Applied Dynamical Systems

This file was processed February 16, 2018.

Contents

1 Introduction

1.1 About this course

Lecturer: Carl Dettmann, contact info at
`https://people.maths.bris.ac.uk/∼macpd/`

Unit home page: See
`https://people.maths.bris.ac.uk/∼macpd/ads/`

Small text (including footnotes) is supplementary non-examinable material. It will improve your understanding of the rest of the course material.

1.2 Books etc

None are essential, and there are very many other good books and internet resources available.

- J. C. Sprott, "Chaos and time series analysis," OUP 2003. Applied and computational approach, also with detail on a very large number of example models appearing in the literature.
- R. L. Devaney "An introduction to chaotic dynamical systems," Westview Press 2003. More rigorous, focused on discrete dynamical systems.
- B. Hasselblatt and A. Katok, "A first course in dynamics," CUP 2003. Also more rigorous. Leads to the more advanced "Introduction to the modern theory of dynamical systems" by the same authors.
- P. Cvitanović et al "Classical and quantum chaos" `www.chaosbook.org`. A free online book detailing periodic orbit methods for classical and quantum chaos; we are mostly interested in the introductory section (Part I).
- G. Teschl "Ordinary differential equations and dynamical systems" available online at `www.mat.univie.ac.at/~gerald/ftp/book-ode/`

You may also want reference/revision works on programming and numerical methods; online information and workshops in Bristol are available at `https://www.acrc.bris.ac.uk/acrc/training.htm` .

Finally, some of the footnotes refer to the primary research literature (journal articles etc). These may be accessed by searching for the journal website and then searching or browsing. Many require a subscription, so must be accessed from a university computer. Arxiv preprints are always available free, but may not (yet) have been refereed. A reference arxiv:1234.5678 refers to the url `arxiv.org/abs/1234.5678` .

1.3 Introduction

If there is a central idea in dynamical systems, it is probably that rather than describing the, often irregular, behaviour $\mathbf{x}(t)$ of some real world variable in time directly, scientific laws often correspond to determining how the state of the system varies, in the form

$$\dot{\mathbf{x}} = f(\mathbf{x})$$

for continuous time $t \in \mathbb{R}$ (the dot denotes differentiation with respect to time t), or

$$\mathbf{x}_{t+1} = \Phi(\mathbf{x}_t)$$

for discrete time $t \in \mathbb{Z}$. We can then hope to understand how the complicated function $\mathbf{x}(t)$ arises from the explicitly known $f(\mathbf{x})$ or $\Phi(\mathbf{x})$.

Dynamical systems really go back to Newton's laws of motion and gravitation (late 17th century), which correspond to ordinary differential equations describing the motions of N massive bodies such as planets:

$$m_i \ddot{\mathbf{q}}_i = \sum_{j \neq i} \frac{G m_i m_j}{|\mathbf{q}_i - \mathbf{q}_j|^3} (\mathbf{q_j} - \mathbf{q}_i)$$

where the index i is over masses, each $\mathbf{q}_i \in \mathbb{R}^3$ and $|\cdot|$ is Euclidean distance on $\mathbb{R}^3$. G is a Newton's gravitational constant. Any second order equation can be written as coupled first order equations by making $d\mathbf{q}_i/dt$ a separate variable. In mechanics we typically specify an initial value problem giving positions $\mathbf{q}_i$ and velocities $\dot{\mathbf{q}}_i$ or momenta $m_i \dot{\mathbf{q}}_i$ at the initial time. Newton's equations are an example of Hamiltonian dynamics, which has a number of consequences such as a conserved quantity (the energy) that we will discuss towards the end of the course.

The two body problem (eg sun and earth) has an exact solution, which describes well the main features of planetary motion (elliptical orbits etc). Laplace (early 19th century) realised some philosophical implications of such an approach:

> *"We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes."*

Mathematically, we could support this by noting that the ODEs together with initial conditions on positions and velocities satisfy conditions for the existence and uniqueness of solutions, at least for short times:

Definition 1.1. *A function $f : M \to M$ is* **Lipschitz continuous** *if there is a constant λ so that $d(f(x), f(y)) < \lambda d(x, y)$ for all x and y.*

Here, M is a metric space with distance $d(x, y)$; if $M = \mathbb{R}^n$, d is usually Euclidean distance. If we want to be more precise, we can describe f as λ-Lipshitz.

Theorem 1.2. Picard-Lindelöf theorem: *Given the initial value problem $x'(t) = f(x(t), t)$, $x(0) = x_0 \in \mathbb{R}^n$, if f is Lipshitz continuous in x and continuous in t, there exists $\epsilon > 0$ so that the solution exists and is unique for $t \in [-\epsilon, \epsilon]$.*

Thus, given regular initial conditions (ie masses at distinct locations), we are guaranteed to have a unique solution until/unless either there is a collision or particles escape to infinity at some finite time.[1] For more than two masses the solutions seemed more difficult to find. In 1887 the king of Sweden offered a prize for the solution to the three body problem, which was won by Poincaré who demonstrated its intractability, in the process developing many of the ideas of modern dynamical systems.[2]

The mathematics of dynamical systems continued to develop throughout the first half of the twentieth century, but the impetus for applications was the advent of computer simulation and visualisation. Lorenz (1963) observed for a simple model of the atmosphere, the **Lorenz equations**

$$\dot{x} = \sigma(y - x), \qquad \dot{y} = x(\rho - z) - y, \qquad \dot{z} = xy - \beta z$$

with constants $\sigma = 10$, $\beta = 8/3$, $\rho = 28$ that a small change in initial conditions led to drastic changes in behaviour at later times, the so-called "butterfly effect" in which a butterfly flapping its wings in Brazil may cause a hurricane some weeks later in Texas.[3]

Mitchell Feigenbaum, using only a hand-held calculator, discovered in 1975 the existence of new mathematical constants controlling the transition from regular to unpredictable behaviour in a whole class of discrete time dynamical systems, including the extremely simple-looking (and simplified) model arising from population biology, the **logistic map**

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

while in the same year, Li and Yorke published a paper[4] in which they showed

[1] Escape to infinity in finite time was shown to be possible for five or more masses in Z. Xia, Ann. Math. **135**, 411-468 (1992).

[2] Sensitivity to initial conditions was previously articulated by Maxwell in his 1873 essay on determinism and free will: *It is manifest that the existence of unstable conditions renders impossible the prediction of future events, if our knowledge of the present state is only approximate and not accurate.* Now, we understand that equations of motion such as those of Newton approximate a probabilistic quantum theory, which prevents exact initial conditions of position and velocity in principle, although classical equations are still a very good approximation for many processes involving more than a few elementary particles and many systems with a few particles can be partly understood "semi-classically," that is, by relating them to corresponding classical systems, the field of quantum chaos.

[3] The Lorenz equations are still under active investigation, for example new results for the mixing properties (covered later in this course) are found in V. Araujo, I. Melbourne and P. Varandas, Commun. Math. Phys. **340**, 901-938 (2015).

[4] T.-Y. Li and J. A. Yorke, "Period three implies chaos" Amer.

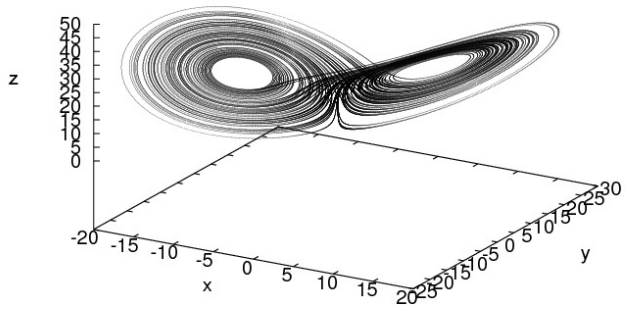

Figure 1: The Lorenz attractor

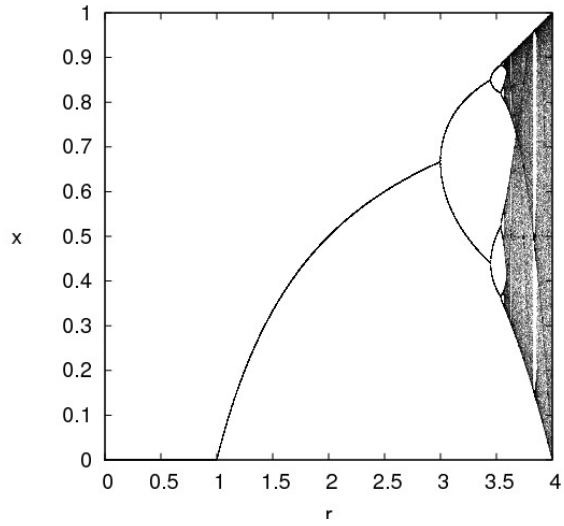

Figure 2: The bifurcation diagram of the logistic map

Theorem 1.3. Period three theorem: *If $I \subset \mathbb{R}$ is an interval and $\Phi : I \to I$ is continuous and has a period three point, then Φ has periodic points of all periods.*

Here, we use notation Φ^n to indicate the n-fold composition $\Phi \circ \Phi \circ \ldots \Phi$, so Φ^0 is the identity transformation and $\Phi^1 = \Phi$. A periodic point x is a point such that $\Phi^p(x) = x$, and its period is the smallest such $p \geq 1$.

This last theorem suggests that chaos is not only possible, it is pervasive. Ulam is quoted as saying

> *Using a term like nonlinear science is like referring to the bulk of zoology as the study of non-elephant animals.*

Math. Month. **82**, 985-992 (1975). The title is the first use of "chaos" in a dynamical context. Later, they discovered that their result was a special case of a result proved in O. Sharkovsky, Ukr. Mat. Zh. **16**, 61-71 (1964).

In fact, very few systems of at least this minimal size - three dimensions in continuous time, one in discrete time - are completely regular, although such solvable systems dominate many introductory courses in mechanics and other fields. A few other systems are completely chaotic, that we will discuss. Much useful understanding can be gained from small perturbations about both of these limits. But the "generic" situation of mixed regularity and chaos, is very incompletely understood.

In the same way as for dynamical systems, geometric structures now associated with chaotic systems were proposed by mathematicians in the late 19th century; note particularly Cantor sets, such as the set of real numbers of the form $\sum_j a_j 3^{-j}$ with $a_j \in \{0, 2\}$, which is uncountable, zero measure, nowhere dense, closed and totally disconnected. According to most methods of defining non-integer dimensions, it has dimension $\ln 2 / \ln 3 \approx 0.631$. Similarly the Koch snowflake (1904) is a nowhere differentiable continuous curve with infinite length and dimension $\ln 4 / \ln 3 \approx 1.262$. The relevance to many physical and biological phenomena was appreciated and popularised by Mandelbrot in the 1960s, again with the aid of computer graphics; he also coined the term **fractal**.

We discussed astronomy, meteorology and population biology as having examples of dynamical systems with interesting behaviour. Many of these have in fact many degrees of freedom, as do systems of many atoms (molecular biology and nanotechnology) and "complex systems"[5] such as social and financial networks. These kinds of systems can involve low dimensional dynamics at two levels - where they are built from microscopic interactions (although may have different collective behaviour), and where the macroscopic behaviour can be well described using only a few well-chosen variables. There is both theoretical and experimental evidence for this, for example "centre manifold theory of infinite dimensional systems"[6] and Libchaber's experiments in the early 80s on Rayleigh-Benard (convective) fluid systems, confirming Feigenbaum's theory for a transition to turbulence.[7]

Thus low dimensional dynamical systems is strongly relevant to current research in

- theory - mathematics, theoretical and computational physics
- few degree of freedom systems - astronomy, population biology

[5] Note "complex dynamics" often refers to dynamics on $\mathbb{C}$. This leads to such interesting objects as the Mandelbrot set, but occurs relatively rarely in applications and unfortunately we will not have time to discuss it. There is some discussion in the recommended texts.

[6] M. Haragus and G. Iooss, "Local bifurcations, center manifolds and normal forms in infinite-dimensional dynamical systems" Springer, 2010.

[7] A Libchaber, C Laroche, S Fauve. "Period doubling cascade in mercury, a quantitative measurement", Journal de Physique Lettres, **43**, 211-216 (1982).

- small quantum systems - quantum chaos
- effectively small systems - transition to turbulence, meteorology, chaotic lasers for secure communication and random number generation
- complex systems - nanoscience, biological, social, financial and communications networks

This is a too-brief summary. For example, biology is a vast source of dynamical problems at all levels considering interaction and movement of atoms, proteins, cells, tissues, organs, organisms and entire species.

In this course we will pose and answer some of the questions: Why does chaos appear in one-dimensional discrete time systems, but need three dimensions if the time is continuous? Why are some systems regular, while others are chaotic? How do fractals arise from dynamics and how to characterise them? What practical analytical and numerical methods are there for understanding dynamical systems and fractals?

Finally, we return to the fundamental questions about determinism and predictability. For regular systems a small perturbation of the initial conditions leads to only to bounded or slowly growing deviations in the trajectory, while for strongly chaotic systems the deviation grows exponentially. However if we consider not the trajectory itself but its average properties, these are perturbed for the regular system, but typically unchanged for the chaotic system. Thus a problem with weather prediction is the presence of chaos, while a problem for climate change prediction is the presence of regularity.

2 Dynamics and time

2.1 Maps and flows

We have seen two types of dynamical systems so far, differential equations such as Newton's or Lorenz's equations, and discrete time systems, such as the logistic map. In each case there is a on-parameter family of maps $\Phi^t : X \to X$ on the space X (say $\mathbb{R}^d$) which moves forward the dynamics in time by an amount t for any initial time s:[8]

$$\mathbf{x}(s+t) = \Phi^t(\mathbf{x}(s))$$

In the map case $t \in \mathbb{N}$ and $\Phi^1 = \Phi$ is just the original map, while for differential equations, $(d/dt)\mathbf{x} = f(\mathbf{x})$, $t \in \mathbb{R}$ and the Φ^t is a called a flow, found (if possible) by solving the differential equation for arbitrary initial condition $\mathbf{x}$:

$$\frac{d}{dt}\Phi^t(\mathbf{x}) = f \circ \Phi^t(\mathbf{x}), \qquad \Phi^0(\mathbf{x}) = \mathbf{x}$$

Example 2.1. *For the harmonic oscillator* $(d/dt)(x, v) = (v, -x)$ *we have* $\Phi^t(x, v) = (x\cos(t) + v\sin(t), -x\sin(t) + v\cos(t))$.

We can easily see that for either a map or flow, Φ satisfies the semigroup property:

$$\Phi^s \circ \Phi^t = \Phi^{s+t}$$

and hence

$$\Phi^0(\mathbf{x}) = \mathbf{x}$$

Saying a map is **invertible** means that $\Phi^{-1} : X \to X$ is uniquely defined, so the Φ^t form a group under composition.[9] Normally, flows are invertible from the Picard-Lindelöf theorem.[10] For either invertible or non-invertible maps, Φ^{-1} is a function on sets $A \subset X$

$$\Phi^{-1}(A) = \{x \in X : \Phi(x) \in A\}$$

giving the **pre-images** of a point or larger set.[11] Invertible systems may also be **(time)-reversible**; this means that there is a transformation $i : X \to X$ satisfying

$$i \circ \Phi^t \circ i = \Phi^{-t}$$

Setting $t = 0$ we note that i^2 is the identity, that is, i is an involution (hence the notation). For example, Newton's law of gravitation is reversible, with i reversing all the velocities.[12]

Often a system can have more than one description, using either a map or a flow. These can be related as follows: A **time-one** or **stroboscopic** map is obtained by considering a flow Φ^t and treating Φ^1 (or more generally some $\Phi^{\Delta t}$) as a map in its own right. An alternative, and probably more useful approach is the **Poincaré** map, defining a hypersurface Y and stopping whenever this is reached. If the time from one event to the next is

$$\tau(\mathbf{y}) := \min_{t>0}\{t : \Phi^t(y) \in Y\}$$

the Poincaré map can be defined as

$$F(\mathbf{y}) = \Phi^{\tau(\mathbf{y})}(\mathbf{y})$$

The reverse process is called a **suspension**: Given a map $F : Y \to Y$ and "roof function" $\tau : Y \to (0, \infty)$, we can construct a flow on the space $X = \{(\mathbf{y}, s) : \mathbf{y} \in Y, 0 \le s < \tau(\mathbf{y})\}$ in the natural way - increase s until $\tau(\mathbf{y})$ is reached, then apply F and set $s = 0$.

Example 2.2. *For the harmonic oscillator above, and the Poincaré section (hypersurface)* $x = 1$ *the Poincaré map is* $F(y) = -y$ *(taking* $y = v$*), but note that not all trajectories of the original flow reach it.*[13] *The roof function is*

$$\tau(y) = \begin{cases} 2\arctan|y| & y > 0 \\ 2\pi & y = 0 \\ 2\pi - 2\arctan|y| & y < 0 \end{cases}$$

A **dynamical billiard** consists of a point particle that moves freely except for mirror-like reflections with the boundary, that is, angle of reflection equals angle of incidence. The flow is discontinuous (in the momentum variable) at collisions, so it is natural to consider a Poincaré section consisting of the boundary.

An **induced map** is obtained starting from a map $\Phi : X \to X$ and a subset Y of full dimension, and proceeding in the same manner. This can help if (almost) all orbits pass through Y and the induced dynamics has more uniform properties.

Example 2.3. *The Farey map is*

$$\Phi(x) = \begin{cases} \frac{x}{1-x} & x < 1/2 \\ \frac{1-x}{x} & x > 1/2 \end{cases}$$

which has very slow behaviour near $x = 0$*. Iterating the left branch we find for small* x

$$\Phi_L^2(x) = \frac{\frac{x}{1-x}}{1 - \frac{x}{1-x}} = \frac{x}{1-2x}$$

[8] We assume here that the system is autonomous; more generally we could consider Φ a function of both s and t. Generalising dynamical results to non-autonomous systems is a popular source of research problems. In some cases it is useful to add the time as an extra variable.

[9] It is also possible to use a semi-group larger than $\mathbb{N}$ or $\mathbb{R}$ as a "time" variable; this comes under the name of "group actions," also a popular research topic. A famous example is Furstenberg's conjecture that the system $\Phi^{(i,j)}(x) = 2^i 3^j x \pmod 1$ has no nontrivial invariant measures. H. Furstenberg Mathematical Systems Theory **1** 1-49 (1967).

[10] Piecewise smooth flows, which need not be invertible, form another rich source of research problems. See for example M. di Bernado and S. J. Hogan, Trans. Roy. Soc. A **368**, 4915-4935 (2010).

[11] Sometimes set-valued dynamics is considered in the forward time direction, too.

[12] More on dynamical reversibility can be found in J. A. G. Roberts and G. R. W. Quispel, Phys. Rep. **216**, 63-177 (1992).

[13] If you start on the Poincaré section, there are often results that show you almost certainly return, for example the Poincaré recurrence theorem discussed in chapter 6.

$$\Phi_L^3(x) = \frac{x}{1-3x}$$

$$\Phi_L^n(x) = \frac{x}{1-nx}$$

Suppose we induce on the right branch, ie $Y = [1/2, 1)$. We take τ iterations to return to the right branch, ie

$$F(y) = \Phi_L^{\tau-1}(\Phi_R(y)) = \frac{\frac{1-y}{y}}{1-(\tau-1)\frac{1-y}{y}} = \frac{1-y}{\tau y-(\tau-1)}$$

Equivalently since we have alternating strings of $\tau-1$ iterations of Φ_L followed by a single Φ_R we can consider the dynamics immediately after the right branch iteration, ie

$$G(x) = \Phi_R(\Phi_L^{\tau-1}(x)) = \frac{1-\frac{x}{1-(\tau-1)x}}{\frac{x}{1-(\tau-1)x}} = \frac{1}{x} - \tau$$

This the famous Gauss map, and the τ values give the continued fraction expansion of x. It is much faster to calculate, and no longer has a region near zero that behaves very differently to the rest of the map, however we have replaced a map with two branches by one with an infinite number.

2.2 Numerical considerations

If we have an explicit equation for a map, it is straightforward to simulate it numerically, although there are issues to do with instability and finite precision that we will discuss later.

Similarly, a flow (ordinary differential equation) can be simulated using standard numerical techniques, which effectively approximate the stroboscopic map using small step size, for example the simplest (Euler) method for approximating $\dot{\mathbf{x}} = f(\mathbf{x})$ is

$$x_{t+h} = x_t + hf(x_t)$$

This requires extremely small step size for accuracy (which then takes longer, and suffers from round-off error); a slightly better algorithm is the midpoint method, which (approximately) calculates the derivative from the midpoint of the interval

$$x_{t+h} = x_t + hf(x_t + \frac{h}{2}f(x_t))$$

There are a variety of more accurate and reliable methods discussed in texts on numerical analysis and used by numerical software.[14]

In order to compute a Poincaré map, we need to do more work, however. We need to simultaneously solve the differential equation, and an algebraic equation in a single variable, the time. Fast methods for algebraic equations $g(t) = 0$ include Newton's method

$$t_{n+1} = t_n - \frac{g(t_n)}{g'(t_n)}$$

— another source of discrete dynamical systems. This converges quadratically (double the number of digits at each step) for good initial guesses, but no guarantee of convergence otherwise.[15] If there is a change of sign, $g(t_1)g(t_2) < 0$, the bisection method is guaranteed to find a solution, with linear convergence. There may however be several changes of sign in the initial interval. So, ideally there should also be a rigorous lower bound on the time.

Example 2.4. *Find the smallest positive solution of $at = \sin t$ for $0 < a < 1$. Choose t_0 to be a small positive value. Since $g(t) = \sin(t) - at$, we know $g''(t) \geq -1$ and so, integrating twice*

$$g(t) \geq -(t-t_n)^2/2 + (t-t_n)g'(t_n) + g(t_n)$$

Thus the desired solution of $g(t) = 0$ is greater than the smallest root (greater than t_n) of this quadratic equation. Also, since we match both the value of $g(t_n)$ and its derivative, the iteration will be quadratically convergent like the Newton method.

An alternative approach is a simple change of variable. Suppose we have a system with d variables, and we can write the system in the form

$$\frac{dg}{dt} = f_1(g, x_2, \ldots, x_d)$$

$$\frac{dx_j}{dt} = f_j(g, x_2, \ldots, x_d), \qquad j \geq 2$$

where the Poincaré section is at $g = 0$. Then, making g the independent variable we have

$$\frac{dt}{dg} = \frac{1}{f_1}$$

$$\frac{dx_j}{dg} = \frac{dx_j}{dt}\frac{dt}{dg} = \frac{f_j}{f_1}$$

In this form we can integrate directly to $g = 0$.

[14] Accuracy is improved by combining linear combinations of the function evaluated at different times, the Runge-Kutta methods. Stability is improved by including x_{t+h} on the RHS, requiring Newton's method or similar at each step, the implicit methods. Often these approaches are combined.

[15] Intriguingly, multiplying the last term by a random variable may sometimes be guaranteed to converge almost surely: H. Sumi arxiv:1608.05230.

3 Local dynamics

3.1 Linearised dynamics

Start with the orbit of a one dimensional map, $x(n+1) = \Phi(x(n))$, assumed to be smooth. We consider a perturbed orbit $x(n+1) + \delta(n+1) = \Phi(x(n) + \delta(n))$. A Taylor expansion gives

$$\Phi(x+\delta) = \Phi(x) + \Phi'(x)\delta + O(\delta^2)$$

and so to linear order in $\delta(0)$

$$\delta(n+1) = \Phi'(x(n))\delta(n)$$

$$\begin{aligned}\delta(n) &= \left(\prod_{j=0}^{n-1} \Phi'|_{x(j)}\right)\delta(0)\\ &= \left(\prod_{j=0}^{n-1} \Phi' \circ \Phi^j|_{x(0)}\right)\delta(0)\\ &= (D\Phi^n)|_{x(0)}\delta(0)\end{aligned}$$

Where $D\Phi^n$ is the derivative of Φ^n.

We can do this also in d dimensions: $x_j(n+1) = \Phi_j(\mathbf{x}(n)), \quad j \in \{1,..,d\}$. A Taylor expansion gives

$$\Phi_j(\mathbf{x}+\delta) = \Phi_j(\mathbf{x}) + \sum_i \frac{\partial \Phi_j}{\partial x_i}\delta_i + O(\delta^2)$$

and so (again to linear order)

$$\delta(n+1) = (D\Phi)\delta(n)$$

$$\begin{aligned}\delta(n) &= (D\Phi)|_{\mathbf{x}(n-1)} \dots (D\Phi)|_{\mathbf{x}(1)}(D\Phi)|_{\mathbf{x}(0)}\delta(0)\\ &= (D\Phi^n)|_{\mathbf{x}(0)}\delta(0)\end{aligned}$$

Now $D\Phi$ is the Jacobian matrix of derivatives evaluated along the orbit, with the sum over i giving matrix multiplications.[16] Note the convention $(D\Phi)_{ji} = \partial_i \Phi_j$.

In the case of a fixed point, we multiply by the same matrix each time, giving $(D\Phi)^n$. If all the eigenvalues of $D\Phi$ are of magnitude less than one, it can be shown that the fixed point is **asymptotically stable** (see the textbook K&H, Lemma 3.6), that is, all orbits in a neighbourhood of it approach it. Note that $D\Phi$ need not contract all vectors, but sufficiently high powers of it do.

Example 3.1. *The matrix*

$$M = \begin{pmatrix} 1 & -1 \\ 1/2 & 0 \end{pmatrix}$$

has eigenvalues

$$\lambda = \frac{1 \pm i}{2}$$

which are both of magnitude less than unity. However it expands the vector $(1,0)^T$ *to* $(1,1/2)^T$*, where superscript* T *denotes transpose.*

This can be expressed concisely in terms of the spectral **matrix norm**

$$\|M\| = \sup_{\mathbf{v}:|\mathbf{v}|=1} |M\mathbf{v}|$$

where $\mathbf{v}$ is a vector. This norm also gives the square-root of the largest eigenvalue of $M^T M$.

We see that $\|M\| > 1$ but $\|M^n\| < 1$ for all sufficiently large n.[17]

Example 3.2. *Find the fixed points and corresponding linearised dynamics for the logistic map* $\Phi(x) = rx(1-x)$ *and determine their stability. A fixed point satisfies* $x = \Phi(x)$ *so we have*

$$x = rx(1-x)$$

$$rx^2 - rx + x = 0$$

$$x = 0, \frac{r-1}{r}$$

A small perturbation around fixed point x^* *evolves to linear order as*

$$\delta(n) = \Phi'(x^*)^n \delta(0)$$

We have $\Phi'(x) = r(1-2x)$*. For* $x^* = 0$ *we have* $\Phi'(x^*) = r$ *so it is stable for* $-1 < r < 1$ *(normally we consider only* $0 \le r \le 4$*). At* $r = \pm 1$ *the fixed point is "marginal" and we cannot determine its stability from a linear analysis. For* $|r| > 1$ *it is unstable;* $\delta(n)$ *grows exponentially, until it is large enough for the linear theory to break down. For* $x^* = (r-1)/r$ *we have* $\Phi'(x^*) = 2 - r$*. Thus it is stable for* $1 < r < 3$*. We see that at* $r = 1$ *both fixed points coincide at* $x^* = 0$ *and change their stability; this is an example of a bifurcation.*

In the case of a periodic point of order p, we can apply the same analysis to Φ^p, the p-composed map, thus we need to study $D\Phi^p$. Note that eigenvalues are invariant under cyclic permutations of matrix products, so we get similar behaviour starting from any of the points $\{\mathbf{x}, \Phi(\mathbf{x}), \Phi^2(\mathbf{x}), \dots \Phi^{p-1}(\mathbf{x})\}$.

For a flow $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ the same analysis gives, again to linear order in $\delta(0)$

$$\frac{d}{dt}\delta(t) = (Df)\delta(t)$$

which may be integrated (analytically or numerically) to

$$\delta(t) = (D\Phi^t)\delta(0)$$

[16] Hence a connection with another active research area, products of random matrices

[17] In this situation a natural approach is to redefine the norm to align with the eigenvectors.

where Φ^t is the corresponding flow. Its derivatives satisfy the matrix differential equation

$$\frac{d}{dt}(D\Phi^t) = (Df \circ \Phi^t)(D\Phi^t) \qquad D\Phi^0 = I$$

which in general must be integrated along the trajectory, an extra d^2 equations.

We can similarly specialise to the case of fixed points ($\Phi^t(\mathbf{x}) = \mathbf{x}$ for all t). Here, we see that

$$D\Phi^t = \exp(t(Df))$$

so that the eigenvalues of $D\Phi^t$ are the exponentials of those of Df.

We can also consider periodic orbits ($\Phi^T(\mathbf{x}) = \mathbf{x}$ and the period is the smallest positive such T). A (non-fixed) periodic orbit may be analysed using $D\Phi^T$, which always has one eigenvalue equal to one corresponding to the flow direction. If all other eigenvalues are of magnitude less than one, the periodic orbit is asymptotically stable and is called a **limit cycle** (K&H, Sec 2.4.3).

Actually, linear and nonlinear dynamics are related for many unstable situations also:

Definition 3.3. *Two dynamical systems (either flows or maps) with Φ^t and Ψ^t are conjugate if there is an invertible function h satisfying*

$$h \circ \Phi^t = \Psi^t \circ h$$

Topological conjugacy requires that h and its inverse be continuous. **Smooth conjugacy** requires that h and its inverse be differentiable (eg C^k). **Local conjugacy** requires that the relation holds only in the neighbourhood of a specific point. A related concept is

Definition 3.4. *Two flows Φ^t and Ψ^t are orbit equivalent if they are related by*

$$h(\Phi^t(h^{-1}(\mathbf{x}))) = \Psi^{\alpha(t,\mathbf{x})}(\mathbf{x})$$

for h invertible and α an increasing function of t.

Theorem 3.5. *(Hartman-Grobman theorem) If a differential equation has a fixed point with Jacobian Df having all eigenvalues with non-zero real part ("hyperbolic"), there is a local topological conjugacy between the linear and nonlinear flows.*[18]

Notes: The change of variables is continuous (normally Hölder continuous) — but derivatives may not match or even exist.[19] The non-zero real part is equivalent to $D\Phi^t$ for the flow having magnitude not equal to one. The equivalent statement also holds for diffeomorphisms (ie differentiably invertible maps).

Warning: The word hyperbolic is used in several different senses in dynamical systems and between authors. The most common other usage would require eigenvalues with real part less and greater than zero, or magnitude less and greater than one, as appropriate.

We can also relate the stability of a flow $\Phi^t(\mathbf{x})$ to its Poincaré map $\Phi^{\tau(\mathbf{x})}(\mathbf{x})$. The chain rule gives

$$\begin{aligned}\delta_j(n+1) &= \sum_i \frac{d}{dx_i}\Phi_j^{\tau(\mathbf{x}(n))}(\mathbf{x}(n))\delta_i(n) \\ &= \sum_i \left[\frac{\partial\tau}{\partial x_i} f(\Phi_j^\tau) + \frac{\partial \Phi_j^\tau}{\partial x_i}\right]\Bigg|_{\mathbf{x}(n)} \delta_i(n)\end{aligned}$$

The second term evolves the system by a fixed τ, however τ is in general a function of $\mathbf{x}$ so this is adjusted by the flow term f to project the perturbation to the Poincaré surface Y.

Example 3.6. *If we consider the previous examples, harmonic oscillator with Poincaré section at $x = 1$, $y = v$, we have as before, $\Phi^t(x, y) = (x\cos t + y\sin t, -x\sin t + y\cos t)$ and $\tau(y) = 2\arctan y, \quad y > 0$. Using the y-component of Φ^t and substituting $x = 1$ we find*

$$\begin{aligned}\frac{\delta(n+1)}{\delta(n)} &= \frac{d\tau}{dy}\frac{d}{d\tau}\Phi^\tau(y) + \frac{d}{dy}\Phi^\tau(y) \\ &= \frac{2}{1+y^2}(-1\cos\tau - y\sin\tau) + 1\cos\tau \\ &= \frac{2}{1+y^2}\left[-\frac{1-y^2}{1+y^2} - \frac{2y^2}{1+y^2}\right] + \frac{1-y^2}{1+y^2} \\ &= -1\end{aligned}$$

which is just the derivative of the Poincaré map $F(y) = -y$.

If the surface Y is given by the solution of

$$g(\mathbf{x}) = 0$$

we can find the above derivatives of τ by implicit differentiation of

$$g(\Phi^{\tau(\mathbf{x})}(\mathbf{x})) = 0$$

with respect to the coordinates.[20] It is clear that (assuming Y is transverse to the orbit and smooth) exponential growth or decay of a perturbation of a periodic orbit of the flow evolves at the same rate for the fixed point of the corresponding Poincaré map.

[18] See Teschl; also note that global versions exist, for example in P. Zgliczynski arXiv:1405.6733

[19] In two dimensions, life is generally C^1 smooth, though: See D. Stowe J. Diff. Eq. **63**, 183-226 (1986) for this, and also a discussion of the resonance conditions (relations between eigenvalues) that inhibit smoothness.

[20] An example where the equation for τ is not analytically solvable, but the Jacobian may be obtained by implicit differentiation is given in J. Lloyd, M. Niemeyer, L. Rondoni and G. P. Morriss, Chaos **5**, 536-551 (1995).

3.2 Linear dynamics

Having seen we can approximate dynamics in the vicinity of fixed (and periodic) points by linearisation, we now need to classify such behaviour.

Consider the map

$$\mathbf{x}_{t+1} = A\mathbf{x}_t$$

where A is a $d \times d$ matrix. We can change coordinates $\mathbf{x} = C\mathbf{y}$ (with C invertible) so that

$$\mathbf{y}_{t+1} = B\mathbf{y}_t$$

with $B = C^{-1}AC$. We choose C so that B is in **Jordan normal form**, that is, zero except for eigenvalues on the diagonal, with possibly 1's immediately above the diagonal where there are equal eigenvalues. Specifically, the columns of C are right eigenvectors of A, solutions of

$$(A - \lambda I)\mathbf{v} = 0$$

If the algebraic multiplicity is greater than the geometric multiplicity, ie not enough linearly independent eigenvectors exist, we find generalised eigenvectors using

$$(A - \lambda I)\mathbf{v}_{j+1} = c\mathbf{v}_j$$

to give a c above the diagonal. The standard Jordan normal form has $c = 1$, but it is also convenient to use $c = \lambda$. Finally if there are complex conjugate eigenvalues and eigenvectors it is convenient to use the real combinations $(\mathbf{v} + \mathbf{v}^*)/2$, $(\mathbf{v} - \mathbf{v}^*)/(2i)$ which lead to a 2×2 block in B of the form

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$$

Alternatively, we can consider the differential equation

$$\dot{\mathbf{y}} = D\mathbf{y}$$

with D in real Jordan normal form, having the solution

$$\mathbf{y}(t) = e^{Dt}\mathbf{y}(0)$$

where the matrix exponential can be defined using the usual power series. Putting $t = 1$ we arrive back at the map case. The exponential of a matrix in Jordan block form can be found explicitly, for example

$$\exp\begin{pmatrix} \lambda t & t \\ 0 & \lambda t \end{pmatrix} = \begin{pmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{pmatrix}$$

Thus we reduce to the map case, with the main difference being that the solution is continuous in t. The boundary between stable and unstable for a flow is where the eigenvalues of D cross the imaginary axis, while for a map it is the exponential of this, ie the unit circle.

Case: $d = 1$. $A = \lambda$ (scalar) and so $x_t = \lambda^n x_0$. The fixed point at 0 is stable if $|\lambda| < 1$, marginal if $|\lambda| = 1$ and unstable if $|\lambda| > 1$. The orbit remains on one side of the fixed point if $\lambda > 0$ and flips if $\lambda < 0$.

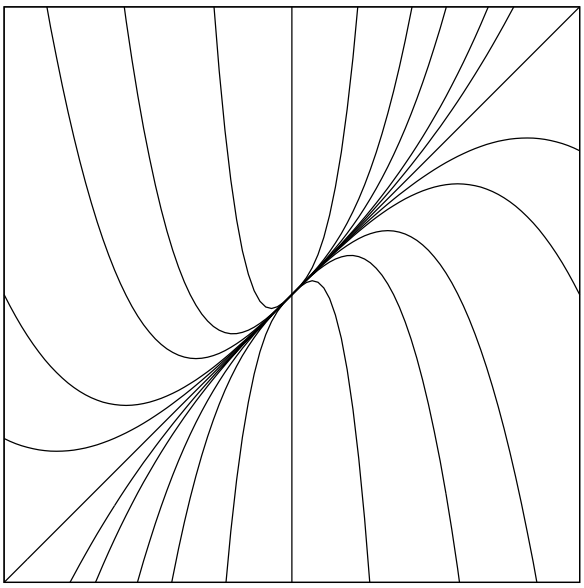

Figure 3: Node

Case: $d = 2$, distinct real eigenvalues λ and μ with $|\lambda| \geq |\mu|$. We have

$$B = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$

$$y_1(n) = \lambda^n y_1(0), \qquad y_2(n) = \mu^n y_2(0)$$

thus points lie on the invariant curves

$$|y_2| = c|y_1|^{\ln|\mu|/\ln|\lambda|}$$

with

$$c = y_2(0)y_1(0)^{\ln|\mu|/\ln|\lambda|}$$

in the map case, and consist of a branch of these curves (ie choice of signs) in the flow case. These are named

- $|\lambda| > 1, |\mu| > 1$: **Unstable node**
- $|\lambda| > 1, |\mu| < 1$: **Saddle**
- $|\lambda| < 1, |\mu| < 1$: **Stable node**

When one of the eigenvalues has magnitude one, the invariant curves become parallel.

Case: $d = 2$, complex conjugate eigenvalues $\lambda e^{\pm i\omega}$. In this case we use the real form of the Jordan normal form.

$$B = \lambda\begin{pmatrix} \cos\omega & \sin\omega \\ -\sin\omega & \cos\omega \end{pmatrix}$$

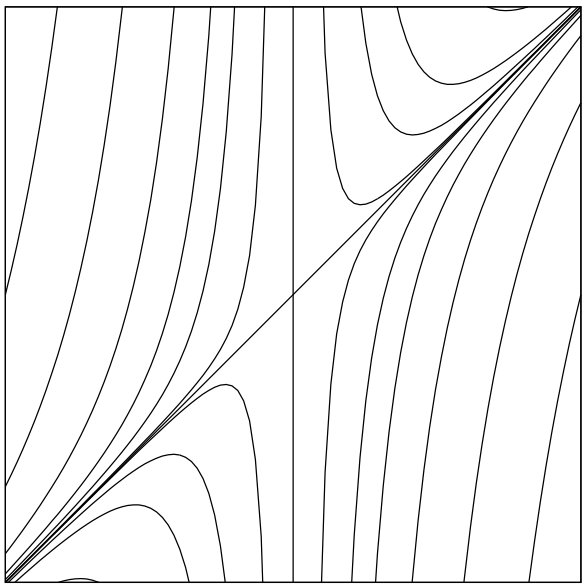

Figure 4: Saddle

so that

$$B^n = \lambda^n \begin{pmatrix} \cos n\omega & \sin n\omega \\ -\sin n\omega & \cos n\omega \end{pmatrix}$$

The invariant curves are spirals, and the fixed point is called a (stable or unstable) **focus**. If $\lambda = 1$, the map is a rotation by angle ω and is called a **centre**. In this case its properties depend sensitively on ω: If ω is a rational multiple of π, all orbits are periodic, while if not, all orbits fill the invariant curves densely and uniformly.

Case: $d = 2$, Equal eigenvalues, proportional to the identity

$$B = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$$

The map now just scales all directions equally, and the invariant curves are directed radially.

Case: $d = 2$, Equal eigenvalues, otherwise. We can transform the problem to

$$B = \begin{pmatrix} \lambda & \lambda \\ 0 & \lambda \end{pmatrix}$$

from which we find

$$B^n = \lambda^n \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}$$

Thus

$$y_1(n) = \lambda^n (y_1(0) + n y_2(0))$$

$$y_2(n) = \lambda^n y_2(0)$$

Solving the second equation for n and substituting into the first, we find

$$y_1(n) = y_2(n) \left[\frac{y_1(0)}{y_2(0)} + \frac{\ln(y_2(n)/y_2(0))}{\ln \lambda} \right]$$

which is a (stable or unstable) **degenerate node**. When $\lambda = \pm 1$ we have (solving from the beginning again)

$$y_1(n) = (\pm 1)^n y_1(0) + n y_2(n)$$

which is a **shear**.

A similar analysis can be carried out in higher dimensions.

Example 3.7. *Classify the fixed point of the map* $(x_1, x_2) \to (3x_1 + x_2, x_2 - x_1)$ *and give an explicit expression for the* n*th iterate. We have*

$$A = \begin{pmatrix} 3 & 1 \\ -1 & 1 \end{pmatrix}$$

There is a repeated eigenvalue $\lambda = 2$ *but only a single eigenvector* $\mathbf{v}_1 = (1, -1)^T$. *Thus we find a generalised eigenvector*

$$(A - \lambda I)\mathbf{v}_2 = 2\mathbf{v}_1$$

$$\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \mathbf{v}_2 = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

for example $\mathbf{v}_2 = (2, 0)^T$. *Thus we can use*

$$C = \begin{pmatrix} 1 & 2 \\ -1 & 0 \end{pmatrix}$$

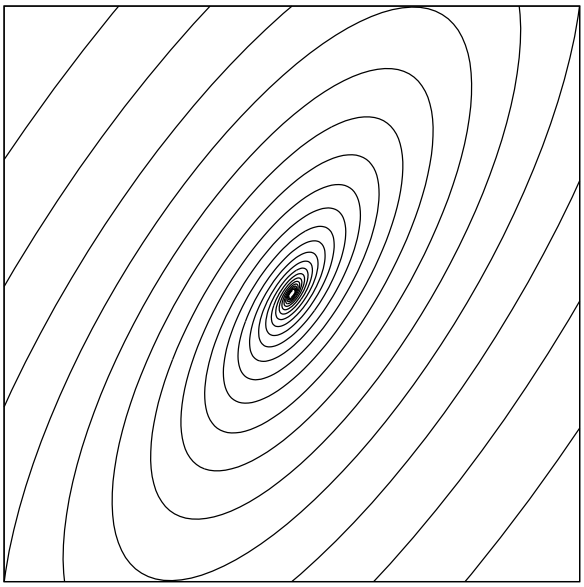

Figure 5: Focus

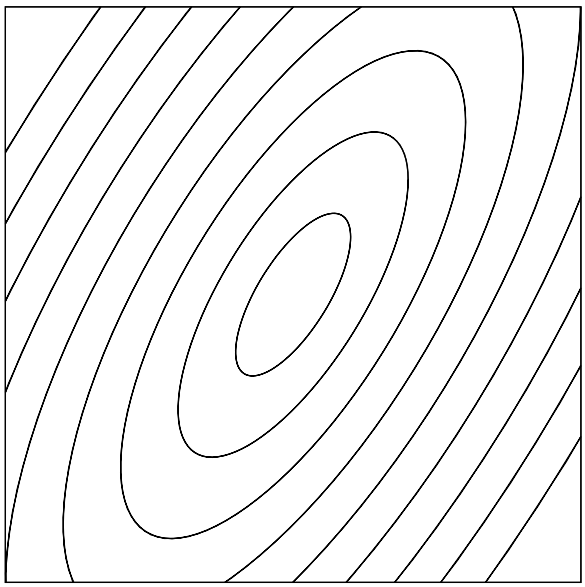

Figure 6: Centre

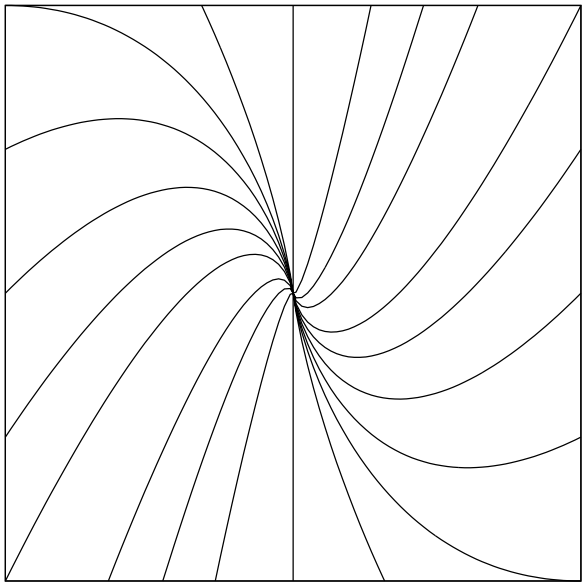

Figure 7: Degenerate node

which reduces the problem to:

$$B = C^{-1}AC = \begin{pmatrix} 2 & 2 \\ 0 & 2 \end{pmatrix}$$

The general solution is (similar to above)

$$\mathbf{y}(n) = 2^n \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \mathbf{y}(0)$$

Using C and C^{-1} to transform back, we find

$$\mathbf{x}(n) = 2^{n-1} \begin{pmatrix} n+2 & n \\ -n & -n+2 \end{pmatrix} \mathbf{x}(0)$$

This is a degenerate unstable node.

Example 3.8. *Classify the fixed point of the damped harmonic oscillator $\dot{x} = v, \quad \dot{v} = -x - \alpha v$ for each possible α. The matrix*

$$D = \begin{pmatrix} 0 & 1 \\ -1 & -\alpha \end{pmatrix}$$

has eigenvalues $\lambda = (-\alpha \pm \sqrt{\alpha^2 - 4})/2$.

For $\alpha = 0$ we have the undamped harmonic oscillator, which is a centre. For $0 < \alpha < 2$ (underdamped) we have two complex conjugate eigenvalues with negative real part; this is a stable focus. For $\alpha > 2$ (overdamped) we have two negative real eigenvalues, a stable node. For $\alpha = 2$ (critical damping) we have both eigenvalues equal to -1. We need to check the geometric multiplicity. An eigenvector satisfies

$$(D - \lambda I)\mathbf{v} = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \mathbf{v} = 0$$

We see there is only a single eigenvector $(1, -1)^T$. Thus we have a stable degenerate node. For negative α (unphysical) we have similar but unstable behaviour: An unstable focus for $-2 < \alpha < 0$, an unstable degenerate node for $\alpha = -2$ and an unstable node for $\alpha < -2$.

3.3 Local bifurcations

Generically we expect the eigenvalues of $D\Phi^t$ for a fixed or periodic point to be hyperbolic, ie differ from magnitude one, and hence expect node, saddle and focus behaviour unless there is a good reason, eg a symmetry in the model. An important exception to this, namely, Hamiltonian dynamics, will be considered at the end of the course.

Hyperbolic fixed or periodic points have the important property of **structural stability**, that is, all sufficiently small perturbations[21] the perturbed system is locally topologically conjugate to the original for maps, or orbit equivalent for flows. More precisely, consider a map $\Phi_\mu : \mathbb{R}^d \to \mathbb{R}^d$ depending on a parameter $\mu \in \mathbb{R}$ and C^1 in $\mathbf{x}$ and μ. Then, by applying the implicit function theorem to the equation $\Psi(\mathbf{x}, \mu) = \Phi_\mu(\mathbf{x}) - \mathbf{x} = 0$ we have

[21] in a specified topology; we need at least C^1 to ensure the Jacobian exists and is close to that of the unperturbed system.

Theorem 3.9. *(Lack of bifurcation) Suppose x_0 is a hyperbolic fixed point of a map Φ_{μ_0} (all eigenvalues of $D\Phi_{\mu_0}$ have magnitude different from unity), then there are neighbourhoods U of μ_0 and V of x_0 and a map $X : U \to V$ so that $x \in V$ is a fixed point of Φ_μ if and only if $x = X(\mu)$.*

Apparently the implicit function theorem is only concerned with eigenvalues equal to one, however a change in stability can be associated with any eigenvalue of magnitude one, and we also want to describe higher iterates of the map, for example an eigenvalue of -1 may be associated with a period doubling (below). For a flow, fixed points may have hyperbolic eigenvalues, however for limit cycles there is always a unit eigenvalue corresponding to the flow direction. In this case it is possible to apply the above theorem to a transverse[22] Poincaré map; the period of the orbit will in general vary with μ.

Conversely, cases where one or more eigenvalues has magnitude one are sensitive to small effects due to nonlinearity or varying a parameter. As mentioned before, bifurcations are variations in qualitative behaviour of a dynamical system due to variation of a parameter. Local bifurcations are due to eigenvalue(s) of $D\Phi^t$ for a fixed or periodic point reaching magnitude one.

In order to analyse bifurcations, we perform Taylor series expansions[23] in both the dynamical variable(s) (eg x) and parameter (eg μ), locating the fixed point without loss of generality at $x = 0$ for all μ and the bifurcation point at $\mu = 0$. As with the linear maps above, coordinate changes (in general nonlinear) can be used to reduce many problems to certain "normal forms", including forcing higher order terms to vanish.[24]

Now, we restrict to one-dimensional maps and flows. The most fundamental bifurcation is that of the **fold** (also called tangent bifurcation) with normal forms

$$\dot{x} = \mu - x^2$$

$$x_{n+1} - x_n = \mu - x_n^2$$

for flow and map respectively. Here μ is the parameter. For $\mu < 0$ there are no fixed points, while for $\mu > 0$ there are two, at $x = \pm\sqrt{\mu}$. Differentiating the right hand sides, we find to linear order

$$\dot{\delta} = -2x\delta$$

$$\delta_{n+1} = (1 - 2x)\delta_n$$

[22] That is, not parallel to the flow.

[23] We assume here that the dynamical system has sufficient smoothness (continuous derivatives). An active area of study is that of bifurcations for non-smooth systems, see eg di Bernado and Hogan, Phil. Trans. Roy Soc A **368**, 4915-4935 (2010).

[24] Mathematical justification for this kind of approach comes from centre manifold theory; see Wiggins, *Introduction to applied nonlinear dynamical systems and chaos*, Springer 2003

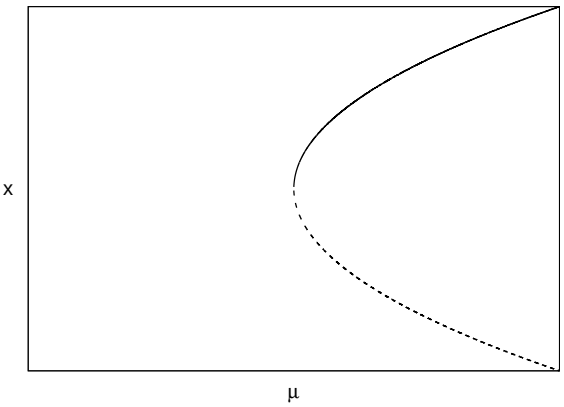

Figure 8: The fold bifurcation. Here, and in the later figures, solid lines indicate stable fixed points and dotted lines unstable fixed points.

Thus the fixed point at $x = \sqrt{\mu}$ is stable, while the fixed point at $x = -\sqrt{\mu}$ is unstable.

This bifurcation is itself structurally stable: A small[25] perturbation in the family of systems can be mapped back to the original by a topological conjugacy together with a reparametrisation of μ. Its generality is represented by a theorem, proved using the implicit function theorem:[26]

Theorem 3.10. *(Fold bifurcation) Suppose*

1. $\Phi_{\mu_0}(0) = 0$
2. $\Phi'_{\mu_0}(0) = 1$
3. $\Phi''_{\mu_0}(0) \neq 0$
4. $\frac{\partial\Phi}{\partial\mu}_{\mu_0}(0) \neq 0$

Then there is an interval I about zero and a smooth function $p : I \to \mathbb{R}$ such that

$$\Phi_{p(x)}(x) = x$$

as well as $p(0) = \mu_0$, $p'(0) = 0$, $p''(0) \neq 0$.

Thus for every x close to the bifurcation value (here, zero) there is a μ giving x as a fixed point. The function has a quadratic shape, giving no fixed points on one side of μ_0 and two on the other. Notice we have used the implicit function theorem with different variables here - for the "no-bifurcation" theorem we could vary μ and always find a fixed point $\mathbf{x}$, while here we vary x and find a parameter $\mu = p(x)$ for which it is fixed. As we know, the μ on one side of μ_0 have no fixed points.

The **transcritical** bifurcation has normal form

$$\dot{x} = \mu x - x^2$$

[25] That is, C^2

[26] See eg Devaney.

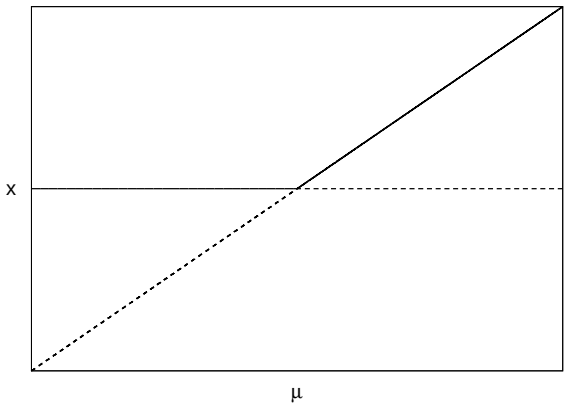

Figure 9: The transcritical bifurcation.

$$x_{n+1} - x_n = \mu x_n - x_n^2$$

from which we find fixed points at $x = 0$ and $x = \mu$, which coincide at $\mu = 0$. To linear order we have

$$\dot{\delta} = (\mu - 2x)\delta$$

$$\delta_{n+1} = (1 + \mu - 2x)\delta_n$$

so that $x = 0$ is stable for $\mu < 0$ while $x = \mu$ is stable for $\mu > 0$. The two fixed points thus switch stability at the bifurcation. This bifurcation is not structurally stable unless restricted to dynamics exhibiting a fixed point for an interval around $\mu = 0$.

The **supercritical pitchfork** bifurcation has normal form

$$\dot{x} = \mu x - x^3$$

$$x_{n+1} - x_n = \mu x_n - x_n^3$$

from which we find fixed points at $x = 0$ and (for $\mu > 0$) $\pm\sqrt{\mu}$. To linear order we have

$$\dot{\delta} = (\mu - 3x^2)\delta$$

$$\delta_{n+1} = (1 + \mu - 3x^2)\delta_n$$

so that $x = 0$ is stable for $\mu < 0$ and the two other fixed points are stable for $\mu > 0$. The **subcritical** case has a plus in the first equation, leading to two fixed points that are unstable and exist at $\mu < 0$. This bifurcation is not structurally stable unless restricted to systems with odd symmetry.

In addition, maps can have fixed points with eigenvalue -1 which has no analogue for flows. This leads to a bifurcation unique to maps (or Poincaré sections of flows), the **flip** or **period doubling** bifurcation, with normal form

$$-x_{n+1} - x_n = \mu x_n - x_n^3$$

Note the extra minus on the left hand side. Again, there is a fixed point at $x = 0$ which is stable for $\mu < 0$. There

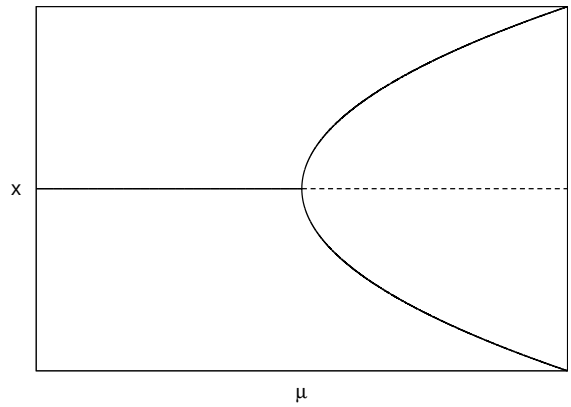

Figure 10: The supercritical pitchfork bifurcation. With the solid curves as period two orbits, it represents a period doubling bifurcation. With the solid curves as stable limit cycles, it represents a Hopf bifurcation.

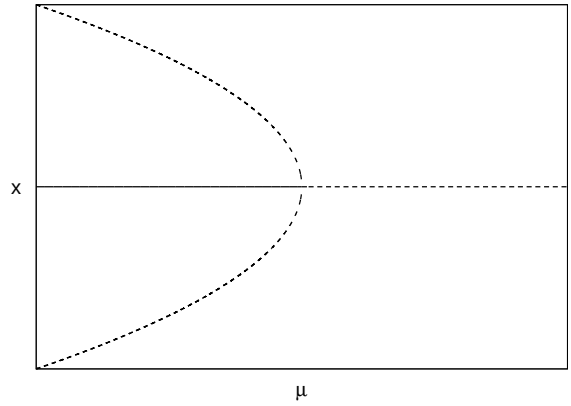

Figure 11: The subcritical pitchfork bifurcation (or period doubling, or Hopf, as in Fig. 10).

are no other fixed points, however the twice iterated map gives

$$x_{n+2} = (1+\mu)^2 x_n - (1+\mu)(2 + 2\mu + \mu^2)x_n^3 + O(x_n^5)$$

which is the correct form for a pitchfork bifurcation. Thus there is a period two orbit present and stable for $\mu > 0$. This bifurcation is structurally stable; in particular a quadratic term can be removed by a conjugation of the form $h(x) = x + ax^2 + \ldots$. A relevant theorem is thus

Theorem 3.11. *(Period doubling) Suppose*

1. $\Phi_\mu(0) = 0$ *for all* μ *in an interval around* μ_0.
2. $\Phi'_{\mu_0}(0) = -1$
3. $\frac{\partial(\Phi_\mu^2)'}{\partial\mu}_{\mu_0}(0) \neq 0$.

Then there is an interval I around zero and a function $p : I \to \mathbb{R}$ such that

$$\Phi_{p(x)}(x) \neq x, \qquad \Phi^2_{p(x)}(x) = x$$

Note that in the period doubling bifurcation, the fixed point changes stability without another fixed point being created or destroyed; the object created is a period two orbit.

There are higher dimensional analogues of these bifurcations, for example adding an expanding or contracting direction as with $\dot{y} = cy$ with $c \neq 0$ (and similarly for a map) to the fold gives a generic **saddle-node** bifurcation in which a saddle and node are created. Sometimes the one-dimensional fold is called a saddle-node for this reason.

There is also one different and commonly encountered bifurcation found in higher dimensional flows.[27] The **Hopf** (or Poincaré-Andronov-Hopf) bifurcation has normal form in polar coordinates

$$\dot{r} = r(\mu - r^2)$$

$$\dot{\theta} = 1$$

We see that in the r variable this is just a pitchfork bifurcation, however $r > 0$ is no longer a point, it is a circle. Hence the stable focus at $\mu < 0$ has become a limit cycle. As with the pitchfork, there is a subcritical version obtained by changing the sign. This bifurcation is structurally stable.

Theorem 3.12. *(Hopf bifurcation) Suppose a flow in $\mathbb{R}^2$ $\dot{\mathbf{x}} = f_\mu(\mathbf{x})$ satisfies $f_\mu(\mathbf{0}) = \mathbf{0}$ for all μ and that Df_μ has eigenvalues $\alpha(\mu) \pm i\beta(\mu)$ with $\alpha(0) = 0$, $\beta(0) \neq 0$, $\alpha'(0) \neq 0$, then any neighbourhood of the origin contains a nontrivial periodic orbit for some μ.*

Example 3.13. *Consider the linear flow*

$$\dot{\mathbf{x}} = \begin{pmatrix} \mu & 1 \\ -1 & \mu \end{pmatrix} \mathbf{x}$$

Then $\alpha = \mu$ and $\beta = 1$. The conditions of the theorem are satisfied, but we find that the periodic orbits exist only for $\mu = 0$ (harmonic oscillator). This is equivalent to replacing $\mu - r^2$ by μ in the normal form.

Example 3.14. *The van der Pol oscillator (used for example in electric circuits) has equations*

$$\ddot{x} + b(x^2 - 1)\dot{x} + x = 0$$

Writing $y = \dot{x}$ we have

$$\dot{x} = y$$

$$\dot{y} = b(1 - x^2)y - x$$

This has a fixed point at the origin with derivative

$$Df = \begin{pmatrix} 0 & 1 \\ -1 & b \end{pmatrix}$$

and hence $\alpha = b/2$, $\beta = \sqrt{1 - b^2/4}$. Thus there is a periodic orbit near $b = 0$. Note however that again this corresponds to $b = 0$ exactly (harmonic oscillator). For $b > 0$ there is a limit cycle in this system, but it is a finite distance from the origin, and so not directly related to the Hopf bifurcation.

[27] The map version is called a Neimark-Sacker bifurcation, but it is significantly more complicated due to resonance phenomena, in particular if the eigenvalue is a kth root of unity for $k \leq 4$.

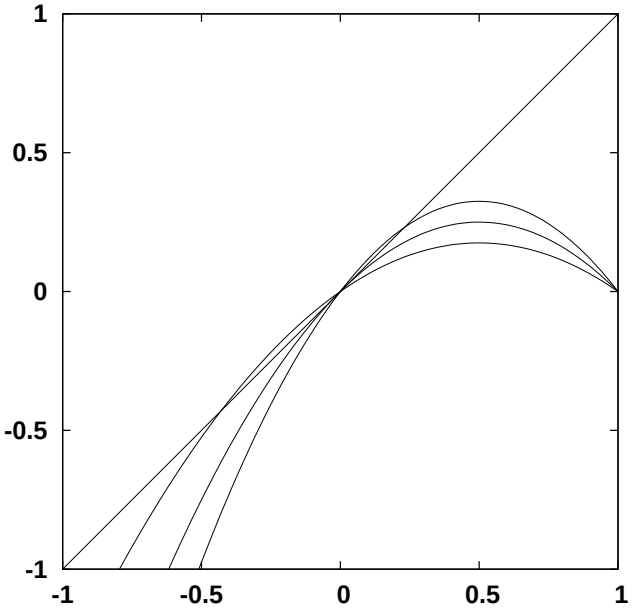

Figure 12: The logistic map Φ_r for $r = 0.7, 1, 1.3$ illustrating the transcritical bifurcation.

3.4 Local bifurcations in the logistic map

The logistic map $rx(1-x)$ provides examples of several of these bifurcations. As discussed previously, there are two fixed points, $x = 0$ and $x = (r-1)/r$ with stability eigenvalues r and $2-r$ respectively. Both are non-hyperbolic at $r = 1$ so we consider the dynamics in that region, writing $x = \delta$, $\mu = r - 1$:

$$\begin{aligned} \delta_{n+1} &= (1+\mu)\delta_n(1-\delta_n) \\ &= (1+\mu)\delta_n - \delta_n^2 + O(\mu\delta_n^2, \delta_n^3) \end{aligned}$$

which corresponds to a transcritical bifurcation.

There is another non-hyperbolic point at $r = 3$. Here we have for $x = x^* + \delta$, $x^* = (r-1)/r$, $\mu = r - 3$:

$$\begin{aligned} \Phi_\mu(x^* + \delta) &= x^* - (1+\mu)\delta - (3+\mu)\delta^2 \\ \Phi_\mu^2(x^* + \delta) &= x^* + (1+\mu)^2\delta - \mu(1+\mu)(3+\mu)\delta^2 + \ldots \end{aligned}$$

so we can see the conditions of the period doubling theorem are satisfied, and we have created a stable period 2 orbit. Analysing this orbit in the same way, we find that it too undergoes a period doubling bifurcation to a period 4 orbit at $r = 1 + \sqrt{6}$.[28]

This mechanism explains the appearance of the periodic orbits which are powers of two, but not the others, for example the period 3 window. The map Φ^3 has, for large r, maxima corresponding to orbits that reach the highest point of the map in the third iteration: $x_0 \to x_1 \to 1/2 \to r/4$. The largest value of $x_0 \approx 0.9$. As r increases, this peak rises until a point near is top is tangent to the line $y = x$, making a fixed point of Φ^3 which is thus a period 3 orbit of Φ. Beyond this point, the peak intersects the line twice, making a stable and an unstable fixed point of Φ^3: A fold bifurcation. The relevant parameter value is $r = 1 + \sqrt{8}$[29]

It is helpful to analyse the behaviour of orbits close to these periodic points. In the case of a stable periodic orbit, $0 < |D\Phi^p| < 1$, we know that the perturbation $\delta_n = x_n - x^*$ evolves as

$$\delta_{n+p} = (D\Phi^p)\delta_n(1 + O(\delta_n))$$

[28]It might be best to enlist the aid of a computer algebra package such as maple or mathematica for this.

[29]This is not easy to derive; if you want to give up, read J. Bechhoefer, Math. Mag. **69**, 115-118 (1996).

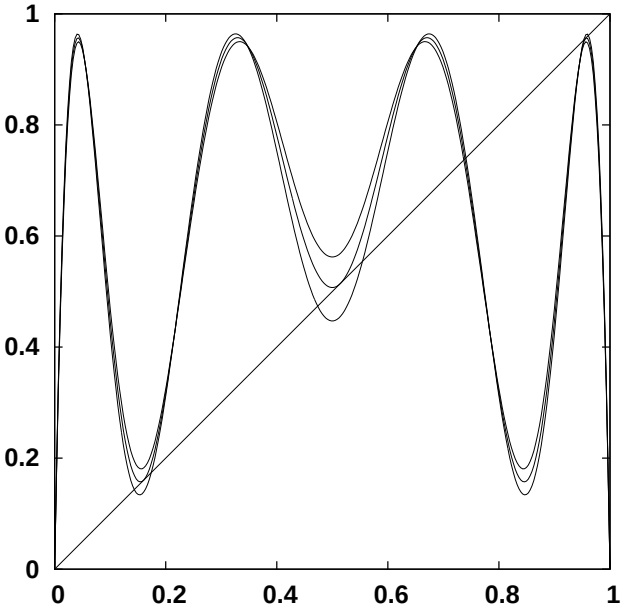

Figure 14: The map Φ_r^3 for $r = 3.8, 1 + \sqrt{8}, 3.856$ illustrating the birth of a stable and unstable pair of period three orbits in a fold bifurcation.

Iterating this we find

$$\delta_{np} = (D\Phi^p)^n\delta_0(1 + O(\delta_0))$$

where the coefficient of the correction term is a convergent

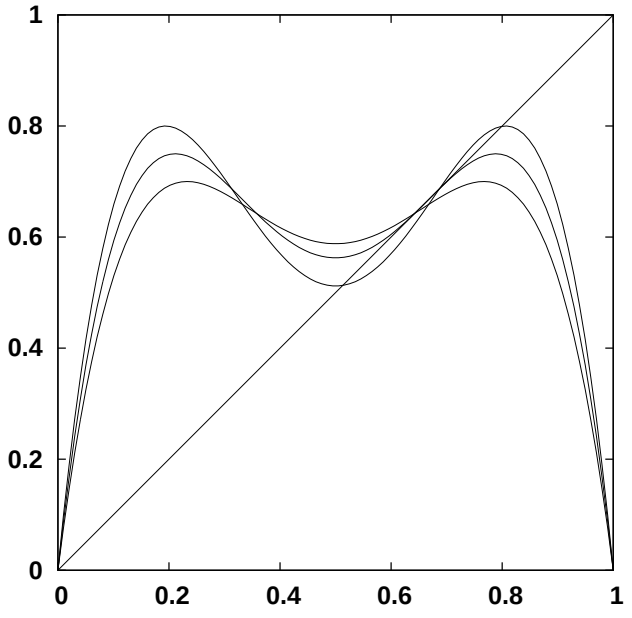

Figure 13: The map Φ_r^2 for $r = 2.8, 3, 3.2$ illustrating the first period doubling of Φ_r, which is a pitchfork bifurcation of Φ_r^2.

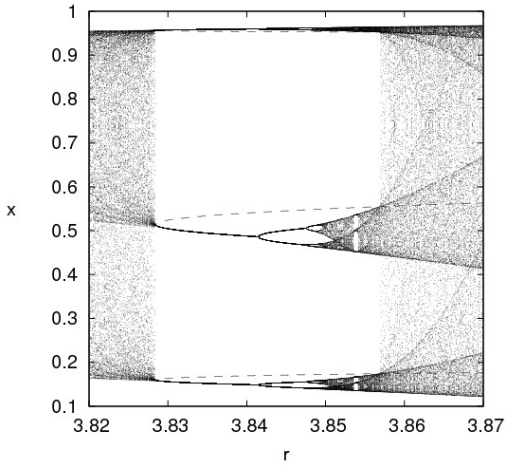

Figure 15: The period 3 window in the logistic map (blowup of Fig 2), together with the unstable period 3 orbit (dotted line).

geometric series.

If $D\Phi^p = 0$ (that is, $1/2$ is one of the points in the orbit) we have a **superstable** orbit;[30] typical behaviour is

$$\delta_{n+p} = C\delta_n^2(1 + O(\delta_n))$$

leading to quadratic convergence, similar to the Newton-Raphson method:

$$\delta_{np} = \exp[2^n \ln \delta_0(1 + O(\delta_0))]$$

If $D\Phi^p = 1$, we have two cases. In the transcritical bifucation at $r = 1$ and later fold bifurcations we have

$$\delta_{n+p} = \delta_n - c\delta_n^2$$

This decreases to zero (the fixed point is still marginally stable), but not exponentially. We see that the equation can be satisfied order by order with

$$\delta_n = cn^{-1} + O(n^{-2})$$

At the left of each period doubling interval we have

$$\delta_{n+p} = \delta_n - c\delta_n^3 + O(\delta_n^5)$$

which is satisfied by

$$\delta_n = \frac{1}{\sqrt{2cn}} + O(n^{-3/2})$$

These effects are important for numerical simulation. If we simulate the map directly with double precision (roughly 16 digit) arithmetic, we cannot expect to get closer than about $10^{-16/3} \approx 10^{-5}$ at period doubling parameters, since the increments δ_n^3 are smaller than the roundoff. Even getting this far will take of order $\delta^{-2} \approx 10^{11}$ iterations. If the application allows superstable parameter values instead, these are clearly preferable.

3.5 General one dimensional maps

The logistic map is interesting as the quadratic (hence perhaps simplest nonlinear) example of one dimensional maps, but it is important to know how many of its properties apply to other examples. The answer is, surprisingly many. We already met the period three theorem in the introduction. For the remaining results the main property we need is that of negative Schwarzian derivative:

Definition 3.15. *The Schwarzian derivative of a function $f(x)$ is*

$$S[f] = \frac{f'''(x)}{f'(x)} - \frac{3}{2}\frac{f''(x)^2}{f'(x)^2}$$

[30] We can generalise superstable orbits to study orbits for which the critical point is pre-periodic rather than periodic. See R. V. Jensen and Christopher R. Myers. "Images of the critical points of nonlinear maps." Phys. Rev. A **32** 1222-1224 (1985).

If a map Φ has negative Schwarzian derivative (this statement always means for all x), so do all its iterates Φ^n for $n \geq 2$. Using this condition we have

Theorem 3.16. *(Singer) If Φ is piecewise monotonic with l intervals and has negative Schwarzian derivative within each interval, Φ has at most $l+1$ stable or marginal periodic orbits, obtained as limits of the orbits of its $l+1$ local extrema.*

For the logistic map, the stable/marginal periodic orbits are either the fixed point at zero (found from iterating the endpoints of the interval) or at most one found by iterating the critical point $x = 1/2$.

Another interesting feature is that of Feigenbaum universality. The period doublings in maps with negative Schwarzian derivative and a single quadratic critical point occur at shorter and shorter intervals in r, such that the ratio of consecutive intervals

$$\delta = \lim_{n\to\infty} \frac{r_n - r_{n-1}}{r_{n+1} - r_n}$$

exists and is equal to $4.6692\ldots$, independent of the map. The reason is that at the endpoint, the map Φ_{r_∞} tends, under the operation of doubling and scaling, to a universal function, the fixed point of the functional dynamical system

$$\mathcal{R}\Phi = \alpha\Phi^2(x\alpha)$$

for a universal constant $\alpha = -2.5029\ldots$. This fixed point (the solution of $\mathcal{R}\Phi = \Phi$) has a single unstable eigenvalue given by δ. The remaining infinitely many dimensions are contracting (hence stable). Thus the fixed point may be reached by varying the single parameter r. Rigorous proofs of these statements exist, but are technical (see K&H, section 11.3). Maps with higher order critical points, such as $r[1/2^k - |1/2 - x|^k]$, $k > 2$ have a separate universality class (hence δ and α constants) for each k. Note that this terminology, "universality," "renormalisation," comes from an analogy with the physics of phase transitions.

Example 3.17. *The map $\Phi(x) = r\sin\pi x$ has negative Schwarzian derivative and a single quadratic critical point on $[0, 1]$ for $r \in (0, 1]$.*

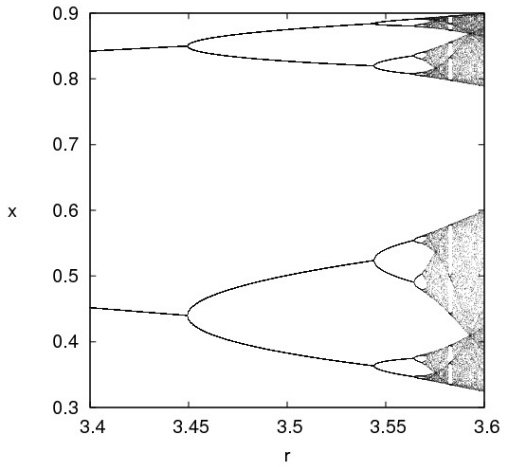

Figure 16: The main bifurcation cascade of the logistic map (blowup of Fig. 2).

4 Global dynamics

4.1 Stable and unstable manifolds

The behaviour of orbits close to nodes and focus points (ie with eigenvalues all stable or all unstable) is straightforward. However for saddle points we want more precise understanding than provided by the Hartman-Grobman theorem. The following applies to either invertible maps or flows:

Definition 4.1. *The stable manifold of a point* $\mathbf{x}$ *is*

$$W^s(\mathbf{x}) = \{\mathbf{y} : |\Phi^t(\mathbf{y}) - \Phi^t(\mathbf{x})| \to 0, t \to \infty\}$$

while its unstable manifold is

$$W^u(\mathbf{x}) = \{\mathbf{y} : |\Phi^{-t}(\mathbf{y}) - \Phi^{-t}(\mathbf{x})| \to 0, t \to \infty\}$$

For linear maps these are linear subspaces E^s and E^u of dimension given by the number of stable and unstable eigenvalues respectively, and the rate of convergence is exponential. The linear spaces are spanned by the relevant eigenvectors and (in the degenerate case) generalised eigenvectors. Given a neighbourhood U of a fixed point, we define local stable and unstable manifolds as

Definition 4.2. *The local stable manifold of a fixed point* $\mathbf{x}$ *is*

$$\begin{aligned} W^s_{loc}(\mathbf{x}) &= \{\mathbf{y} : |\Phi^t(\mathbf{y}) - \Phi^t(\mathbf{x})| \to 0, t \to \infty; \\ &\quad \Phi^t(\mathbf{y}) \in U, t \geq 0\} \end{aligned}$$

while its local unstable manifold is

$$\begin{aligned} W^u_{loc}(\mathbf{x}) &= \{\mathbf{y} : |\Phi^{-t}(\mathbf{y}) - \Phi^{-t}(\mathbf{x})| \to 0, t \to \infty; \\ &\quad \Phi^{-t}(\mathbf{y}) \in U, t \geq 0\} \end{aligned}$$

We have

Theorem 4.3. *(stable manifold theorem) Each hyperbolic fixed point* $\mathbf{x}$ *has a neighbourhood in which* $W^s_{loc}(\mathbf{x})$ *is a manifold of the same dimension as, and tangent to,* $E^s(D\Phi|_{\mathbf{x}})$.

Reversing time, we obtain the same result for unstable manifolds. For non-hyperbolic fixed points, there is also a **centre manifold** tangent to the corresponding linear subspace. It is however in general less smooth than the dynamics, and may not be unique. However the Taylor expansion is unique - centre manifolds may only differ by an amount smaller than any power of distance (eg exponential). The centre manifold is important as it controls bifurcations and its Taylor series expansion is used to derive normal forms for these.

A **manifold** is a set which has locally the same topological and differential structure as Euclidean space. We can apply the dynamics to obtain global manifolds from the local ones, however the theorem does not guarantee these sets to be smooth or continued indefinitely (and hence they may not be manifolds in the usual sense). The global manifolds may also be dense in X.

Example 4.4. *The map* $\Phi(x, y) = (x/2, 2y - 15x^3/8)$ *has a fixed point at* $(x, y) = (0, 0)$. *Its linearisation is*

$$D\Phi = \begin{pmatrix} 1/2 & 0 \\ 0 & 2 \end{pmatrix}$$

thus it is a saddle point. We have $W^u(0, 0)$ *is the y-axis, and* $W^s(0, 0)$ *is the curve* $y = x^3$, *which is tangent to the x-axis, which is the stable space.*

Note that while the definitions of manifolds apply equally to invertible maps and flows, in the map case a single orbit gives only a discrete set of points on the manifold, while for a flow it traces out a one dimensional manifold. This is similar to the invariant curves we saw in the case of linear dynamics. Numerically, a one-dimensional unstable or stable manifold can be estimated by (forward or backward) numerical integration of points near the fixed point, but higher-dimensional manifolds often need more specialised methods[31]

The definitions for stable and unstable manifolds can be applied to more general sets than fixed points and periodic orbits. The linearised map gives an x-dependent linear map $D\Phi|_{\mathbf{x}}$ on perturbation vectors. A **hyperbolic set** $\Lambda \subset X$ is a set for which each point $x \in \Lambda$ has stable and unstable spaces of perturbations[32] $E^{s/u}(\mathbf{x})$ which span the full space of perturbations, and for which perturbations in these spaces decay exponentially (in positive or negative time, respectively).[33] An **Anosov map** is one for which the whole space X is a hyperbolic set, and a an **Anosov flow** is one for which there is also a one-dimensional centre space corresponding to the flow direction.[34]

4.2 Homoclinic and heteroclinic orbits and bifurcations

A **homoclinic orbit** is one contained in both the stable and unstable manifolds of a single fixed point, thus it approaches the fixed point for both limits $t \to \pm\infty$. A **homoclinic point** is a point on such an orbit. Similarly a **heteroclinic orbit** is one that approaches different fixed

[31] See B. Krauskopf, H. M. Osinga, E. J. Doedel, M. E. Henderson, J. Guckenheimer, A. Vladimirsky, M. Dellnitz and O. Junge, Intern. J. Bifur. Chaos, **15**, 763-791 (2005).

[32] Technically, sub-bundles of the tangent bundle.

[33] From this definition it follows that the spaces depend continuously on $\mathbf{x}$ and are invariant under the dynamics, that is, $(D\Phi)_{\mathbf{x}} E^s_{\mathbf{x}} = E^s_{\Phi(\mathbf{x})}$ and the same with s replaced by u. For more details, see the scholarpedia article on hyperbolic dynamics.

[34] Anosov systems are rare in physics; the first mechanical example was probably the triple linkage: T. J. Hunt and R. S. MacKay, Nonlinearity **16** 1499-1510 (2003); M. Kourganoff, Commun. Math. Phys. **344** 831-856 (2016).

points for $t \to \pm\infty$, a **heteroclinic point** is a point on such an orbit, and a **heteroclinic cycle** is a sequence of homoclinic and/or heteroclinic orbits returning to the first fixed point. The consequences of these orbits differ substantially between maps and flows. We have

Theorem 4.5. *(Poincaré-Bendixson tricotomy) For flows in $\mathbb{R}^2$: Suppose a forward orbit $\{\Phi^t(\mathbf{x}), t > 0\}$ is contained in a compact set containing a finite number of fixed points. Then its ω-limit set is either*

- *A fixed point*
- *A periodic orbit*
- *A finite or countable heteroclinic cycle*

Here, the **ω-limit set** is the set of accumulation points of the forward orbit; the time reverse is called the **α-limit set**. Thus flows on the plane cannot be chaotic. The countable case is pathological — very mild conditions are needed to ensure a fixed point has only finitely many homoclinic orbits. Notice that a heteroclinic cycle is not an orbit of the system, however it can be the limit of a orbit that approaches it from the outside or inside. Such an orbit spends increasingly long near the fixed points, and so will have average behaviour - if the relevant limits exist - that is related to these fixed points.

On higher genus surfaces, such as the torus, more complicated behaviour can occur, for example an irrational translation ($(\dot{x}, \dot{y}) = (\alpha, \beta)$ with α and β incommensurate) leads to a dense future orbit, so the ω-limit set is the entire torus.[35] On the other hand, for a map a single homoclinic point with transverse stable and unstable manifolds is sufficient to generate a chaotic "homoclinic tangle"; see Fig. 17.

Many homoclinic and heteroclinic orbits are not structurally stable - a small perturbation will cause the orbit to miss its intended fixed point, or cause the stable and unstable manifolds of a homoclinic point to become transverse: These lead to various kinds of global bifurcations, ie changes to the structure of orbits over a wide region as a result of a parameter change.

Example 4.6. *The Duffing oscillator is given by*

$$\ddot{x} + b\dot{x} + (x^3 - x) = 0$$

It has fixed points at $x = -1, 0, 1$. The fixed point at $x = 0$ is a saddle. The fixed points at $x = \pm 1$ are stable foci for $b > 0$ and unstable foci for $b < 0$. At $b = 0$ the fixed point at $x = 0$ has a pair of homoclinic orbits encircling each of the other fixed points, which become heteroclinic orbits for $b \neq 0$. The structure of the orbits is topologically distinct for $b > 0$ and $b < 0$.

[35] Dynamics of translations on flat surfaces of higher genus (ie with singular points) is a popular research field, related to that of billiards in polygons with angles that are rational multiples of π.

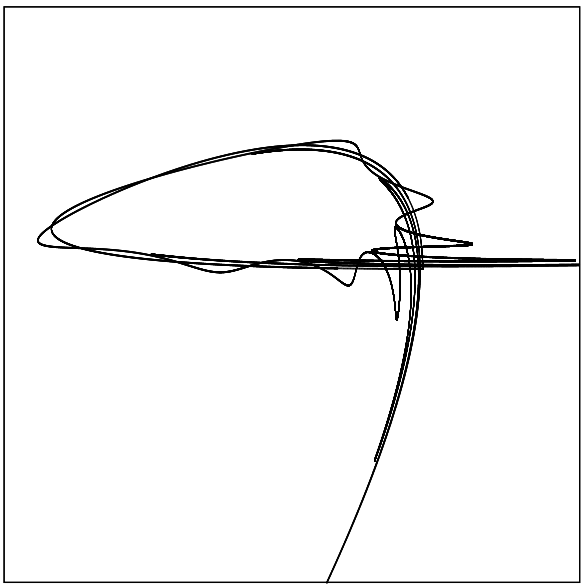

Figure 17: Stable and unstable manifolds for the map $\Phi(x, y) = (3(x + (x - y)^2), (y + (x - y)^2)/3)$ with inverse $\Phi^{-1}(x, y) = (x/3 - (3y - x/3)^2, 3y - (3y - x/3)^2)$, showing a homoclinic tangle.

Instead of a focus, the trapped part of the manifold could also approach a periodic orbit or heteroclinic cycle. Thus a homoclinic bifurcation may also arise as a collision of a periodic orbit and a saddle point. For maps and for $d \geq 3$ flows a wider variety of possibilities occurs.

Example 4.7. *The driven pendulum $(\dot{x}, \dot{v}) = (v, -\omega^2 \sin x + A \sin \Omega t)$ exhibits qualitatively different behaviour for zero and non-zero driving coefficient A. For $A = 0$ there is a homoclinic connection from $x = \pi$ to the equivalent version $x = -\pi$. For non-zero A we can consider the stroboscopic $t = 2\pi\Omega^{-1}$ map, which is now an autonomous two dimensional map. Typical perturbations of the homoclinic connection leads to transverse manifolds and chaos in the vicinity of this orbit for arbitrarily small A.*

Remark: The existence of connections in unperturbed versions of both Duffing and pendulum models is because they are Hamiltonian and so have a conserved energy. The level curves of energy are invariant under the dynamics, so that homoclinic orbits are typical.

4.3 Attractors and crises

A final category of global bifurcations is where the change is related to larger attractors than stable fixed points or limit cycles (recall the Lorenz attactor, Fig. 1); these are

called crises. There are varying definitions of attractor in the literature[36]

Definition 4.8. *Attracting set: A compact invariant set $A \subset X$ that has a neighbourhood U such that $A = \cap_{t=0}^{\infty} \Phi^t(U)$*

If $\Phi^t(U) \subset U$ for all $t > 0$ for a compact set U then the above intersection will always lead to an attracting set. Attracting sets are robust: A map $\tilde{\Phi}$ which is uniformly close to Φ has an attracting set $\tilde{A} \subset U$ which is close to A, similarly obtained by taking the intersection of forward images of U.

Definition 4.9. *Attractor: An attracting set containing a dense orbit.*

This means an attractor does not contain other attractors. Attractors may be (marginally) stable fixed points or limit cycles. They can also be more complicated chaotic and fractal objects such as the Lorenz attractor, and then are often called **strange attractors**. They are chaotic in that there is still sensitive dependence on initial conditions: Any periodic orbits within the attractor will have both stable and unstable directions. The **basin of attraction** $B(A)$ is the set of points whose ω-limit set is contained in A. The **basin boundary** is its boundary $\partial B(A)$.

Similarly

Definition 4.10. *Repelling set: A compact invariant set $R \subset X$ that has a neighbourhood U such that $R = \cap_{t=0}^{\infty} \Phi^{-t}(U)$.*

Definition 4.11. *Repeller: A repelling set containing a dense orbit.*

Note that a repeller for non-invertible map may contain other repellers, for example a chaotic repeller may contain repelling periodic orbits. On the other hand for reversible dynamics the involution maps attractors to repellers and vice versa.

A rough classification of crises is given by

Boundary crisis An attractor touches its basin boundary; beyond this crises orbits will eventually leave the attractor.

Interior crisis An attractor touches an unstable periodic orbit within its basin of attraction, and expands in size.

Attractor merging crisis Two or more attractors touch an unstable periodic orbit on their mutual basin boundary.

[36] See the scholarpedia entry on Attractor

The logistic map has examples of the first two of these. The set displayed on the bifurcation diagram for $0 < r < 4$ is the attractor, whether a stable fixed point or periodic orbit or chaotic set. For $1 < r < 4$ the set U may be taken to be an interval of the form $[1-\epsilon, 1+\epsilon]$. At the Feigenbaum transition point, the set is uncountable.[37] Beyond this point, if the attractor is a chaotic set it contains a zero measure set of unstable periodic orbits (there are other orbits, eg $x = 0$ that are not contained in it, however). The set of periodic orbits is countable, but the closure of the set includes uncountably many aperiodic orbits, some of them dense in the set (we will see this in the following sections). We see that for a map, an attractor need not be connected. For a flow, it is, since it must be invariant under the dynamics.

Boundary crisis At $r = 4$ the attractor fills the whole interval $[0, 1]$ and hence touches its basin boundary. Beyond this, orbits will remain in the interval only if they avoid the interval mapping above one, ie

$$\frac{1}{2}\left(1 - \sqrt{\frac{r-4}{r}}\right) < x < \frac{1}{2}\left(1 + \sqrt{\frac{r-4}{r}}\right)$$

Preimages of this "hole" cover the interval densely, however none of the (now unstable) periodic orbits have been destroyed, so there is a zero measure set of orbits that never escape. The set, now a repeller, is uncountable and comprises the closure of these orbits. It contains contains other repellers, for example the individual unstable periodic orbits. This repeller, and many others, is an example of a hyperbolic set.

Interior crisis Similarly, the period 3 "window" ends at $r \approx 3.8568$ with an interior crisis - the 3-fold attractor touches the unstable period 3 orbit that was created with the fold bifurcation at $r = 1 + \sqrt{8} \approx 3.8284$ and expands to fill the entire region. The unstable period 3 orbit is in fact the edge of a repeller comprised of the closure of the remaining infinitely many unstable periodic orbits, so this crisis may also be viewed as a collision between an attractor and a repeller.

Attractor merging crisis The attractor merging crisis is also called a **symmetry breaking crisis** as normally the two attractors are symmetrically related. For example, the antisymmetric logistic map $\Phi(x) = rx(1 - |x|)$ on behaves like two separate copies of the logistic map for $r \leq 4$, with opposite signs. These merge at $r = 4$, leading to a single attractor located in $|x| < (r+1)/r$ up to

[37] Feigenbaum attractor is a fractal with Hausdorff dimension approximately 0.538. Hausdorff dimension is discussed in chapter 6.

$r = 2 + \sqrt{8} \approx 4.8284$, at which point the orbits of the critical points $x = \pm 1/2$ start to leave the system.[38]

[38] A more subtle example is given by the Lorentz gas in C. P. Dettmann and G. P. Morriss, Phys. Rev. E **54** 4782-4790 (1996). Here the attractor and its time reverse (repeller) both collide with a periodic orbit and merge.

5 Symbolic dynamics

5.1 The binary shift

The special case $r = 4$ of the logistic map, or the equivalent $1 - 2x^2$ on $[-1, 1]$ is sometimes called the Ulam map. We note the similarity with the second Tchebyscheff[39] polynomial $T_2(x) = 2x^2 - 1$. In general Tchebyscheff polynomials are defined by the (rather un-polynomial looking) $T_n(x) = \cos n \arccos(x)$, that is, it gives the formula for $\cos nx$ as a polynomial in $\cos x$. This suggests a trigonometric conjugation; for the logistic version we see that if

$$x_{n+1} = 4x_n(1 - x_n)$$

is transformed according to $x_n = \sin^2(\pi y_n/2)$, we find

$$x_{n+1} = 4\sin^2\frac{\pi y_n}{2}(1 - \sin^2\frac{\pi y_n}{2}) = \sin^2(\pi y_n)$$

Thus

$$y_{n+1} = \pm 2y_n \pmod 1$$

If we take the usual arcsin of a positive number, so in the range $[0, \pi/2]$ we find from this

$$y_{n+1} = r\min(y_n, 1 - y_n) = \frac{r}{2}(1 - |2y - 1|)$$

which is called the **tent map**, for $r = 2$.[40] We will keep $r = 2$ for the tent map for the rest of this section. An even simpler related map is the **doubling map** (also called **Bernoulli map**)

$$y_{n+1} = \{2y_n\}$$

where $\{\}$ denotes the fractional part.

Consider the binary representation of the point $y \in [0, 1)$,[41]

$$y = \sum_{j=0}^{\infty} \omega_j 2^{-(j+1)}$$

where the symbols $\omega_j \in \{0, 1\}$. The doubling map just ignores ω_0 and shifts all the other ω_j by one. The tent map does the same, but if $\omega_0 = 1$ all symbols are flipped as well as shifted. In each case the sequence of ω_0 values, which denotes the "rough location" of the point with respect to the partition $\{[0, 1/2), [1/2, 1)\}$ is called the **symbol sequence**.

We see there is (almost) a 1:1 correspondence between y and $\{\omega_j\}$, with the minor exception being the case of trailing repeated 1s, a countable set. Thus the dynamics $\Phi : [0, 1) \to [0, 1)$ is conjugate to the shift $\Sigma : \Omega_2^R \to \Omega_2^R$ where Ω_2^R is the set of "right" sequences of two symbols. The set Ω_2 denotes bi-infinite sequences $j \in \mathbb{Z}$, and is useful for invertible maps. The metric, ie distance between two sequences can be defined as[42]

$$d(\{\omega_j\}, \{\phi_j\}) = 2^{-\min\{|j|:\omega_j \neq \phi_j\}}$$

which then defines a topology in which Σ is continuous. This topological conjugacy between the shift map and doubling or tent maps (and hence also the Ulam map) has some immediate consequences:

- Periodic points are countable and dense.
- There is a dense orbit; this property is called **topological transitivity**[43]
- There are orbits that are neither periodic nor dense.

Example 5.1. *List and concatenate all possible finite symbol sequences* $\{0, 1, 00, 01, 10, 11, 000, \ldots\}$*:*

$$0100011011000001010011100101110111\ldots$$

Each finite symbol sequence appears infinitely often, so the orbit generated by this sequence is dense.

Example 5.2. *Any aperiodic sequence of* 00 *and* 10 *gives a nowhere dense orbit since there are real numbers with binary expansions containing* 11 *arbitrarily close to any real number.*

This should be compared with the Devaney definition of chaos:[44]

- Periodic points are dense
- The system is topologically transitive
- There is sensitive dependence on initial conditions.

The last condition is that in every neighbourhood of a point, there are initial conditions that eventually separate to a specified distance. It turns out[45] that the last condition follows from the first two. Thus the doubling, tent and Ulam maps are chaotic according to this definition.

We get more specific information about the periodic points — there are clearly 2^n symbol sequences with periods a factor of n for each n, and a dense set of preperiodic

[39] There are other spellings; the initial 'T' makes sense in terms of the usual notation $T_n(x)$

[40] The tent map is also topologically conjugate to the Farey map introduced in chapter 2 using as the conjugation the "Minkowski question mark function". The latter has the property that periodic continued fractions (ie quadratic irrationals) get mapped to periodic binary expansions (ie rationals).

[41] The $j + 1$ is so that j starts at zero, for consistency with the literature for symbolic dynamics.

[42] There are many equivalent metrics used in the literature.

[43] Sometimes the definition is that for any open sets U and V, there is an $n \geq 0$ so that $\Phi^n(U) \cap V$ is non-empty. For a discussion of when these are equivalent, see `http://www.scholarpedia.org/article/Topological_transitivity`

[44] From his textbook, *A first course in chaotic dynamical systems* first published in 1992. This is a popular definition but other definitions are useful in different contexts.

[45] J. Banks, J. Brooks, G. Cairns, G. Davis and P. Stacey, Amer. Math. Mon. **99**, 332-334 (1992).

points for each periodic point. Starting from any finite symbol sequence, say 001, we can construct a periodic symbol sequence $\overline{001}$ and hence calculate its corresponding point in the doubling map $y = 0.\overline{001}_2 = \sum_{j=1}^{\infty} 2^{-3j} = 1/7$ where the subscript denotes binary. For the tent map we ensure that the flips are taken into account, giving $y = 0.\overline{001110}_2 = 14\sum_{j=1}^{\infty} 2^{-6j} = 2/9$. Thus the corresponding point in the Ulam map is $x = \sin^2 \pi/9 \approx 0.116978$.

The stability of a periodic point $D\Phi^p = 2^p$ for the doubling map and $\pm 2^p$ for the tent map depending on the parity of the number of 1s in the symbol sequence. We can see that the conjugation relating the tent and Ulam maps preserves this: If $\Psi = h^{-1} \circ \Phi \circ h$ for some conjugating function h, we see that $\Psi^p = h^{-1} \circ \Phi^p \circ h$ and so for a fixed point x of Ψ^p and corresponding $y = h(x)$ of Φ^p we have

$$D\Psi^p|_x = (Dh^{-1}|_y)(D\Phi^p|_y)(Dh|_x) = D\Phi^p|_y$$

since the first and last terms in the product cancel, assuming both are non-zero. Thus we have $\Psi^p = \pm 2^p$ for the Ulam map also, except for the fixed point $x = 0$ (at which the conjugation is singular) which has $D\Psi = 4$.[46]

Remark: The doubling map is particularly bad to simulate directly on a computer, since most software uses a binary representation of real numbers. After a very small number of doublings the result is zero. It is much better to simulate a (pseudo-)random sequence of binary symbols.

Remark: Sequences with trailing 1s are equivalent in the binary representation to others with trailing 0s. They are just pre-images of the two fixed points $\overline{0}$ and $\overline{1}$ which are identified for the doubling map. Thus there are actually 2^{n-1} points of period a factor of n for the doubling map. In contrast, these are all distinct for the tent and Ulam maps.

5.2 Open binary shifts

For the logistic map with $r > 4$ and tent map for $r > 2$ there are intervals around $1/2$ that map out of $[0, 1]$. However the image of the interval $[0, 1/2]$ still includes the whole space $[0, 1]$, as does the image of $[1/2, 1]$. Thus for any point $x \in [0, 1]$ we can construct two preimages $\Phi_0^{-1}(x)$ and $\Phi_1^{-1}(x)$ and hence 2^n preimages of order n, one for each sequence of n symbols. It can be shown (using the Schwarzian derivative property for the logistic map) that this leads to a 1:1 correspondence between the set of points that remain forever in $[0, 1]$ and the binary shift.

Example 5.3. *Consider the case $r = 3$ for the tent map. The intervals $[0, 1/3]$ and $[2/3, 1]$ are each mapped to $[0, 1]$, so that the shift corresponds to the ternary representation*

$$x = \sum_{j=0}^{\infty} 2\omega_j 3^{-j-1}$$

This is the **middle third Cantor set.**

Properties of these sets follow easily from the shift representation: They are uncountable, complete, nowhere dense and totally disconnected (any two points are in different components). Also, the Cantor set itself is structurally stable - perturbing r does not affect any of these properties.

All the periodic orbits remain unstable as r is increased, so we can use the method of **inverse iteration** to locate them: Start at a convenient point (say, $x = 1/2$) and apply a periodic sequence of $\Phi_\omega^{-1}(x)$ until the result converges.

A natural higher dimensional version of the open binary shift is called the Smale horseshoe. If a (roughly rectangular) set is mapped to a "horseshoe" shaped set that covers the full width of the original in two places and for which the original covers the full width of the horseshoe, then the set surviving for infinite time is a Cantor set (labelled as above by the symbol sequence) in the unstable direction and smooth in the stable direction. The set surviving for both positive and negative infinite time is the intersection of two Cantor sets, itself a Cantor set, and labelled by the shift on the full space Ω_2. Again, it is structurally stable.

An important result is that a homoclinic tangle, ie map with a homoclinic point at which the stable and unstable manifolds are transverse, has horseshoe dynamics in a sufficiently high iterate of the map, and hence the full complexity of the binary shift dynamics. Recall Fig. 17.

Example 5.4. *The* **Henon map**, $(x, y) \to (1 - ax^2 + y, bx)$,[47] *has a good example of a Smale horseshoe. For parameters $a = 12$, $b = 0.8$ it has the form shown in Fig. 18, leading to a Cantor set of points that never escape. For other parameter values, such as the original $a = 1.4$, $b = 3$ it behaves like the logistic map for $r < 4$, having an attractor of a fixed point or a fractal. It is closely related to the logistic map, but less well understood and also an active subject of research.*

Example 5.5. *A billiard system consisting of three circular scatterers with a "non-eclipsing" condition (no scatterer intersects the convex hull of the others) has a trapped set with complete binary symbolic dynamics, with symbols denoting which of the other two scatterers is encountered next.*[48]

[46] The argument can often be reversed - if all the periodic points of two hyperbolic dynamical systems have the same spectra (eigenvalues of $D\Phi^p$), they can often be shown to have a smooth conjugation. See for example Thm 20.4.3 in A. Katok and B. Hasselblatt, "Introduction to the modern theory of dynamical systems." (Cambridge University Press, 1997).

[47] Other trivial variations of the equations can be found in the literature

[48] The dynamics of this system was studied rigorously in A. Lopes and R. Markarian, Siam J. Appl. Math. **56** 651-680 (1996). But it had appeared previously in the physics literature — see chaosbook.org of Cvitanovic et al, where it is called three disk pinball

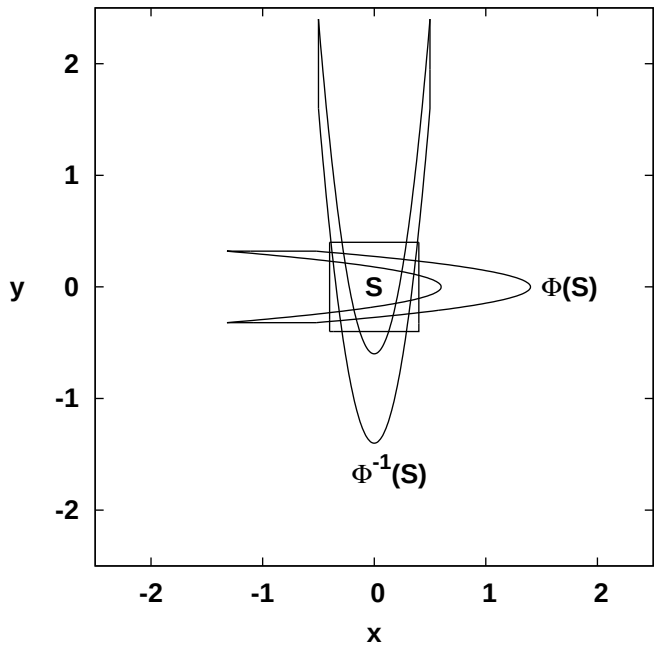

Figure 18: The Henon map, $a = 12$ and $b = 0.8$

If there are both expanding and contracting directions, as in these two dimensional examples, we cannot use inverse iteration to locate the periodic orbits numerically. In boundary value problems in ODEs there are two common approaches: Shooting and relaxation.[49] A shooting method involves finding an approximate initial condition, evolving the system to the end point, and checking the final boundary condition (here, that it is equal to the initial condition). In chaotic systems this is problematic since orbits are often exponentially unstable. Thus we usually need an approximation to the whole orbit, either by running a long trajectory and looking for near recurrences (for example if the periodic orbit is embedded in an attractor) or using known symbolic dynamics. Then we can refine it using one of the following methods:

Damped Multipoint Newton method For a cycle of length p we seek a zero of the function $F : \mathbb{R}^p \to \mathbb{R}^p$

$$F\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_p \end{pmatrix} = \begin{pmatrix} x_1 - \Phi(x_p) \\ x_2 - \Phi(x_1) \\ \dots \\ x_p - \Phi(x_{p-1}) \end{pmatrix}$$

The multidimensional Newton formula, found by taking the Taylor expansion around the zero to linear order, gives

$$(DF)(\mathbf{x}_{n+1} - \mathbf{x}_n) = -\gamma F(\mathbf{x})$$

where a damping parameter $0 < \gamma \leq 1$ is added by hand to increase the basin of attraction; the usual Newton method is $\gamma = 1$. Here we have (writing $\Phi'(x_k) = \Phi'_k$)

$$\begin{pmatrix} 1 & & & -\Phi'_p \\ -\Phi'_1 & 1 & & \\ & \dots & \dots & \\ & & -\Phi'_{p-1} & 1 \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \dots \\ \Delta x_p \end{pmatrix} = -\gamma \begin{pmatrix} F_1 \\ F_2 \\ \dots \\ F_p \end{pmatrix}$$

Row reduction gives

$$\begin{pmatrix} 1 & & -\Phi'_p \\ & 1 & -\Phi'_p\Phi'_1 \\ & \dots & \dots \\ & & 1 - \Phi'_p\Phi'_1 \dots \Phi'_{p-1} \end{pmatrix} \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \dots \\ \Delta x_p \end{pmatrix} = -\gamma \begin{pmatrix} F_1 \\ F_2 + \Phi'_1 F_1 \\ \dots \\ F_p + \Phi'_{p-1}F_{p-1} + \dots + \Phi'_{p-1} \dots \Phi'_1 F_1 \end{pmatrix}$$

from which the solution may be found by dividing through by the last diagonal element and back substituting. Note that the matrix manipulations have been done explicitly - we need only store the vectors used in the intermediate steps.

Variational method Write an action function such as $S = |F|^2$ and minimise using standard multidimensional minimisation routines.[50] In the open billiard example, there is a natural action given by the sum of the path lengths.

5.3 Subshifts

The period three window of the logistic map has an infinite set of unstable orbits, including all the periodic orbits

— and a quantum version had been considered rigorously in S. Sjöstrand, Duke Math. J. **60** 1-57 (1990). The latter proposed what is now called a fractal Weyl law, relating the fractal properties of the classical trapped set to the distribution of quantum resonances.

[49] W. H. Press, S. A. Teukolsky, W. T. Vetterling and B. P. Flannery, *Numerical Recipes* (Cambridge University Press, 1992) gives advice on which of these to try first: "Shoot first, and only then relax." But if the shooting is chaotic, this may not be the best strategy.

[50] However standard, routines for multidimensional minimisation are not guaranteed to work unless you know a lot about your system.

from the original bifurcation cascade, that persist without any bifucations in this parameter region. The period three points delineate two regions, roughly $0.15 < x < 0.5$ which belongs to the left branch of the map (symbol '0') and $0.5 < x < 0.95$ which belongs to the right branch (symbol '1'). Points in '0' map to '1' while points in '1' may map either to '0' or '1'. Thus we have a symbolic dynamics in which only some of the possible transitions occur; here we specifically exclude the sequence '00.'

If there are a finite[51] number of exclusion rules, this is called a **subshift of finite type** or **Topological Markov chain**. This will occur in a one-dimensional map which is expanding ($\Phi'(x) > 1$) if there is a partition of the space for which each element (corresponding to a symbol) is mapped to a union of elements (modulo boundary points); this is called a **Markov partition**[52] In the case of the logistic map, the derivative $\Phi'(x)$ is not always less than one, so the conjugacy with a symbolic system needs justification, and clearly fails for some of the stable orbits.

Such a system can be represented as a directed graph with adjacency matrix A with entries zero or one to denote whether a transition is possible, and a symbol space

$$\Omega_A = \{\omega \in \Omega_n | (A)_{\omega_n \omega_{n+1}} = 1 \quad \text{for} \quad n \in \mathbb{Z}\}$$

It is easy to show that the number of possible paths of length m from symbols i to j are given by the entry $(A^m)_{ij}$ of the matrix A^m. In particular, the number of periodic points of length m is the trace of A^m.

We can get from i to j iff $(A^m)_{ij} > 0$ for some $m \geq 0$, and write $i \to j$, or j is **accessible** from i. We have $i \to i$ automatically. If $i \to j$ and $j \to i$ then i and j **communicate**; this is an equivalence relation, so partitions the symbols into disjoint classes. On the other hand, if there is a j so that $i \to j$ but $j \not\to i$ then i is **inessential**. If all symbols are essential, ie there is a single communication class, the system[53] is **irreducible** and topologically transitive. In this case we also have a dense set of periodic orbits and hance chaos in the sense of Devaney.

The **period** of a symbol i is the greatest common divisor of the times m at which the dynamics can return to i, that is, when $(A^m)_{ii} \neq 0$, and infinite if $A^m_{ii} = 0$ for all $m > 0$. For example if the directed graph is bipartite, all states have even periods. The period is constant on all communication classes. If all symbols have period one, the system is **aperiodic**. If it is both irreducible and aperiodic, then for all sufficiently large m, all entries of A^m are positive, and the system satisfies a stronger property that topological transitivity:

Definition 5.6. *A system is* **topologically mixing** *if for any two open sets U, V, $\Phi^t(U) \cap V$ is nonempty for all sufficiently large t.*

Clearly topological mixing implies topological transitivity.

In this case we can use

Theorem 5.7. *Perron-Frobenius theorem: For a matrix A with non-negative entries, such that some power A^m has all positive entries, there is an eigenvector with positive entries with corresponding eigenvalue real, positive, simple and greater in magnitude that all other eigenvalues.*

Finally the growth of symbol sequences and periodic orbits are both controlled by this largest eigenvalue of A: If A is irreducible and aperiodic we have

$$\lim_{n\to\infty} \frac{1}{n} \ln \sum_{i,j} (A^n)_{ij} = \lim_{n\to\infty} \frac{1}{n} \ln \sum_{i} (A^n)_{ii} = \ln \lambda_{\max}$$

where $\lambda_{\max}$ is the largest eigenvalue.

For a general dynamical system we can define

Definition 5.8. *Let $N(\epsilon, T)$ be the smallest number of points x_k such that for any $x \in X$ we have $|\Phi^t(x) - \Phi^t(x_k)| < \epsilon$ for all $0 \leq t < T$ and some k. Then the* **topological entropy** *is*[54]

$$h_{top} = \lim_{T\to\infty} \limsup_{\epsilon\to 0} \frac{1}{T} \log N(\epsilon, T)$$

The base of the logarithm is arbitrary, often given as 2. The topological entropy is invariant under topological conjugacy, and in the case of an irredicible and aperiodic symbolic system is given by $\log \lambda_{\max}$.

If the largest eigenvalue of a matrix A is unique and simple, as in the irreducible and aperiodic case, it may be found with the **power method**: Apply A repeatedly to an arbitrary positive vector and normalise. The normalisation constant will converge exponentially to $\lambda_{\max}$ at a rate determined by the spectral gap (difference in magnitude between the largest and next largest eigenvalue(s)). The method does not require any reduction of the matrix, and hence can be used with very large sparse matrices.[55]

[51] There are some simply defined generalisations with an infinite number of rules, such as the even shift, in which each 0 is followed by an even number of 1's. This is not a subshift of finite type, but is in a larger category called sofic shifts, represented by directed graphs in which the same symbol may appear in more than one place. So here we allow transitions $0 \to 0$, $0 \to 1$ $1 \to 1'$, $1' \to 1$, $1' \to 0$ and disallow all others. Many typically encountered shifts of infinite type from dynamical systems do not have a simple representation, however.

[52] Markov partitions (with a more involved definition) are also used in higher dimensional dynamics; note that the boundaries can be fractal, see eg Arnoux, Pierre, and Shunji Ito. "Pisot Substitutions and Rauzy fractals." Bulletin of the Belgian Mathematical Society Simon Stevin **8** 181-208 (2001).

[53] "system" depending on context refers to any of the matrix A, the topological Markov chain, the directed graph, the symbolic dynamics, and the original dynamical system.

[54] Lim is equivalent to limsup in this definition, however the same is not true of some alternative but very similar definitions: see arxiv:1707.09052

[55] It is reputedly used in Google PageRank and Twitter Who To Follow algorithms.

The map $\Phi_\beta(x) = \{\beta x\}$ is called the **beta-transformation** (Renyi 1957). For β an integer, we have a (full) shift on β symbols as in the previous section. For other values, dividing the unit interval using multiples of β^{-1} gives the "greedy" representation of a number in inverse powers of β:

$$x = \sum_{j=0}^{\infty} \omega_j \beta^{-(j+1)}$$

For some algebraic values of β, the boundary of the final partition element, 1 maps onto a multiple of β^{-1} and we have a Markov partition.

Example 5.9. *$\beta = g = (1+\sqrt{5})/2$, the golden ratio. $\Phi_\beta(1) = \{g\} = (\sqrt{5}-1)/2 = g^{-1}$ and so we have the transitions $0 \to 0$, $0 \to 1$, $1 \to 0$ analogous to the period three window of the logistic map. The transition matrix is*

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

The higher powers are given by

$$A^n = \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}$$

where F_n is the Fibonacci number, satisfying $F_1 = F_2 = 1$, $F_n = F_{n-1} + F_{n-2}$. This recurrence may be solved explicitly to find

$$F_n = \frac{1}{\sqrt{5}}(g^n - (-g)^{-n})$$

Thus the number of fixed points of order n is $P_n = F_{n+1} + F_{n-1}$. Note that because a matrix satisfies its own characteristic equation we have

$$A^2 - A - I = 0$$

Multiplying by an arbitrary power of A and taking the trace we have

$$P_n = P_{n-1} + P_{n-2}$$

which may be solved together with $P_1 = 1$, $P_2 = 3$ without determining A^n for general n directly. Finally, note that as with the doubling map, the symbolic dynamics is not quite 1:1: The discontinuity $x = g^{-1}$ has two symbol sequences $1\overline{0}$ and $\overline{01}$; similarly for its preimages.

Example 5.10. *Another system with this symbolic dynamics is given by the doubling map $x \to \{2x\}$ but enforcing escape for any x with symbol sequence 11, corresponding to the points $x \in [3/4, 1]$.*

6 Statistical properties

6.1 Probability measures

We now consider statistical properties of dynamical systems. Since for chaotic systems there is sensitive dependence on initial conditions, and it is not practical in physical systems to specify or measure the initial conditions exactly, we can take an approach specifying only the probability that the system is in a given region $A \subset X$ at any time, using a probability measure μ:

$$\mathbb{P}(x \in A) = \mu(A)$$

for which the main properties are that $0 \leq \mu(A) \leq 1$ with $\mu(X) = 1$[56] and $\mu(\cup_i A_i) = \sum_i \mu(A_i)$ for a finite or countable collection of disjoint sets A_i.[57] Often, the measure is given by a density ρ

$$\mu(A) = \int_A \rho(x)dx$$

however we have already met some sets, namely Cantor sets, in which a density does not exist. We can include the deterministic case using the Dirac measure $\delta_x(A)$ which is 1 if $x \in A$ and zero otherwise.

A map or flow Φ^t then causes the probability measure to evolve according to the **transfer operator**

$$\Phi^t_*(\mu)(A) = \mu(\Phi^{-t}(A))$$

where we recall that for noninvertible maps the inverse can be defined on sets. The $*$ is conventional but is sometimes omitted. In terms of densities we have

$$\Phi^t_*(\rho)(x) = \sum_{y \in \Phi^{-t}(x)} \frac{\rho(y)}{|\det(D\Phi^t|_y)|} = \int \delta(x - \Phi^t(y))\rho(y)dy$$

The transfer operator is a (generally infinite dimensional) linear operator on measures or densities. Fixed points of the transfer operator, that is, eigenvectors with eigenvalue one, are called **invariant measures**. They satisfy $\Phi^t_*\mu = \mu$, or equivalently $\mu(\Phi^{-t}(A)) = \mu(A)$. Any convex (ie normalised positive linear) combination of invariant measures is an invariant measure.

Example 6.1. *The map $x \to \{3x\}$ has many invariant measures. The uniform measure $\rho = 1$ is invariant, since each point has three pre-images and the Jacobian factor is $1/3$ everywhere. Each periodic orbit gives invariant delta measures, for example*

$$\frac{1}{4}(\delta_{1/10} + \delta_{3/10} + \delta_{9/10} + \delta_{7/10})$$

There are also fractal measures, for example the uniform measure on the middle third Cantor set that gives each included interval of size 3^{-n} a measure 2^{-n}.

Do invariant measures exist in general? If for a given initial point x the average rate of landing in a (sufficiently) arbitrary set A exists, it generates a measure:

$$\mu_x(A) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \chi_A(\Phi^t(x))$$

Here $\chi_A(x)$ is the characteristic function of A, equal to 1 if $x \in A$ and zero otherwise. Linear combinations of these can approximate any continuous function $\phi : X \to \mathbb{R}$, giving an expression for the **time average**, also called **Birkoff average** of the function with initial point x:

$$\phi_T(x) = \lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \phi(\Phi^t(x))$$

We thus have a measure μ_x defined so that

$$\phi_T(x) = \int_X \phi(y)d\mu_x(y) = \int_X \phi(y)\rho_x(y)dy$$

where the second equality holds if μ_x has a density.

All continuous maps on a compact space have at least one invariant measure, obtained by taking a subsequence in the limit. An important result is the **Birkoff ergodic theorem** which states that for any invariant measure μ, the set of x for which the time average does not exist is of zero μ-measure. The time average may still depend on x, however we have

$$\int \phi_T(x)d\mu(x) = \int \phi(x)d\mu(x)$$

Example 6.2. *The map $x \to \{3x\} + [x]$ (where square brackets indicate integer part) on $[0, 2)$ has an invariant density $\rho(x) = 1/2$. But for $\phi(x) = x$ we find*

$$\phi_T(x) = \begin{cases} 1/2 & x < 1 \\ 3/2 & x > 1 \end{cases}$$

for almost all x. We have

$$\int \phi_T(x)\rho(x)dx = \int x\rho(x)dx = 1$$

as expected.

An important case is that of C^2 expanding circle maps (for example perturbations of the doubling map that continue to identify 0 with 1 and satisfy $|\Phi'(x)| > 1$ everywhere): Here an invariant density exists and is unique.

Note that by introducing probability some philosophical issues have crept in: The deterministic system with uncertain initial conditions now has behaviour that is indistinguishable from a fair die.

[56] An important area of research is that of dynamics with infinite measures, in which $\mu(X) = \infty$ instead.

[57] Measure theory requires some more technical properties, particularly that we define $\mu(A)$ only for some subsets, called measurable sets, which normally include all Borel sets, obtained by complements and countable intersections and unions of closed and open intervals.

6.2 Markov chains

A piecewise linear map with Markov partition can be represented as a topological Markov chain with corresponding symbolic dynamics of finite type, as we saw before. The transfer operator gives further information: Any density which is constant on the partition elements, $\rho(x) = \rho_i$ for $x \in X_i$ evolves to another density of the same type. In particular, we have

$$(\Phi_*\rho)_j = \sum_i \frac{\rho_i A_{ij}}{|D\Phi_i|}$$

or in terms of the total mass in each partition element $\pi_i = \mu(X_i) = \rho_i |X_i|$

$$\pi_j = \sum_i \pi_i P_{ij}$$

with the transition probabilities

$$P_{ij} = \frac{A_{ij}|X_j|}{|D\Phi_i||X_i|}$$

This leads to dynamics on a directed graph, with transition probabilities given by the P_{ij} matrix, which, like A_{ij}, satisfies the Perron-Frobenius theorem in the irreducible aperiodic case. In a closed system, probabilities add to one, so we have

$$\sum_j P_{ij} = 1$$

This ensures the leading eigenvalue is 1, with left eigenvector π^{inv} giving the invariant measure corresponding to the limiting state, with

$$\lim_{n\to\infty} (P^n)_{ij} = \pi_j^{\text{inv}}$$

a projection onto that state.

Example 6.3. *The golden mean beta map $\{gx\}$ above has partition elements $X_0 = [0, g^{-1})$, $X_1 = [g^{-1}, 1)$. We find*

$$P = \begin{pmatrix} g^{-1} & g^{-2} \\ 1 & 0 \end{pmatrix}$$

and hence $\pi^{\text{inv}} = (g, g^{-1})/\sqrt{5}$.

Even when there is no Markov partition, a useful numerical technique called **Ulam's method** can approximate the transfer operator using a similar approach.[58] Divide the space into a fine partition, and assume that the probability of a transition from i to j is given by the proportion of X_i that is mapped to X_j, that is

$$P_{ij} = \frac{|X_i \cap \Phi^{-1}(X_j)|}{|X_i|}$$

Then (for a closed system) these probabilities add to one and if the matrix is irreducible and aperiodic, there is a single invariant measure given by the left eigenvalue. The power method applies here also: Starting with a positive vector π, repeated multiplication by P will then converge to this measure exponentially fast.

6.3 Measures in open systems

In the case of open systems, probability is not conserved, but we may want to know the probability conditional on remaining within the system. Thus the (normalised) measure evolves according to

$$\mu_t = \frac{\Phi_*^t \mu_0}{\Phi_*^t \mu(X)}$$

A fixed point of this operation is called a **conditionally invariant measure**[59] and in this case the total probability decays exponentially

$$\mu(\Phi^{-t}(X)) = e^{-\gamma t}$$

with escape rate

$$\gamma = -\ln \Phi_*^1 \mu(X)$$

The value $e^{-\gamma}$ can be considered an eigenvalue of the transfer operator Φ_*^1.

Example 6.4. *For the map $x \to \{3x\}$ with hole $[1/3, 2/3]$, the uniform measure is conditionally invariant with $\Phi_*\mu(X) = 2/3$.*

It is also useful to consider (fully) invariant measures in these systems, for example supported on the trapped set.

6.4 Fractal dimensions

Measures can also be used to describe the size of fractals such as the Cantor set and Lorenz attractor. The d-dimensional Hausdorff measure[60] is defined by

$$H^d(A) = \frac{\pi^{d/2}}{2^d \Gamma(d/2+1)} \lim_{\delta\to 0} \inf_{\substack{\{U_i\} \\ A \subset \cup_i U_i \\ |U_i| < \delta}} \sum_i |U_i|^d$$

where we cover the set A with sets U_i of maximum diameter δ, take the infimum over covers and the limit $\delta \to 0$. The normalisation constant ensures that for d an integer we get the usual Lebesgue measure. For ordinary sets such as lines, we have that the length may be finite, but

[58] Ulam's method gives as a basis, piecewise constant functions on the X_i. If there is reason to assume the invariant density is smooth, alternative methods can be developed using polynomial or trigonometric basis functions.

[59] See for example, M. F. Demers and L.-S. Young, Nonlinearity **19**, 377-397 (2005).

[60] Mark Pollicott has a nice set of lecture notes on fractal dimensions: `http://homepages.warwick.ac.uk/~masdbl/preprints.html`

the number of points (zero dimensional measure) are infinite and the area (two dimensional measure) is zero. This behaviour applies more generally, and we can define the **Hausdorff dimension**:

$$d_H(A) = \sup\{d : H^d(A) = \infty\} = \inf\{d : H^d(A) = 0\}$$

Note that $H^{d_H}(A)$ may be zero, finite or infinite. The Hausdorff dimension behaves nicely under finite or countable unions

$$d_H(\cup_i A_i) = \sup_i d_H(A_i)$$

and in particular, the Hausdorff dimension of a countable set is zero.

If we insist that all covering sets are the same size, we arrive at a different quantity, the **Minkowski** or **box** dimension:

$$D_B(A) = -\lim_{\delta\to 0} \frac{\ln N(\delta)}{\ln \delta}$$

if the limit exists, where $N(\delta)$ can be either the number of balls or cubes needed to cover A, or the number of cubes containing A in a grid of box length ϵ. If the limit is not defined, the lim sup and lim inf give upper and lower Minkowski dimensions, respectively. In general we have

$$D_H(A) \le D_B(A)$$

Example 6.5. *The set $1/n$ for $n \in \{1, 2, 3, \ldots\}$ is countable, so it has Hausdorff dimension zero, but requires of order $\delta^{-1/2}$ boxes to cover for small δ. Thus it has box dimension $1/2$.*

In the case where the set is a finite union of similar[61] contracted copies of itself, as with the trapped sets of piecewise linear open maps,

$$A = \cup_i f_i(A)$$

with $|f_i(\mathbf{x}) - f_i(\mathbf{y})| = r_i|\mathbf{x} - \mathbf{y}|$ for all $\mathbf{x}$ and $\mathbf{y}$ and with $r_i < 1$, and there is a nonempty open set V satisfying

$$V \subset \cup_i f_i(V)$$

with the union disjoint, the **open set condition**, we have

$$D_B(A) = D_H(A) = D_S(A)$$

where the **similarity dimension** $D_S(A)$ satisfies

$$\sum_i r_i^{D_S(A)} = 1$$

Note that the maps f_i here are contractions; in dynamical contexts they typically correspond to inverse branches of an expanding map Φ.

Example 6.6. *The middle third Cantor set consists of two copies of itself scaled by $1/3$. Thus its dimension satisfies*

$$2(1/3)^D = 1, \qquad D = \frac{\ln 2}{\ln 3}$$

[61] The case of affine maps, where the contraction rates differ in different directions, is more complicated and a subject of current research.

6.5 Ergodic properties

With respect to an invariant measure μ we can define properties analogous to topological transitivity and mixing. The statements in this section hold for any positive measure sets $A, B \subset X$ as $T \to \infty$; for flows the sums are replaced by integrals. We have in increasing order of strength:

$$\mu(A \cap \Phi^{-t}A) \not\to 0 \qquad \textbf{Recurrence}$$

Poincaré's recurrence theorem states that all systems with invariant probability measures are recurrent, an unexpected result since it seems to imply history is (with probability one) destined to repeat itself infinitely many times.[62] Suppose we consider a container partitioned into two sections, and N gas particles initially on the left. If the dynamics is measure preserving (physically realistic, as we will see), we might expect this event to re-occur after roughly a characteristic time scale multiplied by the inverse of the measure of this state, 2^N. However for $N \approx 10^{23}$ this time is unphysically large; real physical experiments do not have access to infinite time limits. Note that in the context of nanotechnology, we often have only a few particles, so timescales may be more reasonable.

$$\frac{1}{T}\sum_{t=0}^{T-1} \mu(A \cap \Phi^{-t}B) - \mu(A)\mu(B) \to 0 \qquad \textbf{Ergodicity}$$

For ergodic measures we have an important result relating time and space averages:

$$\phi_T(x) = \int_X \phi(x) d\mu$$

for all x except a set of measure zero, and ϕ for which the integral is defined. Thus in an ergodic system, varying the initial conditions does not (with probability one) affect the long time average, as claimed in the Introduction.

We can always decompose an invariant measure as a (possibly uncountable) convex combination of ergodic measures. However, it is important (and often difficult) to know whether a natural invariant measure, such as Lebesgue, is ergodic. Sinai showed in 1970 that two disks on a torus is ergodic (on the set defined by the conserved quantities such as energy), however a proof for arbitrary number of balls in arbitrary dimension was finally published by Simanyi in 2013.[63] Since Boltzmann in the mid-19th century, ergodicity has been assumed in statistical mechanics to calculate macroscopic properties of systems of many particles. Here, again, the question of time scales

[62] Recurrence is not guaranteed in infinite measure systems, for example $x \to x + 1$ with $x \in \mathbb{R}$. In such systems ergodicity, defined as the statement that all invariant sets or their complements have zero measure, implies recurrence under mild conditions, however there is no generally agreed definition of mixing.

[63] N. Simanyi, Nonlinearity **26**, 1703-1717 (2013).

arises, and also the likelihood that many systems are not quite ergodic, having very small measure regions in phase space that do not communicate with the bulk.

$$\frac{1}{T}\sum_{t=0}^{T-1}|\mu(A\cap\Phi^{-t}B)-\mu(A)\mu(B)|\to 0 \qquad \textbf{Weak Mixing}$$

Weak mixing is equivalent to the statement that the doubled system on $X\times X$, $(\mathbf{x}_1,\mathbf{x}_2)\to(\Phi(\mathbf{x}_1),\Phi(\mathbf{x}_2))$ is ergodic.

$$\mu(A\cap\Phi^{-T}B)-\mu(A)\mu(B)\to 0 \qquad \textbf{(Strong) Mixing}$$

Mixing is equivalent to correlation decay

$$\int_X f(\mathbf{x})g(\Phi^t(x))d\mu-\int_X f(\mathbf{x})d\mu\int_X g(\mathbf{x})d\mu\to 0$$

for every square integrable f,g. The rate of decay (dependence on t) is important, in that diffusion and similar properties can be expressed as sums over correlation functions. In general it depends on the functions, but it can be shown for example that for systems conjugate to irreducible aperiodi Markov chains, decay is exponential if f and g are Hölder continuous.

Example 6.7. *Rotations $x\to\{x+\alpha\}$ have a uniform invariant measure. They are are ergodic iff $\alpha\notin\mathbb{Q}$. In this case they are* **uniquely ergodic***: There is only a single invariant measure. They are not weakly mixing.*[64]

Example 6.8. *Markov chains are not ergodic if reducible, ergodic but not weakly mixing if irreducible but periodic, and strong mixing if irreducible and aperiodic.*

Example 6.9. *The Chacon shift, generated recursively by the substitution $0\to 0010$, $1\to 1$ and allowing any finite sequence appearing there, is a notable system which is weak mixing but not strong mixing.*

The **metric entropy** or **Kolmogorov-Sinai entropy**,[65] is defined using a finite partition ξ to define the measure of initial conditions leading after n iterations of Φ to a sequence of symbols $\omega\equiv\omega_1\ldots\omega_n$:

$$h(\xi)=\lim_{n\to\infty}\frac{-1}{n}\sum_{\omega}\mu(X_\omega)\ln\mu(X_\omega)$$

Then the KS-entropy is

$$h_{KS}=\sup_{\xi}h(\xi)\le h_{top}$$

The KS entropy of a Markov chain is

$$h_{KS}=-\sum_{ij}\pi_i^{\text{inv}}P_{ij}\ln P_{ij}$$

Example 6.10. *We find for the golden beta map $x\to\{gx\}$*

$$\begin{aligned}h_{KS} &= -\left[\frac{g}{\sqrt{5}}g^{-1}\ln g^{-1}+\frac{g}{\sqrt{5}}g^{-2}\ln g^{-2}\right.\\ &\quad\left.+\frac{g^{-1}}{\sqrt{5}}1\ln 1+\frac{g^{-1}}{\sqrt{5}}0\ln 0\right]\\ &= \frac{1+2g^{-1}}{\sqrt{5}}\ln g\\ &= \ln g\end{aligned}$$

Note that $0\ln 0=0$ (which is the limit as $P_{ij}\to 0$). Also it turns out that all β expansions have topological entropy $\ln\beta$, so in this case it reaches the maximum.

This quantifies the maximum rate of information loss in the system,[66] but a positive value does not require ergodicity. If however, all nontrivial partitions $h(\xi)>0$, it is equivalent to a condition stronger than mixing called **K-mixing** (or Kolmogorov mixing).

The strongest ergodic property, beyond K-mixing is the **Bernoulli** property, which states that there is a partition with respect to which, elements at different times are completely uncorrelated:

$$\mu(X_i\cap\Phi^{-t}X_j)-\mu(X_i)\mu(X_j)=0$$

for all i and j, and all $t>0$. Aperiodic irreducible Markov chains and equivalent systems have this property. Thus we have Bernoulli implies K-mixing implies strong mixing implies weak mixing implies ergodicity. Although Bernoulli is the strongest ergodic property, it is worth recalling that systems containing a subset on which the dynamics is Bernoulli are extremely prevalent, including in particular Smale Horseshoes and therefore in homoclinic tangles.

6.6 Lyapunov exponents

[67] Lyapunov exponents quantify the sensitive dependence on initial conditions. We have already seen that the instability of periodic orbits is determined by the eigenvalues of $D\Phi^T$ where T is the period, and know that in general these eigenvalues can vary widely between orbits. However in ergodic systems, a similar phenomenon occurs as with time averages: Even though there is a wide variety of possible values of the expansion rate, almost all (with respect to the measure) orbits have the same values.

[64] However, a generalisation of rotations, **interval exchange transformations** are weakly mixing (but not strong mixing) for almost all parameter values.

[65] Confusingly, measure-theoretic quantities are called metric properties in dynamical systems, and various thermodynamic terms (entropy, pressure) are used in ways that differ from their physical usage. KS-entropy is more like an entropy per unit time.

[66] For a recent generalisation see R. G. James, K. Burke and J. P. Crutchfield, Phys. Lett. A **378** 2124-2127 (2014).

[67] A recent introductory survey on Lyapunov exponents is found in A. Wilkinson, arxiv:1608.02843.

For one dimensional maps we have for the exponential growth rate of perturbations in the initial condition,

$$\begin{aligned}\lambda_x &= \lim_{n\to\infty}\frac{1}{n}\ln|D\Phi^n|_x| \\ &= \lim_{n\to\infty}\frac{1}{n}\ln\prod_{i=0}^{n-1}|D\Phi|_{\Phi^i x} \\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}\ln|D\Phi|_{\Phi^i x}\end{aligned}$$

which is just an ordinary time average, and hence for an ergodic measure given by

$$\lambda = \int \ln|D\Phi| d\mu$$

for almost all x with respect to μ.

In higher dimensions the evolution of the perturbation is described by the product of the matrices along the orbit, acting on a vector that gives the direction of the initial perturbation; recall the discussion from section 3.1. It is clear that the result is very different if the vector lies in the local stable or unstable manifold. Results given by the **Oseledets theorem** are that

$$\lambda(\mathbf{v}) = \limsup_{t\to\infty}\frac{1}{t}\ln\frac{|D\Phi^t\mathbf{v}|}{|\mathbf{v}|}$$

takes the same finite set of values for almost every $\mathbf{x}$, called **Lyapunov exponents**, giving the largest expansion rate accessible to the linear space containing $\mathbf{v}$: Typically a one dimensional space in the stable manifold for the smallest Lyapunov exponent, and almost the whole space for the largest. The Lyapunov exponents may be found from

$$\{\lambda_i\} = \ln\{\text{e-vals of}\lim_{t\to\infty}((D\Phi^t)^*(D\Phi^t))^{1/2t}\}$$

where $*$ denotes transpose and hence the eigenvalues are all real. The matrix is of size d so there are d Lyapunov exponents (counted with multiplicity). Similar to stability eigenvalues of periodic orbits, the Lyapunov exponents are invariant under smooth conjugations. Note, however that unexpected phenomena can occur if the dynamics is time-dependent.[68]

We have

$$\sum_i \lambda_i = \lim_{t\to\infty}\frac{1}{t}\ln|\det(D\Phi^t)|$$

In the case of **Sinai-Ruelle-Bowen** (SRB) measures (that is, absolutely continuous on unstable manifolds), we have the **Pesin formula**

$$\sum_{i:\lambda_i>0}\lambda_i = h_{KS}$$

Example 6.11. *Consider the open map $x \to \{3x\}$ with escape from the middle third $x \in [1/3, 2/3)$. The natural measure in this case is the uniform measure on the non-escaping set, which is the middle third Cantor set. We have*

$$\begin{aligned}h_{KS} &= \ln 2 \\ \lambda &= \ln 3 \\ \gamma &= \ln(3/2) \\ d_H &= \frac{\ln 2}{\ln 3}\end{aligned}$$

This example strongly suggests the following relations:

$$h_{KS} = \lambda d_H$$

This is a form of the **Ledrappier-Young** formula.

$$\gamma = \lambda - h_{KS}$$

This is called the **escape rate formula**, generalising Pesin's formula. Both formulas have been shown under more general conditions[69]

Calculation of Lyapunov exponents numerically typically uses the **Benettin algorithm.** Evolve both the equations of the original system ($\mathbf{x}\in\mathbb{R}^d$)

$$\frac{d}{dt}\mathbf{x} = f(\mathbf{x}(t))$$

and the linearised equations ($\delta_i \in \mathbb{R}^d, \quad i \in \{1,2,3,\dots l\}$

$$\frac{d}{dt}\delta_i = (Df)|_{\mathbf{x}(t)}\delta_i$$

for as many perturbations l as Lyapunov exponents are required. Thus we solve $d(l+1)$ equations altogether. As some periodic interval T, apply a Gram-Schmidt orthogonalisation to the δ_i vectors,

$$\delta_1' = \delta_1 \qquad \delta_1'' = \frac{\delta_1'}{|\delta_1'|}$$

$$\delta_2' = \delta_2 - \delta_1''\cdot\delta_2 \qquad \delta_2'' = \frac{\delta_2'}{|\delta_2'|}$$

$$\delta_3' = \delta_3 - \delta_1''\cdot\delta_3 - \delta_2''\cdot\delta_3 \qquad \delta_3'' = \frac{\delta_3'}{|\delta_3'|}$$

then the largest l Lyapunov exponents are approximated by the sums

$$\lambda_i \approx \frac{1}{kT}\sum_{j=1}^{k}\ln|\delta_i'|$$

where j sums over the times the orthogonalisation is applied. Note that the vectors obtained are not directly

[68] G. A. Leonov and N. V. Kuznetsov, *Time-varying linearization and the Perron effects,* Intl. J. Bif. Chaos **17**, 1079-1107 (2007).

[69] See M. F. Demers, P. Wright and L.-S. Young, Ergod. Theor. Dyn. Sys. **32**, 1270-1301 (2012).

related to expanding or contracting spaces in the original dynamical systems (which need not be orthogonal), and that a similar algorithm applies to maps[70]

We see that the sum of positive exponents is related to the KS entropy (above). The largest exponent is also important, giving the rate of fastest growth of perturbations. Positivity is often a numerical signal of chaos (except in the case of a single unstable periodic orbit as in the pendulum).

In addition, the largest Lyapunov exponent can be used to estimate the number of iterations before a perturbation at the level of machine round-off (say, $\epsilon = 10^{-16}$) becomes of order unity, $-(\ln \epsilon)/\lambda$. In the case of the doubling map, this is the extent to which a direct numerical simulation gives typical behaviour (then reaching the fixed point at zero and then remaining there). For hyperbolic sets[71] there are **shadowing lemmas** that guarantee the existence of orbits close to approximate (eg numerical) orbits, however they rarely say much about typical behaviour: In particular the numerical doubling map orbit is an exact solution of the dynamics. Most systems do however show good numerical convergence as precision is increased, until a previously visited point is reached, leading to exact periodicity. In general we expect around ϵ^{-d} distinct points where d is a dimension of the attractor or other invariant set. If these are assumed to appear randomly, the time taken to reach a periodic cycle is around $\epsilon^{-d/2}$. The addition of weak noise clearly prevents such periodic cycles and presumably masks any effects due to the finite precision.[72]

The entire Lyapunov spectrum is also a fruitful object of study in many degree of freedom system, such as molecular dynamics models of large numbers (hundreds or thousands) of atoms.[73]

6.7 Cycle expansions

It is possible using the transfer operators discussed at the start of this chapter, to calculate many statistical quantities including averages, escape rates and Lyapunov exponents in terms of unstable periodic orbits, in some cases to incredible accuracy. The presentation here is based on chaosbook.org and restricted to the simplest systems (1D maps); it is possible to extend to higher dimensional maps, flows, semiclassical approach to quantum systems (see the quantum chaos course), and to a more rigorous treatment, especially for uniformly expanding or (in higher dimensions) hyperbolic systems.[74]

We saw that $e^{-\gamma}$ where γ is the escape rate, can be understood as the leading (ie greatest magnitude) eigenvalue of the transfer operator Φ^1_*. Treating this as a matrix,[75] the eigenvalues are inverses of solutions of

$$\begin{aligned} 0 &= \det(1 - z\Phi^1_*) \\ &= \exp(\mathrm{tr}\ln(1 - z\Phi^1_*)) \\ &= 1 - z\mathrm{tr}\Phi^1_* - \frac{z^2}{2}(\mathrm{tr}\Phi^2_* - (\mathrm{tr}\Phi^1_*)^2) + \dots \\ &= \sum_{n=0}^{\infty} Q_n z^n \end{aligned}$$

where we have expanded in powers of z (since looking for the smallest solution). Differentiating leads to the useful relation

$$Q_n = \frac{1}{n}\left(\mathrm{tr}\Phi^n_* - \sum_{m=1}^{n-1} Q_m \mathrm{tr}\Phi^{n-m}_*\right)$$

Now from the previous section we have

$$\Phi^n_*(\rho)(x) = \int \delta(x - \Phi^n(y))\rho(y)dy$$

from which we see that

$$\mathrm{tr}\Phi^n_* = \int \delta(x - \Phi^n(x))dx = \sum_{x:\Phi^n(x)=x} \frac{1}{|D\Phi^n|_x - 1|}$$

which is just a sum over periodic points of length n, including repeats of orbits with lengths that are factors of n. Thus, truncating at some value of n, we obtain an nth degree polynomial, whose roots give an approximation to the spectrum of Φ^1_*. The escape rate is then estimated as

$$\gamma = \ln z_1$$

where z_1 is the root of smallest absolute value. We expect good convergence when the system is hyperbolic (clearly $D\Phi(x) = 1$ is problematic above) and the symbolic dynamics is well behaved, so that long periodic orbits are partly cancelled by combinations of shorter periodic orbits in terms like $\mathrm{tr}\Phi^2_* - (\mathrm{tr}\Phi^1_*)^2$.

Example 6.12. *Using cycle expansions to obtain the escape rate of the open map $\Phi(x) = 5x(1-x)$ we find from the 8 periodic orbits up to length 4, the result $\gamma = 0.5527651$, accurate to 7 digits.*

[70] Also Poincaré sections, for example in billiards; see H. R. Dullin, Nonlinearity, **11**, 151-173 (1998).

[71] If hyperbolicity fails (as in almost all realistic systems), we can no longer expect shadowing. This is the case even if all periodic orbits are hyperbolic but they have different numbers of positive Lyapunov exponents (J. A Yorke, private communication).

[72] Some recent discussion of these issues is in R. Lozi, "Can we trust numerical computations of chaotic solutions of dynamical systems," (unpublished?)

[73] H.-L. Yand and G. Radons, Phys. Rev. Lett. **100** 024101 (2008).

[74] See for example "Rigorous effective bounds on the Hausdorff dimension of continued fraction Cantor sets: A hundred decimal digits for the dimension of E_2," O. Jenkinson and M. Pollicott, arxiv:1611.09276.

[75] The Fredholm determinant is a precise formulation of determinants of some classes of infinite dimensional operators.

7 Hamiltonian dynamics

7.1 Volume preserving systems

In the case of an invertible map or flow, it is easy to check whether the uniform measure $\rho = 1$ is invariant. We just need the Jacobian $|\det(D\Phi^t|_x)| = 1$, independent of x and t. For the flow we can differentiate with respect to t and set $t = 0$ to find $\mathrm{tr}(Df) = 0$. Such systems are called area or volume preserving. It is clear that there can be no attractors in such systems: Any neighbourhood of such a set will have larger volume and cannot shrink to it under the dynamics. Also, we know from above that the sum of the Lyapunov exponents must be zero.

This condition, together with the fact that the entries are real, implies that the product of the eigenvalues of $D\Phi$ is ± 1 or equivalently (for the flow) that the sum of the eigenvalues of Df is zero. For two dimensional systems, this means that fixed points are generically saddles or centres, may also be one of the marginal cases such as shears, but may not be a node or focus. Such fixed points are typically denoted "hyperbolic" (saddle), "elliptic" (centre) or "parabolic" (shear). Bifurcations of volume preserving (and specifically Hamiltonian) systems can be classified accordingly.[76] In two dimensions with $\det(D\Phi) = 1$, these may be distinguished by the trace $T = \mathrm{tr}D\Phi$, which is the sum of the two eigenvalues: $|T| < 2$ is elliptic, $|T| = 2$ is parabolic and $|T| > 2$ is hyperbolic. In the latter case $T < -2$ is sometimes denoted **inverse hyperbolic** or **reflection hyperbolic**. An important class of area preserving maps is given by **linear toral automorphisms**, $\mathbf{x} \to T\mathbf{x}$ considered modulo one in both coordinates, with T an integer matrix with unit determinant.

The condition $\det(D\Phi^t) = 1$ in two dimensions is equivalent to

$$S^* J S = J$$

where the star indicates transpose, $S = D\Phi^t$ and

$$J = \begin{pmatrix} 0_n & I_n \\ -I_n & 0_n \end{pmatrix}$$

in dimension $2n$, where 0_n and I_n are zero and identity matrices respectively as given in the introduction. This condition says the map Φ^t is **symplectic**,[77] and $D\Phi^t$ is a **symplectic matrix**. We see that for general n, areas of two dimensional spaces defined by the corresponding variables are preserved. Symplectic matrices of a given size form a group.

[76] Bifurcations, among others, include the saddle-centre, pitchfork and Hopf bifurcations, roughly corresponding to the one dimensional fold, pitchfork and Hopf bifurcations studied in chapter 3, but with stable/unstable fixed points replaced by centres and saddles. See the lecture notes "Symmetric Hamiltonian bifurcations," by P.-L. Buono, F. Laurent-Polz and J. Montaldi, LMS Lecture Notes **306** 357-402 (2005).

[77] Symplectic maps are reviewed in J. Meiss, Rev. Mod. Phys. **64** 795-848 (1992).

For $n > 1$ the symplectic condition is stronger than unit determinant. The characteristic polynomial of a symplectic matrix S

$$\begin{aligned} p(\lambda) &= \det(S - \lambda I) \\ &= \det(JS - \lambda J) \\ &= \det(JS - \lambda S^* JS) \\ &= \det(-\lambda^{-1} J + S^* J)\det(-\lambda S) \\ &= \lambda^{2n}\det(S - \lambda^{-1} I) \\ &= \lambda^{2n} p(\lambda^{-1}) \end{aligned}$$

noting that both S and J have unit determinant. Thus the eigenvalue spectrum of a symplectic matrix splits into pairs of inverses (quadruples of inverses and complex conjugates where complex). This implies that the Lyapunov spectrum of Φ is symmetric around zero.

7.2 Hamiltonian systems

Symplectic maps arise naturally in physical systems that are derived from a Hamiltonian $H(\mathbf{x})$ with $\mathbf{x} = (\mathbf{q}, \mathbf{p}), \mathbf{q}, \mathbf{p} \in \mathbb{R}^n$, the coordinates and momenta respectively. The integer n is the number of **degrees of freedom**. Note that in general, the Hamiltonian can depend explicitly on time. The momentum variables, called "canonical momenta" are not necessarily mass times velocity, for example if one of the q_i is an angle, the corresponding p_i could be an angular momentum.[78]

Hamilton's equations of motion are as follows:

$$\dot{\mathbf{x}} = J D_{\mathbf{x}} H(\mathbf{x})$$

where $\mathbf{x} = (\mathbf{q}, \mathbf{p}) \in \mathbb{R}^{2n}$, $D_{\mathbf{x}}$ denotes the gradient, J denotes the block matrix

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

where the zero and unit submatrices are here of size n. For a system of N particles in three dimensions, $n = 3N$.

The most common Hamiltonian function is of the form

$$H(\mathbf{q}, \mathbf{p}) = \sum_i \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{q})$$

for particles indexed by i with masses m_i moving with a potential energy function $V(\mathbf{q})$ that depends on the positions of all the particles, for example for Newtonian gravity as discussed in the introduction, we have

$$V(\mathbf{q}) = -\sum_{i<j} \frac{G m_i m_j}{|\mathbf{q}_i - \mathbf{q}_j|}$$

[78] The precise prescription, which we will not need, is as follows: For a given arbitrary set of coordinates q_i which can be any independent functions of positions and velocities, construct a Lagrangian function $L(\mathbf{q}, \dot{\mathbf{q}}, t)$ equal to the kinetic *minus* potential energies. Then invert the expressions $p_i = \partial L/\partial \dot{q}_i$ to write $\dot{q}_i = f_i(\mathbf{q}, \mathbf{p}, t)$ in terms of which the Hamiltonian is $H(\mathbf{q}, \mathbf{p}, t) = \sum_i p_i f_i(\mathbf{q}, \mathbf{p}, t) - L(\mathbf{q}, \mathbf{f}(\mathbf{q}, \mathbf{p}, t), t)$.

Another example we have seen is the simple pendulum, where q represents the angle from the lowest position, p the angular momentum, and

$$H(q,p) = \frac{p^2}{2ml^2} - mgl\cos q$$

where l is the length of the pendulum, m the mass of the bob and g the acceleration of gravity.

Hamilton's equations imply

$$Df = J(\partial^2_{\mathbf{x}} H)$$

with $\partial^2_{\mathbf{x}} H$ a symmetric matrix. Thus

$$(Df)^* J + J(Df) = 0$$

which then implies that $D\Phi^t$ is a symplectic matrix for all t, in particular that the dynamics is volume preserving and the Lyapunov spectrum is symmetric. Also, the involution $i(\mathbf{q},\mathbf{p}) = (\mathbf{q},-\mathbf{p})$ shows that the dynamics is reversible if the Hamiltonian is even in the momentum $H(\mathbf{q},-\mathbf{p}) = H(\mathbf{q},\mathbf{p})$.

Using Hamilton's equations, the time dependence of any phase variable $f(\mathbf{x})$ is

$$\frac{df}{dt} = \{f,H\}$$

where

$$\{f,g\} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} \right)$$

is called the **Poisson bracket**.

It is easy to see that $\{H,H\} = 0$, so Hamiltonian systems have a constant of motion given by the Hamiltonian function itself, normally corresponding to energy. Thus they have at least two zero Lyapunov exponents, one corresponding to the flow direction, and one for perturbations to the energy.

Hamiltonian systems may also have other constants of motion. In the case of Newtonian gravity, the energy, total momentum and total angular momentum are all constants of motion. It follows directly from Hamilton's equations that if H does not depend on q_i for some i, then p_i is constant. Then, properties such as ergodicity make sense only on the invariant surfaces where all the constants of motion are fixed. Also, there are zero Lyapunov exponents corresponding to both the q_i and p_i coordinates: Varying q_i leads to an identical system shifted in this coordinate, while a perturbation in p_i is unchanged.

A Hamiltonian with n degrees of freedom is called **Liouville integrable**, if there are n conserved quantities J_i including the Hamiltonian itself, that are functionally independent and have mutual zero Poisson brackets. In this case, the Hamiltonian may be transformed so it is a function only of the J_i as momenta, and so the corresponding coordinates have constant time-derivative. These are called **action-angle coordinates** and if the energy surface is compact, the dynamics in these coordinates is simply free motion on a torus.

A Poincaré section fixing a coordinate, say q_n of a Hamiltonian flow, and using a given constant energy E to determine p_n at each iteration, also leads to a symplectic map in the other variables. For example, a 2D billiard map is symplectic using the arc length s and component of the momentum parallel to the boundary, $p_\parallel = |p|\sin\theta$, where θ is the angle between the particle direction following a collision and the inward normal to the boundary, and furthermore the normalisation constants are consistent. This fact leads to an exact formula for the mean free path in billiards. Let us calculate the total phase space volume for the speed fixed to unity. Using the flow invariant measure this is $2\pi|D|$ where $|D|$ is the area of the billiard and the 2π corresponds to directions. This must be equal to the same quantity calculated using the above boundary invariant measure $2\bar{\tau}|\partial D|$ where the 2 is the domain of $\sin\theta$, $\bar{\tau}$ is the average time per collision, and $|\partial D|$ is the perimeter. Thus we have

$$\bar{\tau} = \frac{2\pi|D|}{2|\partial D|} = \frac{\pi|D|}{|\partial D|}$$

There are similar formulas is higher dimensions.

The most common Hamiltonian function is of the form

$$H(\mathbf{q},\mathbf{p}) = \sum_i \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{q})$$

for particles indexed by i with masses m_i moving with a potential energy function $V(\mathbf{q})$ that depends on the positions of all the particles, for example Newtonian gravity as discussed in the introduction. Another example we have seen is the simple pendulum, where q represents the angle from the lowest position, p the angular momentum, and

$$H(q,p) = \frac{p^2}{2ml^2} - mgl\cos q$$

where l is the length of the pendulum, m the mass of the bob and g the acceleration of gravity.

Numerical integration of Hamiltonian systems is most commonly performed using **symplectic integrators**. These use exactly symplectic maps that approximate the true (symplectic) dynamics, and hence retain phase space volume conservation, and conservation of a quantity very close to the real energy. In a **splitting method** the Hamiltonian is split into a sum of parts, for example kinetic (T) plus potential (V), each of which can be integrated by an exact symplectic map. The product of such symplectic maps is also symplectic, and approximates the true Hamiltonian dynamics.

If the time step is τ, a purely kinetic Hamiltonian (in the usual form) gives

$$\Phi_T^\tau(\mathbf{q},\mathbf{p}) = (\mathbf{q} + \tau\mathbf{p}/m, \mathbf{p})$$

while a purely potential Hamiltonian gives

$$\Phi_V^\tau(\mathbf{q},\mathbf{p}) = (\mathbf{q}, \mathbf{p} - \tau\nabla_{\mathbf{q}} V)$$

For small time step, the combined evolution can be approximated in terms of these, for example the **Störmer-Verlet algorithm**

$$\Phi^\tau \approx \Phi_T^{\tau/2} \circ \Phi_V^\tau \circ \Phi_T^{\tau/2}$$

7.3 The bouncer model

This was first introduced by Pustylnikov in 1983. Consider a plate vibrating with vertical position $y_0(t) = \epsilon\cos\omega t$ and a particle that moves above it $y(t)$ accelerating downwards due to a gravitational field $-g$ and making perfectly elastic collisions at times t_n with the plate. If the plate is assumed to have infinite mass, conservation of energy and momentum at collision leads to the rule

$$\dot{y}(t^+) = -\dot{y}(t^-) + 2\dot{y}_0(t)$$

Scaling the position and time, we can set $g = \omega = 1$, leaving a single parameter ϵ. We ignore air-resistance.

This system is non-autonomous, but the periodicity of the vibrations allows us to use either a time 2π map or collision map to reduce the problem to an autonomous map. We take the latter approach, describing the collisions by the phase at collision $\phi_n = t_n \bmod 2\pi$ and velocity subsequent to collision $v_n = \dot{y}(t_n^+)$.

The displacement from one collision to the next is

$$\epsilon\cos\phi_{n+1} - \epsilon\cos\phi_n = v_n\Delta_n - \frac{1}{2}\Delta_n^2$$

where $\Delta_n = t_{n+1} - t_n$. Also, the velocity decreases by an amount Δ_n during this time (since $g = 1$). Thus the outgoing velocities are related by

$$v_{n+1} = -(v_n - \Delta_n) - 2\epsilon\sin\phi_{n+1}$$

These equations cannot be solved explicitly, since they involve the intersection of a sinusoidal curve and a parabola. Numerical methods can use a Newton-like iteration to solve the equation numerically.[79]

A simpler and numerically more efficient approach makes an additional assumptions: Assume that v is much greater than ϵ. Then particle is moving fast enough so that it cannot be overtaken by the upwardly moving plate (no repeat collisions), and the distance it travels any time it is near the plate is also much greater than that of the plate. These constitute the **static wall approximation**. Thus we find $\Delta_n \approx 2v_n$, leading to

$$v_{n+1} = v_n - 2\epsilon\sin\phi_{n+1}$$

$$\phi_{n+1} = \phi_n + 2v_n$$

Note the subscript $n+1$ on the right hand side.

The original system appears Hamiltonian[80] but we need to check whether the symplectic (effectively just area preserving) property holds under the static wall approximation (SWA):

$$\begin{aligned} D\Phi &= \begin{pmatrix} \frac{\partial\phi_{n+1}}{\partial\phi_n} & \frac{\partial\phi_{n+1}}{\partial v_n} \\ \frac{\partial v_{n+1}}{\partial\phi_n} & \frac{\partial v_{n+1}}{\partial v_n} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 \\ -2\epsilon\cos(\phi_n + 2v_n) & 1 - 4\epsilon\cos(\phi_n + 2v_n) \end{pmatrix} \end{aligned}$$

which we see has unit determinant as required. The trace is $2 - 4\epsilon\cos\phi_{n+1}$.

Fixed points are of two types: For true fixed points we have ϕ and v multiples of π. These are elliptic if $\epsilon < 1$ and ϕ is an even multiple of π, and hyperbolic otherwise. Then, for $\epsilon > \pi/2$ there are travelling fixed points, called **accelerator modes** where v increases by a multiple of π at each step (if decreasing, clearly the SWA would become invalid). Again, stability is determined by the trace and may either be elliptic or hyperbolic. In the elliptic case, there is very likely a positive measure set where v increases without bound, the phenomenon of **Fermi acceleration**.

There are a number of similar systems in which a particle is injected into a periodically oscillating region after a time dependent on its energy. These include the Fermi-Ulam model of a particle between a fixed and oscillating wall $t \sim v^{-1}$, and the Kepler model of a comet in an eccentric orbit that spends most of its time far from the sun but is perturbed on its closest approach by Jupiter: The diffusing quantity is energy E, with $t \sim |E|^{-3/2}$. Unlike the bouncer model, however, the condition for stability of fixed points (which then influences the rest of phase space) in these models depends on the magnitude of the energy.

7.4 The standard map

We can treat the post-collision velocity v in the SWA bouncer model as an angle modulo 2π. An advantage of this is that the dynamical system now has a finite area (hence normalised invariant measure). A change of variables, $X_n = -\phi_{n+1}$, $Y_n = -2v_n$, $K = 4\epsilon$ leads to the equations of the **Standard map**:[81]

$$X_{n+1} = X_n + Y_{n+1}$$

$$Y_{n+1} = Y_n + K\sin X_n$$

For $K = 0$ the dynamics is a completely regular shear, while for $K \to \infty$ it appears (but has not been proven)

[79] See for example C. P. Dettmann and E. D. Leonel, Physica D **241** 403-408 (2012).

[80] We would need to check that the hard collisions make sense as the limit of a sequence of time-dependent potential energy functions.

[81] B. V. Chirikov, Nuclear Physics Institute of the Siberian section of the USSR Academy of Sciences Report 267 (1969).

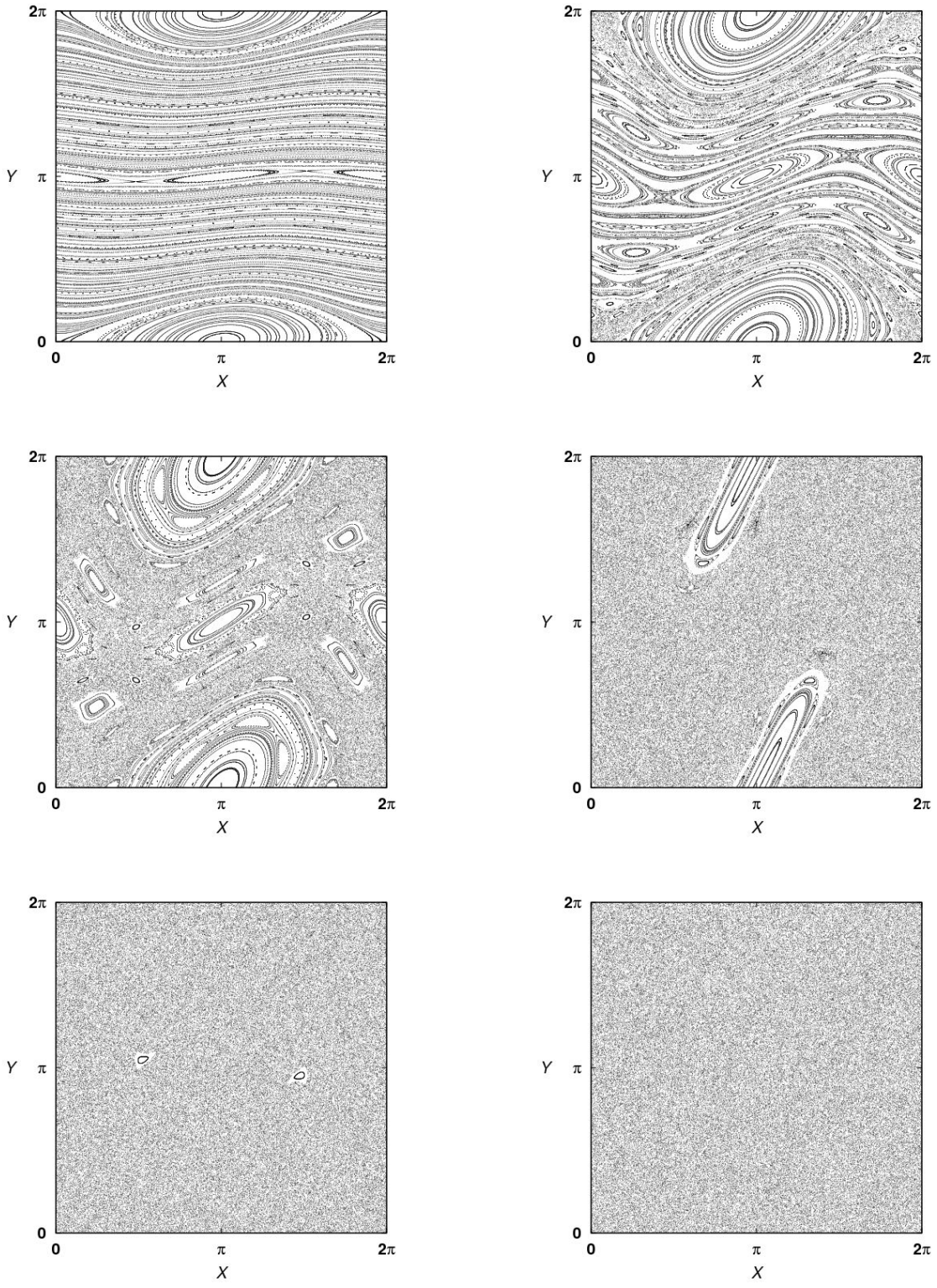

Figure 19: The Standard map, from left to right then top to bottom: $K = 0.2, 0.8, 1.2, 4, 6, 8$.

that the dynamics is completely chaotic. The transition to chaos occurs in several stages (see Fig.19).

At $K = 0$, Y is constant and the system is integrable, with the orbits in X having a rotation number $\omega = Y/(2\pi)$: If this rotation number is rational m/n, the orbits are n-periodic, while otherwise they are dense in the invariant curve $Y = \text{const}$.

For small K, Kolmogorov-Arnold-Moser (KAM) theory shows that most of the invariant curves remain, with exceptions only at resonances. To see this, consider both K and Y small, and replace differences by derivatives:

$$\dot{X} = Y$$

$$\dot{Y} = K \sin X$$

This (with $X = -\theta$) is just a pendulum, with conserved energy and Hamiltonian

$$H(X, Y) = \frac{Y^2}{2} + K \cos X$$

There are two fixed points connected by a separatrix of width $W = 4\sqrt{K}$ in the Y direction. Similarly, we may perturb around $Y = 2\pi m/n$; this leads to island chains of width $W \sim K^{n/2}$. Inside the separatrices all orbits have rotation number m/n. Around the separatrices are chaotic regions arising from homoclinic tangles. Outside the separatrices, there remain invariant curves with rotation numbers that are badly approximable by rationals.[82]

The most "irrational" number is the golden ratio $g = (1+\sqrt{5})/2$, so its invariant curve is the last to be destroyed as K increases. The Chirikov criterion suggests that this happens around when the main island chains overlap, that is $W = 2\pi$ or $K = \pi^2/4 \approx 2.467$. This is however rather inaccurate: The transition actually occurs at $K \approx 0.9716$. Beyond this point, orbits in the chaotic regions can diffuse throughout the system, and in the bouncer model, up to arbitrarily high velocity.

The fixed point at zero becomes unstable at $K = 4$, and beyond this point there are only small elliptic islands visible in phase space. For many large values of K there are no islands visible, but there are many open questions,[83] for example

> **For any fixed K, show that there is a set of positive measure of orbits with positive entropy.**

[82] These have continued fraction expansions with coefficients that are not too large. See for example A. M. Rockett and P. Szüsz *Continued fractions* (World Scientific, 1992).

[83] See for example A. Giorgilli and V. F. Lazutkin, Phys. Lett. A **272**, 359-367 (2000). Note that it is possible to slightly perturb the standard map to obtain positive entropy: P. Berger and D. Turaev, arxiv:1704.02473

