PHY 411 Advanced Classical Mechanics (Chaos) U. of Rochester Spring 2002, Fall 2006

S. G. Rajeev

October 12, 2006

Contents

Chapter 1

Iterations of Maps

1 Given $f : M \to M$, a continuous map of a topological space to itself we define $f^n(x) = f(f^{n-1}(x)), f^1(x) = f(x)$ for $n = 1, 2, \cdots$ and also $f^0(p) = p$.

1.1 The sequence $(x, f(x), f(f(x)), f(f(f(x))), \cdots)$ is called the orbit of p.

2 A point $x \in M$ is a *fixed point* if $f(x) = x$.

2.1 Then its orbit is just the sequence $x, x, \cdots$.

3 A point is x is *periodic with period* p if

$$f^p(x) = x, f^k(x) \neq x \text{ for } k < p.$$

3.1 That is, p is the smallest number of steps in which x will return to itself.

4 We also define $N_n(f)$ to be the number of solutions to the equation

$$f^n(x) = x$$

whenever this is finite.

5 The *dynamical zeta function* of a map is a kind of generating function of the sequence $N_n(f)$:

$$\zeta_f(z) = \exp(\sum_{n=1}^{\infty} \frac{z^n}{n} N_n(f)).$$

5.1 The enumeration of periodic points is a basic problem in the study of iterations of any map. The zeta function is a convenient way of encoding this information.

6 In general it is too difficult to predict where a point will end up after many iterations. The best we can do us to estimate the probability that a particular point will be the outcome.

7 A probability density (also called a probability measure) is a positive function such that

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

Under a change of co-ordinates $x \mapsto f(x)$ it transforms as $f^*(\rho)(x) = \sum_{x=f(y)} \frac{\rho(y)}{|f'(y)|}$. Here $f'(y)$ denotes the Jacobian of the transformation.

7.1 Proof The meaning of the probability density is that the average of any function of the random variable x is

$$< g(x) >= \int \rho(x)g(x)dx.$$

For example the mean value of x itself is $< x >= \int x\rho(x)dx$. Now suppose we set $x_1 = f(x)$ and ask for a density $\rho_1(x)$ such that

$$< g(f(x)) >= \int \rho_1(x)g(x)dx.$$

That is

$$\int g(f(y))\rho(y)dy = \int \rho_1(x)g(x)dx;$$

If $f(y) = x$ has a single solution and $f'(x) > 0$, by a change of variable we get

$$\int g(x)\rho(y)\frac{dy}{dx}dx = \int \rho_1(x)g(x) \Leftrightarrow \rho_1(x) = \frac{\rho(f^{-1}(x))}{f'(x)}.$$

If there is still a single solution but $f'(x) < 0$, we would have to interchange the limits of integration so we would get the abobe answer up to a change of sign:

$$\rho_1(x) = \frac{\rho(f^{-1}(x))}{|f'(x)|}$$

If there are several solutions for $x = f(y)$, we will have to split up the region of integration into intervals within each of which there is one solution; the answer is given by summing over all such regions.Hence in general

$$f^*(\rho)(x) = \sum_{x=f(y)} \frac{\rho(y)}{|f'(y)|}$$

8 We can now look at the effect of the iterations of the map f on the density. If we have a density that is invariant under this map to which any initial density will converge we have a simple situation: we can predict an eventual probability distribution for the points.

9 As an example we study the maps of a circle to itself.

9.1 A circle can be thought of as the set of real numbers taken modulo 1 . Thus a point on the circle can be represented as a real number between 0 and 1 ; any number of magnitudde greater than one will correspond uniquely to a number in this region by adding an integer.

9.2 Now consider the map

$$f(x) = 2x \text{ mod } 1.$$

That is, we double each number and if the outcome is greater than one, we just drop the 'fractional' part.

9.3 Clearly $x = 0$ is a fixed point. $x = \frac{1}{2}$ will converge to this point after one iteration; its orbit is $\frac{1}{2}, 0, 0, \cdots$.

9.4 It will be convenient to represent a real number in its binary expansion:

$$x = 0.a_1 a_2 a_3 \cdots$$

where each a_k is either zero or one. Thus $\frac{1}{2} = 0.1, \frac{3}{4} = 0.11$ etc.

9.5 This representation is unique except for one type of situation: an infinite string of 1's should be replaced by a string of zero's and replacing the entry just before this string by 1. for example,

$$0.01111\cdots = 0.1$$

(Sum this geometric series to prove this.) Every rational number will terminate in a repeating sequence. For example,

$$0.\bar{1}, \quad 0.1101\overline{1}\overline{1}\overline{0}, \cdots$$

are all rational numbers. The bar denotes a sequence that is repeated for ever.

9.6 Multiplication by two just shifts the 'binary point' one step to the right; taking this number modulo 1 simply means we replace everything to the left of the binary point by 0.

9.7 Now it is easy to find the orbit of any point on the circle. If the starting number is rational, after a finite number of iterations we will settle to a repeating pattern. In the simplest cases this will be a fixed point; i.e., when the number terminates with zeros. Otherwise it will be the repetition of some sequence over and over again.

9.8 If the starting point is irrational, there will be no repeating pattern: the orbit will wander all over the circle; indeed there will be no predictability to this. If $x < y$ it does not follow that $f(x) < f(y)$: the numbers will get thoroughly reshuffled. This is a simple example of chaos.

10 The effect of densities is simple to work out:

$$f^{*}\rho(x) = \frac{1}{2}\left[f\left(\frac{x}{2}\right) + f\left(\frac{x}{2} + \frac{1}{2}\right)\right].$$

11 A simple example of a chaotic system was constructed by Arnold as the iteration of a linear transformation on a torus.

11.1 It is the interplay between periodicity and the linearity that leads to chaos.

11.2 The simplest chaotic mechanical system is a linear transformation on a torus:

$$f : \begin{pmatrix} p \\ q \end{pmatrix} \mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} \mod 1$$

This will preserve the area if $ad - bc = 1$. We again think of the co-ordinates modulo 1.

11.3 This is called *Arnold's cat map*. Arnold showed that the outcome of these maps is chaotic by showing its effect on the picture of a cat[1]. For other graphical realizations see for example
http://www.physics.iitm.ac.in/ arul/serc/serc1.html

[1]No cats were harmed in this study.

Chapter 2

Dynamics on paths of a graph

12 It is useful to study a finite model of dynamics represented by a graph.

13 A *transition matrix* is a square matrix whose entries are either 0 or 1 . Each such matrix defines a graph: the numbers of vertices is equal to the dimension of the matrix and there is a directed line connecting the k th point to the l th point $A_{kl} = 1$ but not if $A_{kl} = 0$.

13.1 What we call a graph for short is more properly called a *finite directed graph*.

13.2 A graph can be thought of a crude approximation to a classical mechanical syste,: the vertices represent the states (now reduced to a finite number) and the edges the equation of motion. You could imagine dividing up the phase space into a finite number of boxes; then see if there is any initial condition contained in the j th box that take you inside the k th box; if so define $A_{jk} = 1$, and $A_{jk} = 0$ otherwise. Now, as the boxes become smaller (the number of boxes will become larger) we will get a better approximation to the classical dynamics.

14 An *orbit* on a graph is an infinite sequence $s_0 s_1 \cdots$ such that any pair of neighboring elements is an edge.; i.e., $A_{s_j s_{j+1}} = 1$ for all $j = 0, 1 \cdots$. We will denote the set of all orbits by Λ_A .

15 A graph determines a dynamical system $\sigma_A : \Lambda_A \to \sigma_A$, through the shift map:

$$[\sigma_A(s)]_j = s_{j+1}.$$

16 A fixed point of σ_A^n is simply a closed path of length n repeated over and over:

$$\sigma_A^n(s) = s \iff s = s_0 s_1 \cdots s_{n-1} s_0 s_1 \cdots s_{n-1} \cdots$$

16.1 Thus $N_n(\sigma_A)$ is the number of closed paths of length n .

17 The dynamical zeta function of this system can be shown to be

$$\zeta_{\sigma_A}(z) = \frac{1}{\det[1 - zA]}.$$

17.1 To see this, note for example that

$$\sum_{ij} A_{ij} A_{ji} = \operatorname{tr} A^2$$

is $N_2(\sigma_A)$, the number of closed paths of length two. The summand is either one or zero: for it to be one, ij and ji must both be edges. Similarly $N_n = \operatorname{tr} A^n$. Now we see that

$$\zeta_A(z) = \exp(\sum_{n=1}^{\infty} \frac{z^n}{n} \operatorname{tr} A^n) = \exp(-\operatorname{tr} \log[1 - zA])$$

Now we recall that

$$\operatorname{tr} \log M = \log \det M$$

for any matrix with non-zero determinant.

17.2 Example: If $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ corresponds to the graph with two vertices where any ordered pair of vertices are connected (the complete directed graph with two vertices). Then Λ_A is the space of all inifinite sequences in two letters. The zeta function is easily calculated from the two by two determinant $\frac{1}{1-2z}$. We can then read off that the numbet of fixed points is 2^n . This is correct, as this is the number of sequences of length n .

17.3 If the radius of convergence of $\zeta_{\sigma_A}(z)$ is the reciprocal of $|A|$. ($|A|$ is the magnitude of the largest eigenvalue of A .) Also, $N_n(\sigma_A) \sim |A|^n$ as $n \to \infty$.

Chapter 3

Trace Formulas

18 It is often easier to work with the generating function

$$G_f(z) = \frac{d\log\zeta_f(z)}{d\log z} = \sum_{n=1}^{\infty} z^n N_n(f).$$

18.1 For example, for the discrete dynamical system obtained from a graph,

$$G(z) = \operatorname{tr}\frac{zA}{1-zA}.$$

sums (or traces) are easier than products. Clearly $G(z)$ determines the dynamical zeta function through the above differential equation.

19 When the number of states is continuously infinite, the trace becomes an integral; the eigenvalues are determined by solving an integral equation.

20 We can define an analogue of the transition matrix

$$A_f(x,y) = \delta(x - f(y));$$

this is an integral operator

$$A_f\psi(x) = \int A_f(x,y)\psi(y)dy = \sum_{y \text{ suchthat } f(y)=x} \frac{\psi(y)}{|\det f'(y)|}$$

21 Geometrically this means that ψ transforms as a density.

22 Then we can define

$$\mathrm{tr}A_f = \int A_f(x,x)dx = \sum_{x|f(x)=x} \frac{1}{|\det[f'(x)]|}$$

which makes sense if there are only a finite number of fixed points and if each of them is isolated: $\det f' \neq 0$. This is a sum over fixed points weighted by the 'stretching' of the neighborhood of each fixed point caused by f .

23 We can now define a generating function

$$G_f(z) = \mathrm{tr} zA_f[1 - zA_f]^{-1}.$$

This is a weighted sum over the fixed points of f^n . It can be determined by solving the eigenvalue problem of the above integral operator:

$$\sum_{y|f(x)=y} \frac{\psi(y)}{|\det f'(y)|} = \lambda\psi(x).$$

Assuming that there are a countable number of solutions to this equation,

$$G_f(x) = \sum_r \frac{z\lambda_r}{1 - z\lambda_r}, \quad \zeta_f = \frac{1}{\prod_r [1 - z\lambda_r]}.$$

24 Example: Let $f : [0,1] \to [0,1]$ be the Bernoulli shift, $f(x) = 2x \bmod 1$. Then

$$A_f\psi(x) = \frac{1}{2}\left[\psi(\frac{1}{2}x) + \psi(\frac{x+1}{2})\right];$$

This is the average of the pre-images of x . The eigenvaue equation

$$A_f\psi(x) = \lambda\psi(x) \iff \lambda\psi(x) = \frac{1}{2}\left[\psi(\frac{1}{2}x) + \frac{1}{2}\psi(\frac{x+1}{2})\right].$$

25 The solutions involve the Bernoulli polynomials; they are defined by the generating function

$$Z_t(x) = \sum_{n=0}^{\infty} B_n(x)\frac{t^n}{n!} = \frac{te^{xt}}{e^t - 1}.$$

so that $B_0(x) = 1, B_1(x) = x - \frac{1}{2}$ etc. If we calculate using the definition above, we get

$$A_f Z_t(x) = Z_{\frac{t}{2}}(x), \Rightarrow A_f B_n(x) = \frac{1}{2^n} B_n(x).$$

Thus the Bernoulli polynomials are the eigenfunctions with the eigenvalues 2^{-n} .

26 Thus we get

$$\zeta_f(z) = \prod_{n=0}^{\infty} \frac{1}{1 - z2^{-n}}.$$

Chapter 4

Evolution of Probability Density

27 Given a function that maps a domain of R^d to itself, $f : \Omega \rightarrow \Omega$ we define the Perron-Frobenius operator

$$A_f \rho(x) = \int \delta(x - f(y))\rho(y)dy$$

28 If the Jacobian $f'(x)$ is non-zero everywhere we can write this also as

$$A_f \rho(x) = \sum_{\{y|x=f(y)\}} \frac{\rho(y)}{|f'(y)|}$$

28.1 If ρ is a probability density (i.e., $\rho(x) \geq 0, \int \rho(x)dx = 1$) so will be $A_f \rho$.

28.2 The effect of a large number of iterations on ρ is of interest: it tells us what happens to a randomly chosen initial condition as 'time' goes to infinity.

28.3 If we can find eigenfunctions of A_f

$$A_f \psi_n(x) = \lambda_n \psi_n(x)$$

such that any density can be expanded in this basis

$$\rho(x) = \sum_k c_k \psi_n(x)$$

the time evolution is easy to describe:

$$A_f^n \rho(x) = \sum_k \lambda_k^n c_k \psi_k(x).$$

28.4 The eigenvalues may not be real; the eigenfunctions may not be positive or even real either. Nevertheless the coefficients c_n will appear in complex conjugate pairs so that the sum is real.

28.5 The eigenfunction of the largest eigenvalue has a special significance: it is the one to which a generic initial condition will evolve. It defines the invariant measure of the dynamics.For,if λ_0 is the eigenvalue of the largest magnitude (and if it is non-degenerate; i.e., there is only one eigenvector corresponding to it) we have

$$A_f^n \rho(x) = \lambda_0^n \left[c_0 \psi_0(x) + \sum_{k \neq 0} \left(\frac{\lambda_k}{\lambda_0} \right)^n c_k \psi_k \right]$$

As $n \to \infty$, the coeffcients of the sum will approach zero. Then only the largest eigenvalue matters. Indeed we suspect that in 'good' cases we would have $\lambda_0 = 1$ and $\psi_0(x) \geq 0$.

29 Example: Let $f : [0,1] \to [0,1]$ be the Bernoulli shift, $f(x) = 2x \bmod 1$. Then

$$A_f \psi(x) = \frac{1}{2} \left[\psi(\frac{1}{2}x) + \psi(\frac{x+1}{2}) \right];$$

This is the average of the pre-images of x . The eigenvalue equation

$$A_f \psi_k(x) = \lambda_k \psi_k(x) \iff \lambda_k \psi_k(x) = \frac{1}{2} \left[\psi_k(\frac{1}{2}x) + \frac{1}{2} \psi_k(\frac{x+1}{2}) \right].$$

30 Introduce the generating function

$$Z_t(x) = \sum_k \psi_k(x)\frac{t^k}{k!}$$

Then

$$A_f Z_t(x) = \frac{1}{2}\left[Z_t(x) + Z_t(\frac{x+1}{2})\right]$$

This suggests the ansatz

$$Z_t(x) = z(t)e^{xt}, \Rightarrow A_f Z_t(x) = z(t)\frac{1}{2}[1 + e^{\frac{t}{2}}]e^{\frac{xt}{2}}$$

If we also put

$$\lambda_k = 2^{-k}$$

the eigenvalue equation reduces to

$$\frac{1}{2}[1 + e^{\frac{t}{2}}]z(t) = z(\frac{t}{2})$$

which is solved by

$$z(t) = \frac{t}{e^t - 1}.$$

31 If we expand we see that the eigenfunctions are a family of polynomials called the Bernoulli polynomials:

$$Z_t(x) = \sum_{n=0}^{\infty} B_n(x)\frac{t^n}{n!} = \frac{te^{xt}}{e^t - 1}.$$

so that $B_0(x) = 1, B_1(x) = x - \frac{1}{2}$ etc.

32 Thus we get the generating function for cycles to be

$$\zeta_f(z) = \prod_{n=0}^{\infty} \frac{1}{1 - z2^{-n}}.$$

32.1 Thus in this case the largest eigenvalue is one and the corresponding eigenfunction is positive (the uniform distribution). How general is this? What are some natural sufficient conditions for which a linear operator has a unique eigenvector of eigenvalue one? When will this eigenvector be positive? Will the other eigenvalues be less than one?

33 A is a *Markov matrix* if each of its entries are positive and each row adds up to one: $\sum_i A_{ij} = 1$. A_{ij} can be thought of the probability that system makes a transition to state i given that it is in state j.

34 A Markov matrix of finite dimension has an eigenvector all of whose components are positive and whose eigenvalue is unity.

34.1 Proof One proof of this fact uses the Brower fixed point theorem of topology. Consider the set of vectors whose components are positive and add up to one:

$$\Delta_N = \{(p_1, \cdots p_N)|p_i \geq 0, \sum_i p_i = 1.\}$$

Geometrically this is a *simplex*. For $N = 2$ it is just the closed unit interval; for $N = 3$ it is the triangle *including its boundary*, for $N = 4$ a tetrahedron and so on. An $N \times N$ Markov matrix is a continuous map of Δ_N to itself:

$$p_i \mapsto [Ap]_i = \sum_j A_{ij} p_j.$$

Since each entry of $A_{ij} > 0$ the sum is also positive; also by interchanging sums

$$\sum_i [Ap]_i = \sum_i \sum_j A_{ij} p_j = \sum_j p_j = 1.$$

Now, as a topological space, the simplex is equivalent to a closed ball; (i.e., it is homeomorphic; there is a continuous function with continuous inverse which maps the simplex to the closed unit ball). The Brouwer fixed point theorem then guarantees there is a fixed point; i.e., a vector satisfying

$$Ap = p$$

This is an eigenvector with eigenvalue one and with positive components.

35 Brouwer Fixed Point Theorem

Every continuous function from the closed unit ball D_n to itself has at least one fixed point.

35.1 If the matrix can be broken up into block upper triangular (also called reducible) form $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$ we would not have always have a unique eigenvector with eigenvale one: we could have eigenvectors in each block separately. But such a matrix would describe an evolution that leaves a certain subset of states invariant: the problem can then be broken up into smaller matrices, one for each such subset. Thus, we are interested in matrices which cannot be brought to this form by permutation of co-ordinates: they are said to be *irreducible*.

36 Perron-Frobenius Theorem

If A has nonnegative entries, is irreducible, then

1. one of its eigenvalues is positive and greater than or equal to (in absolute value) all other eigenvalues

2. there is a positive eigenvector corresponding to that eigenvalue

3. that eigenvalue is a simple root of the characteristic equation of A; i.e., its eigenspace is one dimensional.