and the same applies to survival. A predator is hardly likely to catch a prey that is larger, stronger, and quicker than itself. The list could be continued almost forever. In many cases, differences between individuals make a difference for the resultant dynamics. Moreover, it may be important to predict the *composition* and not only the size of a population (think, for instance, of human demography in which estimates of the future age distribution of the population influence socio-economic decision making). In branching processes, such considerations lead to multi-type processes with type spaces of varying complexity. In deterministic approaches, they take us to the realm of *structured populations* (Metz and Diekmann 1986). In this section a short introduction to the modeling of

References

References in the book in which this section is published are integrated in a single list, which appears on pp. 295–305. For the purpose of this reprint, references cited in the section have been assembled below.

Allee WC (1931). *Animal Aggregations: A Study in General Sociology*. Chicago, IL, USA: University of Chicago Press

Case TJ (2000). *An Illustrated Guide to Theoretical Ecology*. New York, NY, USA: Oxford University Press

Hassell MP, Lawton JH & May RM (1976). Patterns of dynamical behaviour in single-species populations. *Journal of Animal Ecology* **45**:471–486

Jagers P & Klebaner FC (2003). Random variation and concentration effects in PCR. *Journal of Theoretical Biology* **224**:299–304

Klebaner FC (1993). Population-dependent branching processes with a threshold. *Stochastic Processes and their Applications* **46**:115–127

Klebaner FC & Nerman O (1994). Autoregressive approximation in branching processes with a threshold. *Stochastic Processes and their Applications* **51**:1–7

May RM (1976). Simple mathematical models with very complicated dynamics. *Nature* **261**:459–467

Schnell S & Mendoza C (1997a). Theoretical description of the polymerase chain reaction (QC–PCR). *Journal of Theoretical Biology* **184**:433–440

Sharkovskii AN, Maistrenko YL & Romanenko EY (1993). *Difference Equations and Their Applications*. Dordrecht, Netherlands: Kluwer

Smith M (1974). *Models in Ecology*. Cambridge, UK: Cambridge University Press

Thompson MT & Stuart HB (1986). *Nonlinear Dynamics and Chaos.* New York, NY, USA: John Wiley & Sons

Watkins J (2000). Consistency and fluctuations theorems for discrete time structured population models having demographic stochasticity. *Journal of Mathematical Biology* **41**:253–271