Dual Degree Dissertation

On Two Control Problems: Stability of Dynamical Systems and Optimal Power Allocation

submitted in partial fulfillment of the
requirements for the degree of

Bachelor of Technology
and
Master of Technology
in *Communications and Signal Processing*
(under the Dual Degree Programme)

by

Mukul Agarwal
Roll no. 99D07042

under the guidance of
Prof. Harish Pillai
and
Prof. Abhay Karandikar

Department of Electrical Engineering
Indian Institute of Technology, Bombay
Powai, Mumbai-400076.
July, 2004.

Acknowledgements

This thesis is divided into two parts, Part I and Part II.

I thank Prof. N. K. Karmarkar, TIFR, Pune, who helped me in the initial stages of Part I. I also thank Prof. Harish Pillai for guiding me through Part I and for the many insightful technical discussions that I have had with him. I also thank Prof. J. K. Verma and Prof. Balwant Singh for their help and ideas.

I thank Prof. Abhay Karandikar for guiding me through Part II of the project. My most profuse acknowledgements go to Prof. V. S. Borkar, TIFR Bombay, without whose ideas, Part II wouldn't have been possible. I also thank Premal Shah, alumnus, IIT Bombay, for introducing me to the problem in Part II, and for the many technical and non-technical discussions that I had with him.

I thank Ashutosh Gore for the endless time he spent proof reading the thesis and the many errors that he pointed out.

I thank many members of the Information Networks Lab, without whom, the sail through the project would have been boring, and who have made my last few months stay at IIT Bombay, very memorable.

Finally, I thank my family members for their constant support and encouragement, and without whose support, the thesis just could not have been possible.

Mukul Agarwal

Abstract

In Part I of the thesis, we consider the problem of Liapunov Stability of the origin for the system given by differential equations

$$\begin{aligned} \dot{x} &= -y + f_1(x, y) \\ \dot{y} &= +x + f_2(x, y) \end{aligned} \tag{1}$$

(where f_1 and f_2 are polynomials with no constant or linear terms). We present a method to construct Liapunov Functions for such systems. We solve the problem of stability/instability for the set of Systems (1), generically, by using this construction methodology.

In part II of the thesis, we consider the problem of scheduling packets on a point to point wireless link with fading. We derive the structural properties of a policy that is optimal with respect to minimizing the average power subject to an average delay constraint. We consider the cases of both memoryless and Markov incoming packet distribution as well as channel fading. We prove that in case of First Order Stochastically Dominant (FSD) incoming packet distribution, and in presence of fading, there exists an optimal policy which is increasing both in the queue length and the number of packet arrivals. In case of memoryless fading, the optimal policy is also decreasing with the *level* of channel fade. Finally, if the fading is FSD, nothing can be said about the nature of the optimal policy with respect to the channel fade.

Contents

List of Figures

PART I

Stability of a Class of Systems using the Method of Liapunov Function Construction

Chapter 1

Introduction, the Problem, and an Example

The stability of planar systems

$$\begin{aligned}\dot{x} &= f_1(x,y)\\ \dot{y} &= f_2(x,y)\end{aligned}$$

is a long standing problem. When both the eigenvalues for the linearization have negative real parts, it is known that the system is stable. When at least one eigenvalue has positive real part, it is known that the system is unstable. The non-trivial cases are those where both the eigenvalues are purely imaginary or at least one of the eigenvalues is zero. In this part of the thesis, we consider the case when both the eigenvalues are purely imaginary.

Without loss of generality, assume that the origin is an equilibrium point. This can be done by a suitable translation of axes. The system of differential equations representing such a system is:

$$\begin{aligned}\dot{x} &= -k_1 y + f_1(x,y)\\ \dot{y} &= k_2 x + f_2(x,y)\end{aligned}$$

where f_1 and f_2 are polynomials with no constant or linear terms in x or y, and k_1 and k_2 are positive reals. The System (1),

$$\begin{aligned}\dot{x} &= -y + f_1(x,y)\\ \dot{y} &= x + f_2(x,y)\end{aligned}$$

is this system with $k_1 = k_2 = 1$. We consider the system (1). We shall see towards the end of this part of the thesis, that the same method works for the case when k_1 and k_2 are arbitrary positive reals.

A lot of work has been done on the problem of Liapunov Stability in the 1960s and 1970s (see for example, [1]). Most of this work tries to attack the problem of stability of the system

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

in its full generality. In this part of the thesis, we consider a special case, and derive much stronger results.

To motivate the method, we present an example.

A theorem presented in the next chapter will be used in this example.

Consider the system

$$\begin{aligned} \dot{x} &= -y - xy \\ \dot{y} &= +x + xy \end{aligned} \tag{1.1}$$

This is the same as a special case of the Lotka-Volterra equations (see [2]) with the stable point shifted to the origin (note that origin is an equilibrium point). For this system, it is easy to prove that a quadratic form will not work as a Liapunov Function.

Had the quadratic terms on the right hand side (RHS) of (1.1) not been there, a Liapunov Function would have been

$$V_2 = \frac{x^2 + y^2}{2}$$

But V_2 fails as a Liapunov Function candidate in this case since

$$\dot{V_2} = -x^2 y + xy^2$$

is not sign semidefinite.

We modify V_2 by adding new terms to it so that the derivative is modified.

Consider

$$V_3 = V_2 - \frac{x^3 + y^3}{3} = \frac{x^2 + y^2}{2} - \frac{x^3 + y^3}{3}$$

$$\dot{V}_3 = x^3y - xy^3$$

which is, again, not sign semidefinite.

But the "advantage" of V_3 over V_2 is that $\dot{V}_3$ has monomials of degree 4, whereas $\dot{V}_2$ had monomials of "***lesser***" degree 3. Thus, in a neighborhood of the origin, $\dot{V}_3$ is "closer" to zero than $\dot{V}_2$. Thus, assuming that V_i has been formed, new terms are added to V_i so that the monomials in $\dot{V}_i$ which hindered its sign semidefiniteness are cancelled. Let this new function be denoted as V_{i+1}. Thus, there might be new monomials in $\dot{V}_{i+1}$, but these monomials will be of higher degree than those of $\dot{V}_i$. With this view, consider

$$\begin{aligned} V_n &= \frac{x^2+y^2}{2} - \frac{x^3+y^3}{3} + \frac{x^4+y^4}{4} - \cdots \\ &\quad +(-1)^n\frac{x^n+y^n}{n} \end{aligned}$$

Then,

$$\dot{V}_n = (-1)^n[x^{n+1}y - xy^{n+1}]$$

and in the limit,

$$\begin{aligned} V &= \frac{x^2+y^2}{2} - \frac{x^3+y^3}{3} + \frac{x^4+y^4}{4} - \cdots \\ &\quad +(-1)^n\frac{x^n+y^n}{n} + \cdots \end{aligned}$$

and

$$\dot{V} = 0$$

Thus, V is a Liapunov Function, which can be written, more simply, as

$$V = x - \log(1+x) + y - \log(1+y)$$

which is the standard Liapunov function for the Lotka-Volterra system mentioned in books, but arrived at by the method of separation of variables. By Theorem 2.1 of Chapter 2, the Lotka-Volterra system is stable at the origin.

As suggested by this example, we will use the following approach. Suppose V_i has been formed. V_{i+1} is formed from V_i by adding a homogeneous polynomial of degree $(i+1)$ to V_i - this homogeneous polynomial is chosen in such a way that V_{i+1} is positive definite and that, the monomials of the least degree in $\dot{V}_i$ which hindered its sign semidefiniteness get cancelled - and such that the monomials in $\dot{V}_{i+1}$ (except the "innocuous" monomials which don't hinder sign semidefiniteness) are of a higher degree than the least degree monomials of $\dot{V}_i$ which hindered its sign semidefiniteness.

We present the method in detail in Chapter 3.

Chapter 2

Mathematical Preliminaries

In this chapter, we state a few theorems which will be used in Part I of this thesis.

By stability/instability, we would mean stability/instability in the sense of Liapunov. See [3] for definitions.

Theorem 2.1 (Liapunov's Stability Criterion) *Consider the system* $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, *and let origin be an equilibrium point. If there exists a function* $V(\mathbf{x})$ *such that*

1. V *is positive definite in a neighborhood of the origin,*
2. $\dot{V}$ *is negative semidefinite in a neighborhood of the origin,*

then the system is stable at the origin.

If

1. V *is positive definite in a neighborhood of the origin,*
2. $\dot{V} = 0$ *in a neighborhood of the origin,*

then the system is stable but not asymptotically stable at the origin.

If

1. V *is positive definite in a neighborhood of the origin,*

2. $\dot{V}$ *is negative semidefinite in a neighborhood of the origin and does not vanish identically on solution trajectories when the system starts in this neighborhood,*

then the system is asymptotically stable at the origin.

Proof A proof can be found in [3]. ■

Theorem 2.2 *If the system* $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ *is asymptotically stable at the origin, the system* $\dot{\mathbf{x}} = -\mathbf{f}(\mathbf{x})$ *is unstable at the origin.*

Proof This follows from the definition of asymptotic stability and instability. ■

Theorem 2.3 (Ratio Test) *Consider the power series* $A = \sum_{n=0}^{\infty} a_n z^n$. *Let*

$$\limsup_{n\to\infty} \frac{a_{n+1}}{a_n} = M$$

Then, the power series has a radius of convergence $R = \frac{1}{M}$.

Proof A proof can be found in [4] ■

Theorem 2.4 *Let* $\sum_n x_n$ *and* $\sum_n y_n$ *converge. Then,* $\sum_n [x_n + y_n]$ *converges too.*

Proof A proof can be found in [4]. ■

Theorem 2.5 (Hilbert Basis Theorem) *Consider an infinite system of polynomial equalities in* n *variables* $x_1, x_2, \ldots, x_n$

$$\begin{aligned} z_1(x_1, x_2, \ldots, x_n) &= 0 \\ z_2(x_1, x_2, \ldots, x_n) &= 0 \\ &\vdots \\ z_k(x_1, x_2, \ldots, x_n) &= 0 \\ &\vdots \end{aligned}$$

Then, there exists an integer k *such that the simultaneous solution set of the above infinite set of equations is the same as that of the first* k *equations.*

Proof A proof can be found in [5]. ■

Remark 2.1 *Note that the theorem* ***does not*** *say that there are no redundancies in the first k equations. It just says that the solution set of the first k equations is exactly the same as that of all the equations.*

In other words, an infinite system of polynomial equalities is equivalent to a finite system of polynomial equalities, and that, the solution is equivalent to the simultaneous solution of the first k equations for some k.

Chapter 3

Main Results and Conclusions

In this chapter, we state and prove an algorithm for determining the stability/instability of a large class of systems of the type (1), and conclude Part I of the thesis.

3.1 The Main Result

We first state some definitions.

Definition 3.1 *Consider the system (1)*

$$\begin{aligned} \dot{x} &= -y + f_1(x,y) \\ \dot{y} &= +x + f_2(x,y) \end{aligned}$$

Let W be an arbitrary function of x and y. The total derivative of W is

$$\begin{aligned} \dot{W} &= \dot{x}\frac{\partial W}{\partial x} + \dot{y}\frac{\partial W}{\partial y} \\ &= [-y + f_1(x,y)]\frac{\partial W}{\partial x} + [x + f_2(x,y)]\frac{\partial W}{\partial y} \\ &= \underbrace{-y\frac{\partial W}{\partial x} + x\frac{\partial W}{\partial y}}_{T_1(W)} + \underbrace{f_1(x,y)\frac{\partial W}{\partial x} + f_2(x,y)\frac{\partial W}{\partial y}}_{T_2(W)} \\ &= T_1(W) + T_2(W) \end{aligned}$$

Thus, define,

$$\begin{aligned} T_1(W) &= -y\frac{\partial W}{\partial x} + x\frac{\partial W}{\partial y} \\ T_2(W) &= f_1(x,y)\frac{\partial W}{\partial x} + f_2(x,y)\frac{\partial W}{\partial y} \end{aligned}$$

Definition 3.2 (Total degree, x-degree, y-degree) *Let $p = x^i y^j$ be a monomial. The total degree of p is $(i+j)$. The x-degree of p is i. The y-degree of p is j.*

We will refer to the total degree as just the degree.

Definition 3.3 ($Hom_n(W)$) *Let W be an arbitrary polynomial in x and y. $Hom_n(W)$= homogeneous polynomial of degree n in W.*

Thus, $Hom_n(W)$ is a polynomial of degree n such that $[W - Hom_n(W)]$ has no monomials of degree n.

Lemma 3.1 *Let W be a homogeneous polynomial in x and y of degree n.*

If n is odd, there exists a homogeneous polynomial V_W of degree n such that $W + T_1(V_W) = 0$.

If n is of the form $4k$ for some integer k, there exist a homogeneous polynomial V_W of degree n and a real number k_W such that $W + T_1(V_W) = k_W x^{2k} y^{2k}$.

If n is of the form $4k+2$ for some integer k, there exist a homogeneous polynomial V_W of degree n and a real number k_W such that $W + T_1(V_W) = k_W x^{2k} y^{2k+2}$.

V_W thus obtained will be denoted by $Cancel(W)$.

Proof Let

$$W = \sum_{i=0}^{n} a_i x^{n-i} y^i$$

We first state a procedure to find $Cancel(W)$.

Begin

1. $Cancel(W) = 0$

2. If n is odd,

   (a) For($i = 0$; $i \leq n - 1$; $i = i + 2$)

      i. $T = \frac{-a_i}{i+1} x^{n-(i+1)} y^{i+1}$

      ii. $Cancel(W) = Cancel(W) + T$

      iii. $W = W + T_1(T)$

      End for

   (b) For($i = n$; $i \geq 1$; $i = i - 2$)

      i. $T = \frac{a_i}{n-(i-1)} x^{n-(i-1)} y^{i-1}$

      ii. $Cancel(W) = Cancel(W) + T$

      iii. $W = W + T_1(T)$

      End for

3. If n is divisible by 4,

   (a) For($i = 0$; $i \leq n/2 - 1$; $i = i + 2$)

      i. $T = \frac{-a_i}{i+1} x^{n-(i+1)} y^{i+1}$

      ii. $Cancel(W) = Cancel(W) + T$

      iii. $W = W + T_1(T)$

      End for

   (b) For($i = n/2 + 1$; $i \leq n - 1$; $i = i + 2$)

      i. $T = \frac{-a_i}{i+1} x^{n-(i+1)} y^{i+1}$

      ii. $Cancel(W) = Cancel(W) + T$

      iii. $W = W + T_1(T)$

      End for

   (c) For($i = n/2 - 1$; $i \geq 1$; $i = i - 2$)

      i. $T = \frac{a_i}{n-(i-1)} x^{n-(i-1)} y^{i-1}$

      ii. $Cancel(W) = Cancel(W) + T$

      iii. $W = W + T_1(T)$

      End for

   (d) For($i = n$; $i \geq n/2 + 1$; $i = i - 2$)

      i. $T = \frac{a_i}{n-(i-1)} x^{n-(i-1)} y^{i-1}$

ii. $Cancel(W) = Cancel(W) + T$

iii. $W = W + T_1(T)$

End for

4. If $n-2$ is divisible by 4,

   (a) For($i=0$; $i \leq n/2-1$; $i=i+2$)

      i. $T = \frac{-a_i}{i+1} x^{n-(i+1)} y^{i+1}$

      ii. $Cancel(W) = Cancel(W) + T$

      iii. $W = W + T_1(T)$

      End for

   (b) For($i = n/2+2$; $i \leq n-1$; $i=i+2$)

      i. $T = \frac{-a_i}{i+1} x^{n-(i+1)} y^{i+1}$

      ii. $Cancel(W) = Cancel(W) + T$

      iii. $W = W + T_1(T)$

      End for

   (c) For($i = n/2$; $i \geq 1$; $i = i-2$)

      i. $T = \frac{a_i}{n-(i-1)} x^{n-(i-1)} y^{i-1}$

      ii. $Cancel(W) = Cancel(W) + T$

      iii. $W = W + T_1(T)$

      End for

   (d) For($i = n$; $i \geq n/2+2$; $i = i-2$)

      i. $T = \frac{a_i}{n-(i-1)} x^{n-(i-1)} y^{i-1}$

      ii. $Cancel(W) = Cancel(W) + T$

      iii. $W = W + T_1(T)$

      End for

5. $Cancel(W)$ is the required polynomial.

End

Note that the value of W is updated many times in the procedure, and that, a_i refers to the updated coefficient of $x^{n-i}y^i$ in W.

We now prove that the $Cancel(W)$ generated by this procedure indeed satisfies the claims stated in the lemma. We do this for the case when n is odd - a similar proof can be given when n is even.

Let $Cancel(W) = \sum_{i=0}^{n} q_i x^{n-i} y^i$. From the algorithm, it is easy to see that

$$\begin{aligned} q_1 &= -a_0, \text{ and recursively,} \\ q_{2k+1} &= \frac{[n-(2k-1)]q_{2k-1} - a_{2k}}{2k+1} \end{aligned}$$

$$\begin{aligned} q_{n-1} &= a_n, \text{and recursively,} \\ q_{n-(2k+1)} &= \frac{[n-(2k-1)]q_{n-(2k-1)} + a_{n-2k}}{2k+1} \end{aligned}$$

From these relations, it follows in one step by direct calculation that $W + T_1(Cancel(W)) = 0$.

Hence proved. ■

Definition 3.4 *Let $m_1 = x^i y^j$ and $m_2 = x^p y^q$ be two monomials. Then, m_1 is greater than m_2, or,*

$$m_1 > m_2 \tag{3.1}$$

if $i \leq p$, $j \leq q$, and $(i+j) < (p+q)$.

Similarly,

$$m_1 < m_2 \tag{3.2}$$

if $i \geq p$, $j \geq q$, and $(i+j) > (p+q)$.

The motivation of this definition is that if i and j are even, $x^i y^j > x^p y^q$ means that there exists a neighborhood of the origin in which the value of $x^i y^j$ is greater than the value of $x^p y^q$ at each point.

Theorem 3.1 *Consider the system (1). For each $n \geq 2$, there exists a polynomial V_n of degree n such that V_n is positive definite in a neighborhood of the origin, and that, the only monomials in $\dot{V}_n$ of degree $\leq n$ are of the form $x^{2k}y^{2k}$ and $x^{2l}y^{2l+2}$.*

Proof We prove the theorem by induction. For $n = 2$, it is easy to check that $V_2 = \frac{x^2+y^2}{2}$ works. Assume, that the theorem is true for $n = k$; thus, there exists a V_k such that the hypothesis of the theorem is satisfied for $n = k$.

Let $W = Hom_{k+1}(\dot{V}_k)$. Consider $V_{k+1} = V_k + Cancel(W)$. Note that $Hom_2(V_{k+1}) = \frac{x^2+y^2}{2}$ is positive definite. Also, all monomials other than x^2 and y^2 in V_{k+1} have degree greater than or equal to 3, and hence, all other monomials are less than atleast one of x^2 or y^2. Hence, the sign of V_{k+1} is determined by the sign of $\frac{x^2+y^2}{2}$ in a neighborhood of the origin, which is positive. Hence, V_{k+1} is positive definite in a neighborhood of the origin.

The hypothesis of the theorem regarding $\dot{V}_{k+1}$ follows by Lemma 3.1.

Hence, V_{k+1} satisfies the hypothesis of the theorem, and by induction, the theorem follows.
■

Assume that $V = \lim_{n\to\infty} V_n$ exists. Then, V is positive definite in a neighborhood of the origin such that

$$\dot{V} = k_{22}x^2y^2 + k_{24}x^2y^4 + k_{44}x^4y^4 + \cdots \tag{3.3}$$

is sign semidefinite in a neighborhood of the origin. We can then answer the question of stability/instability of the system at the origin by use of Theorems 2.1 and 2.2. Unfortunately, it is not clear that V thus obtained would be convergent or not.

We now prove an important result:

Theorem 3.2 *Consider the system (1)*

$$\begin{aligned}\dot{x} &= -y + f_1(x,y)\\ \dot{y} &= x + f_2(x,y)\end{aligned}$$

For this system, compute V_ns successively as suggested by the proof of Theorem 3.1. Thus, V_n is positive definite, and the only monomials in $\dot{V}_n$ which are not of the form $x^{2k}y^{2k}$ or $x^{2k}y^{2k+2}$ have degree greater than n. Let n_0 be the least integer such that monomial of degree n_0 in $\dot{V}_{n_0}$ of the form $x^{2k}y^{2k}$ or $x^{2l}y^{2l+2}$ ($4k = n_0$ or $4l+2 = n_0$ respectively) has a non-zero coefficient. Let this monomial be $x^{p_x}y^{p_y}$ ($p_x + p_y = n_0$ and either $p_x = p_y$ or $p_x = p_y - 2$). Then, the system is asymptotically stable if this coefficient is negative, and unstable, if this coefficient is positive.

Remark 3.1 *Note that n_0 might not exist, that is, coefficient of monomials $x^{2k}y^{2k}$ or $x^{2l}y^{2l+2}$ is zero in each $\dot{V}_n$. In this theorem, we assume that n_0 exists, and is finite. The case when n_0 does not exist is considered later in this section.*

The proof of this theorem is based on the fact that we can obtain a convergent series for V such that $\dot{V}$ is sign semidefinite. Once an n_0 (in the notation of the above theorem) is found, the procedure for adding more terms to obtain the relevant V is changed, so that a convergent series results. We elaborate this procedure in the algorithm below. This new procedure rests on the fact that at any stage n, all monomials in $\dot{V}_n$ which are less than $x^{p_x}y^{p_y}$ can be neglected, and hence, V_{n+1} that is formed does not have "a large number" of extra terms as compared to V_n, not large enough to make the infinite series, $V = \lim_{n\to\infty} V_n$, divergent. The rest of this section builds towards the proof of the theorem, and the proof itself is at the end of the section. First, we state the algorithm for constructing the modified V_ns. The required Liapunov Function, V, will be $\lim_{n\to\infty} V_n$.

Begin Algorithm

1. Following the procedure of Theorem 1, get either
   - V_{4k} such that V_{4k} is positive definite and $\dot{V}_{4k} = mx^{2k}y^{2k}$ + (monomials of order higher than $4k$), $m \neq 0$, or,
   - V_{4l+2} such that V_{4l+2} is positive definite and $\dot{V}_{4l+2} = mx^{2l}y^{2l+2}$ + (monomials of order higher than $4l+2$), $m \neq 0$

2. If the former happens, define $n_0 = 4k$, $p_x = 2k$ and $p_y = 2k$. If the latter happens, define $n_0 = 4l+2$, $p_x = 2l$ and $p_y = 2l+2$.

3. For $n = p_x + p_y$ to ∞ do

   (a) $V_{n+1} = V_n$

   (b) $W = Hom_{n+1}(\dot{V}_n)$

   (c) For each term $a_{ij}x^iy^j$ in W do

      i. If $x^iy^j < x^{p_x}y^{p_y}$, define, $Y_{ij} = 0$

      ii. If NOT$(x^iy^j < x^{p_x}y^{p_y})$, $i > j$ and j is odd, define,

$$Y_{ij} = \sum_{k=j+1}^{p_y} r_k x^{i+j-k} y^k \tag{3.4}$$

      where, r_ts $(j+1 \leq t \leq p_y)$ satisfy:

      $r_t = 0$ for t odd, and for t even, define r_ts recursively by

$$\begin{aligned} r_{j+1} &= -\frac{a_{ij}}{j+1} \\ &\vdots \\ r_{j+2m+1} &= \frac{r_{j+2m-1}(i-2m+1)}{j+2m+1} \\ &\vdots \end{aligned}$$

iii. If NOT$(x^iy^j < x^{p_x}y^{p_y})$, $i > j$ and j is even, define,

$$Y_{ij} = \sum_{k=j+1}^{p_y-1} r_k x^{i+j-k} y^k \tag{3.5}$$

where r_ts $(j+1 \le t \le p_y - 1)$ satisfy:

$r_t = 0$ for t even, and for t odd, define r_t recursively by

$$\begin{aligned} r_{j+1} &= -\frac{a_{ij}}{j+1} \\ &\vdots \\ r_{j+2m+1} &= \frac{r_{j+2m-1}(i-2m+1)}{j+2m+1} \\ &\vdots \end{aligned}$$

iv. If NOT$(x^iy^j < x^{p_x}y^{p_y})$, $i < j$ and i is odd, define,

$$Y_{ij} = \sum_{k=i+1}^{p_x} r_k x^k y^{i+j-k} \tag{3.6}$$

where r_ts $(i+1 \le t \le p_x)$ satisfy:

$r_t = 0$ for t odd, and for t even, define r_t recursively by

$$\begin{aligned} r_{i+1} &= \frac{a_{ij}}{i+1} \\ &\vdots \\ r_{i+2m+1} &= \frac{r_{i+2m-1}(j-2m+1)}{i+2m+1} \\ &\vdots \end{aligned}$$

v. If NOT$(x^iy^j < x^{p_x}y^{p_y})$, $i < j$ and i is even, define,

$$Y_{ij} = \sum_{k=i+1}^{p_x-1} r_k x^k y^{i+j-k} \tag{3.7}$$

where r_ts $(i+1 \le t \le p_x - 1)$ satisfy: $r_t = 0$ for t even, and for t odd, define r_t recursively by

$$\begin{aligned} r_{i+1} &= \frac{a_{ij}}{i+1} \\ &\vdots \\ r_{i+2m+1} &= \frac{r_{i+2m-1}(j-2m+1)}{i+2m+1} \\ &\vdots \end{aligned}$$

vi. Compute $V_{n+1} = V_{n+1} + Y_{ij}$

End step 3(c)

End step 3

End Algorithm

The algorithm rests on the fact that the V_n computed at each stage is positive definite (in a neighborhood of the origin) and that, $\dot{V}_n$ (for $n \geq n_0$, n_0 as in Theorem 3.2) has terms, only of the following type:

- The monomial $x^{p_x}y^{p_y}$ (p_x and p_y are defined in Theorem 3.2 and also, in the Algorithm Step 2 above. Just to emphasize, p_x and p_y are constants for the system at hand, and bear no relation to n).
- Monomials which are less than $x^{p_x}y^{p_y}$.
- Monomials with degree $(n+1)$ or higher.

To prove this, note that V_n consists of the following terms:

- x^2, y^2.
- Monomials which are less than at least one of x^2 or y^2 (This is because all other monomials are of degree ≥ 3).

Thus, there exists a neighborhood of the origin in which the sign of V_n is dominated by the sign of $x^2 + y^2$, which is positive. Hence, there exists a neighborhood around the origin in which V_n is positive definite.

We use induction to prove the assertion regarding $\dot{V}_n$. By the Algorithm Step 1, it is clear that for $n = n_0$, $\dot{V}_{n_0}$ contains only the monomial $x^{p_x}y^{p_y}$ and monomials of degree greater than n_0, and hence, satisfies the assertion. Assume the assertion is true for $n = k$. We, now, prove it for $n = k+1$.

By the Algorithm, for each term $a_{ij}x^iy^j$ of degree $(k+1)$ in $\dot{V}_k$, a polynomial Y_{ij} is formed which is homogeneous of degree $(i+j)$. It can be easily checked from the Algorithm Step 3c that Y_{ij} is formed in such a way that

$$a_{ij}x^iy^j + T_1(Y_{ij}) = M \tag{3.8}$$

where M is a term which is less than $x^{p_x}y^{p_y}$.

To form V_{k+1} from V_k, for each term $a_{ij}x^iy^j$ of degree $k+1$ $(i+j=k+1)$ in $\dot{V}_k$, Y_{ij} is added to V_k. Thus, V_{k+1} is such that the only terms in $\dot{V}_{k+1}$ which have degree less than or equal to $(k+1)$ are less than $x^{p_x}y^{p_y}$. Thus, $\dot{V}_{k+1}$ contains terms, only of the following type:

- The monomial $x^{p_x}y^{p_y}$.
- Monomials which are less than $x^{p_x}y^{p_y}$.
- Monomials with degree (k + 2) or higher.

The induction step is complete, and the assertion is proved.

Consider

$$P = \sum_{k=1}^{n} Hom_k(\dot{V}_n) \tag{3.9}$$

From the assertion proved above, it follows that P is sign semidefinite in a neighborhood of the origin — its sign being determined by that of the coefficient of the monomial $x^{p_x}y^{p_y}$. This is because monomials in P other than $x^{p_x}y^{p_y}$ are less than $x^{p_x}y^{p_y}$.

Now, consider $V = \lim_{n\to\infty} V_n$. Assuming that the radius of convergence of V is greater than zero, V is positive definite in a neighborhood of the origin and $\dot{V}$ is sign definite in a neighborhood of the origin. This is for the following reason. V consists of terms in x^2, y^2, and all other terms in V are less than at least one of x^2, y^2. Thus, V can be written in the form

$$\begin{aligned} V &= x^2\Big[\frac{1}{2} + xg_1(x,y) + yg_2(x,y)\Big] \\ &\quad + y^2\Big[\frac{1}{2} + xg_3(x,y) + yg_4(x,y)\Big] \end{aligned}$$

where g_1, g_2, g_3, g_4 are convergent in a neighborhood of the origin. Clearly, there exists a neighborhood of the origin where V is positive definite (for example, where $|xg_1|$, $|yg_2|$, $|xg_3|$, $|yg_4|$ is each less than $\frac{1}{8}$).

Similarly, $\dot{V}$ has all terms (except $x^{p_x}y^{p_y}$) less than $x^{p_x}y^{p_y}$. Thus, $\dot{V}$ can be written as

$$\dot{V} = mx^{p_x}y^{p_y}[1 + xh_1(x,y) + yh_2(x,y)]$$

Note that p_x and p_y are even. Clearly, $\dot{V}$ is sign semidefinite in a neighborhood of the origin — its sign being determined by the sign of m. Also, since $x^{p_x}y^{p_y}$ is zero only on the x and y axes, same is true for $\dot{V}$.

To sum up, assuming that the V is convergent in a neighborhood of the origin, ,

- V is positive definite in a neighborhood of the origin
- $\dot{V}$ is sign-semidefinite in a neighborhood of the origin
- $\dot{V}$ is zero only on the x and y axes.

Now, we use Theorems 2.1 and 2.2 to prove asymptotic stability/instability First, we consider the case when the coefficient m of the monomial $x^{p_x}y^{p_y}$ in $\dot{V}$ is negative. Hence, V is positive definite and $\dot{V}$ is negative semidefinite in a neighborhood of the origin. By Theorem 2.1, the system is stable in a neighborhood of the origin. Hence, for every $\epsilon > 0$, $\exists\ \delta_\epsilon > 0$ such that if the trajectory starts in δ_ϵ ball around the origin, it stays forever, in the ϵ ball. To prove asymptotic stability, we need to prove (by Theorem 2.1) that $\dot{V}$ does not vanish identically on solution trajectories. The only points where $\dot{V}$ vanishes are the x and y axes. Thus, we need to show that if the trajectory reaches the x or the y axis, it leaves the same. Assume that the trajectory has reached the x axis. On this axis,

$$\begin{aligned} \dot{x} &= g_1(x) \\ \dot{y} &= x + g_2(x) \end{aligned} \tag{3.10}$$

where g_1 and g_2 are polynomials with each monomial of degree ≥ 2. Choose ϵ_0 small enough so that the polynomial $[x + g_2(x)]$ has no roots except zero in the interval $[-\epsilon_0, +\epsilon_0]$ (since a polynomial has a finite number of roots, this can obviously be done). Now, if the trajectory started in a δ_{ϵ_0} ball around the origin, the trajectory stays forever in the ϵ_0 ball. In this ball, on the x axis, as suggested by equation (3.10), $\dot{y} \neq 0$ (except at the origin), and this implies that if the trajectory reaches the x axis, it leaves it (except if it started at the origin in which case it stays, forever, at the origin). Similar reason can be given that a neighborhood around the origin can be found such that if the trajectory started in this neighborhood then whenever it reaches the y axis, it is bound to leave it. To sum up, a neighborhood around the origin can be chosen such that a trajectory starting in that neighborhood, on reaching the x or y axis, leaves the same. Hence, $\dot{V}$ is not identically zero on solution trajectories (since the only points where $\dot{V}$ is zero are the x and the y axes. By use of Theorem 2.1, the system is asymptotically stable at the origin.

Next consider the case when the coefficient m of $x^{p_x}y^{p_y}$ in $\dot{V}$ is positive.

Consider the system

$$
\begin{aligned}
\dot{x} &= -[-y + f_1(x,y)] \\
\dot{y} &= -[+x + f_2(x,y)]
\end{aligned} \tag{3.11}
$$

This is the original system with the signs of the RHSs reversed. For this system (3.11), if the same V is used, it is easy to see that $\dot{V}$ is the negative of what it is for the original system, and hence, for this system (3.11), $\dot{V}$ has the coefficient of $x^{p_x}y^{p_y}$ negative. The system (3.11) is thus, asymptotically stable. By Theorem 2.2, the original system is unstable at the origin.

We now only need to prove that V is convergent in a neighborhood of the origin.

First we consider some examples to illustrate the idea of the proof of convergence.

Example 3.1 *Consider the system*

$$
\begin{aligned}
\dot{x} &= -y + a_1x^2 + 2h_1xy + b_1y^2 \\
\dot{y} &= x + a_2x^2 + 2h_2xy + b_2y^2
\end{aligned} \tag{3.12}
$$

Assume that for this system, with the notation of the algorithm, $p_x = p_y = 2$ (that is, the coefficient of the x^2y^2 monomial in $\dot{V}_4$ is non zero). Then, the system is asymptotically stable or unstable, respectively as the coefficient of x^2y^2 monomial in $\dot{V}_4$ is negative or positive.

Assume that V_i, $i \geq 4$, has been formed. Then, it is easy to see (from the construction Step 3b in the above Algorithm), that

$$
V_{i+1} = V_i + ax^iy + bx^{i-1}y^2 + cx^2y^{i-1} + dxy^i
$$

for some a, b, c, d. That is, monomials x^iy^j where both $i, j > 2$ are not present. Intuitive reason for this is that only the x^{i+1}, x^iy, xy^i, y^{i+1} terms in $\dot{V}_i$ need to be cancelled (since other monomials are less than x^2y^2), and this can be done, respectively, by introducing x^iy, $x^{i-1}y^2$, x^2y^{i-1}, xy^i terms in V_{i+1}. Let

$$
M > \max(|a_1|, |b_1|, |h_1|, |a_2|, |b_2|, |h_2|). \tag{3.13}
$$

Assume, that for a particular degree i (for example, take $i = 5$),

$$
Hom_i(V_i) = p_ix^{i-1}y + q_ix^{i-2}y^2 + r_ix^2y^{i-2} + s_ixy^{i-1}
$$

Then, from the above Algorithm, Step 3c, it follows that

$$\begin{aligned} V_{i+1} &= V_i - a_2 p_i x^i y - \left[a_2 q_i + \frac{h_2 p_i}{2} + \frac{(i-1)a_1 p_i}{2}\right] x^{i-1} y^2 \\ &\quad + \text{similar terms for } xy^i \text{ and } x^2 y^{i-1} \end{aligned}$$

The idea is to prove the absolute convergence of V, and for this, we place upper bounds on $|p_{i+1}|$ and $|q_{i+1}|$ as follows.

Recalling the definition of M given above, it is easy to see that

$$\begin{aligned} |p_{i+1}| &< M|p_i| \\ |q_{i+1}| &< M\left[|q_i| + \frac{i|p_i|}{2}\right] \end{aligned}$$

This implies that

$$\begin{aligned} |p_{i+1}| &< M|p_i| \\ |q_{i+1}| &< M[|q_i| + i|p_i|] \end{aligned}$$

Consider the following pair of recurrence relations

$$\begin{aligned} p'_{i+1} &= Mp'_i \\ q'_{i+1} &= M[q'_i + ip'_i] \end{aligned}$$

Assume, that $p'_5 = |p_5|$ and $q'_5 = |q_5|$. Then, it is clear that $p'_n > |p_n|$ and $q'_n > |q_n|$ for $n > 5$.

Assume, for simplicity of presentation, that $M = 1$. Then,

$$p'_{i+k} = p'_i, \forall k$$

and hence,

$$\begin{aligned} q'_{i+1} - q'_i &= ip'_i \\ q'_{i+2} - q'_{i+1} &= (i+1)p'_i \\ q'_{i+3} - q'_{i+2} &= (i+2)p'_i \\ &\vdots \\ q'_{i+n} - q'_{i+(n-1)} &= (i+n-1)p'_i \end{aligned}$$

Adding all these equations,

$$q'_{i+n} = q'_i + nip'_i + \frac{n(n-1)p'_i}{2}$$

For a general M , with the same initial p_i and q_i, it is easy to check that,

$$\begin{aligned} p'_{i+n} &= M^n p'_i \\ q'_{i+n} &= M^n \left[q'_i + nip'_i + \frac{n(n-1)p'_i}{2} \right] \end{aligned}$$

This can be proved by induction and the proof is simple and hence, omitted. Now,

$$\begin{aligned} \lim_{n\to\infty} \frac{p'_{i+n+1}}{p'_{i+n}} &= M \\ \lim_{n\to\infty} \frac{q'_{i+n+1}}{q'_{i+n}} & \\ = \lim_{n\to\infty} \frac{M^{n+1}\left[q'_i + (n+1)ip'_i + \frac{n(n+1)p'_i}{2}\right]}{M^n\left[q'_i + nip'_i + \frac{(n-1)np'_i}{2}\right]} &= M \end{aligned}$$

and M is a constant (note that p'_i and q'_i are constants). Hence, by the ratio test (Theorem 2.3), the series, $\sum p'_n z^n$ and $\sum q'_n z^n$ are absolutely convergent in a neighborhood of the origin with radius of convergence say R_1 and R_2 respectively, where $R_1 = R_2 = \frac{1}{M}$.

Hence, clearly, the series

$$X'_1 = \sum_{n=i}^{\infty} q'_n x^{n-2} y^2; \quad X'_2 = \sum_{n=i}^{\infty} p'_n x^{n-1} y$$

are absolutely convergent. X'_1 being convergent for $|x| \leq R_1$ and arbitrary y. X'_2 is convergent for $|x| \leq R_2$, and arbitrary y.

Hence,

$$X_1 = \sum_{n=i}^{\infty} q_n x^{n-2} y^2; \quad X_2 = \sum_{n=i}^{\infty} p_n x^{n-1} y$$

are also convergent in the same regions (since $|p'_n| \geq |p_n|$ and $|q'_n| \geq |q_n|$.

In a similar way, we can prove that

$$X_3 = \sum_{n=i}^{\infty} r_n y^{n-2} x^2; \quad X_4 = \sum_{n=i}^{\infty} s_n y^{n-1} x$$

are also absolutely convergent. X_3 is convergent in a region $|y| \leq R_3$ (for some R_3), x arbitrary, and X_4 is convergent in a region $|y| \leq R_4$ (for some R_4), x arbitrary. Let $R_0 = \min(R_1, R_2, R_3, R_4)$. Then, it is clear that X_1, X_2, X_3, X_4 are all convergent in a 2 dimensional R_0 ball around the origin.

Now that

$$V = P + \sum_{n=i}^{\infty}[p_n x^{n-1} y + q_n x^{n-2} y^2 + r_n y^{n-2} x^2 + s_n y^{n-1} x]$$

where P is a polynomial. The following Lemma will be used to prove the convergence of V.

Lemma 3.2 *Let $\sum_n x_n$ and $\sum_n y_n$ be convergent series. Then so is the series $\sum_n [x_n + y_n]$.*

Proof See [4]. ■

Fix x_0, y_0 in the R_0 ball around the origin. Then, each of the series

$$\sum_{n=i}^{\infty} q_n x_0^{n-2} y_0^2, \quad \sum_{n=i}^{\infty} p_n x_0^{n-1} y_0, \quad \sum_{n=i}^{\infty} r_n y_0^{n-2} x_0^2, \quad \sum_{n=i}^{\infty} s_n y_0^{n-1} x_0$$

is convergent, and hence, by the above lemma, so is the series,

$$\sum_{n=i}^{\infty}[p_n x_0^{n-1} y_0 + q_n x_0^{n-2} y_0^2 + r_n y_0^{n-2} x_0^2 + s_n y_0^{n-1} x_0]$$

This holds for arbitrary x_0, y_0 in the neighborhood of radius R_0 around the origin, and hence, V is convergent in a neighborhood of radius R_0 around the origin. Hence, the example is complete. ■

Next, another example is considered.

Example 3.2 *Consider the system*

$$\begin{aligned} \dot{x} &= -y + a_1 x^2 + 2h_1 xy + b_1 y^2 \\ \dot{y} &= x + a_2 x^2 + 2h_2 xy + b_2 y^2 \end{aligned}$$

Assume that for this system, with the notation of the Theorem 3.2, $p_x = p_y = 4$ (that is, the coefficient of the $x^4 y^4$ monomial in $\dot{V}_8$ is non zero). Then, the system is asymptotically stable or unstable, respectively as the coefficient of $x^4 y^4$ monomial in $\dot{V}_8$ is negative or positive.

A method similar to the one used in the previous example is used.

Assume, that V_i has been formed, $i > 7$. Then, it can be seen (from the construction step 3c in the above algorithm), that

$$\begin{aligned} V_{i+1} &= V_i + ax^i y + bx^{i-1}y^2 + cx^{i-2}y^3 + dx^{i-3}y^4 \\ &\quad + ey^i x + fy^{i-1}x^2 + gy^{i-2}x^3 + hy^{i-3}x^4 \end{aligned}$$

for some a, b, c, d, e, f, g, h. Let M be as before. Let, for some i

$$\begin{aligned} Hom_i(V_i) &= \\ & p_i x^{i-1}y + q_i x^{i-2}y^2 + r_i x^{i-3}y^3 + s_i x^{i-4}y^4 \\ & +(\text{terms for } xy^{i-1}, x^2y^{i-2}, x^3y^{i-3}, x^4y^{i-4}) \end{aligned}$$

Aim is to derive upper bounds for $|p_i|, |q_i|, |r_i|, |s_i|$. By using an approach similar to the previous example, it can be proved that

$$\begin{aligned} p_{i+1} &\leq M|p_i| \\ q_{i+1} &\leq M\left[2i|p_i| + |q_i|\right] \\ r_{i+1} &\leq M\left[2i|p_i| + 2i|q_i| + |r_i|\right] \\ s_{i+1} &\leq M\left[2i^2|p_i| + 4i|q_i| + 3i|r_i| + |s_i|\right] \end{aligned}$$

From this, again, by using a method of proof similar to the last example, it can be proved that V is convergent. Basically, it can be proved that if the above inequalities are replaced by equalities, then each of p_n, q_n, r_n, s_n is $k^n \times$ (polynomial in n), for some k. The convergence of this series can be proved by using the ratio test (theorem 2.3) and method similar to that used in the previous example. Hence, the example is complete. ■

The examples considered until now had no monomials of degree greater than 2 in the RHS of the equation (1). This simplified the proof because taking the total derivative of $\dot{V}_i$ yielded terms of degree less than or equal to $(i + 1)$.

A more complicated example is considered next.

Example 3.3 *Consider the system*

$$\begin{aligned} \dot{x} &= -y + f_1(x, y) \\ \dot{y} &= x + f_2(x, y) \end{aligned}$$

where f_1 and f_2 are polynomials with monomials of total degree 2 or 3. Assume that for this system, with the notation of the above theorem, $p_x = p_y = 2$ (that is, the coefficient of the x^2y^2 monomial in $\dot{V}_4$ is non zero). Then, the system is asymptotically stable or unstable, respectively as the coefficient of x^2y^2 monomial in $\dot{V}_4$ is negative or positive.

Given V_i, $V_{i+1} = V_i + rx^ky + sx^{k-1}y^2 + ty^kx + uy^{k-1}x^2$. This is because of reasons similar to Example 3.1. But now, both V_{i-1} and V_i will determine V_{i+1}(in the last two examples, only V_i determined V_{i+1}). Let $M > \max(1,$ coefficients of monomials appearing in f_1 and $f_2)$.

Let, for a particular i,

$$\begin{aligned} Hom_i(V_i) &= p_ix^{i-1}y + q_ix^{i-2}y^2 \\ &\quad +r_iy^{i-1}x + s_iy^{i-2}x^2 \\ Hom_{i+1}(V_{i+1}) &= p_{i+1}x^iy + q_{i+1}x^{i-1}y^2 \\ &\quad +r_{i+1}y^ix + s_{i+1}y^{i-1}x^2 \end{aligned}$$

Now, upper bounds will be derived on p_{i+2}, q_{i+2} in terms of $p_i, q_i, p_{i+1}, q_{i+1}$. Note that there is a dependency on two previous levels because of the presence of third order monomials in f_1 and f_2. In general, if f_1 and f_2 have a monomial of maximum degree d, there will be a dependency on $(d-1)$ previous levels. By using the formulae of step 3c from the algorithm, and by using arguments similar to Example 3.1, it can be easily checked that

$$\begin{aligned} p_{i+2} &< M[p_i + p_{i+1}] \\ q_{i+2} &< M\Big[\frac{ip_i + (i+1)p_{i+1} + 2q_i + 2q_{i+1}}{2}\Big] \end{aligned}$$

Consider the equations

$$\begin{aligned} p'_{i+2} &= M[p'_i + p'_{i+1}] \\ q_{i+2} &= M\Big[\frac{ip'_i + (i+1)p'_{i+1} + 2q'_i + 2q'_{i+1}}{2}\Big] \end{aligned} \tag{3.14}$$

Assume that for initial conditions, p'_i, p'_{i+1}, q'_i, q'_{i+1} are moduli of p_i, p_{i+1}, q_i and q_{i+1} respectively. Then, it is easy to see that $p'_i > |p_i|$, $p'_{i+1} > |p_{i+1}|$, $q'_i > |q_i|$, $q'_{i+1} > |q_{i+1}|$. Also, since $M \geq 1$, it follows that $p'_{i+1} > p'_i$ and $q'_{i+1} > q'_i$. Thus, it follows from equations (3.14) that

$$\begin{gathered} p'_{i+2} < 2Mp'_{i+1} \\ q'_{i+2} < M\Big[\frac{2i+1}{2}p'_{i+1} + 2q'_{i+1}\Big] \end{gathered} \tag{3.15}$$

These recurrances are similar to Example 3.1 and by using methods similar to Example 3.1, it can be proved that the series $\sum p'_i z^i$ and $q'_i z^i$ are convergent, and hence, again, by arguments similar to Example 3.1, it can be proved that V is convergent.

Hence, the example is complete. ■

These examples illustrate the results in Theorem 3.2. The basic idea is the following. Suppose that f_1 and f_2 (in the notation of Equation (1)) have a maximum monomial degree of d. Also, assume that the $x^{2k}y^{2k}$ is the least degree monomial in $\dot{V}$ which has non zero coefficient (argument will be similar in case when the least degree monomial with non-zero coefficient is $x^{2l}y^{2l+2}$ for some l). Now, in Example 3.3, there was a dependence of p_{i+2} on p_i and p_{i+1}, and similarly for q_{i+2}, r_{i+2} and s_{i+2}. This was because the maximum monomial degree in f_1 and f_2 was 3. In general, when the maximum monomial degree d, it is clear that there will be a dependence on the previous $(d-1)$ levels. Also, this dependence on the last $(d-1)$ levels can be reduced to a dependence on just the previous level as was done in Example 3.3 in going from Equation (3.14) to Equation (3.15).

Also, the only monomials that were introduced in Example 3.3 while forming V_{i+1} from V_i were x^iy, $x^{i-1}y^2$, xy^i, x^2y^{i-1} with coefficients p_{i+1}, q_{i+1}, r_{i+1}, s_{i+1} respectively. These were added to cancel the x^{i+1}, x^iy, y^{i+1}, y^ix terms respectively in $\dot{V}_i$. The reason that only these terms needed to be cancelled in the above example was because these were the only monomials **not** less than x^2y^2. In case that $x^{2k}y^{2k}$ monomial in $\dot{V}_{4k}$ is non-zero (note that $i > 4k$; for $i < 4k$, Step 1 of the Algorithm was used to construct V_is) , the monomials in $\dot{V}_i$ that will be needed to be cancelled will be x^{i+1}, x^iy, $\ldots$, $x^{i-2k+2}y^{2k-1}$; y^{i+1}, y^ix, $\ldots y^{i-2k+2}x^{2k-1}$, and hence, the monomials in $\dot{V}_i$ that will be needed to be introduced will be $x^{i+1-j}y^j$, $1 \le j \le 2k$ and $y^{i+1-j}x^j$, $1 \le j \le 2k$. Thus,

$$V_{i+1} \quad = \quad V_i + \sum_{j=1}^{2k} p^j_{i+1} x^{i+1-j} y^j + \sum_{j=1}^{2k} s^j_{i+1} y^{i+1-j} x^j$$

Then, by arguments similar to the three examples considered above, the recurrance relations in p_r^j will be of the form,

$$\begin{aligned}
p_{r+1}^1 &< c_1 p_r^1 \\
p_{r+1}^2 &< \alpha_{2,1}(r)p_r^1 + c_2 p_r^2 \\
&\vdots \\
p_{r+1}^{j+1} &< \alpha_{j+1,1}(r)p_r^1 + \ldots + \alpha_{j+1,j}(r)p_r^j + c_{j+1}p_r^{j+1} \\
&\vdots \\
p_{r+1}^{2k} &< \alpha_{2k,1}(r)p_r^1 + \ldots + \alpha_{2k,2k-1}(r)p_r^{2k-1} + c_{2k}p_r^{2k}
\end{aligned}$$

where c_is are constants and $\alpha_{i,j}(r)s$ are polynomials in r. The important thing is that c_is are constants and not polynomials (this follows from the algorithm step 3c — the way the functions V_is are constructed succesively). Note that superscripts **do not** denote powers. From these, it easily follows that

$$p_r^j < l_j^r \beta_j(r)$$

where l_rs are constants (the superscript r on l_j denotes power!) and $\beta_j(r)$s are polynomials in r. This can be proved by induction, first proving it for p_r^1, using this to prove it for p_r^2, and so on. The proof is simple, and hence, omitted.

Now, by the ratio rule, and argument similar to that used in Example 3.1, it can be proved that this will give rise to a convergent series for V.

The proof of the theorem is complete. ■

In Theorem 3.2, it had been assumed that p_x and p_y are finite. That is, after following the procedure of the Algorithm, Step 1, a stage is reached when $x^{2k}y^{2k}$ or $x^{2l}y^{2l+2}$ coefficient in $\dot{V}_n$ (n=$2k$ or $(2l+2)$ respectively) is non-zero. That this would not always be the case follows from the next theorem. Before the theorem, a definition is stated.

Definition 3.5 *Call a system **Marginally Stable** if it is stable, but not asymptotically stable.*

Theorem 3.3 *For a marginally stable system, Step 1 of the Algorithm does not terminate.*

Proof A ■ ssume to the contrary. Thus, there exist finite p_x and p_y. By Theorem 3.2, if the coefficient of the monomial $x^{p_x}y^{p_y}$ in $\dot{V}_{n_0}$ ($n_0 = p_x + p_y$) is positive, the system is unstable, and if it is negative, the system is asymptotically stable. This contradicts the assumption that the system is Marginally Stable, since unstable, asymptotically stable and marginally stable systems form disjoint sets.
An example of a marginally stable system is the Lotka-Volterra system (equations (1.1)) considered in the first section. A more general example is of the system

$$\begin{aligned} \dot{x} &= -y + kxy \\ \dot{y} &= x + lxy \end{aligned}$$

As is easy to check, this system is marginally stable for all k, l.

Assume that Step 1 of the Algorithm does not terminate. This can be stated as a problem of satisfiability of an infinite system of polynomial equations in the coefficients of the monomials of f_1 and f_2 (f_1 and f_2 are as in equation (1)). This follows from the way the successive V_ns are constructed in the proof of Theorem 3.1.

For example, consider the system

$$\begin{aligned} \dot{x} &= -y + a_1x^2 + h_1xy + b_1y^2 \\ \dot{y} &= x + a_2x^2 + h_2xy + b_2y^2 \end{aligned} \tag{3.16}$$

For this system, $Hom_4(\dot{V}_4) = 0$, that is, the coefficient of x^2y^2 in $\dot{V}_4$ is equal to zero translates to

$$a_1h_1 + h_1b_1 - 2a_1a_2 - a_2h_2 + 2b_1b_2 - h_2b_2 = 0$$

This can be checked by direct but a rather lengthy calculation. For the system (3.16), it is easy to see that $Hom_{4k}(\dot{V}_{4k}) = 0$ results in a $4k$-form in the variables $a_1, h_1, b_1, a_2, h_2, b_2$ being zero and $Hom_{4l+2}(\dot{V}_{4l+2}) = 0$ results in a $4l+2$-form being zero. In general, for the system (1), it is easy to see that for all k, l, $Hom_{4k}(\dot{V}_{4k}) = 0$ or $Hom_{4l+2}(\dot{V}_{4l+2}) = 0$, each results in a polynomial (in the coefficients of the monomials of f_1 and f_2) being equal to zero. Therefore, the non-termination of the procedure of Algorithm Step 1 is equivalent to the satisfiability of an infinite set of polynomial equations. The variables in these polynomial equations are the coefficients of the monomials of f_1 and f_2.

This infinite system of equations can be converted into a finite system by use of the theorem 2.5 In other words, an infinite system of polynomial equalities is equivalent to a finite system of polynomial equalities, and that, the solution is equivalent to the simultaneous solution of the first k equations for some k.

For the system (3.16), simulations with Macaulay-2 (see [6]) suggest that k (in the notation of Theorem 2.5) is equal to 3. This means that the coefficients of x^2y^2, x^2y^4, x^4y^4 in $\dot{V}_4$, $\dot{V}_6$, $\dot{V}_8$, respectively, are zero implies $Hom_{4k}(\dot{V}_{4k}) = 0$ and $Hom_{4l+2}(\dot{V}_{4l+2}) = 0$ for all k and l. This has **not** been proved rigorously and efforts are underway to do the same.

Consider the system (3.16). For this set of systems, the space of coefficients $(a_1, h_1, b_1, a_2, h_2, b_2)$ is 6-dimensional.

Theorem 3.4 *The dimension of the set where all the polynomial equations are satisfied is 4.*

Proof C ■ onsider the system

$$\begin{aligned} \dot{x} &= -y + a_1x^2 - 2b_2xy + b_1y^2 \\ \dot{y} &= x + a_2x^2 - 2a_1xy + b_2y^2 \end{aligned} \tag{3.17}$$

Clearly, this set of systems is a subset of dimension 4 in the original set of systems (3.16). For this system, it is easy to check that if

$$V = \frac{x^2+y^2}{2} + \frac{1}{3}a_2x^3 - a_1x^2y + b_2xy^2 - \frac{1}{3}b_1y^3$$

then

$$\dot{V} = 0$$

Note that V is positive definite in a neighborhood of the origin. By Theorem 2.1, this set of systems is marginally stable. By Theorem 3.3, the set of marginally stable systems is a subset of the set where all polynomial equations are satisfied, and hence, the set where all the polynomial equations are satisfied has dimension greater than or equal to 4. Next, it is easy to check using Macaulay-2 that the dimension of the set where only the first 2 polynomial equations are satisfied is also 4. Hence, the set where all the polynomial equations are satisfied has dimension less than or equal to 4. Hence, the set where all polynomial equations are satisfied has dimension precisely equal to 4. Hence proved.

Thus, the set on which the question of stability/instability has not yet been resolved has dimension less than 6 (infact, less than 5). Thus, the problem has been solved **generically**

— for any point in the 6 dimensional space, there is a point arbitrarily close to it where the question of stability/instability has been resolved.

The previous 2 paragraphs considered set of systems (3.16). For the general set of systems (1), it can be proved that the dimension of the set where all polynomial equations are satisfied is less than the dimension of the space equal to that of the number of variables. Note that the variable space is the same as the coefficient space of the monomials of f_1 and f_2. Thus, for the set of systems (1), the problem has been solved generically.

3.2 Conclusions

We have considered the problem of stability/instability of origin for the system

$$\dot{x} = -y + f_1(x, y)$$
$$\dot{y} = +x + f_2(x, y)$$

Note that a similar method can be used for the system

$$\dot{x} = -k_1 y + f_1(x, y)$$
$$\dot{y} = +k_2 x + f_2(x, y)$$

where $k_1, k_2 > 0$. There will be minor changes in the construction of V.

We resolved the question of stability/instability except for systems whose coefficients satisfy a set of polynomial equalities.

For the system

$$\begin{aligned}\dot{x} &= -y + a_1 x^2 + h_1 xy + b_1 y^2 \\ \dot{y} &= x + a_2 x^2 + h_2 xy + b_2 y^2\end{aligned}$$

it seems checking the infinite set of polynomial equations for their satisfiability is the same as checking the first 3 equations for their satisfiability, though we have not been able to prove this. If this is true, we can generate these equations easily using Macaulay-2, and check for their satisfiability. If all the three equations are satisfied, the method presented here works, else, it does not.

For the general set of systems (1), the set of polynomial equations can again be generated using Macaulay-2. The dimension of the simultaneous solution set of these equations and the number of equations that suffice to check whether all the equations are satisfied are interesting but as yet unresolved questions.

3.3 Important Note

After this work was done, it was found that similar work was done in [7] and [8], [9], [10], [11]

PART II

Optimal Scheduling
for
Power Efficient Transmission

Chapter 4

Introduction and Related work

4.1 Introduction

Power efficient communication has been an important design challenge for wireless communication systems. In wireless as well as in wired channels with Additive White Guassian Noise (AWGN), the power increases as a convex function of the transmission rate. The convex nature of the relationship between power and rate allows one to save energy by choosing the rate at various stages in transmission in an appropriate fashion, i.e., to transmit the data in opportunistic chunks. This, however, would lead to an increase in buffered data and hence the average delay. The wireless channel being time varying offers another opportunity of power savings where the scheduler can simply defer the transmission of packets during 'bad' channel states to 'good' channel states.

We provide structural results for an optimal power scheduling problem with delay constraints for a single user transmitting on a link. The setting is the following:

There is a scheduler at which packets arrive and get buffered. The time is slotted and that, in each time slot, packets arrive with a given probability distribution. Since the packets get buffered, each packet suffers a delay. There is a constraint on the *average delay* a packet can suffer. By Little's Theorem[12], the average delay is proportional to the average buffer occupancy, and hence, equivalently, there is a constraint on the average buffer occupancy. The scheduler needs to transmit the packets. The energy required to transmit k packets is a convex function of k. This means that the power requirement is a convex function of the rate at which we want to transmit. We assume fluid packets, and hence, k is any non-negative real number.

A convex energy function has the implication that from an "energy saving" point of view, it is better to transmit packets in small number in different time slots rather than transmit a large number of packets in one time slot. However, transmitting small number of packets in different time slots leads to a higher average packet delay.

The objective is to find a scheduling policy which minimizes the power consumption but conforms to a constraint on the average buffer occupancy.

4.2 Related Work

The problem of energy efficient scheduling has been considered in [13, 14, 15, 16, 17, 18, 19, 20].

In [15], the authors have considered the problem of power-efficient scheduling under average and absolute delay constraints. They have used a slotted arrival system and considered an arbitrary i.i.d. packet arrival process. A characterization of the optimal scheduler has been provided in terms of a smaller class of deterministic schedulers. [15] and [13] do not consider Markov packet arrivals and the presence of fading.

On the other hand, [16] has dealt with energy-efficient scheduling under an average delay constraint for time-varying channels. However, the authors have assumed a linear relationship between power (P) and rate (R), and thus do not take into account the gains that can be realized by varying the transmission rates.

Recently, in [17, 18], the tradeoff between the average delay and the average power in a fading channel has been analyzed. The delay-power tradeoff has also been quantified in the region of asymptotically large delays. More recently, [19] has also considered the problem of energy-efficient scheduling taking into account both the fluctuating channel conditions and the convex $P-R$ relationship. [20] also contains structural results for a policy which minimizes the average delay subject to a constraint on the average power, in presence of channel fading. It is proved in [17, 20] that there exists an optimal stationary policy which increases as the buffer occupancy increases, and decreases, as the level of channel fading goes from good to bad.

The main contribution of this part of the thesis is to provide structural results related to the dependence of number of transmissions in a time slot on the number of packet arrivals in the previous slot. We also provide structural results concerning the dependence of optimal policy on buffer occupancy and channel fade when the channel fade is i.i.d. These results have been proved in [17, 20] too. We also state a negative result in the case of FSD Markov channel fade (see Chapter 5 for definition). This result is hinted at, though not precisely stated in [17].

Chapter 5

Mathematical Preliminaries

In this chapter, we state some results that will be used in Part II of the thesis. We also mention some concepts which will be used, and give references as to where they are elaborated.

We have made extensive use of the Dynamic Programming technique. Dynamic Programming deals with optimization problems where decisions are made in stages, and the objective is to minimize a certain cost. The key point is that in most cases, the desire for a low present cost might lead to high future costs, and hence, stagewise decisions cannot be viewed in isolation. The Dynamic Programming technique captures this tradeoff. The readers are referred to [21] for details of the technique.

Probability

Definition 5.1 (First Order Stochastic Dominance, $\succ$, FSD) *Let p_1 and p_2 be probability measures on $\mathbb{R}$. Let $p_1((-\infty, x)) \leq p_2((-\infty, x))$, $\forall x \in \mathbb{R}$. Then, p_1 is said to be First Order Stochastically Dominant with respect to p_2.*

We will abbreviate First Order Stochastic Dominance as FSD in this thesis. First Order Stochastic Dominance of p_1 with respect to p_2 will be denoted by $p_1 \succ p_2$, or by $p_2 \prec p_1$.

Theorem 5.1 *Let $p_1 \succ p_2$ be probability measures on $\mathbb{R}$. Let $f : \mathbb{R} \to \mathbb{R}$ be an increasing function. Then,*

$$\int_{\mathbb{R}} f(w) p_1(dw) \geq \int_{\mathbb{R}} f(w) p_2(dw) \tag{5.1}$$

Proof A proof can be found in [22]. ■

Theorem 5.2 *Let $p_1 \succ p_2$ be probability measures on $\mathbb{R}$. Let X_1 and X_2 be random variables, $X_1 \sim p_1$, $X_2 \sim p_2$. Then, there exist random variables $\bar{X}_1$ and $\bar{X}_2$ such that $\bar{X}_1 \sim p_1$, $\bar{X}_2 \sim p_2$, and $\bar{X}_1 \geq \bar{X}_2$ almost surely.*

Proof A proof can be found in [22]. ■

Theorem 5.3 (Hoeffding's Inequality) *Let $M =< X_n >_{n \geq 0}$ be a Markov Chain on $[a, b]$ with transition kernel $p(x, A)$. Assume that there exists a positive measure ν on $[a, b]$ and a real number δ such that $p(x, A) - \delta\nu(A) \geq 0$ for all $x \in [a, b]$, $A \in \mathcal{B}([a, b])$. Let $S_n = (X_1 + \cdots + X_n)$. Then, for all x, $n \geq \frac{2(b-a)}{\delta\epsilon}$,*

$$\Pr_x(S_n - ES_n \geq n\epsilon) \leq \exp\left[-\frac{\delta^2 \left\{n\epsilon - \frac{2(b-a)}{\delta}\right\}^2}{2n(b-a)^2}\right] \tag{5.2}$$

Proof A proof can be found in [23] ■ .

Real Analysis

The following theorems from real analysis will be used:

Theorem 5.4 *Let S be a compact subset of $\mathbb{R}$. Let $f : S \to \mathbb{R}$ be a continuous function. Then, f attains a minimum.*

Proof A proof can be found in [4]. ■

Theorem 5.5 (Monotone Convergence Theorem) *Let μ be a measure on $\mathbb{R}$. Let $f_n : \mathbb{R} \to \mathbb{R}$ be an increasing sequence of non-negative measurable functions. Let $f_n \uparrow f$. Then ,*

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu \tag{5.3}$$

Proof A proof can be found in [24]. ■

Theorem 5.6 (Dominated Convergence Theorem) *Let μ be a measure on $\mathbb{R}$. Let $f_n : \mathbb{R} \to \mathbb{R}$ be a sequence of measurable functions converging to a measurable function f. Let $|f_n| \leq g$ for all n where g is μ-integrable. Then, f is integrable and*

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu \tag{5.4}$$

Proof A proof can be found in [24]. ■

Theorem 5.7 (Dini's Theorem) *Let $f_n : \mathbb{R} \to \mathbb{R}$ be an increasing sequence of continuous functions converging to a continuous function f. Then, the convergence is uniform on compact subsets of $\mathbb{R}$.*

Proof A proof can be found in [24]. ■

Theorem 5.8 (Arzela-Ascoli Theorem) *Let $f_n : \mathbb{R} \to \mathbb{R}$ be a family of functions which are equicontinuous on every compact set. Then, there exists a function f and a sequence of functions $< f_{n_k} >$ such that $f_{n_k} \to f$ uniformly on $\mathbb{R}$.*

Proof A proof can be found in [4]. ■

Theorem 5.9 *Let $f_n : \mathbb{R} \to \mathbb{R}$ be a sequence of continuous functions converging uniformly to a continuous function f. Let S be a compact subset of $\mathbb{R}$. Then,*

$$\lim_{n\to\infty} \min_{x\in S} f_n(x) = \min_{x\in S} \lim_{n\to\infty} f_n(x) = \min_{x\in S} f(x) \tag{5.5}$$

Proof A proof can be found in [4]. ■

Theorem 5.10 *Let $f : \mathbb{R} \to \mathbb{R}$ be a decreasing convex function. Then,*

i. f is continuous.

ii. The left and the right derivatives of f exist at each point. Let the left and the right derivatives at the point x be denoted by $D_l(x)$ and $D_r(x)$ respectively.

iii. If $a < b$, $D_l(a) \leq D_r(a) \leq D_l(b) \leq D_r(b)$.

iv. If $a < b$, $-D_l(b)(b-a) \leq f(a) - f(b) \leq -D_l(a)(b-a)$.

Proof A proof can be found in any book which deals with convex functions. ■

Optimization

Lemma 5.1 *Let h be a function of $(n+2)$ variables, viz., scalars y, x, and n-dimensional vector $\mathbf{s}$. Let S be a sublattice of $\mathbb{R}^n$. The sublattice structure defines an order $>$ on S. Let h be defined on the set (x, y, S), where $x \in \mathbb{R}$, and $0 \leq y \leq x$. Let h be a continuous function. Let h satisfy decreasing differences in $(y, x, \mathbf{s})$, that is, for $x' > x$, $\mathbf{s}' > \mathbf{s}$, $y' > y$,*

$$h(y', x', \mathbf{s}') - h(y, x', \mathbf{s}') \leq h(y', x, \mathbf{s}) - h(y, x, \mathbf{s}) \tag{5.6}$$

Let

$$D^*(x, \mathbf{s}) = \operatorname{argmin}\{h(y, x, \mathbf{s}) | y \in [0, x]\}$$

Then,

- *For each $(x, \mathbf{s})$, $D^*(x, \mathbf{s})$ is a non-empty compact sublattice of R, and admits a greatest element denoted by $y^*(x, \mathbf{s})$.*
- *$y^*(x', \mathbf{s}') \geq y^*(x, \mathbf{s})$ whenever $x' \geq x$, $\mathbf{s}' \geq \mathbf{s}$.*

Proof This follows from Theorems 10.7 and 10.12 of [25]. ■

Pseudo Atom Construction

Pseudo Atom Construction is a method of *splitting* a Markov Chain so that its essential properties are retained but a set called an atom gets introduced. Pseudo atom construction is discussed in detail in Chapter 9. A good reference is [26].

Chapter 6

Model, Problem Formulation, and Approach

6.1 System Model

As stated in Section 4.1, the system consists of a single transmitter (source) and a single receiver (sink) with a point to point link. The other assumptions of the model are given below.

1. A slotted time system is considered. A slot is assumed to be of unit length. The slot $[n, n+1)$ $(n \geq 0)$ is numbered as slot n. Packets are assumed to arrive just before the end of the slot. Packets are transmitted just after the beginning of a slot. Thus, in slot n, packets arrive and are transmitted, respectively, at times $(n+1)^-$ and n^+ respectively. These are labelled as W_{n+1} and Y_n respectively (see Figure 6.1). W_n and Y_n are random variables. A particular realization of W_n and Y_n will be denoted by w_n and y_n respectively. The arrival and transmission process are assumed to be fluid. Accordingly, w_n and y_n are real numbers.

2. The buffer is assumed to be of infinite size, though the results will hold for a finite-sized buffer too. Let X_n denote the buffer occupancy or queue length at time n, that is, at the beginning of slot n. The buffer occupancy evolves as:
$$X_{n+1} = X_n - Y_n + W_{n+1} \tag{6.1}$$

3. The channel is assumed to be time-varying with fading. The channel fade is assumed to remain constant over a slot duration (slow fading). The state of the channel (called

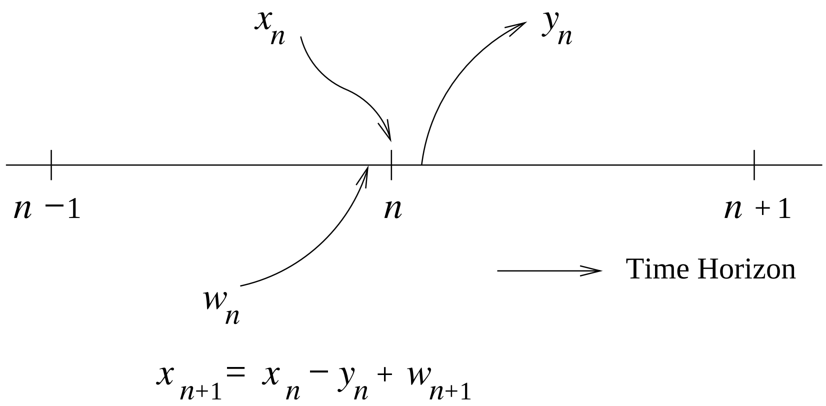

Figure 6.1: The System Model

level of channel fade) in slot n will be denoted by the random variable Ψ_n. A particular realization of Ψ_n will be denoted by μ_n. The scheduler is assumed to sense the state of the channel at the beginning of a slot, and adapt the transmission accordingly.

The state of the above system can be completely characterized by the 3-tuple, $\mathbf{v_n} = (x_n, w_n, \mu_n)$, comprising the buffer occupancy or queue length, the number of packet arrivals in the previous time slot and the *level* of channel fade. At the beginning of slot 0, we assume that the buffer occupancy (the initial buffer occupancy) is x_0. For mathematical convenience, we assume that w_0 is the number of packets that arrived just before slot 0.

In slot n, the control (or scheduling) action corresponds to the number of packets transmitted, y_n. The control policy is a sequence of functions $\{\pi_1, \pi_2, \ldots\}$ where π_n specifies the conditional law of y_n given the past history of the system state $\mathbf{v_n}$.

We further assume a stationary packet arrival process and a stationary channel state evolution process.

- Both i.i.d. and Markov arrivals are considered in this paper; the probability measure for the arrival process (conditional probability measure given a arrivals in the previous slot respectively) will be denoted by $p(dw)$ ($p(a, dw)$ respectively).
- We consider two models of channel fading - i.i.d. and Markov. The probability measure for the channel fade (conditional probability measure given channel fade ν in the previous slot respectively) will be denoted by $q(d\mu)$ ($q(\nu, d\mu)$ respectively).
- When the *level* of channel fade is μ, the energy required to transmit y packets is μ times a convex function, E, of y. There is a real number y_{max}, which is the maximum

number of packets that can be transmitted in a slot. We assume that $y_{max} < \infty$. We also assume that y_{max} is strictly greater that the expected number of packet arrivals in a slot.

6.2 Problem Formulation

By Little's theorem, the limit superior of the average packet delay, D, is related to the limit superior of the average queue length, Q as:

$$D = \frac{Q}{\bar{w}} \tag{6.2}$$

Here $\bar{w}$ denotes the average arrival rate. In our problem, Q can be approximated as:

$$Q = \limsup_{M \to \infty} \frac{1}{M} \Big[\sum_{n=1}^{M} x_n \Big] \tag{6.3}$$

Actually, the average queue length will be less than this because the above formula assumes that the packet transmission takes place towards the end of a slot. We make the assumption that even the packets that are transmitted in a slot suffer a delay equal to slot length - with this assumption, the above formula becomes exact. Thus,

$$D = \frac{1}{\bar{w}} \limsup_{M \to \infty} \frac{1}{M} \Big[\sum_{n=1}^{M} x_n \Big] \tag{6.4}$$

In the rest of the paper, we ignore the proportionality constant $\bar{w}$ and treat average delay D as synonymous with average queue length Q.

Thus, we define:

Definition 6.1 (Expected Average Delay, D)

$$D = \limsup_{M \to \infty} \frac{1}{M} E \Big[\sum_{n=1}^{M} x_n \Big] \tag{6.5}$$

In slot n, we transmit y_n packets. Thus, we define:

Definition 6.2 (Expected Power, P)

$$P = \limsup_{M \to \infty} \frac{1}{M} E \Big[\sum_{n=1}^{M} \mu_n E(y_n) \Big] \tag{6.6}$$

We would want to expend as little energy as possible. But expending less energy corresponds to transmitting lesser number of packets, and hence, higher delay.

Thus, we have *The Average Cost Optimal Scheduling Problem* that can be stated as:

Given the observation of the state of the system up to the present time, determine a scheduling policy that minimizes P subject to a constraint on D.

This is the problem that we want to address.

6.3 Approach

This is a constrained Markov decision process problem. To solve this problem, we first need to solve a corresponding unconstrained Markov decision process problem. Let λ be a positive real number. Let $C_\lambda = P + \lambda D$, where, P and D are defined in the previous section. *The Unconstrained Average Cost Optimal Scheduling Problem* can be stated as:

Given the observation of the state of the system up to the present time, determine a scheduling policy that minimizes C_λ.

To solve this problem, we first need to solve the corresponding infinite horizon discounted cost problem. Let $0 < \beta < 1$ be a real number. Define:

Definition 6.3 (Infinite Horizon Expected Discounted Delay)

$$D_\infty = E\Big[\sum_{n=1}^{\infty} \beta^n x_n\Big] \tag{6.7}$$

Definition 6.4 (Infinite Horizon Expected Discounted Energy)

$$P_\infty = E\Big[\sum_{n=1}^{M} \beta^n \mu_n E(y_n)\Big] \tag{6.8}$$

Note that if $\beta = 1$, both the above expressions will become ∞. Define $C_{\infty,\lambda} = P_\infty + \lambda D_\infty$.

The Infinite Horizon Unconstrained Discounted Cost Optimal Scheduling Problem can be stated as:

Given the observation of the state of the system up to the present time, determine a scheduling policy that minimizes $C_{\infty,\lambda}$.

To solve this problem, we first need to solve the corresponding finite horizon problem. Let N be a non-negative integer. For this N-step problem, P_N and D_N are defined exactly as in the infinite horizon problem, except that the summation extends from 0 to N instead of 0 to ∞. The corresponding cost function is denoted by $C_{N,\lambda}$. Thus, given the observation of the state of the system up to the present time, the intention is to determine a scheduling policy that minimizes $C_{N,\lambda}$.

We thus summarize the above procedure below:

1. Solve the unconstrained finite horizon problem for all N.
2. Solve the unconstrained infinite horizon problem by taking limit as $N \to \infty$. We have to justify the limiting procedure.
3. Solve the unconstrained average cost problem by taking limit as $\beta \to 1$. We have to justify the limiting procedure.
4. Solve the constrained average cost problem.

These problems will be solved in the chapters that follow.

Chapter 7

The Finite Horizon Unconstrained Problem

In this chapter, we derive some properties of the optimal cost functions for the finite horizon problem. These properties will be useful in the next chapter, where we will derive structural results about the optimal policies for the infinite horizon problem.

Let $V_n^\beta(\mathbf{v})$ denote the n-step optimal cost when the initial state is $\mathbf{v}$ and the discount factor is β. In the rest of this chapter, β will be a constant, and we do not explicitly show the superscript β. Thus, $V_n^\beta(\mathbf{v})$ will be denoted as $V_n(\mathbf{v})$. Let $J_{n,\pi}^\beta(\mathbf{v})$ denote the n-step cost when the initial state is $\mathbf{v}$ and the policy used is π. As before, we supress the superscript β in this chapter.

7.1 i.i.d. Arrivals and No Channel Fade

In this section, we assume the channel to be constant and packet arrivals to be i.i.d. As a result, the state of the system, $\mathbf{v} = x$, the buffer occupancy. $V_0(x)$, the zero step cost is equal to λx. Note that we do not apriori have a guarantee that V_ns exist. Define:

$$g_n(x) = \min_{y \le x}[\lambda x + F(y) + \beta \int p(dw) g_{n-1}(x - y + w)] \tag{7.1}$$

This is the Bellman's equation.

Define

$$h_n(y, x) = \lambda x + F(y) + \beta \int p(dw) g_{n-1}(x - y + w) \tag{7.2}$$

Note that $h_n(y,x) \geq g_n(x)$, assuming that $h_n(x)$ and $g_n(x)$ exist.

Lemma 7.1 *Let $g_0(x) \triangleq V_0(x) = \lambda x$. If $g_i(x)$, $1 \leq i \leq n$ satisfying (7.1) exist, $V_i(x)$, $1 \leq i \leq n$ exist, and $V_i(x) = g_i(x)$, $0 \leq i \leq n$.*

Proof A proof can be found in [21]. ■

Lemma 7.2 *If $V_n(x)$ exists, as $x \uparrow$, $V_n(x) \uparrow$, and $V_n(x)$ is a continuous function of x.*

Proof Let $x < x'$. On a probability space $(\Omega, \mathcal{F}, P)$, we have the arrival process sequence $< W_k >$, where each W_k is independent of the other and identically distributed with law $p(dw)$. Let $< Y'_k >$ be optimal (optimal in the sense that the expected discounted n-step cost is minimized) for $< X'_k >$ given by

$$X'_{k+1} = X'_k - Y'_k + W_{k+1}, X'_0 = x', 0 \leq k \leq n \tag{7.3}$$

Thus, for each sample path $< w_k >$ of the packet arrival process, there is a corresponding sample path $< y'_k >$ of the packet transmission process. For each k, y'_k depends only on $x'_i, 0 \leq i \leq k, w_i, 1 \leq i \leq k, y'_i, 0 \leq i \leq k-1$. Construct inductively

$$X_{k+1} = X_k - Y_k + W_{k+1}, X_0 = x, 0 \leq k \leq n \tag{7.4}$$

where Y_k is the process $\min(X_k, Y'_k)$. It is easy to check that for each sample path $< w_k >$ of the arrival process, $y_k = \min(x_k, y'_k) \leq y'_k$, and $x_k < x'_k$. Note that $< Y_k >$ need not be optimal for the process $< X_k >$. Thus,

$$\begin{aligned} V_n(x) \leq J_{n,<Y_k>}(x) = E\left(\sum_{i=0}^{n} \beta^i [\lambda X_i + F(Y_i)]\right) \\ < E\left(\sum_{i=0}^{n} \beta^i [\lambda X'_i + F(Y'_i)]\right) = V_n(x') \end{aligned} \tag{7.5}$$

Note that $Y_n = Y'_n = 0$ since we are considering the n-step problem, but X_n and X'_n might be non-zero. Hence, the assertion regarding increasing nature of V_n is proved.

Next, we prove the continuity of V_n. Consider the same process $< W_k >$. Let $< Y_k >$ be optimal for $< X_k >$ given by

$$X_{k+1} = X_k - Y_k + W_{k+1}, X_0 = x, 0 \leq k \leq n \tag{7.6}$$

Consider the process

$$X'_{k+1} = X'_k - Y_k + W_{k+1}, X_0 = x', 0 \le k \le n \tag{7.7}$$

It is easy to check that for each sample path $< w_k >$ of the arrival process, $(x'_k - x_k) = (x' - x)$. Also, the transmission process $< Y_k >$ is the same in both cases. By using the fact that $V_n(x') > V_n(x)$, it is easy to see that

$$V_n(x') - V_n(x) \le J_{n,<Y_k>}(x') - V_n(x) = \sum_{i=0}^{n} \beta^i \lambda(x' - x) = \frac{\lambda(1-\beta^{n+1})}{1-\beta}(x' - x) \tag{7.8}$$

This proves the continuity of V_n. ■

Lemma 7.3 *With $g_0(x) = V_0(x) = \lambda x$, g_n exists, and is continuous, for all n.*

Proof g_0 satisfies the assertions of the lemma. Assume that g_i, $1 \le i \le (n-1)$, satisfy the assertions of the lemma. Since g_{n-1} is continuous, $h_n(y, x)$ is continuous. Since the set $[0, x]$ is compact, $\min_{y \le x} h_n(y, x)$ exists for all x. Hence, g_n exists. By Lemma 7.1, V_n exists, and is equal to g_n. By Lemma 7.2, V_n is continuous, and hence, g_n is continuous. Thus, g_0 satisfies the assertions of the lemma, and with the assumption that g_i, $1 \le i \le (n-1)$ satisfy the assertions of the lemma, we have proved that g_n satisfies the assertions of the lemma. By induction, the lemma is proved. ■

Theorem 7.1 *V_ns exist. V_ns satisfy (7.1). As $x \uparrow$, $V_n(x) \uparrow$, for all n. V_n is continuous for all n. V_ns satisfy:*

$$|V_n(x') - V_n(x)| \le \frac{\lambda}{1-\beta}(x' - x) \tag{7.9}$$

Proof The proof follows by Lemmas 7.1, 7.2, 7.3. ■

Theorem 7.2 *V_n is a convex function of x, for all n.*

Proof $V_0(x) = \lambda x$ is a convex function. Assume that $V_{n-1}(x)$ is a convex function. Consider the Bellman's equation,

$$V_n(x) = \min_{y \le x}[\lambda x + F(y) + \beta \textstyle\int p(dw) V_{n-1}(x - y + w)] \tag{7.10}$$

Note that by Theorem 7.1, V_ns satisfy (7.1). Let, corresponding to $x = x_1, x_2$, minimum in (7.10) be attained, respectively, at $y = y_1, y_2$.

Thus,

$$
\begin{aligned}
V_n(x_1) &= \lambda x_1 + F(y_1) + \beta \textstyle\int p(dw) V_{n-1}(x_1 - y_1 + w) \\
V_n(x_2) &= \lambda x_2 + F(y_2) + \beta \textstyle\int p(dw) V_{n-1}(x_2 - y_2 + w) \\
V_n(x_1) + V_n(x_2) &= \lambda(x_1 + x_2) + F(y_1) + F(y_2) \\
&\qquad + \beta \textstyle\int p(dw)[V_{n-1}(x_1 - y_1 + w) + V_{n-1}(x_2 - y_2 + w)] \\
&\geq^* 2\Big[\lambda\Big(\frac{x_1 + x_2}{2}\Big) + F\Big(\frac{y_1 + y_2}{2}\Big) \\
&\qquad + \beta \int p(dw) V_{n-1}\Big(\frac{x_1 + x_2}{2} - \frac{y_1 + y_2}{2} + w\Big)\Big] \\
&= 2h_n\Big(\frac{y_1 + y_2}{2}, \frac{x_1 + x_2}{2}\Big) \\
&\geq 2V_n\Big(\frac{x_1 + x_2}{2}\Big)
\end{aligned}
$$

(*) follows by the convexity of V_{n-1} and F. This proves the convexity of V_n assuming the convexity of V_{n-1}. By induction, V_n is convex for all n. Hence, proved. ■

We do not state any structural results about optimal policies in this section since we are mainly interested in the infinite horizon discounted cost problem and the average cost problem, for which such results are proved later. It is, however, possible to provide such results for the finite horizon problem as well. We will use the convexity of V_n crucially, when we consider the infinite horizon discounted cost problem.

7.2 Markov Arrivals, No Channel Fade

In this case, the state of the system, $\mathbf{v} = (x, a)$, where x is the buffer occupancy, and a is the number of arrivals in the previous slot. $V_0(x, a)$, the optimal zero step cost is equal to λx. We do not know apriori that V_ns exist. Define:

$$g_n(x, a) = \min_{y \leq x}[\lambda x + F(y) + \beta \textstyle\int p(a, dw) g_{n-1}(x - y + w, w)] \tag{7.11}$$

This is the Bellman's equation.

Define

$$h_n(y, x, a) = \lambda x + F(y) + \beta \textstyle\int p(a, dw) g_{n-1}(x - y + w, w) \tag{7.12}$$

Note that $h_n(y, x, a) \geq g_n(x, a)$.

We make the following assumptions for $p(a, dw)$.

1. $p(a, dw)$ is first order stochastically dominant (FSD) with respect to $p(b, dw)$, whenever $a > b$. This will be denoted by $p(a, dw) \succ p(b, dw)$, or equivalently, by $p(b, dw) \prec p(a, dw)$.

2. $p(x, dw)$ is continuous in x. This means that $\int |p(x, dw) - p(x', dw)| \to 0$ as $x' \to x$.

Note 7.1 *$\int |p(x, dw) - p(x', dw)|$ is the total variation norm.*

Lemma 7.4 *If V_n exists, for all $x' \geq x, a' \geq a$, $V_n(x', a') \geq V_n(x, a)$. Also, $V_n(x, a)$ is a continuous function of x, a.*

Proof We first prove that $V_n(x', a') \geq V_n(x, a)$. To this end, we first prove that $V_n(x, a') \geq V_n(x, a)$. Let $< W_k >$ and $< W_k' >$ denote arrival processes corresponding to a and a' arrivals before slot zero. Thus, $W_0 = a$ and $W_0' = a'$. To each sample path $< w_k >$ of $< W_k >$, we associate a sample path $< w_k' >$ of $< W_k' >$ as follows:

$w_0' = a' \geq a = w_0$. Assume inductively that $w_k' \geq w_k$. Now, $w_{k+1}' \sim p(w_k', dw) \succ p(w_k, dw) \sim w_{k+1}$. By Theorem 5.2, we can have an association wherein $w_{k+1}' \geq w_{k+1}$ *a.s.*

Thus, we have arrival sequences $< W_k >$ and $< W_k' >$ such that $W_0 = a$, $W_0' = a'$, and $W_k' \geq W_k$ *a.s.*

Let $< Y_k' >$ be optimal for $< X_k' >$, defined inductively by

$$X_{k+1}' = X_k' - Y_k' + W_{k+1}', X_0' = x, W_0' = a', 0 \leq k \leq n \tag{7.13}$$

Construct, inductively

$$X_{k+1} = X_k - Y_k + W_{k+1}, X_0 = x, W_0 = a, 0 \leq k \leq n \tag{7.14}$$

where $< Y_k >= \min(X_k, Y_k')$.

As in Lemma 7.2, for each sample path $< w_k >$ of $< W_k >$ and the corresponding sample path $< w_k' >$ of $< W_k' >$, $y_k \leq y_k'$ and $x_k \leq x_k'$.

It follows that

$$V_n(x,a) \leq J_{n,<Y_k>}(x,a) = E\left(\sum_{i=0}^{n} \beta^i[\lambda X_i + F(Y_i)]\right)$$
$$\leq E\left(\sum_{i=0}^{n} \beta^i[\lambda X_i' + F(Y_i')]\right) = V_n(x,a') \quad (7.15)$$

The proof that $V_n(x',a') \geq V_n(x,a')$ is exactly the same as the proof in Lemma 7.2. Thus, $V_n(x',a') \geq V_n(x,a') \geq V_n(x,a)$. Hence proved.

We next prove the continuity of V_n.

Consider the same processes $< W_k >, < W_k' >$. Let $< Y_k >$ be optimal for $< X_k >$ given inductively by

$$X_{k+1} = X_k - Y_k + W_{k+1}, X_0 = x, W_0 = a, 0 \leq k \leq n \quad (7.16)$$

Consider the process

$$X_{k+1}' = X_k' - Y_k + W_{k+1}', X_0' = x', W_0' = a', 0 \leq k \leq n \quad (7.17)$$

It is easy to check that for each sample path $< w_k >$ of $< W_k >$, and the corresponding sample path $< w_k' >$ of $< W_k' >$,

$$(x_k' - x_k) = (x' - x) + \sum_{i=1}^{k}(w_i' - w_i) \quad (7.18)$$

Now,

$$\begin{aligned} E(w_k' - w_k) &= \int [wp^{(k)}(a',dw) - wp^{(k)}(a,dw)] \\ &= \int\int wp^{(k-1)}(y,dw)[p(a',dy) - p(a,dy)] \\ &\leq \int\int wp^{(k-1)}(y,dw)|p(a',dy) - p(a,dy)| \\ &= \int |p(a',dy) - p(a,dy)| \int wp^{(k-1)}(y,dw) \\ &\leq \int w_{max}|p(a',dy) - p(a,dy)| \\ &= w_{max} \int |p(a',dy) - p(a,dy)| \end{aligned} \quad (7.19)$$

Recall that w_{max} is the maximum number of packet transmissions in a slot.

Coming back, note that the transmission process $< Y_k >$ is the same in both cases. It is easy to see by use of (7.18) and (7.19) that

$$V_n(x',a') - V_n(x,a) \leq J_{n,<Y_k>}(x',a') - V_n(x,a)$$
$$\leq \frac{\lambda(x'-x)}{1-\beta} + \frac{w_{max}\lambda\beta}{(1-\beta)^2} \int |p(a',dy) - p(a,dy)| \quad (7.20)$$

The continuity of V_n follows by use of the Assumption 2 about $p(a, dw)$ stated at the beginning of the section. ■

Theorem 7.3 *V_ns exist. V_ns satisfy (7.11). For $x' \geq x$, $a' \geq a$, $V_n(x', a') \geq V_n(x, a)$ for all n. V_n is continuous for all n. V_ns satisfy:*

$$V_n(x', a') - V_n(x, a) \leq \frac{\lambda(x' - x)}{1 - \beta} + \frac{w_{max}\lambda\beta}{(1-\beta)^2} \int |p(a', dy) - p(a, dy)| \tag{7.21}$$

Proof There exist trivial extensions of Lemma 7.1 and Lemma 7.3 for the case of Markov arrivals. Using these extensions and Lemma 7.4, the theorem follows. ■

We now prove some more properties of V_n.

Theorem 7.4 *$V_n(x, a)$ is convex in x for each a, for all n.*

Proof The proof is essentially the same as the proof of Theorem 7.2. ■

Lemma 7.5 *Let $f(x, a)$ be a function which is convex in x for each a, and satisfies*

$$f(x + \Delta x, a + \Delta a) + f(x, a) \geq f(x, a + \Delta a) + f(x + \Delta x, a), \tag{7.22}$$

where Δx and Δa are arbitrary non-negative real numbers. Then,

$$f(x + \Delta x + w, w) - f(x + w, w)$$

increases with w.

Proof Let $\Delta w > 0$. Consider

$$\begin{aligned}
&\Big[f(x + \Delta x + w + \Delta w, w + \Delta w) - f(x + w + \Delta w, w + \Delta w)\Big] \\
&\qquad\qquad\qquad\qquad - \Big[f(x + \Delta x + w, w) - f(x + w, w)\Big] \\
= &\Big[f(x + \Delta x + w + \Delta w, w + \Delta w) - f(x + \Delta x + w, w + \Delta w) \\
&\qquad - f(x + w + \Delta w, w + \Delta w) + f(x + w, w + \Delta w)\Big] \\
&\qquad\qquad + \Big[f(x + \Delta x + w, w + \Delta w) - f(x + w, w + \Delta w) \\
&\qquad\qquad\qquad - f(x + \Delta x + w, w) + f(x + w, w)\Big] \\
\geq^* &\ 0 + 0 \\
= &\ 0,
\end{aligned}$$

where (*) follows by the hypothesis of the lemma. Hence proved. ■

Theorem 7.5 *Let Δx and Δa be arbitrary non-negative real numbers. Then, for each n, V_n satisfies,*

$$V_n(x+\Delta x, a+\Delta a) + V_n(x,a) \quad \geq \quad V_n(x, a+\Delta a) + V_n(x+\Delta x, a) \tag{7.23}$$

Proof $V_0(x,a) = \lambda x$ satisfies the assertion. Assume that the assertion is true for V_{n-1}. We will prove that the assertion is true for V_n. Consider the Bellman's equation,

$$V_n(x,a) = \min_{y \leq x}[\lambda x + F(y) + \beta \textstyle\int p(a,dw) V_{n-1}(x-y+w,w)] \tag{7.24}$$

Let $\mathbf{v_1} = (x+\Delta x, a+\Delta a)$, $\mathbf{v_2} = (x,a)$. Let, corresponding to $\mathbf{v_1}$, $\mathbf{v_2}$, the minimum in (7.24) be attained, respectively, at y_1, y_2.

Case 1: $y_1 - y_2 \leq \Delta x$.

$$\begin{aligned}
& y_1 - y_2 \leq \Delta x \\
\Rightarrow \quad & x + \Delta x - y_1 + w \geq x - y_2 + w \\
\Rightarrow^{*^1} \quad & [V_{n-1}(x+\Delta x - y_1 + w, w) - V_{n-1}(x - y_2 + w, w)] \uparrow \text{ as } w \uparrow \\
\Rightarrow^{*^2} \quad & \int p(a+\Delta a, dw)[V_{n-1}(x+\Delta x - y_1 + w, w) - V_{n-1}(x-y_2+w,w)] \\
& \geq \int p(a,dw)[V_{n-1}(x+\Delta x - y_1 + w, w) - V_{n-1}(x-y_2+w,w)] \\
\Rightarrow \quad & V_n(x+\Delta x, a+\Delta a) + V_n(x,a) \\
& = \lambda(x+\Delta x) + F(y_1) + \beta \int p(a+\Delta a, dw) V_{n-1}(x+\Delta x - y_1 + w, w) \\
& \qquad\qquad + \lambda x + F(y_2) + \beta \int p(a,dw) V_{n-1}(x-y_2+w,w) \\
& \geq \lambda x + F(y_2) + \beta \int p(a+\Delta a, dw) V_{n-1}(x-y_2+w,w) \\
& \qquad\qquad + \lambda(x+\Delta x) + F(y_1) + \beta \int p(a,dw) V_{n-1}(x+\Delta x - y_1 + w, w) \\
& = h_n(y_2, x, a+\Delta a) + h_n(y_1, x+\Delta x, a) \\
& \geq V_n(x, a+\Delta a) + V_n(x+\Delta x, a)
\end{aligned}$$

where the ($*^1$) follows by the induction hypothesis that (7.23) holds for V_{n-1} and Lemma 7.5. ($*^2$) follows by Theorem 5.1.

Case 2: $y_1 - y_2 \geq \Delta x$.

$$
\begin{aligned}
& y_1 - y_2 \geq \Delta x \\
\Rightarrow^* \; & F(y_1) + F(y_2) \geq F(y_1 - \Delta x) + F(y_2 + \Delta x) \\
\Rightarrow \quad & V_n(x + \Delta x, a + \Delta a) + V_n(x, a) \\
& = \lambda(x + \Delta x) + F(y_1) + \beta \int p(a + \Delta a, dw) V_{n-1}(x + \Delta x - y_1 + w, w) \\
& \qquad\qquad + \lambda x + F(y_2) + \beta \int p(a, dw) V_{n-1}(x - y_2 + w, w) \\
& \geq \lambda x + F(y_1 - \Delta x) + \beta \int p(a + \Delta a, dw) V_{n-1}(x + \Delta x - y_1 + w, w) \\
& \qquad\qquad + \lambda(x + \Delta x) + F(y_2 + \Delta x) + \beta \int p(a, dw) V_{n-1}(x - y_2 + w, w) \\
& = h_n(y_1 - \Delta x, x, a + \Delta a) + h_n(y_2 + \Delta x, x + \Delta x, a) \\
& \geq V_n(x, a + \Delta a) + V_n(x + \Delta x, a)
\end{aligned}
$$

(*) follows since F is convex. Thus, V_0 satisfies the assertion of the theorem, and V_n satisfies the assertion of the theorem on the hypothesis that V_{n-1} does. By induction, the theorem is proved. ■

We will use these properties of V_n in the section where we consider the infinite horizon discounted cost problem to derive structural results about the optimal policy.

7.3 Markov Arrivals, i.i.d. Channel Fade

Define:

$$
g_n(x, a, \mu) = \min_{y \leq x} [\lambda x + \mu F(y) + \beta \textstyle\int p(a, dw) q(d\nu) g_{n-1}(x - y + w, w, \nu)] \tag{7.25}
$$

This is the Bellman's equation.

Lemma 7.6 *If V_n exists, $V_n(x, a, \mu)$ is a continuous function of μ and an increasing function of μ.*

Proof Let $\mu' > \mu$. We first prove that $V_n(x, a, \mu') \geq V_n(x, a, \mu)$.

Let $< W_k >$ be the arrival process with $W_0 = a$. Recall that $< \Psi_k >$ is the channel fade process.

Let $< Y'_k >$ be optimal for the process $< X'_k >$ given inductively by

$$X'_{k+1} = X'_k - Y'_k + W_{k+1}, X'_0 = x, W_0 = a, \Psi_0 = \mu', 0 \le k \le n \tag{7.26}$$

Construct, inductively

$$X_{k+1} = X_k - Y'_k + W_{k+1}, X_0 = x, W_0 = a, \Psi_0 = \mu, 0 \le k \le n \tag{7.27}$$

It is easy to check that

$$V_n(x, a, \mu) \le J_{n,<Y'_k>}(x, a, \mu) \le V_n(x, a, \mu') \tag{7.28}$$

Next, we prove the continuity of $V_n(x, a, \mu)$ in μ. Let $< Y_k >$ be optimal for $< X_k >$ defined inductively as

$$X_{k+1} = X_k - Y_k + W_{k+1}, X_0 = x, W_0 = a, \Psi_0 = \mu, 0 \le k \le n \tag{7.29}$$

Let $< X'_k >$ be defined inductively as

$$X'_{k+1} = X'_k - Y_k + W_{k+1}, X'_0 = x', W_0 = a, \Psi_0 = \mu', 0 \le k \le n \tag{7.30}$$

It is easy to check that

$$V_n(x, a, \mu') - V_n(x, a, \mu) \le J_{n,<Y_k>}(x, a, \mu') - V_n(x, a, \mu) \le (\mu' - \mu)F(x)$$

This proves the continuity of V_n in μ. ■

Theorem 7.6 *i. V_ns exist. V_ns satisfy (7.25). $V_n(x, a, \mu)$ increases if one or more of x, a, μ increase. V_n is a continuous function. V_ns satisfy (there is no order relation between x', x; a', a; μ', μ in the equation below):*

$$|V_n(x', a', \mu') - V_n(x, a, \mu)| \le \frac{\lambda}{1-\beta}|x' - x| + \frac{w_{max}\lambda\beta}{(1-\beta)^2} \int |p(a', dw) - p(a, dw)| + |\mu' - \mu| F(\max(x', x)) \tag{7.31}$$

ii. For fixed a, μ, $V_n(x, a, \mu)$ is convex in x.

iii. For all μ, $x' \ge x$, $a' \ge a$,

$$V_n(x', a', \mu) + V_n(x, a, \mu) \ge V_n(x, a', \mu) + V_n(x', a, \mu) \tag{7.32}$$

Proof

i) We first prove the continuity of V_n assuming its existence.

$$\begin{aligned}
& |V_n(x',a',\mu') - V_n(x,a,\mu)| \\
\leq \; & |V_n(x',a',\mu') - V_n(x',a',\mu)| + |V_n(x',a',\mu) - V_n(x',a,\mu)| + |V_n(x',a,\mu) - V_n(x,a,\mu)| \\
\leq^* \; & |\mu' - \mu| F(x') + \frac{w_{max}\lambda\beta}{(1-\beta)^2} \int |p(a',dw) - p(a,dw)| + \frac{\lambda}{1-\beta}|x'-x|
\end{aligned}$$

(*) follows by use of bounds derived in Theorems 7.6, 7.3 and 7.1. The arguments used to derive the bounds for $|V_n(x',a',\mu) - V_n(x',a,\mu)|$ and $|V_n(x',a,\mu) - V_n(x,a,\mu)|$ are similar to those used to derive the bounds in Lemmas 7.2 and 7.4 respectively.

The other parts follow by arguments similar to Theorem 7.3.

ii) The proof is essentially the same as the proof of Theorem 7.2. We omit the proof.

iii) The proof is essentially the same as the proof of Theorem 7.5. We omit the proof.

The theorem is proved. ■

Chapter 8

Infinite Horizon Discounted Cost Problem

In this section, we consider the infinite horizon discounted cost problem for the case of Markov arrival process and i.i.d. channel fade.

Recall that $V^\beta(x, a, \mu)$ denotes the infinite horizon optimal β-discounted cost when the initial state is (x, a, μ). We do not know, apriori, that V^β exists.

Lemma 8.1 *There exists a function $M(x, a, \mu)$ such that $V_n^\beta(x, a, \mu) \leq M(x, a, \mu)$ for all n.*

Proof Let $\pi_{\mathbf{0}}$ be the policy which transmits nothing in any time slot, irrespective of the state. Then,

$$V_n^\beta(x, a, \mu) \leq J_{n,\pi_{\mathbf{0}}}^\beta(x, a, \mu) \leq J_{\pi_{\mathbf{0}}}^\beta(x, a, \mu) = \sum_{i=0}^{\infty} \beta^i (x + i w_{max}) \triangleq M(x, a, \mu) < \infty$$

Hence proved. ■

Note that since V_n^β is the optimal n-step cost, V_n^β increases as n increases. Using this fact and Lemma 8.1, it follows that there exists a function W^β such that $V_n^\beta \uparrow W^\beta$.

By use of Theorem 7.6 and a limiting argument, it follows that

$$W^\beta(x', a', \mu') - W^\beta(x, a, \mu) \leq |\mu' - \mu| F(\max(x', x)) + \frac{w_{max}\beta\lambda}{(1-\beta)^2} \int |p(a', dy) - p(a, dy)| + \frac{\lambda}{1-\beta}|x' - x| \quad (8.1)$$

This implies that W^β is continuous.

Thus, $V_n^\beta \uparrow W^\beta$, and V_n^β, W^β are continuous functions. By Dini's theorem (see Chapter 5), $V_n^\beta \uparrow W^\beta$ uniformly on compact sets.

Consider the Bellman's equation,

$$\begin{aligned} V_n^\beta(x,a,\mu) &= \min_{y\leq x}[\lambda x + \mu F(y) + \beta \textstyle\int p(a,dw)q(d\nu)V_{n-1}^\beta(x-y+w,w,\nu)] \qquad (8.2) \\ &\triangleq \min_{y\leq x} h_n^\beta(y,x,a,\mu) \end{aligned}$$

Define:

$$h^\beta(y,x,a,\mu) \triangleq \lambda x + \mu F(y) + \beta \textstyle\int p(a,dw)q(d\nu)W^\beta(x-y+w,w,\nu) \qquad (8.3)$$

Let $n \to \infty$ on both sides in (8.2).

LHS: $\lim_{n\to\infty} V_n^\beta(x,a,\mu) = W^\beta(x,a,\mu)$

RHS: By Monotone Convergence Theorem (see Chapter 5),

$$\lim_{n\to\infty} \textstyle\int p(a,dw)q(d\nu)V_{n-1}^\beta(x-y+w,w,\nu) = \int p(a,dw)q(d\nu)W^\beta(x-y+w,w,\mu)$$

Note, also, that since $V_{n-1}^\beta \to W^\beta$ uniformly on compact sets, it is easy to prove that $h_n^\beta(y,x,a,\mu) \to h^\beta(y,x,a,\mu)$ uniformly on compact sets. By Theorem 5.9,

$$\lim_{n\to\infty} \min_{y\leq x} h_n^\beta(y,x,a,\mu) = \min_{y\leq x} h^\beta(y,x,a,\mu)$$

Thus, W^β satisfies:

$$W^\beta(x,a,\mu) = \min_{y\leq x}[\lambda x + \mu F(y) + \beta \textstyle\int p(a,dw)q(d\nu)W^\beta(x-y+w,w,\nu)] \qquad (8.4)$$

(8.4) is the Bellman's equation for the Infinite Horizon Problem. We know that if the infinite horizon Bellman's equation is solvable, its solution is the optimal discounted infinite horizon cost. Thus, indeed, $W^\beta = V^\beta$, and,

$$V^\beta(x,a,\mu) = \min_{y\leq x}[\lambda x + \mu F(y) + \beta \textstyle\int p(a,dw)q(d\nu)V^\beta(x-y+w,w,\nu)] \qquad (8.5)$$

Note also, that since the infinite horizon Bellman's equation is satisfied, there exists an optimal stationary policy for the infinite horizon β-discounted problem. This policy is given by $y^*(x) = \arg\min_{y\leq x}[\lambda x + \mu F(y) + \beta \int p(a,dw)q(d\nu)V^\beta(x-y+w,w,\nu)]$.

Theorem 8.1 *i. For fixed a, μ, $V^\beta(x,a,\mu)$ is convex in x.*

ii. For all μ, $x' \geq x$, $a' \geq a$,

$$V^\beta(x', a', \mu) + V^\beta(x, a, \mu) \geq V^\beta(x, a', \mu) + V^\beta(x', a, \mu) \tag{8.6}$$

Proof Let $n \to \infty$ in Theorem 7.6. ■

We now prove structural results about the optimal stationary policies. Let y be a stationary policy. Let $y(x, a, \mu)$ denote the number of packet transmissions in state (x, a, μ) when the policy y is being used. Note that since y is a stationary policy, the action taken depends only on the present state.

Theorem 8.2 *There exists an optimal stationary policy y_{opt} such that $y_{opt}(x', a', \mu') \geq y_{opt}(x, a, \mu)$, for all $x' \geq x$, $a' \geq a$, $\mu' \leq \mu$.*

Proof Let $\chi = -\mu$, $\chi' = -\mu'$. Thus, $x' \geq x$, $a' \geq a$, $\chi' \geq \chi$.

Let $y' \geq y$.

Theorem 8.1ii and Lemma 7.5 imply that

$$\left[V^\beta(x' - y + w, w, \mu) - V^\beta(x' - y' + w, w, \mu)\right] \uparrow \text{ as } w \uparrow \forall \mu \tag{8.7}$$

Theorem 5.1 and (8.7) imply

$$\begin{aligned} &\int p(a', dw) q(d\nu) \left[V^\beta(x' - y + w, w, \nu) - V^\beta(x' - y' + w, w, \nu)\right] \\ &\qquad \geq \int p(a, dw) q(d\nu) \left[V^\beta(x' - y + w, w, \nu) - V^\beta(x' - y' + w, w, \nu)\right] \end{aligned} \tag{8.8}$$

Theorem 8.1i implies

$$\begin{aligned} &V^\beta(x' - y + w, w, \mu) - V^\beta(x' - y' + w, w, \mu) \\ &\qquad \geq V^\beta(x - y + w, w, \mu) - V^\beta(x - y' + w, w, \mu), \quad \forall \mu \end{aligned} \tag{8.9}$$

Using (8.9) in (8.8) we get

$$\begin{aligned} &\int p(a', dw) q(d\nu) \left[V^\beta(x' - y + w, w, \nu) - V^\beta(x' - y' + w, w, \nu)\right] \\ &\qquad \geq \int p(a, dw) q(d\nu) \left[V^\beta(x - y + w, w, \nu) - V^\beta(x - y' + w, w, \nu)\right] \end{aligned} \tag{8.10}$$

Manipulating (8.10) by using (8.3), we get

$$[h^\beta(y', x', a', \chi') - h^\beta(y, x', a', \chi')] \leq [h^\beta(y', x, a, \chi) - h^\beta(y, x, a, \chi)] \tag{8.11}$$

In the notation of Theorem 5.1, $\mathbf{s} = (a, \chi)$, and h satisfies decreasing differences. By Theorem 5.1, the theorem is proved. ■

Chapter 9

The Average Cost Problem

In this section we prove that the unconstrained average cost problem has a solution.

9.1 Markov Arrivals, i.i.d. Channel Fade

We first prove a result for the case of i.i.d. arrivals:

Lemma 9.1 *For all β,*

$$|V^\beta(x') - V^\beta(x)| = O[|x' - x| + |x' - x|(x' + x)] \tag{9.1}$$

Proof Let $x' > x$ and $(x' - x) < \frac{y_{max} - w_0}{2} \triangleq \Delta$.

Let $< W_k >$ be the arrival process sequence. Let $< Y_k >$ be optimal for the process:

$$X_{k+1} = X_k - Y_k + W_{k+1}, X_0 = x$$

Consider the process $< X'_k >$ defined, inductively as:

$$X'_{k+1} = X'_k - Y'_k + W_{k+1}, X_0 = x'$$

Let

$$Y'_k = \begin{cases} x' - x + Y_k & Y_k \leq y_{max} - \Delta \text{ for the first time} \\ Y_k & \text{otherwise} \end{cases}$$

Let τ be the first slot in which $Y_\tau < y_{max} - \Delta$. Then, it is easy to check that for $n \leq \tau$ $(X'_n - X_n) = (x' - x)$, and for $n > \tau$, $(X'_n - X_n) = 0$. Using the fact that $< Y'_k >$ need not be optimal for $< X'_k >$ and that, $V^\beta(x') > V^\beta(x)$, it is easy to see that

$$V^\beta(x') - V^\beta(x) \leq E[\tau\lambda(x'-x) + F(x'-x+Y_\tau) - F(Y_\tau)] \leq \lambda(x'-x)E\tau + \eta(x'-x)$$

where η is the Lipschitz constant for $F(y)$, $0 \leq y \leq y_{max}$.

We next prove that $E\tau \leq O(x)$. For $k \leq \tau$,

$$\begin{aligned} E(X_{k+1}|X_k) - X_k &= E(W_{k+1}) - E(Y_k|X_k) \\ &\leq w_0 - [y_{max} - \Delta] \\ &= -\Delta \end{aligned}$$

Thus, $V_{liap}(x) = x$ is a Liapunov function for the process $< X_k >$ for $k \leq \tau$. Note also that if $X_k \leq (y_{max} - \Delta)$, $Y_k \leq (y_{max} - \Delta)$. It follows that $E\tau \leq O(x)$. Thus, there exists α such that $E\tau < \alpha x$. Thus,

$$V^\beta(x') - V^\beta(x) \leq (\lambda\alpha x + \eta)(x'-x)$$

We had assumed that $x' - x < \Delta$. Now we remove this assumption.

$$\begin{aligned} & V^\beta(x') - V^\beta(x) \\ = & \sum_{n=0}^{\lfloor\frac{x'-x}{\Delta}\rfloor - 1} \left[V^\beta(x+(n+1)\Delta) - V^\beta(x+n\Delta)\right] + \left[V^\beta(x') - V^\beta\left(x + \left\lfloor\frac{x'-x}{\Delta}\right\rfloor\Delta\right)\right] \\ \leq & \sum_{n=0}^{\lfloor\frac{x'-x}{\Delta}\rfloor - 1} [\lambda\alpha(x+n\Delta)+\eta]\Delta + (\lambda\alpha x' + \eta)\left(x' - x - \left\lfloor\frac{x'-x}{\Delta}\right\rfloor\Delta\right) \\ = & (\lambda\alpha x + \eta)\Delta\left\lfloor\frac{x'-x}{\Delta}\right\rfloor + \frac{\lambda\alpha\Delta^2}{2}\left(\left\lfloor\frac{x'-x}{\Delta}\right\rfloor - 1\right)\left(\left\lfloor\frac{x'-x}{\Delta}\right\rfloor\right) \\ & + (\lambda\alpha x' + \eta)\left(x' - x - \left\lfloor\frac{x'-x}{\Delta}\right\rfloor\Delta\right) \\ \leq & (\lambda\alpha x + \eta)(x'-x) + \frac{\lambda\alpha}{2}(x'-x)^2 + (\lambda\alpha x' + \eta)(x'-x) \\ = & \eta(x'-x) + \lambda\alpha(x'+x)(x'-x) + \frac{\lambda\alpha}{2}(x'-x)^2 \\ = & \mathrm{O}\left[|x'-x| + |x'-x|(x'+x)\right] \end{aligned}$$

This proves the lemma. ■

Next, we state some assumptions and derive some results for the case of Markov arrivals when there is no channel fade. Let the space of packet arrivals, $[0, w_{max}]$ be denoted by W. The packet arrival process is a Markov chain M with transition kernel $p(x, A)$, $x \in W$,

$A \in \mathcal{B}(W)$, the Borel sigma field on W. We assume that this transition kernel satisfies the following:

Minorization Condition

There exists a probability measure ν on W, and a real number $\delta > 0$ such that for all $x \in W$, $A \in \mathcal{B}(W)$,

$$p(x, A) - \delta\nu(A) \geq 0 \tag{9.2}$$

Note 9.1 *The minorization condition is a rather restrictive condition on the transition probabilities $p(a, dw)$. Loosely speaking, it can be justified, if we assume, for example, that there is a neighborhood N of 0 such that $p(x, N) \geq \epsilon$ for all x, where $\epsilon > 0$ is a constant. Intuitively this means that there is a finite probability of "close to" zero packet arrivals irrespective of the number of packet arrivals in the previous slot.*

Theorem 9.1 *There exists a stationary distribution π for the Markov Chain M. Also, there exists $c > 0$, $0 < \gamma < 1$ such that*

$$|\textstyle\int w[p^{(n)}(a, dw) - \pi(dw)]| \leq c\gamma^n, \qquad \forall n > 0 \tag{9.3}$$

Proof A proof can be found in [26]. ■

Lemma 9.2 *For all a,*

$$|V^\beta(x', a) - V^\beta(x, a)| = O[|x' - x| + |x' - x|(x' + x)] \tag{9.4}$$

Proof By a method of proof similar to Lemma 9.1 and the Hoeffding's Inequality (see Chapter 5), the lemma can be proved. We omit the proof. ■

We also assume the following about the transition probabilities $p(a, dw)$:

$$\textstyle\int |p(a', dw) - p(a, dw)| \leq c_0|a' - a| \tag{9.5}$$

for some constant c_0.

Definition 9.1 $\zeta^{(n)} = \int w[p^{(n)}(a', dw) - p^{(n)}(a, dw)]$

Lemma 9.3 *i.* $\zeta^{(n)} \leq k_1|a' - a|$

ii. $\zeta^{(n)} \leq k_2\gamma^n$

for some constants k_1 and k_2, and γ is as in Theorem 9.1.

Proof i)

$$\begin{aligned}
\zeta^{(n)} &= \int [wp^{(n)}(a', dw) - wp^{(n)}(a, dw)] \\
&= \int\int wp^{(n-1)}(y, dw)[p(a', dy) - p(a, dy)] \\
&\leq \int\int wp^{(n-1)}(y, dw)|p(a', dy) - p(a, dy)| \\
&= \int |p(a', dy) - p(a, dy)| \int wp^{(n-1)}(y, dw) \\
&\leq \int w_{max}|p(a', dy) - p(a, dy)| \\
&\leq c_0 w_{max}|a' - a|
\end{aligned}$$

Thus, first part of the lemma is true with $k_1 = c_0 w_{max}$.

ii) Follows trivially by use of Theorem 9.1 with $k_2 = 2c$. ∎

Lemma 9.4 *Let $a' > a$. Then,*

$$\sum_{n=1}^{\infty} n^k \zeta^{(n)} \leq f_k(a' - a)$$

where $f_k(x) \to 0$ as $x \to 0$, for all k.

Proof Let n_0 be such that

$$k_2\gamma^{n_0} \geq k_1|a' - a| \geq k_2\gamma^{n_0+1} \tag{9.6}$$

It is easy to check that

$$n_0 = \left\lfloor \frac{\log \frac{k_1}{k_2}(a' - a)}{\log \gamma} \right\rfloor$$

$$\begin{aligned}
\sum_{n=1}^{\infty} n^k \zeta^{(n)} &= \sum_{n=1}^{n_0} n^k \zeta^{(n)} + \sum_{n=n_0+1}^{\infty} n^k \zeta^{(n)} \\
&\leq \sum_{n=1}^{n_0} n^k k_1|a' - a| + \sum_{n=n_0+1}^{\infty} n^k k_2\gamma^n \\
&= \underbrace{k_1 P_k(n_0)|a' - a|}_{T_1} + \underbrace{k_2 F_k(n_0)}_{T_2}
\end{aligned}$$

where P_k is a polynomial and F_k is a function. It is well known that $x^n \log x \to 0$ as $x \to 0$, for all n. Using the value of n_0 from (9.6) in the above equation, it follows that $T_1 \to 0$ as $|a' - a| \to 0$. Also, note that as $|a' - a| \to 0$, $n_0 \to \infty$, and hence, $T_2 \to 0$. Also, clearly, T_1 and T_2 are functions of $|a' - a|$. The lemma follows. ■

Pseudo Atom Construction

Here, we describe in brief, the Athreya-Ney-Neummelin construction (see [26]) of pseudo atom to construct another Markov chain $\check{M}$ related to M. The construction is as follows:

Let $\check{W} = W \times \{0, 1\}$, where, $W^0 = W \times \{0\}$ and $W^1 = W \times \{1\}$ are thought of as copies of W, equipped with copies $\mathcal{B}(W^0)$ and $\mathcal{B}(W^1)$ of the Borel sigma field $\mathcal{B}(W)$.

Let $\mathcal{B}(\check{W})$ denote the sigma field of subsets of $\check{W}$ generated by $\mathcal{B}(W^0)$ and $\mathcal{B}(W^1)$. Thus, $\mathcal{B}(\check{W})$ is the smallest sigma field containing sets of the form $A^0 := A \times \{0\}$, $A^1 := A \times \{1\}$; $A \in \mathcal{B}(W)$. Let x^i, w^i, a^i, $i = 0, 1$, denote the elements of $\check{W}$. x_0, w^0, a_0 denote elements in W^0 and x_1, w^1, a_1 denotes an element in W^1.

For each measure λ on W, we define a measure λ^* on $\check{W}$ as follows:

$$\begin{aligned} \lambda^*(A^0) &= (1 - \delta)\lambda(A) \\ \lambda^*(A^1) &= \delta\lambda(A) \end{aligned}$$

We now define the Markov chain $\check{M}$ on $\check{W}$ by listing down the single step transition kernel $\check{p}$.

$$\begin{aligned} \check{p}(x^0, \cdot) &= \frac{p^*(x, \cdot) - \delta\nu^*(\cdot)}{1 - \delta} \\ \check{p}(x^1, \cdot) &= \nu^*(\cdot) \end{aligned} \tag{9.7}$$

More clearly,

$$\begin{aligned} \check{p}(x^0, A^0) &= p(x, A) - \delta\nu(A) \\ \check{p}(x^0, A^1) &= \frac{\delta}{1 - \delta}[p(x, A) - \delta\nu(A)] \\ \check{p}(x^1, A^0) &= (1 - \delta)\nu(A) \\ \check{p}(x^1, A^1) &= \delta\nu(A) \end{aligned}$$

Note that these probabilities are well defined since the Minorization Condition, (9.2) holds.

It is easy to prove (see [26]) that

$$\begin{aligned}
\breve{p}^{(n)}(x^i, A^1) &= \frac{\delta}{1-\delta}\breve{p}^{(n)}(x^i, A^0) \\
p^{(n)}(x, A) &= (1-\delta)\breve{p}^{(n)}(x^0, A^0 \cup A^1) + \delta\breve{p}^{(n)}(x^1, A^0 \cup A^1)
\end{aligned} \tag{9.8}$$

Thus, clearly, the Markov chain M can be recovered from $\breve{M}$ as follows:

Let the initial distribution on M be $\lambda(dw)$. Let the initial distribution on $\breve{M}$ be $\lambda^*(dw)$.

Then, it is easy to prove by use of (9.8) that

$$\begin{aligned}
&\int_W \lambda(dw) \Pr(w_n \in A_n, \ldots, w_1 \in A_1 | w_0 = w) \\
&\qquad = \int_{\breve{W}} \lambda^*(dw) \Pr(w_n \in A_n^0 \cup A_n^1, \ldots, w_1 \in A_1^0 \cup A_1^1 | w_0 = w)
\end{aligned} \tag{9.9}$$

The intuition behind this formation of the chain $\breve{M}$ is that if n is the smallest time at which the chain is in W^1, the behavior of the chain from time $(n+1)$ onwards is independent of where the chain started. This is because, as can be checked from (9.7), $p(x^1, \cdot)$ is independent of x^1. This property about $\breve{M}$ helps derive many useful properties about $\breve{M}$, and corresponding properties about M can be proved by use of (9.8).

The set W^1 is called *Pseudo Atom.*

Next, we make a few definitions and observations.

We define a partial order, $\geq$, on $\breve{W}$ as follows:

Definition 9.2 (Partial Order, $\geq$, on $\breve{W}$) *$w_1^i \geq w_2^j$ if $i = j$ and $w_1 \geq w_2$. Note that if $i \neq j$, w_1^i and w_2^j are not related by the partial order $\geq$.*

Definition 9.3 (Stochastic Dominance, $\succ$, on $\breve{W}$) *Let $\breve{p}$ and $\breve{q}$ be probability measures on $\breve{W}$. We say that $\breve{p} \succ \breve{q}$ if $\breve{p}|_{W^0} \succ \breve{q}|_{W^0}$ and $\breve{p}|_{W^1} \succ \breve{q}|_{W^1}$. Equivalently, $\breve{p}([w^i, w^i_{max}]) \geq \breve{q}([w^i, w^i_{max}])$ for $i = 0, 1$, for all $w^i \in \breve{W}$.*

Lemma 9.5 *Let $\breve{X}_1$ and $\breve{X}_2$ be random variables with sample space $\breve{W}$. Let $\breve{X}_1 \sim \breve{p}$, $\breve{X}_2 \sim \breve{q}$. Let $\breve{p} \succ \breve{q}$. Then, there exist random variables $\breve{X}_1'$ and $\breve{X}_2'$ such that $\breve{X}_1' \sim \breve{p}$, $\breve{X}_2' \sim \breve{q}$, and $\breve{X}_1' \geq \breve{X}_2'$ almost everywhere.*

Proof This is a trivial extension of Theorem 5.2. ■

Lemma 9.6 *$\check{p}(x'^0, \cdot) \succ \check{p}(x^0, \cdot)$, $\check{p}(x'^1, \cdot) = \check{p}(x^1, \cdot)$, where $x' \geq x$*

Proof This is easy to prove. The proof is omitted ■

The packet arrival process is the Markov chain M. Since M can be recovered from $\check{M}$, we assume that our packet arrival process is the chain $\check{M}$. By using the structure on $\check{M}$, we will relate different sample paths of the packet arrival process and derive useful results.

Let $< W_k >$ and $< W'_k >$ be packet arrival processes with $W_1 \sim (1-\delta)a^0 + \delta a^1$, $W'_1 \sim (1-\delta)a'^0 + \delta a'^1$. For each sample path $< w_k >^{i_k}$ of the process $< W_k >$, we associate a corresponding sample path $< w'^{i'_k}_k >$ of the process $< W'_k >$ as follows:

- Case I: k is such that $i_r = 0$, $0 \leq r \leq (k-1)$. Assume, inductively, that $w'^{i'_r}_r \geq w^{i_r}_r$, $0 \leq r \leq (k-1)$, and $i'_r = i_r$, $0 \leq r \leq (k-1)$. Now, $\check{p}(w'^{i'_{k-1}}_{k-1}, \cdot) = \check{p}(w'^0_{k-1}, \cdot) \succ \check{p}(w^0_{k-1}, \cdot) = \check{p}(w^{i_{k-1}}_{k-1}, \cdot)$. By Lemma 9.5, we can have an association wherein $w'^{i'_k}_k > w^{i_k}_k$, which also implies that $i'_k = i_k$.
- Case II: There exists $k' < k$ such that $i'_k = 1$. In this case, associate, $w'^{i'_k}_k = w^{i_k}_k$.

Thus, to each path $< w_k{}^{i_k} >$, we associate a corresponding sample path $< w'^{i'_k}_k >$ such that $w'^{i'_k}_k \geq w_k{}^{i_k}$, and equality holds for $n \geq (n_0 + 1)$, where n_0 is the smallest integer for which $i_{n_0} = 1$.

Definition 9.4 (Coupling Time, τ_c) *The time when $< W_k >$ (and also $< W'_k >$) hits W^1 for the first time is called the coupling time, and is denoted by τ_c.*

The reason for this nomenclature is that $< W_k >$ and $< W'_k >$ get "coupled" at this time in the sense that for all future time, $W'_k = W_k$.

We first state two lemmas:

Lemma 9.7 *The probability that a sample path of $< W_k >$ (or $< W'_k >$) has coupling time n is given by:*

$$\Pr(\tau_c = n) \quad = \quad \delta(1-\delta)^n$$

Proof Note that $W_0 \sim (1-\delta)a^0 + \delta a^1$. Thus,

$$\begin{aligned}
\Pr(W_0 \in W^0) &= 1-\delta \\
\Pr(W_0 \in W^1) &= \delta
\end{aligned}$$

From (9.7), it is easy to see that

$$\begin{aligned}
\Pr(W_{k+1} \in W^0 \mid W_k \in W^0) &= 1-\delta \\
\Pr(W_{k+1} \in W^1 \mid W_k \in W^0) &= \delta
\end{aligned}$$

Now,

$$\Pr(\tau_c = 0) = \Pr(W_0 \in W^1) = \delta$$

$$\begin{aligned}
& \Pr(\tau_c = n) \\
= \; & \Pr(W_0 \in W^0, W_1 \in W^0, \ldots, W_{n-1} \in W^0, W_n \in W^1) \\
= \; & \Pr(W_0 \in W^0)\Big[\prod_{k=0}^{n-2} \Pr(W_{k+1} \in W^0 \mid W_k \in W^0)\Big] \Pr(W_n \in W^1 \mid W_{n-1} \in W^0) \\
= \; & (1-\delta)(1-\delta)^{n-1}\delta \\
= \; & \delta(1-\delta)^n
\end{aligned}$$

Hence, proved. ■

Definition 9.5 $\check{p}_\tau^{(n)}(a^i, A^j) \triangleq \Pr(W_n \in A^j \mid W_0 = a^i, \tau_c = \tau)$

Lemma 9.8 *Given* $\tau_c = \tau \geq 1$,

$$\begin{aligned}
\check{p}_\tau^{(n)}(a^0, A^0) &= \frac{p^{(n)}(a, A) - \delta\eta_n(A)}{(1-\delta)^n} && 0 \leq n \leq \tau - 1 \\
\check{p}_\tau^{(n)}(a^0, A^1) &= 0 && 0 \leq n \leq \tau - 1 \\
\check{p}_\tau^{(n)}(a^0, A^0) &= 0 && n = \tau \\
\check{p}_\tau^{(n)}(a^0, A^1) &= \frac{p^{(n)}(a, A) - \delta\eta_n(A)}{(1-\delta)^n} && n = \tau
\end{aligned}$$

where η_n *are positive measures satisfying*

$$\int_W \frac{p^{(n)}(a, dw) - \delta\eta_n(dw)}{(1-\delta)^n} = 1$$

Proof We first prove the theorem for $n = 1$.

Case 1: $\tau > 1$. This means that $W_1 \in W^0$. This implies that $\check{p}_\tau^{(1)}(a^0, A^1) = 0$.

$$
\begin{aligned}
\check{p}_\tau^{(1)}(a^0, A^0) &= \Pr(W_1 \in A^0 \mid W_0 = a^0, \tau_c = \tau) \\
&= \frac{\Pr(W_1 \in A^0, \tau_c = \tau \mid W_0 = a^0)}{\Pr(\tau_c = \tau \mid W_0 = a^0)} \\
&= \frac{\Pr(\tau_c = \tau \mid W_1 \in A^0, W_0 = a^0)\Pr(W_1 \in A^0 \mid W_0 = a^0)}{\Pr(\tau_c = \tau \mid W_0 = a^0)} \\
&= \frac{\Pr(\tau_c = \tau \mid W_1 \in A^0)\check{p}(a^0, A^0)}{\Pr(\tau_c = \tau \mid W_0 = a^0)} \\
&= \frac{(1-\delta)^{\tau-1}[p(a, A) - \delta\nu(A)]}{(1-\delta)^\tau} \\
&= \frac{p(a, A) - \delta\nu(A)}{1-\delta}
\end{aligned}
$$

Thus, the theorem is true in the case $\tau > 1$, $n = 1$, with $\eta_1 \equiv \nu$.

Case 2: $\tau = 1$. The proof is similar to Case 1, and indeed, $\eta_1 \equiv \nu$.

Let the theorem be true for $n = k$. We prove the theorem for $n = (k+1)$.

Case 1: $\tau > (k+1)$. Clearly, $\check{p}_\tau^{(k+1)}(a^0, A^1) = 0$.

$$
\begin{aligned}
& \check{p}_\tau^{(k+1)}(a^0, A^0) \\
=\ & \int_{W^0} \check{p}_\tau^{(1)}(a^0, dy^0) p_\tau^{(k)}(dy^0, A^0) \\
=\ & \int_W \left[\frac{p(a, dy) - \delta\nu(dy)}{(1-\delta)} \cdot \frac{p^{(k)}(y, A) - \delta\eta_k(A)}{(1-\delta)^k}\right] \\
=\ & \frac{p^{(k+1)}(x, A) - \delta\left[\eta_k(A) - \delta\eta_k(A) + \int_W \nu(dy)p^{(k)}(y, A)\right]}{(1-\delta)^{k+1}}
\end{aligned}
$$

Thus, with $\eta_{k+1}(A) = [\eta_k(A) - \delta\eta_k(A) + \int_W \nu(dy)p^{(k)}(y, A)]$, Case 1 follows.

Case 2: $\tau = (k+1)$. Similar proof can be given, and indeed, η_{k+1} is the same as in Case 1.

The theorem follows by induction. ■

Next, we estimate $V^\beta(x, a') - V^\beta(x, a)$.

Let $< Y_k >$ be optimal for the process $< X_k >$, defined inductively as:

$$X_{k+1} = X_k - Y_k + W_{k+1}, X_0 = x, W_0 = (1-\delta)a^0 + \delta a^1$$

Consider the process $< X'_k >$ defined inductively, for $0 \leq k \leq \tau_c$ as:

$$X'_{k+1} = X'_k - Y'_k + W'_{k+1}, X_0 = x, W_0 = (1-\delta)a^0 + \delta a^1, 0 \leq k \leq \tau_c$$

where, to each sample path $< w_k^{i_k} >$ of $< W_k >$ we have associated the sample path $< {w'}_k^{i'_k} >$ of $< W'_k >$ as before. Since $w'_k > w_k$, the policy $< Y_k >$ is admissible in the process $< X'_k >$. Note also, that $< Y_k >$ need not be optimal for $< X'_k >$. Hence,

$$\begin{aligned} V^\beta(x, a') - V^\beta(x, a) \leq E\Big[& \sum_{k=0}^{\tau_c} \beta^k (X'_k - X_k) \\ & + \beta^{\tau_c+1} \left\{ V^\beta(X'_{\tau_c+1}, W'_{\tau_c+1}) - V^\beta(X_{\tau_c+1}, W_{\tau_c+1}) \right\} \Big] \end{aligned} \tag{9.10}$$

$$\begin{aligned} & E\left[\sum_{k=0}^{\tau_c} \beta^k (X'_k - X_k)\right] \\ \leq \quad & E\left[\sum_{k=0}^{\tau_c} (X'_k - X_k)\right] \\ = \quad & E\left[\sum_{k=1}^{\tau_c} (\tau_c - k + 1)(W'_k - W_k)\right] \\ = \quad & \sum_{n=1}^{\infty} p(\tau_c = n) E\left[\sum_{k=1}^{n} (n - k + 1)(W'_k - W_k) \;\middle|\; \tau_c = n\right] \\ = \quad & \sum_{n=1}^{\infty} p(\tau_c = n) \sum_{k=1}^{n} (n - k + 1) E(W'_k - W_k \mid \tau_c = n) \\ =^{*^1} \quad & \sum_{n=1}^{\infty} \delta(1-\delta)^n \sum_{k=1}^{n} (n - k + 1) \int_{W^0} w \left[\check{p}_{\tau_c}^{(k)}({a'}^0, dw^0) - \check{p}_{\tau_c}^{(k)}(a^0, dw^0)\right] \\ =^{*^2} \quad & \sum_{n=1}^{\infty} \delta(1-\delta)^n \sum_{k=1}^{n} (n - k + 1) \frac{\zeta^{(k)}}{(1-\delta)^k} \\ =^{*^3} \quad & \mathrm{O}\left(\sum_{n=1}^{\infty} \zeta^{(n)}\right) \\ =^{*^4} \quad & \mathrm{O}[f_0(a' - a)] \to 0 \text{ as } a' \to a \end{aligned} \tag{9.11}$$

- $(*^1)$ follows from Theorems 9.7 and 9.8.
- $(*^2)$ follows from Definition 9.1.
- $(*^3)$ follows by interchange of summations.
- $(*^4)$ follows by Lemma 9.4.

$$
\begin{aligned}
&E[V^\beta(X'_{\tau_c+1}, W'_{\tau_c+1}) - V^\beta(X_{\tau_c+1}, W_{\tau_c+1})] \\
=^{*1} \quad & E[V^\beta(X'_{\tau_c+1}, W_{\tau_c+1}) - V^\beta(X_{\tau_c+1}, W_{\tau_c+1})] \\
= \quad & \sum_{n=0}^{\infty} p(\tau_c = n) E\Bigg[V^\beta \left(x - \sum_{k=0}^{n} Y_k + \sum_{k=1}^{n} W'_k + W_{n+1},\ W_{n+1} \right) \\
& \qquad - V^\beta \left(x - \sum_{k=0}^{n} Y_k + \sum_{k=1}^{n+1} W_k,\ W_{n+1} \right) \Bigg| \tau_c = n \Bigg] \\
\leq^{*2} \quad & \sum_{n=1}^{\infty} p(\tau_c = n) E\Bigg[\left(\sum_{k=1}^{n} W'_k - \sum_{k=1}^{n} W_k \right) \\
& \qquad + \left(\sum_{k=1}^{n} W'_k + \sum_{k=1}^{n} W_k \right) \left(\sum_{k=1}^{n} W'_k - \sum_{k=1}^{n} W_k \right) \Bigg| \tau_c = n \Bigg] \\
=^{*3} \quad & \sum_{n=1}^{\infty} p(\tau_c = n) \left[\sum_{k=1}^{n} \frac{\zeta^{(k)}}{(1-\delta)^k} + 2n w_{max} \sum_{k=1}^{n} \frac{\zeta^{(k)}}{(1-\delta)^k} \right] \\
=^{*4} \quad & \mathrm{O}\left(\sum_{n=0}^{\infty} n \zeta^{(n)} \right) \\
=^{*5} \quad & [f_1(a'-a)] \to \ 0 \text{ as } a' \to a
\end{aligned}
\tag{9.12}
$$

where f_1 is as in Lemma 9.4

- $(*^1)$ follows because for $n > \tau_c$, $W'_n > W_n$.
- $(*^2)$ follows by Theorem 9.2.
- $(*^3)$ Here, we have used the fact that if X and Y are non-negative random variables and the maximum possible value of X is x_{max}, $E(XY) \leq x_{max} EY$. $X = [\sum_{k=1}^{n} W'_k + \sum_{k=1}^{n} W_k]$ and $Y = [\sum_{k=1}^{n} W'_k - \sum_{k=1}^{n} W_k]$ in this case.
- $(*^4)$ follows by interchanging summations.
- $(*^5)$ follows by Lemma 9.4.

Using (9.11) and (9.12) in (9.10), we get:

$$
V^\beta(x, a') - V^\beta(x, a) \leq \mathrm{O}[f_0(a'-a) + f_1(a'-a)] = \mathrm{O}[f_1(a'-a)] \tag{9.13}
$$

We now consider the general case when the packet arrival process is Markov and the channel fade is i.i.d.

Lemma 9.9

$$V^\beta(x',a',\mu') - V^\beta(x,a,\mu)$$
$$\leq \underbrace{k_1[|x'-x| + |x'-x|(x'-x)]}_{T_1} + \underbrace{k_2 f_1(a'-a)}_{T_2} + \underbrace{(\mu'-\mu)F(\min(y_{max},x))}_{T_3} \quad (9.14)$$

Proof T_1 and T_2 follow by Lemma 9.2 and (9.13). T_3 follows by arguments similar to Lemma 7.6. ■

Definition 9.6 ($\bar{V}^\beta$) $\bar{V}^\beta(x,a,\mu) = V^\beta(x,a,\mu) - V^\beta(0,0,0)$

Lemma 9.10 *i. For all x, a, μ, there exists M depending on x, a, μ such that $\bar{V}^\beta(x,a,\mu) \leq M$ for all β.*

ii. $\bar{V}^\beta$ is an equicontinuous family of functions over β.

Proof Both parts of the lemma follow by Lemma 9.9. ■

Lemma 9.11 *There exists M such that $(1-\beta)V^\beta(0,0,0) \leq M$ for all β.*

Proof

$$\begin{aligned}
V^\beta(0,0,0) &=^{*^1} \beta \textstyle\int p(0,dw)q(d\nu)V^\beta(w,w,\mu) \\
&\leq \beta V^\beta(w_{max}, w_{max}, \mu_{max}) \\
&\leq^{*^2} \beta[V^\beta(0,0,0) + k_1(w_{max} + w_{max}^2) + k_2 f_1(w_{max}) \\
&\qquad\qquad + (\mu'-\mu)F(\min(w_{max}, y_{max}))]
\end{aligned}$$

$$\Rightarrow (1-\beta)V^\beta(0,0,0) \leq \beta[k_1(w_{max} + w_{max}^2) + k_2 f_1(w_{max}) + (\mu'-\mu)F(\min(w_{max}, y_{max}))]$$

Here, k_1 and k_2 are some constants. ($*^1$) follows by Bellman's equation and ($*^2$) follows by Lemma 9.9. The lemma follows. ■

Lemma 9.12 *There exists a sequence $<\beta_n>$, $\beta_n \uparrow 1$, $\beta_n < 1$, such that*

1. $\lim_{n\to\infty} \bar{V}^{\beta_n} = \bar{V}$ *for some function* $\bar{V}$*. In fact, the convergence is uniform.*

2. $\lim_{n\to\infty}(1-\beta_n)V^{\beta_n}(0,0,0) = \rho$*, for some* ρ*.*

Proof By Lemma 9.10, $\bar{V}^{\beta}$ is a bounded equicontinuous family of functions. By the Arzela-Ascoli Theorem (see Chapter 5), there exists sequence $< \beta_{k_n} > \uparrow 1$ such that $\lim_{n\to\infty} \bar{V}^{\beta_{k_n}} = \bar{V}$, and the convergence is uniform.

By Lemma 9.11, $(1-\beta)\bar{V}^{\beta}(0,0,0)$ is bounded. Hence, there exists a subsequence $< \beta_n >$ of $< \beta_{k_n} >$, $\beta_n \uparrow 1$, such that $\lim_{n\to\infty}(1-\beta_n)\bar{V}^{\beta_n}(0,0,0) = \rho$, for some ρ.

The sequence $< \beta_n >$ is the required sequence. ■

Consider the Bellman's equation, (8.5),

$$V^{\beta}(x,a,\mu) = \min_{y\le x}[\lambda x + \mu F(y) + \beta \textstyle\int p(a,dw)q(d\nu)V^{\beta}(x-y+w,w,\nu)]$$

On some manipulation and using Definition 9.6,

$$\bar{V}^{\beta}(x,a,\mu) = \min_{y\le x}[\lambda x + \mu F(y) - (1-\beta)\bar{V}^{\beta}(0,0,0) \\ + \beta \textstyle\int p(a,dw)q(d\nu)\bar{V}^{\beta}(x-y+w,w,\nu)]$$

Letting $\beta \to 1$ on both sides along β_n, we get:

$$\bar{V}(x,a,\mu) = \min_{y\le x}[\lambda x + \mu F(y) - \rho + \textstyle\int p(a,dw)q(d\nu)\bar{V}(x-y+w,w,\nu)] \tag{9.15}$$

This follows by arguments similar to those used to prove the convergence of V_n^{β} to V^{β} in Chapter 8, except that in place of the Monotone Convergence Theorem, we have to use the Dominated Convergence Theorem (see Chapter 5). We have to use the Dominated Convergence Theorem because $\bar{V}^{\beta_n}$ need not be an increasing sequence of functions. The applicability of the Dominated Convergence Theorem follows by use of Lemma 9.10i. Also we do not need to use Dini's Theorem since uniform convergence on compact sets follows by equicontinuity.

This is the Bellman's equation for the Average Cost Problem, and hence, there exists a stationary policy which is optimal for the Average Cost Problem.

Theorem 9.2 *i. For fixed* a, μ, $\bar{V}(x,a,\mu)$ *is convex in* x*.*

ii. For all μ, $x' \ge x$, $a' \ge a$,

$$\bar{V}(x',a',\mu) + \bar{V}(x,a,\mu) \ge \bar{V}(x,a',\mu) + \bar{V}(x',a,\mu) \tag{9.16}$$

Proof Using Definition 9.6 it is easy to see that in Theorem 8.1, V^{β} can be replaced everywhere by $\bar{V}^{\beta}$. Let $\beta \to 1$ along β_n. ■

Theorem 9.3 *There exists an optimal stationary policy y_{opt} such that $y_{opt}(x', a', \mu') \geq y_{opt}(x, a, \mu)$, for all $x' \geq x$, $a' \geq a$, $\mu' \leq \mu$.*

Proof The proof is similar to that of Theorem 8.2 ■ .

9.2 Markov Arrivals, Markov Channel Fade

In case of Markov channel fade, by using the methods of the previous two chapters and Section 9.1, it can be proved that there exists an optimal stationary policy which is increasing in buffer occupancy and the number of packet arrivals in the previous time slot. However, unlike the case of i.i.d. channel fade where the policy was decreasing in the level of channel fade, nothing can be said, in general, when the channel fade is Markov. We give an intuitive reason for this:

Let $q(\nu_1, d\mu) \succ q(\nu_2, d\mu)$ if $\nu_1 > \nu_2$. Intuitively, this means that if the channel fade in the current slot is high, it is expected that the channel fade in the next slot will be higher than what it would be if the channel fade in the current slot is low. Thus, even if the channel fade in the current slot is high, we might be better off transmitting a larger number of packets in anticipation of a very bad channel fade in the future, compared to when the channel fade in the current slot is low. For this reason, the number of transmissions in a bad channel state might be higher compared to the number of transmissions in a good channel state.

Chapter 10

The Constrained Average Cost Problem

In this chapter, we prove that there is a stationary, deterministic policy which optimizes the power, given a constraint on the average delay.

We assume that the packet arrival process is FSD Markov and the channel fade is i.i.d.

We make the following assumption: $F(y)$ is a *strictly* convex function of y.

Lemma 10.1 *For each λ, there is a unique policy which minimizes $P + \lambda D$. Also, this policy is stationary and deterministic.*

Proof Consider the Bellman's Equation, (9.15),

$$\bar{V}(x, a, \mu) = \min_{y \leq x}[\lambda x + \mu F(y) - \rho + \textstyle\int p(a, dw) q(d\nu) \bar{V}(x - y + w, w, \nu)] \triangleq \min_{y \leq x} \bar{h}(y, x, a, \mu)$$

Now, for fixed x, a, μ, it follows by the assumption that F is strictly convex and Theorem 9.2i that $\bar{h}(y, x, a, \mu)$ is a strictly convex function of y. Hence, for fixed, x, a, μ, the minimum in (9.15) is attained at a unique y. It follows that there exists a unique policy which minimizes $P + \lambda D$. Clearly, this policy is stationary and deterministic. ■

Lemma 10.2 *The $P - D$ region achieved by all possible schedulers is a convex region, bounded from top, left and bottom, and unbounded from the right.*

Proof This convexity follows from a standard result in Markov Decision Processes (see [27]). The other conclusions of the lemma are easy to prove. ■

Definition 10.1 (Boundary, B) *A point (D_0, P_0) lies on the boundary, B, of the $P - D$ region if and only if (D_1, P_0) does not lie in the $P - D$ region for all $D_1 < D_0$.*

Note 10.1 *B is a function, $P = P(D)$.*

Lemma 10.3 *$P(D)$ is a convex, decreasing function of D.*

Proof This is easy to prove. The proof is omitted. ■

Let the left derivative of the function P at the point D be denoted by $D_l(D)$. D_l exists by Theorem 5.10ii. Note that $D_l(D) \leq 0$ for all D.

Lemma 10.4 *For all achievable D_1, D_2, $D_1 \neq D_2$, $P(D_1) - D_l(D_1)D_1 < P(D_2) - D_l(D_1)D_2$.*

Proof That $P(D_1) - D_l(D_1)D_1 \leq P(D_2) - D_l(D_1)D_2$ follows by Theorem 5.10iv. Also, if $P(D_1) - D_l(D_1)D_1 = P(D_2) - D_l(D_1)D_2$, for $\lambda = -D_l(D_1)$, $P + \lambda D$ is optimized for atleast two schedulers with (D, P) equal to $(D_1, P(D_1))$ and $(D_2, P(D_2))$ respectively, which contradicts Lemma 10.1. This proves the lemma. ■

Theorem 10.1 *Given an achievable bound D_0 on the average packet delay D, there exists a stationary deterministic policy which satisfies the delay constraint D_0 and optimizes the power P. This policy is increasing in the buffer occupancy, increasing in the number of packet arrivals in the previous time slot, and decreasing in the* level *of channel fade.*

Proof Clearly, by Lemma 10.4, the scheduler which yields $(D, P) = (D_0, P(D_0))$ is the same as the scheduler which optimizes $P + \lambda D$ for $\lambda = -D_l(D_0)$. The latter scheduler satisfies the hypothesis of the theorem, and hence, so does the former. This proves the theorem. ■

Chapter 11

Conclusions

We have considered the problem of power efficient scheduling, wherein, we have to minimize the power requirement, P, subject to a constraint on the average delay. We conclude that:

- When the packet arrival process is FSD Markov and the channel fade is i.i.d., there exists an optimal stationary deterministic policy which is increasing in buffer occupancy and number of packet arrivals in the previous slot, and decreasing in the level of channel fade.
- When the packet arrival process is FSD Markov and the channel fade is Markov, there exists an optimal stationary deterministic policy which is increasing in buffer occupancy and number of packet arrivals in the previous time slot. Nothing can, in general, be said about the nature of the optimal policy with respect to channel fade.

Bibliography

[1] V. I. Zubov, *Methods of A. M. Lyapunov and their Application.* Noordhoff, 1964.

[2] Morris W. Hirsch, Stephen Smale, *Differential Equations, Dynamical Systems, and Linear Algebra.* Academic Press, 1974.

[3] J. L. Willems, *Stability Theory of Dynamical Systems.* Wiley Interscience, 1970.

[4] Walter Rudin, *Principles of Mathematical Analysis.* McGraw-Hill, 1976.

[5] David Cox, John Little, Donal O'Shea, *Ideals, Varieties and Algorithms.* Springer Verlag, 1992.

[6] David Eisenbud, Daniel R. Grayson, Michael Stillman, Bernd Sturmfels, *Computations in Algebraic Geometry with Macaulay 2.* Algorithms and Computations in Mathematics, Volume 8, Springer Verlag, 2002.

[7] V. V. Nemytskii, V. V. Stepanov, *Qualitative Theory of Differential Equations.* Princeton University Press, 1960.

[8] N. G. Lloyd, J. M. Pearson, "Symmetry in Planar Dynamical Systems," *Journal of Symbolic Computation*, pp. 357–366, 2002.

[9] N. G. Lloyd, J. M. Pearson, "A Cubic Kolmogorov System with Six Limit Cycles," *Computers and Mathematics with Applications 44*, pp. 445–445, 2002.

[10] N. G. Lloyd, J. M. Pearson, "Cubic Systems and Abel Equations," *Journal of Differential Equations 147*, pp. 435–454, 1998.

[11] N. G. Lloyd, J. M. Pearson, "Computing Integrability Conditions for a Cubic Differential System," *Computer Mathematics Applications, Vol. 32, No. 10*, pp. 99–107, 1996.

[12] R. Gallager, D. Bertsekas, *Data Networks.* Prentice Hall, 1991.

[13] Balaji Prabhakar, Elif Uysal-Biyikoglu and Abbas El Gamal, "Energy-efficient Transmission over a Wireless Link via Lazy Packet Scheduling," in *Proceedings of the IEEE INFOCOM*, (Anchorage, Alaska), 2001.

[14] Abbas El Gamal, Chandra Nair, Balaji Prabhakar, Elif Uysal-Biyikoglu and Sina Zahedi, "Energy-efficient Scheduling of Packet Transmissions over Wireless Networks," *INFOCOM*, 2002.

[15] Dinesh Rajan, Ashutosh Sabharwal and Behnaam Aazhang, "Delay-Bounded Packet Scheduling of Bursty Traffic Over Wireless Channels," *IEEE Transactions on Information Theory*, November 2001.

[16] B. E. Collins and R. L. Cruz, "Transmission Policies for Time Varying Channels with Average Delay Constraints," *Proc. 1999 Allerton Conf. on Commun., Control, & Comp.*

[17] Randall A. Berry, *Power and Delay Trade-offs in Fading Channels.* PhD thesis, Massachusetts Institute of Technology, June 2000. http://www.ece.northwestern.edu/~rberry/pubs.html.

[18] Rendall A. Berry and R. G. Gallager, "Communications over Fading Channels with Delay Constraints," *IEEE Transactions on Information Theory*, vol. 43, pp. 1986–1992, November 1997.

[19] Heng Wang, *Opportunistic Transmission for Wireless Data over Fading Channels under Energy and Delay Constraints, Chapter 4.* PhD thesis, Graduate School-New Brunswick Rutgers, The State University of New Jersey, January 2003. http://www.winlab.rutgers.edu/~hwang/papers/thesis.pdf.

[20] Munish Goyal, Anurag Kumar, Vinod Sharma, "Scheduling Transmission over a Fading Channel," in *IEEE INFOCOM*, 2003.

[21] Dimitri P. Bertsekas, *Dynamic Programming and Optimal Control, Vol. I and II.* Prentice-Hall, Englewoods Cliffs, 1995.

[22] Moshe Shaked and J. George Shanthikumar, *Stochastic Orders and their Applications.* Boston:Academic Press, 1994.

[23] Peter W. Glynn and Dirk Ormoneit, "Hoeffding's Inequality for Uniformly Ergodic Markov Chains," *Statistics and Probability Letters*, vol. 56, pp. 143–146, January 2002.

[24] H. L. Royden, *Real Analysis.* Macmillan Publishing Company, 1968.

[25] Rangarajan K Sundaram, *A First Course in Optimization Theory.* Cambridge University Press, 1996.

[26] Meyn and Tweedie, *Markov Chains and Stochastic Stability.* Springer-Verlag, 1993.

[27] Eitan Altman, *Constrained Markov Decision Processes.* Chapman and Hall/CRC, 1999.