ASYMPTOTIC PERIODICITY NOTES
10 MARCH, 1995
FILE: APNOTES.TEX

Michael C. Mackey

Centre for Nonlinear Dynamics in Physiology and Medicine
Departments of Physiology, Physics and Mathematics
McGill University
Montreal, Quebec, Canada

I. WHAT DON'T I KNOW?
NOTES OF 12 MARCH, 1993

(1) Can the period ever depend on the initial density f_0?
(2) Can the period ever depend on ϵ in a CML? [YES–SEE WORK WITH JEROME LOSSON.]
(3) How many of the asymptotic periodicity properties carry over to a lattice?
(4) Could I ever have asymptotic periodicity with period 3^n, 5^n, etc? [YES–LUBKIN (KIEV) HAD AN EXAMPLE IN COMO.]
(5) Is it possible to have asymptotic periodicity when the basis densities $g_i(x)$ do not have disjoint support? It violates the current definition of asymptotic periodicity as now given.

WHAT DO I KNOW?
13 MARCH, 1993

(1) Period $T \leq r!$.
(2) If $\exists\, h \ni \lim_{n\to\infty} P^{nT} f \leq h$, then $r \leq ||h||$ (LM p89-90].
(3) $\exists$ a stationary density $f_* = \frac{1}{r}\sum g_i$ [LM, Proposition 5.4.1].
(4) If P has a constant stationary density f_*, then

$$P^{t+1}f(x) = \sum_{i=1}^{r} \lambda_{\alpha^{-t}(i)}(f)\bar{1}_{A_i}(x) + Q_t f(x)$$

with

$$\bar{1}_{A_i}(x) = \frac{1_{A_i}(x)}{\mu(A_i)}.$$

[LM Proposition 5.4.2].
(5) P ergodic $\iff$ permutation is cyclical [LM Theorem 5.5.1].
(6) $r = 1 \implies$ asymptotic stability of P [LM Proposition 5.5.2].
(7) P mixing $\implies r = 1$ [LM Theorem 5.5.3].
(8) $\lim_{t\to\infty} H_c(P^t f|f_*) = H_{max}(f|f_*)$ [Arrow, Theorem 6.5].
(9) Correlation function is periodic [Arrow, Chapter 6D and Provatas Mackey papers].
(10) H_{BG} has period T

Proof. *For large times we can write*

$$P^{t+1}f(x) = \sum_{i=1}^{r} \lambda_{\alpha^{-t}(i)}(f)g_i(x)$$

so

$$\begin{aligned}
H(P^{t+1}f) &= -\int_X \sum_{i=1}^{r} \lambda_{\alpha^{-t}(i)}(f) g_i(x) \log\left[\sum_{i=1}^{r} \lambda_{\alpha^{-t}(i)}(f) g_i(x)\right] dx \\
&= -\sum_{i=1}^{r} \lambda_{\alpha^{-t}(i)}(f) \int_{A_i} g_i(x) \log\left[\lambda_{\alpha^{-t}(i)}(f) g_i(x)\right] dx \\
&= -\sum_{i=1}^{r} \lambda_{\alpha^{-t}(i)}(f) \left[\int_{A_i} g_i(x) \log \lambda_{\alpha^{-t}(i)}(f) dx + \int_{A_i} g_i(x) \log g_i(x)\right] dx \\
&= -\sum_{i=1}^{r} \lambda_{\alpha^{-t}(i)}(f) \left[\log \lambda_{\alpha^{-t}(i)}(f) - \int_{A_i} g_i(x) \log g_i(x)\right] dx \\
&= -\sum_{i=1}^{r} \lambda_{\alpha^{-t}(i)}(f) \log \lambda_{\alpha^{-t}(i)}(f) + \sum_{i=1}^{r} \lambda_{\alpha^{-t}(i)}(f) H(g_i)
\end{aligned}$$

Now the first term on the last line is independent of time, but the second isn't, and in fact it oscillates with period T.

N ote the strange fact that although H_{BG} oscillates, H_c approaches a stationary value [Arrow, Theorem 6.5–see item 8 above].

****

CHAPTER 6.
ASYMPTOTIC PERIODICITY AND ENTROPY EVOLUTION

In this chapter we turn to an investigation of the fascinating property of asymptotic periodicity in the evolution of densities. This behaviour is the statistical analog for densities of the more common periodicity found in some time series. The existence of asymptotic periodicity will allow us to prove a weak form of the Second Law in which the conditional entropy increases to (at least) a local maximum.

In Section A we introduce a class of Markov operators known as smoothing. Smoothing operators have three characteristics that are important for our ultimate understanding of the basis of the Second Law of thermodynamics. First, the sequence of densities evolving under the action of a smoothing Markov operator has the property of asymptotic (or statistical) periodicity. This is illustrated in Section B using the hat and quadratic maps. Second, any smoothing Markov operator has at least one stationary density thus ensuring that there is a state (perhaps not unique) of thermodynamic equilibrium. In Section C we show how, for asymptotically periodic systems, the entropy of the sequence of densities always increases to a maximum. This maximum, however, may only be relative and less than the maximum possible entropy value, thus corresponding to a metastable state. The relative maximum of entropy which asymptotically periodic systems approach usually depends on the initial density of the system (the way in which the system was prepared). In Section D we show that the correlation function for an asymptotically periodic system is made up of a stochastic component and a strictly periodic (nondecreasing) component.

A. ASYMPTOTIC PERIODICITY.

First, we define a smoothing Markov operator. A Markov operator P^t is said to be **smoothing** if there exists a set A of finite measure, and two positive constants $k < 1$ and $\delta > 0$ such that for every set E with $\mu_L(E) < \delta$ and every density f there is some integer $t_0(f, E)$ for which

$$\int_{E\cup(X\setminus A)} P^t f(x)\,dx \le k \qquad \text{for } t \ge t_0(f,E).$$

This definition implies that any initial density, even if concentrated on a small region of the phase space X, will eventually be smoothed out by P^t and not end up looking looking like a delta function. Notice that if X is a finite phase space we can take $X = A$ so the smoothing condition looks simpler:

$$\int_E P^t f(x)\,dx \le k \qquad \text{for } t \ge t_0(f,E).$$

Smoothing operators are important because of a theorem of Komornik and Lasota (1987), first proved in a more restricted situation by Lasota, Li, and Yorke (1984).

Theorem 6.1. *Spectral Decomposition Theorem (Komornik and Lasota, 1987). Let P^t be a smoothing Markov operator. Then there is an integer $r > 0$, a sequence of nonnegative densities g_i, a sequence of bounded linear functionals λ_i, $i = 1, \dots, r$, and an operator $Q : L^1 \to L^1$ such that for all densities f, Pf has the form*

$$Pf(x) = \sum_{i=1}^{r} \lambda_i(f) g_i(x) + Qf(x). \tag{6.1}$$

The densities g_i and the transient operator Q have the following properties:

1. *The g_i have disjoint support (i.e. are mutually orthogonal and thus form a basis set), so $g_i(x)g_j(x) = 0$ for all $i \neq j$.*
2. *For each integer i there is a unique integer $\alpha(i)$ such that $Pg_i = g_{\alpha(i)}$. Furthermore, $\alpha(i) \neq \alpha(j)$ for $i \neq j$. Thus the operator P permutes the densities g_i.*
3. *$\| P^t Qf \| \to 0$ as $t \to \infty$, $\quad t \in N$.*

Notice from equation (6.1) that $P^{t+1}f$ may be written in the form

$$P^{t+1}f(x) = \sum_{i=1}^{r} \lambda_i(f) g_{\alpha^t(i)}(x) + Q_t f(x), \qquad t \in N \tag{6.2}$$

where $Q_t = P^t Q$, $\| Q_t f \| \to 0$ as $t \to \infty$, and $\alpha^t(i) = \alpha(\alpha^{t-1}(i)) = \cdots$. The density terms in the summation of (6.2) are just permuted by each application of P. Since r is finite, the series

$$\sum_{i=1}^{r} \lambda_i(f) g_{\alpha^t(i)}(x) \tag{6.3}$$

must be periodic with a period $T \leq r!$. Further, as

$$\{\alpha^t(1), \dots, \alpha^t(r)\}$$

is just a permutation of $1, \cdots, r$ the summation (6.3) may be written in the alternative form

$$\sum_{i=1}^{r} \lambda_{\alpha^{-t}(i)}(f) g_i(x),$$

where $\alpha^{-t}(i)$ is the inverse permutation of $\alpha^t(i)$.

This rewriting of the summation portion of (6.2) makes the effect of successive applications of P completely transparent. Each application of P simply permutes the set of scaling coefficients associated with the densities $g_i(x)$ [remember that these densities have disjoint support].

Since T is finite and the summation (6.3) is periodic (with a period bounded above by $r!$), and $\| Q_t f \| \to 0$ as $t \to \infty$, we say that for any smoothing Markov operator the sequence $\{P^t f\}$ is **asymptotically periodic** or, more briefly, that P is asymptotically periodic. Komornik (1991) has recently reviewed the subject of asymptotic periodicity.

One interpretation of equation (6.2) is that *any asymptotically periodic system is quantized from a statistical point of view.* Thus if t is large enough, which simply means that we have observed the system longer than its relaxation time so $\| Q_t f \|$ is approximately zero, then

$$P^{t+1}f(x) \simeq \sum_{i=1}^{r} \lambda_i(f) g_{\alpha^t(i)}(x).$$

Asymptotically, $P^t f$ is either equal to one of the basis densities g_i of the i^{th} pure state, or to a mixture of the densities of these states, each weighted by $\lambda_i(f)$. It is important to also realize that the limiting sequence $\{P^t f\}$ is, in general, dependent on the choice of the initial density f.

How would the property of asymptotic periodicity be manifested in a continuous time system? If t is continuous, $t \in R^+$, then for every t we can find a positive integer m and a number $\theta \in [0, 1]$ such that $t + 1 = m + \theta$. Then, asymptotically

$$P^{t+1}f(x) = P^m(P^\theta f) \simeq \sum_{i=1}^{r} \lambda_{\alpha^m(i)}(P^\theta f) g_i(x).$$

Now, in the continuous time case we expect that there will be a periodic modulation of the scaling coefficients λ dependent on the initial density f, and the asymptotic limiting density will continue to display the quantized nature characteristic of the discrete time situation. This behaviour has been discovered and studied by Losson (1991) in differential delay equations.

Asymptotically periodic Markov operators always have at least one stationary density given by

$$f_*(x) = \frac{1}{r}\sum_{i=1}^{r} g_i(x), \tag{6.4}$$

where r and the $g_i(x)$ are defined in Theorem 6.1. It is easy to see that $f_*(x)$ is a stationary density, since by Property 2 of Theorem 6.1 we also have

$$Pf_*(x) = \frac{1}{r}\sum_{i=1}^{r} g_{\alpha(i)}(x),$$

and thus f_* is a stationary density of P^t. Therefore, for any smoothing Markov operator the stationary density (6.4) is just the average of the densities g_i.

Our next theorem will be very useful in Chapter 10 when we study the entropy behaviour of discrete time systems placed in contact with a heat bath.

Theorem 6.2. *Let P be a Markov operator. If there exists an $h \in L^1$ and $\gamma < 1$ such that*

$$\limsup_{t\to\infty} ||(P^t f - h)^+|| \le \gamma \qquad \text{for} \quad f \in D, \tag{6.5}$$

then $\{P^t f\}$ is asymptotically periodic.

Proof. Let $\epsilon = \frac{1}{4}(1-\gamma)$ and take $\mathcal{F} = \{h\}$. Since $\mathcal{F}$, which contains only one element, is evidently weakly precompact, then by WPC3 of Chapter 3 there exists a $\delta > 0$ such that

$$\int_E h(x)\,\mu(dx) < \epsilon \qquad \text{for} \quad \mu(E) < \delta. \tag{6.6}$$

Furthermore, there is a measurable set A of finite measure for which

$$\int_{X\setminus A} h(x)\,\mu(dx) < \epsilon. \tag{6.7}$$

Now fix $f \in D$. From (6.5) we may choose an integer $n_0(f)$ such that

$$||\,(P^t f - h)^+\,|| \le \gamma + \epsilon \qquad \text{for} \quad t \ge t_0(f),$$

and, as a consequence

$$\int_C P^t f(x)\,\mu(dx) \le \int_C h(x)\,\mu(dx) + \gamma + \epsilon \qquad \text{for} \quad t \ge t_0(f) \tag{6.8}$$

for an arbitrary set C. Setting $C = E \cup (X \setminus A)$ in (6.8) and using (6.6) and (6.7) we have

$$\begin{aligned}\int_{E\cup(X\setminus A)} P^t f(x)\,\mu(dx) &\le \int_E h(x)\,\mu(dx) + \int_{X\setminus A} h(x)\,\mu(dx) + \gamma + \epsilon \\ &< 3\epsilon + \gamma = 1 - \epsilon \qquad \text{for} \quad t \ge t_0(f).\end{aligned}$$

Thus P is smoothing. This, in conjunction with Theorem 6.1, completes the proof. □

The interpretation of Theorem 6.2 is straightforward. Namely, for those regions where $P^t f > h$ for sufficiently large t, if the area of the difference between $P^t f$ and h is bounded above by $\gamma < 1$, then $\{P^t f\}$ is asymptotically periodic.

We close this section with the statement and proof of a necessary and sufficient condition for the ergodicity of a smoothing Markov operator.

Theorem 6.3. *Let P be an asymptotically periodic Markov operator in a normalized measure space. Then P is ergodic if and only if the permutation $\alpha(i)$ of the Spectral Decomposition Theorem 6.1 is cyclical.*

Proof. We start with the proof that when P is ergodic then $\alpha(i)$ must be a cyclical permutation. Suppose that the disjoint supports of the r densities $g_i(x)$ are labeled by A_i, $i = 1, \cdots, r$. Assume that $\alpha(i)$ is not cyclical so there is an invariant subset $I \in \{\alpha(i)\}$. As a consequence, there is at least one set A_i that is invariant, and since the supports of the densities $g_i(x)$ have positive measure we conclude that there is an invariant subset of the phase space X that is not trivial. This contradicts the definition of ergodicity, so when P is ergodic the permutation $\alpha(i)$ must be cyclical.

To prove the converse, that if $\alpha(i)$ is a cyclical permutation then P is ergodic, we first use the spectral decomposition of $P^t f$ given by equation (6.1) to write the system state density average (3.15a) as

$$A_t f(x) = \sum_{i=1}^{r} g_i(x) \frac{1}{t} \sum_{k=0}^{t-1} \lambda_{\alpha^{-k}(i)}(f) + \frac{1}{t} \sum_{k=0}^{t-1} Q_k f(x).$$

Now the limit

$$\bar{\lambda}_i(f) = \lim_{t \to \infty} \frac{1}{t} \sum_{k=0}^{t-1} \lambda_{\alpha^{-k}(i)}(f)$$

exists because the cyclicity of the permutation $\alpha(i)$ of the set $\{1, \cdots, r\}$ implies the periodicity of the $\lambda_{\alpha^{-k}(i)}(f)$. Furthermore, since every portion of this summation of length r consists of exactly the same set of numbers but in a different order for each different i, it is clear that the limit $\bar{\lambda}_i(f)$ is, in fact, independent of i. Call it $\bar{\lambda}(f)$. Thus, from the Spectral Decomposition Theorem 6.1, we have that

$$\lim_{t \to \infty} A_t f(x) = \bar{\lambda}(f) \sum_{i=1}^{r} g_i(x).$$

Since $\lim_{t \to \infty} A_t f$ is a density, integrating over the entire phase space X gives

$$\int_X \lim_{t \to \infty} A_t f(x)\, dx = r\bar{\lambda}(f) = 1,$$

so $\bar{\lambda}(f) = \frac{1}{r}$ and

$$\lim_{t \to \infty} A_t f(x) = \frac{1}{r} \sum_{i=1}^{r} g_i(x) \equiv f_*(x),$$

which is a stationary density of the asymptotically periodic Markov operator P. Thus, $\{P^t f\}$ is Cesàro convergent to a unique stationary density f_* and P is ergodic by Theorem 4.7. This finishes the proof. □

This theorem tells us that for an asymptotically periodic system, cyclicity of the permutation $\alpha(i)$ is necessary and sufficient for the existence of a unique state of thermodynamic equilibrium characterized by the stationary density f_*.

B. ASYMPTOTIC PERIODICITY ILLUSTRATED.

Asymptotic periodicity may be either inherent to a dynamical system, or induced by stochastic perturbations of a system (Chapter 10). For dynamics described by maps on the unit interval, the following theorem (Lasota and Mackey, 1994) is sometimes useful in establishing the existence of inherent asymptotic periodicity.

Theorem 6.4. *Let $S : [0,1] \to [0,1]$ be a nonsingular transformation satisfying the following three conditions:*

(1) *There exists a partition $0 = b_0 < b_1 < \cdots < b_m = 1$ of $[0,1]$ such that for each integer $i = 1, \cdots, m$ the restriction of $S(x)$ to $[b_{i-1}, b_i]$ is a C^2 function.*
(2) $|S'(x)| \geq \vartheta > 1$, $x \neq b_i$.
(3) *There exists a real constant c such that* $\frac{|S''(x)|}{|S'(x)|^2} \leq c < \infty$, $x \neq b_i$, $i = 0, 1, \cdots, m$.

Further, let P be the Frobenius-Perron operator corresponding to S. Then for all densities f, the sequence $\{P^t f\}$ is asymptotically periodic.

Example 6.1. To examine the properties of an asymptotically periodic system, choose a generalization of the tent map (3.5),

$$S(x) = \begin{cases} ax & 0 \le x < \frac{1}{2} \\ a(1-x) & \frac{1}{2} \le x \le 1, \end{cases} \tag{6.9}$$

where $1 < a < 2$ (see Provatas and Mackey, 1991a).

To investigate how the map (6.9) transforms densities, we must first derive an expression for the operator P that corresponds to this transformation. Proceeding as in Example 3.6 where the tent map with $a = 2$ was introduced, it is a simple calculation to show that the Frobenius-Perron operator corresponding to (6.9) is given by

$$Pf(x) = \tfrac{1}{a}\left[f\left(\tfrac{1}{a}x\right) + f\left(1 - \tfrac{1}{a}x\right)\right]. \tag{6.10}$$

For $1 < a \le 2$, and for the partition $b_0 = 0 < b_1 = \frac{1}{2} < b_2 = 1$, the generalized hat map (6.9) satisfies the conditions of Theorem 6.4. Thus, the hat map is asymptotically periodic and the evolution of densities *via* the operator (6.10) can be expressed through the spectral decomposition (6.1).

Ito *et al.* (1979) have shown that the hat map is ergodic, thus possessing a unique invariant density f_* of the form (6.4). Its form has been derived in the parameter window $a_{n+1} = 2^{1/2^{1/(n+1)}} < a \le 2^{1/2^{1/n}} = a_n$, $n = 1, 2, \cdots$ by Yoshida *et al.* (1983). Provatas and Mackey (1991a) have proved the asymptotic periodicity of (6.9) with period $T = (n+1)$ for $2^{1/2^{1/(n+1)}} < a \le 2^{1/2^{1/n}}$. Thus, for example, $\{P^t f\}$ has period 1 for $2^{1/2} < a \le 2$, period 2 for $2^{1/4} < a \le 2^{1/2}$, period 4 for $2^{1/8} < a \le 2^{1/4}$, *etc.*

To analytically illustrate the eventual dependence of the sequence $\{P^t f\}$ on the initial density f for asymptotically periodic systems, pick $a = \sqrt{2}$ which is the upper boundary of the range of a values for which (6.9) is asymptotically periodic with period 2. For this value of a, the unique stationary density (6.4) satisfying $Pf_* = f_*$, where P is given by (6.10), takes the explicit form

$$f_*(x) = u1_{J_1}(x) + v1_{J_2}(x) \tag{6.11}$$

where $u = \frac{1}{2}[3 + 2\sqrt{2}]$, $v = \frac{1}{2}[4 + 3\sqrt{2}]$, and the sets J_1 and J_2 are defined by

$$J_1 = [\sqrt{2} - 1, 2 - \sqrt{2}] \qquad \text{and} \qquad J_2 = [2 - \sqrt{2}, \frac{1}{2}\sqrt{2}] \tag{6.12}$$

respectively (*cf.* Provatas and Mackey (1991a)). S maps the set J_1 into J_2 and *vice versa.*

It can be shown analytically that picking $f_*(x)$ given by (6.11) as an initial density simply results in a sequence of densities all equal to the starting density. This is quite different from what happens with an initially uniform density

$$f(x) = (2 + \sqrt{2})1_{J_1 \cup J_2}(x). \tag{6.13}$$

In this case, the first iterate $f_1 = Pf$ is given by

$$f_1(x) = (1 + \sqrt{2})1_{J_1}(x) + 2(1 + \sqrt{2})1_{J_2}(x) \tag{6.14}$$

and iteration of $f_1(x)$ leads, in turn to an $f_2(x) = f(x)$ and thus the cycling of densities repeats indefinitely with period 2 (*cf.* Figure 6.1a).

This effect of the choice of the initial density on the sequence of subsequent densities can be further illustrated by choosing an initial density

$$f(x) = [3 + 2\sqrt{2}]1_{J_1}(x) \tag{6.15}$$

totally supported on the set J_1. In this case,

$$f_1(x) = Pf(x) = [4 + 3\sqrt{2}]1_{J_2}(x), \tag{6.16}$$

and $f_2 = f$, $f_3 = f_1$, $\cdots$ so once again the densities cycle between f and f_1 with period 2 (*cf.* Figure 6.1b). Figure 6.1c illustrates the behaviour of $\{P^t f\}$ for an initially nonuniform density.•

Example 6.2. Sharkovski (1965) has shown that maps like (3.7),

$$S(x) = rx(1-x) \tag{6.17}$$

with a single quadratic maximum display period doubling in the number of fixed points as the parameter r is increased. For example, with $0 \le r < 1$ the single fixed point of (5.17) is $x^* = 0$, while for $1 < r \le 3$, equation (6.17) has one stable fixed point given by $x^* = 1 - \frac{1}{r}$. For r between $3 < r \le r_c \simeq 3.57\cdots$ there is a cascade of parameters which sequentially give rise to 2 unstable fixed points, then 4, 8 *etc.* The periodicity in each of these intervals is equal to the number of fixed points. At r_c, also known as the accumulation point, there are an infinite number of unstable fixed points.

On the other side of the critical parameter, $r_c < r \le 4$ the quadratic map (and maps like it with a single quadratic maximum) has a spectrum of parameter values, labeled by r_n, $n = 1, 2, \cdots$ where so-called "banded chaos" has been reported by Crutchfield *et al.* (1980), Lorenz (1980), and Grossman and Thomae (1977, 1981) based on numerical work. At these values the unit interval $X = [0,1]$ partitions into 2^n subintervals, labeled J_l, $l = 1, 2, \cdots, 2^n$. These are such that $S^{2^n} : J_l \to J_l$ maps J_l *onto* J_l. As well each J_l is mapped cyclically through the whole sequence of $\{J_l\}$ after 2^n applications of S. The iterates of a time series are attracted to these J_l subsets, returning to any J_l every 2^n iterations. These iterates form an aperiodic sequence with a positive Liapunov exponent [Devaney, 1986]. The procedure whereby which one obtains the parameter values r_n at which 2^n banded chaos occurs is given by Grossman and Thomae (1981).

The Frobenius-Perron operator corresponding to the quadratic map (6.17) is

$$Pf(x) = \frac{1}{\sqrt{1 - \dfrac{4x}{r}}} \left[f\left(\frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{4x}{r}} \right) + f\left(\frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{4x}{r}} \right) \right]. \tag{6.18}$$

At values of $r = r_n$, the iterates of any initial density f supported on $[0,1]$, acted on by (6.18), will eventually decompose so they are supported on disjoint sets J_l. Subsequent to the contraction of density supports onto the sequence of sets $\{J_l\}$, the evolution of the sequence $\{P^t f\}$ becomes periodic in time. At these values, the observed periodic evolution of ensemble densities, is, in fact, asymptotically periodic (Provatas and Mackey, 1991a).

The parameter values $r = r_n$ define a reverse sequence to the period doubling sequence for $r \le r_c$. For the latter sequence, we talk of a period doubling in the number of unstable fixed points. When $r = r_n$ however, fixed points are replaced by "chaotic bands" and going from r_n to r_{n+1} involves a doubling in the number of bands.

As with the hat map of Example 6.1, the scaling coefficients $\lambda_1(f)$, $\lambda_2(f)$ can be analytically determined for period two asymptotic periodicity for the quadratic map when $r = r_1$, and the attracting phase space consists of the subsets J_1 and J_2. These are disjoint and connected at the fixed point of (6.17), and $S : J_1 \to J_2$, $S : J_2 \to J_1$. The coefficients $\lambda_1(f)$, $\lambda_2(f)$ may be obtained for any arbitrary initial density f supported on the phase space $X = [0,1]$.

Figure 6.2 illustrates the period 2 asymptotic evolution of $\{P^t f\}$ after 20 transients, for $r = r_1$. In Figure 6.2a the initial density is uniform on the region of $J_1 \cup J_2$ given by $[0.7, 0.8]$. Figure 6.2b shows an asymptotic cycle of $\{P^t f\}$ with $f(x) = 200(x - 0.9)$ supported on $[0.9, 1]$. Figure 6.3a illustrates a period 4 cycle in $\{P^t f\}$ when $r = r_2$, with the initial density f uniform on $[0.5, 0.85]$. Figure 6.3b shows one period 4 cycle of $P^t f$ with $f(x) = 200(x - 0.91)$ supported on $[0.9, 1]$. All of the illustrated sequences are dependent on the initial density. •

C. THE WEAK FORM OF THE SECOND LAW.

The fact that asymptotically periodic Markov operators have a stationary density given by (6.4) does not guarantee the uniqueness of this stationary density. Regardless of whether or not asymptotically periodic systems have unique stationary densities, they have the important property that their conditional entropy is an increasing function that approaches a maximum. This result is formulated more precisely in

Theorem 6.5. *Let P be an asymptotically periodic Markov operator with stationary density f_*. Then as $t \to \infty$ the conditional entropy $H_c(P^t f | f_*)$ of $P^t f$ with respect to f_* approaches a limiting value $H_{max}(f|f_*) \le 0$, where*

$$H_{max}(f|f_*) = -\sum_i \int_X \lambda_i(f) g_i(x) \log \left\{ \frac{1}{f_*(x)} \sum_i \lambda_i(f) g_i(x) \right\} dx. \tag{6.19}$$

Proof. Since P is asymptotically periodic, the representation of the Spectral Decomposition Theorem 6.1 is valid, and more precisely equation (6.2) for $P^t f$. Write equation (6.2) in the form

$$P^{t+1} f(x) = \Sigma_t(f, x) + Q_t f(x),$$

where $\Sigma_t(f,x)$ denotes the summation portion of (6.2). Remember that since P is asymptotically periodic, for large times t, $\| Q_t f \| \simeq 0$ and thus $P^{t+1}f(x) \simeq \Sigma_t(f,x)$, so the long time conditional entropy is given by

$$\begin{aligned} H_c(P^{t+1}f|f_*) &\simeq -\int_X \Sigma_t(f,x) \log\left\{\frac{\Sigma_t(f,x)}{f_*(x)}\right\} dx \\ &= H_c(\Sigma_t(f)|f_*). \end{aligned}$$

However, $\Sigma_t(f,x)$ is periodic with finite period T. Since by Theorem 3.1 we also know that $H_c(P^t f|f_*) \geq H_c(f|f_*)$ [the conditional entropy can never decrease], it follows that the approach of $H_c(P^t f|f_*)$ to $H_c(\Sigma_t(f)|f_*)$ must be uniform. Even though $\Sigma_t(f,x)$ is periodic with a finite period T, $H_c(\Sigma_t(f)|f_*)$ is a constant independent of t. In fact we have

$$\begin{aligned} H_c(\Sigma_t(f)|f_*) &= -\int_X \sum_i \lambda_i(f) g_i(x) \log\left\{\frac{1}{f_*(x)}\sum_i \lambda_i(f) g_i(x)\right\} dx \\ &\equiv H_{max}(f|f_*) \leq 0 \end{aligned}$$

for large t. The nonpositivity of $H_{max}(f|f_*)$ is a consequence of the integrated Gibbs inequality (1.5). □

Note that if the stationary density f_* of P is given by (6.4), then the expression for $H_{max}(f|f_*)$ becomes even simpler. Namely, with

$$f_*(x) = \frac{1}{r}\sum_{i=1}^{r} g_i(x),$$

$H_{max}(f|f_*)$ as given by (6.19) becomes

$$H_{max}(f|f_*) = -\log r - \sum_{i=1}^{r} \lambda_i(f) \log \lambda_i(f) \tag{6.20}$$

when we use the orthogonality of the densities $g_i(x)$. Since $0 \leq \lambda_i(f) \leq 1$ for all i, we may also place a lower bound on $H_{max}(f|f_*)$:

$$-\log r \leq H_{max}(f|f_*) \leq 0.$$

This weak form of the Second Law of thermodynamics is the strongest result that we have yet encountered. It demonstrates that as long as the density evolves under the action of a Markov operator that is smoothing, the conditional entropy of that density converges to a maximum. There are two important facets of this evolution that should be recognized:

(1) The convergence of the entropy is due to the fact that $\| Q^t f \| \to 0$ as $t \to \infty$ in the representation (6.2) of Theorem 6.1.
(2) The maximum value of the entropy, $H_{max}(f|f_*)$, as made explicit by the notation, is generally dependent on the choice of the initial density f and, thus, the method of preparation of the system. This indicates that systems with asymptotically periodic dynamics may have a discrete or continuous spectrum of metastable states of thermodynamic equilibrium, each with an associated maximal entropy.

Example 6.3. To illustrate the evolution of the conditional entropy of an asymptotically periodic system we return to our example of the tent map (6.9) with $a = \sqrt{2}$. For this value of a, the stationary density f_* is given by equation (6.11). If we pick an initial density given by f_*, then the conditional entropy $H_c(P^t f_*|f_*) = 0$, its maximal value, for all t. However, if we pick an initially uniform density (6.13), $f(x) = (2+\sqrt{2})1_{J_1 \cup J_2}$, then it is straightforward to show that

$$H_c(f|f_*) = H_c(f_1|f_*) \simeq -0.01479,$$

where $f_1 = Pf$ is given by equation (6.14). Thus by choosing an initial density given by (6.13) or (6.14), the limiting conditional entropy approaches a value less than its maximal value of zero.

This effect of the choice of the initial density affecting the limiting value of the conditional entropy can be further illustrated by choosing an initial density

$$f(x) = [3 + 2\sqrt{2}]1_{J_1}(x)$$

totally supported on the set J_1. In this case, as we have shown,

$$f_1(x) = Pf(x) = [4 + 3\sqrt{2}]1_{J_2}(x),$$

and $f_2 = f$, $f_3 = f_1$, *etc.* so once again the densities cycle between f and f_1 with period 2. The limiting value of the conditional entropy is given by

$$H_c(f|f_*) = H_c(f_1|f_*) = -\log(2) \simeq -0.69316.$$

Thus, with three different choices of an initial density f we have shown that the conditional entropy of the asymptotically periodic system (6.9) may have at least three different limiting asymptotic values. •

Example 6.4. The continuous functional dependence of $H_{max}(f|f_*)$ on the initial density f can be illustrated analytically for the maps (6.9) and (6.17) when they generate period 2 asymptotic periodicity. In particular consider a class of initial densities given by (*cf.* Provatas and Mackey, 1991a)

$$f(x) = \begin{cases} \frac{1}{\xi} & x \in [\gamma_1, \gamma_1 + \xi] \\ 0 & \text{otherwise,} \end{cases} \tag{6.21}$$

where γ_1 is the solution of $\gamma_1 = S^2(\gamma_1)$ and is given by

$$\gamma_1 = \frac{1}{a+1}$$

for the hat map and by

$$\gamma_1 = \frac{r_1^2}{4}\left(1 - \frac{r_1}{4}\right)$$

for the quadratic map.

A plot of $H_{max}(f|f_*)$ for the hat map is shown in Figure 6.4. A remarkable feature of Figure 6.4 is the existence of a sequence of ξ values at which the limiting conditional entropy values are equal. For these values of ξ the asymptotic decomposition of $P^t f$ is identical and the limiting conditional entropy is $H_{max}(f|f_*) \simeq -0.01479$ as we calculated in the previous example with an initial density given by (6.13). Note also the local minima in the limiting conditional entropy as the spreading parameter ξ increases.

A similar comparison of the limiting conditional entropy can be made for the asymptotic periodicity of the quadratic map at $r = r_1$. The same set of initial densities defined by (6.21) is considered. Figure 6.5 is the plot analogous to Figure 6.4 for the hat map. Note that for the quadratic map the maxima in the limiting conditional entropy do not define isoentropic points, although $H_{max}(f|f_*) \simeq -0.093$ as $\xi \to 1$. Moreover, a zig-zag pattern similar to Figure 6.4 emerges but on a much smaller scale, as shown by the insets. •

We close this section with the statement and proof of a sufficient condition for the weak form of the Second Law of thermodynamics.

Theorem 6.6. *Let P be a Markov operator in a normalized measure space, and assume that there is a stationary density $f_* > 0$ of P. If there is a constant $c > 0$ such that for every bounded initial density f*

$$H_c(P^t f|f_*) \geq -c$$

for sufficiently large t, then P^t is asymptotically periodic and

$$\lim_{t\to\infty} H_c(P^t f|f_*) = H_{max}(f|f_*) \leq 0.$$

This theorem assures us that if we are able to find some time t_1 such that the conditional entropy is bounded below for times $t > t_1$, then the entropy is evolving under the action of an asymptotically periodic Markov operator and, as a consequence of Theorem 6.5, the conditional entropy of $P^t f$ approaches a maximum that is generally dependent on the initial density with which the system was prepared.

Proof. Pick a subset E of the phase space X with nonzero Lebesgue measure $\mu_L(E)$. From the definition of the conditional entropy $H_c(P^t f|f_*)$ and our hypothesis, for all sufficiently large times t we have

$$\begin{aligned} H_c(P^t f|f_*) &\equiv -\int_E P^t f(x) \log\left(\frac{P^t f(x)}{f_*(x)}\right) dx - \int_{X\setminus E} P^t f(x) \log\left(\frac{P^t f(x)}{f_*(x)}\right) dx \\ &\geq -c. \end{aligned}$$

Remembering the definition of the function η from equation (1.2), it follows that

$$\begin{aligned}\int_E P^t f(x) \log\left(\frac{P^t f(x)}{f_*(x)}\right) dx &\leq c - \int_{X\setminus E} P^t f(x) \log\left(\frac{P^t f(x)}{f_*(x)}\right) dx \\ &= c + \int_{X\setminus E} \eta\left(\frac{P^t f(x)}{f_*(x)}\right) \mu_*(dx) \\ &\leq c + \eta_{max} \int_{X\setminus E} \mu_*(dx) \\ &\leq c + \frac{\mu_*(X)}{e}.\end{aligned}$$

Further specify the set E by

$$E = \left\{ x \in X : \left(\frac{P^t f(x)}{f_*(x)}\right) > N \right\},$$

where the constant N is selected to make $\mu_L(E) < \delta$. Then,

$$\log N \int_E P^t f(x)\, dx \leq \int_E P^t f(x) \log\left(\frac{P^t f(x)}{f_*(x)}\right) dx \leq c + \frac{\mu_*(X)}{e},$$

or

$$\int_E P^t f(x)\, dx \leq \frac{c + \frac{\mu_*(X)}{e}}{\log N} \equiv \epsilon.$$

Next, pick a second subset $A \subset X$ of nonzero measure so

$$\int_{X\setminus A} P^t f(x)\, dx = \int_X P^t f(x)\, dx - \int_A P^t f(x)\, dx = 1 - \mu_L(A).$$

Thus,

$$\int_{E\cup(X\setminus A)} P^t f(x)\, dx \leq 1 - \mu_L(A) + \epsilon \equiv k.$$

It is clear that we may always select the set A in such a way that $\epsilon < \mu_L(A) < 1$ and, hence, $0 < k < 1$. Therefore, P is smoothing by definition. The rest of the proof is a direct consequence of the Spectral Decomposition Theorem 6.1 and Theorem 6.5 concerning the convergence of the conditional entropy under the action of an asymptotically periodic Markov operator. □

D. ASYMPTOTIC PERIODICITY AND CORRELATIONS.

In the previous chapter we showed that temporal correlations in mixing systems decay to zero in spite of the fact that entropy is absolutely constant when the system is invertible. Suppose that instead of a mixing transformation we have an asymptotically periodic transformation with a unique stationary density f_* of the corresponding Markov operator P, and, as a consequence, the system is ergodic. In this case the behaviour of the correlation function is quite different.

Since P is asymptotically periodic and Theorem 6.1 also holds for L^1 functions, we can choose $f = \eta$ to obtain

$$P^{\tau+1}\eta(x) = \sum_{i=1}^{r} \lambda_i(\eta) g_{\alpha^\tau(i)}(x) + Q_\tau \eta(x).$$

Further, because of the ergodicity of P we can write the correlation function as

$$R_{\sigma,\eta}(\tau+1) = < P^{\tau+1}\eta, \sigma >$$

or, more explicitly,

$$R_{\sigma,\eta}(\tau+1) = \sum_{i=1}^{r} \lambda_i(\eta) \int_X g_{\alpha^\tau(i)}(x)\sigma(x)\, dx + \int_X \sigma(x) Q_\tau \eta(x)\, dx. \tag{6.22}$$

Due to the asymptotic periodicity of P, the first term in (6.22) is periodic in τ with period $T \leq r!$. Furthermore, because of the convergence properties of the transient operator Q the second term will decay to zero as $\tau \to \infty$. Therefore, for asymptotically periodic dynamics the correlation function naturally separates into sustained periodic and decaying stochastic components.

This decoupling of the time correlation function into two independent components can be understood as follows. Asymptotically periodic systems have r disjoint attracting regions of their phase space X whose union is given by

$$\bigcup_{i=1}^{r} \operatorname{supp}\{g_i\}.$$

Each of the regions $\operatorname{supp}\{g_i\}$ map onto each other cyclically according to $\alpha(i)$. All ensembles of initial conditions will asymptotically map into these regions (*i.e.*, all densities will decompose). Thus a time series will also visit these supports periodically, and we expect a periodic component in the time correlation function. However, iterates of the time series which return into any one of the $\operatorname{supp}\{g_i\}$, are described by a density g_i, and so there must exist a stochastic component of the correlation function (the second term of (6.22)).

Thus, asymptotically periodic systems have temporal correlations which are a combination of both periodic and stochastic elements and which never decay to constant values as $t \to \infty$ in spite of the fact that their conditional entropy does approach a local maximum as $t \to \infty$. This is to be compared with mixing systems whose entropy is forever fixed by the mode of preparation of the system, but which nevertheless show an approach of the correlations in the system to zero at long times. The contrasting nature of these two results indicates that there is no connection to be drawn between the limiting behaviour of entropy in a system and the limiting behaviour of temporal correlations.

E. SUMMARY.

In this chapter we have shown how the property of smoothing for Markov operators is equivalent to asymptotic periodicity of sequences of densities (Theorem 6.1), and that asymptotic periodicity is sufficient to guarantee the existence of at least one state of thermodynamic equilibrium (with density given by equation (6.4)) as well as the increase of the entropy to a maximum as time progresses (Theorem 6.5). Interestingly, the maximum entropy to which asymptotically periodic systems evolve in this circumstance (equation (6.19)) may be less than the absolute maximum value corresponding to equation (6.4), and usually depends on the initial density with which the system is prepared. Thus the entropy of the final thermodynamic state of an asymptotically periodic system depends, in general, on the initial state. Theorem 6.6 gives a sufficient condition for this behaviour in the form of the existence of a finite lower bound on the conditional entropy. Further, the behaviour of the entropy and correlations in asymptotically periodic systems is opposite to that of mixing systems, indicating that there is no connection to be drawn between entropy evolution and the limiting behaviour of correlations.

In the next chapter we introduce a dynamical property even stronger than asymptotic periodicity which is both necessary and sufficient for the evolution of system entropy to its unique maximal value of zero.

Figure 6.1. The evolution of $P^t f$ in the period two window under the action of the hat map, with $a = \sqrt{2}$. In (a) f is uniform over $J_1 \cup J_2$. Since the g_i are uniform over J_i, $i = 0, 1$, $P^t f$ sets into immediate oscillations without transients. In (b) f is uniform over the subspace J_1. Again $P^t f$ sets into immediate oscillations through the states g_1 and g_2. In (c) $f(x) = \frac{4(5+\sqrt{2})}{7} x$, restricted to $J_1 \cup J_2$. Now $P^t f$ evolves through two transient densities before settling into a periodic oscillation.

Figure 6.2. A numerical illustration of one periodic cycle of the asymptotic sequence $\{P^t f\}$ under the action of the quadratic map for the parameter $r = r_1 = 3.678573508$. A transient of 20 densities has been discarded, and the iterates $P^{21}f$, $P^{22}f$, and $P^{23}f$ are shown. Since $P^{21}f = P^{23}f$, the sequence $\{P^t f\}$ asymptotically repeats with period 2. In (a) the initial density f, shown in the inset, is uniform over $[0.7, 0.8]$. In (b), $f(x) = 200(x - 0.9)$ over $[0.9, 1]$.

Figure 6.3. Two period 4 cycles of the asymptotic sequence $\{P^t f\}$ for the quadratic map when $r = r_2 = 3.592572184$. In this figure 40 transients have been discarded and the iterates $P^{41}f$, $P^{42}f$, $P^{43}f$, $P^{44}f$ and $P^{45}f$ are shown. Since $P^{41}f = P^{45}f$, the sequence $\{P^t f\}$ asymptotically repeats with period 4. In (a) the initial density (inset) f is uniform over $[0.5, 0.85]$. In (b) $f(x) = 200(x - 0.9)$ over $[0.9, 1]$.

Figure 6.4. The limiting conditional entropy, $H_{max}(f|f_*)$, versus the spreading parameter ξ for the hat map at $a = \sqrt{2}$. ξ is equal to the width of the support of an initial uniform density f. The local maxima in the figure correspond to equal limiting conditional entropy values.

Figure 6.5. A graph of the limiting conditional entropy $H_{max}(f|f_*)$ versus ξ for the quadratic map at $r = r_1$. The parameter ξ plays the same role as in Figure 6.4. Variations in $H_{max}(f|f_*)$ occur over a smaller ξ scale for the quadratic map. (ii) is a blow-up of the inset box in frame (i). (iii) is a blow-up of the inset box in (ii).