Chapter 3: Tools for Analyzing the Interactions of Procedural Radios[1]

"Before thinking outside the box, one should know what's in the box. The box tends to have a lot of good ideas – that's how they came to be in the box."

In this chapter we consider the problem of analyzing the interactions procedural radios based on the model presented in Chapter 2. In general, we can study the interactions of procedural radios via a reduced model that excludes their goals, namely the tuple $\left\langle N, A, \left\{d_j\right\}, T\right\rangle$. As this chapter shows, many traditional analysis techniques from engineering can be applied to the analysis of procedural radios, including dynamical systems theory, optimization theory, parallel processing (contraction mappings), and Markov chain theory. Before cognitive radio, before SDR, these techniques were being applied to the analysis of wireless algorithms. And as we show in this chapter, they are still useful for the analysis of procedural cognitive radios. Further, when we turn to the analysis of ontological radios in subsequent chapters many of the concepts presented in this chapter resurface.

The remainder of this chapter is organized as follows. Section 3.1 considers dynamical systems and describes how the *evolution function* can be used to determine steady-states, optimality, convergence, and stability. Section 3.2 presents variants on contraction mappings, including the standard interference function and pseudo-contractions, and describes how they can be used to determine steady-states, optimality, convergence, and stability. Section 3.3 presents Markov chain theory which can be used to determine steady-states, optimality, convergence, and stability for non-deterministic procedural radios.

3.1 A Dynamical Systems Approach

Dynamical systems theory is concerned with analyzing the behavior of dynamical systems and designing mechanisms so the systems act in a desirable manner. Typical

[1] This chapter is based on a section in J. Neel. J. Reed, A. MacKenzie, "Cognitive Radio Network Performance Analysis" in **Cognitive Radio Technology**, B. Fette, ed., Elsevier, July 28 2006.

analysis goals of dynamical systems theory are similar to the ones that we set out in Chapter 2: determining the expected behavior, convergence, and stability of the system.

Formally, a dynamical system is a system whose change in state is determined by a function of the current state and time. In other words, a dynamical system is any system of the form given by (3.1) which describes the change in the state of a system as a function of the system state, *a*, and time, *t*. Implicitly the system is assumed to be at state *a*(0) at time *t*=0.

$$\dot{a} = g(a,t) \tag{3.1}$$

When (3.1) is not directly dependent on *t*, i.e., $\dot{a} = g(a)$, the system is said to be *autonomous*. For our purposes, it makes sense to treat synchronous systems as autonomous, but for random and asynchronous systems, it is difficult to eliminate the time dependency.

The first goal of a dynamical systems analyst is to solve (3.1) to yield the *evolution function* that describes the state of the system as a function of time. This typically involves solving an ordinary differential equation – a task that we would preferably not undertake without knowing that a solution exists. Given a dynamical systems model, we can be assured that such a solution exists by the Picard-Lindelöf theorem [Walker_80].

Theorem 3.1: Picard-Lindelöf Theorem
Given an open set $D \subset A \times T$ and g as in (3.1), if g is continuous on D and *locally Lipschitz* with respect to a for every $a \in D$, then there is a unique solution, d^t, to the dynamical system for every $a(0)$ while d^t remains in *D*.

Note that Theorem 3.1 requires that g is not only continuous, but also locally Lipschitz – a term we define in Definition 3.1. Note that any function that is Lipschitz continuous is also continuous.

Definition 3.1: *Lipschitz continuity*
A function, $d^t : A \times T \to A, A \subset \mathbb{R}^n$ [2] is said to be *Lipschitz continuous*[3] at (a,t) if there exists a $K < \infty$ such that $\left\| d^t\left(a^1,t\right) - d^t\left(a^2,t\right)\right\| \le K\left\|a^1 - a^2\right\|$ for all $a^1, a^2 \in A$; d^t would be *locally Lipschitz continuous* if this condition were only satisfied for some open set $D \subset A \times T$. Similarly the function d is *Lipschitz continuous* if it is Lipschitz continuous for all $(a,t) \in A \times T$.

In general, the solution to (3.1) will take the form of the decision update rule, d^t, which we assumed existed as part of our model. So this section primarily serves the purpose of connecting our model to the model traditionally assumed in dynamical systems. However, Theorem 3.1 foreshadows the importance of fixed point theorems to the steady-states of procedural radio networks.

3.1.1 Fixed Points and Solutions to Cognitive Radio Networks

A solution for the evolution function d^t may imply a system that is changing states over time, perhaps bounded within a certain region or wandering over the entire action space. For some systems, continual adaptations may not be an issue and may even be desirable. However, continual adaptations for a cognitive radio network implies that significant bandwidth is being consumed to support the signaling overhead required to support these adaptations.

For a cognitive radio network, we would prefer that the network settle down to a particular steady state and only adapt as the environment changes. Identifying these steady-states also allows a cognitive radio designer to predict network performance. In the context of our state equation, such a steady state is a *fixed point* of d^t.

Definition 3.2: Fixed Point
A point $a^* \in A$ is said to be a *fixed point* of $d^t : A \to A$ if $a^* = d^t\left(a^*\right) \forall t \ge t^*$.

---

[2] d^t is a function that maps from the Cartesian Product of the action space with the set of all update times to the action space, where the action space is a subset of all real *n*-tuples; that is, given an initial action state, it forms the action space over all time.
[3] A function, d^t, is Lipschitz continuous if there exists a finite real K, such that for all action states a^1 and a^2 in the action space, the Euclidean distance between $d^t(a^1)$ and $d^t(a^2)$ is less than K times the distance between a^1 and a^2.

For one dimensional sets, it is convenient to envision a fixed point of a function as a point where the function intersects the line $x = f(x)$. Figure 3.1 illustrates a function, $f(x)$, that has three fixed points.

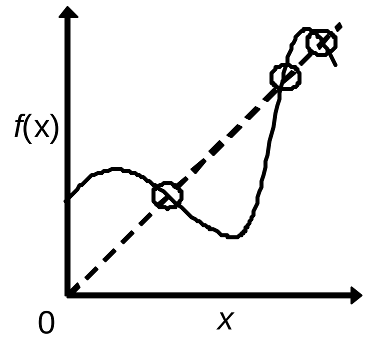

Figure 3.1: A function with three fixed points (circled). For functions on a one dimensional sets, the points at which the function intersect the line $f(x) = x$ (dashed) are fixed points.

Solving for fixed points can be tedious as it may involve a search over the entire action space (an impossibility for an infinite action space, and a considerable undertaking for most realistic finite action spaces), so we would like to know if a fixed point exists before we begin our search. Fortunately, this can be readily established by the Leray-Schauder-Tychonoff fixed point theorem given by Proposition 1.3 in Chapter 3 of [Bertsekas_97].[4]

<u>**Theorem 3.2**</u>: Leray-Schauder-Tychonoff Fixed Point Theorem
If $A \subset \mathbb{R}^n$ is <u>nonempty</u>, <u>convex</u>, and <u>compact</u> (see Appendix B), and if $d : A \rightarrow A$ is a <u>continuous</u> function, then there exists some $a^* \in A$ such that $a^* = d^t(a^*)$.[5]

However, there are several limitations to this theorem. First, the theorem is inappropriate for finite action sets – a likely condition – as while finite sets are compact, they are not convex. Second, d may not be a continuous function. Third, actually solving for a fixed point under such general conditions can be much harder, though under these conditions we simultaneous solution of (3.2) is appropriate for identifying steady-states.

[4] The game theorist may recognize this as a variant of Brouwer's fixed point theorem.
[5] Leray-Schauder-Tychonoff actually considers continuous *mappings* instead of continuous function, but a continuous function is a continuous mapping.

$$a_i^* = d_i\left(a^*\right) \forall i \in N \quad (3.2)$$

3.1.2 Establishing Optimality

Perhaps the easiest way to establish that a solution to a cognitive radio network is optimal is to show that it max(min)imizes some objective function $J: A \to \mathbb{R}$. For networks with a finite action space we can perform an exhaustive search and evaluate *J* at each point in *A*.

However, this approach is impractical for infinite action spaces. But when *J* is differentiable and *A* is a compact interval of $\mathbb{R}^n$, we can reduce the search space by noting that if a particular action vector, a^*, is optimal, then a^* must either be a boundary point or a point where $\nabla J\left(a^*\right) = 0$ where $\nabla J\left(a\right) = \frac{\partial J\left(a\right)}{\partial a_1}\hat{a}_1 + \frac{\partial J\left(a\right)}{\partial a_2}\hat{a}_2 + \cdots + \frac{\partial J\left(a\right)}{\partial a_n}\hat{a}_n$ [6] where each $\hat{a}_j$ is a dimension of *A*. So in effect, this condition says that for a^* to optimize *J*, there must be no direction that can be followed from a^* that increases *J*. If *J* is *pseudo-concave*, we can change this to a sufficient condition, i.e., if there exists some point such that $\nabla J\left(a^*\right) = 0$, then it is optimal.[7] [Zangwill_69]

> **<u>Definition 3.3</u>**: *Pseudo-concavity*
> A function $J: A \to \mathbb{R}$ is said to be *pseudo-concave* if $\nabla J(a'') \cdot (a' - a'') \leq 0 \Rightarrow J\left(a'\right) \leq J\left(a''\right)$ for all points $a', a'' \in A$.

More familiarly, a function that is *concave* is also pseudo-concave.

> **<u>Definition 3.4</u>**: *Concavity*
> A function, $J: A \to \mathbb{R}$, is concave on the set *A* if for all $a_1, a_2 \in A$, $J\left(l\, a_1 + (1-l)\, a_2\right) \geq l\, J\left(a_1\right) + (1-l)\, J\left(a_2\right)$ for all $l \in [0,1]$.

[6] The gradient of the cost function *J*, is in general a vector valued function that when evaluated at a particular point, *a''* indicates the magnitude and direction of greatest increase for *J* at *a''*. When *J* is a function of a single dimension, then the gradient of *J* is equivalent to the slope of *J*.
[7] This is a variant on the Karush-Kuhn-Tucker theorem.

Equivalently, a function is concave if it is impossible to join two points in the function with a line that contains points above the function.

Figure 3.2 shows an example of a function that is pseudo-concave, but not concave. This function can be verified to not be concave by considering a line joining the points (0, 0) and (1,1) (shown as a dashed line); except for the endpoints, all of the points in this line lie above the function.

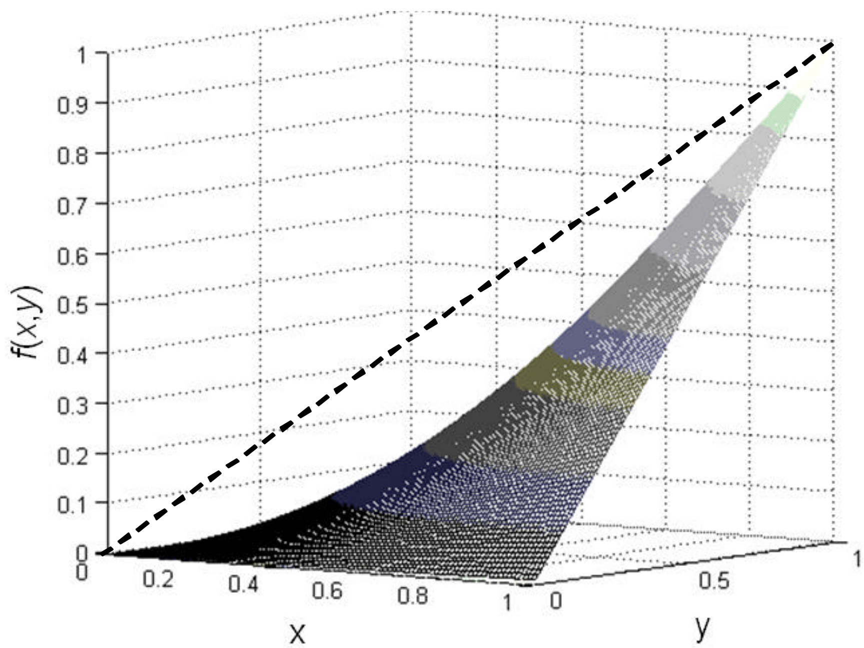

Figure 3.2: $f(x, y) = xy$, $x, y > 0$ – A function that is pseudo-concave, but not concave.

3.1.3 Convergence and Stability

When discussing convergence and stability of a decision rules fixed point, it is convenient to make use of two forms of stability: *Lyapunov stability* and *attractivity.*

Definition 3.5: *Lyapunov stability*
We say that an action vector, a^*, is *Lyapunov stable* if for every $e > 0$ there is a $d > 0$ such that for all $t \geq t^0$, $\left\| a\left(t^0\right), a^* \right\| < d \Rightarrow \left\| a\left(t\right), a^* \right\| < e$ [8].

[8] Equivalently, the action vector a^* is said to be Lyapunov stable if for every arbitrarily sized $\varepsilon>0$, it is possible to identify a $\delta>0$ such that after a pertubation to any point $a(t^0)$, all subsequent action vectors are no more than a Euclidean distance of ε away from a^*.

While no particular relation between d and e can be inferred from this definition, an engineer may be more comfortable thinking of Lyapunov stability as akin to Bounded-Input-Bounded-Output stability wherein after a bounded "stimulus" of d is added to a system operating at a^*, the system remains within a bounded distance e of a^*.

<u>Definition 3.6</u>: *Attractvity*
The action vector a^* is said to be *attractive* over the region $S \subset A$, $S = \left\{ a \in A \mid \left\| a, a^* \right\| < M \right\}$, if given any $a(t_0) \in S$, the sequence $\left\{ a(t) \right\}$ converges to a^* for $t \geq t_0$.

Tying both concepts together is the *asymptotic stability.*

<u>Definition 3.7</u>: *Asymptotic Stability*
The action vector a^* is said to be *asymptotically stable* if it is both Lyapunov stable and attractive.

Note that Lyapunov stability does not imply attractivity nor does attractivity imply Lyapunov stability. For instance, the fixed point (0,0) in Figure 3.3 is Lyapunov stable, but not attractive; meanwhile the fixed point (0,0) in Figure 3.4 is attractive, but not Lyapunov stable. However, the intuition that both stability and attractivity are frequently found together is borne out by Lyapunov's Direct Method.

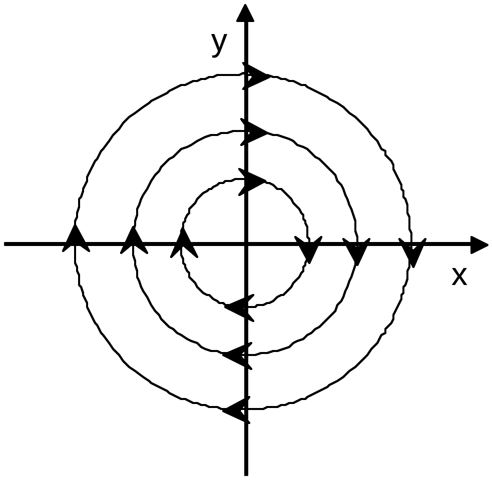

Figure 3.3: Paths (formed by recursive application of d^t with direction indicated by arrows) for a system that is Lyapunov stable but not attractive.

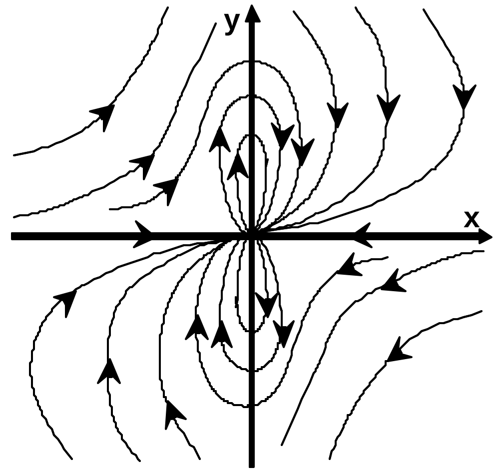

Figure 3.4: Paths for a fixed point that is attractive but not Lyapunov stable.

Instead of attempting to directly apply the definitions of Lyapunov stability and asymptotic stability, we can use Lyapunov's direct method. The discrete time version of Lyapunov's direct method is given in Theorem 3.4 of [Medio_01] as follows.

Theorem 3.3: *Lyapunov's Direct Method for Discrete Time Systems*

Given a recursion $a\left(t^{k+1}\right)=d^t\left(a\left(t^k\right)\right)$ *with fixed point* a^*, *we know that* a^* *is Lyapunov stable if there exists a continuous function (known as a* Lyapunov function*) that maps a neighborhood of* a^* *to the real numbers, i.e.,* $L:N(a^*)\to\mathbb{R}$, *such that the following three conditions are satisfied:*

1) $L\left(a^*\right)=0$
2) $L(a)>0\ \forall a\in N\left(a^*\right)\backslash a^*$
3) $\Delta L\left(a\left(t\right)\right)\equiv L\left[d^t\left(a\left(t\right)\right)\right]-L\left(a\left(t\right)\right)\le 0\ \ \forall a\in N\left(a^*\right)\backslash a^*$

Further, if conditions 1-3 hold and

a) $N\left(a^*\right)=A$, *then* a^* *is globally Lyapunov stable;*
b) $\Delta L\left(a\left(t\right)\right)<0\ \forall a\in N\left(a^*\right)\backslash a^*$, *then* a^* *is asymptotically stable;*
c) $N\left(a^*\right)=A$ *and* $\Delta L\left(a\left(t\right)\right)<0\ \forall a\in N\left(a^*\right)\backslash a^*$, *then* a^* *is globally asymptotically stable.*

Lyapunov's direct method says, in effect, that if we can find a function that strictly decreases along all paths created by the adaptations of a cognitive radio network, then that cognitive radio network is asymptotically stable.

It is also interesting to note that the existence of a Lyapunov function can be used to establish the existence and identify the network's steady-states, namely all points where $L\left(a^*\right)=0$. Further, Lyapunov's direct method can be readily applied to both synchronous and asynchronous cognitive radio networks – the only requirement being each adaptation must decrease the value of the Lyapunov function. Of course, there are cognitive radio algorithms with many steady-states which are so closely spaced that it is impossible to identify any combination of neighborhood and Lyapunov function that meets these definitions. Such a scenario is considered in Chapter 7.

While this section does not present a particular example analysis of an evolution equation for fixed points, convergence, optimality, or Lyapunov stability, these concepts will be repeatedly applied throughout the remainder of this document.

3.2 Contraction Mappings and the General Convergence Theorem

In the preceding discussion, we assumed a closed form expression for the next network state as a function of current network state. Now suppose that after one recursion of the network update rule we are unable to precisely predict the next network state. However, we are able to bound the network state within a particular set of states $A(t^1)$. Then suppose that armed with the knowledge that the network starts in $A(t^1)$, we could say that after the second iteration, the network state would have to be within another set $A(t^2)$, which is a subset of $A(t^1)$. Extending this concept, suppose that given any set of network states, $A(t^k)$, we know that the decision update rule always results in a network state in the set, $A(t^{k+1})$, which is a subset of $A(t^k)$.

In effect, this process is saying that as the recursion continues finer and finer approximations on the operating point of the network are possible, perhaps resulting in a prediction of a specific steady-state for the network. Such a sequence of finer approximations might look as shown in Figure 3.5 where the recursion of subsets, $A(t^k)$, converges to a single point. This iterative restriction on a recursion's possible points forms the basis of numerous valuable algorithms and is a characteristic of special class of algorithms known as *contraction mappings*.

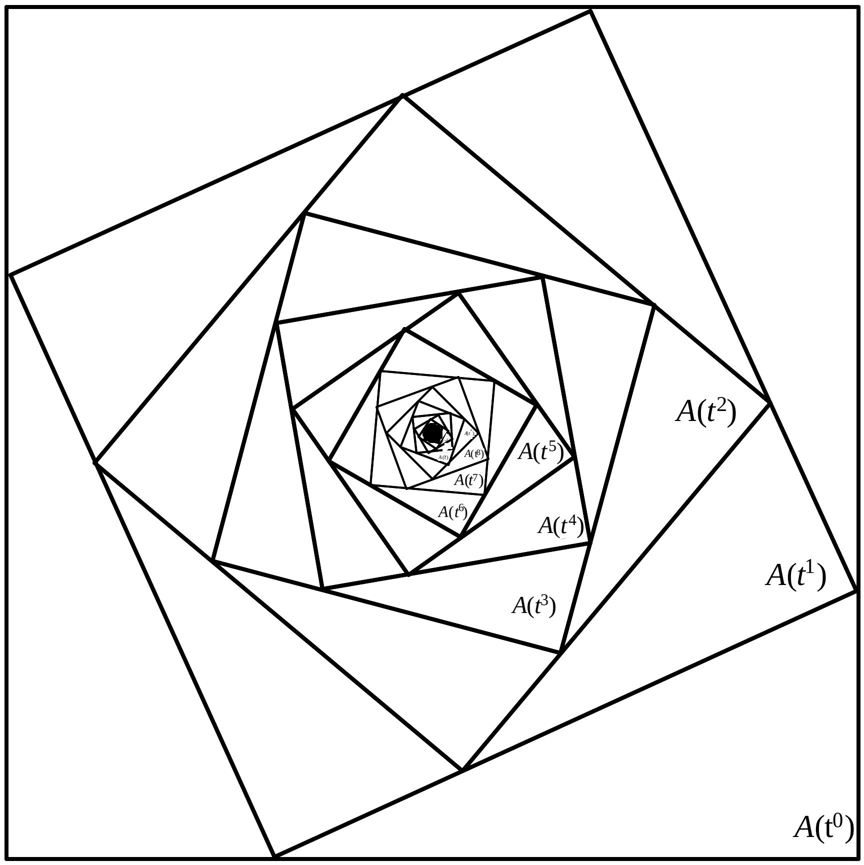

Figure 3.5: A sequence of contracting sets, $\cdots \subset A\left(t^2\right) \subset A\left(t^1\right) \subset A\left(t^0\right)$.

3.2.1 Contraction Mappings

<u>Definition 3.8</u>: *Contraction mapping*

Given a recursion $a\left(t^{k+1}\right) = d\left(a\left(t^k\right)\right)$, *d* is said to be a *contraction mapping* with modulus α if there is an $a \in [0,1)$ such that $\left\| d(a), d(b) \right\| \le a \left\| a, b \right\| \ \forall b, a \in A$.

While applying this definition to a decision rule can be difficult, we can show that an arbitrary recursion, *d*, is a contraction mapping if the following two conditions are satisfied: [Blackwell_65]

Theorem 3.4: *Blackwell's conditions*
Given recursion, $a\left(t^{k+1}\right)=d\left(a\left(t^{k}\right)\right)$, *d* forms a contraction mapping if *d* satisfies monotonicity and discounting.
1) *Monotonicity* – Given bounded functions $g_1, g_2 : A \rightarrow \mathbb{R}$ where $g_1(a) \leq g_2(a)$ $\forall a \in A$, *d* must satisfy $d\left(g_1(a)\right) \leq d\left(g_2(a)\right)\ \forall a \in A$.
2) *Discounting* – There exists a $b \in (0,1)$ such that $d\left(g_1(a)+c\right)=d\left(g_1(a)\right)+b\,c$ for all bounded $g_1 : A \rightarrow \mathbb{R}$, $c \geq 0$, $a \in A$.

3.2.2 Analysis Insights

Knowing that our decision rule constitutes a contraction mapping immediately provides us with several valuable insights. From Banach's contraction mapping theorem[9] given in [Sundaram_99], we know that *d* has a unique fixed point to which the recursion *f* converges from any starting point. After k iterations, a bound on the distance of the current state from the fixed point is given by (3.3).

$$\left\| a\left(t^{k}\right), a^{*} \right\| \leq \frac{\alpha^{k}}{1-\alpha} \left\| a\left(t^{1}\right), a\left(t^{0}\right) \right\| \tag{3.3}$$

(3.3) is also useful for bounding the error in estimating d 's fixed point by recursively evaluating *d*. Additionally, a Lyapunov function for any contraction mapping with fixed point a^* is given by (3.4). Thus every contraction mapping, *d*, has a unique stable fixed point to which *d* converges at a predictable rate.

$$L(a)=\left\| a, a^{*} \right\| \tag{3.4}$$

3.2.3 Pseudo-Contractions

A pseudo-contraction eliminates the contraction mapping's requirement that all points move closer to each other after each iteration but still requires that after each iteration, all points move closer to a unique fixed point.

Definition 3.9: *Pseudo-contraction*
Given mapping $d : A \rightarrow A$ with fixed point, a^*, we say d is a *pseudo-contraction* if there is an $\alpha \in [0,1)$ such that $\left\| d(a), d\left(a^{*}\right) \right\| \leq \alpha \left\| a, a^{*} \right\|\ \forall a \in A$.

[9] Banach's fixed point theorem is simply that every contraction mapping has a unique fixed point.

By definition, d has a unique fixed point, a^*, to which d converges at a rate given by (3.5). Note that evaluation of (3.5) requires knowledge of the fixed point, so unlike (3.3), it is not appropriate for bounding the error on an estimate of the system's fixed point while iterating to solve for the fixed point. Also note that (3.4) serves as a Lyapunov function for a pseudo-contraction and that a^* is globally asymptotically stable.

$$\left\| a\left(t^k\right), a^* \right\| \le \alpha^k \left\| a(0), a^* \right\| \tag{3.5}$$

3.2.4 General Convergence Theorem

For most contraction mappings, it is assumed that the updating process occurs synchronously (recall the discussion of decision timings in Chapter 2). We can relax this assumption by introducing the general convergence theorem presented in Proposition 2.1 of Chapter 6 in [Bertsekas_97].

> **Theorem 3.5**: *General Convergence Theorem*
> *Suppose we know that* $\cdots \subset A\left(t^{k+1}\right) \subset A\left(t^k\right) \subset \cdots \subset A\left(t^0\right)$ *where* $A(t^k)$ *represents the possible states of the network after k iterations and* $A\left(t^0\right)$ *represents all possible initial states for the network. Then if the following two conditions hold, then f also converges asynchronously.*
> 1)*Synchronous Convergence Condition*
> *(a)* $d(a) \in A\left(t^{k+1}\right)\ \forall k, a \in A\left(t^k\right)$
> *(b) If* $\{a(t^k)\}$ *is a sequence such that* $a\left(t^k\right) \in A\left(t^k\right)$ *for every k, then every limit point of* $\{a(t^k)\}$ *is a fixed point of d.*
> *2) Box Condition*
> *For every k, there exist sets* $A_j\left(t^k\right) \subset A_j$ *such that* $A\left(t^k\right) = A_1\left(t^k\right) \times \cdots A_n\left(t^k\right)$.

For our purposes, the general convergence theorem states that under an assumption that each radio's action sets are independent (thereby implying the action space satisfies the box condition), any contraction or pseudo-contraction mapping that converges synchronously also converges asynchronously. However, we can also apply the general convergence theorem to algorithms that are not obviously contraction mappings as seen in the following extended example.

3.2.4.1 Standard Interference Function Model

Many traditional analyses consider specific decision rules that model specific applications. The following discusses such an analysis that is also an example of a non-obvious contraction mapping. [Yates_95] considers a power control algorithm operating on the uplink of a cellular system with uniform frequency reuse.[10] For this algorithm, there is a set of N mobiles where each mobile, j, attempts to achieve a target received SINR, $\hat{g}_j$. The development of this algorithms assumes that each mobile is capable of observing its received SINR (perhaps via feedback from a base station) which is generally given by (3.6) where g_{kj} can be the link budget gain from mobile k to the base station of j, p_k is the transmit power of mobile k, and N_j is the noise power at the base station that is receiving mobile j's signal.

$$g_j = \frac{g_{jj} p_j}{\sum_{k \in N} g_{kj} p_j + N_j} \tag{3.6}$$

Based on observations of (3.6), the mobiles compute a scenario dependent *interference function*, $I_j(\mathbf{p})$, which is formed as the ratio of the target SINR, $\hat{g}_j$, and the effective SINR, g_j, i.e., as shown in (3.7)

$$I_j(\mathbf{p}) = \hat{g}_j / g_j \tag{3.7}$$

where $\mathbf{p}$ is the vector of transmit powers, $\mathbf{p} = (p_1, p_2, \ldots, p_n)$, drawn from the power vector space $\mathbf{P}$.

Generalizing beyond this ratio formalization, Yates defines any interference function to be *standard* if it satisfies the conditions given in Definition 3.10 where we write $\mathbf{p}^1 \geq \mathbf{p}^2$ if $p_j^1 \geq p_j^2 \; \forall j \in N$ and $I(\mathbf{p})$ is the synchronous evaluation of all $I_j(\mathbf{p})$.

[10] It is not stated in [Yates_95] that uniform frequency reuse is an assumption; rather this is the result of a conversation with Yates at DySPAN 05 about [Yates_95].

Definition 3.10: Standard Interference Function
An interference function, $I:\mathbf{P}\rightarrow\mathbb{R}^n$, is said to be standard if it satisfies the following three conditions:

1. *Positivity* - $I(\mathbf{p})>0$
2. *Monotonicity* – If $\mathbf{p}^1\geq\mathbf{p}^2$ then $I(\mathbf{p}^1)\geq I(\mathbf{p}^2)$
3. *Scalability* – For all $a>1$, $a\,I(\mathbf{p})>I(a\mathbf{p})$

Assuming the existence of a standard interference function, Yates defines a synchronous updating process of the form $\mathbf{p}(t^{k+1})=d(\mathbf{p}(t^k))$ where $d(\mathbf{p})=d_1(\mathbf{p})\times\cdots\times d_n(\mathbf{p})$ and f_j is given by (3.8).

$$d_j(\mathbf{p}(t^k))=p_j(t^k)\,I_j(\mathbf{p}(t^k)) \qquad (3.8)$$

[Yates_95] then considers the situation where the target SINR vector, $\hat{\mathrm{g}}=(\hat{g}_1,\ldots,\hat{g}_n)$, is *feasible*.

Definition 3.11: Feasible SINR Vector
A target SINR vector, $\hat{\mathrm{g}}$, is said to be *feasible* if there exists a $\mathbf{p}\in\mathbf{P}$ such that $g_j\geq\hat{g}_j\ \forall j\in N$.

When the, [Yates_95] is able to show that an algorithm updating the power vector according to (3.8) has the following properties:

1. A fixed point exists, i.e., there is some $\mathbf{p}^*$ such that $\mathbf{p}^*=d(\mathbf{p}^*)$
2. This fixed point is unique
3. Starting from any initial power vector, d converges to $\mathbf{p}^*$.

While [Yates_95] shows these results in an ad-hoc manner, [Berggren_01] shows that this updating process constitutes a pseudo-contraction which could be used to establish these same results by applying the results of Section 3.2.3. Further we would also know that d is stable. The fact that d constitutes a pseudo-contraction implies that $\cdots\subset\mathbf{P}(t^{k+1})\subset\mathbf{P}(t^k)\subset\cdots\subset\mathbf{P}(t^0)$ where $\mathbf{P}(t^k)$ is the power vector space for iteration k. Coupled with the just established synchronous convergence of f and implicit satisfaction of the box condition, this means that d has satisfied the conditions for the general convergence theorem. Thus it is known that d converges both synchronously and asynchronously. These results are proven in a more rigorous fashion using different techniques in Chapter 9.

3.2.4.2 Further Insights from the Standard Interference Function

Assuming the SINR feasibility criterion is satisfied, [Yates_95] also shows that the following target SINR arrangements of base stations and mobiles have standard interference functions and thus converge synchronously and asynchronously to a unique power vector when the decision update rule is given as in (3.8).

- Fixed assignment – each mobile is assigned to a particular base station;
- Minimum power assignment – each mobile is assigned to the base station in the network where the mobile's SINR is maximized
- Macro diversity – all base stations in the network combine the signals of the mobiles;
- Limited diversity – a subset of the base stations combine the signals of the mobiles; and
- Multiple connection reception – the target SINR must be maintained at a number of base stations.

Feasible SINR

Previously we defined a target SINR vector, $\hat{g}$, as being feasible if there exists a $\mathbf{p} \in \mathbf{P}$ such that $g_j \geq \hat{g}_j \; \forall j \in N$. Rather than performing an exhaustive search over **P**, [Zander_01] presents the following approach for determining if $\hat{g}$ is feasible.

Consider a network with link gain matrix **G** formed as $\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{bmatrix}$ where g_{jk} is the link gain as used in (3.8). Now form the normalized link matrix H as $h_{ij} = g\dfrac{g_{ji}}{g_{ii}}, i \neq j$ with $h_{ii} = 0$. Then p. 155 of [Zander_01] tells that the uniform target

SINR vector $\hat{g}_u = (\hat{g}, \hat{g}, \ldots, \hat{g})$ is achievable if the spectral radius (largest eigenvalue)[11] of **H** is less than or equal to one. When the spectral radius is exactly 1, then $\hat{g}$ is achievable only when there is no noise in the system. Interestingly, [Berggren_01] states that for the pseudo-contraction formed by the standard interference function, $a = r(\mathbf{H})$ which allows us to estimate the convergence rate as well.

A similar expression can be found for the nonuniform target SINR scenario where $\hat{g} = (\hat{g}_1, \hat{g}_2, \ldots, \hat{g}_n)$ as follows where the link matrix, ***H'***, is formed as $h'_{ij} = \frac{\hat{g}_i}{\hat{g}_{\max}} \frac{g_{ji}}{g_{ii}}, i \neq j$ where $\hat{g}_{\max} = \max_{i \in N} \{\hat{g}_i\}$.

Assuming the target SINRs are feasible, then the power vector corresponding to the unique fixed point specific can be found by solving (3.9)

$$\mathbf{Zp} = \vec{g} \tag{3.9}$$

where $\mathbf{Z} = \begin{bmatrix} g_{11} & -\hat{g}_1 g_{21} & \cdots & -\hat{g}_1 g_{n1} \\ -\hat{g}_2 g_{1n} & g_{22} & \cdots & -\hat{g}_2 g_{n2} \\ \vdots & & \ddots & \vdots \\ -\hat{g}_n g_{1n} & -\hat{g}_n g_{2n} & \cdots & g_{nn} \end{bmatrix}$ and $\vec{g} = [\hat{g}_1 N_1 \quad \hat{g}_2 N_2 \quad \cdots \quad \hat{g}_n N_n]^T$.

3.3 Markov Models

Perhaps because of uncertainty in the order of adaptation (as would be the case for a randomly or asynchronously timed process) or because of uncertainties in the decision rules (either from noise or a non-deterministic procedural radio), it may be impossible to derive a closed-form expression for an evolution equation or to even to bound the adaptations into sequential subsets. Instead, suppose we can model the changes of the cognitive radio network from one state to another as a sequence of probabilistic events conditioned on past states that the system may have passed through. When the probability

[11] Due to the work of Hilbert, spectral theory refers to a set of theories relating to matrices, eigenvalues and eigenvectors. In spectral theory, the set of eigenvalues for a matrix is said to be its spectrum.

distribution for the next state in time, $a(t^{k+1})$, is conditioned solely on the most recent state as shown in (3.10), the random sequence of states, $\{a(t)\}$ is said to be a *Markov chain*. A model of a system whose states form a Markov chain is said to be a *Markov model*. Throughout the remainder of this section we use these two terms interchangeably.

$$P\left(a\left(t^{k+1}\right)=a^{k} \mid a(0),\ldots,a(t)\right)=P\left(a\left(t^{k+1}\right)=a^{k} \mid a\left(t^{k}\right)\right) \qquad (3.10)$$

Formalizing our model, let us assume that our state space is finite. This is not a requirement for a Markov chain, but the assumption is useful for the subsequent discussion. Further, let us assume that if the network is in state $a^m \in A$ at time t^k, then at time t^{k+1}, the network transitions to state $a^n \in A$ with probability p_{mn} where $p_{mn} \geq 0 \ \forall a^m, a^n \in A$ and $\sum_{j \leq |A|} p_{mj} = 1$. Of course, it is also permitted that the system remains in state a^m which it does with probability p_{mm}. To simplify notation, we make use of a *transition matrix*, which we represent with symbol $\mathbf{P}$. The transition matrix is formed by assigning p_{mn} to the entry corresponding to the m^{th} row and n^{th} column.

3.3.1 Markov Model Analysis Insights

From $\mathbf{P}$ we can then form $\mathbf{P}^2$ as the matrix product $\mathbf{PP}$. Now entry p_{mn}^2 in the m^{th} row and n^{th} column of $\mathbf{P}^2$ represents the probability that system is in state a^n two iterations after being in state a^m. Similarly, if we consider the matrix $\mathbf{P}^k$ formed as $\mathbf{P}^k = \mathbf{P}\mathbf{P}^{k-1}$ (an example of a Chapman-Kolmogorov equation for a Markov chain [Stewart_94]), then entry p_{mn}^k in the m^{th} row and n^{th} column of $\mathbf{P}^k$ represents the probability that system is in state a^n k iterations after being in state a^m.

A similar relationship can be found when the initial state is specified by a random probability distribution arranged as a column vector p where $p_m \in [0,1]$ and $\sum_{m=1}^{|A|} p_m = 1$ where p_m represents the probability of starting in state a^m. For such a situation, the state probability distribution after k iterations is given by $\mathrm{p}^{T}\mathbf{P}^{k}$ where the superscripted T denotes the transpose operation.

3.3.2 Ergodic Markov Chains

Tying back into our analysis objectives of steady-states and convergence, we are particularly interested in determining the *stationary distributions* and *limiting distribution* of a Markov chain that models a cognitive radio network.

> **Definition 3.12**: *Stationary Distribution*
> A probability distribution such that p^{*} such that $\mathrm{p}^{*T}\mathbf{P} = \mathrm{p}^{*T}$ is said to be a stationary distribution for the Markov chain defined by **P**.

Note that solving for a stationary distribution is equivalent to solving the eigenvector equation given in (3.11) where $\lambda=1$.

$$\mathrm{p}^{*T}\mathbf{P} = \lambda \mathrm{p}^{*T} \tag{3.11}$$

> **Definition 3.13**: *Limiting Distribution*
> Given initial distribution p^{0} and transition matrix **P**, the *limiting distribution* is the distribution that results from evaluating $\lim_{k\to\infty} \mathrm{p}^{0\,T}\mathbf{P}^{k}$.

While not generally a steady state as we considered in the previous discussion, showing that a Markov chain has a unique distribution that is both stationary and limiting would permit us to characterize the behavior of the network. Specifically, given the unique stationary limiting distribution p^{*}, we could predict that at a particular instance in time and after a sufficient number of iterations, the network would be in state a^{m} with probability π_{m}. Thus it is desirable to be able to identify when such a unique stationary limiting distribution exists as is done in the ergodicity theorem given in [Syski_92].

> **Theorem 3.6**: *Ergodicity Theorem*[12]
> If a Markov chain is *ergodic*, then there exists a unique limiting and stationary distribution for all initial distributions p^{0}.

[12] This is also called the "Fundamental Theorem of Markov Chains."

This theorem is in reality just a restatement of the definition of an ergodic Markov chain. However the theorem emphasizes a valuable insight – an ergodic Markov chain converges to the same limiting distribution regardless of the initial distribution. Thus when we can show that a cognitive radio network can be modeled as an ergodic Markov chain we gain the following insights:

- The network has a unique "steady-state" distribution p^*
- This distribution can be found by solving the eigenvalue problem $p^{*T}\mathbf{P} = \lambda p^{*T}$ where $\lambda=1$.
- From all initial distributions, the network converges to p^*.

[Stewart_94] states that a Markov chain is ergodic if it is a Markov chain is ergodic if it is a) *irreducible*, b) *positive recurrent*, and c) *aperiodic*.

<u>Definition 3.14</u>: *Irreducibility*
A Markov chain is *irreducible* if $\forall a^m, a^n \in A$, there exist sequences of state transitions with nonzero probability that lead to every state.

<u>Definition 3.15</u>: *Positive Recurrence*
A Markov chain is *positive recurrent* if $\forall a^m \in A$, the expected number of iterations to return to state a^m is less than ∞.

<u>Definition 3.16</u>: *Aperiodicity*
A Markov chain is *aperiodic* if $\forall a^m \in A$, there there is no integer, $n > 1$, such that once the system leaves the state, it can only return to the state in multiples of n iterations.

Note that a network with round-robin timing will not satisfy aperiodicity as an adaptation away from state a^m on radio *i's* turn can only return to a^m on one of radio *i*'s turns which by definition only occurs every n iterations. However, if we treat an entire round as an iteration, then the aperiodicity can be satisfied by a network with round-robin timing. In general, we prefer not to have to apply definitions to the identification process as it tends to be quite time-consuming. Fortunately theorem 4.1.2 in [Kemeny_60] provides a readily applied identification criterion.

Theorem 3.7: *Ergodocity Criteria*
A finite Markov chains with transition matrix **P** is ergodic if and only if there is some k such that $\mathbf{P}^k$ has no zero entries.

Thus by identifying this simple condition, we know that a unique identifiable stationary limiting distribution exists.

Example 3.1: Markov Model of Cognitive Radio Adaptations

Consider a network consisting of two cognitive radios where each radio can choose between two actions. This network would have four possible states which we could label $\{a^1, a^2, a^3, a^4\}$. Suppose that from experimental observation, we observe the probability transition matrix shown in (3.12) and illustrated in Figure 3.6 where each state is represented as a vertex (circle) and each transition is represented as a weighted and directed edge labeled with its associated transition probability.

$$\mathbf{P} = \begin{array}{c} \\ a^1 \\ a^2 \\ a^3 \\ a^4 \end{array} \begin{array}{c} \begin{array}{cccc} a^1 & a^2 & a^3 & a^4 \end{array} \\ \begin{bmatrix} 0.1 & 0.3 & 0.1 & 0.5 \\ 0.4 & 0.0 & 0.3 & 0.3 \\ 0.4 & 0.1 & 0.3 & 0.2 \\ 0.1 & 0.4 & 0.3 & 0.2 \end{bmatrix} \end{array} \tag{3.12}$$

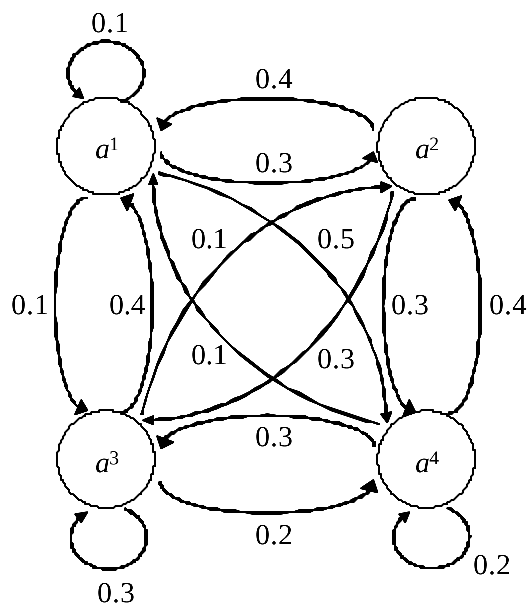

Figure 3.6 Digraph Representation of (3.12)

As specified by (3.12), **P** gives the probability of transitioning from state a^2 to state a^3 as 0.3. After calculating $\mathbf{P}^2$ as shown below, we can immediately determine the probability of the system operating in state a^4 after two iterations after starting in a^3 ($p^2_{34} = 0.33$).

$$\mathbf{P}^2 = \begin{array}{c} \\ a^1 \\ a^2 \\ a^3 \\ a^4 \end{array} \begin{array}{c} \begin{array}{cccc} a^1 & a^2 & a^3 & a^4 \end{array} \\ \begin{bmatrix} 0.22 & 0.24 & 0.28 & 0.26 \\ 0.19 & 0.27 & 0.22 & 0.32 \\ 0.22 & 0.23 & 0.22 & 0.33 \\ 0.31 & 0.14 & 0.28 & 0.27 \end{bmatrix} \end{array}$$

Similarly, given an initial distribution of states $\text{p} = \begin{bmatrix} 0.1 & 0.2 & 0.3 & 0.4 \end{bmatrix}^T$, after two iterations, the probability of being in each state is given by $\text{p}\,\mathbf{P}^2 = \begin{bmatrix} 0.25 & 0.203 & 0.25 & 0.297 \end{bmatrix}^T$. Because all elements in $\mathbf{P}^2$ are positive, there exists a stationary distribution p^* which we can find by solving the eigenvector equation $\text{p}^{*T}\,\mathbf{P} = \text{p}^{*T}$ to yield $\text{p}^{*T} = \begin{bmatrix} 0.2382 & 0.2352 & 0.2272 & 0.2938 \end{bmatrix}$.

3.3.3 Absorbing Markov Chains

For cognitive radio networks that we can model as ergodic chains we can readily find the unique limiting distribution. However, this "steady-state" is somewhat unsatisfying as the network will not remain at a single state and all states will have nonzero probability of being occupied and thus the "steady-state" of an ergodic Markov chain does not conform to our expectations from Chapter 2. However, this is not a problem for absorbing Markov chains.

A state a^k in a Markov chain is said to be an *absorbing state* if there are no paths that leave a^k. This is defined more formally in Definition 3.17.

Definition 3.17: *Absorbing State*
Given a Markov chain with transition matrix **P**, a state a^k is said to be an absorbing state if $p_{kk} = 1$.

Definition 3.18: *Absorbing Markov chain*
A Markov chain is said to be an *absorbing Markov chain* if

a) it has at least one absorbing state and
b) from every state in the Markov chain there exists a sequence of state transitions with nonzero probability that leads to an absorbing state. These nonabsorbing states in are called *transient states*.

For example, (3.13) gives a transition matrix for an absorbing Markov chain where a^4 (note that p_{44}=1) is the absorbing state and a^1, a^2, and a^3 are the transient states where all transient states have a nonzero probability of transitioning directly to a^4. As we will see in later examples, the existence of a direct transition to an absorbing state is not a requirement of an absorbing Markov chain nor must a transition matrix have only a single absorbing state.

$$\mathbf{P} = \begin{array}{c} \\ a^1 \\ a^2 \\ a^3 \\ a^4 \end{array} \begin{array}{c} \begin{array}{cccc} a^1 & a^2 & a^3 & a^4 \end{array} \\ \begin{bmatrix} 0.1 & 0.3 & 0.1 & 0.5 \\ 0.4 & 0.0 & 0.3 & 0.3 \\ 0.4 & 0.1 & 0.3 & 0.2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \qquad (3.13)$$

3.3.3.1 Absorbing Markov Chains Analysis Insights

Within the context of our analysis objectives, an absorbing state is a fixed point or steady state that once reached, the system never leaves. Similarly, valuable convergence insights can be gained when the system can be modeled as an absorbing Markov chain. However, establishing these convergence results requires the introduction of some additional matrices based on our transition matrix $\mathbf{P}$.

First let us write our Markov chain transition matrix in *canonical form* which is given by the modified transition matrix, $\mathbf{P}'$ shown in (3.14) where $\mathbf{I}^{ab}$ is the identity matrix corresponding to the state transitions between the absorbing states of the chain, $\mathbf{Q}$ represents the state transitions between the nonabsorbing states of the chain, $\mathbf{0}$ is a rectangular matrix filled with all zeros representing the probability of transition from absorbing states to nonabsorbing states, and $\mathbf{R}$ represents the rectangular matrix of state transition probabilities from nonabsorbing states to absorbing states. At this point we have not performed any operations on $\mathbf{P}$, merely relabeled the states in a way which we will find convenient.

(*canonical form*)
$$\mathbf{P}' = \left[\begin{array}{c|c} \mathbf{Q} & \mathbf{R} \\ \hline \mathbf{0} & \mathbf{I}^{ab} \end{array}\right] \tag{3.14}$$

Given $\mathbf{P}'$, Markov theory provides us with information on convergence and the expected frequency that the system visits a transitory state. First, we know that recursive evaluation of $\mathbf{P'}^k$ yields $\lim_{k\to\infty} \mathbf{Q}^k \to \mathbf{0}$. Recall that we earlier said that the entry p_{mn} in $\mathbf{P}^k$ represented the probability of the system initially occupying state p^k_{mn} in the m^{th} row and n^{th} column of $\mathbf{P}^k$ represents the probability that system is in state a^n k iterations after being in state a^m. Thus $\lim_{k\to\infty} \mathbf{Q}^k \to \mathbf{0}$ implies that that the probability of the system not being "absorbed", i.e., not terminating in one of the absorbing states of the chain, goes to zero.

Given an absorbing chain with a modified transition matrix as in (3.14), the *fundamental matrix* is given by (3.15).

(*fundamental matrix*) $$\mathbf{N} = \left(\mathbf{I} - \mathbf{Q}\right)^{-1} \tag{3.15}$$

Solving for the fundamental matrix **N** permits a number of valuable analytic insights. First, theorem 3.2.4 in [Kemeny_60] states that the entry n_{km} gives the expected number of times that the system will pass through state a_m given that the system starts in state a_k. Second, theorem 3.3.5 in [Kemeny_60] states that if we evaluate $\mathbf{t} = \mathbf{N}\mathbf{1}$ where **1** is a column vector of all ones, then t_k gives the expected number of iterations before the state is absorbed when the system starts in state a_k. Finally, theorem 3.3.7 in [Kemeny_60] states that if we evaluate (3.16)

$$\mathbf{B} = \mathbf{N}\mathbf{R} \tag{3.16}$$

where **R** is as given in (3.14), then entry b_{km} in **B** specifies the probability the system ends up in absorbing state a_m if the system starts in state a_k.

Thus, once we show that a Markov model for a network of cognitive radios with transition matrix **P** is an absorbing Markov chain, the following insights are readily gained:

- Steady-states for the system can be identified by finding those states a^m for which $p_{mm} = 1$.
- Convergence to one of these steady-states is assured, and the expected distribution of states can be found by solving for **B.**
- Given an initial state, a^k, convergence rate information is given by solving for **t**.

Example 3.2 describes a procedural cognitive radio DFS algorithm that can be modeled and analyzed using Markov models. It is interesting to note that with the additional stipulation that when adapting channels are chosen at random, the algorithm described in Example 3.2 can be readily scaled to any network of n radios with $c \geq n$ channels and still remain an absorbing Markov chain. However for $n > c$, the network is no longer an absorbing Markov chain and instead becomes an ergodic Markov chain. One approach to overcoming this limitation is to adjust the decision rule so that no radio switches to a channel which would be predicted to receive the same amount of interference. In such a case, the network can again be modeled as an absorbing Markov chain. Chapter 7 will consider additional techniques for ensuring desirable behavior for a DFS algorithm where $n > c$.

Example 3.2: DFS as an Absorbing Markov chain

Consider two cognitive radios implementing dynamic frequency selection over the two channel set $F=\{f_1, f_2\}$. Assume that these two radios are seeking to minimize the interference their signal and that both are implementing the simple decision rule that if an interfering signal is detected, then the radio switches to the other frequency.

Using the model from Chapter 2, this system can be modeled as $N = \{1,2\}$, $A = \{(f_1,f_1), (f_1,f_2), (f_2,f_1), (f_2,f_2)\}$, $u_j(a) = \begin{cases} 1 & f_j \neq f_{-j} \\ -1 & f_j = f_{-j} \end{cases}$, $d_j(f_j, f_{-j}) = \begin{cases} f_j & u_j(a) = 1 \\ f \in F \setminus f_j & u_j(a) = -1 \end{cases}$, and T is asynchronous where due to a random timer for each $t \in T$ each radio gets a chance of 0.5.

This model can then be converted into a Markov model with the transition matrix shown in (3.17) and illustrated in Figure 3.7.

$$\mathbf{P} = \begin{array}{c} \\ (f_1,f_1) \\ (f_1,f_2) \\ (f_2,f_1) \\ (f_2,f_2) \end{array} \begin{array}{c} \begin{array}{cccc} (f_1,f_1) & (f_1,f_2) & (f_2,f_1) & (f_2,f_2) \end{array} \\ \begin{bmatrix} 0.25 & 0.25 & 0.25 & 0.25 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0.25 & 0.25 & 0.25 & 0.25 \end{bmatrix} \end{array} \tag{3.17}$$

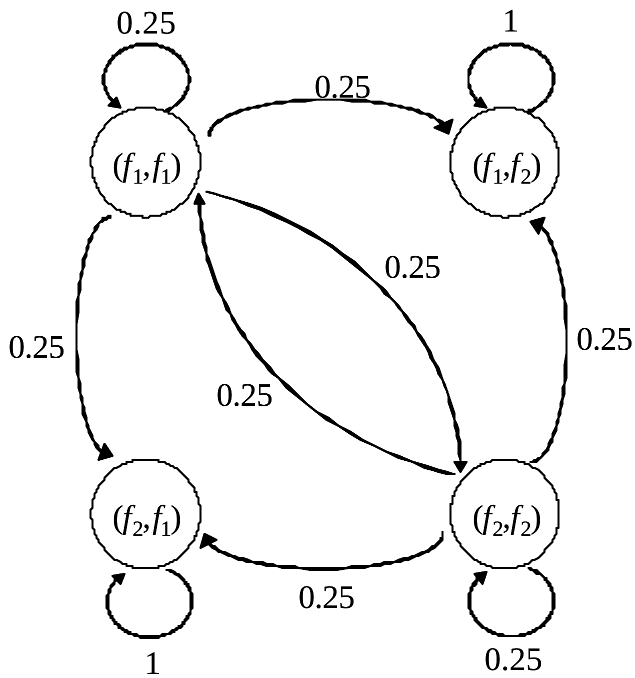

Figure 3.7: Digraph of DFS Example

.

Note that this Markov chain forms an absorbing Markov chain with (f_1,f_2) and (f_2,f_1) as absorbing states and {(f_1,f_1) and (f_2,f_2) as transient states. Thus we immediately know that this network has two steady-states [(f_1,f_2) and (f_2,f_1)] and that the network will converge to these states. Further, by evaluating (3.15)and (3.16) for **N**, **t**, and **B**, respectively, we can determine how long we can expect to remain in a transition state and how what the distribution of steady-states will be given an initial choice of channels.

So we know from (3.18) that with repeated trials of the network starts from (f_1,f_1), the system will on average pass through (f_1,f_1) 1.5 times and (f_1,f_1) 0.5 times, and we know from (3.19) that the system is equally likely to end up in either absorbing state.

$$\mathbf{N} = \begin{array}{c} \\ (f_1,f_1) \\ (f_2,f_2) \end{array} \begin{array}{c} \begin{array}{cc} (f_1,f_1) & (f_2,f_2) \end{array} \\ \begin{bmatrix} 1.5 & 0.5 \\ 0.5 & 1.5 \end{bmatrix} \end{array} \qquad (3.18)$$

$$\mathbf{B} = \begin{matrix} & (f_1,f_2) & (f_2,f_1) \\ (f_1,f_1) & 0.5 & 0.5 \\ (f_2,f_2) & 0.5 & 0.5 \end{matrix} \tag{3.19}$$

3.4 Summary

This chapter introduced several powerful techniques for analyzing the interactions of procedural cognitive radios based on the knowledge of an evolution function, d to determine steady-states, optimality, convergence, and stability. These models and the techniques for establishing if a cognitive radio network satisfies the conditions of the model are summarized in Table 3.1.

Table 3.1: Presented Models

Model (Section number)	Basic model	Identification
Dynamical Systems (3.1)	evolution equation $a\left(t^{k+1}\right)=d^t\left(a\left(t^k\right)\right)$	Assumed to exist
Contraction Mappings (3.2)	$\left\\|f(a), f(b)\right\\| \leq a\left\\|a,b\right\\|$ $\forall b,a\in A$	Blackwell's conditions
Standard Interference Function Power Control (3.2.4.1)	$d_j\left(\mathbf{p}\left(t^k\right)\right)=p_j\left(t^k\right)I_j\left(\mathbf{p}\left(t^k\right)\right)$	$I(\mathbf{p})$ satisfies positivity, montonicity, and scalability
Ergodic Markov Chain (3.3.2)	$P\left(a\left(t^{k+1}\right)=a^k \mid a(0),\ldots,a(t)\right)$ $=P\left(a\left(t^{k+1}\right)=a^k \mid a\left(t^k\right)\right)$	$\exists\, k$ such that $\mathbf{P}^k$ has all positive entries
Absorbing Markov Chain (3.3.3)	$\mathbf{P}'=\left[\begin{array}{c\|c} \mathbf{Q} & \mathbf{R} \\ \hline \mathbf{0} & \mathbf{I}^{ab} \end{array}\right]$	Apply model definition

For these models, this chapter presented analysis insights that can be gleaned by demonstrating that a procedural cognitive radio network satisfies the modeling conditions for one of the models listed in Table 3.1. The steady-state properties, the convergence properties, and the stability properties for each of these models are summarized in Table

3.2 Table 3.3, and Table 3.4, respectively. We also presented an approach to determining the desirability of network behavior –evaluation of a network objective function.

Table 3.2 Steady-State Properties by Model

Model (Section number)	Existence	Identification
Dynamical Systems (3.1)	Maybe, evaluate Leray-Schauder-Tychonoff theorem on evolution equation	Exhaustive Search, Solve $d\left(a^*\right) = a^*$
Contraction Mappings (3.2)	Yes (Banach's Theorem)	Recursion (Unique steady-state)
Standard Interference Function Power Control (3.2.4.1)	Yes ([Yates_95])	Recursion (Unique steady-state), $\mathbf{Z}\mathbf{p} = \vec{g}$
Ergodic Markov Chain (3.3.2)	Yes (Ergodocity theorem)	Recursion (Unique distribution), Solve $p^{*T}\mathbf{P} = p^{*T}$
Absorbing Markov Chain (3.3.3)	Yes (Definition)	$p_{mm} = 1$

Table 3.3 Convergence Properties by Model

Model (Section number)	Sensitivity	Rate
Dynamical Systems (3.1)	Apply Lyapunov's direct method (when possible)	No general technique
Contraction Mappings (3.2)	Everywhere convergent	$\left\| a\left(t^k\right), a^* \right\| \le \frac{a^k}{1-a} \left\| a\left(t^1\right), a\left(t^0\right) \right\|$
Standard Interference Function Power Control (3.2.4.1)	Everywhere convergent	$\left\| \mathbf{p}\left(t^k\right), \mathbf{p}^* \right\| \le a^k \left\| \mathbf{p}(0), \mathbf{p}^* \right\|$ $a = r\left(\mathbf{H}\right)$
Ergodic Markov Chain (3.3.2)	Converges to distribution from all starting distributions	Transition matrix dependent
Absorbing Markov Chain (3.3.3)	$\mathbf{B} = \mathbf{N}\mathbf{R}$	$\mathbf{t} = \mathbf{N}\mathbf{1}$

Table 3.4 Stability Properties by Model

Model (Section number)	Lyapunov Stability	Attractivity
Dynamical Systems (3.1)	Apply Lyapunov's direct method (when possible)	Apply Lyapunov's direct method (when possible)
Contraction Mappings (3.2)	Global	Global
Standard Interference Function Power Control (3.2.4.1)	Global	Global
Ergodic Markov Chain (3.3.2)	No	No
Absorbing Markov Chain (3.3.3)	Not guaranteed.	If unique absorbing state

As we saw with the Standard Interference Function, sometimes cognitive radio networks satisfy the conditions of multiple models. In these cases, the analytic insights from each of the applicable multiple models are available. While this Chapter presents a significant number of useful analytic results, the reader should be aware that this Chapter was only able to include a brief treatment of these extensive models. In fact, many of these models have entire disciplines dedicated to their analysis and application. Accordingly, the interested reader is encouraged to explore the texts listed in the references for further study.

3.5 References

[Berggren_01] F. Berggren, "Power Control, Transmission Rate Control and Scheduling in Cellular Radio Systems," PhD Dissertation Royal Institute of Technology, Stockholm, Sweden, May, 2001.

[Bertsekas_97] D. Bertsekas, J. Tsitsiklis, **Parallel and Distributed Computation: Numerical Methods**, Athena Scientific, Belmont MA, 1997.

[Blackwell_65] D. Blackwell, "Discounted Dynamic Programming," *The Annals of Mathematical Statistics* vol. 36, no 1, Feb. 1965, pp. 226-235.

[Kemeny_60] J. Kemeny and J. Snell, **Finite Markov Chains**, D. Van Nostrand Company, Princeton, New Jersey, 1960.

[Medio_01] A. Medio and M. Lines, **Nonlinear Dynamics: A Primer**, Cambridge University Press, Cambridge, UK, 2001.

[Stewart_94] W. Stewart, **Introduction to the numerical solution of Markov chains**, Princeton, NJ, Princeton University Press, 1994.

[Sundaram_99] R. Sundaram **A First Course in Optimization**, Cambridge University Press, New York, New York, 1999.

[Syski_92] R. Syski, **Passage Times for Markov Chains**, IOS Press,Washington DC

[Walker_80] J. Walker, **Dynamical Systems and Evolution Equations: Theory and Applications**, Plenum Press, New York, New York, 1980.

[Yates_95] R. Yates, "A Framework for Uplink Power Control in Cellular Radio Systems," *IEEE Journal on Selected Areas in Communications*, Vol. 13, No 7, September 1995, pp. 1341-1347.

[Zander_01] J. Zander, S. Kim, **Radio Resource Management for Wireless Networks**, Artech House, Norwood, MA 2001.

[Zangwill_69] W. Zangwill, **Nonlinear Programming: A Unified Approach**, Prentice Hall, Englewood Cliffs NJ, 1969.