

## Topic 1: Chaos and the Logistic Map

This set of exercises examines a simple population model which exhibits a number of interesting features that can be studied both analytically and numerically: **Equilibrium points**, **bifurcations** and **chaotic behaviour**. The model is **discrete**, making it particularly easy to implement in a computer.

### The Logistic Map

Let  $x_n$  be a population at a time  $n$ , and assume that the population changes according to the equation

$$x_{n+1} = r x_n (1 - x_n), \quad (1)$$

where  $r > 0$  is a parameter. For this model, we're measuring time in rough units, say years, so  $x_1$  is the population after one year,  $x_2$  is the population after two years, etc. That is, the time  $n$  is an integer.

If we know the population at  $x_0$ , we can compute the population at  $x_1$ , and then  $x_2$ , etc., by repeatedly calculating the RHS and substituting the result into the next equation:

$$\begin{aligned} x_1 &= r x_0 (1 - x_0), \\ &\quad \searrow \\ x_2 &= r x_1 (1 - x_1), \\ &\quad \searrow \\ x_3 &= r x_2 (1 - x_2), \\ &\quad \vdots \end{aligned}$$

An important question we might like to answer is: What happens to  $x_n$  as  $n \rightarrow \infty$ ? Does it settle on a single value? Does it go off to infinity?

- If we think of the equation as describing a **series** of values  $x_n$ , we are asking whether the series **converges**. See MAM102.
- In terms of dynamical systems, we are asking whether it has **equilibrium points** which are **stable**. See MAP312.

A further important question is how this behaviour depends on the parameter  $r$ .

### The logistic equation

We are already familiar with a related (but different!) population model expressed as a differential equation, the *logistic equation*:

$$\dot{x} = rx(1 - x). \quad (2)$$

This is a population model which exhibit nearly exponential growth for small  $x$  (due to the  $rx$  term), but has an equilibrium point at  $x = 1$ .

**Exercise:** Review your dynamical systems notes to determine the following behaviour of  $x(t)$ :

- How do we find the equilibrium points of this model?
- Where are the equilibrium points located?
- Classify any equilibrium points as stable or unstable.
- How does the location of the equilibrium points depend on  $r$ ?
- Plot  $x_*$  vs.  $r$ , where  $x_*$  are the equilibrium points of the system.

## Discrete vs. continuum models

If we compare Eqs. (1) and (2), we see they are of a similar form on the RHS. The main difference is that in Eq. (1) the variable  $x_n$  is only defined at integer values of  $t$ :  $x_0, x_1, x_2, \dots$ . Therefore, we call this a **discrete** model. There are gaps between neighbouring values of  $t$ .

For the logistic equation, Eq. (2), however, the function  $x = x(t)$  has a continuous time parameter  $t$  which can take any real value. This is called a **continuum** model.

- Computers are naturally good at studying discrete models, since we evaluate them by simply **iterating**. Given Eq. (1) and an initial value for  $x_0$ , it is easy to write a **loop** to calculate  $x_1$ , then  $x_2$ , and so on.
- Computers are not able to deal with continuum models as easily, since we can't take infinitely small steps in a loop. The main reason for this is the digital nature of computers – numbers are stored as a finite number of digits (**floating point numbers**) and there are gaps between neighbouring numbers. So it's not possible to represent a continuous parameter  $t$ , or continuous function  $x(t)$ , or a derivative  $dx/dt$ . The best we can do is **approximate** these ideas.

A significant goal of **numerical analysis** is to develop methods to approximate continuum models in a way that the inevitable errors are limited, or at least understood and quantified.

Consider the following:

- Many problems in physics are expressed as differential equations (ODEs or PDEs), such as Newton's equation for the force of gravity, the Navier-Stokes equations for fluid flow, or the Schrödinger equation of quantum mechanics.

- Even population models are often expressed in terms of continuous variables, such as the logistic model above, or the Lotka-Volterra equations. Continuum models are useful for studying a broad class of systems.
- However, some of these continuum models have discrete systems at their core. A population is not really continuous – there might be 100 animals or 101 animals, but not 100.37 animals or  $2\pi$  animals. The water in a river seems like a continuous fluid, but actually it is made up of discrete molecules and the forces between them. It only looks continuous if we stand back and look at length scales much larger than the individual particles.

We often use continuous functions to represent systems of many individual particles. This is valid as long as the length scales we are studying  $L$  are much larger than the length scales of the component particles,  $\ell$ :

$$L \gg \ell.$$

For instance,

- If we are studying water waves, we are interested in a scale of metres:  $L \simeq 1$  m. Water molecules are approximately  $\ell \simeq 0.27$  nm across. So  $L \simeq 3.7 \times 10^9 \ell$ . We comfortably satisfy  $L \gg \ell$ , and therefore the continuum Navier-Stokes PDEs can be used to model ocean water.
- A single  $\ell = 1$  zebra is “small” compared to a population of  $L = 1000$  zebras, so we might consider using the continuum logistic equations to model this population. However, if it’s a smaller group, say  $L = 100$  zebras, the continuum approximation might not work as well and we need to treat any predictions with caution.

The reason we often prefer to build models out of continuous equations is that they are often a more efficient way of representing a problem. For instance, a single function “density” encapsulates the position of billions of molecules without having to keep track of each particle individually.

And some physical systems, e.g. Newtonian gravity, Schrödinger’s equation or general relativity, are manifestly continuous, without underlying particles that we know of.

When we treat a system composed of many particles as a single continuous function, we call this a **continuum approximation**.

Ironically, when we want to study these problems in a computer, we need to go back to a discrete version of the problem for the reasons discussed above.

## Exercises — Exploring the logistic map

1. Write a computer program to evaluate the logistic map, Eq. (1). Given an initial value  $x_0$ , and growth parameter  $r$ , you should be able to evaluate  $x_n$  for any integer value of  $n$  by iteration:

$$x_0 \rightarrow x_1 \rightarrow x_2 \rightarrow \cdots \rightarrow x_n.$$

Try it out with  $r = 2$  and  $x_0 = 0.5$ . You should get  $x_n = 0.5$  for all values of  $n$ .

2. **Finding equilibrium points analytically.** In dynamical systems, an equilibrium point occurs when the system no longer changes, i.e.  $\dot{x} = 0$ . For a discrete system, we require the same thing, that the system doesn't change,  $x_{n+1} \simeq x_n$ . For  $r = 2$ , find a pair of equilibrium points for the logistic map, Eq. (1).
3. **Finding equilibrium points numerically.** For the following questions, use  $r = 2$  with the code developed in Q1 and  $x_0 = 0.01$ .
  - a) Does the value of  $x_n$  approach an equilibrium point? What is the equilibrium value, approximately?
  - b) How many iterations does it take to get within 0.1% error of the equilibrium point? (Three digits of accuracy.)
  - c) Store the first 20 values of  $x_n$  and make a plot of  $x_n$  vs.  $t$  to visually demonstrate the solution converging to the equilibrium point.
4. **Verifying that the fixed point is independent of the initial data.** Again with  $r = 2$ , evaluate the logistic map for each of the following initial data values:

$$x_0 = \{ 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 \}$$

- Run for a sufficient number of iterations that the convergence to a fixed value is clear (e.g. three decimal places)
- Store the data at each iteration so that you can plot  $x_n$  vs.  $t$
- Plot all the curves for each  $x_0$  on a single axis.

The resulting plot should look something like Fig. 1. The curves bounce around at the beginning, but eventually all settle on a common value. What is that value?

(But what happens if we use  $x_0 \leq 0$  or  $x_0 \geq 1$ ?)

5. **Examining equilibrium point position  $x^*$  as a function of  $r$ .** The location of the equilibrium point is a function of the parameter  $r$ . We can examine this by examining various values of  $r$  and examining the value to which the function converges.

- Set up an array of  $r$  values between  $r = 0$  and  $r = 3$ , spaced by 0.05, i.e.

$$r = [ 0.0, 0.05, 0.10, 0.15, \dots, 2.90, 2.95, 3.0 ]$$

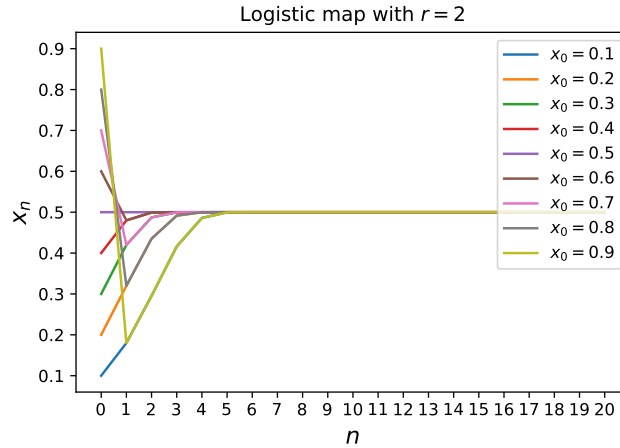


Figure 1: The logistic map with  $r = 2$  converges to the same value independent of the initial data (Exercise 4).

- For each value of  $r$ , iterate the logistic map to determine the value to which it converges. For instance, keep iterating until the values stop changing by very much:

$$\text{Stop iterating when: } x_n = 0, \quad \text{or} \quad \frac{x_{n+1} - x_n}{x_n} < 10^{-6}$$

and set the equilibrium point location to be

$$x^* \simeq \frac{x_{n+1} + x_n}{2}.$$

That is, we can approximate the location of the equilibrium point using the average of the last two points. It should be accurate to 0.0001%, which is good enough for our purposes. Store this value of  $x^*$ .

- Keep track of  $x^*$  for each value of  $r$ , and then plot  $x^*$  vs.  $r$ .

Note that based on the results of the previous exercise, it doesn't matter which initial data value we use, so pick any value  $0 < x_0 < 1$ .

You should see the equilibrium point  $x^*$  is always zero for small values of  $r$ . At some particular value of  $r$ , the behaviour changes — what is this value? Can we understand this from the analytical results from Exercise 2?

- 6. Bifurcations in the logistic map.** The previous exercise examines values of  $r$  between  $0 \leq r \leq 3$ . What happens if we go outside this range?

Iterate the logistic map using  $r = 3.2$ , and plot  $x_n$  vs.  $n$ . Your plot should look something like Fig. 2b.

We see that the values of  $x_n$  don't settle to a single value, but rather alternates between a pair of values. Further, if we try a few different values of  $r$  between  $3 < r < 3.449$ , the distance between the two values grows.

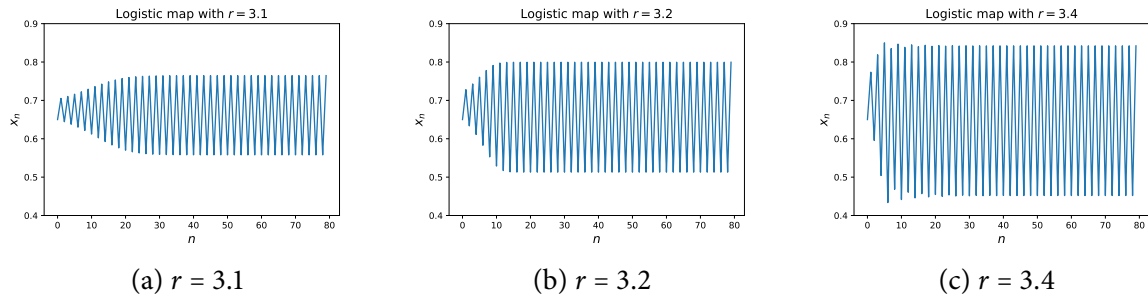


Figure 2: For values of  $r > 3$ , the equilibrium point structure of the logistic map changes. Between  $0 < r < 3.5$ , it settles down to an alternation between two different values. (Exercise 6)

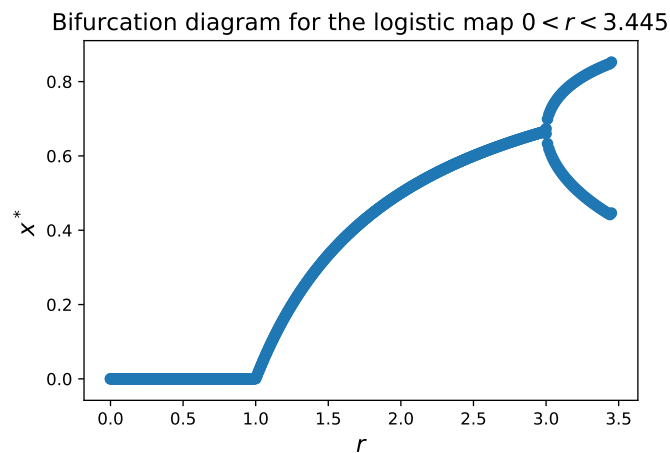


Figure 3: At  $r = 3$ , the logistic map bifurcates to oscillate between two stable equilibrium points. (Exercise 7)

The transition from one to two equilibrium points at  $r = 3$  is called a **period doubling**. Since we jump back and forth, we have to move two steps to return to the same equilibrium value.

7. Extend your plot from Exercise 5 to the range  $0 < r < 3.45$ . To keep things simple:
  - For each value of  $r$ , iterate the logistic map 1000 times.
  - Store the **last two values** of  $x_n$ . For  $r < 3$ , these will be almost identical and appear as a single point on our plot. But for  $r > 3$  each of these points will correspond to a different branch of the bifurcation.
  - Add these to a list of  $(r, x_n)$  pairs.
  - Use a **scatter plot** to display the results. It should look something Fig. 4b. (Use a scatter plot because a regular plot will usually try to connect points with lines, which will look confusing if you have multiple  $x_n$  for each  $r$  value.)

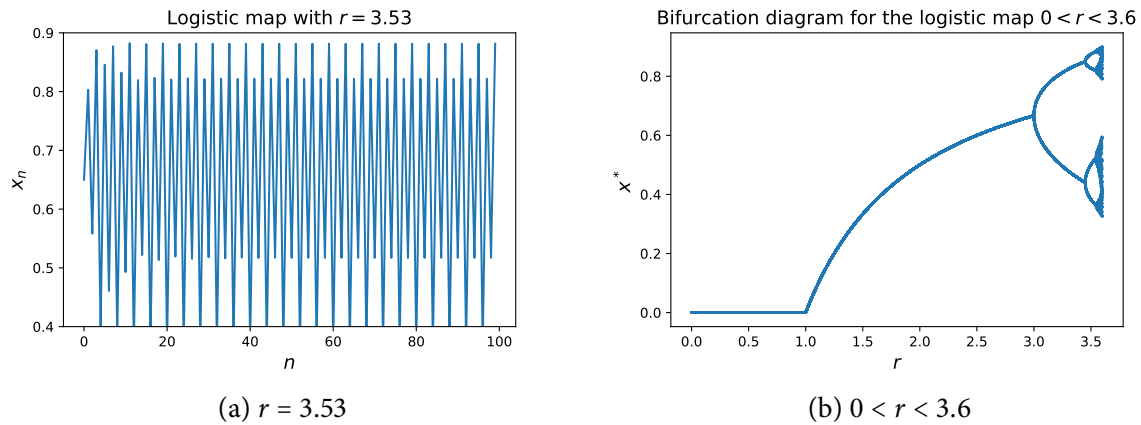


Figure 4: For  $r > 3.449$ , we see another period doubling in the logistic map. The values of  $x_n$  fluctuate between four different roots of the equation, as seen in the left-hand plot. The right-hand plot shows the corresponding bifurcation diagram for  $0 < r < 3.6$ .

Why did we limit the range to  $r < 3.45$ ? Analytically, we can show that at the value  $r = 1 + \sqrt{6} \approx 3.449$ , another bifurcation occurs. If we were to continue the plot a bit further, it would look like Fig. 4b. We see another period doubling, followed shortly by even more period doublings towards the right side of the plot.

8. **Finding the onset of chaos in the logistic map.** If we continue the bifurcation diagram to larger values of  $r$ , we find that more bifurcations occur and at an increasing rate. Eventually it is not possible to see any pattern in the values of  $x_n$ .

Continue the bifurcation diagram to the left, including values  $0 < r < 4$ . Use the following procedure, which extends what we did in the previous exercise:

- Check a number of evenly spaced  $r$  values between  $r = 0$  and  $r = 4$ , e.g.

$$r = \{ 0.00, 0.01, 0.02, \dots, 3.97, 3.98, 4.00 \}$$

- Evolve the logistic map for 1000 iterations.
- Record the next 32 values of  $x_n$  (i.e.  $x_{1001}$  to  $x_{1100}$ ), along with the corresponding  $r$  value.
- Use a scatter plot to display the results. It should look something like Fig. 5.

## What is happening?

**The onset of chaos.** For  $r > 3$  the logistic map experiences an increasing number of bifurcations until eventually it is not possible to distinguish any periodic behaviour in the values of  $x_n$ .

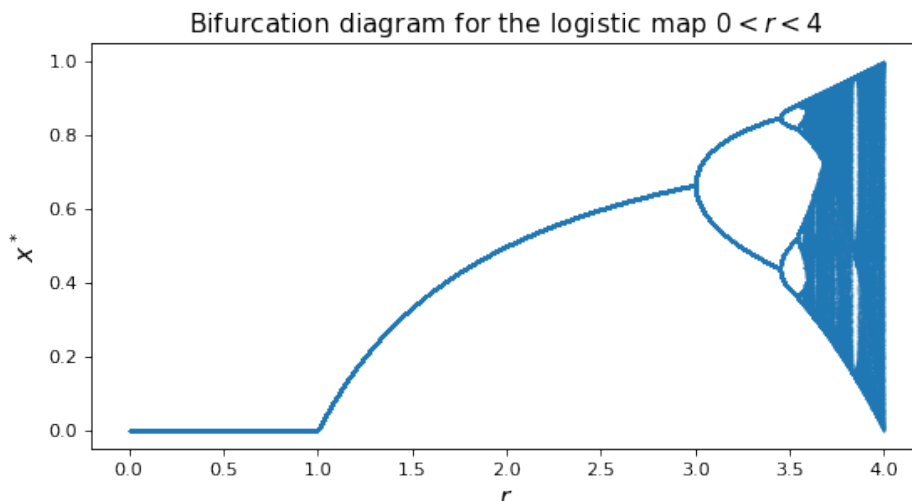


Figure 5: A bifurcation for the logistic map for the range  $0 \leq r \leq 4$ . For  $r < 3$ , the values of  $x_n$  approach a single value. Between  $3 < r < 1 + \sqrt{6}$ , it alternates between two values. At  $r = 1 + \sqrt{6} \simeq 3.445$  there is a further bifurcation and  $x_n$  alternates between four values. Further bifurcations occur at larger values of  $r$  until eventually no patterns can be discerned in the values of  $x_n$ . Interestingly, there are brief windows of regularity for larger values of  $r$ , for instance near  $r = 3.84$ .

Further, if we choose even slightly different initial data we end up with a completely different sequence  $x_n$ . This is the definition of **chaos**: A system is **chaotic** when it is sensitive to the choice of initial data. Small changes of the starting point lead to completely different evolutions.

For example, Fig. 6 plots results for the logistic map for  $r = 3.9$ . The first thing we can notice is that the evolutions don't appear to be periodic and don't settle on any repeating values.<sup>1</sup>

We can also see that the results of the sequence are no longer independent of the initial data. This behaviour is very different from what we observed in Exercise 4, where for  $r = 3.2$  we found that regardless of  $x_0$  the logistic map always converged to the same result. In the case of  $r = 3.9$ , rather, the two displayed sequences with  $x_0 = 0.65$  and  $x_0 = 0.6501$  diverge after only a few iterations, and appear to be completely uncorrelated.

**Fixed points of the logistic map.** The logistic map in Eq. (1) has the form

$$x_{n+1} = f(x_n),$$

<sup>1</sup>This is true at least within the shown interval, but it's possible a pattern emerges if we go further along in  $n$ .



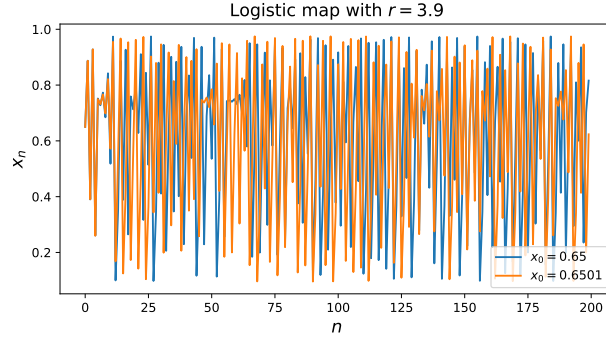


Figure 6: “Chaos” in a mathematical system is defined by a sensitive dependence on initial data. Here we see that two nearby values of the initial value  $x_0$  of the logistic map lead to very different evolutions when  $r = 3.9$ . Neither seem to converge to any fixed points, and after only a few iterations the two series for  $x_n$  appear to be completely uncorrelated.

where the specific RHS is given by  $f(x_n) = rx_n(1 - x_n)$ . As we iterate the logistic map, we have

$$\begin{aligned}
 x_{n+1} &= f(x_n) \\
 x_{n+2} &= f(x_{n+1}) = f(f(x_n)) = (f \circ f)(x_n) = f^2(x_n) \\
 x_{n+3} &= f(x_{n+2}) = f(f(f(x_n))) = (f \circ f \circ f)(x_n) = f^3(x_n) \\
 &\vdots \\
 x_{n+k} &= f^k(x_n)
 \end{aligned} \tag{3}$$

In Exercise 2 we identified the fixed points by solving the first of these equations for  $x_{n+1} = x_n$ :

$$x_n = f(x_n), \quad \text{i.e.} \quad x_n = rx_n(1 - x_n).$$

This is a quadratic equation for  $x_n$  and therefore has two roots, one of which is at  $x^* = 0$  and the other at  $x^* = 1 - 1/r$ .

But if  $x^*$  is actually an equilibrium point, it actually needs to satisfy each of the equations in (3):

$$x^* = f(x^*), \quad x^* = f^2(x^*), \quad x^* = f^n(x^*), \quad \dots \quad x^* = f^n(x^*)$$

The first of these equations is quadratic, with two roots. The second equation is quartic, with four roots, etc., and the order of the polynomial increases as we move to the right. As  $n \rightarrow \infty$ , we end up with an equation that has an infinite number of roots.

It is remarkable that for  $r \leq 3$ , all of these equations happen to have the same pair of repeated roots. For  $3 < r < 1 + \sqrt{6}$ , this changes and we end up with four roots of the quartic equation  $x^* = f^4(x^*)$ . All of the higher order equations have the same four roots. But for larger values of  $r$ , this symmetry is broken until we arrive at a situation where the equations  $x^* = f^n(x^*)$  have an arbitrarily large number of unique notes.

**Implications of chaos.** If we are using the logistic map as a model for making predictions, this is an important concern. In a physical model, we'd normally set our initial data based on some measurement. But measurements can only be known to a finite level of accuracy. Since the chaotic system is sensitive to small inaccuracies, we can't use it to make predictions, since even a small inaccuracy in the initial data leads to a very different series of  $x_n$ .

In popular culture, this is often referred to as the **butterfly effect**: A tiny event, such as the death of a butterfly, can have implications which change the course of history.

A chaotic system is still **deterministic**. Given perfect knowledge of the initial conditions, the future evolution is completely determined. But it is still **unpredictable** because we never have perfect knowledge of the initial conditions.

This feature of the logistic equation was first recognized by Robert May, a 26-year old post-doc in theoretical physics at Princeton, in the early 1970s. See, for example, [1, 2]. Although the phenomenon of “chaos” had been previously observed in other equations, the remarkable feature of this result was how such complex behaviour could arise from an equation as simple as the logistic map.

## References

- [1] Robert M May. Biological populations with nonoverlapping generations: stable points, stable cycles, and chaos. *Science*, 186(4164):645–647, 1974.
- [2] Robert M May. Simple mathematical models with very complicated dynamics. *Nature*, 261:459–467, 1976.

Here is the research paper in which Robert May first identified chaos in the logistic map. It's fairly straightforward and doesn't rely on anything beyond 3rd year math, but was fundamental in highlighting the importance of chaos in differential equations.

R May, Science, vol 186, 1974.

## Biological Populations with Nonoverlapping Generations: Stable Points, Stable Cycles, and Chaos

**Abstract.** *Some of the simplest nonlinear difference equations describing the growth of biological populations with nonoverlapping generations can exhibit a remarkable spectrum of dynamical behavior, from stable equilibrium points, to stable cyclic oscillations between 2 population points, to stable cycles with 4, 8, 16, . . . points, through to a chaotic regime in which (depending on the initial population value) cycles of any period, or even totally aperiodic but bounded population fluctuations, can occur. This rich dynamical structure is overlooked in conventional linearized analyses; its existence in such fully deterministic nonlinear difference equations is a fact of considerable mathematical and ecological interest.*

In some biological populations (for example, man), growth is a continuous process and generations overlap; the appropriate mathematical description involves nonlinear differential equations. In other biological situations (for example, in 13-year periodical cicadas), population growth takes place at discrete intervals of time and generations are completely nonoverlapping; the appropriate mathematical description is in terms of nonlinear difference equations. For a single species, the simplest such differential equations, with no time delays, lead to very simple dynamics: a familiar example is the logistic,  $dN/dt = rN(1 - N/K)$ , with a globally stable equilibrium point at  $N = K$  for all  $r > 0$ .

It is the purpose of this report to point out that many of the corresponding difference equations of population biology have been discussed inadequately, as having either a stable equilibrium point or being unstable, with growing oscillations (1, 2). In fact, some of the very simplest nonlinear difference equations even for single species exhibit a spectrum of dynamical behavior which, as the intrinsic growth rate  $r$  increases, goes from a stable equilibrium point, to stable cyclic oscillations between 2 population points, to stable cycles with 4 points, then 8 points, and so on, through to a regime which can only be described as chaotic (a term coined by J. A. Yorke). For any given value of  $r$  in this chaotic regime there are cycles of period 2, 3, 4, 5, . . . ,  $n$ , . . . , where  $n$  is any positive integer, along with an uncountable number of initial points for which the system does not eventually settle into any finite cycle; whether the system converges on a cycle, and, if so, which cycle, depends on the initial population point (and of course some of the cycles may be attained only from infinitely unlikely initial points). Figure 1 aims to illustrate this range of behavior.

Specifically, consider the simple nonlinear equation

$$N_{t+1} = N_t \exp[r(1 - N_t/K)] \quad (1)$$

This is considered by some people (2, 3) to be the difference equation analog of the logistic differential equation, with  $r$  and  $K$  the usual growth rate and carrying capacity, respectively. The stability character of this equation, as a function of increasing  $r$ , is set out in Table 1 and illustrated by Fig. 1.

Another example is

$$N_{t+1} = N_t[1 + r(1 - N_t/K)] \quad (2)$$

This quadratic form is probably the simplest nonlinear equation one could write. Although discussed by various people (4, 5) as the analog of the logistic differential equation, Eq. 2 is less satisfactory than Eq. 1 by virtue of its unbiological feature that the population can become negative if at any point  $N_t$  exceeds  $K(1 + r)/r$ . Thus, stability properties here refer to stability within some specific neighborhood, whereas in Eq. 1, for example, the stable equilibrium point at  $N = K$  is globally stable (for all  $N > 0$ ) for  $2 > r > 0$ . With this proviso, the stability behavior of Eq. 2 is strikingly similar to that of Eq. 1; see Table 1.

That such single species difference equations should describe populations going from stable equilibrium points to stable cycles as  $r$  increases is not sur-

prising, in view of the general engineering precept that excessively long time delays in otherwise stabilizing feedback mechanisms can lead to "instability" or, more precisely, to stable limit cycles (5, chapter 4; 6). What is remarkable, and disturbing, is that the simplest, purely deterministic, single species models give essentially arbitrary dynamical behavior once  $r$  is big enough ( $r > 2.692$  for Eq. 1,  $r > 2.570$  for Eq. 2). Such behavior has previously been noted in a meteorological context (7), and doubtless has other applications elsewhere. For population biology in general, and for temperate zone insects in particular, the implication is that even if the natural world were 100 percent predictable, the dynamics of populations with "density dependent" regulation could nonetheless in some circumstances be indistinguishable from chaos, if the intrinsic growth rate  $r$  were large enough.

The detailed analysis substantiating these remarks, and deriving Table 1, will be set out in the technical literature. A very brief outline is as follows: (i) For the general nonlinear difference equation

$$N_{t+1} = f(N_t) \quad (3)$$

the locally stable equilibrium point or points can be found by the conventional techniques of linearized stability analysis. For Eq. 1, a fully nonlinear analysis can be given by observing that  $V_t = (N_t - K)^2$  is a Lyapunov function, with the properties  $V_t \geq 0$  and  $\Delta V_t \equiv V_{t+1} - V_t \leq 0$  for all  $N_t > 0$ , for  $2 > r > 0$ : this ensures that the equilibrium point is globally stable. (ii) Next, the possible occurrence of cycles with period 2 may be studied for the equation

$$N_{t+2} = f(f(N_t)) \quad (4)$$

For Eqs. 1 and 2 this has a unique nontrivial equilibrium solution,  $N^* = K$ , for  $r < 2$ , corresponding to the above stable point; as  $r$  increases above 2 this solution of Eq. 4 becomes unstable, and (as

Eqs (1) and (2) introduce his models. They're different, but quite similar, to what we've studied in the notes.

Wow, Lyapunov functions.

Now I'm very glad I learned about those in 3rd year.

For our model we found a 2-point cycle for  $3 < r < 3.445$ .

Table 1. Dynamics of a population described by the difference equations 1 or 2.

Dynamical behavior	Value of the growth rate, $r$	
	Equation 1	Equation 2
Stable equilibrium point	$2 > r > 0^*$	$2 > r > 0$
Stable cycles of period 2 <sup>a</sup>		
2-point cycle	$2.526 > r > 2.000^\dagger$	$2.449 > r > 2.000$
4-point cycle	$2.656 > r > 2.526^\ddagger$	$2.544 > r > 2.449$
8-point cycle	$2.685 > r > 2.656$	$2.564 > r > 2.544$
16, 32, 64, . . .	$2.692 > r > 2.685$	$2.570 > r > 2.564$
Chaotic behavior. (Cycles of arbitrary period, or aperiodic behavior, depending on initial condition.)	$r > 2.692^\S$	$r > 2.570$

\* See Fig. 1a. † See Fig. 1b. ‡ See Fig. 1c. § See Fig. 1, d, e, and f.

Compare these to figures from the Exercises.

(a) has a single asymptotic point.

(b) has a 2-point cycle

(c) has a 4-point cycle

(d) looks like there isn't a pattern

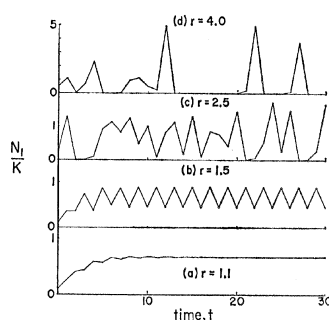
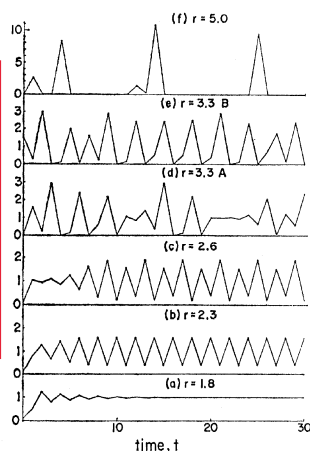


Fig. 1 (left). (a to f) Spectrum of dynamical behavior of the population density,  $N_t/K$ , as a function of time,  $t$ , as described by the difference Eq. 1 for various values of  $r$ . Specifically: (a)  $r=1.8$ , stable equilibrium point; (b)  $r=2.3$ , stable 2-point cycle; (c)  $r=2.6$ , stable 4-point cycle; (d to f) in the chaotic regime, where the detailed character of the solution depends on the initial population value, with (d)  $r=3.3$  ( $N_0/K=0.075$ ), (e)  $r=3.3$  ( $N_0/K=1.5$ ), (f)  $r=5.0$  ( $N_0/K=0.02$ ). Fig. 2 (right). Stability character of the difference equation model of two-species competition, Eq. 5. Specifically, the figure is for  $r_1=r_2=2r$ ,  $K_1=K_2=K$ ,  $\alpha_{11}=\alpha_{22}=1$ ,  $\alpha_{12}=\alpha_{21}=\alpha$ ; under these conditions the criterion for a stable point, Eq. 7, reduces to the requirements  $\alpha < 1$  (as for the analogous Lotka-Volterra differential equation), together with  $[3 - (1 + 8\alpha^2)^{1/2}] / [2(1 - \alpha)] > r > 0$ . The first population, expressed as  $N_1/K$ , is shown as a function of time for  $\alpha=0.5$  and several values of  $r$ : (a)  $r=1.1$ ; (b)  $r=1.5$ ; (c)  $r=2.5$ ; (d)  $r=4.0$ .

point cycle; (d to f) in the chaotic regime, where the detailed character of the solution depends on the initial population value, with (d)  $r=3.3$  ( $N_0/K=0.075$ ), (e)  $r=3.3$  ( $N_0/K=1.5$ ), (f)  $r=5.0$  ( $N_0/K=0.02$ ). Fig. 2 (right). Stability character of the difference equation model of two-species competition, Eq. 5. Specifically, the figure is for  $r_1=r_2=2r$ ,  $K_1=K_2=K$ ,  $\alpha_{11}=\alpha_{22}=1$ ,  $\alpha_{12}=\alpha_{21}=\alpha$ ; under these conditions the criterion for a stable point, Eq. 7, reduces to the requirements  $\alpha < 1$  (as for the analogous Lotka-Volterra differential equation), together with  $[3 - (1 + 8\alpha^2)^{1/2}] / [2(1 - \alpha)] > r > 0$ . The first population, expressed as  $N_1/K$ , is shown as a function of time for  $\alpha=0.5$  and several values of  $r$ : (a)  $r=1.1$ ; (b)  $r=1.5$ ; (c)  $r=2.5$ ; (d)  $r=4.0$ .

illustrated by Fig. 1b) bifurcates into a pair of points, between which the population alternates in a 2-point cycle which is stable provided  $2 < r < 2.526$  for Eq. 1, and  $2 < r < 2.449$  for Eq. 2. Beyond this, the 2-point cycle in turn becomes unstable and each of the points bifurcates into 2 further points, giving a stable 4-point cycle (for example, Fig. 1c), and so on. (iii) As  $r$  continues to increase, there is a limit to this process whereby cycles of period  $2^n$  become unstable and bifurcate into stable cycles of period  $2^{n+1}$ . This limiting value of  $r$ ,  $r_c$ , say, may be calculated [either by brute force, or by analytic methods developed in (8)], and is as set out in the final line in Table 1. For  $r > r_c$ , there ensues a regime of chaos, in which there exist an uncountable number of initial points  $N_0$  for which the system does not eventually settle into any cycle (that is, is not "asymptotically periodic"). (iv) In particular, at yet larger values of  $r$  ( $r > 3.102$  for Eq. 1, and  $r > 2.828$  for Eq. 2), Eqs. 1 and 2 may be shown to have cycles with period 3; that is, solutions such that  $N_{t+3} = N_t \neq N_{t+1} \neq N_{t+2}$ . But Li and Yorke (9) have recently proved an elegant and abstract mathematical theorem, which states that if the general difference equation, Eq. 3, has a 3-point cycle, it necessarily follows that for the same parameter

values there are cycles of period  $n$ , where  $n$  is any positive integer, and furthermore there exist an uncountable number of initial points for which the system is not even asymptotically periodic. Li and Yorke's general theorem for cycles of period 3 may be extended (8) to show that equations of the generic form of 1 and 2 will enter a regime of chaos, with an uncountable number of cycles of integral period along with an uncountable number of aperiodic solutions, beyond the limiting value  $r_c$  defined above.

The dynamical behavior of Eqs. 1 and 2 in this chaotic regime,  $r > r_c$ , is illustrated in Fig. 1, d, e, and f. Figure 1, d and e, are for the same value of  $r$ , and differ only in their initial population value. Note that either of these figures, if looked at only over particular short time intervals, could convey the impression of being locked into a 3-point cycle; around this value of  $r$  there is a tendency to be "captured" into almost-periodic 3-point cycles, in between episodes of apparently chaotic behavior. A detailed understanding of these properties remains an interesting mathematical problem, related to that of determining what fraction of the totality of initial points converge to a 3-point cycle, what fraction to a 5-point cycle, and so on, ending with a determination of the fraction of initial points

which lead to aperiodic behavior. For relatively large values of  $r$  beyond  $r_c$  (for example, Fig. 1f) the population variations become more severe, although the mean population value may be shown to remain around  $K$ ; as  $r$  becomes larger, this mean value is increasingly constituted of a few fairly large population values, together with long sequences of very low population values.

The above discussion is restricted to single species systems obeying difference equations. However, similar considerations are likely to apply, a fortiori, to multispecies situations.

As one among many possible examples, consider a simple difference equation model for competition between two species

$$N_1(t+1) = N_1(t) \exp\{r_1[K_1 - \alpha_{11}N_1(t) - \alpha_{12}N_2(t)]/K_1\} \quad (5a)$$

$$N_2(t+1) = N_2(t) \exp\{r_2[K_2 - \alpha_{21}N_1(t) - \alpha_{22}N_2(t)]/K_2\} \quad (5b)$$

Just as Eq. 1 may be regarded as a difference equation analog of the logistic, Eq. 5 may be regarded as an analog of the familiar Lotka-Volterra differential equation model for two-species competition. As usual,  $r_i$  are the growth rates,  $K_i$  the carrying capacities, and  $\alpha_{ij}$  the competition coefficients. The dynamical properties of such Lotka-Volterra differential equations are straightforward: the two species coexist, with a globally stable equilibrium point, if and only if

$$D > 0 \quad (6)$$

where  $D$  is defined as  $D = \alpha_{11}\alpha_{22} - \alpha_{12}\alpha_{21}$ . Failing this, one or the other species is extinguished. But for the system of difference equations, Eq. 5, the criterion for the existence of a stable two-species equilibrium point is more restrictive, namely

$$A > D > 0 \quad [\text{if } A < 2] \quad (7a)$$

$$A > D > 2A - 4 \quad [\text{if } A \geq 2] \quad (7b)$$

Here  $D$  is as above, and  $A$  is defined as  $A = (\alpha_{11}K_2/r_2N_2^*) + (\alpha_{22}K_1/r_1N_1^*)$ , with  $N_1^*$  and  $N_2^*$  the equilibrium solutions of Eq. 5. [The methodology for stability analysis of such two-species difference equations is indicated elsewhere (6)]. If the right-hand side of Eq. 7a is violated, one of the species is eliminated, as in the differential equation model, Eq. 6. If any of the other inequalities in Eq. 7 is transgressed, the two species continue to coexist, but there is no longer a stable equilibrium point. Numerical studies reveal a regime of stable cycles, giving way to one

"elegant and abstract" probably means "complicated"

Sounds like a useful result, though, might be worth looking up.



of apparent chaos, as for the single species systems discussed in detail above. The behavior of the system of Eq. 5 in these various regimes is illustrated in Fig. 2.

Equations 1 and 2 are two of the simplest nonlinear (density dependent) difference equations that can be written down. Their rich dynamical structure, and in particular the regime of apparent chaos wherein cycles of essentially arbitrary period are possible, is a fact of considerable mathematical and ecological interest, which deserves to be more widely appreciated. Without an understanding of the range of behavior latent in such deterministic difference equations, one could be hard put to make sense of computer simulations or time-series analyses in these models.

ROBERT M. MAY

Biology Department, Princeton University, Princeton, New Jersey 08540

#### References and Notes

1. Previous work in this general area of population biology consists largely of remarks on the relation between differential equation models and difference equation models [H. R. Van der Vaart, *Bull. Math. Biophys.* **35**, 195 (1973); R. M. May, *Am. Nat.* **107**, 46 (1972); J. M. Smith, *Models in Ecology* (Cambridge Univ. Press, Cambridge, 1974)] and linearized analyses of Eqs. 1 and 2 showing the equilibrium point to be locally stable only if  $2 > r > 0$  [L. M. Cooke (2) for Eq. 1; J. M. Smith, *Models in Ecology*, for Eq. 2].
2. L. M. Cooke, *Nature (Lond.)* **207**, 316 (1965).
3. A. Macfadyen, *Animal Ecology: Aims and Methods* (Pitman, London, ed. 2, 1963).
4. J. M. Smith, *Mathematical Ideas in Biology* (Cambridge Univ. Press, Cambridge, 1968); J. R. Krebs, *Ecology: The Experimental Analysis of Distribution and Abundance* (Harper & Row, New York, 1972).
5. R. M. May, *Stability and Complexity in Model Ecosystems* (Princeton Univ. Press, Princeton, N.J., 1973).
6. —, M. P. Hassell, G. R. Conway, T. R. E. Southwood, *J. Anim. Ecol.*, in press.
7. E. N. Lorenz, *Tellus* **16**, 1 (1964); *J. Atmos. Sci.* **20**, 448 (1963).
8. R. M. May and G. F. Oster, in preparation.
9. T.-Y. Li and J. A. Yorke, *SIAM (Soc. Ind. Appl. Math.) J. Math. Anal.*, in press.
10. I am indebted to many people, and particularly to J. A. Yorke, for stimulating discussions.

3 June 1974

### Hycanthone Analogs: Dissociation of Mutagenic Effects from Antischistosomal Effects

**Abstract.** N-Oxidation at the diethylamino group of hycanthone, of lucanthone, and of two chlorobenzothiopyranoindazoles resulted in a marked reduction in mutagenic activity, while antischistosomal activity was retained or even enhanced. Introduction of chlorine into the 8-position of benzothiopyranoindazoles reduced acute toxicity but had no effect on chemotherapeutic potency. These dissociations of biological activities indicate that safer antischistosomal compounds of this class can be developed.

The geographical distribution and the nature of human schistosomiasis require special care in the selection of chemotherapeutic agents for the treatment of this infection. More than 200 million human subjects are infected with schistosomes and the incidence is on the increase. Even a low frequency of delayed serious complications, produced by mutagenic, teratogenic, and carcinogenic actions of a drug, can involve a large absolute number of individuals. Populations protected by schistosomes are not protected by national drug laws or regulatory agencies. Moreover, in an undetermined number, and possibly the majority, of subjects infected with *Schistosoma hematobium*, overt clinical and pathological manifestations disappear in adulthood (1). This must be taken into account when considering a drug for the mass treatment of children whose life expectancies are longer and whose reproductive potentials are greater than those of adults. As was stated by Rubidge *et al.* (2), "urinary tract bilharziasis is

a relatively mild disease in South Africa and serious sequelae are rare. Hence, therapy must be safe."

It is estimated that during the past 6 years, in Brazil, Africa, and the Middle East, at least 700,000 human subjects infected with *S. hematobium* and *S. mansoni* have been treated with the antischistosomal thioxanthene derivative hycanthone (the drug is ineffective in infections produced by *S. japonicum* prevalent in mainland China and the Philippines) (3). Reports from a variety of laboratories have indicated that

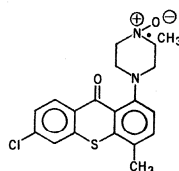


Fig. 1. Mutagenic activity: none detectable (less than 0.1 percent as active as hycanthone). Antischistosomal activity: intramuscular, 0.4; oral, 0.3.

hycanthone is mutagenic (4) and teratogenic (5), and that it induces prophage (6), mitotic crossing-over (7), cytogenic changes (8), and malignant transformations (9); hycanthone is carcinogenic in mice infected with *S. mansoni* (10). As pointed out by Firminger (11), a report (12) which seemingly did not support the last observation was based on such a small number of animals that no significant negative results could have been obtained. Since a number of compounds chemically related to hycanthone exhibit antischistosomal activity, the question arose whether structural alterations can bring about a dissociation of undesirable toxicological properties from chemotherapeutic activity. Data summarized below indicate that this is the case.

A chloroindazole analog (IA-4, structure in Table 1) of hycanthone has the same antischistosomal activity in mice as hycanthone (13), while its acute toxicity and its hepatotoxicity are lower (13, 14). Compound IA-4 failed to induce demonstrable malignant transformations in cells infected with Rauscher virus (9). Its mutagenic activity was found to be lower in *Salmonella* (15), bacteriophage T4 (15), and mouse lymphoblasts (16); no mutagenic effects were detected in yeast (17); no cytogenetic effects were detected in rat bone marrow cells (18). Furthermore, in contrast to hycanthone and to a number of chemical carcinogens, IA-4 failed to induce breaks in rat liver DNA (19). Another indazole analog (IA-3) had lower antischistosomal activity; but since there is decreased acute toxicity, the chemotherapeutic index of IA-3 approximately equals that of IA-4 (13).

We found that chloro substitution in position 8 produced a marked decrease in the acute toxicity of the indazole analogs for mice. For example, the median intramuscular lethal dose ( $LD_{50}$ ) of IA-3 and of IA-4 was more than seven times higher than that of the corresponding deschloro derivatives.

In further studies of the effect of structural modifications on antischistosomal activity and on mutagenicity, N-oxides of active thioxanthenones and benzothiopyranoindazoles were prepared. The parent bases were oxidized with *m*-chloroperbenzoic acid in dichloromethane solution, and after chromatography ( $Al_2O_3$ ) the N-oxides so obtained were converted to their water-soluble methanesulfonate salts. N-Oxidation at the diethylaminoethyl group consistently resulted in a marked reduction in mutagenicity for *Salmonella*

15 NOVEMBER 1974

647