Part I

Dynamical systems – general
Introduction to Part I

Not chaos-like, together crushed and bruised,
But, as the world harmoniously confused:
Where order in variety we see,
And where, though all things differ, all agree.
 Alexander Pope, *Windsor Forest.*

In this part of the book we provide the basic mathematical background for dynamical systems and chaos theory. Some of the ideas introduced here will be applied to various astrophysical systems in the second part of the book. Our discussion here, while not particularly rigorous, will, however, be rather theoretical and abstract. I believe that a reasonable precision in building the basis for further understanding and research is mandatory. Throughout the discussion we continually give specific examples and often return to them in other places in the book. These examples, including some systems that are important by themselves, illustrate the various abstract concepts.

I have made an effort to interest readers, whose background is astronomy and astrophysics, by starting with astrophysical examples. After all, dynamical system theory and chaos have their origins in the studies of the three-body problem and celestial mechanics by Poincaré at the end of the nineteenth century. Fluid turbulence, an important unsolved scientific problem, is now being approached using methods from chaos and dynamical system theory. It has also had many important applications in astrophysics. Readers who are interested more in applications and less in theory and mathematical structure are particularly encouraged to become acquainted with the main concepts and results of this part of the book. Technical details may be skipped, certainly during first reading. When dealing with the second part (applications), the interested reader may return to the relevant material in the first part and study it more deeply.

The contemporary intensive study of chaotic behaviour in nonlinear dynamical systems owes its existence to the availability of fast digital computers. Irregular and
aperiodic solutions of ordinary differential equations were discovered long ago, but they were, in a way, forgotten. In contrast, enormous analytical progress has been made in the study of linear systems. This progress is obviously reflected in the subjects taught in mathematics and the physical sciences at all levels. The bias towards the linear is so strong that a student or a scientist naturally attacks every problem with linear ‘weapons’ like the various techniques of eigenvalue problems, normal mode analysis, Fourier and other transforms, linear perturbations etc. These techniques often fail when the interest is the studied system’s nonlinear behaviour. It may then seem that the only choice left is computer simulation and thus one may think that the ‘linear paradise’ of powerful analytical methods is lost and what remains is just a ‘numerical hell’. I have often encountered the term nonlinear being used in astrophysics as a synonym for numerical simulation. This should not be so, as substantial analytical knowledge on nonlinear systems is now available. A lot of information on a nonlinear system may be deduced by analytical and perturbative methods, accompanied by specific numerical computations that are, in general, much easier than brute force direct numerical simulations. In any case, any such knowledge is very useful in devising the right numerical method for a full scale computer simulation and in understanding its result. For the student of modern methods in nonlinear dynamics ‘paradise’ can not only be regained, but it also reveals new and unexpected beauty. It may be manifested in astounding fractal structures and in fundamental invariances, symmetries and some other more general properties.

One such property, for example, is both very interesting and useful. It is the fact that nonlinear systems have a generic behaviour. Some very different-looking systems, describing completely unrelated natural phenomena, fall into classes that have identical behaviour. It seems that though all nonlinear systems differ, all agree in their fundamental properties. A typical example is the finding that transitions between different types of behaviour in a layer of fluid heated from below (transition to convective motions) have exactly the same properties as bifurcations (behaviour changes) of a class of quadratic mappings (like the logistic map obtained in population dynamics). As a result, the essence of very complex phenomena may often be described by a simple model.

We devote the remainder of this chapter to the introduction of the most basic mathematical notion used in the book, that of a dynamical system. It may be loosely defined as a set of rules by application of which the state of a physical (or some other, well defined) system can be found, if an initial state is known. Symbolically, we specify the state of a system by its being a well defined point, \( x \), in some space. The vector notation reflects the fact that several variables may be needed to completely specify the state of a system. The space in which the state \( x \) lives is naturally called the state space.
The previously mentioned set of rules may be thought of as an *evolution operator*, $T$, acting on a state variable and transforming it into some other state, $x'$. This is expressed formally by

\[ x' = T x \]

The evolution may be continuous in time or performed in discrete steps. In the former case the state variable is a function of time and the continuous time evolution operator (labelled by a subscript $t$) is understood to carry the system from an initial condition (the state at $t = 0$) to the state at time $t$. In the latter case the discrete operator (with subscript $n$) evolves the system from an initial state $x_0$ through $n$ discrete steps. We can thus formally write for the two cases

\[ x(t) = T_t x(0) \] (1.1)

and

\[ x_n = T_n x_0 \] (1.2)

respectively.

The continuous case (1.1) can be realised, for example, by a set of two *ordinary differential equations* (ODE) like

\[ \frac{du}{dt} = f(u, v) \quad \frac{dv}{dt} = g(u, v) \] (1.3)

where the state variable is two dimensional, $x = (u, v)$, and where $f$ and $g$ are some well-behaved functions. The evolution operator $T_t$ symbolises the solution of (1.3), that is, the unique determination of $u(t)$ and $v(t)$ from the initial values $u(0)$ and $v(0)$.

The discrete case (1.2) can be illustrated by an example of an *iterated mapping* (or map) e.g.

\[ x_{j+1} = F(x_j) \quad \text{with} \quad j = 0, 1, ... \] (1.4)

where $F$ is some well-defined function. With the help of (1.4) the state variable (in this example it is one-dimensional) can be propagated in discrete steps from an initial value $x_0$ to its value after $n$ iterations, $x_n$. The operator $T_n$ reflects here the repeated ($n$ times) application of $F$.

Thus, the action of the evolution operator on an initial state (point in state space) transforms the latter continuously, or in discrete steps, into another state. Thus, we may imagine an extra ‘axis’ (e.g. time $t$ or the iteration number $n$), along which the state space is being transformed. The state space plus the evolution axis is called the *configuration space*.

So far we have considered only finite-dimensional state spaces, but this is clearly not the case when the dynamical system is represented by a *partial differential*
Introduction to Part I

Such equations can be regarded as the limit \( (m \to \infty) \) of a set of \( m \) ordinary differential equations. Let \( \{x^1(t), x^2(t), \ldots, x^m(t)\} \) be the components of the state variable \( x \) satisfying such a set of \( m \) ordinary differential equations. The system can be compactly written, using vector notation as

\[
\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}, t)
\]

or explicitly, using components

\[
\frac{dx^1}{dt} = F^1(x^1, x^2, \ldots, x^m, t) \\
\frac{dx^2}{dt} = F^2(x^1, x^2, \ldots, x^m, t) \\
\vdots \\
\frac{dx^m}{dt} = F^m(x^1, x^2, \ldots, x^m, t)
\]

The vector state components can also be written in the form \( x(\xi, t) \equiv x^k(t) \), with \( k = 1, 2, \ldots, m \). The limit \( m \to \infty \) is realised by replacing the discrete set of natural numbers \( 1 \leq k \leq m \) by a continuous ‘space’ variable, say \( \xi \). The state variable is then written as \( x(\xi, t) \) and is a continuous function of the space variable, \( \xi \), and time. The evolution operator acts, in this case, on an infinite-dimensional function space, whose ‘points’ (defining states of the system) are functions. A function at some initial time \( x(\xi, 0) \), is transformed by the evolution operator into \( x(\xi, t) \),

\[
x(\xi, t) = T_t(\xi) x(\xi, 0)
\]

where we have explicitly stressed that the evolution operator depends now on \( \xi \). In most cases it includes derivatives of \( x \) with respect to \( \xi \), thus actually it depends on the function values in the neighbourhood of \( \xi \) as well.

As an example, consider the linear diffusion equation in one space dimension (\( \xi \) is a scalar) for the state function \( x(\xi, t) \), defined on some interval \( a \leq \xi \leq b \).

\[
\frac{\partial x}{\partial t} = \kappa \frac{\partial^2 x}{\partial \xi^2} \tag{1.5}
\]

where \( \kappa \) is a constant.

The operator \( T_t(\xi) \) symbolises the evolution of the state function from its initial value \( x(\xi, 0) \) to its value at time \( t - x(\xi, t) \), i.e., the solutions of (1.5). Equation (1.5) has a unique solution, and thus the evolution operator is well defined, if suitable boundary conditions at the interval edges, \( \xi = a \) and \( \xi = b \), are specified.

A discrete version of a partial differential equation is obtained if all the independent variables are discretised. In the case of (1.5) these variables are \( \xi \) and \( t \). Thus
the function $x$ has to be replaced by an array, $x_{kj} = x(\xi_k, t_j)$ say. This can be done by setting a spatial interval as $\Delta\xi \equiv (b - a)/m$ with some integer $m$, for example, and any time interval $\Delta t$ and defining recursively

$$
\begin{align*}
\xi_{k+1} &= \xi_k + \Delta\xi \\
\xi_0 &= a \\
t_{j+1} &= t_j + \Delta t \\
t_0 &= 0
\end{align*}
$$

As an example we express the derivatives in (1.5) by finite differences. The equation is thus written as

$$
x_{kj+1} = x_{kj} + \alpha(x_{k-1,j} - 2x_{kj} + x_{k+1,j}) 
$$

for $j = 0, 1, 2, \ldots$ and any $k = 0, 1, \ldots, m$ (so that $\xi_k$ is inside the definition region) and where $\alpha \equiv \kappa/Delta1t/(\Delta\xi)^2$.

Equations like (1.6) can be regarded as a mapping from a time step $j$ to the time step $j + 1$. The discrete evolution operator in this case, $T_n(\xi)$, advances the system from an initial state $x_{k0}$ to the state at the $n$th time step, $x_{kn}$, through the repeated ($n$ times) application of the map (1.6) and we can write

$$
x_{kn} = T_n^{[k,k+1]} x_{k0} 
$$

In this case the discrete (time) evolution at a particular position, $\xi_k$, depends explicitly also on the state function values in the neighbourhood of $\xi_k$. This fact is reflected by writing the evolution operator in (1.7) as dependent on $k$ and the neighbouring indices $k - 1$ and $k + 1$. Such discrete dynamical systems are called *cellular automata*. The question if the discrete system (1.6) can lead to an approximate solution of (1.5) is irrelevant here.

The above discussion covers essentially all the types of dynamical systems that will be discussed in this book in both of its parts.
Astrophysical examples

Few things are harder to put up with than the annoyance of a good example.
Mark Twain, *Pudd’nhead Wilson*.

Rather than starting from abstract mathematical definitions related to dynamical systems and the concepts used to analyse them, I prefer to start from the outset with a few familiar examples. The systems described below are related to the paradigms of deterministic chaos, some of which have indeed been the ones leading to the discovery, definition and understanding of chaotic behaviour. Instead of repeating here the so often quoted examples such as biological population growth, nonlinearly driven electrical oscillations, weather unpredictability, three-body Hamiltonian dynamics and chemical reaction oscillations and patterns, I shall attempt to motivate the reader by trying to find such examples among simplistic models of astrophysical systems. Obviously, the underlying mathematical structure of these will be very similar to the above mentioned paradigms. This only strengthens one of the primary lessons of nonlinear dynamics, namely that this is a generic, universal approach to natural phenomena.

Examples and analogies may sometimes be misleading and decide nothing, but they can make one feel more at home. This was, at least, the view of Sigmund Freud, the father of psychology, whose advice on matters didactic should not be dismissed. Indeed, as stressed before, these examples are the readers’ old acquaintances from their astrophysics educational ‘home’. In the next chapter, where the basic notions characterising chaotic behaviour will be dealt with in detail, these examples will sometimes be used again for demonstrating abstract concepts.

2.1 Stellar population dynamics

The primary tool of the physical scientist is the infinitesimal calculus. Consequently, the relevant system variables are described as *continuous* functions of time.
2.1 Stellar population dynamics

The physical laws are posed as differential equations, either ordinary (ODE) or partial (PDE). In the biological and social sciences, however, the dynamic laws are often formulated as relationships between the variables at discrete time intervals whose size relates to typical time scales of change or sampling. Moreover, discretisation naturally appears when differential equations are being solved on digital computers. The basic dynamic relations are, thus, expressed as difference equations.

A typical example of this appears in the investigation of population dynamics. May (1976) was the first to recognise the possible complexity arising from a very simple model in this context. We shall choose stars here as the ‘organisms’ in the population dynamics model.

The number of stars of a particular type, all identical with mass \( m \), all born together as the result of a supernova-triggered star formation event, say, and all ending their lives in a supernova explosion, is denoted by \( n_i \). The subscript \( i \) refers to the \( i \)th generation and we view this stellar population evolution as a discrete series of events. We further assume that the region where this stellar population lives is well mixed and contains a fixed amount of matter, stars plus gas. Assume now that the number of stars in the next generation, \( n_{i+1} \) is proportional to \( n_i \). This is a very reasonable assumption, on average, with the proportionality constant, \( c_1 \), say, being the average number of stars formed as the result of a single supernova explosion.

The efficiency of star formation must depend also on gas density, or in the context of our fixed volume zone, on the gas mass. The total amount is constant, say, \( M \) and therefore the number \( n_{i+1} \) depends on the available gas mass in the \( i \)th generation, which is \( M - m n_i \). In general, this dependence may take the form of a power law \( c_2 (M - mn_i)^a \), but to simplify the discussion we take \( a = 1 \). The constant \( c_2 \) reflects some complex properties of star formation. The stellar population dynamical law might therefore look something like

\[
\frac{n_{i+1}}{n_i} = c n_i (M - mn_i)
\]

(2.1)

where the two proportionality constants have been combined into one \( c \). For \( c, M \) and \( m \) constant, this equation can be rewritten as

\[
x_{i+1} = 4rx_i(1 - x_i)
\]

(2.2)

where \( x_i \equiv mn_i / M \) and \( r \equiv cM/4 \).

This is the well known logistic map. It is a typical, well studied nonlinear quadratic map (or mapping) and was the equation that led to the discovery of the period-doubling bifurcation route to chaos with its universal properties, in the celebrated work of Feigenbaum (1978) in the United States. Coullet & Tresser (1978) worked on this problem independently in France and reached similar conclusions.
More sophisticated models of stellar population dynamics can also be made, taking into account the fact that a prescribed fraction of matter is not returned to the interstellar medium, allowing for different stellar types and using a suitable power-law dependence on the amount of gas. Such models could lead to two-dimensional (or multidimensional) mappings, namely two (or more) equations with a suitable number of variables. The simplest model, leading to the mapping (2.2), is however sufficient for our purposes here.

Equation (2.2) can be easily iterated with the help of a computer (or even a programmable calculator), starting from some given $x_0 < 1$ (see below). The result depends, as may be expected, on the value of the constant $r$. We would like $x_k$ to be non-negative for any $k$; therefore its value must also always be $\leq 1$ (to prevent a negative value in the next generation) and it is also necessary that the constant $r$ be $r < 1$. The logistic equation (2.2) is then a mapping of the interval $[0, 1]$ onto itself. It is qualitatively obvious that a very small value of $r$ (too small star formation efficiency or too little gas in the system) will lead to extinction. What happens for larger values of $r$ is less clear.

Since the logistic difference equation exhibits many of the important properties of nonlinear systems bifurcations, it may be seen as a prototype for one-dimensional (and even multidimensional) nonlinear mappings. Its study and the related concepts will be elaborated on in the next two chapters. Here, we shall just introduce a few simple but important notions and the graphical procedure, which are very useful in the qualitative study of iterated maps and bifurcation theory.

In general, a one-dimensional mapping, like (2.2), can be written in the functional form

$$x_{i+1} = F(x_i) \quad (2.3)$$

The successive iterations give rise to a sequence of $x$ values called *iterates*. The iterates can be followed by plotting $y = F(x)$ and the line $y = x$ on the same graph (see Figure 2.1) and moving successively vertically and horizontally between these two curves. The points at which the line and the curve intersect (i.e., $F(x) = x$) correspond to values of $x_i$ for which the iteration sequence values remain fixed. These are referred to as *fixed points* of the map. As is apparent from Figure 2.1 our mapping has two fixed points, $x^*$, which can be found analytically by solving

$$x^* = 4r x^* (1 - x^*) \quad (2.4)$$

The solutions are: $x^* = 0$ and $x^* = 1 - 1/4r$.

Note that only for $r > 0.25$ do both fixed points lie in the interval $[0, 1]$.

A fixed point of a mapping can be *stable* (or *attracting*), if an iteration starting in its neighbourhood will bring the iterates closer and closer to the point. Conversely, if an iteration of the map starting close to a fixed point results in the iterates moving...
2.1 Stellar population dynamics

Fig. 2.1. Geometrical construction for finding successive iterates and fixed points of a mapping. The first three iterations are shown in (a). These are followed by spiralling into (shown in (b)), or out of (shown in (c)), the fixed point $x^*$, depending on its stability.

away from it, it is unstable (or repelling). Without entering into a complete analysis of the fixed point stability of the logistic map (this will be done in Chapter 4), we just note that the stability here depends on the value of $r$. More specifically:

(i) $x^* = 0$ is stable for $0 < r < 0.25$ and unstable for $r > 0.25$.
(ii) $x^* = 1 - 1/4r$ is unstable for $0 < r < 0.25$ and stable for $0.25 < r < 0.75$.
(iii) both fixed points are unstable for $r > 0.75$.

It is therefore clear that iterations, starting from any initial $x_0$, will approach the constant value $x^* = 0$ (extinction) as long as $r < 0.25$, as expected. For any $0.25 < r < 0.75$ the second fixed point will be approached as closely as desired after a sufficient number of iterations. The first fixed point is thus the attracting point of the map for $r < 0.25$, while the second one becomes the attracting point for $0.25 < r < 0.75$. At $r = 0.25$ the map changes its qualitative behaviour. Such a point on the parameter axis (here $r$) is called a bifurcation point.

At the value of $r$ at which the second fixed point loses its stability, $r = 0.75$, a more interesting bifurcation occurs. Both fixed points are now unstable and successive iterations will result in the values of $x_i$ alternating repeatedly between two fixed values. This happens only if $r$ is not too far from 0.75. This periodic cycle of the map, which includes just two points, is now attracting; it is called a limit cycle.

We can construct the bifurcation diagram of our map by plotting the limiting values of the iterate (for large $i$) as a function of the control parameter $r$, starting from $r = 0.75$; this is shown in Figure 2.2. We note that the bifurcation at $r = r_1 = 0.75$ is not the last one. Quite the contrary, bifurcation points keep appearing with increasing density. The aforementioned limit cycle loses its stability and a more complicated limit cycle, with four values, appears at some value of $r$ (actually at