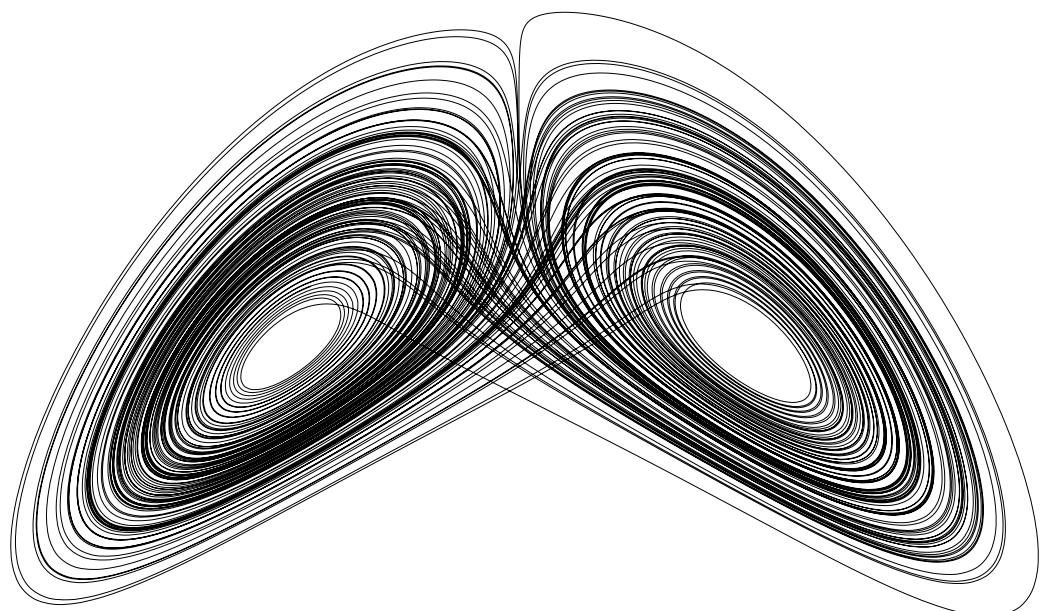


Dynamics and Chaos

MA30060 2021–2022
University of Bath



MA30060: Dynamics and chaos

Semester 1, 2021–22

Lecturer: Philippe H. Trinh

Introduction to the course

Chaos theory refers to the branch of mathematics that studies particular dynamical systems that exhibit great sensitivity to perturbations. Here, the popular example is that a butterfly's wings flapping in a faraway place like Tokyo can cause a tornado in London. In this course, we shall put this (somewhat inaccurate and sensationalised) phrase on a firm mathematical foundation; at the same time, we also explore some of the wider social and scientific consequences of chaos. In particular, we will focus on developing and studying the simplest possible mathematical models that exhibit chaos (one-dimensional discrete-time difference equations). From that topic, numerical and analytical tools are developed, and several possible definitions of chaos are then proposed.

Proofs indicated in brackets: **[** ... **]** are not examinable. Proofs *NOT* enclosed by **[** ... **]** are examinable. If in doubt ask for clarification, and assume examinable.

Acknowledgements and changes

These typeset notes have been written and added-to by multiple authors through the years, notably J.H.P. Dawes, B. Adams, J. Hook, and P. Trinh (since 2018).

Notes about the lecture format

For ease of counting, we will refer to each day of the course as "Lecture 1, Lecture 2, ..." and not distinguish between lecture and problem class. There are three lectures per week, and the term runs for twelve weeks. This course is heavy on visualisation and drawings—please make sure to bring pens or pencils of different colours and a ruler so that you can produce your own set of beautiful notes.

CONTENTS

Contents	v
1 On the origins of chaos	1
2 Dynamical systems, cobwebbing, and the logistic equation	9
3 Orbits, fixed/periodic points, and invariant sets	13
4 Introduction to asymptotic analysis	17
5 Local analysis of bifurcations I	21
6 Local analysis of bifurcations II	27
7 Stability and chaos	31
8 Topological conjugacy	37
9 Symbolic dynamics	39
10 The logistic map with $\mu > 4$	41
11 Conjugacy between the logistic map, F , and shift map, σ	43
12 Subshifts of Finite Type (SSFT)	49
13 From continuous maps, $F : I \rightarrow \mathbb{R}$, to symbolic maps	53
14 Horseshoes, N -cycles, and chaos	59
15 The tent map	63
16 Lorenz papers	69

CHAPTER 1

ON THE ORIGINS OF CHAOS

‘Chaos’ is a topic that has garnered a great deal of popular attention in books, television, and general discourse. We often refer to certain real-life phenomena as ‘chaotic’ without any firm definition of the term. For instance, would you consider any of the following as exhibiting chaos?

- The value of a stock on the stock market;
- The chance of rain in a week;
- Whether the sun rises tomorrow;
- The swinging of a pendulum.

In fact, all of these examples can be modelled in systems that do exhibit chaos in general, but not all such phenomena are chaotic. Some questions that we might ask are:

- What is the exact definition of ‘chaos’?
- Is there a difference between being chaotic and being random or unpredictable?

One of the main objectives of this course is to define very precisely what we mean by mathematical chaos.

1.1 Chaos in the N -body problem

Let there be three bodies (or planets) with positions $\mathbf{r}_i(t) = (x_i(t), y_i(t), z_i(t))$ for $i = 1, 2, 3$ with respective masses m_i . By Newton’s law of gravity, each planet is subjected to an attractive force by the other planets in the system. The forces are proportional to their masses and inversely proportional to the square of their distance. Hence by Newton’s second law (mass times acceleration is equal to the sum of forces)

$$\begin{aligned}\widetilde{m_1 \mathbf{r}_1''(t)} &= \widetilde{-Gm_1 m_2 \frac{\mathbf{r}_1 - \mathbf{r}_2}{|\mathbf{r}_1 - \mathbf{r}_2|^3}} - \widetilde{Gm_1 m_3 \frac{\mathbf{r}_1 - \mathbf{r}_3}{|\mathbf{r}_1 - \mathbf{r}_3|^3}}, \\ m_2 \mathbf{r}_2''(t) &= -Gm_2 m_1 \frac{\mathbf{r}_2 - \mathbf{r}_1}{|\mathbf{r}_2 - \mathbf{r}_1|^3} - Gm_2 m_3 \frac{\mathbf{r}_2 - \mathbf{r}_3}{|\mathbf{r}_2 - \mathbf{r}_3|^3}, \\ m_3 \mathbf{r}_3''(t) &= -Gm_3 m_1 \frac{\mathbf{r}_3 - \mathbf{r}_1}{|\mathbf{r}_3 - \mathbf{r}_1|^3} - Gm_3 m_2 \frac{\mathbf{r}_3 - \mathbf{r}_2}{|\mathbf{r}_3 - \mathbf{r}_2|^3}.\end{aligned}$$

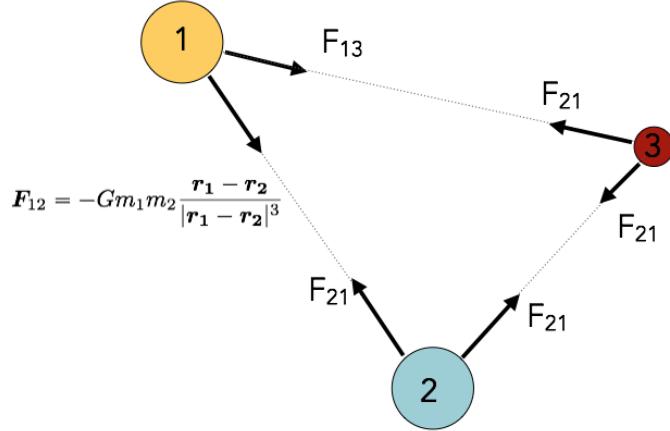


Figure 1.1: Forces in the three-body problem

The basic configuration of the three-body problem is shown in Fig. 1.1.

This system of three differential equations is solved with initial positions and initial velocities of the three bodies:

$$\mathbf{r}_i(0) = \text{const.} \quad \text{and} \quad \mathbf{r}'_i(0) = \text{const.}$$

Together, the solution of the set of differential equations and initial conditions is known as the $N = 3$ -body problem. This is not an easy problem to solve (in fact, it is analytically intractable).

The case of the $N = 2$ or two-body problem can be completely solved. The analysis indicates that each member planet of a two-body system will orbit about the system's center of mass in either an elliptic, parabolic, or hyperbolic motion (essentially the conic sections). However, the case of $N > 2$ was discovered to be much more difficult. An example of the numerical solution is given in Fig. 1.2.

1.2 Chaos in the double pendulum

Another classic example of chaos occurs in the double pendulum—essentially two pendulums coupled to one another. Place the origin at the pivot of the upper pendulum. Let the upper pendulum mass be m_1 and located a position (x_1, y_1) and with an angle θ_1 measured with respect to the vertical. Similarly (x_2, y_2) is the position of the lower pendulum with mass m_2 and angle θ_2 . All the position/angle quantities are functions of time, t .

In order to derive the differential equations for the pendula, we must apply Newton's law and write down a balance of horizontal and vertical forces in the system. For example, a balance of vertical forces on mass m_1 gives

$$\begin{aligned} m_1 x_1''(t) &= -T_1 \sin \theta_1 + T_2 \sin \theta_2, \\ m_1 y_1''(t) &= T_1 \cos \theta_1 - T_2 \cos \theta_2 - m_1 g, \\ m_2 x_2''(t) &= -T_2 \sin \theta_2, \\ m_2 y_2''(t) &= T_2 \cos \theta_2 - m_2 g, \end{aligned}$$

where g is gravity, and $T_{1,2}$ are the tensions in the respective rods. By relating the coordinates with the angles, we can re-write the above system as a set of two differential equations:

$$\begin{aligned} \theta_1''(t) &= f_1(\theta_1, \theta_2) \\ \theta_2''(t) &= f_2(\theta_1, \theta_2), \end{aligned}$$

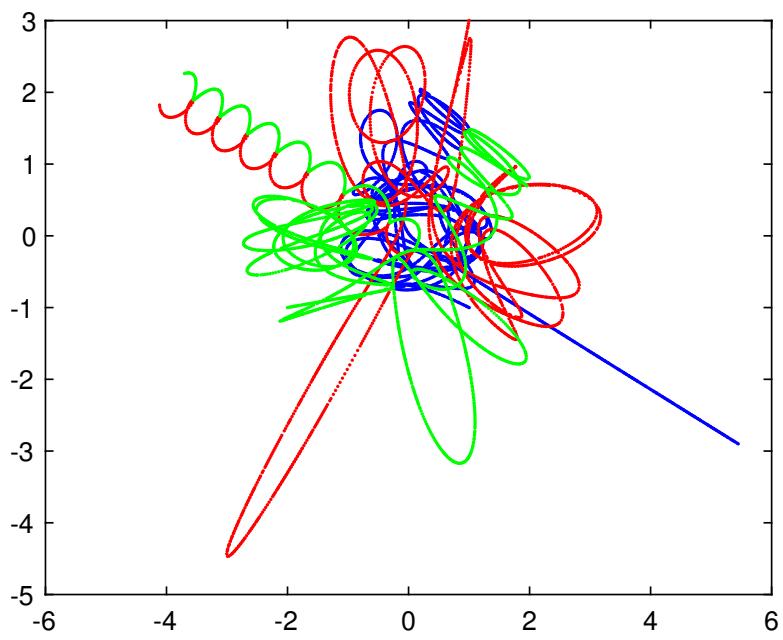
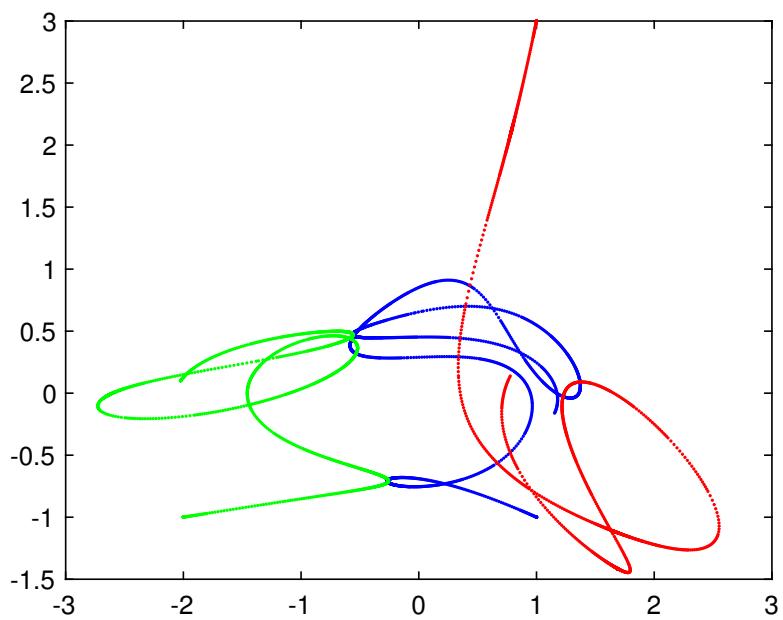


Figure 1.2: Simulation of three-body problem in a frame of reference where the positions are planar; shown at early time (top) and later time (bottom). The orbits are shown only in the plane of motion.

that also depend on the specification of parameters L_1 , L_2 , and m_1 and m_2 . This is then solved with initial conditions on the angles and velocities. For instance, beginning the system at rest and with

$$\begin{aligned}\theta_1(0) &= \theta_{10} & \theta'_1(0) &= 0 \\ \theta_2(0) &= \theta_{20} & \theta'_2(0) &= 0.\end{aligned}$$

Like the N -body problem, this is another difficult, analytically intractable problem to solve. In the lecture, we will examine some numerical simulations of the double pendulum.

1.3 Chaos in the Lorenz system

The Lorenz system, as proposed by Lorenz in 1963, is given by the misleadingly simple-looking set of differential equations:

$$\begin{aligned}\frac{dx}{dt} &= \sigma(y - x), \\ \frac{dy}{dt} &= \rho x - y - xz, \\ \frac{dz}{dt} &= xy - \beta z,\end{aligned}$$

along with three initial conditions on the positions of (x, y, z) at $t = 0$. In the original meteorological context, the equations relate to a two-dimensional fluid layer uniformly warmed from below and cooled from above. Here x is proportional to the rate of convection, y to the horizontal temperature variation, and z to the vertical temperature variation. The constants σ , ρ , and β are the Prandtl number, Rayleigh number, and a number characterizing the physical dimensions of the layer. A typical simulation result is shown in Fig. 1.3. An excerpt from his paper is given in Fig. 1.4.

1.4 Forwards!

During the lecture, and through our examination of the three above systems, we will formulate a number of questions about the nature of chaos and its properties. Perhaps the two most crucial questions we ask are:

1. What is the **simplest** mathematical model for chaos?
2. What does analysis of this model show us?

In essence, this is the focus of the Dynamics and Chaos course. These questions are also of the sort that inspired a 1976 review article by May (Fig. 1.5), which was very influential in the development of ‘simplified’ chaos courses like the one you are currently taking.

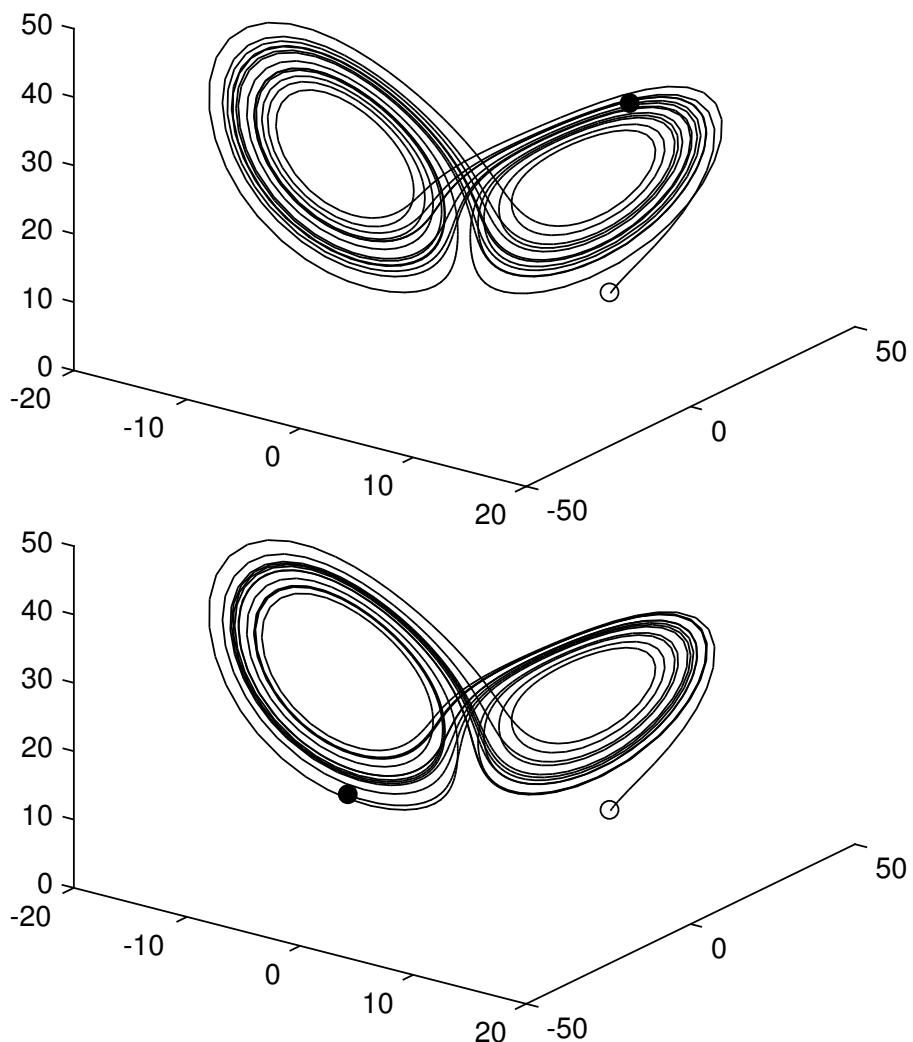


Figure 1.3: Simulations of the Lorenz system with $\beta = 10$, $\sigma = 8/3$, and $\rho = 28$. (Top) Initial condition of $(10, 10, 10)$; (bottom) initial condition of $(10 + 10^{-4}, 10, 10)$.

Deterministic Nonperiodic Flow¹

EDWARD N. LORENZ

Massachusetts Institute of Technology

(Manuscript received 18 November 1962, in revised form 7 January 1963)

ABSTRACT

Finite systems of deterministic ordinary nonlinear differential equations may be designed to represent forced dissipative hydrodynamic flow. Solutions of these equations can be identified with trajectories in phase space. For those systems with bounded solutions, it is found that nonperiodic solutions are ordinarily unstable with respect to small modifications, so that slightly differing initial states can evolve into considerably different states. Systems with bounded solutions are shown to possess bounded numerical solutions.

A simple system representing cellular convection is solved numerically. All of the solutions are found to be unstable, and almost all of them are nonperiodic.

The feasibility of very-long-range weather prediction is examined in the light of these results.

1. Introduction

Certain hydrodynamical systems exhibit steady-state flow patterns, while others oscillate in a regular periodic fashion. Still others vary in an irregular, seemingly haphazard manner, and, even when observed for long periods of time, do not appear to repeat their previous history.

These modes of behavior may all be observed in the familiar rotating-basin experiments, described by Fultz, *et al.* (1959) and Hide (1958). In these experiments, a cylindrical vessel containing water is rotated about its axis, and is heated near its rim and cooled near its center in a steady symmetrical fashion. Under certain conditions the resulting flow is as symmetric and steady as the heating which gives rise to it. Under different conditions a system of regularly spaced waves develops, and progresses at a uniform speed without changing its shape. Under still different conditions an irregular flow pattern forms, and moves and changes its shape in an irregular nonperiodic manner.

Lack of periodicity is very common in natural systems, and is one of the distinguishing features of turbulent flow. Because instantaneous turbulent flow patterns are so irregular, attention is often confined to the statistics of turbulence, which, in contrast to the details of turbulence, often behave in a regular well-organized manner. The short-range weather forecaster, however, is forced willy-nilly to predict the details of the large-scale turbulent eddies—the cyclones and anticyclones—which continually arrange themselves into new patterns.

Thus there are occasions when more than the statistics of irregular flow are of very real concern.

In this study we shall work with systems of deterministic equations which are idealizations of hydrodynamical systems. We shall be interested principally in nonperiodic solutions, i.e., solutions which never repeat their past history exactly, and where all approximate repetitions are of finite duration. Thus we shall be involved with the ultimate behavior of the solutions, as opposed to the transient behavior associated with arbitrary initial conditions.

¹ The research reported in this work has been sponsored by the Geophysics Research Directorate of the Air Force Cambridge Research Center, under Contract No. AF 19(604)-4969.

Figure 1.4: Excerpt from Lorenz's 1963 paper.

review article

Simple mathematical models with very complicated dynamics

Robert M. May*

First-order difference equations arise in many contexts in the biological, economic and social sciences. Such equations, even though simple and deterministic, can exhibit a surprising array of dynamical behaviour, from stable points, to a bifurcating hierarchy of stable cycles, to apparently random fluctuations. There are consequently many fascinating problems, some concerned with delicate mathematical aspects of the fine structure of the trajectories, and some concerned with the practical implications and applications. This is an interpretive review of them.

THERE are many situations, in many disciplines, which can be described, at least to a crude first approximation, by a simple first-order difference equation. Studies of the dynamical properties of such models usually consist of finding constant equilibrium solutions, and then conducting a linearised analysis to determine their stability with respect to small disturbances: explicitly nonlinear dynamical features are usually not considered.

Recent studies have, however, shown that the very simplest nonlinear difference equations can possess an extraordinarily rich spectrum of dynamical behaviour, from stable points, through cascades of stable cycles, to a regime in which the behaviour (although fully deterministic) is in many respects "chaotic", or indistinguishable from the sample function of a random process.

This review article has several aims.

First, although the main features of these nonlinear phenomena have been discovered and independently rediscovered by several people, I know of no source where all the main results are collected together. I have therefore tried to give such a synoptic account. This is done in a brief and descriptive way, and includes some new material: the detailed mathematical proofs are to be found in the technical literature, to which signposts are given.

Second, I indicate some of the interesting mathematical questions which do not seem to be fully resolved. Some of these problems are of a practical kind, to do with providing a probabilistic description for trajectories which seem random, even though their underlying structure is deterministic. Other problems are of intrinsic mathematical interest, and treat such things as the pathology of the bifurcation structure, or the truly random behaviour, that can arise when the nonlinear function $F(X)$ of equation (1) is not analytical. One aim here is to stimulate research on these questions, particularly on the empirical questions which relate to processing data.

Third, consideration is given to some fields where these notions may find practical application. Such applications range from the abstractly metaphorical (where, for example, the transition from a stable point to "chaos" serves as a metaphor for the onset of turbulence in a fluid), to models for the dynamic behaviour of biological populations (where one can seek to use field or laboratory data to estimate the values of the parameters in the difference equation).

Fourth, there is a very brief review of the literature pertaining to the way this spectrum of behaviour—stable points, stable cycles, chaos—can arise in second or higher order difference equations (that is, two or more dimensions; two or more interacting species), where the onset of chaos usually requires less severe nonlinearities. Differential equations are also surveyed in this light; it seems that a three-dimensional system of first-order ordinary differential equations is required for the manifestation of chaotic behaviour.

The review ends with an evangelical plea for the introduction of these difference equations into elementary mathematics courses, so that students' intuition may be enriched by seeing the wild things that simple nonlinear equations can do.

*King's College Research Centre, Cambridge CB2 1ST; on leave from Biology Department, Princeton University, Princeton 08540.

CHAPTER 2

DYNAMICAL SYSTEMS, COBWEBBING, AND THE LOGISTIC EQUATION

2.1 General Definitions

Let us begin with basic object of study.

Definition 2.1 (Dynamical system). *A dynamical system is an object whose state at a future time depends deterministically on*

- its present state, and
- a rule that governs its evolution through time.

Definition 2.2 (State space, evolution operator). *The state space (also called ‘phase space’) of the dynamical system is the set X of all possible states x of the system. The evolution operator is a map $F : X \rightarrow X$ that specifies the evolution rule. In the first part of this course we will take $X = \mathbb{R}$. In the second part we will take $X = S^1$ (the circle).*

The map $F : X \rightarrow X$ will usually be continuous (and as differentiable as we need to make sense of the calculations at hand), even if this is not stated. We will often think of F as depending (in a similar r -times differentiable manner) on a real parameter μ . Then, given the state of the system x_n at some time point n , the dynamical system is described by the difference equation

$$x_{n+1} = F(x_n, \mu)$$

The collection of maps $F(x, \mu)$ for $\mu_1 < \mu < \mu_2$ is often referred to as a ‘family’ of maps when emphasising how properties of the map vary with μ .

‘Dynamical systems’ viewpoint

We take the ‘dynamical systems’ viewpoint in the course: we focus our attention on the following.

1. The asymptotic behaviour of solutions at long times, after transients have decayed.
2. Qualitative features of the dynamics on X , that are independent of our choice of co-ordinate system on X .
3. What happens to these qualitative features as parameters vary (bifurcations).

We will also consider

1. How to define ‘chaotic’ behaviour.
2. Whether there are kinds of behaviour that look complicated but are not actually chaotic.
3. Why evolution operators with very different functional forms produce very similar looking bifurcation diagrams.

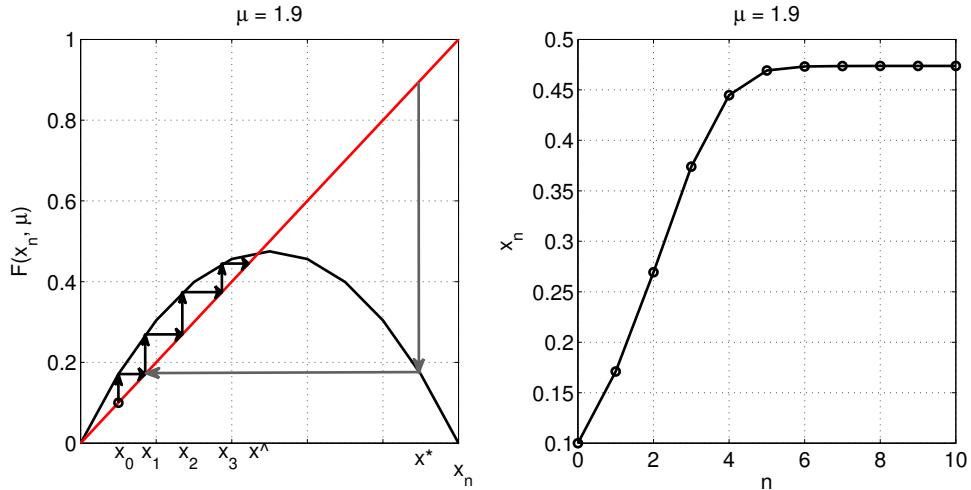
2.2 Revision of the logistic map

In this section summarise the dynamics of the logistic map

$$x_{n+1} = F(x_n, \mu) := \mu x_n (1 - x_n)$$

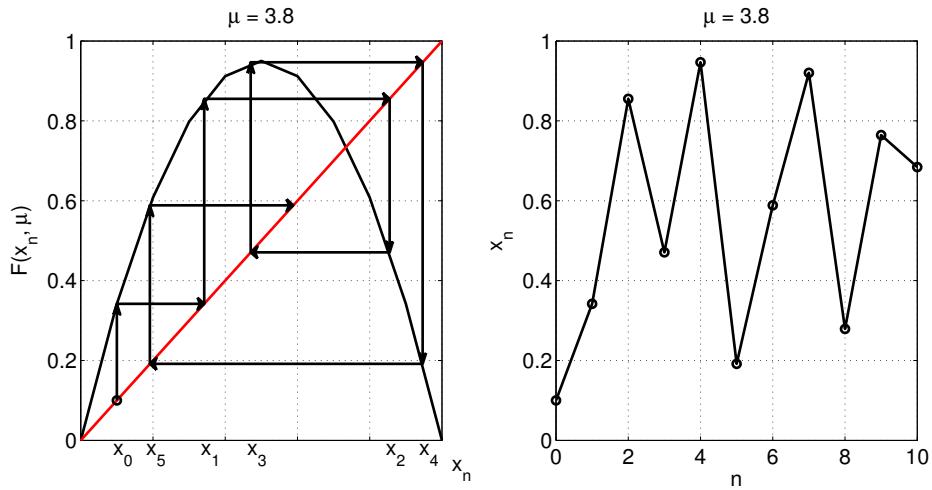
for $X = [0, 1]$ and (usually) $0 \leq \mu \leq 4$. Iterates of F , i.e. repeated composition of F with itself, maps points around X .

The dynamics of F can be described informally by sketching a ‘cobweb diagram’, for example when $\mu \approx 2$ we have something like:



Note that when $0 \leq \mu \leq 1$ we have $F(x) \leq x$ and so all iterates starting in X are attracted to $x = 0$. When $1 \leq \mu \leq 3$ we find that iterates are attracted to a different point $x = \hat{x} = 1 - 1/\mu \in X$ for $\mu \geq 1$. Clearly the map F is not invertible here, for example the point x_1 has two pre-images x_0 and x^* .

For larger values of μ , say $\mu \approx 3.8$ we find something more like the following:



and iterates do not appear to settle down to anything that we can describe by eye.

We also recall the idea of linear stability of a fixed point. Fixed points satisfy $\hat{x} = F(\hat{x})$. To see if nearby points move towards or away from a fixed point we look at small perturbations. Let $x_n = \hat{x} + \delta_n$. Then (suppressing any dependence on parameters)

$$\begin{aligned} x_{n+1} &= \hat{x} + \delta_{n+1} = & F(\hat{x} + \delta_n) \\ &= & F(\hat{x}) + F'(\hat{x})\delta_n + \text{h.o.t.} \end{aligned}$$

by Taylor series. Since $\hat{x} = F(\hat{x})$ we are left with

$$\delta_{n+1} = F'(\hat{x})\delta_n + \text{h.o.t.}$$

so if $|F'(\hat{x})| < 1$ then we expect sufficiently small perturbations to decay to zero, and so we would expect that \hat{x} is ‘stable’ in some sense.

For the logistic map we have $F'(x) = \mu(1 - 2x)$, so $F'(\hat{x}) = 2 - \mu$, so

$$|F'(\hat{x})| < 1 \quad \text{when} \quad 1 < \mu < 3.$$

At $\mu = 3$ we have $F'(\hat{x}) = -1$ which (recall) indicates a period-doubling bifurcation. Numerically we observe the existence of a period=2 orbit (2-cycle) consisting of points $\{x_1, x_2\}$ such that $x_2 = F(x_1)$ and $x_1 = F(x_2)$, $x_1 \neq x_2$.

We can find the 2-cycle explicitly by solving $F(F(x)) \equiv F \circ F(x) \equiv F^2(x) = x$. We obtain a quartic expression that contains the fixed points as roots as well as the 2-cycle we want. Factoring them out we obtain

$$x \left(x - 1 + \frac{1}{\mu} \right) \left(x^2 - x - \frac{x}{\mu} + \frac{1}{\mu} \left(1 + \frac{1}{\mu} \right) \right) = 0$$

so the new period-2 points are

$$x_{1,2} = \frac{1}{2} \left(1 + \frac{1}{\mu} \right) \pm \frac{1}{2} \left[\left(1 + \frac{1}{\mu} \right) \left(1 - \frac{3}{\mu} \right) \right]^{1/2}.$$

Since these points are fixed under F^2 it makes sense to think about their stability as fixed points of a new map $G := F^2$. Note that, by the chain rule

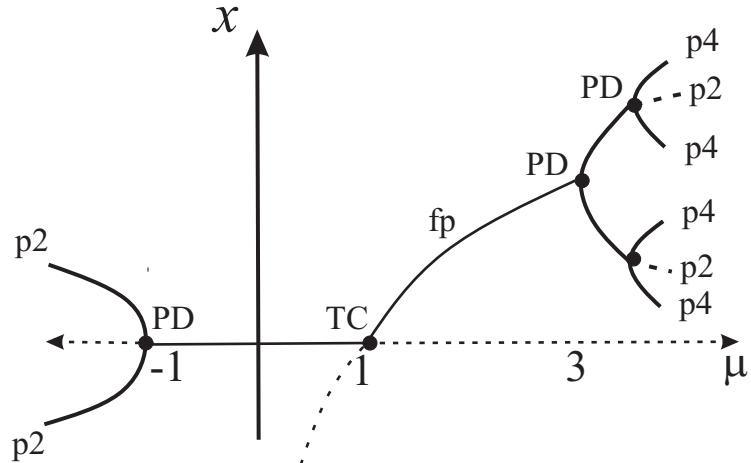
$$G'(x_1) = F'(F(x_1)) \quad F'(x_1) = F'(x_2)F'(x_1) = G'(x_2)$$

so it doesn’t matter which point on the 2-cycle we take. Using the expression above for $F'(x)$ we find

$$\begin{aligned} G'(x_1) &= \mu^2(1 - 2x_1)(1 - 2x_2) \\ &= 4 + 2\mu - \mu^2 \end{aligned}$$

It is easy to show, for example by considering the graph of $G'(x_1)$, that $|G'(x_1)| < 1$ for $3 < \mu < 1 + \sqrt{6}$, and passes through -1 when $\mu = 1 + \sqrt{6} \approx 3.45$ which means there is another period doubling bifurcation.

Overall we have deduced (parts of) the following bifurcation diagram for the logistic map:



CHAPTER 3

ORBITS, FIXED/PERIODIC POINTS, AND INVARIANT SETS

Throughout we will be considering the discrete map defined by

$$x_{n+1} = F(x_n). \quad (3.1)$$

Often the domain and range of F will be from $\mathbb{R} \rightarrow \mathbb{R}$, but we may later restrict this (the context is generally clear). In this course, we are generally working with functions F that are piecewise differentiable.

Definition 3.1 (Forward orbit, orbit). *The forward orbit of a point x_0 is the set $\mathcal{O}^+(x_0) = \{x_0, x_1, x_2, \dots\}$. If F is invertible we can define the bi-infinite sequence $\{\dots, x_{-n}, x_{-n+1}, \dots, x_{-1}, x_0, x_1, x_2, \dots, x_n, \dots\}$ to be the orbit $\mathcal{O}(x_0)$.*

We denote repeated composition of F by powers, i.e. $F(F(x)) =: F^2(x)$ and similarly for longer compositions: $F^N(x)$ denotes the N -fold composition.

Definition 3.2 (Fixed point, periodic point, N -cycle). *The point x_0 is a fixed point if $F(x_0) = x_0$. x_0 is a periodic point with least period N if $F^N(x_0) = x_0$ but $x_0 \notin \{x_1, x_2, \dots, x_{N-1}\}$. In this case we refer to the set $\{x_0, \dots, x_{N-1}\}$ as an N -cycle.*

Definition 3.3 (Invariant). *A set $S \subset \mathbb{R}$ is invariant under F if, whenever $x \in S$ then $\mathcal{O}(x) \subseteq S$.*

Fixed points, N -cycles and forward orbits are all examples of invariant sets.

3.1 Stability

2021-22 note: We have changed this section to be simpler and without introducing Lyapunov stability and quasi-asymptotic stability. The definitions are less general, but they work better for this course.

Based on our investigations of the logistic equation, we believe that stability of fixed points can be understood by examining the gradient of the function at that point. This motivates the following definitions.

Definition 3.4 (Attracting/repelling/neutral). *Suppose x_0 is a fixed point for F . Then x_0 is an attracting fixed point if $|F'(x_0)| < 1$. The point x_0 is a repelling fixed point if $|F'(x_0)| > 1$. Finally if $|F'(x_0)| = 1$, then x_0 is a neutral fixed point.*

We are now led to two theorems.

Theorem 3.1 (Attracting fixed point). *Suppose x_0 is an attracting fixed point for F . Then there is an interval I that contains x_0 in its interior and in which the following condition is satisfied: (i) if $x \in I$, then $F^n(x) \in I$ for all n and moreover, (ii) $F^n(x) \rightarrow x_0$ as $n \rightarrow \infty$.*

We can intuitively describe (i) as “starts near stays near” (this is what is known as Luapunov stability) and (ii) as “eventually tends to” (this is quasi-asymptotic stability).

Proof: Since $|F'(x_0)| < 1$, then there is a $\lambda > 0$ such that $|F'(x_0)| < \lambda < 1$. So choose $\delta > 0$ so that $|F'(x)| < \lambda$ if $x \in I = [x_0 - \delta, x_0 + \delta]$. Now choose a point $p \in I$. By the Mean Value Theorem, all secant lines have gradients bounded by λ ,

$$\frac{|F(p) - F(x_0)|}{|p - x_0|} < \lambda,$$

so we have that

$$|F(p) - F(x_0)| < \lambda|p - x_0|.$$

Note that x_0 is a fixed point so

$$|F(p) - x_0| < \lambda|p - x_0|.$$

Since $\lambda < 1$, this means that $F(p)$ is as close to x_0 as p . Thus $F(p)$ lies in I . We can then continue with the same argument applied to $F(p)$ and $F(x_0)$. We would find that

$$\begin{aligned} |F^2(p) - x_0| &= |F^2(p) - F^2(x_0)| \\ &< \lambda|F(p) - F(x_0)| \\ &< \lambda^2|p - x_0|. \end{aligned}$$

Since we have $\lambda < 1$, then $\lambda^2 < \lambda$. Then $F^2(p)$ and x_0 are even closer together than $F(p)$ and x_0 . We continue this procedure to find that, for any $n > 0$,

$$|F^n(p) - x_0| < \lambda^n|x - x_0|.$$

Since $\lambda^n \rightarrow 0$ as $n \rightarrow \infty$, then $F^n(p) \rightarrow x_0$ as $n \rightarrow \infty$. QED.

The following theorem is proved in an analogous manner.

Theorem 3.2 (Repelling fixed point). *Suppose x_0 is a repelling fixed point for F . Then there is an interval I that contains x_0 in its interior and in which the following condition is satisfied: if $x \in I$ and $x \neq x_0$, then there is an integer $n > 0$ such that $F^n(x) \notin I$.*

CHAPTER 4

INTRODUCTION TO ASYMPTOTIC ANALYSIS

In the next few chapters, our goal is to study the solutions of the equation,

$$F(x, \mu) = 0, \quad (4.1)$$

when we know that both x and μ are localised near some point, say $x = x_0$, $\mu = \mu_0$. Only in very simple cases is the above equation solvable, so our emphasis will turn to developing *approximations*. *Asymptotic analysis* or *perturbation theory* refers to the set of mathematical techniques that allow us to systematically approximate functions in some limit. Typically, we discuss this in the context of small parameters, say ε or δ , as these parameters tend to zero (or equivalently, are considered to be small numbers). In this notation, we would then seek approximations of (4.1) as $\varepsilon = x - x_0$ and $\delta = \mu - \mu_0$ tends to zero.

4.1 Taylor series

The Taylor series is an example of an asymptotic approximation, developed for a function $f(x)$ where $\varepsilon = x - x_0$ is considered to be small.

Theorem 4.1 (Taylor's theorem in 1D). *Let f and its derivatives, f' , f'' , f''' , ... exist and be continuous on some interval of the point $x = x_0$. Then we may approximate f near x_0 using the Taylor series*

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n \quad \text{where } a_n = \frac{f^{(n)}(x_0)}{n!}.$$

In practice, we may choose to use only a few terms (or code a computer program to compute many terms). Here, the small parameter is $\varepsilon = x - x_0$, and measures the distance from the origin of the approximation. We might believe that better approximations are obtained if we take ε smaller and smaller, or if we include more and more terms. This is indeed the case.

Note that, if you only need a few terms from the series, you often do not need to go through the formal procedure of differentiation. Many examples use the trick of applying the formula for the geometric series. We will use the formula for the geometric series often in this course, so it is worth a reminder:

Theorem 4.2 (Geometric series).

$$\frac{1}{1-x} = 1 + x + x^2 + \dots$$

The other series that is most often used in series manipulations is the binomial expansion,

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha+1)}{2} x^2 + \dots$$

Example 4.1. Find the first four terms of a Taylor series expansion of the function $1/(1+x^2)$ when $x = 0$.

We have from the formula of the geometric series,

$$\frac{1}{1+x^2} = \frac{1}{1-(-x^2)} = 1 + (-x^2) + (-x^2)^2 + (-x^2)^3 + (-x^2)^4 + \dots$$

It is possible to extend Taylor's theorem to higher dimensions.

Theorem 4.3 (Taylor's theorem in 2D). Let f be a function of two variables x and y . If all partial derivatives of order n are continuous in a closed region, then we may approximate f around the point (x_0, y_0) using the series

$$f(x, y) = f(x_0, y_0) + \left[\varepsilon \frac{\partial}{\partial x} + \delta \frac{\partial}{\partial y} \right] f(x_0, y_0) + \frac{1}{2!} \left[\varepsilon \frac{\partial}{\partial x} + \delta \frac{\partial}{\partial y} \right]^2 f(x_0, y_0) + \dots$$

where $\varepsilon = x - x_0$ and $\delta = y - y_0$.

Again, these are examples of asymptotic approximations, now in two variables, with $\varepsilon, \delta \rightarrow 0$. You would be able to develop the approximation of f to a certain number of terms, and as long as ε and δ are sufficiently small, the sum of such terms should be close to the true value of f . In this course, we do not discuss the issue of *series convergence*. In fact, most of the interesting series you encounter in the real world are divergent!

4.2 Order notation and the tilde sign for asymptotic

Next, we want to define precisely what we mean when we say that two functions, say f and g , exhibit the same behaviour in some limit, say $\varepsilon \rightarrow 0$ or $x \rightarrow x_0$ or $x \rightarrow \infty$ and so forth. For instance, we claim that the graphs of $\sin(x)$ and x look very similar as $x \rightarrow 0$. Thus we might write

$$\sin(x) \sim x \quad \text{as } x \rightarrow 0. \tag{4.2}$$

This notion of *similarity* allows us to specify functional behaviours at a deeper level than just limits. As you can see, it is not as useful to specify that

$$\lim_{x \rightarrow 0} \sin x = \lim_{x \rightarrow 0} x.$$

In contrast, (4.2) is much more prescriptive about the way that the functions are approaching the limit.

Definition 4.1 (\sim , \gg , \ll). *First, the notation*

$$f(x) \ll g(x), \quad x \rightarrow x_0,$$

is read as “ $f(x)$ is much smaller than $g(x)$ as $x \rightarrow x_0$ ” and means

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = 0.$$

We may analogously use $g(x) \gg f(x)$ for “much greater than...”.

Second, the notation

$$f(x) \sim g(x), \quad x \rightarrow x_0,$$

is read as “ $f(x)$ is asymptotic to $g(x)$ as $x \rightarrow x_0$ ”, and means that the error between f and g tends to zero as $x \rightarrow x_0$, or

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = 1.$$

We will often say “ f is like g ” or “ f behaves like g ”.

Example 4.2. *Here are some examples:*

- $\sin x \sim x \sim \tan x$ as $x \rightarrow 0$
- $x^2 + x + 1 \sim \frac{x^3 + \sin x}{1 + x}$ as $x \rightarrow \infty$
- $\sin x \ll \cos x$ as $x \rightarrow 0$

In the examination of limiting processes, often the main issue of consideration is the relative sizes of quantities defined according to their powers. For example, if x is a very small number, with $x = 10^{-5}$, then x^5 is much smaller than x (in terms of our notation, $x^5 \ll x$ as $x \rightarrow 0$). On the other hand, we might not care so much about the difference between

$$x^5 \quad \text{vs.} \quad 5x^5$$

The point is that the *order* of x^5 and $5x^5$ is the same as $x \rightarrow 0$. The “Big-Oh” notation formalises this distinction.

Definition 4.2 (Big-Oh). *We write $f = O(g)$ as $x \rightarrow x_0$ to mean that there exists constants $K > 0$ and $x^* > 0$ such that*

$$|f| < K|g| \quad \text{for all } |x - x_0| < x^*.$$

In practice, the use of the order symbol is very natural and you will not need to work with the technical definition. For example, when you derive the terms of the Maclaurin/Taylor series, you are

naturally clustering all the terms of the same order (power) together. For us, the O symbol provides a very convenient way of separating terms of different sizes.

4.3 The method of dominant balance

We want to now explain the procedure from which we can estimate solutions of equations using an asymptotic approximation.

Let us explain the procedure with a simple example. It is best to select an example from which we already know the solution. We seek solutions of the equation

$$\underbrace{x^2}_{\textcircled{1}} + \underbrace{\varepsilon x}_{\textcircled{2}} - \underbrace{1}_{\textcircled{3}} = 0, \quad (4.3)$$

when the parameter, ε is a small number. When ε is a small number, we argue that at least two of the three terms must be balanced. The procedure in which we establish which terms are dominant is called the method of dominant balance.

- Case 1: $\textcircled{1} \sim \textcircled{3} \gg \textcircled{2}$

This yields the balance of $x^2 \sim 1$ and thus $x \sim \pm 1$.

If it is indeed the case that $x = O(1)$, then we can check that $\textcircled{2} = O(\varepsilon)$ is indeed smaller than $\textcircled{1}, \textcircled{3}$. The balance is therefore consistent.

- Case 2: $\textcircled{1} \sim \textcircled{2} \gg \textcircled{3}$

Balancing the first two terms yields $x = O(\sqrt{\varepsilon})$ which is small.

However, the last term is $O(1)$, which turns out to be larger than $\textcircled{1}$ or $\textcircled{2}$. So this is a contradiction that the first two terms establish the dominant balance.

- Verify that all remaining cases (of $\textcircled{2} \sim \textcircled{3}$ or all three terms in balance) yield contradictions.

The above seems all pretty obvious, but this systematic method will be important for less obvious problems. Once the dominant balance has been established, we can systematically develop approximations. We first set

$$x = x_0 + \varepsilon x_1 + \varepsilon^2 x_2 + \dots$$

Substitute into (4.3) and collect terms in increasing powers of ε . This yields

$$\left[x_0^2 + (2x_0 x_1)\varepsilon + O(\varepsilon^2) \right] + \varepsilon \left[x_0 + \varepsilon x_1 + \dots \right] - 1 = 0,$$

for which we solve to get

$$x_0 = \pm 1 \quad x_1 = -\frac{1}{2}.$$

Thus we have to the first two orders of approximation:

$$x = \pm 1 - \frac{\varepsilon}{2} + \dots$$

The above can be verified and contrasted with expansion of the exact solution:

$$x = \frac{-\varepsilon \pm \sqrt{\varepsilon^2 + 4}}{2} = \frac{-\varepsilon \pm (2 + \varepsilon^2/4 + \dots)}{2} = \pm 1 - \frac{\varepsilon}{2} + \dots$$

The most important lesson to grasp (beyond following the manipulations above) is that this method of dominant balance, followed by asymptotic approximations, gives you a very powerful and very general way of estimating solutions to problems that you would not ordinarily be able to solve. You will get more practice in the problem set, but the above is sufficient to follow the next chapter.

 CHAPTER 5

LOCAL ANALYSIS OF BIFURCATIONS I

Let us now return to the main task. Suppose we are given a dynamical system,

$$x_{n+1} = F(x_n, \mu).$$

We have established that the fixed points, where $F(x^*) = 0$ are important. We have then seen an examples where the behaviour of the fixed point changes suddenly as some critical value of $\mu = \mu^*$ is passed. This is what is known as a bifurcation. Our objective is to then study near this bifurcation, where $x - x^*$ and $\mu - \mu^*$ are small.

Definition 5.1 (Bifurcation). *A bifurcation is a change in the structure of the periodic orbits of a dynamical system as a system parameter varies continuously through a critical value.*

Notice that according to the above definition, we are interested in the behaviours of fixed points, 2-cycles, 3-cycles, etc. (since these are periodic orbits). Typically the bifurcation occurs at a point, say $x = x^*$ and $\mu = \mu^*$, and we call this the **bifurcation point**. Therefore, by assumption of a change in stability, we have

$$|F_x(x^*, \mu^*)| = 1. \quad (5.1)$$

Without loss of generality we can shift coordinates, say with

$$x = x^* + \hat{x} \quad \text{and} \quad \mu = \mu^* + \hat{\mu},$$

so that the bifurcation occurs at $(\hat{x}, \hat{\mu}) = (0, 0)$. We henceforth assume that such a transformation has been made and re-label variables with hats to normal variables, so that now $F(x, \mu)$ is assumed to have a bifurcation point at $(0, 0)$.

We look at the behaviour of iterates near $x = 0$ and expand $F(x, \mu)$ in a Taylor series in both variables:

$$F(x, \mu) = F(0, 0) + a_0 x + a_1 \mu + b_0 x^2 + b_1 x \mu + b_2 \mu^2 + c_0 x^3 + c_1 x^2 \mu + c_2 x \mu^2 + c_3 \mu^3 + O(x^4, \mu^4), \quad (5.2)$$

where

$$\begin{aligned} a_0 &= F_x(0, 0), & a_1 &= F_\mu(0, 0), \\ b_0 &= \frac{1}{2} F_{xx}(0, 0), & b_1 &= F_{x\mu}(0, 0), & b_2 &= \frac{1}{2} F_{\mu\mu}(0, 0), \end{aligned}$$

and we have written $O(x^4, \mu^4)$ to mean terms of quartic order in x and μ , including combinations such as $O(x^3\mu)$.

Since $x = 0$ is a fixed point at $\mu = 0$, $F(0, 0) = 0$. Since $x = 0$ changes stability, either $a_0 = +1$ (Case A) or $a_0 = -1$ (Case B).

Case A: $F_x(0, 0) = +1$

At a fixed point $F(x, \mu) = x$ and using the Taylor series for $F(x, \mu)$:

$$0 = a_1\mu + b_0x^2 + b_1x\mu + b_2\mu^2 + c_0x^3 + c_1x^2\mu + c_2x\mu^2 + c_3\mu^3 + O(x^4, \mu^4). \quad (5.3)$$

We estimate fixed points near to $x = \mu = 0$ by looking for solutions of the leading-order approximation to (5.3) (and then, formally, applying the Implicit Function Theorem to assert that the fixed point(s) $x(\mu)$ that we locate do indeed continue to exist over some small open interval in μ containing zero).

Note we will do the dominant balances carefully in lectures.

Case A1: $a_1 \neq 0, b_0 \neq 0$; saddle-node bifurcation

For x and μ small, we assume that the dominant balance is given by the first two terms of the right hand-side of (5.3), giving

$$a_1\mu \sim -b_0x^2$$

and hence fixed points occur when

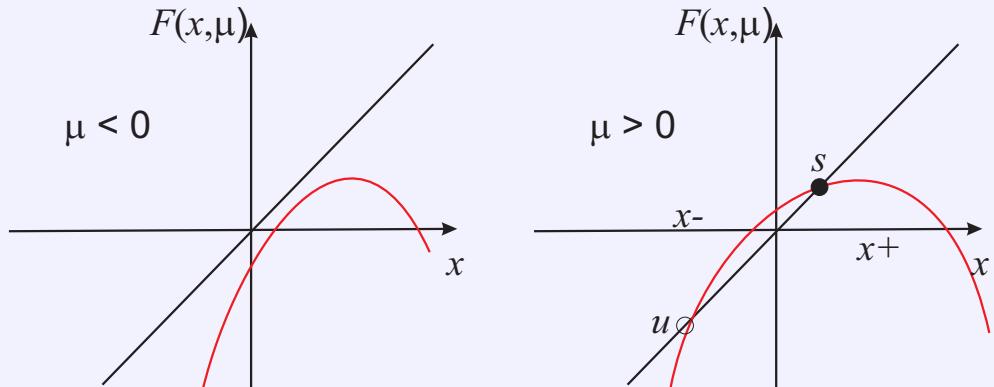
$$x \sim \pm \sqrt{-\frac{a_1}{b_0}\mu}.$$

So there are two fixed points that exist on the same side of the bifurcation point $\mu = 0$, depending on the sign of a_1/b_0 . This bifurcation is called a **saddle-node bifurcation**.

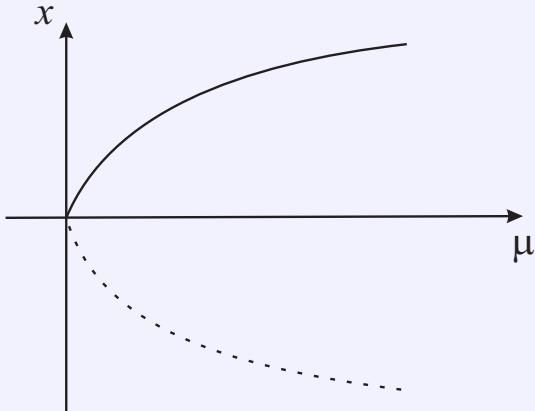
Example 5.1. The ‘canonical (archetypal) form’ for a saddle-node bifurcation is

$$x_{n+1} = \mu + x_n - x_n^2 =: F(x_n, \mu).$$

Fixed points are located at $x_{\pm} = \pm\sqrt{\mu}$ so they exist only in $\mu \geq 0$, and $F'(x_{\pm}, \mu) = 1 \mp 2\sqrt{\mu}$ so the fixed point x_+ is stable and x_- is unstable. The graphs of F for negative and positive μ look like:



We can summarise the locations of the fixed points in a bifurcation diagram: this is a plot of the location of invariant sets $x(\mu)$ as a function of μ . The general convention for bifurcation diagrams is that solid lines denote stable solutions dashed lines unstable solutions.:



Case A2: $a_1 = 0, b_0 \neq 0$; transcritical bifurcation

Now with the first term of (5.3) zero, we assume that the dominant balance is given by all the quadratic terms together,

$$b_0 x^2 \sim - (b_1 x \mu + b_2 \mu^2),$$

and hence solving by the quadratic formula,

$$x \sim \frac{-b_1 \mu \pm \mu [b_1^2 - 4b_0 b_2]^{1/2}}{2b_0} = -\mu \left(\frac{b_1 \pm \Delta}{2b_0} \right)$$

where $\Delta^2 = b_1^2 - 4b_0 b_2$. If $b_1^2 - 4b_0 b_2 < 0$ then there are no branches of fixed points $x(\mu)$ near $\mu = 0$. But if $\Delta^2 > 0$ then there are two branches of fixed points that exist in both $\mu < 0$ and $\mu > 0$. This is therefore different to the saddle-node bifurcation case. It is called a **transcritical bifurcation**.

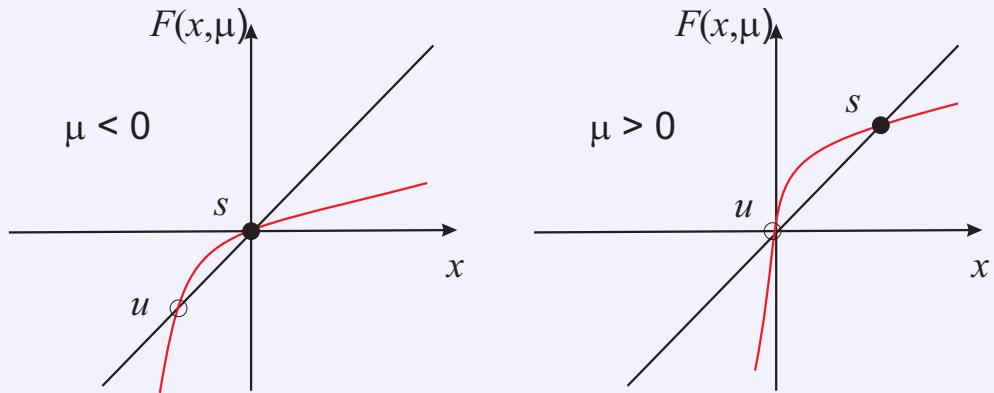
Example 5.2. The ‘canonical form’ for a transcritical bifurcation is

$$x_{n+1} = (1 + \mu)x_n - x_n^2 =: F(x_n, \mu).$$

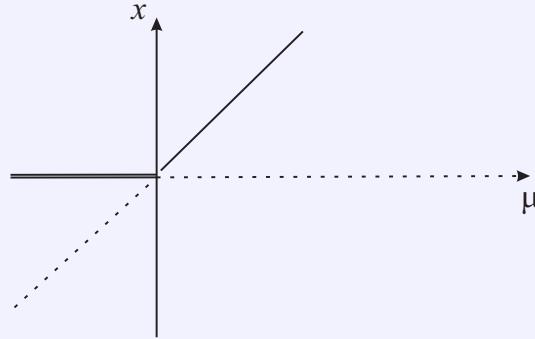
Fixed points exist at $x = 0$ and $x = \mu$. To check stability, we compute $F_x(x, \mu) = 1 + \mu - 2x$, so

- $F_x(0, \mu) = 1 + \mu$: stable in $\mu < 0$ and unstable in $\mu > 0$.
- $F_x(\mu, \mu) = 1 - \mu$: unstable in $\mu < 0$ and stable in $\mu > 0$.

The graphs of F for negative and positive μ look like this:



The corresponding bifurcation diagram is



Note that there are two fixed points on both sides of the bifurcation point. This provides a useful quick way to distinguish between the saddle-node and transcritical cases.

No case: $a_1 \neq 0, b_0 = 0$

Then (5.3) gives

$$0 = a_1\mu + b_1x\mu + b_2\mu^2 + O(x^3, \mu^3),$$

For small μ , this equation has a solution $x \sim -a_1/b_1$ which is (typically, without extra information) not close to $x = 0$. So it does not provide a branch of fixed points that connects with/emerges from $x = \mu = 0$ which was our supposed fixed point to start with. So this combination of constraints is not labelled, or taken any further.

Case A3: $a_1 = 0, b_0 = 0$; pitchfork bifurcation

Then (5.3) gives

$$0 = b_1x\mu + b_2\mu^2 + c_0x^3 + c_1x^2\mu + c_2x\mu^2 + c_3\mu^3 + O(x^4, \mu^4).$$

This case yields an interesting twist, similar to the twist encountered in studying the cubic equation in the problem set: there are, in fact, two possible distinct leading-order balances. First, we use the first two terms as the dominant balance, giving

$$b_1x\mu \sim -b_2\mu^2,$$

and so

$$x \sim -\frac{b_2}{b_1}\mu.$$

Again, the reader should check, by verifying the other terms that the above approximation produces a consistent dominant balance.

However, there is in fact another legitimate balance:

$$c_0x^3 \sim -b_1x\mu,$$

and so

$$x \sim \pm \left(-\frac{b_1\mu}{c_0} \right)^{1/2}.$$

Again, using the Implicit Function Theorem (implicitly!) we can assert that fixed points near to these values of x exist as long as $b_1 \neq 0$ and $c_0 \neq 0$. The first fixed point exists in both μ negative

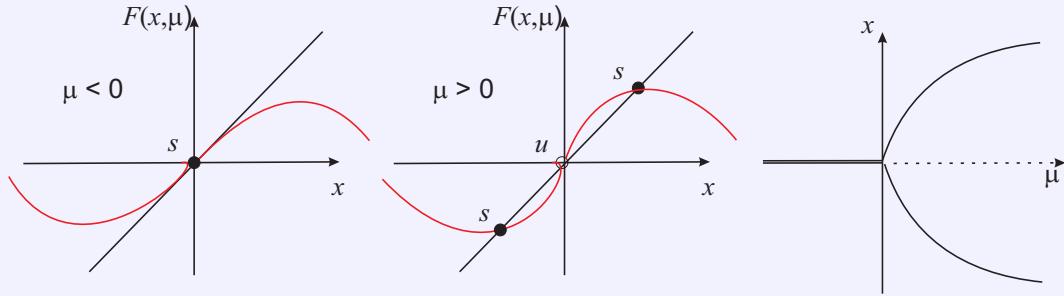
and positive. The other two fixed points exist only for one sign of μ . This is called a **pitchfork bifurcation**.

Example 5.3. The ‘canonical forms’ for pitchfork bifurcations

Supercritical

$$x_{n+1} = (1 + \mu)x_n - x_n^3 =: F(x_n, \mu).$$

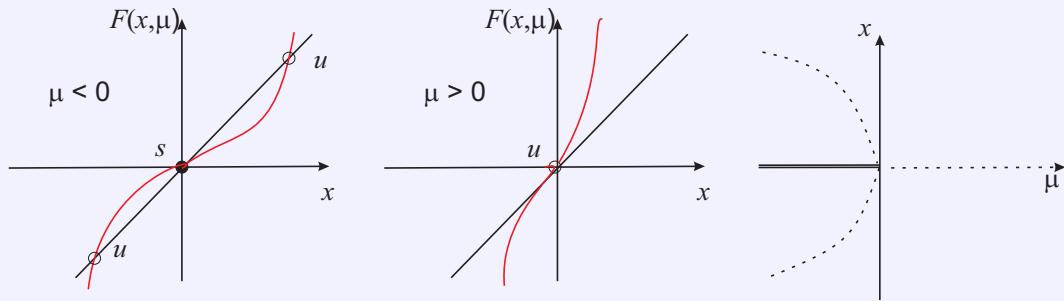
The fixed points are at $x = 0$ and $x_{\pm} = \pm\sqrt{\mu}$. $F_x(x) = 1 + \mu - 3x^2$ so $x = 0$ is stable in $\mu < 0$ (and near zero), and unstable in $\mu > 0$. The non-zero fixed points x_{\pm} exist only in $\mu > 0$, above the critical point, where they are stable. The graphs of F for negative and positive μ and the corresponding bifurcation diagram look like



Subcritical

$$x_{n+1} = (1 + \mu)x_n + x_n^3 =: F(x_n, \mu).$$

This change of the sign of the cubic term in (5.3) leads to **unstable** non-zero fixed points. The fixed points are $x = 0$ and $x_{\pm} = \pm\sqrt{-\mu}$ so that x_{\pm} exist in $\mu < 0$, below the critical point. The derivative of F is $F_x(x) = 1 + \mu + 3x^2$. We see that the stability of $x = 0$ is as before, but the non-zero fixed points x_{\pm} have $F_x(x_{\pm}) = 1 - 2\mu$ which is greater than 1 when x_{\pm} exist. This behaviour is fundamentally different to the supercritical case, i.e. there is no combination of change of variable and parameter re-labelling that would turn one into the other. The graphs of F for negative and positive μ and the corresponding bifurcation diagram look like



Pitchfork bifurcations have one fixed point on one side of the bifurcation point, and three on the other. This provides a quick way to distinguish between pitchfork and saddle-node or transcritical cases.

Codimension

One could continue demanding more and more conditions on the coefficients and investigating the different cases that arise ad infinitum. But having more conditions to satisfy implies more constraints on the problem which in turn implies that the bifurcation is ‘less typical’. This idea of ‘typicality’, or ‘genericity’ is captured by the idea of the ‘codimension’ of a bifurcation. For our purposes the codimension of the bifurcation is simply the number of constraints that we have to introduce in order to see it. We can summarise the cases A1-A3 above in a table as follows.

Name	Fixed points	Constraints	Codimension	Genericity conditions
Saddle-node	$0 \longleftrightarrow 2$	$a_0 = +1$	1	$a_1, b_0 \neq 0$
Transcritical	$2 \longleftrightarrow 2$	$a_0 = +1, a_1 = 0$	2	$b_0 \neq 0$
Pitchfork	$1 \longleftrightarrow 3$	$a_0 = +1, a_1 = b_0 = 0$	3	$b_1, c_0 \neq 0$

Since the coefficients a_j, b_j, c_j are multiples of the partial derivatives of $F(x, \mu)$ at the bifurcation point, we can more properly express the constraints in terms of partial derivatives of F :

Name	Fixed points	Constraints	Codimension	Genericity conditions
Saddle-node	$0 \longleftrightarrow 2$	$F_x = +1$	1	$F_\mu, F_{xx} \neq 0$
Transcritical	$2 \longleftrightarrow 2$	$F_x = +1, F_\mu = 0$	2	$F_{xx} \neq 0$
Pitchfork	$1 \longleftrightarrow 3$	$F_x = +1, F_\mu = F_{xx} = 0$	3	$F_{x\mu}, F_{xxx} \neq 0$

where F_x means $\partial F / \partial x$ evaluated at $x = \mu = 0$ and F_μ means $\partial F / \partial \mu$ evaluated at $x = \mu = 0$, etc.

LOCAL ANALYSIS OF BIFURCATIONS II

We have completed the analysis near the bifurcation points of $F(x, \mu)$ on the assumption that $F_x(0, 0) = 1$ (these were called ‘Case A’). In doing so, we derived cases of transcritical, saddle-node, and pitchfork bifurcations. Let us now turn to the case of $F_x(0, 0) = -1$.

6.1 Case B: $F_x(0, 0) = -1$; period-doubling bifurcation

As before $F(0, 0) = 0$ since $x = 0$ is a fixed point when $\mu = 0$. As before we estimate fixed points near to $x = 0, \mu = 0$ by looking for solutions of the leading order approximation to $F(x, \mu) = x$. From the Taylor Series (5.3), we obtain

$$x = -x + a_1\mu + b_0x^2 + b_1x\mu + b_2\mu^2 + c_0x^3 + O(x^3, \mu^3).$$

We may verify that the most straightforward dominant balance is consistent, found by balancing the x -term with the μ -term, or

$$x \sim \frac{a_1}{2}\mu.$$

Is the above asymptotic approximation stable or unstable? Using the Taylor series for $F(x, \mu)$,

$$\frac{\partial F}{\partial x} = a_0 + 2b_0x + b_1\mu + \dots$$

We substitute $x = a_1\mu/2$, yielding

$$\frac{\partial F}{\partial x} = -1 + [a_1b_0 + b_1]\mu + \dots$$

Hence, if $\Delta = a_1b_0 + b_1 > 0$, $x \sim a_1\mu/2$ is stable if $\mu > 0$, but unstable if $\mu < 0$. These stability conditions are reversed if $a_1b_0 + b_1 < 0$. In either case, if $a_1b_0 + b_1 \neq 0$ stability switches at $\mu = 0$. No other scalings lead to natural balances. There is exactly one fixed point and it exists on both sides of the bifurcation point $\mu = 0$. This is a somewhat strange thing to happen. Consider the case $\Delta > 0$. Iterates were, for $\mu > 0$, attracted to the stable fixed point at the origin. As soon as $\mu < 0$, the origin is no longer stable, but one might expect that the iterates would then be attracted to a nearby set.

Analysis of the 2-cycle

After a fixed point, the next simplest possibility for an invariant set is a 2-cycle. A 2-cycle occurs when $F^2(x) = F(F(x)) = x$. Near to $x = \mu = 0$ the Taylor series (5.3) gives an approximation for this

equation (note notation)

$$\begin{aligned} x &= -F(x) + a_1\mu + b_0(F(x))^2 + b_1F(x)\mu + b_2\mu^2 + c_0(F(x))^3 + \dots, \\ &= x - b_1x\mu - b_2\mu^2 + b_0(a_1^2\mu^2 - 2a_1x\mu) - b_1x\mu + a_1b_1\mu^2 + b_0b_1\mu x^2 \\ &\quad + b_1x\mu^2 + b_1b_2\mu^3 + b_2\mu^2 - 2c_0x^3 + O(x^3, \mu^3), \end{aligned}$$

where the $+O(x^3, \mu^3)$ term indicates other cubic terms that do not include any x^3 terms: the only x^3 term is the penultimate one in the expansion. Substituting the Taylor series for $F(x, \mu)$ and omitting terms of order $x\mu^2, \mu x^2, x^4$ or smaller, gives the following equation for points satisfying $F^2(x) = x$:

$$0 = -(2b_1 + 2a_1b_0)x\mu + (a_1^2b_0 + a_1b_1)\mu^2 - 2(b_0 + c_0)x^3 + \dots$$

This equation does not contain terms in μ or x^2 since they cancel out. So it is similar in form to the equation we derived in Case A3 for the pitchfork bifurcation. There are two distinct leading-order balances. When $x = O(\mu)$, at leading order

$$-(2a_1b_0 + 2b_1)x\mu + (a_1^2b_0 + a_1b_1)\mu^2 = \dots$$

and so

$$x \sim \frac{a_1(a_1b_0 + b_1)}{2(a_1b_0 + b_1)}\mu = \frac{a_1}{2}\mu.$$

When $x = O(\sqrt{\mu})$, at leading order

$$-(2a_1b_0 + 2b_1)x\mu - 2(b_0 + c_0)x^3 = \dots$$

and so

$$x \sim \pm \sqrt{\frac{-(b_1 + a_1b_0)\mu}{b_0 + c_0}}.$$

The first balance gives the continuation of the fixed point $x = \mu = 0$ which we also found from $F(x, \mu) = x$. The second balance gives points that are fixed points of F^2 but not of F . These points lie on a 2-cycle. The 2-cycle emerges from the fixed point at $x = \mu = 0$ on exactly one side of the bifurcation point, generically, as long as $b_0 + c_0 \neq 0$. This bifurcation is therefore known as a **period-doubling bifurcation**. Like the pitchfork bifurcation, it occurs in two varieties: supercritical (where the new 2-cycle is stable) and subcritical (where the new 2-cycle is unstable).

Remark 6.1. *Codimension of the period-doubling bifurcation.* We didn't need to specify any additional constraints on the coefficients a_j, b_j etc above, and we supposed implicitly that important coefficients did not vanish (for example c_0). So it follows that the period-doubling bifurcation is just as typical in maps as the saddle-node bifurcation i.e. the period-doubling bifurcation has codimension 1.

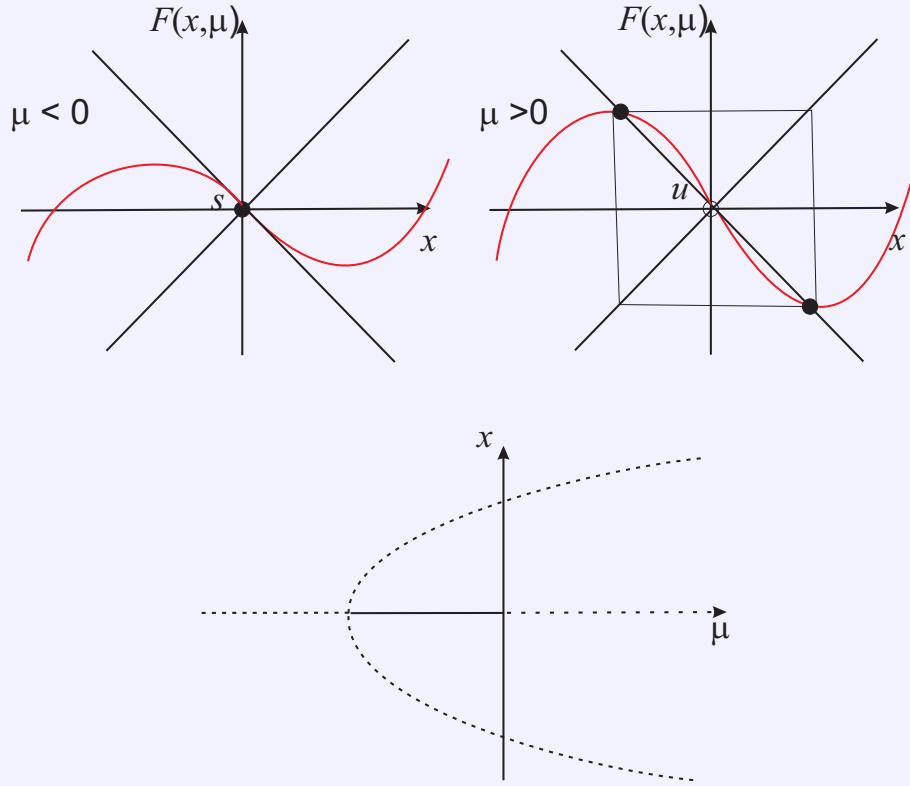
Example 6.1. The 'canonical form' for the period-doubling bifurcation

Supercritical

$$x_{n+1} = -(1 + \mu)x_n + x_n^3 =: F(x_n, \mu).$$

Looking at all real x , not just x near zero, there are fixed points $x = 0$ and $x_{\pm} = \pm\sqrt{2 + \mu}$. We have that $F_x(x) = -(1 + \mu) + 3x^2$. So the fixed point $x = 0$ is stable for $-2 < \mu < 0$. At x_{\pm} we have $F_x(x_{\pm}) = 5 + 2\mu$ which is greater than 1 when x_{\pm} exist (i.e. for $\mu > -2$). So when $\mu > 0$ all the fixed points of the map are unstable.

The graphs of $F(x)$ and preliminary bifurcation diagram look like:

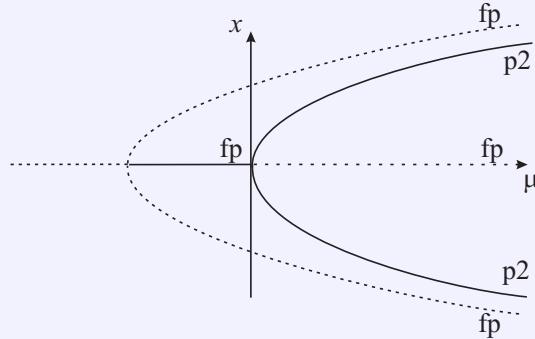


This bifurcation diagram appears incomplete. When $\mu > 0$ iterates are repelled both from zero and from x_{\pm} . There may be some stable object ‘between’ the unstable branches.

For this example $a_1 = b_0 = 0$, $b_1 = -1$ and $c_0 = 1$ in the general estimate for x such that $F^2(x) = x$ given in (6.1). So the new 2-cycle contains points $x \approx \pm\sqrt{\mu}$. We can also see this directly by solving $x = F^2(x)$ which gives

$$x = (1 + 2\mu + \mu^2)x - 2x^3 + O(x^4, \mu^4),$$

from which we also conclude that period 2 points exist near $x = \pm\sqrt{\mu}$. So there is a stable (check) 2-cycle in $\mu > 0$ to complete the bifurcation diagram.

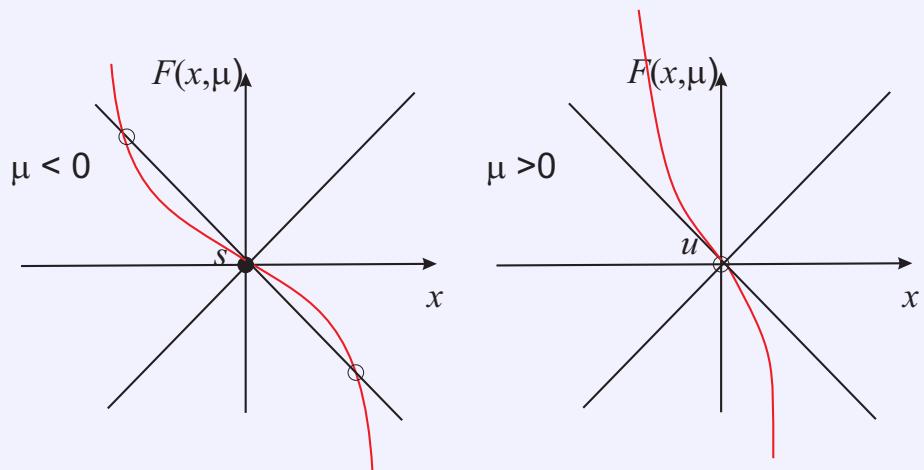


Since the new 2-cycle is stable this is an example of a **supercritical** period-doubling bifurcation.

Subcritical

$$x_{n+1} = -(1 + \mu)x_n - x_n^3 =: F(x_n, \mu),$$

In this case the graphs of $F(x)$ look like:



The bifurcation diagram follows in a similar way to the diagram for the subcritical pitchfork bifurcation.

STABILITY AND CHAOS

2021–22: The stability chapter has been improved to be simpler, and now to provide a preliminary definition of ‘chaos’. This will be one of two definitions presented in this course (the second will be topological in nature and will wait until later chapters).

In this chapter we will investigate the generation of complicated dynamics in discrete time maps of the interval. We take the point of view that ‘complicated dynamics means’ ‘orbit complexity’ i.e. the guaranteed generation of large numbers of periodic orbits. We will begin by discussing two motivating examples. Along with the investigation of the examples we will develop the definitions we need, and prove various straightforward results. Then we will define a more abstract class of dynamical systems: the shift map acting on spaces of sequences of symbols.

‘Symbolic dynamics’ turn out to provide good models for ‘complicated dynamics’. In particular we can count the numbers of N -cycles that are guaranteed to arise. We can then relate the symbolic dynamics to the iteration of continuous maps of an interval. We can prove the existence of N -cycles in the map using the various properties of the symbolic dynamical systems. This is a powerful and very general idea. Our reasoning will be largely ‘topological’ in nature. In the nicest cases we can actually demonstrate a topological conjugacy between the symbolic dynamics and iteration of the map. Even in cases where the symbolic and map dynamics are not topologically conjugate, we can use symbol sequences to derive useful insights.

7.1 The sawtooth map

Example 7.1 (The sawtooth map). *The sawtooth map $F : [0, 1] \rightarrow [0, 1]$ is defined by*

$$x_{n+1} = 2x_n \mod 1 \tag{7.1}$$

which appears to have interesting dynamics, and many periodic points. For example $x = 0$ is a fixed point, $\{\frac{1}{3}, \frac{2}{3}\}$ is a 2-cycle, $\{\frac{1}{5}, \frac{2}{5}, \frac{4}{5}, \frac{3}{5}\}$ is a 4-cycle. We expect all of these invariant sets to be unstable since the gradient $F'(x)$ takes the value 2 everywhere (except at the discontinuity).

In Fig. 7.1, you can see the cobweb path for the sawtooth map $x_{n+1} = 2x_n \mod 1$. Note that iterations around the 4-cycle $\{\frac{1}{5}, \frac{2}{5}, \frac{4}{5}, \frac{3}{5}\}$ starting from the point $x = \frac{1}{5}$ are indicated by the blue arrows.

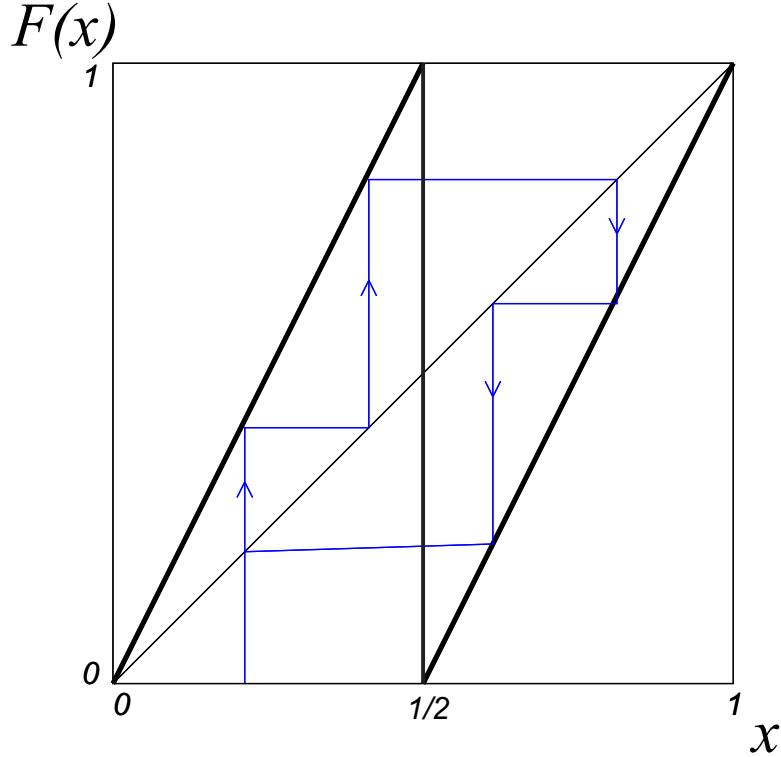


Figure 7.1: Sawtooth map

A nice way to represent the dynamics, that turns out to be the key to understanding how to simplify the analysis of the sawtooth map, is to write points $x \in [0, 1]$ in terms of their base-2 (binary) expansions e.g.

$$\frac{1}{3} = 0 \cdot \overset{\frac{1}{2}}{0} \overset{\frac{1}{4}}{1} \overset{\frac{1}{8}}{0} \overset{\frac{1}{16}}{1} \overset{\frac{1}{32}}{0} \overset{\frac{1}{64}}{1} \dots = \frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \dots = \frac{\frac{1}{4}}{1 - \frac{1}{4}} = \frac{1}{3}.$$

Then the dynamics of applying the map F to a point x , i.e. the operation $x \rightarrow 2x \bmod 1$ map corresponds to shifting the binary sequence one place to the left and then discarding the leading ‘1’ (if any); this second part of the operation is the ‘modulo 1’ part of the operation. Checking this explicitly:

$$F\left(\frac{1}{3}\right) = \frac{2}{3} = 0 \cdot \overset{\frac{1}{2}}{1} \overset{\frac{1}{4}}{0} \overset{\frac{1}{8}}{1} \overset{\frac{1}{16}}{0} \overset{\frac{1}{32}}{1} \overset{\frac{1}{64}}{0} \dots = \frac{1}{2} + \frac{1}{8} + \frac{1}{32} + \dots = \frac{\frac{1}{2}}{1 - \frac{1}{4}} = \frac{2}{3}.$$

Observe that the successive symbols in the base-2 expansion describe which side of $1/2$ the successive iterates land. So a symbol ‘0’ in the j^{th} place to the right of the ‘decimal point’ indicates that the $j - 1^{th}$ iterate of x lies in the interval $[0, 1/2)$, while a ‘1’ indicates that the $j - 1^{th}$ iterate of x lies in the interval $(1/2, 1]$. Also, periodic points have periodic symbol sequences. Applying a map F a total of N times to a point x corresponds to shifting the symbol sequence to the left N times. So, if the symbol sequence is periodic with period N , after N shifts left, we will be back to the sequence we started with.

For example, the symbol sequence

$$0 \cdot \overset{\frac{1}{2}}{0} \overset{\frac{1}{4}}{1} \overset{\frac{1}{8}}{1} \overset{\frac{1}{16}}{1} \overset{\frac{1}{32}}{0} \overset{\frac{1}{64}}{1} \overset{\frac{1}{128}}{1} \overset{\frac{1}{256}}{1} \dots$$

is periodic with period 4 (as indicated by the spacing). So it must correspond to a point x that is on

a 4-cycle. Summing the geometric progressions we have

$$\begin{aligned}
0 \cdot \overset{\frac{1}{2}}{0} \overset{\frac{1}{4}}{1} \overset{\frac{1}{8}}{1} \overset{\frac{1}{16}}{1} \cdots &= \left(\frac{1}{4} + \frac{1}{64} + \cdots \right) + \left(\frac{1}{8} + \frac{1}{128} + \cdots \right) + \left(\frac{1}{16} + \frac{1}{256} + \cdots \right) \\
&= \frac{1/4}{1 - \frac{1}{16}} + \frac{1/8}{1 - \frac{1}{16}} + \frac{1/16}{1 - \frac{1}{16}} \\
&= \frac{7}{16} \cdot \frac{16}{15} = \frac{7}{15}.
\end{aligned}$$

So this symbol sequence corresponds to the point $x = 7/15$ which has iterates under F as follows:

$$\frac{7}{15} \rightarrow \frac{14}{15} \rightarrow \frac{13}{15} \rightarrow \frac{11}{15} \rightarrow \frac{7}{15}.$$

So $x = 7/15$ is part of a 4-cycle; this is a different 4-cycle to the one we found above: $\{\frac{1}{5}, \frac{2}{5}, \frac{4}{5}, \frac{3}{5}\}$. And the successive 0s and 1s in the base-2 expansion show that three of the points on this new 4-cycle lie in $x > 1/2$ and only one in $x < 1/2$.

Note: There are potential problems with this nice correspondence between the symbolic dynamics of ‘base-2 expansions’ and the dynamics of points of the form $p/2^n$ for integers p and n . Binary expansions are not unique for points of the form $p/2^n$ since $0 \cdot 011111\cdots$ codes for the same point ($x = \frac{1}{2}$) as $0 \cdot 100000\cdots$, etc. This is exactly the same issue as saying $0 \cdot 9999999 = 1 \cdot 0000000$ in decimal expansions. We will return to this point at the end of the next section

7.2 Three properties of a chaotic system

The sawtooth map exhibits chaos!

This is your first taste of an incredibly simple dynamical system that is “easily” proven to exhibit infinite complexity. We will introduce you to three essential criteria and then explain how the sawtooth map satisfies these criteria.

Definition 7.1 (Dense). Suppose X is a set and $Y \subseteq X$. We say that Y is dense in X if, for any point $x \in X$, there is a point $y \in Y$ arbitrarily close to x .

Equivalently, Y is dense in X if, for any $x \in X$, there exists a sequence of points $\{y_n\} \in Y$ that converges to x .

For example, take $X = \mathbb{R}$ and let Y be the rational numbers. To prove that Y is dense in X , we must show, for every $x \in \mathbb{R}$, there is a sequence of rational numbers that approach x .

First, if x is itself rational, then we are done. Assume x is irrational and take $x = a_n \cdots a_0.b_1b_2b_3\cdots$. We can now design the rational sequence to be

$$\begin{aligned}
y_1 &= a_n \cdots a_0.b_1, \\
y_2 &= a_n \cdots a_0.b_1b_2, \\
y_3 &= a_n \cdots a_0.b_1b_2b_3,
\end{aligned}$$

and so forth. Clearly $y_n \rightarrow x$ as $n \rightarrow \infty$.

We set aside the definition of denseness for now.

We propose the following definition for sensitive dependence on initial conditions.

Definition 7.2 (Sensitive dependence on initial conditions (SDIC)). *A dynamical system F has sensitive dependence on initial conditions (SDIC) if $\exists \delta > 0$ such that, for any x and for any $\varepsilon > 0$, there exists a point y within the neighbourhood, $|x - y| < \varepsilon$, and $n \geq 0$ such that $|F^n(x) - F^n(y)| > \delta$.*

i.e. near any x there is always some point that separates to at least a distance δ away under iteration by F . This is a formalised version of the notion that even the smallest errors in initial conditions (or in a numerical integration scheme) inevitably grow to become as large as the true value of the state, meaning that it becomes impossible to predict the future behaviour of the system, even though it remains entirely deterministic. This is often referred to as the ‘butterfly effect’ and makes prediction of, for example, the weather, extremely uncertain once one looks more than roughly five days ahead.

Definition 7.3 (Topologically transitive (TT)). *A dynamical system is transitive (or topologically transitive) if for any pair of points x and y and any $\varepsilon > 0$ there is a third point z with $|z - x| < \varepsilon$ whose orbit comes within ε of y , i.e. $|F^n(z) - y| < \varepsilon$ for some $n \geq 0$.*

This ensures that even the smallest open sets eventually intersect and ‘mix together’ under iteration of the map.

Proposition 7.1 (Equivalency between dense orbits and transitivity). *(i) A dynamical system that has a dense orbit is transitive and (ii) a transitive dynamical system has a dense orbit.*

The proof of statement (i) is essentially immediate. The existence of a dense orbit posits a point z where the orbit $O(z)$ gets arbitrarily close to every other point in the space. Hence F is transitive. The converse (ii) is harder to prove¹.

Remark: The two properties SDIC and TT are independent. For example the irrational rotation $\theta_{n+1} = \theta_n + \omega$ for ω irrational is TT on the circle $S^1 \simeq \mathbb{R}/\mathbb{Z} \simeq [0, 1]/\sim$ (i.e. identify the endpoints 0 and 1), but points do not move away from (or towards) each other, so this rotation map does not have SDIC. On the other hand, the map $x_{n+1} = 2x_n$ on \mathbb{R} displays SDIC (the distance between points doubles with each iteration) but is not TT.

We thus state one possible definition of chaos.

Definition 7.4 (The D+TT+SDIC criterion of chaos). *A dynamical system F is chaotic if:*

1. (D) Periodic points for F are dense.
2. (TT) F is transitive.
3. (SDIC) F exhibits sensitive dependence on initial conditions.

¹And is known as the Baire Category Theorem from real analysis

7.3 Chaos in the sawtooth map

Proposition 7.2 (Chaos in the sawtooth map). *The sawtooth map (7.1) is chaotic on $\Lambda = [0, 1]$.*

Proof of D:

We wish to show that the periodic points are dense. To do this, we must show that near any point $x \in [0, 1]$, it is possible to find a periodic point. Then construct the binary expansion for x , say

$$x = 0 \cdot a_0 a_1 a_2 \cdots a_{n-1} a_n a_{n+1} \cdots$$

We can then construct a periodic point that matches x to $n + 1$ entries,

$$p_n = 0 \cdot \overline{a_0 a_1 a_2 \cdots a_{n-1} a_n}$$

So p_n is a periodic point of period $n + 1$. Notice that $|x - p_n| \leq \frac{1}{2^n}$. So taking $n \rightarrow \infty$ means we can find a periodic point arbitrarily close to x [one can set $\varepsilon < 1/2^n$].

Proof of SDIC:

Set $\delta = \frac{1}{4}$. Given $\varepsilon > 0$ and $x \in [0, 1]$, pick n such that $2^{-(n+1)} < \varepsilon$. Then construct the binary expansion for x , say

$$0 \cdot a_0 a_1 a_2 \cdots a_{n-1} a_n a_{n+1} \cdots$$

Take y to be the point with symbol sequence

$$0 \cdot a_0 a_1 a_2 \cdots a_{n-1} \bar{a}_n a_{n+1} \cdots$$

where $\bar{a}_n = 1 - a_n$ means change the symbol a_n from a 0 to a 1 or vice-versa as appropriate. Then we see that $|x - y| = 2^{-(n+1)} < \varepsilon$ but $|F^n(x) - F^n(y)| = \frac{1}{2} > \delta$.

Proof of TT:

Construct a point x which has a dense orbit, i.e. comes arbitrarily close to any given point $y \in [0, 1]$. Let x be the point given by the binary expansion

$$0 \cdot \underbrace{0 1}_{\text{length 1}} \underbrace{00 01}_{\text{length 2}} \underbrace{10 11}_{\text{length 3}} \underbrace{000 001 010 011}_{\text{length 4}} \underbrace{100 101 110 111}_{\text{length 5}} \cdots$$

taking all blocks of lengths $1, 2, 3, \dots$ in order. Then for any point y we can compute the corresponding symbol sequence $0 \cdot a_0 a_1 a_2 \cdots a_{n-1} a_n a_{n+1} \cdots$. Then for any $n > 0$ there exists a $k > 0$ such that the binary expansion of $F^k(x)$ agrees with the expansion of y on at least the first n places, implying $|F^k(x) - y| < 2^{-n}$, i.e. the forward orbit of x comes arbitrarily close to any point y and this is a dense orbit.

TOPOLOGICAL CONJUGACY

Before we explain the relationship of the sawtooth map to other maps, we want to move into the direction of establishing a formal correspondence between symbolic dynamics and maps. In order for us to do that, we should define the notion of topological conjugacy, which is a way of explaining why one map may resemble another.

Definition 8.1 (Topological conjugacy). *Two maps $F : X \rightarrow X$ and $G : Y \rightarrow Y$ are topologically conjugate if \exists a homeomorphism $h : X \rightarrow Y$ (i.e. a continuous bijection with a continuous inverse) such that*

$$h \circ F(x) = G \circ h(x) \quad \forall x \in X$$

i.e. diagrammatically we have the situation shown in Fig.8.1.

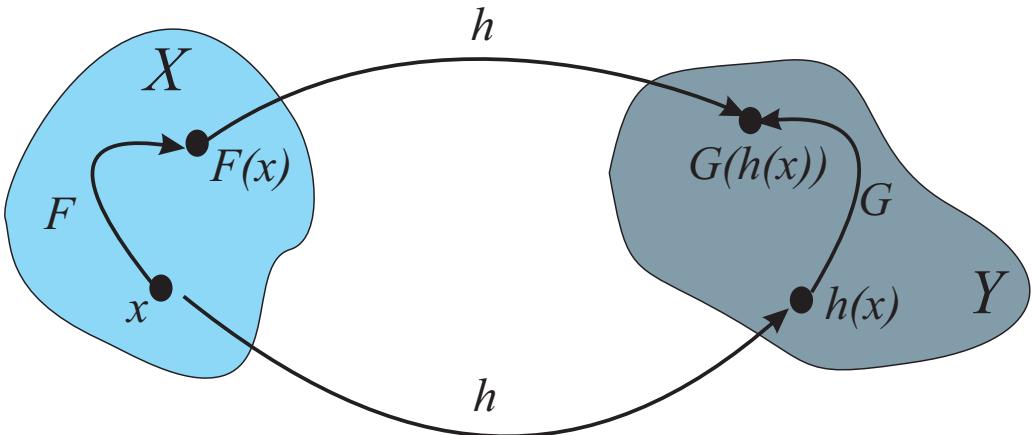


Figure 8.1: topologicalconjugacy

The key is that topological conjugacy should hopefully preserve many important properties. For example, a topological conjugacy maps fixed points to fixed points and N -cycles to N -cycles.

Topological conjugacy can also allow us to define more rigorous notions of stability. For instance, let us informally say that a map F is *structurally stable* if we cannot create new invariant sets through small perturbations. Thus, if we perturb the map via $\tilde{F}(x) = F(x) + f(x)$ where f is small, then the question of whether \tilde{F} is topologically conjugate to F relates to the issue of structural stability

(since topological conjugacy preserves fixed points and N -cycles). The following example reinforces this notion of structural stability.

Example 8.1 (Non-structurally stable). *The map $F(x) = 1 - x$ on $X = [0, 1]$ is not structurally stable.*

The above map is not structurally stable because all points in $[0, 1]$ except $1/2$ lie on 2-cycles of the form $\{a, 1 - a\}$ for $0 \leq a < 1/2$. But the perturbed map

$$\tilde{F}(x) := F(x) + f(x) = 1 - \frac{\delta}{2} - (1 - \delta)x$$

for any small positive δ (certainly we require $0 < \delta < 1$) has the property that all orbits starting in $[0, 1]$ asymptotically tend to $x = 1/2$. We can see this informally by ‘cobwebbing’.

Analytically, let $y_n = x_n - 1/2$, then

$$y_{n+1} = \tilde{F}(x_n) - \frac{1}{2} = -(1 - \delta)y_n$$

and thus

$$\Rightarrow |y_n| = |y_0|(1 - \delta)^n \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Note that we could have previously defined the bifurcation from previous chapters as follows:

Example 8.2. *For a family of maps $F(x, \mu)$, a bifurcation occurs at a parameter value $\mu = \mu_0$ at which $F(x, \mu_0)$ is not structurally stable.*

However, this more technical definition of a bifurcation is not needed in our course.

CHAPTER 9

SYMBOLIC DYNAMICS

In this section we will define a new class of dynamical systems and prove that they have nice properties. We will then set up a topological conjugacy between one of these nice dynamical systems and the logistic map when μ is large enough. This enables us to understand completely the dynamics of the logistic map.

Recall that, in our investigation of the sawtooth map we used binary sequences and the action of F was equivalent to shifting the binary sequence along one place. This motivates the following definition.

Definition 9.1 (Sequence space on N symbols). *The sequence space on N symbols Σ_N is the collection of semi-infinite sequences of symbols, each symbol drawn from the set $\{0, 1, \dots, N - 1\}$:*

$$\Sigma_N = \{\mathbf{a} = (a_0 a_1 a_2 \dots) : a_i \in \{0, 1, \dots, N - 1\} \forall i \geq 0\}$$

Definition 9.2 (Sequence space distance). *The distance between two symbol sequences \mathbf{a}, \mathbf{b} in Σ_N is defined by*

$$d(\mathbf{a}, \mathbf{b}) = \sum_{n=0}^{\infty} \frac{\gamma(a_n, b_n)}{3^n},$$

where $\gamma(p, q) = 0$ if $p = q$ and $\gamma(p, q) = 1$ if $p \neq q$.

Under this distance measure (metric), two elements of the set (or ‘points’ in the state space) Σ_N are close together if they agree on a long initial segment.

Proposition 9.1. *Let $\mathbf{a}, \mathbf{b} \in \Sigma_N$ and $a_i = b_i$ for $0 \leq i < m$ and then $a_m \neq b_m$. Then*

$$3^{-m} \leq d(\mathbf{a}, \mathbf{b}) \leq \frac{3}{2}3^{-m}.$$

Proof: The first $m - 1$ symbols of \mathbf{a} are the same in \mathbf{b} so $\gamma(a_i, b_i) = 0$ for $0 \leq i < m$. Then symbol m differs, so $\gamma(a_m, b_m) = 1$. The minimum distance between \mathbf{a} and \mathbf{b} occurs when all of the remaining symbols agree, so $\gamma(a_i, b_i) = 0$ for $i > m$. In this case $d(\mathbf{a}, \mathbf{b}) = 3^{-m}$. The maximum distance between \mathbf{a} and \mathbf{b} occurs when none of the remaining symbols agree, so $\gamma(a_i, b_i) = 1$ for $i > m$. In this case $d(\mathbf{a}, \mathbf{b}) = \sum_{n=m}^{\infty} 3^{-n} = \frac{3}{2}3^{-m}$.

Definition 9.3 (shift map). *The shift map $\sigma : \Sigma_N \rightarrow \Sigma_N$ acts by dropping the first symbol from a sequence and shifting the remainder to the left $\sigma(a_0a_1a_2\cdots) = (a_1a_2a_3\cdots)$.*

Properties of the shift map

1. $\sigma : \Sigma_N \rightarrow \Sigma_N$ is continuous.

Proof: For any two symbol sequences \mathbf{a}, \mathbf{b} ,

$$\begin{aligned} d(\mathbf{a}, \mathbf{b}) &= \gamma(a_0, b_0) + \frac{1}{3}d(\sigma(\mathbf{a}), \sigma(\mathbf{b})) \\ \Rightarrow d(\sigma(\mathbf{a}), \sigma(\mathbf{b})) &\leq 3d(\mathbf{a}, \mathbf{b}) \end{aligned}$$

Hence, given $\varepsilon > 0$ set $\delta = \varepsilon/3$. Then, if $d(\mathbf{a}, \mathbf{b}) < \delta$, $d(\sigma(\mathbf{a}), \sigma(\mathbf{b})) < \varepsilon$. So, as $\mathbf{a} \rightarrow \mathbf{b}$, $\sigma(\mathbf{a}) \rightarrow \sigma(\mathbf{b})$.

2. σ^k has N^k fixed points.

Proof: $\sigma^k(\mathbf{a}) = \mathbf{a} \iff a_{k+j} = a_j \forall j \geq 0$. So \mathbf{a} is determined by the initial block $(a_0 \cdots a_{k-1})$. Since there are N independent choices for each symbol, there are exactly N^k distinct blocks of length k .

3. The set of periodic points of σ , $\text{Per}(\sigma)$, is dense in Σ_N (i.e. periodic points exist arbitrarily close to any given symbol sequence).

Proof: Given $\mathbf{a} \in \Sigma_N$ and $\varepsilon > 0$, take n such that $\frac{3}{2}3^{-n} < \varepsilon$. Let $\mathbf{b} = (a_0a_1\cdots a_{n-1}a_0a_1\cdots a_{n-1}a_0\cdots)$. Then, by Proposition 9.1, $d(\mathbf{a}, \mathbf{b}) < \varepsilon$ and \mathbf{b} is clearly a periodic symbol sequence.

4. $\sigma : \Sigma_N \rightarrow \Sigma_N$ is TT because there exists a point $\mathbf{a} \in \Sigma_N$ with a dense orbit.

Proof: Let \mathbf{a} be the symbol sequence given by listing all blocks of length 1, then all blocks of length 2, and so on, e.g. for $N = 2$:

$$\mathbf{a} = (\underbrace{0}_{} \underbrace{1}_{} \underbrace{00}_{} \underbrace{01}_{} \underbrace{10}_{} \underbrace{11}_{} \underbrace{000}_{} \underbrace{001}_{} \underbrace{010}_{} \underbrace{011}_{} \underbrace{100}_{} \underbrace{101}_{} \underbrace{110}_{} \underbrace{111}_{} \cdots).$$

Then, given a point $\mathbf{b} \in \Sigma_N$ and $\varepsilon > 0$ pick k large enough such that $\frac{3}{2}3^{-k} < \varepsilon$. Then, because \mathbf{a} contains all length- k symbol strings, there exists $n > 0$, depending on k , such that the shifted version of \mathbf{a} , $\sigma^n(\mathbf{a})$, agrees with \mathbf{b} in the first k places. It follows from Proposition 9.1 that $d(\sigma^n(\mathbf{a}), \mathbf{b}) \leq \frac{3}{2}3^{-k} < \varepsilon$.

5. $\sigma : \Sigma_N \rightarrow \Sigma_N$ has SDIC.

Proof: Let $\delta = 1/2$. Given $\mathbf{a} \in \Sigma_N$ and $\varepsilon > 0$ pick n such that $3^{-n} < \frac{3}{2}3^{-n} < \varepsilon$. Let \mathbf{b} differ from \mathbf{a} for the first time in the n^{th} place in the symbol sequence. Then, by Proposition 9.1, $3^{-n} \leq d(\mathbf{a}, \mathbf{b}) \leq \frac{3}{2}3^{-n} < \varepsilon$. But $d(\sigma^n(\mathbf{a}), \sigma^n(\mathbf{b})) \geq 1 > \delta$ since at least the first symbols of the shifted sequences differ.

THE LOGISTIC MAP WITH $\mu > 4$

The goal of this mini-chapter is essentially to explain the occurrence of the Cantor middle thirds set within the context of the logistic map when $\mu > 4$. This is a very visual procedure, so instead of typeset notes, we have included previous visualiser notes, and the student should refer to the video lecture for the explanations.

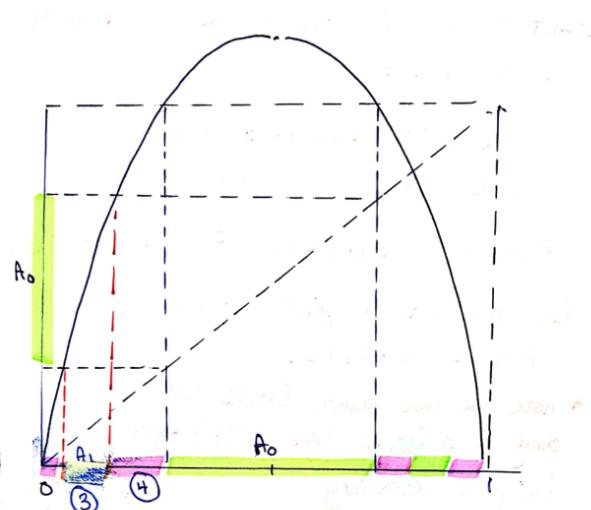
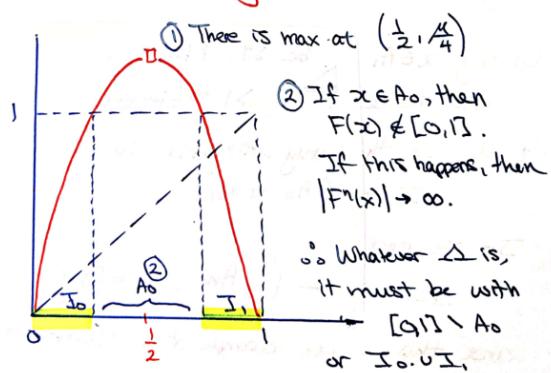
(CHAP.8) Studying the logistic map for $\mu > 4$ 11 NOV. 19.

The goal is to understand the invariant set I for the logistic map,

$$F(x) = \mu x(1-x), \quad F: [0,1] \rightarrow [0,1]$$

i.e. the set of points where $F^n(x) \in [0,1]$.

* note we care about $[0,1]$ s.t. any $x \notin [0,1]$ has $|F^n(x)| \rightarrow \infty$, i.e. not interesting.



③ Any $x \in A_1$ will be s.t. $F(x) \in A_0$ and hence $F^2(x) > 1 \Rightarrow$ ejected.

④ We are thus only interested in $I - (A_0 \cup A_1)$

In the end

$$I = I - \bigcup_{n=0}^{\infty} A_n, \quad I = [0,1].$$

and this is an example of a Cantor set.

⑤ Consider $I - (A_0 \cup A_1)$.

- This has 4 closed intervals and
- F maps each one monotonically to I_0 or I_1 ,
- But F maps I_0 to I and similarly I_1 to I.
- $\therefore F^2$ maps to 4 intervals to I.
- \therefore Each (pink) interval has a centre section mapped to A_0 .

Let us call each centre section (four of them) A_2 .

Now consider $I - (A_0 \cup A_1 \cup A_2)$

⑥ Inductively, each A_n (centre section) has 2^n disjoint open intervals.

Hence $I - (A_0 \cup \dots \cup A_n)$ consists of 2^{n+1} closed intervals

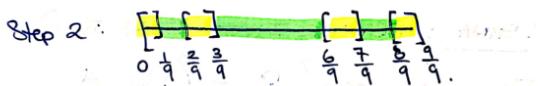
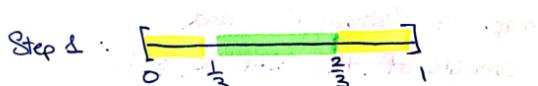
Why?

$$\text{size of } \bigcup_{j=0}^n A_j = 1 + 2 + 2^2 + \dots + 2^n = 2^{n+1} - 1$$

Example : (CANTOR MIDDLE THIRDS)

Start with $I = [0, 1]$. Remove the middle third $(\frac{1}{3}, \frac{2}{3})$; remove middle third of remaining intervals, ...

(continued) ... to infinity



THEOREM : If $\mu > 2 + \sqrt{5}$, then Λ is a CANTOR SET.

PF : We've seen for $\mu > 4$, but in fact true for slightly adjusted μ [see: Devaney (2008)].

Also done in Chap. 9.

⑦ Thus we consider

$$\Lambda = I - \bigcup_{n=0}^{\infty} A_n$$

i.e. Λ consists of removing the "middles" of closed intervals

This is similar to the construction of the Cantor Middle Thirds Set.

Def'n : (CANTOR SET)

A set Λ is a CANTOR SET if it is

(i) closed

(ii) totally disconnected [it contains no intervals]

(iii) perfect subset of I

[every point is a limit point or an accumulation point of other points in Λ].

(we don't need this def'n).

CONJUGACY BETWEEN THE LOGISTIC MAP, F , AND SHIFT MAP, σ

Having proved various nice properties of $\sigma : \Sigma_N \rightarrow \Sigma_N$ (a rather abstract dynamical system), we now make a direct link between the abstract and the ‘real’ dynamics of a continuous map F . The idea is to be able to translate our results about σ acting on Σ_N back into results about F acting on the unit interval I .

Lemma 11.1 (Cantor Intersection Theorem). *The intersection $S_\infty = \cap_{i=0}^\infty S_i$ of an infinite sequence $S_0 \supseteq S_1 \supseteq S_2 \supseteq \dots$ of nested non-empty closed bounded subsets $S_i \subset \mathbb{R}^n$ is non-empty.*

[[Proof: Consider the sequence of real numbers $\{a_k\}$ where a_k is the infimum over the non-empty set S_k . Because S_k is closed (and so it contains all its limit points), $a_k \in S_k$. Also, because the sets S_k are nested inside each other, the sequence $\{a_k\}$ is monotonically increasing. The sequence $\{a_k\}$ is also bounded (all values in the sequence are contained in the bounded set S_0), so it therefore must converge to some limit value L . Now choose any $j \geq 0$; the subsequence $\{a_k\}$ for $k \geq j$ is contained in S_j and converges to L . Since S_j is closed, L lies in S_j . And this is true for all j , so the limit point L lies in all S_j , and so L is contained in their intersection. \square **]]**

Remark 11.1. *If in addition $\text{diam}(S_k) < \alpha \text{ diam}(S_{k-1})$ for some $0 < \alpha < 1$, where $\text{diam}(S_k) := \sup\{|x - y| : x, y \in S_k\}$ is the diameter of the set S_k , then S_∞ , the infinite intersection of the closed nested sets S_k , is a single point.*

These following definitions should already be known to you from other courses, but we will need them in the Theorem to come. The definition of conjugacy is the same as that which you have encountered in the previous chapters. We introduce in addition the definition of semi-conjugacy which will be needed in future chapters.

The following definition on injectivity should already be known to you, but it is repeated here:

Definition 11.1 (Surjective, injective). Recall a map $h : \Lambda \rightarrow Y$ is surjective (onto) if for all $y \in Y$ there exists a point $x \in \Lambda$ such that $h(x) = y$; h is injective (one-one) if $h(x_1) = h(x_2) \iff x_1 = x_2$.

Now rather than going straight to conjugacy, let us define a (topological) semi-conjugacy, which is a weaker version.

Definition 11.2 (Semiconjugate). Let $F : I \rightarrow \mathbb{R}$ be a continuous map of the interval, and let $\Lambda \subseteq I$ be an invariant set. Let $G : Y \rightarrow Y$ be a continuous map on a (metric) space Y . If there exists a continuous surjection $h : \Lambda \rightarrow Y$ such that $h \circ F = G \circ h$ then F is semiconjugate to G via h .

Finally, we can obtain the same definition of topological conjugacy as in a previous chapter.

Definition 11.3 (Conjugate). If F is semiconjugate to G via h and h is injective and h^{-1} is continuous, so that h is a homeomorphism, then F is (topologically) conjugate to G via h .

11.1 Conjugacy between the logistic and shift maps

Now having reviewed our definition of topological conjugacy and now added in the weaker version of semi-conjugacy, there is a Lemma that we require that essentially ensures that the logistic map has an important bound on the steepness of the gradient when μ is sufficiently large.

Lemma 11.2. Let $F : I \rightarrow I$ be the logistic map $F(x) = \mu x(1 - x)$. If $\mu > 2 + \sqrt{5}$ then there exists $\lambda > 1$ such that $|F'(x)| > \lambda$ for all $x \in I \cap F^{-1}(I)$.

For $\mu > 2 + \sqrt{5}$ the magnitude of the slope of the logistic map is greater than unity within the parts of $[0, 1]$ that are mapped inside $[0, 1]$. The proof is a straightforward exercise.

Theorem 11.1. The logistic map $F(x) = \mu x(1 - x)$, with $\mu > 2 + \sqrt{5}$, has an invariant set $\Lambda \subset [0, 1]$ on which $F|_{\Lambda}$ is conjugate to $\sigma|_{\Sigma_2}$.

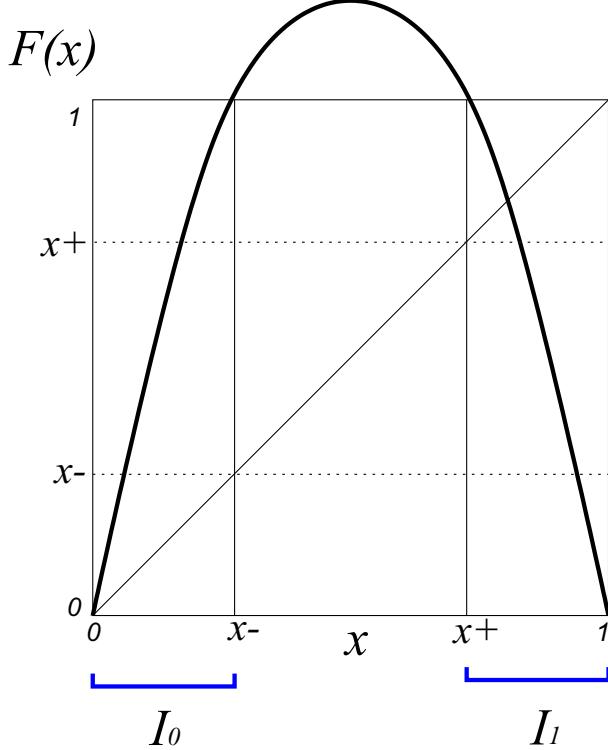
Recall that the invariant set Λ is the collection of points that, when iterated under F , remain in $[0, 1]$ for all time, i.e. $\Lambda = \{x \in [0, 1] : F^n(x) \in [0, 1] \forall n \geq 0\}$.

Proof

To prove the theorem we need to

- construct a map $h : \Lambda \rightarrow \Sigma_2$

- show that h is injective
- show that h is surjective
- show that h is continuous with a continuous inverse.



The Logistic Map $x_{n+1} = \mu x_n(1 - x_n)$ for $\mu > 2 + \sqrt{5}$.

Proof: (Construct h). By observation, the set $I \cap F^{-1}(I)$ is a disjoint union of two closed intervals $I_0 = [0, x_-]$ and $I_1 = [x_+, 1]$ where $x_\pm = (1 \pm \sqrt{1 - 4/\mu})/2$ are the points at which $F(x) = 1$. In the open interval (x_-, x_+) , F maps points above $x = 1$ so they cannot be part of the invariant set (in fact, iterates move rapidly off to $-\infty$). Define the symbol sequence \mathbf{a} corresponding to a point $x \in \Lambda$ by

$$\begin{aligned} a_j &= 0 && \text{if } F^j(x) \in I_0 \\ a_j &= 1 && \text{if } F^j(x) \in I_1 \end{aligned}$$

Then define the map $h : \Lambda \rightarrow \Sigma_2$ is defined by setting $h(x) = \mathbf{a}$. □

Proof: (h is injective). Given $x, y \in \Lambda$, show that $h(x) = h(y) \iff x = y$. Suppose there exists $x, y \in \Lambda$ with $x \neq y$ and $h(x) = h(y) = \mathbf{a}$. Then (from the definition of \mathbf{a}) $F^j(x)$ and $F^j(y)$ are always on the same side of $x = \frac{1}{2}$ as each other (they are always in the same interval I_0 or I_1). Hence $|F^j(x) - F^j(y)| < \frac{1}{2} \forall j$. But, by Lemma 11.2 $|F'(x)| > \lambda > 1$ so $|F^j(x) - F^j(y)| > \lambda^j |x - y|$ for all j . The right-hand side eventually becomes greater than $\frac{1}{2}$ resulting in a contradiction unless $|x - y| = 0$. Hence $x = y$. □

Proof: (h is surjective). Given a symbol sequence $\mathbf{a} = (a_0 a_1 a_2 \dots)$, show there exists $x \in \Lambda$ such that $F^j(x) \in I_{a_j} \forall j$. Let $J \subset I$ be a closed interval. Let $F^{-1}(J) = \{x : F(x) \in J\}$ denote the preimage of J . From the graph of F , $F^{-1}(J)$ is the disjoint union of a pair of closed subintervals, one in each of I_0 and I_1 . Define

$$\begin{aligned} I_{a_0 a_1 \dots a_n} &= \{x : x \in I_{a_0}, F(x) \in I_{a_1}, \dots, F^n(x) \in I_{a_n}\} \\ &= I_{a_0} \cap F^{-1}(I_{a_1}) \cap F^{-2}(I_{a_2}) \cap \dots \cap F^{-n}(I_{a_n}). \end{aligned}$$

Combining these definitions

$$I_{a_0 a_1 \dots a_n} = I_{a_0} \cap F^{-1}(I_{a_1 a_2 \dots a_n})$$

because these are both exactly the sets of points for which $F^j(x) \in I_{a_j}$ for $0 \leq j \leq n$.

Assume that $I_{a_k \dots a_n}$ is a non-empty closed interval. This is certainly true for $k = n$ i.e. I_{a_n} . Then

$F^{-1}(I_{a_k \dots a_n})$ consists of a pair of closed intervals, one in I_0 and one in I_1 . Hence

$I_{a_{k-1} \dots a_n} = I_{a_{k-1}} \cap F^{-1}(I_{a_k \dots a_n})$ is exactly one non-empty closed interval. It follows by induction that $I_{a_0 a_1 \dots a_n} = I_{a_0} \cap F^{-1}(I_{a_1 \dots a_n})$ is a single non-empty closed interval for all n .

Now,

$$I_{a_0 \dots a_n} = I_{a_0 \dots a_{n-1}} \cap F^{-n}(I_{a_n}) \subset I_{a_0 \dots a_{n-1}}$$

So the intervals $I_{a_0 \dots a_n}$ are nested closed intervals. Hence the Cantor Intersection Theorem implies

$$I_{\mathbf{a}} := \bigcap_{n=0}^{\infty} I_{a_0 \dots a_n}$$

is non-empty. So there exists x such that $F^j(x) \in I_{a_j}$ for all j . Since h is one-one, $I_{\mathbf{a}}$ contains exactly one point. We can also see this since the length of the intervals $I_{a_0 \dots a_n}$ tends to zero at least as fast as λ^{-n} as $n \rightarrow \infty$ because of the condition that $|F'(x)| > \lambda > 1$ everywhere. \square

Proof: (h and h^{-1} are continuous). For any pair of points $x, y \in \Lambda$ we have established that there exist unique symbol sequences $\mathbf{a} = h(x)$ and $\mathbf{b} = h(y)$. Now, $I_{a_0 \dots a_n}$ is a closed bounded interval for all n . So x tends to y ($x \rightarrow y$) \iff $F^j(x)$ and $F^j(y)$ remain on the same side of $\frac{1}{2}$ for a larger number of iterates j (by continuity of F) \iff \mathbf{a} and \mathbf{b} agree on a longer and longer initial segment of symbols $\iff d(\mathbf{a}, \mathbf{b}) \rightarrow 0 \iff h(x) \rightarrow h(y)$. \square

Remark 11.2 (on Theorem 11.1).

1. Since $h \circ F|_{\Lambda} = \sigma \circ h$, $F|_{\Lambda}$ is conjugate to $\sigma|_{\Sigma_2}$ so the dynamics of $F|_{\Lambda}$ has all the properties of $\sigma|_{\Sigma_2}$, including dense periodic points and TT. There also both have SDIC, although this does not necessarily follow from conjugacy.
2. The invariant set Λ is an example of a Cantor set: it is closed, contains no intervals (a set with this property is called ‘totally disconnected’) and every point in Λ is a limit point of a sequence of points in Λ (a set with this property is called ‘perfect’). The classic Cantor set is the ‘middle-third’ construction which arises in a closed related problem.
3. Conjugacy between $F|_{\Lambda}$ and $\sigma|_{\Sigma_2}$ holds in fact for any $\mu > 4$. But the proof requires more careful estimates of the rates of separation of nearby points to guarantee that h is injective.

Remark 11.3. For the sawtooth map, the existence of two symbol sequences for points $p/2^n$ implies h cannot be injective, and therefore could not be a conjugacy. But the non-uniqueness of binary expansions for points of the form $x = p/2^n$ for integer p and n prevents the dynamics of the sawtooth map being even semiconjugate to the shift map acting on Σ_2 . The problem is that the natural map h taking points x to the symbol sequence of labels for the location of future iterates is not continuous. To see this pick

two sets of points $x^{(j)}, y^{(j)} \in [0, 1]$ that converge to $1/2$ but whose corresponding symbol sequences begin with a 0 and a 1 respectively, say $\mathbf{a}^{(j)} = 0 \cdot 0 \dots$ and $\mathbf{b}^{(j)} = 0 \cdot 1 \dots$. Then $d(\mathbf{a}^{(j)}, \mathbf{b}^{(j)}) \geq 1$ for all j even if the sequences of points $|x^{(j)} - y^{(j)}| \rightarrow 0$. So h is not continuous at $1/2$, or indeed at any of the dense set of points of the form $p/2^n$ by the same argument. More fundamentally, this problem is caused by the overlap of the end points of the two intervals $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$. If we work with disjoint intervals this problem disappears and we can indeed find a semiconjugacy h .

Remark 11.4. In general, to show conjugacy (rather than just semiconjugacy) between symbolic dynamics and map iterates needs a guarantee of injectivity. Usually this demands an expansivity property of the map, i.e. points move apart after every iterate. Usually such a property can only be guaranteed given an explicit definition of a map - see the logistic map example with $\mu > 2 + \sqrt{5}$ where we know $|F'| > \lambda > 1$. This kind of metric condition also often guarantees SDIC.

SUBSHIFTS OF FINITE TYPE (SSFT)

A useful refinement of the sequence space Σ_N is to restrict the sequences of symbols that are allowed to occur. Allowed symbol sequences are encoded by an $N \times N$ transition matrix A .

Definition 12.1 (transition matrix). *A transition matrix for the sequence space Σ_N is an $N \times N$ matrix A with $A_{ij} = 1$ if symbol j is allowed to follow symbol i and $A_{ij} = 0$ if symbol j is not allowed to follow symbol i .*

Note: matrix indices now range over $0 \leq i, j \leq N - 1$ rather than $1 \leq i, j \leq N$. This will not cause notational or other difficulties but it is worth pointing out in order to avoid confusion.

This idea of transitions between states is similar to that developed in the theory of discrete time Markov chains.

Example 12.1. Let $N = 2$ (binary sequence space). If the transitions, or symbol pairs, 01, 10 and 11 are allowed but 00 is not, then

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Example 12.2. Let $N = 2$. If only the transitions, or symbol pairs, 01 and 11 are allowed, then

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

Definition 12.2 (Shift space). *The shift space $\Sigma_{N,A}$ is the set of allowed sequences, a closed subset of Σ_N :*

$$\Sigma_{N,A} = \{ \mathbf{a} \in \Sigma_N : A_{a_n a_{n+1}} = 1 \ \forall n \geq 0 \}.$$

Note $\Sigma_{N,A}$ is invariant under the shift σ .

Definition 12.3 (Subshift of Finite Type (SSFT)). *A subshift of finite type is a dynamical system defined by the action of $\sigma : \Sigma_{N,A} \rightarrow \Sigma_{N,A}$. Write σ_A as shorthand for $\sigma|_{\Sigma_{N,A}}$.*

Definition 12.4 (Irreducible). *The transition matrix A is irreducible if, for all ordered pairs of symbols (i, j) there exists an $n \geq 0$ such that $(A^n)_{ij} \geq 1$ i.e. for all symbol pairs (i, j) there exists an allowed symbol sequence of some length that starts with i and ends with j .*

Definition 12.5 (Non-trivial). *The transition matrix A is non-trivial if, for some i there exists j_1 and j_2 ($j_1 \neq j_2$) such that the sequences ij_1 and ij_2 are allowed.*

Note that this excludes permutation matrices such as

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \quad (12.1)$$

Properties of the SSFT σ_A

1. The number $N_{ij}^{(n)}$ of allowed sequences of length $n+1$ from symbol i to symbol j , say $ia_1a_2 \cdots a_{n-1}j$, is given by $(A^n)_{ij}$.

Proof. The result holds for $n = 1$ since $N_{ij}^{(1)} = A_{ij} = 1$ if ij is allowed, 0 if not. Assume, for induction, that the result holds for $n - 1$. So the number of allowed sequences of length n from symbol i to symbol m is $N_{im}^{(n-1)} = (A^{n-1})_{im}$ by assumption. All sequences of length $n + 1$ from i to j are composed of sequences of length n from i to some m , followed by j . If the transition mj is allowed then $A_{mj} = 1$. Hence the total number of sequences from i to j of length $n + 1$ is

$$\begin{aligned} \sum_{A_{mj}=1} N_{im}^{(n-1)} &= \sum_{A_{mj}=1} (A^{n-1})_{im} \\ &= \sum_m (A^{n-1})_{im} A_{mj} = (A^{n-1}A)_{ij} = (A^n)_{ij} \end{aligned}$$

The number of points in (not necessarily least) period- n orbits P_n (i.e. fixed points of σ_A^n) is given by $\text{tr}(A^n)$.

Proof. Period- n orbits are exactly those sequences where $i = j$. So

$$P_n = \sum_i N_{ii}^{(n)} = \sum_i (A^n)_{ii} = \text{tr}(A^n).$$

Let N_q be the number of q -cycles (i.e. periodic orbits of least period q). Then $P_n = \sum_{q|n} qN_q$, where the sum is taken over integers q that are factors of n .

Proof. If n contains q as a factor then (repeated) q -cycles are period- n orbits. Hence each q -cycle contributes q points to every set of points in period- n orbits, P_n .

Note that $\text{tr}(A^n)$ can be computed from the recurrence relation given by the Cayley–Hamilton Theorem.

Example 12.3. $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ has characteristic polynomial $P(\lambda) = \lambda^2 - \lambda - 1$. So, by Cayley–Hamilton, $A^2 - A - I = 0$ and for any non-negative integer m

$$\text{tr}(A^{m+2}) = \text{tr}(A^{m+1}) + \text{tr}(A^m).$$

Hence

$$\begin{aligned} P_1 &= 1, & P_2 &= 3, & P_3 &= 4, & P_4 &= 7, & P_5 &= 11. \\ N_1 &= 1, & N_2 &= 1, & N_3 &= 1, & N_4 &= 1, & N_5 &= 2. \end{aligned}$$

So there are exactly two distinct 5-cycles for σ_A , i.e. exactly two distinct allowed symbol sequences with least period-5. They are $(01011\cdots)$ and $(01111\cdots)$.

If A is irreducible then σ_A is TT.

**** Proof:** To show TT, construct a dense orbit. As in the proof that the shift map is TT, construct a symbol sequence \mathbf{a} by writing a list of all allowed symbol sequences of length 1, then of length 2, and so on, then appending these sequences into a single infinite sequence. If joining two sequences results in a symbol pair (i, j) that is not allowed, insert a transition sequence \mathbf{a}_{ij} such that $i\mathbf{a}_{ij}j$ is allowed. Such a sequence is guaranteed to exist by the irreducibility of A . \square **]

If A is irreducible and non-trivial then σ_A has SDIC.

**** Proof:** Given a sequence $\mathbf{a} = a_0a_1\cdots \in \Sigma_{N,A}$ and $\varepsilon > 0$ choose $M \geq 0$ such that $\frac{3}{2}3^{-(M+1)} < \varepsilon$. Construct $\mathbf{b} = a_0a_1\cdots a_M b_{M+1}b_{M+2}\cdots b_K$ where $a_M b_{M+1}\cdots b_K$ is an allowed sequence from a_M to $b_K = i$ (this exists by irreducibility). Let $ij_1 ij_2$ with $j_1 \neq j_2$ be allowed sequences (these exist by non-triviality). Let $b_{K+1} = j_1$ unless $a_{K+1} = j_1$ in which case set $b_{K+1} = j_2$. Then \mathbf{b} differs from \mathbf{a} either at some place $M+1 \leq n < K+1$ or at place $n = K+1$. So, by Proposition 9.1, $d(\mathbf{a}, \mathbf{b}) \leq \frac{3}{2}3^{-(M+1)} < \varepsilon$ and $d(\sigma_A^n(\mathbf{a}), \sigma_A^n(\mathbf{b})) \geq 1$. So σ_A has SDIC. \square **]

FROM CONTINUOUS MAPS, $F : I \rightarrow \mathbb{R}$, TO SYMBOLIC MAPS

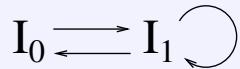
Here we use results for the abstract SSFT dynamical systems to prove results about maps on the real line. First:

1. Let $\{I_i : i = 0, \dots, N - 1\}$ be a collection of disjoint closed bounded intervals in \mathbb{R} and $I = \bigcup_{i=0}^{N-1} I_i$.
2. Let $F : I \rightarrow \mathbb{R}$ be a continuous map of the real line.
3. Let $\Lambda = \{x \in I : F^n(x) \in I \forall n \geq 0\}$ be the invariant set for F .

Definition 13.1 (F -covering). *Let I_i and I_j be intervals in \mathbb{R} . I_i F -covers I_j ($I_i \rightarrow I_j$) if $F(I_i) \supseteq I_j$.*

Definition 13.2 (Transition graph). *A transition graph Γ is a graph with N vertices, each corresponding to an interval I_i , and directed edges between these vertices indicating the F -covering relations.*

Example 13.1. Suppose there are two intervals I_0 and I_1 with I_0 F -covering I_1 and I_1 F -covering I_0 and I_1 . Then the transition graph is shown as follows.



Given the above transition graph, we can attempt to sketch a map and interval pair with the F -coverings. An example of such a graph is given in Figure 13.1.

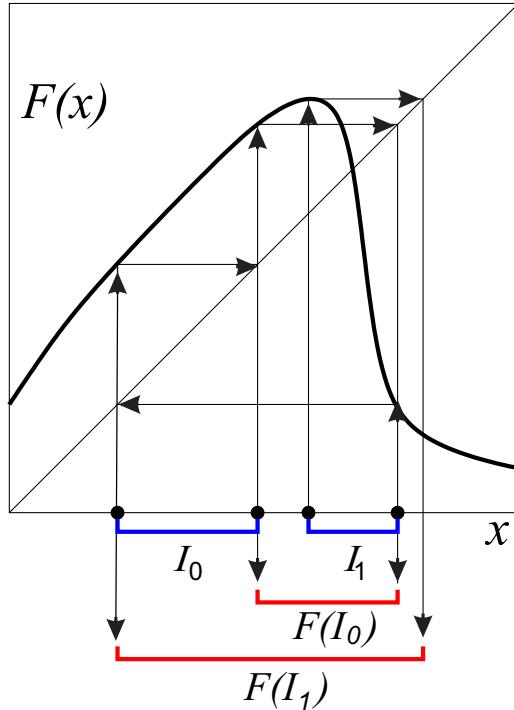
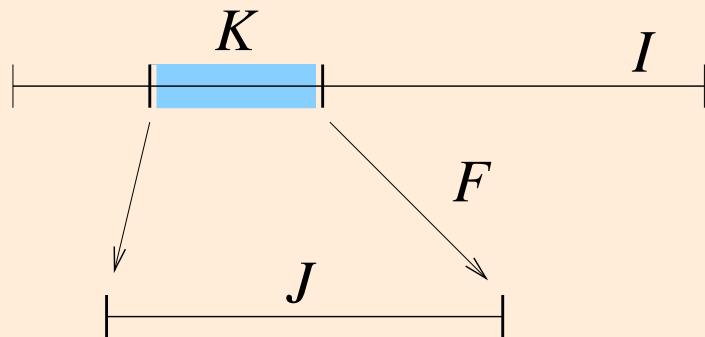


Figure 13.1: An example of a map and interval pair with a specified set of F -coverings.

Remark 13.1. *F -covering relations can be summarised in a transition matrix A defined by $A_{ij} = 1$ if $I_i \rightarrow I_j$ and $A_{ij} = 0$ if not.*

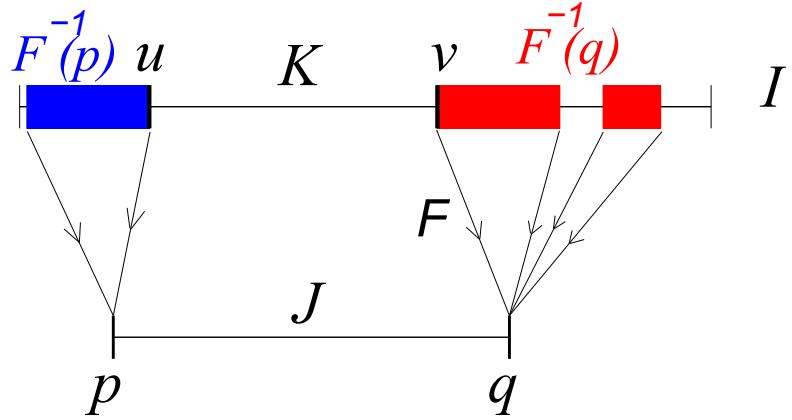
If $I_0 \xrightarrow{F} I_1 \xrightarrow{F} I_2$ then $I_0 \xrightarrow{F^2} I_2$ by continuity of F . This construction is sometimes referred to as the ‘induced graph’ of F^2 -covering relations.

Lemma 13.1. *Let I, J be closed intervals. If I F -covers J then there exists a closed subinterval $K \subseteq I$ such that $F(K) = J$. i.e.*



A map $F : I \rightarrow \mathbb{R}$ F -covers an interval J , and there is a closed interval K such that $F(K) = J$.

Proof. Let $J = [p, q]$ be a closed interval. Then $F^{-1}(p)$ and $F^{-1}(q)$ are closed and non-empty. If the picture looks like this,



Proof: Set $u = \sup\{F^{-1}(p)\}$ (u is the largest $x \in I$ such that $F(x) = p$) and set $v = \inf\{F^{-1}(q)\}$ (v is the smallest $x \in I$ such that $F(x) = q$). If all of $F^{-1}(q)$ lies to the left of all of $F^{-1}(p)$; take $u = \sup\{F^{-1}(q)\}$ and $v = \inf\{F^{-1}(p)\}$ instead. Set $K = [u, v]$, a closed subinterval of I . Now, for any $s \in J = [p, q]$ let $g(x) = F(x) - s$. Then

$$\begin{aligned} g(u) &= F(u) - s = p - s \leq 0 \\ g(v) &= F(v) - s = q - s \geq 0 \end{aligned}$$

and g is continuous. Hence, by the Intermediate Value Theorem, there exists $\hat{x} \in K$ such that $g(\hat{x}) = 0$ i.e. $F(\hat{x}) = s$, so every $s \in J$ has a preimage in K and $F(K) = J$. \square

Theorem 13.1. Let $\{I_i : i = 0, \dots, N-1\}$ be a collection of disjoint closed bounded intervals in \mathbb{R} and $I = \bigcup_{i=0}^{N-1} I_i$. Let $F : I \rightarrow \mathbb{R}$ be a continuous map of the real line. Let $\Lambda = \{x \in I : F^n(x) \in I \forall n \geq 0\}$ be the invariant set for F . Let A be the transition matrix of F -covering relationships for the intervals I_i . Then $F|_\Lambda$ is semiconjugate to σ_A .

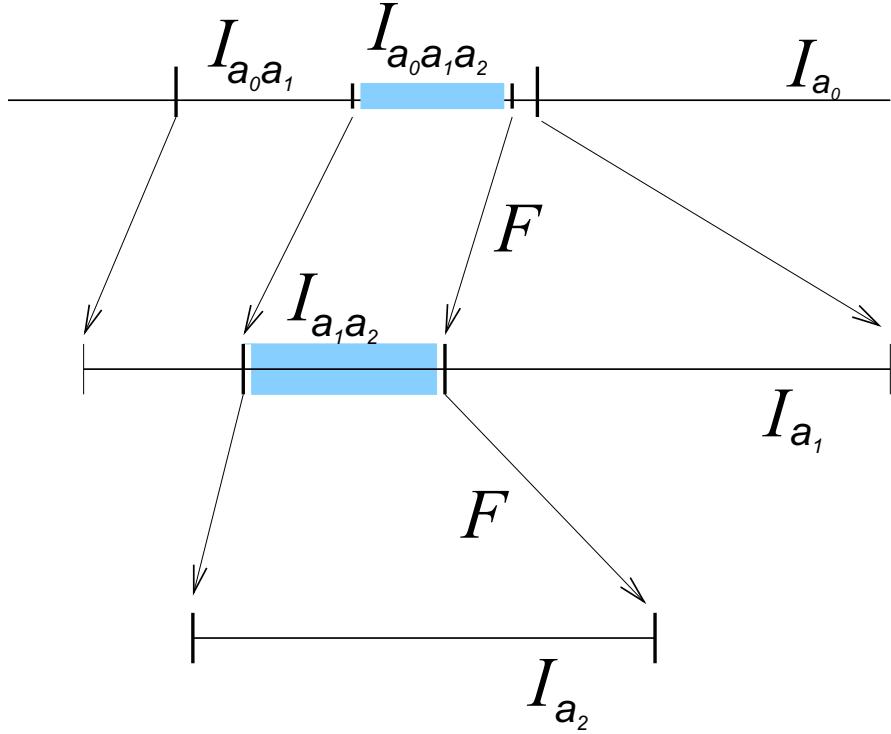
Proof: Define $h : \Lambda \rightarrow \Sigma_{N,A}$ by $x \mapsto \mathbf{a}$ where \mathbf{a} is the symbol sequence of $\{F^n(x)\}_{n \geq 0}$ such that $F^n(x) \in I_{a_n}$ for all $n \geq 0$. Then (i) h is continuous and (ii) h is surjective and hence h is a semiconjugacy.

(i) h is continuous. Given given $\varepsilon > 0$ choose M such that $\frac{3}{2}3^{-M+1} < \varepsilon$. By continuity of F , there exists δ such that $|x - y| < \delta$ implies $F^n(x), F^n(y) \in I_{a_n}$ for $0 \leq n \leq M$, i.e. x and y remain in the same intervals as each other for (at least) the first M iterates. Hence, by Proposition 9.1, $d(h(x), h(y)) < \frac{3}{2}3^{-M+1} = \varepsilon$.

(ii) h is surjective. Let $\mathbf{a} = a_0a_1a_2 \dots \in \Sigma_{N,A}$ be an allowed sequence. So, by definition of A , if the symbol pair $a_i a_j$ appears in \mathbf{a} then I_{a_i} F -covers I_{a_j} . Hence I_{a_0} F -covers I_{a_1} . It follows by Lemma 13.1 that there exists $I_{a_0a_1} \subseteq I_{a_0}$ such that $F(I_{a_0a_1}) = I_{a_1}$. Similarly, I_{a_1} F -covers I_{a_2} so there exists $I_{a_1a_2} \subseteq I_{a_1}$ such that $F(I_{a_1a_2}) = I_{a_2}$. Now, $I_{a_1a_2} \subseteq I_{a_1}$ and $I_{a_0a_1}$ F -covers I_{a_1} . So, $I_{a_0a_1}$ F -covers $I_{a_1a_2}$. So, by Lemma 13.1 again, there exists $I_{a_0a_1a_2} \subseteq I_{a_0a_1} \subseteq I_{a_0}$ such that $F(I_{a_0a_1a_2}) = I_{a_1a_2}$ and $F^2(I_{a_0a_1a_2}) = F(I_{a_1a_2}) = I_{a_2}$. Continuing this construction inductively generates a nested sequence of M closed intervals

$$I_{a_0} \supseteq I_{a_0a_1} \supseteq I_{a_0a_1a_2} \supseteq \dots \supseteq I_{a_0a_1\dots a_M}$$

such that $x \in I_{a_0a_1\dots a_M}$ implies $F^n(x) \in I_{a_n}$ for all $0 \leq n \leq M$. This is illustrated in the figure below:



Nested intervals $I_{a_0 a_1 a_2} \subseteq I_{a_0 a_1} \subseteq I_{a_0}$ and their images under F .

Now we apply the Cantor Intersection Theorem to assert that

$$S_{\mathbf{a}} := \cap_{M=0}^{\infty} I_{a_0 a_1 \dots a_M}$$

is non-empty, so there exists a point $x \in \Lambda$ such that $h(x) = \mathbf{a}$, i.e. h is surjective. \square

Remark 13.2.

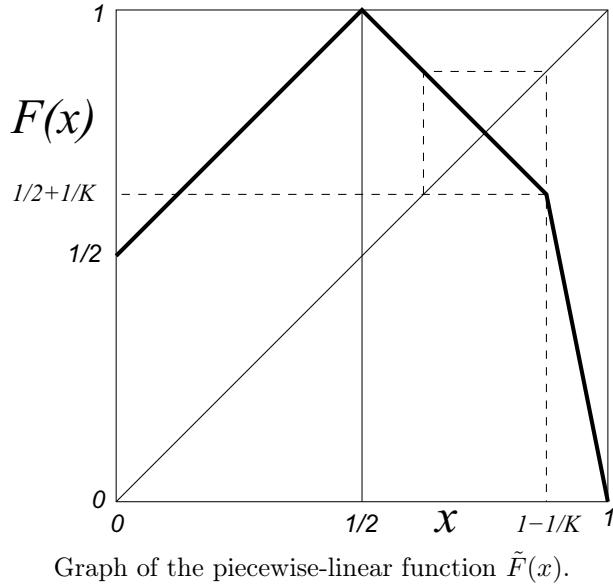
1. The iterates of x must lie in exactly the sequence of intervals $\{I_{a_n}\}$.
2. A crucial difference from the conjugacy between the logistic map and the binary sequence space (Theorem 11.1) is that it might not be the case that $\text{diam}(S_{\mathbf{a}}) \rightarrow 0$ for all \mathbf{a} . So that it might happen that $S_{\mathbf{a}}$ contains a whole interval of points x in which case h could not be a conjugacy.
3. $F|_{\Lambda}$ semiconjugate to σ_A implies that $F|_{\Lambda}$ has a periodic orbit $\{x_0, x_1, \dots, x_{N-1}\}$ corresponding to every periodic symbol sequence $(a_0 a_1, \dots, a_{N-1} a_0 a_1 \dots)$ in $\Sigma_{N,A}$. So $F|_{\Lambda}$ has at least as many period orbits as $\Sigma_{N,A}$. Moreover $x_j \in I_{a_j}$ for all $0 \leq j < N$ and hence this periodic orbit corresponds exactly to a closed path $I_{a_0} \rightarrow I_{a_1} \rightarrow \dots \rightarrow I_{a_{N-1}} \rightarrow I_{a_0}$ in the graph Γ .
4. $F|_{\Lambda}$ semiconjugate to σ_A with A irreducible and nontrivial does not imply that $F|_{\Lambda}$ has SDIC or TT. This is because F could have intervals on which it is non-expanding, see the example below. If, however, we have additional information that shows F is suitably ‘expanding’ then we can show, in a similar fashion to the proof of Theorem 11.1, that h is a conjugacy which would imply that $F|_{\Lambda}$ has SDIC and is TT.

[** Example: A map $F(x)$ for which $\Lambda = [0, 1]$ is an invariant set, and for which the map $h : \Lambda \rightarrow \Sigma_2$ is surjective but which does not have SDIC or TT.

Define $\tilde{F}(x)$ to be the piecewise-linear map

$$\tilde{F}(x) = \begin{cases} x + \frac{1}{2} & \text{if } 0 \leq x \leq \frac{1}{2}, \\ \frac{3}{2} - x & \text{if } \frac{1}{2} \leq x \leq 1 - \frac{1}{K}, \\ \left(\frac{K}{2} + 1\right)(1-x) & \text{if } 1 - \frac{1}{K} \leq x \leq 1. \end{cases}$$

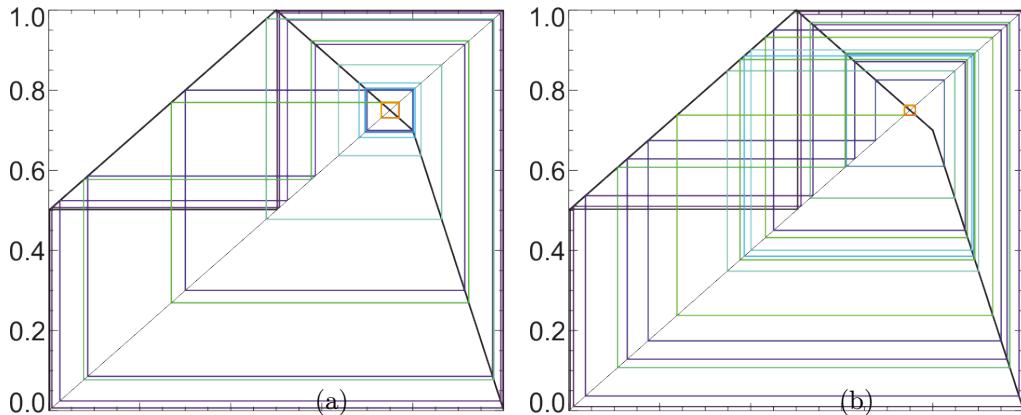
Fix $K > 4$, e.g. for concreteness $K = 5$. The graph of \tilde{F} is as in the figure below:



Graph of the piecewise-linear function $\tilde{F}(x)$.

Then, all points in the interval $\frac{1}{2} + \frac{1}{K} \leq x \leq 1 - \frac{1}{K}$ lie on 2-cycles (except $x = \frac{3}{4}$ which is clearly a fixed point!), so these points do not move apart from, or towards, each other (so $\tilde{F}|_{\Lambda}$ does not have SDIC) and the 2-cycles do not ‘mix’ under iteration, so there is no TT in this part of $[0, 1]$ and hence $\tilde{F}|_{\Lambda}$ is not TT.

It turns out that the preimages of the interval $\frac{1}{2} + \frac{1}{K} \leq x \leq 1 - \frac{1}{K}$ are dense in $[0, 1]$, so almost all initial conditions end up on a 2-cycle, as the following figure shows:



The first 48 (or so) iterates of the map \tilde{F} , taking $K=5$, starting from (a) $x_0 = 0.002$ and (b) $x_0 = 0.003$. The colour of the iterates x_n darkens as n increases. Eventually, in both cases, the iterates settle to a 2-cycle (shown in red) near $x = \frac{3}{4}$.

Note, however, that there exists a smaller invariant subset $\tilde{\Lambda} \subset \Lambda$ on which F does have SDIC and TT (remove all the preimages of the set of nonhyperbolic 2-cycles from $[0, 1]$ and we are left with a Cantor set of N -cycles and their preimages). Note also that h is not continuous here so the map is not a semiconjugacy. This arises from the fact that the intervals $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$ are not disjoint as we assumed the collection $\{I_i\}$ was at the start of this section. **]

HORSESHOES, N -CYCLES, AND CHAOS

At last in this chapter we will offer a definition of chaos. The definition we use in this course is topological in nature and uses the notions of F -coverings that we previously learned about.

Now in essence, we want a definition that involves the following two ideas:

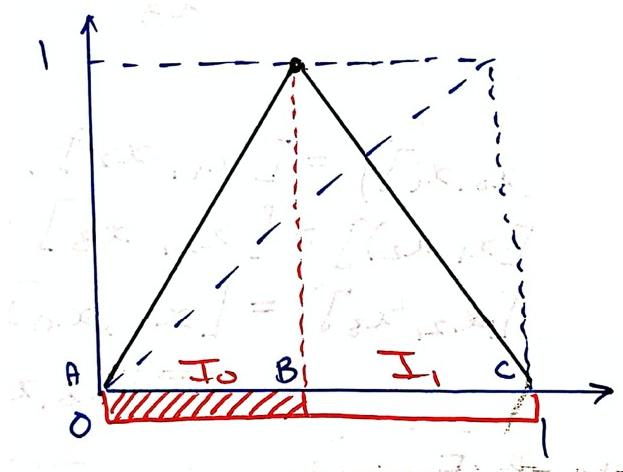
1. Stretching
2. Folding or mixing

You may think of stretching as a characteristic that accompanies SDIC (sensitive dependence on initial conditions) since this behaviour is associated with points evolving away from one another. On the other hand, folding or mixing is a characteristic that is associated with TT (topological transitivity).

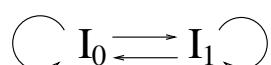
Why are those two ideas essential for the sort of maps we have studied previously? Consider the case of the full-height tent map, which is given by

$$F(x) = \begin{cases} 2x & x \in [0, 1/2] \\ 2(1-x) & x \in [1/2, 1] \end{cases}$$

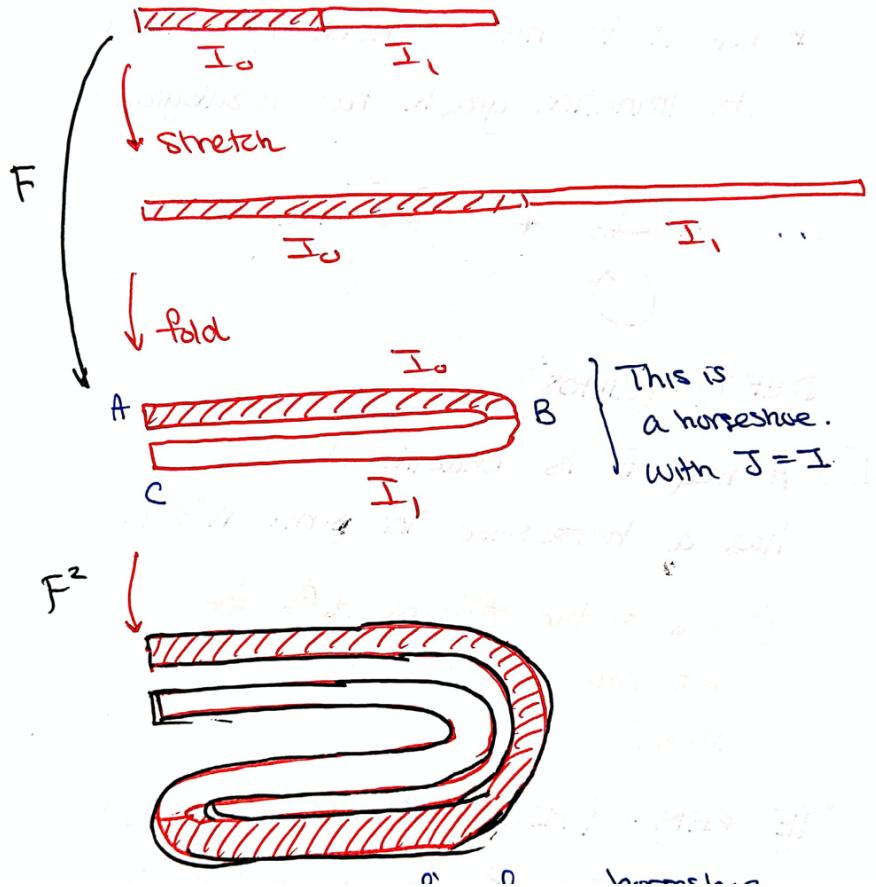
The tent map looks like this:



We have shaded in the interval I_0 and left I_1 unshaded. Notice that the map exhibits the following covering structure:



We can now imagine the operation that F performs on the two intervals, which is composed of a two-times stretch and then a fold about $x = 1/2$.



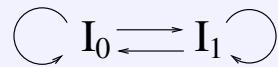
In the above visualisation, the final image of F and F^2 should be ‘squeezed’ onto the real line.

So this inspires the following definition.

Definition 14.1 (Horseshoe). $F : I \rightarrow \mathbb{R}$ has a horseshoe if there exists a closed interval $J \subseteq I$ and closed subintervals $I_0, I_1 \subset J$ with disjoint interiors such that $F(I_0) = F(I_1) = J$.

Remark: In this course, it is useful to use an equivalent definition that does not require the intermediary interval J . We will find it more useful to use this one:

Definition 14.2 (Horseshoe (alternative)). $F : I \rightarrow \mathbb{R}$ has a horseshoe if there exists closed subintervals $I_0, I_1 \subset I$ with disjoint interiors such that $F(I_0) \supseteq I_0 \cup I_1$ and $F(I_1) \supseteq I_0 \cup I_1$. Hence the transition graph for F contains the subgraph:



Hence we now define chaos as the following.

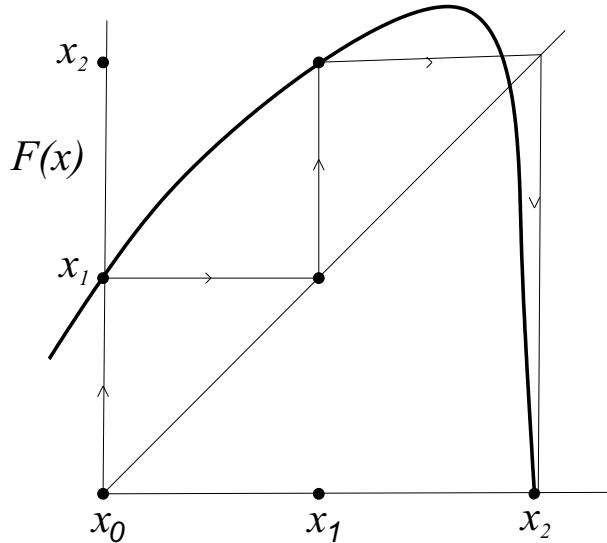
Definition 14.3 (chaotic). A map F is chaotic if F^n has a horseshoe for some $n \geq 1$.

Remark 14.1. This definition of chaos is based on a fundamental topological property. There are many other definitions of chaos. Most commonly, a map F is defined to be chaotic if it has the properties of SDIC and TT. In the majority of cases a map will be chaotic under both definitions. However, there are some examples of maps with SDIC and TT that are not chaotic according to the horseshoe definition.

Theorem 14.1. Let $F : I \rightarrow \mathbb{R}$ be a continuous map on a closed bounded interval $I \subset \mathbb{R}$. If F has a 3-cycle then

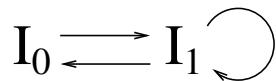
- F is chaotic (F^2 has a horseshoe)
- F has an N -cycle for all $N \geq 1$.

Proof: (F is chaotic). Let $x_0 < x_1 < x_2$ be the 3-cycle. Assume $F(x_0) = x_1, F(x_1) = x_2, F(x_2) = x_0$. The only alternative is $F(x_0) = x_2, F(x_2) = x_1, F(x_1) = x_0$. If this is the case consider $G(x) = -F(-x)$. G is conjugate to F and has a 3-cycle of the form assumed for F above.

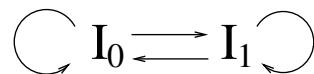


An example of a continuous map F that has a 3-cycle.

Let $I_0 = [x_0, x_1]$ and $I_1 = [x_1, x_2]$. Since $F(x_0) = x_1, F(x_1) = x_2$ and F is continuous, $F(I_0) \supseteq I_1$. Similarly, $F(x_1) = x_2, F(x_2) = x_0$ means that $F(I_1) \supseteq I_0 \cup I_1$. So there are (at least) the F -covering relations indicated by the transition graph:



There are (at least) the F^2 -covering relations indicated by the graph



So F^2 has a horseshoe and F is chaotic. □

Proof: (F has an N -cycle for all $N \geq 1$). Let A be the transition matrix summarising the F -covering relations for I_0 and I_1 . By Theorem 13.1, $F|_{\Lambda}$ is semiconjugate to $\sigma_A : \Sigma_{2,A} \rightarrow \Sigma_{2,A}$. From the transition graph, the symbol sequence

$$\mathbf{a}^{(N)} = \underbrace{01111 \cdots 1}_{\text{ }}, \underbrace{01111 \cdots 1}_{\text{ }}, \underbrace{01111 \cdots 1}_{\text{ }}, \dots$$

is allowed for any $N \geq 2$. So there exists a point $x^{(N)}$ which corresponds to the symbol sequence $\mathbf{a}^{(N)}$ and which lies on an N -cycle, for any $N \geq 2$. Since I_0 is visited only once per period of the period- N symbol sequence, the orbit must have least period N .

For $N = 1$ the allowed sequence $\mathbf{a}^{(1)} = 111 \cdots$ implies the existence of a fixed point in I_1 . \square

This result is associated with Li & Yorke. But it is actually only a special case of a more general result due to Sharkovsky which was proved earlier (around 1964).

Theorem 14.2 (Sharkovsky's Theorem). *Consider the ordering of the integers defined by*

$$\begin{aligned} 1 &\triangleleft 2 \triangleleft 4 \triangleleft 2^3 \triangleleft \cdots \triangleleft 2^n \triangleleft 2^{n+1} \triangleleft \cdots \\ \cdots &\triangleleft 2^{n+1}.9 \triangleleft 2^{n+1}.7 \triangleleft 2^{n+1}.5 \triangleleft 2^{n+1}.3 \triangleleft \cdots \\ \cdots &\triangleleft 2^n.9 \triangleleft 2^n.7 \triangleleft 2^n.5 \triangleleft 2^n.3 \triangleleft \cdots \\ \cdots &\triangleleft 9 \triangleleft 7 \triangleleft 5 \triangleleft 3. \end{aligned}$$

Let $F : I \rightarrow \mathbb{R}$ be a continuous map of the interval. If F has an N -cycle then F has a k -cycle for all $k \triangleleft N$ in the above ordering.

Remark 14.2. *Sharkovsky's theorem implies:*

1. *Existence of a 3-cycle implies the existence of an N -cycle for all $N \geq 1$ (i.e. the Li & Yorke result).*
2. *Existence of a 4-cycle implies the existence of a 1-cycle (fixed point) and a 2-cycle but maybe no more (see problem sheet).*
3. *If m is an odd integer, the existence of a $2^n.m$ -cycle implies the existence of a $2^{n+1}.3$ -cycle which implies that $F^{2^{n+1}}$ has a 3-cycle, which implies F is chaotic.*

THE TENT MAP

The tent map is the piecewise-linear continuous map $x_{n+1} = F(x_n)$ of the unit interval defined by

$$F(x) = \begin{cases} F_L(x) \equiv \mu x & \text{if } 0 \leq x \leq \frac{1}{2} \\ F_R(x) \equiv \mu(1-x) & \text{if } \frac{1}{2} \leq x \leq 1. \end{cases}$$

Chaotic orbits of the tent map

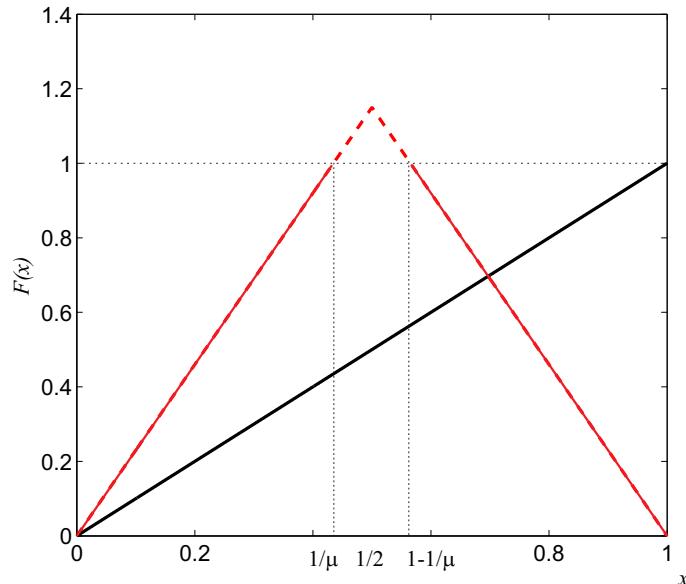


Figure 15.1: Tent map for $\mu = 2.3$.

- If $\mu \geq 2$ then F has a horseshoe and so is chaotic. Let $I_0 = [0, 1/\mu]$, $I_1 = [1 - 1/\mu, 1]$. Then $F(I_0) = I$ and $F(I_1) = I$.
- If $0 \leq \mu \leq 2$ then F maps $[0, 1]$ into itself. There is always a fixed point at $x = 0$.
- If $0 \leq \mu < 1$ the fixed point at $x = 0$ is stable and attracts all initial conditions; therefore no interesting dynamics is possible for μ in this range.

- If $1 < \mu \leq 2$ there is another fixed point $x_0 = \mu/(1 + \mu) > 1/2$. Both of the fixed points $x = 0$ and $x = x_0$ are unstable ($F'(x_0) = -\mu < -1$). So all orbits are bounded but the fixed points are unstable.

Proposition 15.1. F is chaotic throughout the range $\sqrt{2} < \mu < 2$.

Proof. The objective is to show that F^2 has a horseshoe for $\sqrt{2} < \mu < 2$.

Step 1. Establish bounds on the orbits. All orbits (except $x = 0$) eventually enter and stay in the interval $A := [F^2(\frac{1}{2}), F(\frac{1}{2})] = [\mu(1 - \mu/2), \mu/2]$. This is because $F(1/2) > 1/2$ since $\mu > 1$, $F(x) \leq F(1/2)$ if $0 \leq x \leq 1$; if $1/2 \leq x \leq F(1/2)$ then $F_x < 0$ and so $F(x) > F^2(1/2)$; if $F(1/2) \leq x \leq 1$ then $0 \leq F(x) \leq F^2(1/2) < 1/2$; if $x \leq 1/2$ then $F_x > 0$ so $F^n(x) > F^2(1/2)$ for some $n \geq 0$.

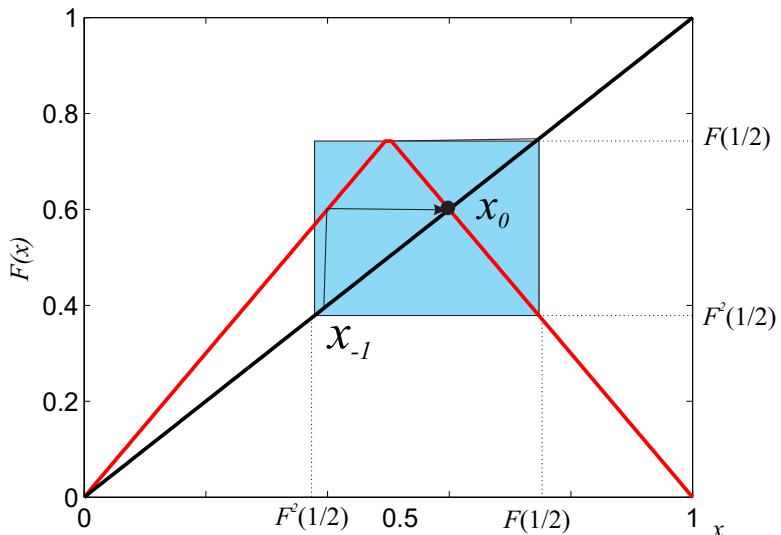


Figure 15.2: Tent map for $\mu = 1.5$ showing attracting region A

Remark 15.1. As points iterate around within A , they can come arbitrarily close to the fixed point x_0 only if the preimage x_{-1} of x_0 (i.e. the point $x_{-1} \neq x_0$ that satisfies $F(x_{-1}) = x_0$) lies in A . This happens when $1/(1 + \mu) > \mu(1 - \mu/2)$, i.e. $\mu > \sqrt{2}$. For $\mu < \sqrt{2}$ orbits remain in A but do not come close to x_0 .

Step 2. Construct $F^2(x)$. This can be done explicitly from F . The form of F^2 depends on the location of x and $F(x)$. These lie either to the left or to the right of $1/2$, and all possibilities need to be considered. So

$$F^2(x) = \begin{cases} F_L \circ F_L(x) & \text{if } 0 \leq x \leq \frac{1}{2} \text{ and } 0 \leq F_L(x) \leq \frac{1}{2} \\ F_R \circ F_L(x) & \text{if } 0 \leq x \leq \frac{1}{2} \text{ and } \frac{1}{2} \leq F_L(x) \leq 1 \\ F_R \circ F_R(x) & \text{if } \frac{1}{2} \leq x \leq 1 \text{ and } \frac{1}{2} \leq F_R(x) \leq 1 \\ F_L \circ F_R(x) & \text{if } \frac{1}{2} \leq x \leq 1 \text{ and } 0 \leq F_R(x) \leq \frac{1}{2} \end{cases}$$

which yields F^2 as a piecewise-affine map in four sections:

$$F^2(x) = \begin{cases} \mu^2 x & \text{if } 0 \leq x \leq \frac{1}{2} \text{ and } 0 \leq \mu x \leq \frac{1}{2} \\ \mu[1 - \mu x] & \text{if } 0 \leq x \leq \frac{1}{2} \text{ and } \frac{1}{2} \leq \mu x \leq 1 \\ \mu[1 - \mu(1 - x)] & \text{if } \frac{1}{2} \leq x \leq 1 \text{ and } \frac{1}{2} \leq \mu(1 - x) \leq 1 \\ \mu^2(1 - x) & \text{if } \frac{1}{2} \leq x \leq 1 \text{ and } 0 \leq \mu(1 - x) \leq \frac{1}{2}. \end{cases}$$

The inequalities that accompany each of these expressions can be carefully considered to see which provide the strongest constraints given $1 < \mu \leq 2$. So

$$F^2(x) = \begin{cases} \mu^2 x & \text{if } 0 \leq x \leq \frac{1}{2\mu} \\ \mu[1 - \mu x] & \text{if } \frac{1}{2\mu} \leq x \leq \frac{1}{2} \\ \mu[1 - \mu(1 - x)] & \text{if } \frac{1}{2} \leq x \leq 1 - \frac{1}{2\mu} \\ \mu^2(1 - x) & \text{if } 1 - \frac{1}{2\mu} \leq x \leq 1. \end{cases}$$

Remark 15.2. The graph of F^2 is sketched for two values of μ in the figures below. Note that the peaks in the graph of F^2 have the same height as that in F since $F(1/(2\mu)) = F(1 - 1/(2\mu)) = 1/2$ implies $F^2(1/(2\mu)) = F^2(1 - 1/(2\mu)) = F(1/2)$ which is the value at the peak.

The pre-images x_{-1} and x_{-2} of x_0 are defined by $F^2(x_{-2}) = F(x_{-1}) = x_0$, and the requirement that x_{-2} is the preimage of x_{-1} that lies in $x > 1/2$. We obtain the explicit values $x_{-1} = 1/(1 + \mu)$ and $x_{-2} = \frac{\mu^2 + \mu - 1}{\mu(\mu + 1)}$.

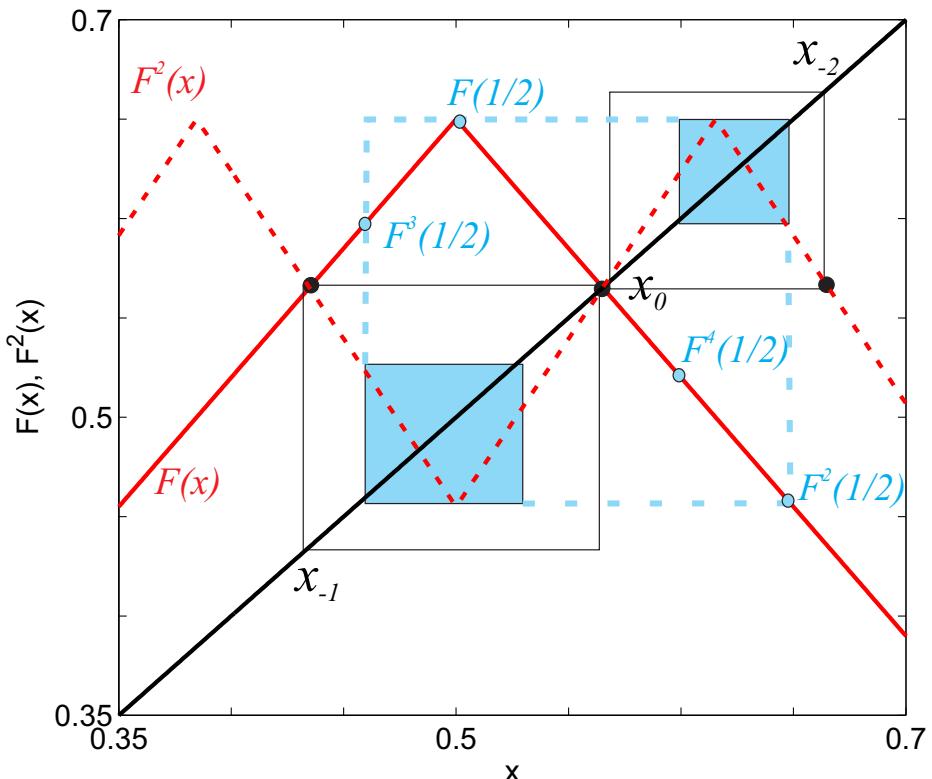


Figure 15.3: For $\mu = 1.3 < \sqrt{2}$, F^2 gives tent maps with parameter μ^2 on the intervals $[x_{-1}, x_0]$ and $[x_0, x_{-2}]$, and the attracting set (shaded) has two components.

Step 3. Look for horseshoes in F^2 . The importance of x_{-1} and x_{-2} is indicated by the figures. F^2 acts as a tent map when restricted to the intervals $I_L := [x_{-1}, x_0]$ or $I_R := [x_0, x_{-2}]$. The only difference to the original tent map is the gradient parameter which has been scaled from μ to μ^2 .

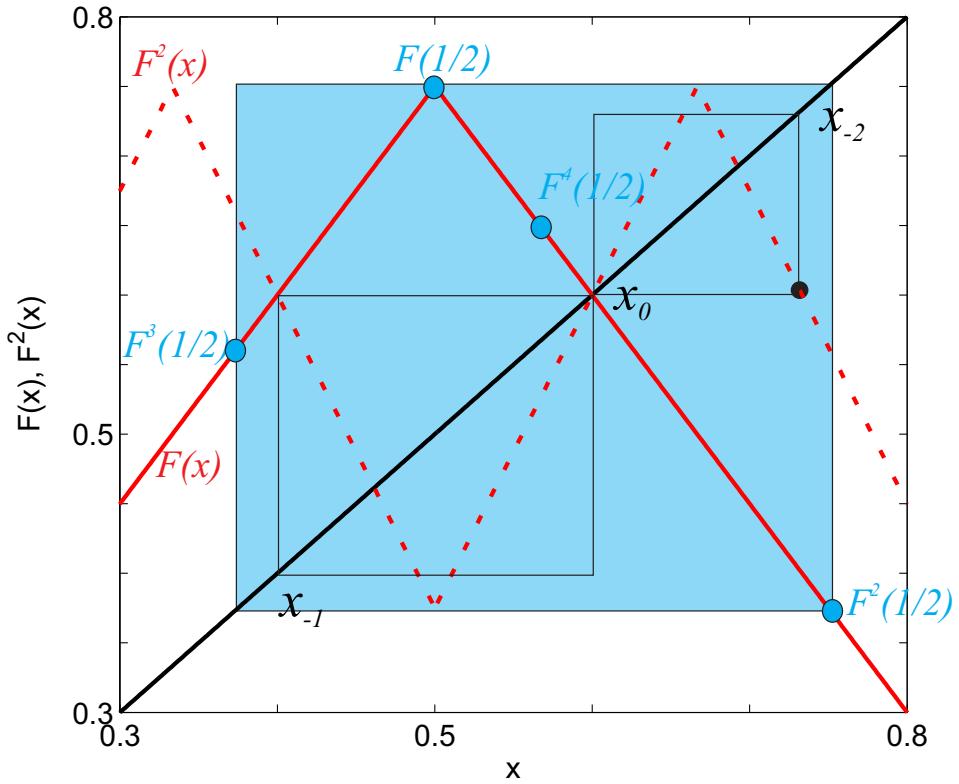


Figure 15.4: For $\mu = 1.5 > \sqrt{2}$, F^2 gives horseshoes on the intervals $[x_{-1}, x_0]$ and $[x_0, x_{-2}]$, and the attracting set (shaded) has one component.

The two figures, for $\mu = 1.3$ and $\mu = 1.5$ indicate the two possibilities for the dynamics. When $1 < \mu < \sqrt{2}$, F^2 maps both I_L and I_R into themselves, respectively. So F^2 does not have a horseshoe and there are attracting subintervals $A_L := [F^2(1/2), F^4(1/2)] \subset I_L$ and $A_R := [F^3(1/2), F(1/2)] \subset I_R$. These attracting subintervals satisfy $F(A_R) = A_L$, $F(A_L) \subseteq A_R$ but $x_0 \notin A_L \cup A_R$: there is an interval of points around x_0 that are not reached by iterates. Therefore the attracting set, that iterates eventually settle close to, must be contained in $A_L \cup A_R$, and it must consist of (at least) two pieces.

When $\mu \geq \sqrt{2}$

$$F^2\left(\frac{1}{2}\right) = \mu\left(1 - \frac{\mu}{2}\right) < x_{-1} = \frac{1}{1+\mu},$$

and

$$F\left(\frac{1}{2}\right) = \frac{\mu}{2} > x_{-2} = \frac{\mu^2 + \mu - 1}{\mu(\mu + 1)}.$$

Hence F^2 has horseshoes on both I_L and I_R , which implies that F is chaotic. Moreover, the attracting set is the whole of A since $x_0 \in A$ (there is no interval of points around x_0 which iterates cannot reach).

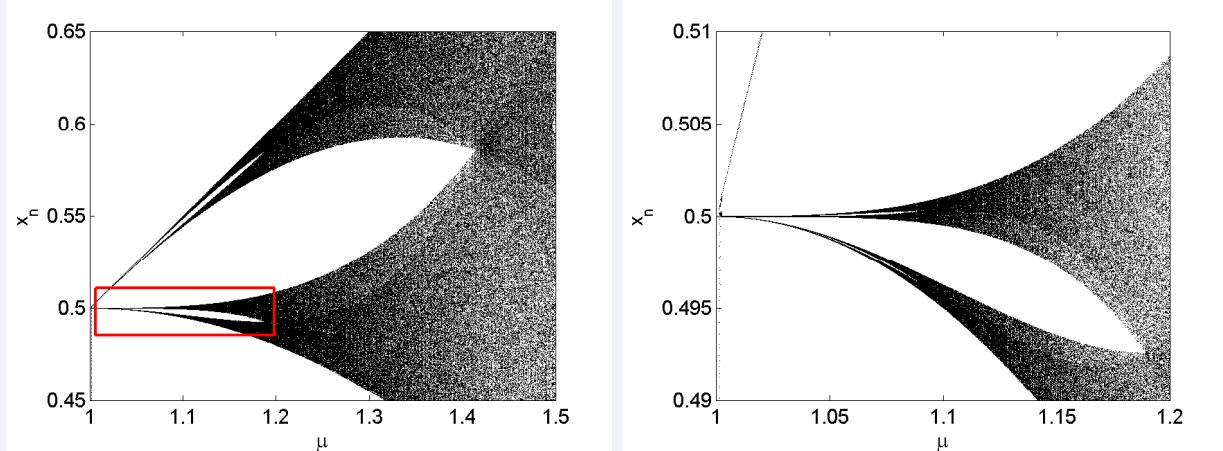
Theorem 15.1. Let $n \geq 0$ be an integer. If $2^{1/2^{n+1}} \leq \mu < 2^{1/2^n}$ then $F^{2^{n+1}}$ has a horseshoe (and so F is chaotic) and the chaotic attractor consists of 2^n intervals that are permuted by F .

Remark 15.3. Proposition 15.1 is a special case of the theorem with $n = 0$, showing that F^2 has a horseshoe for $2^{1/2} \leq \mu < 2$ on which the attractor is a single interval.

The theorem can be proved by an inductive argument based on the proof of Proposition 15.1.

Since the lower limits $2^{1/2^{n+1}}$ tend to 1 from above as $n \rightarrow \infty$, the theorem implies that F is chaotic for all μ in the range $1 < \mu \leq 2$.

The following images of the (μ, x) plane were produced by iterating the tent map for 1000 iterations to allow the orbit to settle close to the chaotic attractor, and then plotting the next 500 points. Observe that the attracting set contains 2^n intervals in $2^{1/2^{n+1}} \leq \mu < 2^{1/2^n}$. The image of the right is an enlargement of the image on the left.



These images suggest that the Tent Map is also chaotic under the definition that requires it to display SDIC and TT. However, direct proofs of these properties are very fiddly because the invariant set Λ is composed of many pieces.

CHAPTER 16

LORENZ PAPERS

Extracts of two of Lorenz's papers (one 'expository' one in 1972 and a famous research article in 1963) are reproduced here.

CLASSICS

Predictability: Does the Flap of a Butterfly's Wings in Brazil Set off a Tornado in Texas?

*Edward U. Lorenz, Sc.D. Professor of Meteorology
Massachusetts Institute of Technology, Cambridge*

Lest I appear frivolous in even posing the title question, let alone suggesting that it might have an affirmative answer, let me try to place it in proper perspective by offering two propositions.

1. If a single flap of a butterfly's wings can be instrumental in generating a tornado, so also can all the previous and subsequent flaps of its wings, as can the flaps of the wings of millions of other butterflies, not to mention the activities of innumerable more powerful creatures, including our own species.
2. If the flap of a butterfly's wings can be instrumental in generating a tornado, it can equally well be instrumental in preventing a tornado.

More generally, I am proposing that over the years minuscule disturbances neither increase nor decrease the frequency of occurrence of various weather events such as tornados; the most that they may do is to modify the sequence in which these events occur. The question which really interests us is whether they can do even this – whether, for example, two particular weather situations differing by as little as the immediate influence of a single butterfly will generally after sufficient time evolve into two situations differing by as much as the presence of a tornado. In more technical language, is the behavior of the atmosphere *unstable* with respect to perturbations of small amplitude?

The connection between this question and our ability to predict the weather is evident. Since we do not know exactly how many butterflies there are, nor where they are all located, let alone which ones are flapping their wings at any instant, we cannot, if the answer to our question is affirmative, accurately predict the occurrence of tornados at a sufficiently distant future time. More significantly, our general failure to detect systems even as large as thunderstorms when they slip between weather stations may impair our ability to predict the general weather pattern even in the near future.

How can we determine whether the atmosphere is unstable? The atmosphere is not a



CLASSICS

controlled laboratory experiment; if we disturb it and then observe what happens, we shall never know what would have happened if we had not disturbed it. Any claim that we can learn what would have happened by referring to the weather forecast would imply that the question whose answer we seek has already been answered in the negative.

The bulk of our conclusions are based upon computer simulation of the atmosphere. The equations to be solved represent our best attempts to approximate the equations actually governing the atmosphere by equations which are compatible with present computer capabilities. Generally two numerical solutions are compared. One of these is taken to simulate the actual weather, while the other simulates the weather which would have evolved from slightly different initial conditions, i.e., the weather which would have been predicted with a perfect forecasting technique but imperfect observations. The difference between the solutions therefore simulates the error in forecasting. New simulations are continually being performed as more powerful computers and improved knowledge of atmospheric dynamics become available.

Although we cannot claim to have proven that the atmosphere is unstable, the evidence that it is so is overwhelming. The most significant results are the following.

1. Small errors in the coarser structure of the weather pattern – those features which are readily resolved by conventional observing networks – tend to double in about three days. As the errors become larger the growth-rate subsides. This limitation alone would allow us to extend the range of acceptable prediction by three days every time we cut the observation error in half, and would offer the hope of eventually making good forecasts several weeks in advance.
2. Small errors in the finer structure – e.g., the positions of individual clouds – tend to grow much more rapidly, doubling in hours or less. This limitation alone would not seriously reduce our hopes for extended-range forecasting, since ordinarily we do not forecast the finer structure at all.
3. Errors in the finer structure, having attained appreciable size, tend to induce errors in the coarser structure. This result, which is less firmly established than the previous ones, implies that after a day or so there will be appreciable errors in the coarser structure, which will thereafter grow just as if they had been present initially. Cutting the observation error in the finer structure in half – a formidable task – would extend the range of acceptable prediction of even the coarser structure only by hours or less. The hopes for predicting two weeks or more in advance are thus greatly diminished.



CLASSICS

4. Certain special quantities such as weekly average temperatures and weekly total rainfall may be predictable at a range at which entire weather patterns are not.

Regardless of what any theoretical study may imply, conclusive proof that good day-to-day forecasts can be made at a range of two weeks or more would be afforded by any valid demonstration that any particular forecasting scheme generally yields good results at that range. To the best of our knowledge, no such demonstration has ever been offered. Of course, even pure guesses will be correct a certain percentage of the time.

Returning now to the question as originally posed, we notice some additional points not yet considered. First of all, the influence of a single butterfly is not only a fine detail – it is confined to a small volume. Some of the numerical methods which seem to be well adapted for examining the intensification of errors are not suitable for studying the dispersion of errors from restricted to unrestricted regions. One hypothesis, unconfirmed, is that the influence of a butterfly's wings will spread in turbulent air, but not in calm air.

A second point is that Brazil and Texas lie in opposite hemispheres. The dynamical properties of the tropical atmosphere differ considerably from those of the atmosphere in temperate and polar latitudes. It is almost as if the tropical atmosphere were a different fluid. It seems entirely possible that an error might be able to spread many thousands of miles within the temperate latitudes of either hemisphere, while yet being unable to cross the equator.

We must therefore leave our original question unanswered for a few more years, even while affirming our faith in the instability of the atmosphere. Meanwhile, today's errors in weather forecasting cannot be blamed entirely nor even primarily upon the finer structure of weather patterns. They arise mainly from our failure to observe even the coarser structure with near completeness, our somewhat incomplete knowledge of the governing physical principles, and the inevitable approximations which must be introduced in formulating these principles as procedures which the human brain or the computer can carry out. These shortcomings cannot be entirely eliminated, but they can be greatly reduced by an expanded observing system and intensive research. It is to the ultimate-purpose of making not exact forecasts but the best forecasts which the atmosphere is willing to have us make that the Global Atmospheric Research Program is dedicated.

Acknowledgement: This work has been supported by the Atmospheric Sciences Section, National Science Foundation, under Grant GA28203X.



Deterministic Nonperiodic Flow¹

EDWARD N. LORENZ

Massachusetts Institute of Technology

(Manuscript received 18 November 1962, in revised form 7 January 1963)

ABSTRACT

Finite systems of deterministic ordinary nonlinear differential equations may be designed to represent forced dissipative hydrodynamic flow. Solutions of these equations can be identified with trajectories in phase space. For those systems with bounded solutions, it is found that nonperiodic solutions are ordinarily unstable with respect to small modifications, so that slightly differing initial states can evolve into considerably different states. Systems with bounded solutions are shown to possess bounded numerical solutions.

A simple system representing cellular convection is solved numerically. All of the solutions are found to be unstable, and almost all of them are nonperiodic.

The feasibility of very-long-range weather prediction is examined in the light of these results.

1. Introduction

Certain hydrodynamical systems exhibit steady-state flow patterns, while others oscillate in a regular periodic fashion. Still others vary in an irregular, seemingly haphazard manner, and, even when observed for long periods of time, do not appear to repeat their previous history.

These modes of behavior may all be observed in the familiar rotating-basin experiments, described by Fultz, *et al.* (1959) and Hide (1958). In these experiments, a cylindrical vessel containing water is rotated about its axis, and is heated near its rim and cooled near its center in a steady symmetrical fashion. Under certain conditions the resulting flow is as symmetric and steady as the heating which gives rise to it. Under different conditions a system of regularly spaced waves develops, and progresses at a uniform speed without changing its shape. Under still different conditions an irregular flow pattern forms, and moves and changes its shape in an irregular nonperiodic manner.

Lack of periodicity is very common in natural systems, and is one of the distinguishing features of turbulent flow. Because instantaneous turbulent flow patterns are so irregular, attention is often confined to the statistics of turbulence, which, in contrast to the details of turbulence, often behave in a regular well-organized manner. The short-range weather forecaster, however, is forced willy-nilly to predict the details of the large-scale turbulent eddies—the cyclones and anticyclones—which continually arrange themselves into new patterns.

Thus there are occasions when more than the statistics of irregular flow are of very real concern.

In this study we shall work with systems of deterministic equations which are idealizations of hydrodynamical systems. We shall be interested principally in nonperiodic solutions, i.e., solutions which never repeat their past history exactly, and where all approximate repetitions are of finite duration. Thus we shall be involved with the ultimate behavior of the solutions, as opposed to the transient behavior associated with arbitrary initial conditions.

A closed hydrodynamical system of finite mass may ostensibly be treated mathematically as a finite collection of molecules—usually a very large finite collection—in which case the governing laws are expressible as a finite set of ordinary differential equations. These equations are generally highly intractable, and the set of molecules is usually approximated by a continuous distribution of mass. The governing laws are then expressed as a set of partial differential equations, containing such quantities as velocity, density, and pressure as dependent variables.

It is sometimes possible to obtain particular solutions of these equations analytically, especially when the solutions are periodic or invariant with time, and, indeed, much work has been devoted to obtaining such solutions by one scheme or another. Ordinarily, however, nonperiodic solutions cannot readily be determined except by numerical procedures. Such procedures involve replacing the continuous variables by a new finite set of functions of time, which may perhaps be the values of the continuous variables at a chosen grid of points, or the coefficients in the expansions of these variables in series of orthogonal functions. The governing laws then become a finite set of ordinary differential

¹ The research reported in this work has been sponsored by the Geophysics Research Directorate of the Air Force Cambridge Research Center, under Contract No. AF 19(604)-4969.

equations again, although a far simpler set than the one which governs individual molecular motions.

In any real hydrodynamical system, viscous dissipation is always occurring, unless the system is moving as a solid, and thermal dissipation is always occurring, unless the system is at constant temperature. For certain purposes many systems may be treated as conservative systems, in which the total energy, or some other quantity, does not vary with time. In seeking the ultimate behavior of a system, the use of conservative equations is unsatisfactory, since the ultimate value of any conservative quantity would then have to equal the arbitrarily chosen initial value. This difficulty may be obviated by including the dissipative processes, thereby making the equations nonconservative, and also including external mechanical or thermal forcing, thus preventing the system from ultimately reaching a state of rest. If the system is to be deterministic, the forcing functions, if not constant with time, must themselves vary according to some deterministic rule.

In this work, then, we shall deal specifically with finite systems of deterministic ordinary differential equations, designed to represent forced dissipative hydrodynamical systems. We shall study the properties of nonperiodic solutions of these equations.

It is not obvious that such solutions can exist at all. Indeed, in dissipative systems governed by finite sets of linear equations, a constant forcing leads ultimately to a constant response, while a periodic forcing leads to a periodic response. Hence, nonperiodic flow has sometimes been regarded as the result of nonperiodic or random forcing.

The reasoning leading to these conclusions is not applicable when the governing equations are nonlinear. If the equations contain terms representing advection—the transport of some property of a fluid by the motion of the fluid itself—a constant forcing can lead to a variable response. In the rotating-basin experiments already mentioned, both periodic and nonperiodic flow result from thermal forcing which, within the limits of experimental control, is constant. Exact periodic solutions of simplified systems of equations, representing dissipative flow with constant thermal forcing, have been obtained analytically by the writer (1962a). The writer (1962b) has also found nonperiodic solutions of similar systems of equations by numerical means.

2. Phase space

Consider a system whose state may be described by M variables X_1, \dots, X_M . Let the system be governed by the set of equations

$$dX_i/dt = F_i(X_1, \dots, X_M), \quad i=1, \dots, M, \quad (1)$$

where time t is the single independent variable, and the functions F_i possess continuous first partial derivatives. Such a system may be studied by means of phase space—

an M -dimensional Euclidean space Γ whose coordinates are X_1, \dots, X_M . Each point in phase space represents a possible instantaneous state of the system. A state which is varying in accordance with (1) is represented by a moving particle in phase space, traveling along a trajectory in phase space. For completeness, the position of a stationary particle, representing a steady state, is included as a trajectory.

Phase space has been a useful concept in treating finite systems, and has been used by such mathematicians as Gibbs (1902) in his development of statistical mechanics, Poincaré (1881) in his treatment of the solutions of differential equations, and Birkhoff (1927) in his treatise on dynamical systems.

From the theory of differential equations (e.g., Ford 1933, ch. 6), it follows, since the partial derivatives $\partial F_i/\partial X_j$ are continuous, that if t_0 is any time, and if X_{10}, \dots, X_{M0} is any point in Γ , equations (1) possess a unique solution

$$X_i = f_i(X_{10}, \dots, X_{M0}, t), \quad i=1, \dots, M, \quad (2)$$

valid throughout some time interval containing t_0 , and satisfying the condition

$$f_i(X_{10}, \dots, X_{M0}, t_0) = X_{i0}, \quad i=1, \dots, M. \quad (3)$$

The functions f_i are continuous in X_{10}, \dots, X_{M0} and t . Hence there is a unique trajectory through each point of Γ . Two or more trajectories may, however, approach the same point or the same curve asymptotically as $t \rightarrow \infty$ or as $t \rightarrow -\infty$. Moreover, since the functions f_i are continuous, the passage of time defines a continuous deformation of any region of Γ into another region.

In the familiar case of a conservative system, where some positive definite quantity Q , which may represent some form of energy, is invariant with time, each trajectory is confined to one or another of the surfaces of constant Q . These surfaces may take the form of closed concentric shells.

If, on the other hand, there is dissipation and forcing, and if, whenever Q equals or exceeds some fixed value Q_1 , the dissipation acts to diminish Q more rapidly than the forcing can increase Q , then $(-dQ/dt)$ has a positive lower bound where $Q \geq Q_1$, and each trajectory must ultimately become trapped in the region where $Q < Q_1$. Trajectories representing forced dissipative flow may therefore differ considerably from those representing conservative flow.

Forced dissipative systems of this sort are typified by the system

$$dX_i/dt = \sum_{j,k} a_{ijk} X_j X_k - \sum_j b_{ij} X_j + c_i, \quad (4)$$

where $\sum a_{ijk} X_j X_k$ vanishes identically, $\sum b_{ij} X_j$ is positive definite, and c_1, \dots, c_M are constants. If

$$Q = \frac{1}{2} \sum_i X_i^2, \quad (5)$$

and if e_1, \dots, e_M are the roots of the equations

$$\sum_i (b_{ij} + b_{ji})e_j = c_i, \quad (6)$$

it follows from (4) that

$$dQ/dt = \sum_{i,j} b_{ij}e_i e_j - \sum_{i,j} b_{ij}(X_i - e_i)(X_j - e_j). \quad (7)$$

The right side of (7) vanishes only on the surface of an ellipsoid E , and is positive only in the interior of E . The surfaces of constant Q are concentric spheres. If S denotes a particular one of these spheres whose interior R contains the ellipsoid E , it is evident that each trajectory eventually becomes trapped within R .

3. The instability of nonperiodic flow

In this section we shall establish one of the most important properties of deterministic nonperiodic flow, namely, its instability with respect to modifications of small amplitude. We shall find it convenient to do this by identifying the solutions of the governing equations with trajectories in phase space. We shall use such symbols as $P(t)$ (variable argument) to denote trajectories, and such symbols as P or $P(t_0)$ (no argument or constant argument) to denote points, the latter symbol denoting the specific point through which $P(t)$ passes at time t_0 .

We shall deal with a phase space Γ in which a unique trajectory passes through each point, and where the passage of time defines a continuous deformation of any region of Γ into another region, so that if the points $P_1(t_0), P_2(t_0), \dots$ approach $P_0(t_0)$ as a limit, the points $P_1(t_0+\tau), P_2(t_0+\tau), \dots$ must approach $P_0(t_0+\tau)$ as a limit. We shall furthermore require that the trajectories be uniformly bounded as $t \rightarrow \infty$; that is, there must be a bounded region R , such that every trajectory ultimately remains with R . Our procedure is influenced by the work of Birkhoff (1927) on dynamical systems, but differs in that Birkhoff was concerned mainly with conservative systems. A rather detailed treatment of dynamical systems has been given by Nemytskii and Stepanov (1960), and rigorous proofs of some of the theorems which we shall present are to be found in that source.

We shall first classify the trajectories in three different manners, namely, according to the absence or presence of transient properties, according to the stability or instability of the trajectories with respect to small modifications, and according to the presence or absence of periodic behavior.

Since any trajectory $P(t)$ is bounded, it must possess at least one *limit point* P_0 , a point which it approaches arbitrarily closely arbitrarily often. More precisely, P_0 is a limit point of $P(t)$ if for any $\epsilon > 0$ and any time t_1 there exists a time $t_2(\epsilon, t_1) > t_1$ such that $|P(t_2) - P_0| < \epsilon$. Here

absolute-value signs denote distance in phase space. Because Γ is continuously deformed as t varies, every point on the trajectory through P_0 is also a limit point of $P(t)$, and the set of limit points of $P(t)$ forms a trajectory, or a set of trajectories, called the *limiting trajectories* of $P(t)$. A limiting trajectory is obviously contained within R in its entirety.

If a trajectory is contained among its own limiting trajectories, it will be called *central*; otherwise it will be called *noncentral*. A central trajectory passes arbitrarily closely arbitrarily often to any point through which it has previously passed, and, in this sense at least, separate sufficiently long segments of a central trajectory are statistically similar. A noncentral trajectory remains a certain distance away from any point through which it has previously passed. It must approach its entire set of limit points asymptotically, although it need not approach any particular limiting trajectory asymptotically. Its instantaneous distance from its closest limit point is therefore a transient quantity, which becomes arbitrarily small as $t \rightarrow \infty$.

A trajectory $P(t)$ will be called *stable at a point* $P(t_1)$ if any other trajectory passing sufficiently close to $P(t_1)$ at time t_1 remains close to $P(t)$ as $t \rightarrow \infty$; i.e., $P(t)$ is stable at $P(t_1)$ if for any $\epsilon > 0$ there exists a $\delta(\epsilon, t_1) > 0$ such that if $|P_1(t_1) - P(t_1)| < \delta$ and $t_2 > t_1$, $|P_1(t_2) - P(t_2)| < \epsilon$. Otherwise $P(t)$ will be called *unstable* at $P(t_1)$. Because Γ is continuously deformed as t varies, a trajectory which is stable at one point is stable at every point, and will be called a *stable* trajectory. A trajectory unstable at one point is unstable at every point, and will be called an *unstable* trajectory. In the special case that $P(t)$ is confined to one point, this definition of stability coincides with the familiar concept of stability of steady flow.

A stable trajectory $P(t)$ will be called uniformly stable if the distance within which a neighboring trajectory must approach a point $P(t_1)$, in order to be certain of remaining close to $P(t)$ as $t \rightarrow \infty$, itself possesses a positive lower bound as $t_1 \rightarrow \infty$; i.e., $P(t)$ is uniformly stable if for any $\epsilon > 0$ there exists a $\delta(\epsilon) > 0$ and a time $t_0(\epsilon)$ such that if $t_1 > t_0$ and $|P_1(t_1) - P(t_1)| < \delta$ and $t_2 > t_1$, $|P_1(t_2) - P(t_2)| < \epsilon$. A limiting trajectory $P_0(t)$ of a uniformly stable trajectory $P(t)$ must be uniformly stable itself, since all trajectories passing sufficiently close to $P_0(t)$ must pass arbitrarily close to some point of $P(t)$ and so must remain close to $P(t)$, and hence to $P_0(t)$, as $t \rightarrow \infty$.

Since each point lies on a unique trajectory, any trajectory passing through a point through which it has previously passed must continue to repeat its past behavior, and so must be *periodic*. A trajectory $P(t)$ will be called *quasi-periodic* if for some arbitrarily large time interval τ , $P(t+\tau)$ ultimately remains arbitrarily close to $P(t)$, i.e., $P(t)$ is quasi-periodic if for any $\epsilon > 0$ and for any time interval τ_0 , there exists a $\tau(\epsilon, \tau_0) > \tau_0$ and a time $t_1(\epsilon, \tau_0)$ such that if $t_2 > t_1$, $|P(t_2+\tau) - P(t_2)|$

$< \epsilon$. Periodic trajectories are special cases of quasi-periodic trajectories.

A trajectory which is not quasi-periodic will be called *nonperiodic*. If $P(t)$ is nonperiodic, $P(t_1+\tau)$ may be arbitrarily close to $P(t_1)$ for some time t_1 and some arbitrarily large time interval τ , but, if this is so, $P(t+\tau)$ cannot remain arbitrarily close to $P(t)$ as $t \rightarrow \infty$. Non-periodic trajectories are of course representations of deterministic nonperiodic flow, and form the principal subject of this paper.

Periodic trajectories are obviously central. Quasi-periodic central trajectories include multiple periodic trajectories with incommensurable periods, while quasi-periodic noncentral trajectories include those which approach periodic trajectories asymptotically. Non-periodic trajectories may be central or noncentral.

We can now establish the theorem that a trajectory with a stable limiting trajectory is quasi-periodic. For if $P_0(t)$ is a limiting trajectory of $P(t)$, two distinct points $P(t_1)$ and $P(t_1+\tau)$, with τ arbitrarily large, may be found arbitrary close to any point $P_0(t_0)$. Since $P_0(t)$ is stable, $P(t)$ and $P(t+\tau)$ must remain arbitrarily close to $P_0(t+t_0-t_1)$, and hence to each other, as $t \rightarrow \infty$, and $P(t)$ is quasi-periodic.

It follows immediately that a stable central trajectory is quasi-periodic, or, equivalently, that a nonperiodic central trajectory is unstable.

The result has far-reaching consequences when the system being considered is an observable nonperiodic system whose future state we may desire to predict. It implies that two states differing by imperceptible amounts may eventually evolve into two considerably different states. If, then, there is any error whatever in observing the present state—and in any real system such errors seem inevitable—an acceptable prediction of an instantaneous state in the distant future may well be impossible.

As for noncentral trajectories, it follows that a uniformly stable noncentral trajectory is quasi-periodic, or, equivalently, a nonperiodic noncentral trajectory is not uniformly stable. The possibility of a nonperiodic non-central trajectory which is stable but not uniformly stable still exists. To the writer, at least, such trajectories, although possible on paper, do not seem characteristic of real hydrodynamical phenomena. Any claim that atmospheric flow, for example, is represented by a trajectory of this sort would lead to the improbable conclusion that we ought to master long-range forecasting as soon as possible, because, the longer we wait, the more difficult our task will become.

In summary, we have shown that, subject to the conditions of uniqueness, continuity, and boundedness prescribed at the beginning of this section, a central trajectory, which in a certain sense is free of transient properties, is unstable if it is nonperiodic. A noncentral trajectory, which is characterized by transient properties, is not uniformly stable if it is nonperiodic, and,

if it is stable at all, its very stability is one of its transient properties, which tends to die out as time progresses. In view of the impossibility of measuring initial conditions precisely, and thereby distinguishing between a central trajectory and a nearby noncentral trajectory, all nonperiodic trajectories are effectively unstable from the point of view of practical prediction.

4. Numerical integration of nonconservative systems

The theorems of the last section can be of importance only if nonperiodic solutions of equations of the type considered actually exist. Since statistically stationary nonperiodic functions of time are not easily described analytically, particular nonperiodic solutions can probably be found most readily by numerical procedures. In this section we shall examine a numerical-integration procedure which is especially applicable to systems of equations of the form (4). In a later section we shall use this procedure to determine a nonperiodic solution of a simple set of equations.

To solve (1) numerically we may choose an initial time t_0 and a time increment Δt , and let

$$X_{i,n} = X_i(t_0 + n\Delta t). \quad (8)$$

We then introduce the auxiliary approximations

$$X_{i(n+1)} = X_{i,n} + F_i(P_n)\Delta t, \quad (9)$$

$$X_{i((n+2))} = X_{i(n+1)} + F_i(P_{(n+1)})\Delta t, \quad (10)$$

where P_n and $P_{(n+1)}$ are the points whose coordinates are

$$(X_{1,n}, \dots, X_{M,n}) \text{ and } (X_{1(n+1)}, \dots, X_{M(n+1)}).$$

The simplest numerical procedure for obtaining approximate solutions of (1) is the forward-difference procedure,

$$X_{i,n+1} = X_{i(n+1)}. \quad (11)$$

In many instances better approximations to the solutions of (1) may be obtained by a centered-difference procedure

$$X_{i,n+1} = X_{i,n-1} + 2F_i(P_n)\Delta t. \quad (12)$$

This procedure is unsuitable, however, when the deterministic nature of (1) is a matter of concern, since the values of $X_{1,n}, \dots, X_{M,n}$ do not uniquely determine the values of $X_{1,n+1}, \dots, X_{M,n+1}$.

A procedure which largely overcomes the disadvantages of both the forward-difference and centered-difference procedures is the double-approximation procedure, defined by the relation

$$X_{i,n+1} = X_{i,n} + \frac{1}{2}[F_i(P_n) + F_i(P_{(n+1)})]\Delta t. \quad (13)$$

Here the coefficient of Δt is an approximation to the time derivative of X_i at time $t_0 + (n + \frac{1}{2})\Delta t$. From (9) and (10), it follows that (13) may be rewritten

$$X_{i,n+1} = \frac{1}{2}(X_{i,n} + X_{i((n+2))}). \quad (14)$$

A convenient scheme for automatic computation is the successive evaluation of $X_{i(n+1)}$, $X_{i((n+2))}$, and $X_{i,n+1}$ according to (9), (10) and (14). We have used this procedure in all the computations described in this study.

In phase space a numerical solution of (1) must be represented by a jumping particle rather than a continuously moving particle. Moreover, if a digital computer is instructed to represent each number in its memory by a preassigned fixed number of bits, only certain discrete points in phase space will ever be occupied. If the numerical solution is bounded, repetitions must eventually occur, so that, strictly speaking, every numerical solution is periodic. In practice this consideration may be disregarded, if the number of different possible states is far greater than the number of iterations ever likely to be performed. The necessity for repetition could be avoided altogether by the somewhat uneconomical procedure of letting the precision of computation increase as n increases.

Consider now numerical solutions of equations (4), obtained by the forward-difference procedure (11). For such solutions,

$$Q_{n+1} = Q_n + (dQ/dt)_n \Delta t + \frac{1}{2} \sum_i F_i^2(P_n) \Delta t^2. \quad (15)$$

Let S' be any surface of constant Q whose interior R' contains the ellipsoid E where dQ/dt vanishes, and let S be any surface of constant Q whose interior R contains S' .

Since $\sum F_i^2$ and dQ/dt both possess upper bounds in R' , we may choose Δt so small that P_{n+1} lies in R if P_n lies in R' . Likewise, since $\sum F_i^2$ possesses an upper bound and dQ/dt possesses a negative upper bound in $R - R'$, we may choose Δt so small that $Q_{n+1} < Q_n$ if P_n lies in $R - R'$. Hence Δt may be chosen so small that any jumping particle which has entered R remains trapped within R , and the numerical solution does not blow up. A blow-up may still occur, however, if initially the particle is exterior to R .

Consider now the double-approximation procedure (14). The previous arguments imply not only that $P_{(n+1)}$ lies within R if P_n lies within R , but also that $P_{((n+2))}$ lies within R if $P_{(n+1)}$ lies within R . Since the region R is convex, it follows that P_{n+1} , as given by (14), lies within R if P_n lies within R . Hence if Δt is chosen so small that the forward-difference procedure does not blow up, the double-approximation procedure also does not blow up.

We note in passing that if we apply the forward-difference procedure to a conservative system where $dQ/dt = 0$ everywhere,

$$Q_{n+1} = Q_n + \frac{1}{2} \sum_i F_i^2(P_n) \Delta t^2. \quad (16)$$

In this case, for any fixed choice of Δt the numerical solution ultimately goes to infinity, unless it is asymptotically approaching a steady state. A similar result holds when the double-approximation procedure (14) is applied to a conservative system.

5. The convection equations of Saltzman

In this section we shall introduce a system of three ordinary differential equations whose solutions afford the simplest example of deterministic nonperiodic flow of which the writer is aware. The system is a simplification of one derived by Saltzman (1962) to study finite-amplitude convection. Although our present interest is in the nonperiodic nature of its solutions, rather than in its contributions to the convection problem, we shall describe its physical background briefly.

Rayleigh (1916) studied the flow occurring in a layer of fluid of uniform depth H , when the temperature difference between the upper and lower surfaces is maintained at a constant value ΔT . Such a system possesses a steady-state solution in which there is no motion, and the temperature varies linearly with depth. If this solution is unstable, convection should develop.

In the case where all motions are parallel to the $x-z$ -plane, and no variations in the direction of the y -axis occur, the governing equations may be written (see Saltzman, 1962)

$$\frac{\partial}{\partial t} \nabla^2 \psi = -\frac{\partial(\psi, \nabla^2 \psi)}{\partial(x, z)} + \nu \nabla^4 \psi + g \alpha \frac{\partial \theta}{\partial x}, \quad (17)$$

$$\frac{\partial}{\partial t} \theta = -\frac{\partial(\psi, \theta)}{\partial(x, z)} + \frac{\Delta T}{H} \frac{\partial \psi}{\partial x} + \kappa \nabla^2 \theta. \quad (18)$$

Here ψ is a stream function for the two-dimensional motion, θ is the departure of temperature from that occurring in the state of no convection, and the constants g , α , ν , and κ denote, respectively, the acceleration of gravity, the coefficient of thermal expansion, the kinematic viscosity, and the thermal conductivity. The problem is most tractable when both the upper and lower boundaries are taken to be free, in which case ψ and $\nabla^2 \psi$ vanish at both boundaries.

Rayleigh found that fields of motion of the form

$$\psi = \psi_0 \sin(\pi a H^{-1} x) \sin(\pi H^{-1} z), \quad (19)$$

$$\theta = \theta_0 \cos(\pi a H^{-1} x) \sin(\pi H^{-1} z), \quad (20)$$

would develop if the quantity

$$R_a = g \alpha H^3 \Delta T \nu^{-1} \kappa^{-1}, \quad (21)$$

now called the *Rayleigh number*, exceeded a critical value

$$R_c = \pi^4 a^{-2} (1 + a^2)^3. \quad (22)$$

The minimum value of R_c , namely $27\pi^4/4$, occurs when $a^2 = \frac{1}{2}$.

Saltzman (1962) derived a set of ordinary differential equations by expanding ψ and θ in double Fourier series in x and z , with functions of t alone for coefficients, and

substituting these series into (17) and (18). He arranged the right-hand sides of the resulting equations in double-Fourier-series form, by replacing products of trigonometric functions of x (or z) by sums of trigonometric functions, and then equated coefficients of similar functions of x and z . He then reduced the resulting infinite system to a finite system by omitting reference to all but a specified finite set of functions of t , in the manner proposed by the writer (1960).

He then obtained time-dependent solutions by numerical integration. In certain cases all except three of the dependent variables eventually tended to zero, and these three variables underwent irregular, apparently nonperiodic fluctuations.

These same solutions would have been obtained if the series had at the start been truncated to include a total of three terms. Accordingly, in this study we shall let

$$a(1+a^2)^{-1}\kappa^{-1}\psi = X\sqrt{2} \sin(\pi aH^{-1}x) \sin(\pi H^{-1}z), \quad (23)$$

$$\begin{aligned} \pi R_c^{-1}R_a \Delta T^{-1}\theta &= Y\sqrt{2} \cos(\pi aH^{-1}x) \sin(\pi H^{-1}z) \\ &\quad - Z \sin(2\pi H^{-1}z), \end{aligned} \quad (24)$$

where X , Y , and Z are functions of time alone. When expressions (23) and (24) are substituted into (17) and (18), and trigonometric terms other than those occurring in (23) and (24) are omitted, we obtain the equations

$$X' = -\sigma X + \sigma Y, \quad (25)$$

$$Y' = -XZ + rX - Y, \quad (26)$$

$$Z' = XY - bZ. \quad (27)$$

Here a dot denotes a derivative with respect to the dimensionless time $\tau = \pi^2 H^{-2}(1+a^2)\kappa t$, while $\sigma = \kappa^{-1}\nu$ is the *Prandtl number*, $r = R_c^{-1}R_a$, and $b = 4(1+a^2)^{-1}$. Except for multiplicative constants, our variables X , Y , and Z are the same as Saltzman's variables A , D , and G . Equations (25), (26), and (27) are the convection equations whose solutions we shall study.

In these equations X is proportional to the intensity of the convective motion, while Y is proportional to the temperature difference between the ascending and descending currents, similar signs of X and Y denoting that warm fluid is rising and cold fluid is descending. The variable Z is proportional to the distortion of the vertical temperature profile from linearity, a positive value indicating that the strongest gradients occur near the boundaries.

Equations (25)–(27) may give realistic results when the Rayleigh number is slightly supercritical, but their solutions cannot be expected to resemble those of (17) and (18) when strong convection occurs, in view of the extreme truncation.

6. Applications of linear theory

Although equations (25)–(27), as they stand, do not have the form of (4), a number of linear transformations

will convert them to this form. One of the simplest of these is the transformation

$$X' = X, \quad Y' = Y, \quad Z' = Z - r - \sigma. \quad (28)$$

Solutions of (25)–(27) therefore remain bounded within a region R as $\tau \rightarrow \infty$, and the general results of Sections 2, 3 and 4 apply to these equations.

The stability of a solution $X(\tau)$, $Y(\tau)$, $Z(\tau)$ may be formally investigated by considering the behavior of small superposed perturbations $x_0(\tau)$, $y_0(\tau)$, $z_0(\tau)$. Such perturbations are temporarily governed by the linearized equations

$$\begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}' = \begin{bmatrix} -\sigma & \sigma & 0 \\ (r-Z) & -1 & -X \\ Y & X & -b \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}. \quad (29)$$

Since the coefficients in (29) vary with time, unless the basic state X , Y , Z is a steady-state solution of (25)–(27), a general solution of (29) is not feasible. However, the variation of the volume V_0 of a small region in phase space, as each point in the region is displaced in accordance with (25)–(27), is determined by the diagonal sum of the matrix of coefficients; specifically

$$V_0' = -(\sigma + b + 1)V_0. \quad (30)$$

This is perhaps most readily seen by visualizing the motion in phase space as the flow of a fluid, whose divergence is

$$\frac{\partial X'}{\partial X} + \frac{\partial Y'}{\partial Y} + \frac{\partial Z'}{\partial Z} = -(\sigma + b + 1). \quad (31)$$

Hence each small volume shrinks to zero as $\tau \rightarrow \infty$, at a rate independent of X , Y , and Z . This does not imply that each small volume shrinks to a point; it may simply become flattened into a surface. It follows that the volume of the region initially enclosed by the surface S shrinks to zero at this same rate, so that all trajectories ultimately become confined to a specific subspace having zero volume. This subspace contains all those trajectories which lie entirely within R , and so contains all central trajectories.

Equations (25)–(27) possess the steady-state solution $X = Y = Z = 0$, representing the state of no convection. With this basic solution, the characteristic equation of the matrix in (29) is

$$[\lambda + b][\lambda^2 + (\sigma + 1)\lambda + \sigma(1 - r)] = 0. \quad (32)$$

This equation has three real roots when $r > 0$; all are negative when $r < 1$, but one is positive when $r > 1$. The criterion for the onset of convection is therefore $r = 1$, or $R_c = R_a$, in agreement with Rayleigh's result.

When $r > 1$, equations (25)–(27) possess two additional steady-state solutions $X = Y = \pm \sqrt{b(r-1)}$, $Z = r - 1$.

For either of these solutions, the characteristic equation of the matrix in (29) is

$$\lambda^3 + (\sigma + b + 1)\lambda^2 + (r + \sigma)b\lambda + 2\sigma b(r - 1) = 0. \quad (33)$$

This equation possesses one real negative root and two complex conjugate roots when $r > 1$; the complex conjugate roots are pure imaginary if the product of the coefficients of λ^2 and λ equals the constant term, or

$$r = \sigma(\sigma + b + 3)(\sigma - b - 1)^{-1}. \quad (34)$$

This is the critical value of r for the instability of steady convection. Thus if $\sigma < b + 1$, no positive value of r satisfies (34), and steady convection is always stable, but if $\sigma > b + 1$, steady convection is unstable for sufficiently high Rayleigh numbers. This result of course applies only to idealized convection governed by (25)–(27), and not to the solutions of the partial differential equations (17) and (18).

The presence of complex roots of (34) shows that if unstable steady convection is disturbed, the motion will oscillate in intensity. What happens when the disturbances become large is not revealed by linear theory. To investigate finite-amplitude convection, and to study the subspace to which trajectories are ultimately confined, we turn to numerical integration.

TABLE 1. Numerical solution of the convection equations. Values of X , Y , Z are given at every fifth iteration N , for the first 160 iterations.

N	X	Y	Z
0000	0000	0010	0000
0005	0004	0012	0000
0010	0009	0020	0000
0015	0016	0036	0002
0020	0030	0066	0007
0025	0054	0115	0024
0030	0093	0192	0074
0035	0150	0268	0201
0040	0195	0234	0397
0045	0174	0055	0483
0050	0097	-0067	0415
0055	0025	-0093	0340
0060	-0020	-0089	0298
0065	-0046	-0084	0275
0070	-0061	-0083	0262
0075	-0070	-0086	0256
0080	-0077	-0091	0255
0085	-0084	-0095	0258
0090	-0089	-0098	0266
0095	-0093	-0098	0275
0100	-0094	-0093	0283
0105	-0092	-0086	0297
0110	-0088	-0079	0286
0115	-0083	-0073	0281
0120	-0078	-0070	0273
0125	-0075	-0071	0264
0130	-0074	-0075	0257
0135	-0076	-0080	0252
0140	-0079	-0087	0251
0145	-0083	-0093	0254
0150	-0088	-0098	0262
0155	-0092	-0099	0271
0160	-0094	-0096	0281

7. Numerical integration of the convection equations

To obtain numerical solutions of the convection equations, we must choose numerical values for the constants. Following Saltzman (1962), we shall let $\sigma = 10$ and $a^2 = \frac{1}{2}$, so that $b = 8/3$. The critical Rayleigh number for instability of steady convection then occurs when $r = 470/19 = 24.74$.

We shall choose the slightly supercritical value $r = 28$. The states of steady convection are then represented by the points $(6\sqrt{2}, 6\sqrt{2}, 27)$ and $(-6\sqrt{2}, -6\sqrt{2}, 27)$ in phase space, while the state of no convection corresponds to the origin $(0,0,0)$.

We have used the double-approximation procedure for numerical integration, defined by (9), (10), and (14). The value $\Delta\tau = 0.01$ has been chosen for the dimensionless time increment. The computations have been performed on a Royal McBee LGP-30 electronic com-

TABLE 2. Numerical solution of the convection equations. Values of X , Y , Z are given at every iteration N for which Z possesses a relative maximum, for the first 6000 iterations.

N	X	Y	Z	N	X	Y	Z
0045	0174	0055	0483	3029	0117	0075	0352
0107	-0091	-0083	0287	3098	0123	0076	0365
0168	-0092	-0084	0288	3171	0134	0082	0383
0230	-0092	-0084	0289	3268	0155	0069	0435
0292	-0092	-0083	0290	3333	-0114	-0079	0342
0354	-0093	-0083	0292	3400	-0117	-0077	0350
0416	-0093	-0083	0293	3468	-0125	-0083	0361
0478	-0094	-0082	0295	3541	-0129	-0073	0378
0540	-0094	-0082	0296	3625	-0146	-0074	0413
0602	-0095	-0082	0298	3695	0127	0079	0370
0664	-0096	-0083	0300	3772	0136	0072	0394
0726	-0097	-0083	0302	3853	-0144	-0077	0407
0789	-0097	-0081	0304	3926	0129	0072	0380
0851	-0099	-0083	0307	4014	0148	0068	0421
0914	-0100	-0081	0309	4082	-0120	-0074	0359
0977	-0100	-0080	0312	4153	-0129	-0078	0375
1040	-0102	-0080	0315	4233	-0144	-0082	0404
1103	-0104	-0081	0319	4307	0135	0081	0385
1167	-0105	-0079	0323	4417	-0162	-0069	0450
1231	-0107	-0079	0328	4480	0106	0081	0324
1295	-0111	-0082	0333	4544	0109	0082	0329
1361	-0111	-0077	0339	4609	0110	0080	0334
1427	-0116	-0079	0347	4675	0112	0076	0341
1495	-0120	-0077	0357	4741	0118	0081	0349
1566	-0125	-0072	0371	4810	0120	0074	0360
1643	-0139	-0077	0396	4881	0130	0081	0376
1722	0140	0075	0401	4963	0141	0068	0406
1798	-0135	-0072	0391	5035	-0133	-0081	0381
1882	0146	0074	0413	5124	-0151	-0076	0422
1952	-0127	-0078	0370	5192	0119	0075	0358
2029	-0135	-0070	0393	5262	0129	0083	0372
2110	0146	0083	0408	5340	0140	0079	0397
2183	-0128	-0070	0379	5419	-0137	-0067	0399
2268	-0144	-0066	0415	5495	0140	0081	0394
2337	0126	0079	0368	5576	-0141	-0072	0405
2412	0137	0081	0389	5649	0135	0082	0384
2501	-0153	-0080	0423	5752	0160	0074	0443
2569	0119	0076	0357	5816	-0110	-0081	0332
2639	0129	0082	0371	5881	-0113	-0082	0339
2717	0136	0070	0395	5948	-0114	-0075	0346
2796	-0143	-0079	0402				
2871	0134	0076	0388				
2962	-0152	-0072	0426				

puting machine. Approximately one second per iteration, aside from output time, is required.

For initial conditions we have chosen a slight departure from the state of no convection, namely $(0,1,0)$. Table 1 has been prepared by the computer. It gives the values of N (the number of iterations), X , Y , and Z at every fifth iteration for the first 160 iterations. In the printed output (but not in the computations) the values of X , Y , and Z are multiplied by ten, and then only those figures to the left of the decimal point are printed. Thus the states of steady convection would appear as 0084, 0084, 0270 and $-0084, -0084, 0270$, while the state of no convection would appear as 0000, 0000, 0000.

The initial instability of the state of rest is evident. All three variables grow rapidly, as the sinking cold fluid is replaced by even colder fluid from above, and the rising warm fluid by warmer fluid from below, so that by step 35 the strength of the convection far exceeds that of steady convection. Then Y diminishes as the warm fluid is carried over the top of the convective cells, so that by step 50, when X and Y have opposite signs, warm fluid is descending and cold fluid is ascending. The motion thereupon ceases and reverses its direction, as indicated by the negative values of X following step 60. By step 85 the system has reached a state not far from that of steady convection. Between steps 85 and 150 it executes a complete oscillation in its intensity, the slight amplification being almost indetectable.

The subsequent behavior of the system is illustrated in Fig. 1, which shows the behavior of Y for the first 3000 iterations. After reaching its early peak near step 35 and then approaching equilibrium near step 85, it undergoes systematic amplified oscillations until near step 1650. At this point a critical state is reached, and thereafter Y changes sign at seemingly irregular intervals, reaching sometimes one, sometimes two, and sometimes three or more extremes of one sign before changing sign again.

Fig. 2 shows the projections on the X - Y - and Y - Z -planes in phase space of the portion of the trajectory corresponding to iterations 1400–1900. The states of steady convection are denoted by C and C' . The first portion of the trajectory spirals outward from the vicinity of C' , as the oscillations about the state of steady convection, which have been occurring since step 85, continue to grow. Eventually, near step 1650, it crosses the X - Z -plane, and is then deflected toward the neighborhood of C . It temporarily spirals about C , but crosses the X - Z -plane after one circuit, and returns to the neighborhood of C' , where it soon joins the spiral over which it has previously traveled. Thereafter it crosses from one spiral to the other at irregular intervals.

Fig. 3, in which the coordinates are Y and Z , is based upon the printed values of X , Y , and Z at every fifth iteration for the first 6000 iterations. These values determine X as a smooth single-valued function of Y and Z over much of the range of Y and Z ; they determine X

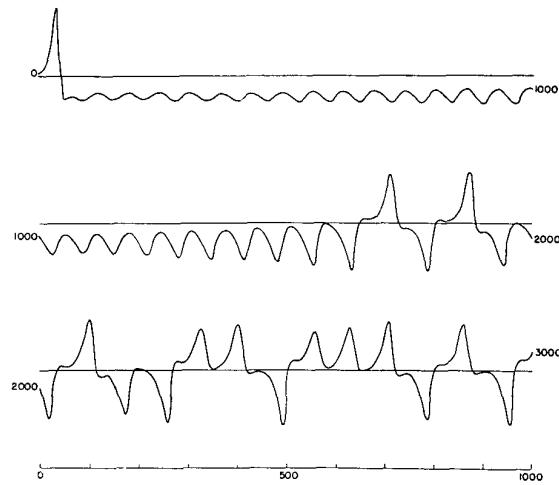


FIG. 1. Numerical solution of the convection equations. Graph of Y as a function of time for the first 1000 iterations (upper curve), second 1000 iterations (middle curve), and third 1000 iterations (lower curve).

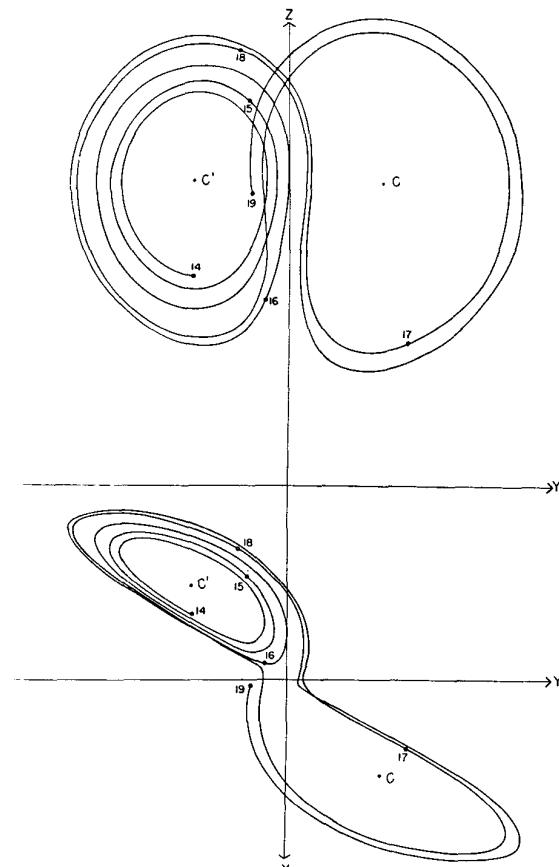


FIG. 2. Numerical solution of the convection equations. Projections on the X - Y -plane and the Y - Z -plane in phase space of the segment of the trajectory extending from iteration 1400 to iteration 1900. Numerals "14," "15," etc., denote positions at iterations 1400, 1500, etc. States of steady convection are denoted by C and C' .

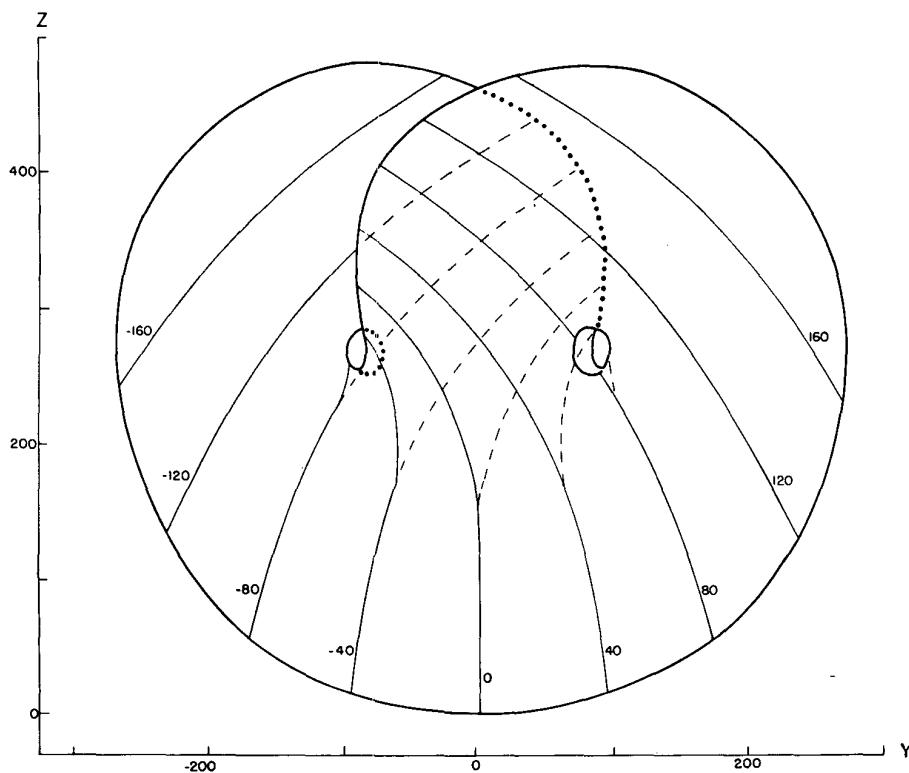


FIG. 3. Isopleths of X as a function of Y and Z (thin solid curves), and isopleths of the lower of two values of X , where two values occur (dashed curves), for approximate surfaces formed by all points on limiting trajectories. Heavy solid curve, and extensions as dotted curves, indicate natural boundaries of surfaces.

as one of two smooth single-valued functions over the remainder of the range. In Fig. 3 the thin solid lines are isopleths of X , and where two values of X exist, the dashed lines are isopleths of the lower value. Thus, within the limits of accuracy of the printed values, the trajectory is confined to a pair of surfaces which appear to merge in the lower portion of Fig. 3. The spiral about C lies in the upper surface, while the spiral about C' lies in the lower surface. Thus it is possible for the trajectory to pass back and forth from one spiral to the other without intersecting itself.

Additional numerical solutions indicate that other trajectories, originating at points well removed from these surfaces, soon meet these surfaces. The surfaces therefore appear to be composed of all points lying on limiting trajectories.

Because the origin represents a steady state, no trajectory can pass through it. However, two trajectories emanate from it, i.e., approach it asymptotically as $\tau \rightarrow -\infty$. The heavy solid curve in Fig. 3, and its extensions as dotted curves, are formed by these two trajectories. Trajectories passing close to the origin will tend to follow the heavy curve, but will not cross it, so that the heavy curve forms a natural boundary to the region which a trajectory can ultimately occupy. The

holes near C and C' also represent regions which cannot be occupied after they have once been abandoned.

Returning to Fig. 2, we find that the trajectory apparently leaves one spiral only after exceeding some critical distance from the center. Moreover, the extent to which this distance is exceeded appears to determine the point at which the next spiral is entered; this in turn seems to determine the number of circuits to be executed before changing spirals again.

It therefore seems that some single feature of a given circuit should predict the same feature of the following circuit. A suitable feature of this sort is the maximum value of Z , which occurs when a circuit is nearly completed. Table 2 has again been prepared by the computer, and shows the values of X , Y , and Z at only those iterations N for which Z has a relative maximum. The succession of circuits about C and C' is indicated by the succession of positive and negative values of X and Y . Evidently X and Y change signs following a maximum which exceeds some critical value printed as about 385.

Fig. 4 has been prepared from Table 2. The abscissa is M_n , the value of the n th maximum of Z , while the ordinate is M_{n+1} , the value of the following maximum. Each point represents a pair of successive values of Z taken from Table 2. Within the limits of the round-off

in tabulating Z , there is a precise two-to-one relation between M_n and M_{n+1} . The initial maximum $M_1=483$ is shown as if it had followed a maximum $M_0=385$, since maxima near 385 are followed by close approaches to the origin, and then by exceptionally large maxima.

It follows that an investigator, unaware of the nature of the governing equations, could formulate an empirical prediction scheme from the "data" pictured in Figs. 2 and 4. From the value of the most recent maximum of Z , values at future maxima may be obtained by repeated applications of Fig. 4. Values of X , Y , and Z between maxima of Z may be found from Fig. 2, by interpolating between neighboring curves. Of course, the accuracy of predictions made by this method is limited by the exactness of Figs. 2 and 4, and, as we shall see, by the accuracy with which the initial values of X , Y , and Z are observed.

Some of the implications of Fig. 4 are revealed by considering an idealized two-to-one correspondence between successive members of sequences M_0, M_1, \dots , consisting of numbers between zero and one. These sequences satisfy the relations

$$\begin{aligned} M_{n+1} &= 2M_n && \text{if } M_n < \frac{1}{2} \\ M_{n+1} &\text{ is undefined} && \text{if } M_n = \frac{1}{2} \\ M_{n+1} &= 2 - 2M_n && \text{if } M_n > \frac{1}{2}. \end{aligned} \quad (35)$$

The correspondence defined by (35) is shown in Fig. 5, which is an idealization of Fig. 4. It follows from repeated applications of (35) that in any particular sequence,

$$M_n = m_n \pm 2^n M_0, \quad (36)$$

where m_n is an even integer.

Consider first a sequence where $M_0 = u/2^p$, where u is odd. In this case $M_{p-1} = \frac{1}{2}$, and the sequence terminates. These sequences form a denumerable set, and correspond to the trajectories which score direct hits upon the state of no convection.

Next consider a sequence where $M_0 = u/2^p v$, where u and v are relatively prime odd numbers. Then if $k > 0$, $M_{p+1+k} = u_k/v$, where u_k and v are relatively prime and u_k is even. Since for any v the number of proper fractions u_k/v is finite, repetitions must occur, and the sequence is periodic. These sequences also form a denumerable set, and correspond to periodic trajectories.

The periodic sequences having a given number of distinct values, or phases, are readily tabulated. In particular there are a single one-phase, a single two-phase, and two three-phase sequences, namely,

$$\begin{aligned} &2/3, \dots, \\ &2/5, 4/5, \dots, \\ &2/7, 4/7, 6/7, \dots, \\ &2/9, 4/9, 8/9, \dots. \end{aligned}$$

The two three-phase sequences differ qualitatively in that the former possesses two numbers, and the latter only one number, exceeding $\frac{1}{2}$. Thus the trajectory corresponding to the former makes two circuits about C , followed by one about C' (or vice versa). The trajectory corresponding to the latter makes three circuits about C , followed by three about C' , so that actually only Z varies in three phases, while X and Y vary in six.

Now consider a sequence where M_0 is not a rational fraction. In this case (36) shows that M_{n+k} cannot equal

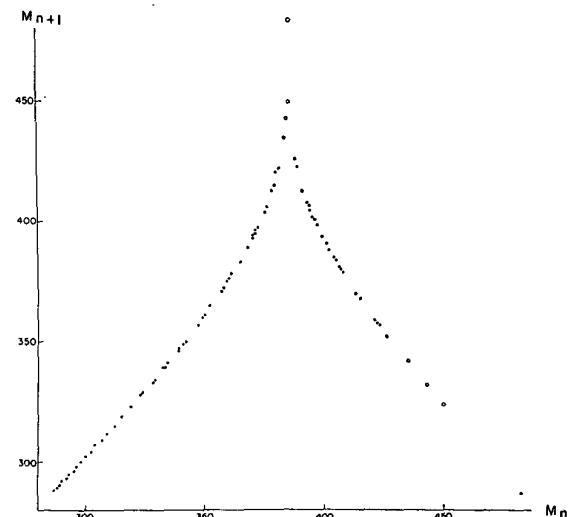


FIG. 4. Corresponding values of relative maximum of Z (abscissa) and subsequent relative maximum of Z (ordinate) occurring during the first 6000 iterations.

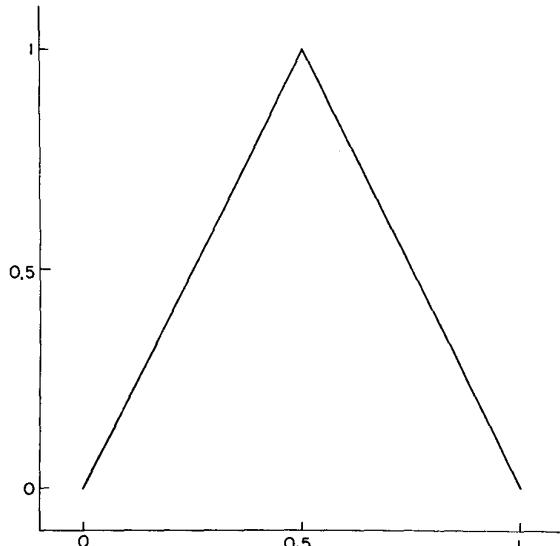


FIG. 5. The function $M_{n+1} = 2M_n$ if $M_n < \frac{1}{2}$, $M_{n+1} = 2 - 2M_n$ if $M_n > \frac{1}{2}$, serving as an idealization of the locus of points in Fig. 4.

M_n if $k > 0$, so that no repetitions occur. These sequences, which form a nondenumerable set, may conceivably approach periodic sequences asymptotically and be quasi-periodic, or they may be nonperiodic.

Finally, consider two sequences M_0, M_1, \dots and M'_0, M'_1, \dots , where $M'_0 = M_0 + \epsilon$. Then for a given k , if ϵ is sufficiently small, $M'_k = M_k \pm 2^k \epsilon$. All sequences are therefore unstable with respect to small modifications. In particular, all periodic sequences are unstable, and no other sequences can approach them asymptotically. All sequences except a set of measure zero are therefore nonperiodic, and correspond to nonperiodic trajectories.

Returning to Fig. 4, we see that periodic sequences analogous to those tabulated above can be found. They are given approximately by

$$\begin{aligned} & 398, \dots, \\ & 377, 410, \dots, \\ & 369, 391, 414, \dots, \\ & 362, 380, 419, \dots \end{aligned}$$

The trajectories possessing these or other periodic sequences of maxima are presumably periodic or quasi-periodic themselves.

The above sequences are temporarily approached in the numerical solution by sequences beginning at iterations 5340, 4881, 3625, and 3926. Since the numerical solution eventually departs from each of these sequences, each is presumably unstable.

More generally, if $M'_n = M_n + \epsilon$, and if ϵ is sufficiently small, $M'_{n+k} = M_{n+k} + \Lambda \epsilon$, where Λ is the product of the slopes of the curve in Fig. 4 at the points whose abscissas are M_n, \dots, M_{n+k-1} . Since the curve apparently has a slope whose magnitude exceeds unity everywhere, all sequences of maxima, and hence all trajectories, are unstable. In particular, the periodic trajectories, whose sequences of maxima form a denumerable set, are unstable, and only exceptional trajectories, having the same sequences of maxima, can approach them asymptotically. The remaining trajectories, whose sequences of maxima form a nondenumerable set, therefore represent deterministic nonperiodic flow.

These conclusions have been based upon a finite segment of a numerically determined solution. They cannot be regarded as mathematically proven, even though the evidence for them is strong. One apparent contradiction requires further examination.

It is difficult to reconcile the merging of two surfaces, one containing each spiral, with the inability of two trajectories to merge. It is not difficult, however, to explain the *apparent* merging of the surfaces. At two times τ_0 and τ_1 , the volumes occupied by a specified set of particles satisfy the relation

$$V_0(\tau_1) = e^{-(\sigma+b+1)(\tau_1-\tau_0)} V_0(\tau_0), \quad (37)$$

according to (30). A typical circuit about C or C' requires about 70 iterations, so that, for such a circuit,

$\tau_2 = \tau_1 + 0.7$, and, since $\sigma+b+1 = 41/3$,

$$V_0(\tau_1) = 0.00007 V_0(\tau_0). \quad (38)$$

Two particles separated from each other in a suitable direction can therefore come together very rapidly, and appear to merge.

It would seem, then, that the two surfaces merely appear to merge, and remain distinct surfaces. Following these surfaces along a path parallel to a trajectory, and circling C or C' , we see that each surface is really a pair of surfaces, so that, where they appear to merge, there are really four surfaces. Continuing this process for another circuit, we see that there are really eight surfaces, etc., and we finally conclude that there is an infinite complex of surfaces, each extremely close to one or the other of two merging surfaces.

The infinite set of values at which a line parallel to the X -axis intersects these surfaces may be likened to the set of all numbers between zero and one whose decimal expansions (or some other expansions besides binary) contain only zeros and ones. This set is plainly nondenumerable, in view of its correspondence to the set of all numbers between zero and one, expressed in binary. Nevertheless it forms a set of measure zero. The sequence of ones and zeros corresponding to a particular surface contains a history of the trajectories lying in that surface, a one or zero immediately to the right of the decimal point indicating that the last circuit was about C or C' , respectively, a one or zero in second place giving the same information about the next to the last circuit, etc. Repeating decimal expansions represent periodic or quasi-periodic trajectories, and, since they define rational fractions, they form a denumerable set.

If one first visualizes this infinite complex of surfaces, it should not be difficult to picture nonperiodic deterministic trajectories embedded in these surfaces.

8. Conclusion

Certain mechanically or thermally forced nonconservative hydrodynamical systems may exhibit either periodic or irregular behavior when there is no obviously related periodicity or irregularity in the forcing process. Both periodic and nonperiodic flow are observed in some experimental models when the forcing process is held constant, within the limits of experimental control. Some finite systems of ordinary differential equations designed to represent these hydrodynamical systems possess periodic analytic solutions when the forcing is strictly constant. Other such systems have yielded nonperiodic numerical solutions.

A finite system of ordinary differential equations representing forced dissipative flow often has the property that all of its solutions are ultimately confined within the same bounds. We have studied in detail the properties of solutions of systems of this sort. Our principal results concern the instability of nonperiodic solutions. A nonperiodic solution with no transient com-

ponent must be unstable, in the sense that solutions temporarily approximating it do not continue to do so. A nonperiodic solution with a transient component is sometimes stable, but in this case its stability is one of its transient properties, which tends to die out.

To verify the existence of deterministic nonperiodic flow, we have obtained numerical solutions of a system of three ordinary differential equations designed to represent a convective process. These equations possess three steady-state solutions and a denumerably infinite set of periodic solutions. All solutions, and in particular the periodic solutions, are found to be unstable. The remaining solutions therefore cannot in general approach the periodic solutions asymptotically, and so are nonperiodic.

When our results concerning the instability of nonperiodic flow are applied to the atmosphere, which is ostensibly nonperiodic, they indicate that prediction of the sufficiently distant future is impossible by any method, unless the present conditions are known exactly. In view of the inevitable inaccuracy and incompleteness of weather observations, precise very-long-range forecasting would seem to be non-existent.

There remains the question as to whether our results really apply to the atmosphere. One does not usually regard the atmosphere as either deterministic or finite, and the lack of periodicity is not a mathematical certainty, since the atmosphere has not been observed forever.

The foundation of our principal result is the eventual necessity for any bounded system of finite dimensionality to come arbitrarily close to acquiring a state which it has previously assumed. If the system is stable, its future development will then remain arbitrarily close to its past history, and it will be quasi-periodic.

In the case of the atmosphere, the crucial point is then whether analogues must have occurred since the state of the atmosphere was first observed. By analogues, we mean specifically two or more states of the atmosphere, together with its environment, which resemble each other so closely that the differences may be ascribed to errors in observation. Thus, to be analogues, two states must be closely alike in regions where observations are accurate and plentiful, while they need not be at all alike in regions where there are no observations at all, whether these be regions of the atmosphere or the environment. If, however, some unobserved features are implicit in a succession of observed states, two successions of states must be nearly alike in order to be analogues.

If it is true that two analogues have occurred since atmospheric observation first began, it follows, since the atmosphere has not been observed to be periodic, that the successions of states following these analogues must eventually have differed, and no forecasting scheme could have given correct results both times. If, instead,

analogues have not occurred during this period, some accurate very-long-range prediction scheme, using observations at present available, may exist. But, if it does exist, the atmosphere will acquire a quasi-periodic behavior, never to be lost, once an analogue occurs. This quasi-periodic behavior need not be established, though, even if very-long-range forecasting is feasible, if the variety of possible atmospheric states is so immense that analogues need never occur. It should be noted that these conclusions do not depend upon whether or not the atmosphere is deterministic.

There remains the very important question as to how long is "very-long-range." Our results do not give the answer for the atmosphere; conceivably it could be a few days or a few centuries. In an idealized system, whether it be the simple convective model described here, or a complicated system designed to resemble the atmosphere as closely as possible, the answer may be obtained by comparing pairs of numerical solutions having nearly identical initial conditions. In the case of the real atmosphere, if all other methods fail, we can wait for an analogue.

Acknowledgments. The writer is indebted to Dr. Barry Saltzman for bringing to his attention the existence of nonperiodic solutions of the convection equations. Special thanks are due to Miss Ellen Fetter for handling the many numerical computations and preparing the graphical presentations of the numerical material.

REFERENCES

- Birkhoff, G. O., 1927: *Dynamical systems*. New York, Amer. Math. Soc., Colloq. Publ., 295 pp.
- Ford, L. R., 1933: *Differential equations*. New York, McGraw-Hill, 264 pp.
- Fultz, D., R. R. Long, G. V. Owens, W. Bohan, R. Kaylor and J. Weil, 1959: Studies of thermal convection in a rotating cylinder with some implications for large-scale atmospheric motions. *Meteor. Monog.*, 4(21), Amer. Meteor. Soc., 104 pp.
- Gibbs, J. W., 1902: *Elementary principles in statistical mechanics*. New York, Scribner, 207 pp.
- Hide, R., 1958: An experimental study of thermal convection in a rotating liquid. *Phil. Trans. Roy. Soc. London*, (A), 250, 441–478.
- Lorenz, E. N., 1960: Maximum simplification of the dynamic equations. *Tellus*, 12, 243–254.
- , 1962a: Simplified dynamic equations applied to the rotating-basin experiments. *J. atmos. Sci.*, 19, 39–51.
- , 1962b: The statistical prediction of solutions of dynamic equations. *Proc. Internat. Symposium Numerical Weather Prediction*, Tokyo, 629–635.
- Nemytskii, V. V., and V. V. Stepanov, 1960: *Qualitative theory of differential equations*. Princeton, Princeton Univ. Press, 523 pp.
- Poincaré, H., 1881: Mémoire sur les courbes définies par une équation différentielle. *J. de Math.*, 7, 375–442.
- Rayleigh, Lord, 1916: On convective currents in a horizontal layer of fluid when the higher temperature is on the under side. *Phil. Mag.*, 32, 529–546.
- Saltzman, B., 1962: Finite amplitude free convection as an initial value problem—I. *J. atmos. Sci.*, 19, 329–341.