A New Perspective on Multi-Echelon Inventory Systems

by

Alp Muharremoğlu


Submitted to the Sloan School of Management in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Operations Research

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2002

© Massachusetts Institute of Technology 2002. All rights reserved.

Author

Sloan School of Management

July 15, 2002

Certified by

John N. Tsitsiklis
Professor of Electrical Engineering and Computer Science Co-Director, Operations Research Center Thesis Supervisor

Accepted by

James B. Orlin
Edward Pennell Brooks Professor of Operations Research Co-Director, Operations Research Center
A New Perspective on Multi-Echelon Inventory Systems

by

Alp Muharremoğlu

Submitted to the Sloan School of Management
on July 15, 2002, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Operations Research

Abstract

We present a new methodology for analyzing multi-echelon inventory systems. The methodology relies on decomposing complicated multi-echelon inventory control problems into much smaller and manageable subproblems, whose solutions in turn help us either characterize the structure of optimal policies for the corresponding overall problems and/or to compute optimal policies efficiently.

We analyze four multi-echelon systems through this perspective. The first system is a serial system with stochastic leadtimes and Markov modulated demand. Here, the methodology amounts to focusing on a single unit as it travels through the supply chain and showing that the original problem is simply a series of single unit problems that are essentially decoupled. We are able to show that state dependent echelon base stock policies are optimal in this setting, both in finite and infinite horizon.

A serial system with expediting options is analyzed next. A stage is not restricted to order items from the next upstream stage but can place orders at stages further upstream in the supply chain, by incurring certain extra costs. We show that given a restriction on the expediting cost structure that we call supermodularity, the system decomposes into single unit subproblems. We characterize the structure of optimal policies as extended echelon base stock policies, which is a generalization of echelon base stock policies.

Next we study a serial system with batch size constraints. We show that the problem can be decomposed into subproblems, each of which has a single batch. We then show that \((R, nQ)\) policies are optimal for this problem, which can be interpreted as echelon base stock policies that incorporate the batch size restrictions.

In addition to providing a simple proof technique, the new approach gives rise to efficient algorithms for the calculation of the policy parameters, for all the systems described above.

Finally we analyze an assembly system with stochastic leadtimes. We show that the problem can be decomposed into a series of subproblems, each with a single kit of parts. This enhances our understanding about optimal policies in this setting and we develop a relatively efficient algorithm for the computation of optimal policies.
Thesis Supervisor: John N. Tsitsiklis
Title: Professor of Electrical Engineering and Computer Science
Co-Director, Operations Research Center
Acknowledgments

First and foremost, I would like to thank my advisor, Prof. John Tsitsiklis. This thesis would not have been completed without his guidance, his wisdom and his constant support. More importantly, I have learned a great deal from him and it was truly an honor to have the opportunity to work so closely with someone of his caliber.

Special thanks go to my committee members, Prof. Stephen Graves and Prof. David Simchi-Levi, for their time, effort and advice. Their suggestions and comments significantly improved the quality of this dissertation.

I am grateful to the Ministry of National Education of the Republic of Turkey for sponsoring my undergraduate education at the University of Michigan. At Michigan, the guidance of Prof. Tava Lennon-Olsen and Prof. Izak Duenyas were invaluable. At MIT, I wrote my Master’s thesis under the supervision of Prof. Amedeo Odoni and his support, in terms of funding and academic issues, is greatly appreciated.

I really enjoyed the last five years that I spent as a graduate student at MIT, to a large extent due to the friendships I formed with fellow students. Hernan Alperin, Fernando Ordonez, Ozlem Ergun, Ramazan Demir, Martin Haugh, Amy Cohn, John Bossert, Emre Koksal, Volkan Muslu, Marc Coumeri, Nicholas Stier, Rene Caldentey, Marina Zaretsky, Beril Toktay, Dessi Pachamanova, Hasan Arslan, Soulaymane Kachani, Dushyant Sharma, Jeremie Gallien, Ayla Ergun and Murat Seyhan are some of the names that come to mind.

Words are not enough to express my gratitude and my appreciation for my family. My parents, Ismet and Güler Muharremoğlu, my brother Ali Muharremoğlu and my sister Ebru Saravin have always supported me, even when I was thousands of miles away from them. Finally and most importantly, my wife Pelin has been a constant support and a source of encouragement for many years. I feel very fortunate to have her companionship. I am forever grateful to her for standing beside me through good and bad, for being my best friend, for her trust and for her love.

This research was supported by the National Science Foundation under grant ACI-9873339.
Contents

1 Introduction .......................................................... 11

2 Preliminaries: Decomposable Systems .......................... 15

3 Serial Systems with Stochastic Leadtimes and Markovian Demand ......................................................... 21

3.1 Introduction ....................................................... 21

3.2 Problem Formulation ............................................ 25

3.2.1 State and Control Variables ............................... 27

3.2.2 Policy Classification ......................................... 32

3.3 Finite Horizon Analysis ........................................ 35

3.4 Infinite Horizon Analysis ....................................... 40

3.4.1 Discounted Cost Criterion ................................. 42

3.4.2 Average Cost Criterion ..................................... 44

3.5 Algorithmic Issues ............................................. 52

3.5.1 Subproblem Formulation ................................ 53

3.5.2 Finite Horizon Complexity ............................... 56

3.5.3 Infinite Horizon Complexity ............................. 56

4 Serial Systems with Expediting ................................ 61

4.1 Introduction ..................................................... 61

4.2 Problem Formulation .......................................... 63

4.3 Finite Horizon Analysis ....................................... 65
List of Figures

3-1 Illustration of the system dynamics and the various definitions. .................. 30

4-1 The supermodularity assumption ......................................................... 67

4-2 Example construction to demonstrate the contradiction described in the proof of Lemma 4.3.2 .......................................................... 72

4-3 The structure of an optimal solution of the single unit subproblem for serial systems with expediting. ................................................... 75

4-4 The sequence of events in Lawson & Porteus [27] .................................. 82

4-5 Example system with artificial locations to model availability to serve customer demand within the period. .......................................... 85

6-1 An assembly system. ........................................................................... 111
Chapter 1

Introduction

Inventories are an integral part of our lives, whether personal or business. We all keep inventories of food or supplies in our homes or offices. Businesses keep inventories of raw materials, work in process or finished good inventories. Even businesses in service industries need to keep inventories of the supporting equipment and supplies in order to function. In many inventory systems, the goods go through more than one stage. These may be the different stages of a supply chain, different operations of a production facility or different facilities in a multi-plant production operation. For many decades, multi-echelon inventory control has attracted the attention of numerous researchers and there is a large body of literature on the topic. The literature includes many valuable contributions, nevertheless, in our opinion, the field is far from saturated. Many interesting problems have not yet been fully addressed. For example, an exact analysis of serial or assembly systems with stochastic leadtimes or systems with capacities have eluded the efforts of many researchers. However, the topic of inventory control remains to be of tremendous importance. A recent case study about Cisco Systems by Berinato [6] discusses how ineffective inventory management, in particular Cisco’s failure to take possible market fluctuations into account, contributed to the recent nearly disastrous performance of the company.

The traditional approach in multi-echelon inventory control has been to decompose the original problem into a series of single stage problems. This approach has been very fruitful,
starting from the seminal work of Clark & Scarf [16] up to recent contributions such as Chen & Song [12]. In this dissertation, we introduce a new methodology for the analysis of uncapacitated multi-echelon inventory systems. In particular, the methodology relies on decomposing complicated multi-echelon inventory control problems into much smaller and manageable subproblems, whose solution then helps us either characterize the structure of optimal policies for the overall problem and/or to compute optimal policies efficiently. In Chapters 3 and 4, we focus on a single unit as it travels through the supply chain and show that under various settings, the original problem is simply a series of single unit problems that are essentially decoupled. In Chapters 5 and 6, the problem again decomposes into small subproblems, this time including not a single unit, but still a small number of units or components. This new way of thinking allows us to provide an exact analysis of several important extensions to the multi-echelon inventory control literature.

We next briefly describe the different problems that are analyzed in the dissertation. A more detailed problem description and the status of the current research about each particular problem is given in corresponding chapters.

i) We demonstrate the method in detail for a serial inventory system with Markov modulated demand and Markov modulated, non-order crossing, stochastic leadtimes. We show that state dependent echelon base stock policies are optimal in this setting.

ii) Another serial inventory system, with expediting options, is analyzed next. In such a system, a stage is not restricted to order items from the next upstream stage but can place orders at stages further upstream in the supply chain, by incurring certain extra costs. We show that extended echelon base stock policies are optimal for this problem, which is a generalization of echelon base stock policies.

iii) This is followed by an analysis of serial systems with batch size constraints. We show that \((R, n)Q\) policies are optimal for this problem, which can be interpreted as echelon base stock policies modified to respect the batch size constraints.

iv) The last system we analyze is an assembly system with stochastic leadtimes. For
such a system, we show that the original problem can be decomposed into a series of subproblems, each with a single *kit of parts*. This result enhances our understanding about the optimal policies for this problem, and allows us to develop a relatively efficient algorithm for the computation of optimal policies.

The contributions of the dissertation are as follows:

i) We present a new approach to multi-echelon inventory control, relying on the decomposition of the problem into a series of smaller subproblems (single unit, single batch or single kit problems), rather than into a sequence of single stage problems. This methodology provides a new perspective on multi-echelon inventory systems. Furthermore, using this approach, once some concepts are defined, certain optimality proofs can be carried out without the need for complicated dynamic programming equations.

ii) We show the optimality of state dependent echelon base stock policies for a multi-echelon inventory system with a stochastic leadtime model, in finite horizon, and in infinite horizon with either discounted or average cost criteria.

iii) We extend the results of Chen & Song [12] to show the optimality of state dependent base stock policies for multi-echelon systems with Markov modulated demand to finite horizon problems and to infinite horizon discounted cost problems.

iv) We characterize the structure of optimal policies in serial inventory systems with expediting, under a quite general expediting cost structure, that we call supermodularity. The structure corresponds to *extended echelon base stock policies*, which is a generalization of echelon base stock policies.

v) We extend the result of Chen [11] about the optimality of \((R, nQ)\) policies for serial inventory systems with batch ordering, to finite horizon problems and to infinite horizon discounted cost problems.
vi) For the serial inventory problems that are analyzed, we provide very efficient algorithms for the calculation of the relevant quantities such as base stock levels or extended base stock levels, for both finite and infinite horizon problems.

vii) We show that for assembly systems with a stochastic leadtime model, the problem can be decomposed into smaller subproblems each of which involves a single kit of parts. We then develop an exact and relatively efficient solution method for such systems.

The rest of this dissertation has six chapters. Chapter 2 introduces the notion of decomposable systems and gives a result about such systems that is used in the subsequent chapters. The first system that is analyzed is a serial system with stochastic leadtimes and Markov modulated demand, which is the topic of Chapter 3. Chapter 4 and 5 investigate serial systems with expediting and with batch size constraints, respectively. Assembly systems are studied in Chapter 6. Finally, Chapter 7 concludes the dissertation.
Chapter 2

Preliminaries: Decomposable Systems

In this chapter, we introduce the problem of optimal control of a decomposable system and point out the decoupled nature of the resulting optimal policies. The result we provide is rather obvious, but we find it useful to state it explicitly, both for ease of exposition and also because it is a key building block for our subsequent analysis.

Following the notation in Bertsekas [7], we consider a generic stationary discrete time dynamical system of the form

\[ x_{t+1} = f(x_t, u_t, w_t), \quad t = 0, 1, \ldots, T - 1. \]  

(2.1)

Here, \( x_t \) is the state of the system at time \( t \), \( u_t \) a control to be applied at time \( t \), \( w_t \) a stochastic disturbance, and \( T \) is the time horizon. We assume that \( x_t, u_t, \) and \( w_t \) are elements of given sets \( X, U, \) and \( W, \) respectively, and that \( f : X \times U \times W \mapsto X \) is a given mapping. Finally, we assume that, given \( x_t \) and \( u_t \), the random variable \( w_t \) is conditionally independent from the past and has a given conditional distribution.

We define a policy \( \pi \) as a sequence of functions, \( \pi = (\mu_0, \mu_1, \ldots, \mu_{T-1}) \), where each \( \mu_t : X \mapsto U \) maps the state \( x \) into a control \( u = \mu_t(x) \). Let \( \Pi \) be the set of all policies.

Given an initial state \( x_0 \) and a policy \( \pi \), the sequence \( x_t \) becomes a Markov chain with a well-defined probability distribution. For any time \( t < T \) and any state \( x \in X \), we define
the cost-to-go $J^*_{t,T}(x)$ (from time $t$ until the end of the horizon) by

$$J^*_{t,T}(x) = E\left\{ \sum_{\tau=t}^{T-1} \alpha^{\tau-t} \cdot g(x_\tau, \mu_\tau(x_\tau)) \mid x_t = x \right\},$$

where $g : X \times U \mapsto [0, \infty]$ is a given cost-per-stage function and $\alpha$ is a discount factor ($0 < \alpha \leq 1$). The optimal cost-to-go function $J^*_{t,T}$ is defined by

$$J^*_{t,T}(x) = \inf_{\pi \in \Pi} J^\pi_{t,T}(x).$$

(Note that $J^\pi_{t,T}(x)$ and $J^*_{t,T}(x)$ can be infinite at certain states.) A policy $\pi^*$ is said to be optimal if

$$J^*_{t,T}(x) = J^\pi_{t,T}(x), \quad \forall \ t, \forall \ x \in X.$$

When $t = 0$, we will use the simpler notation $J^\pi_T(x)$ and $J^*_{0,T}(x)$ instead of $J^\pi_{0,T}(x)$ and $J^*_{0,T}(x)$, respectively.

We now introduce the notion of a decomposable system. Loosely speaking, this is a system consisting of multiple (countably infinite) non-interacting subsystems, that are driven by a common source of uncertainty, that evolves independently of the subsystems and is modulated by a Markov process $s_t$.

**Definition 2.0.1.** A discrete time system of the form (2.1) is said to be decomposable if it admits a representation with the following properties:

A1. The state space is a Cartesian product of the form $X = S \times \hat{X} \times \hat{X} \cdots$, so that any $x \in X$ can be represented as $x = (s, x^1, x^2, \ldots)$ with $s \in S$ and $x^i \in \hat{X}$, for every $i \geq 1$.

A2. There is a set $\bar{U}$ so that the control space $U$ is the Cartesian product of countably many copies of $\bar{U}$, that is, any $u \in U$ can be represented as $u = (u^1, u^2, \ldots)$ with $u^i \in \bar{U}$, for all $i \geq 1$.

A3. For each $t$, the conditional distribution of $w_t$ given $x_t$ and $u_t$, depends only on $s_t$.

A4. The evolution equation (2.1) for $x_t$ is of the form

$$s_{t+1} = f^s(s_t, w_t), \quad x^i_{t+1} = \hat{f}^i(x^i_t, u^i_t, w_t), \quad \forall \ i \geq 1, \forall \ t,$$
for some functions $f^s : S \times W \mapsto S$ and $\hat{f} : \hat{X} \times \hat{U} \times W \mapsto \hat{X}$.

A5. The cost function $g$ is additive, of the form

$$g(x_t, u_t) = \sum_{i=1}^{\infty} \hat{g}(x^i_t, u^i_t),$$

for some function $\hat{g} : \hat{X} \times \hat{U} \mapsto [0, \infty)$.

A6. The sets $\hat{X}$ and $W$ are countable. The sets $S$ and $\hat{U}$ are finite.

Note that the state space of a decomposable system is infinite-dimensional. But it still falls within the Borel model treated in Chapters 8 and 9 of [8], which takes care of any measurability issues that might arise.

In a decomposable system, the control variable is a vector. Therefore, any policy $\pi$ can be represented in terms of component mappings $\mu^i_t : \hat{X} \mapsto \hat{U}$, so that

$$u^i_t = \mu^i_t(x_t), \quad \forall \ i, \ t.$$

We are especially interested in those policies in which the control $u^i_t$ that affects the $i$th subsystem is chosen locally, without considering the state of the other subsystems, and using a mapping $\mu^i_t$ which is the same for all $i$.

**Definition 2.0.2.** A policy $\pi$ for a decomposable system is said to be *decoupled* if it can be represented in terms of mappings $\hat{\mu}_t : S \times \hat{X} \mapsto \hat{U}$, so that

$$u^i_t = \hat{\mu}_t(s_t, x^i_t), \quad \forall \ i, \ t.$$

For a decomposable system, the various state components $x^1_t, x^2_t, \ldots$ evolve independently of each other, the only coupling arising through the exogenous processes $s_t$ and $u_t$. Since the costs are also additive, it should be clear that each subsystem can be controlled separately (that is, using a decoupled policy) without any loss of optimality. Furthermore, since all subsystems are identical, the same mappings $\hat{\mu}_t$ can be used in each subsystem. The required notation and a formal statement is provided below.
Each subsystem $i$ defines a subproblem, with dynamics

$$s_{t+1} = f^s(s_t, w_t), \quad x_{t+1}^i = f^i(x_t^i, u_t^i, w_t),$$

and costs per stage $g(x_t^i, u_t^i)$. A policy $\hat{\pi}$ for a subproblem is of the form $\hat{\pi} = (\hat{\mu}_0, \hat{\mu}_1, \ldots, \hat{\mu}_{T-1})$, where each $\hat{\mu}_t$ is a mapping from $S \times \hat{X}$ into $\hat{U}$:

$$u_t^i = \hat{\mu}_t(s, x_t^i).$$

Let

$$J^*_{t,T}(s, x^i)$$

be the optimal cost-to-go function for a subsystem that starts at time $t$ from state $(s, x^i)$ and evolves until the end of the horizon $T$. Note that this function is the same for all $i$, because we have assumed the subsystems to have identical dynamics and cost functions. Furthermore, since the control set $\hat{U}$ is finite, an optimal policy is guaranteed to exist.

**Lemma 2.0.1.** Consider a decomposable system.

1) For any $x = (s, x^1, x^2, \ldots) \in X$ and any $t \leq T$, we have

$$J^*_{t,T}(x) = \sum_{i=1}^{\infty} J^*_{t,T}(s, x^i).$$

2) There exists a decoupled policy $\pi^*$ which is optimal, that is,

$$J^*_{t,T}(x) = J^*_{t,T}(x), \quad \forall \ t, \ \forall \ x \in X.$$

3) For any $s$, $x^i$, and any remaining time $k$, let $\hat{U}^*_k(s, x^i) \subset \hat{U}$, be the set of all decisions that are optimal for a subproblem, if the state of the subproblem at time $T-k$ is $(s, x^i)$. A policy $\pi = \{\mu_t^i\}$ is optimal if and only if for every $i, t$, and any $x = (s, x^1, x^2, \ldots) \in X$ for which $J^*_{t,T}(x) < \infty$, we have

$$\mu_t^i(x) \in \hat{U}^*_{T-t-i}(s, x^i).$$
The proof of the above result is straightforward and is omitted. Suffice to say that we can pick an optimal policy for the subproblem and replicate it for all subsystems to obtain a decoupled and optimal policy. The last part of the lemma simply states that for any given $x$ and $t$, a decision vector $u_t = (u_1^t, u_2^t, \ldots)$ is optimal if and only if each component $u_i^t$ of the decision is optimal for the $i$th subsystem viewed in isolation (except of course if the cost-to-go $J_{t,T}^*(x)$ is infinite, in which case all possible decisions are optimal). Let us also remark that the sets $\hat{U}_k^*(s, x^t)$ of optimal decisions only depend on the remaining time $k$, but do not depend on the value of $T$. This is an immediate consequence of the stationarity of the problem.
Chapter 3

Serial Systems with Stochastic Leadtimes and Markovian Demand

3.1 Introduction

This chapter deals with serial (multi-echelon) inventory systems of the following type. There are $M$ stages. Stage 1 receives stock from stage 2, stage 2 from stage 3, etc., and stage $M$ receives stock from an outside supplier with ample stock. Demands originate at stage 1, and unfilled demand is backlogged. There are holding, ordering, and backorder costs, and a central controller has the objective of minimizing these costs in the appropriate time frame.

Clark & Scarf [16] characterize optimal policies for multi-echelon inventory systems. They considered finite horizon problems and proved that base stock policies are optimal when the demands are independent identically distributed (i.i.d.) and the leadtimes between stages are deterministic. Their proof technique involves a decomposition of the multi-echelon problem into a series of single stage problems. This general approach guided much of the subsequent literature, and many extensions were obtained using the same stage-by-stage decomposition. In particular, Federgruen & Zipkin [20] extended the results to the stationary infinite horizon setting, and Chen & Zheng [14] provided an alternative proof that is also valid in continuous time. All of these papers assumed i.i.d. demands and deterministic leadtimes.
Our work relaxes two important restrictions of the model in Clark & Scarf [16], namely the i.i.d. demand and deterministic leadtime assumptions. We assume instead that demands and leadtimes are stochastic and are affected (modulated) by an exogenous Markov process. Such a model can capture many phenomena such as seasonalities, exchange rate variations, fluctuating market conditions, demand forecasts, etc. Our assumption that the demand at each time depends on the state of a modulating Markov chain is certainly not new to the inventory control literature. Song & Zipkin [33], Beyer & Sethi [9], Sethi & Cheng [31], Cheng & Sethi [15] all investigate a single stage system with such a demand model and prove the optimality of state dependent \((s, S)\) policies under different time horizon assumptions and with and without backlogging assumptions. Song & Zipkin [32] and Song & Zipkin [34] evaluate the performance of base stock policies in serial inventory systems, with state-independent and state-dependent policies, respectively. Chen & Song [12] have recently shown the optimality of state dependent echelon base stock policies for serial systems with Markov modulated demand and deterministic leadtimes, under an infinite horizon average cost criterion. To our knowledge, this reference is the only existing work that characterizes optimal policies in a serial system under a Markov modulated demand assumption.

The study of stochastic leadtimes in inventory control dates back to the early days of the literature. Hadley & Whitin [25] investigated the subject for a single stage problem, and suggested that two seemingly contradictory assumptions are needed, namely, that orders do not cross each other and that leadtimes are independent. Kaplan [26] came up with a simple model of stochastic leadtimes that prevents order crossing, while keeping the probability that an outstanding order arrives in the next time period independent of the current status of other outstanding orders. He showed that the deterministic leadtime results carry over to his model of stochastic leadtimes. Nahmias [28] and Ehrhardt [19] streamlined Kaplan's results. Zipkin [38] investigated stochastic leadtimes in continuous-time single stage inventory models. Svoronos & Zipkin [35] evaluated one-for-one replenishment policies in the serial system setting. However, we are not aware of any optimality results for serial systems under any type of stochastic leadtimes. Our stochastic leadtime model incorporates the same two im-
portant features of Kaplan's stochastic leadtime model, i.e., the non-order crossing property and the independence from the current status of other outstanding orders. In our model, just like in Kaplan's, an exogeneous random variable determines which outstanding orders are going to arrive at a given stage. However, we additionally allow the stochastic leadtimes to depend on the state of a modulating Markov chain, and we also allow for dependencies between the leadtime random variables corresponding to different stages in the system.

What allows us to handle such a general demand and leadtime model is a new approach to the uncapacitated serial inventory problem. The standard approach is a decomposition into a series of single stage problems. Our approach instead relies on a decomposition into a series of unit-customer pairs. Consider a single unit and a single customer. Assume that the distribution of time until the customer arrives to the system is known and the goal is to move the unit through the system in a way that optimizes the holding versus backorder cost trade-off. Since only a single unit and a single customer are present, this problem is much simpler than the original one. We show that under the assumptions of this chapter, the original problem is equivalent to a series of decoupled single-unit single-customer problems. This approach allows us to handle several extensions to the standard model in a simple manner. In particular, we can bypass tedious inductive arguments based on complicated dynamic programming equations. We use a simple qualitative argument to establish a monotonicity property of optimal policies for the single-unit single-customer problem. For finite horizon problems, the optimality of echelon base stock policies is an immediate corollary. The same is true for infinite horizon problems, once some required limiting arguments are carried out.

Axäter [2] made an observation, that is related to our method. In particular, he observed that in a distribution system with a single depot and multiple retailers that follow base stock policies with base stock levels $S_i$, any unit ordered by retailer $i$ is used to fill the $S_i$th demand following this order. He then matches this unit with that demand and evaluates the expected cost for this unit and "its demand". Using this approach, he develops an efficient method to evaluate the cost of a given base stock policy for a two-echelon distribution system in continuous time with Poisson demand and with the infinite horizon, average cost
criterion. In Axsäter [3] he extends his results to batch ordering policies and in Axsäter [4] he investigates the system with periodic review, using the virtual allocation rule suggested by Graves [24] and a base stock policy. However, this series of papers by Axsäter, even though they provided a good computational approach to evaluate the cost of a base stock policy for a two-echelon distribution system with the average cost criterion, did not go much beyond that. On the other hand, as we show in this dissertation, the insight that Axsäter had, namely that units and demands can be matched, when used properly and rigorously, leads to the idea of decomposing the problem into single unit single customer problems, which is a quite powerful technique for developing optimality results and algorithms for many multi-echelon systems.

A related work is the Masters thesis by Achy-Brou [1] (supervised by the same thesis supervisor, concurrently with this work) who studies the single-unit single-customer subproblem for the case of i.i.d. demands and deterministic leadtimes and a discounted cost criterion. This work formulates the subproblem as a dynamic program, describes and implements the associated dynamic programming algorithm, analyzes structural properties of the solution, and discusses the relationship between the subproblem and basestock policies in the overall inventory system.

Besides providing a simple proof technique, the single-unit single-customer problem leads to simple and efficient algorithms for calculating the base stock levels. Even for some special cases of our model for which computational methods are already available, our algorithms are more intuitive, easier to implement, and at least as efficient.

To recapitulate, the contributions of this chapter are the following:

i) We show that state dependent echelon base stock policies are optimal for a multi-echelon inventory system with stochastic leadtimes, in finite horizon and in infinite horizon with either discounted or average cost criteria.

ii) We extend the results of Chen & Song [12] to show the optimality of state dependent base stock policies for multi-echelon systems with Markov modulated demand to finite horizon problems and to infinite horizon discounted cost problems.
iii) We provide an efficient algorithm for the calculation of the base stock levels for both finite and infinite horizon problems.

iv) We present a new approach to uncapacitated multi-echelon inventory control, relying on a decomposition into single-unit single-customer problems, rather than a sequence of single stage problems.

The rest of the chapter has four sections. Section 3.2 provides a mathematical formulation of the problem and the necessary notation. Sections 3.3 and 3.4 contain the results and their proofs for finite and infinite horizon versions of the problem, respectively. Section 3.5 discusses the resulting algorithms for evaluating the base stock levels.

3.2 Problem Formulation

We consider a single-item serial inventory system consisting of $M$ stages, indexed by $1, \ldots, M$. Customer demand can only be satisfied by units at stage 1. Any demand that is not immediately satisfied is backlogged. The inventory at stage $i$ ($i = 1, \ldots, M - 1$) is replenished by placing an order for units stored at stage $i + 1$. Stage $M$ receives replenishments from an outside supplier with unlimited stock. For notational simplicity, we label the outside supplier as stage $M + 1$. We assume that the system is periodically reviewed and, therefore, a discrete-time model can be employed.

To describe the evolution of the system, we need to specify the sources of uncertainty, the statistics of the demand, and the statistics of the leadtimes for the various orders.

(a) Markovian exogenous uncertainty: We assume that the customer demands and the order leadtimes are influenced by an exogeneous finite-state Markov chain $s_t$, assumed to be time-homogeneous and ergodic (irreducible and aperiodic).

(b) Demand model: The (nonnegative integer) demand $d_t$ during period $t$ is assumed to be Markov modulated. In particular, the probability distribution of $d_t$ depends on the state $s_t$ of the exogeneous Markov chain and, conditioned on that state, is independent
of the history of the process until now. We also assume that $E[d_t \mid s_t = s] < \infty$, for every $s \in S$.

(c) **Leadtime model**: We assume that the leadtime between stage $i + 1$ and stage $i$ is upper bounded by some integer $l_i$. We assume that the probability that an outstanding order arrives during the current period depends only on the amount of time since the order was placed, the exogenous state $s_t$, and the destination stage $i$ and, given these, it is conditionally independent of the history of the process until now. Finally, we assume that orders cannot overtake each other: an order cannot arrive at its destination before an earlier order does.

The leadtime model introduced above includes the obvious special case of deterministic leadtimes. It also includes a stochastic model of the type we describe next, and which extends the model of Kaplan [26]. At each time period $t$, there is a random variable $\rho_t^i$ that determines which outstanding orders will arrive at stage $i$. More precisely, an outstanding order will be delivered at stage $i$ if and only if it was placed $\rho_t^i$ or more time units ago. Note that such a mechanism ensures that orders cannot overtake each other. Let $\rho_t = (\rho_t^1, \rho_t^2, \ldots, \rho_t^M)$ be the vector of leadtime random variables associated with the various stages. We assume that the statistics of $\rho_t$ are given in terms of a conditional probability distribution, given $s_t$. Notice that such a model allows for dependencies between the leadtime random variables corresponding to the same period but different stages. Furthermore, it can also capture intertemporal dependencies through the dynamics of $s_t$.

The cost structure that we use is fairly standard and consists of linear holding, ordering, and backorder costs. Because of this linearity, the cost can be broken down into components that are ascribed to particular units or customers, which is one of the requirements for a decomposable system. In more detail, we assume:

(a) For each stage $i$, there is an inventory holding cost rate $h_i$ that gets charged at each time period to each unit at that stage. We assume that the holding cost rate $h_{M+1}$ at the external supplier is zero. For concreteness, we also assume that after a unit is
ordered and during its leadtime, the holding cost rate charged for this unit is the rate corresponding to the destination echelon.

(b) For each stage $i$, there is a cost $c_i$ for initiating the shipment of a unit from stage $i + 1$ to stage $i$.

(c) There is a backorder cost rate $b$ which is charged at each time step for each unit of backlogged demand.

We assume that the holding cost and backorder cost parameters are positive and that the shipping cost is non-negative.

The detail-oriented reader may have noticed that the model has not been specified in full detail: we would still need to describe the relative timing of observing the demand, fulfilling the demand, placing orders, receiving orders, and charging the costs. Different choices with respect to these details result, in general, to slightly different optimal costs and policies. Whatever specific choices are made, the arguments used for our subsequent results remain unaffected. For specificity, however, we make one assumption about delivery of units to customers. We assume that if a customer arrives during period $t$, a decision to give a unit to the customer can only be made at time $t + 1$, or later.

3.2.1 State and Control Variables

In Chapter 2, we described a generic discrete time dynamic system. In this subsection, we define our choices for the state, control, and disturbance variables for the inventory control system being studied.

The traditional approach would be the following. The state would consist of a vector whose components are the number of units at each stage, the number of units that have been released by stage $i$ and have been in transit for $k$ time units (one component for each $i$-$k$ pair), the size of the backlogged demand, and the state of the modulating Markov chain. The control would be the number of units to be released from each stage to the next, or to be given to already arrived customers. The demand in a given period, the various
random variables associated with the random leadtimes (e.g., the random vector $\rho_t$ in our earlier example), and the transition of the modulating chain $s_t$ would constitute the random disturbance. Obviously, such a choice is sufficient for optimization purposes, since one does not need to distinguish between units that are at the same stage or between units that have been in transit for the same amount of time. However, we approach the problem differently. We treat each individual unit and each individual customer as distinct objects and then show that this results in a decomposable problem, with each unit-customer pair viewed as a separate subsystem. Towards this goal, we start by associating a unique label with each unit and customer.

At any given time, there will be a number of units at each stage or on order between two given stages. In addition, conceptually, we have a countably infinite number of units at the outside supplier, which we call stage $M + 1$. We will now introduce a set of conceptual locations in the system that can be used to describe where a unit is found and, if it is part of an outstanding order, how long ago it was ordered.

**Definition 3.2.1.** The location of a unit: First, each of the actual stages in the system will constitute a location. Next, we insert $l_i - 1$ artificial locations between the locations corresponding to stages $i$ and $i + 1$, for $i = 1, \ldots, M$, in order to model the units in transit between these two stages. If a unit is part of an order between stages $i$ and $i + 1$ that has been outstanding for $k$ periods, $1 \leq k \leq l_i - 1$, then it will be in the $k^{th}$ location between stages $i + 1$ and $i$. Finally, for any unit that has been given to a customer, we define its location to be 0. Thus, the set of possible locations is $\{0, 1, \ldots, N + 1\}$, where $N = \sum_{i=1}^{M} l_i$. We index the locations starting from location 0. Location 1 corresponds to stage 1. Location 2 corresponds to units that have been released $l_1 - 1$ times ago from stage 2. Location $l_1$ corresponds to units that have been released from stage 2 one time step earlier. Location $l_1 + 1$ corresponds to stage 2, etc. Location $N + 1$ corresponds to the outside supplier (stage $M + 1$). For any unit, we define its location as the index of the location at which the unit can be found. For example, in Figure 3-1(a), unit 5 is in location 2 at time $t$, which means that this unit has been released from stage 2 (location 4) at time $t - 2$.  

28
We index the countably infinite pool of units by the nonnegative integers. We assume that the indexing is chosen at time 0 in increasing order of their location, breaking ties arbitrarily.

Let us now turn to the customer side of the model, which we describe using a countably infinite pool of past and potential future customers, with each such customer treated as a distinct object. At any given time, there is a finite number of customers that have arrived and whose demand is either satisfied or backlogged. In addition, conceptually, there is a countably infinite number of potential customers that may arrive to the system at a future period. Consider the system at time 0. Let \( k \) be the number of customers that have arrived, whose demand is already satisfied. We index them as customers \( 1, \ldots, k \), in any arbitrary order. Let \( l \) be the number of customer that have arrived, whose demand is backlogged. We index them as customers \( k + 1, k + 2, \ldots, k + l \) in any order. The remaining (countably infinite) customers are assigned indices \( k + l + 1, k + l + 2, \ldots \), in order of their arrival times to the system, breaking ties arbitrarily, starting with the earliest arrival time. Of course, we do not know the exact arrival times of future customers, but we can conceptually talk about a “next customer,” a “second to next customer,” etc. This way, we index the past and potential future customers. We now define a quantity that we call “the distance of a customer.”

**Definition 3.2.2.** The distance of a customer: Suppose that at time \( t \) a customer \( i \) has arrived and its demand is satisfied. We define the distance of such a customer to be \( 0 \). Suppose that the customer has arrived but its demand is backlogged. Then, we define the distance of the customer to be \( 1 \). If on the other hand, customer \( i \) has not yet arrived but customers \( 1, 2, \ldots, m \) have arrived, then the distance of customer \( i \) at time \( t \) is defined to be \( i - m + 1 \). In particular, a customer whose distance at time \( t \) is \( k \), will have arrived by the end of the current period if and only if \( d_t \geq k - 1 \). For example, in Figure 3-1(b), customer 5 has a distance of 3 at time \( t \). Customer 5 will have arrived at time \( t + 1 \) if and only if 2 or more customers (customers 4 and 5) arrive during period \( t \).

Now that we have labeled every unit and every customer, we can treat them as distin-
Figure 3-1: Illustration of the system dynamics and the various definitions. Consider a system with $M = 2$ stages and deterministic leadtimes $l_1 = l_2 = 3$. This results in a total of $N + 2 = 8$ locations.

Part (a) shows a configuration at some time $t$, in which there are 7 units in the system and an infinity of additional units at the outside supplier. Units 1 and 2 have already been given to customers, in previous periods.

Part (b) shows a customer configuration. Customers 1 and 2 arrived in previous periods and have already been given a unit. Customer 3 has arrived but its demand is backlogged. Customers 4, 5, etc., have not yet arrived.

Suppose that the following decisions are made: $u_t^4 = u_t^7 = 1$ and $u_t^i = 0$ for $i \neq 4, 7$. Suppose also that the new demand $d_t$ turns out to be 2.

(c) Since $u_t^2 = 0$ and $u_t^4 = 1$, unit 4 is given to customer 3. Hence this unit moves to location 0. Also, the released unit 7 moves one location.

(d) Customer 3 obtains a unit and moves to distance 0. The two newly arrived customers (4 and 5) move to distance 1.
guishable objects and, furthermore, we can think of unit $i$ and customer $i$ as forming a pair. This pairing is established at time 0, when indices are assigned, taking into account the initial locations and distances, and is to be maintained throughout the planning horizon.

We are now ready to specify the state and control variables for the problem of interest. For each unit-customer pair $i, i \in \mathbb{N}$, we have a vector $(z^i_t, y^i_t)$, with $z^i_t \in Z = \{0, 1, \ldots, N + 1\}$ and $y^i_t \in Y = \mathbb{N}_0$, where $z^i_t$ is the location of unit $i$ at time $t$, and $y^i_t$ is the distance of customer $i$ at time $t$. The state of the system consists of a countably infinite number of such vectors, one for each unit-customer pair, and finally a variable $s_t \in S$ for the state of the modulating Markov chain, i.e.,

$$x_t = \{s_t, (z^1_t, y^1_t), (z^2_t, y^2_t), \ldots \}.$$  

The control vector is an infinite binary sequence, $u_t = (u^1_t, u^2_t, \ldots)$, where the $i^{th}$ component $u^i_t$ corresponds to a “release” or “hold” decision for the $i^{th}$ unit. If unit $i$ is at an artificial location (in between stages), then the decision $u^i_t$ does not affect the movement of the unit (which is solely governed by the randomness in the leadtimes). If unit $i$ is at location 0, it is already delivered to a customer and is unaffected by $u^i_t$. If unit $i$ is at a non-artificial location (that is, at one of the stages), then $u^i_t = 0$ corresponds to holding the unit at its current location, and $u^i_t = 1$ corresponds to releasing it. Finally, if unit $i$ is at stage 1, a decision $u^i_t = 1$ releases this unit so that it can be given to a customer. In case the number $k$ of units released from stage 1 is larger than the number $m$ of customers whose demand is backlogged, only $m$ of these units are given to customers (i.e., move to location 0), and the remaining $k - m$ units stay at location 1. Otherwise, all $k$ units are given to customers. The rules about which units are given to customers and which customers receive a unit are the following: If unit $i$ is released to be given to a customer and customer $i$'s demand is backlogged, then unit $i$ is given to customer $i$. After all such matchings are done, if there are extra units and customers, the units and customers with the lowest indices are chosen until one side is empty.

Finally, in terms of the general model discussed in Chapter 2, the random disturbance $w_t$ at time $t$ consists of the demand $d_t$, random variables that model the uncertainty in
the leadtimes (e.g., the vector $\rho_t$ of leadtime random variables in our earlier example), and whatever additional exogenous randomness is needed to drive the Markov chain $s_t$.

Clearly, our choice for the state, control, and disturbance variables is a sufficient description of the overall system, albeit not the most compact one.

### 3.2.2 Policy Classification

We now define various classes of policies, state-dependent echelon base stock policies being one particular class. In the next section, we will show that the search for an optimal policy can be restricted to any one of these policy classes, without sacrificing performance.

As a first step, we define a class of states that we call monotonic states.

**Definition 3.2.3.** A state $x_t = \{s_t, (z^1_t, y^1_t), (z^2_t, y^2_t), \ldots \}$ is called monotonic if and only if the unit locations are monotonic functions of the unit labels, that is,

$$i < j \Rightarrow z^i_t \leq z^j_t.$$

Note that $x_0$ is always a monotonic state, by construction. The state $x_t$ shown in Figure 3-1 is monotonic, but the state $x_{t+1}$ is not.

**Definition 3.2.4.** Here, we define four classes of policies. Let $x_t = \{s_t, (z^1_t, y^1_t), (z^2_t, y^2_t), \ldots \}$ and $u_t = (u^1_t, u^2_t, \ldots)$ be the state and control variables as defined in Subsection 3.2.1. Consider a policy $\pi = \{\mu^i_t\}$, where each $\mu^i_t : X \mapsto \hat{U} = \{0, 1\}$ prescribes the control for the $i$th unit at time $t$, according to $u^i_t = \mu^i_t(x_t)$.

**Monotonic policies:** Given a policy $\pi$, let $\mu_t = (\mu^1_t, \mu^2_t, \ldots)$ be the part of the policy that applies at time $t$. The policy $\pi$ is monotonic if it guarantees that a monotonic state $x_t$ always results in a next state $x_{t+1}$ that is monotonic. That is, for every $t$, if $x_t$ is monotonic, then $x_{t+1} = f(x_t, \mu_t(x_t), w_t)$ is monotonic, for every possible realization of $w_t$. Intuitively, a policy is monotonic if and only if units can never overtake each other. (The policy applied at time $t$ in the example of Figure 3-1 is not monotonic, since unit 7 overtakes unit 6 resulting in a nonmonotonic state at time $t + 1$.)
Committed policies: These are policies such that if \( z_i^t = 1 \) and \( y_i^t > 1 \), then \( \mu_i^t(x_t) = 0 \). In words, if unit \( i \) is at the last stage, it can only be released to satisfy the corresponding customer \( i \), and this can only happen if the customer has already arrived and is backlogged. (The policy applied at time \( t \) in the example of Figure 3-1 is not committed because unit 4 is released, whereas customer 4 has not yet arrived and hence unit 4 is given to customer 3.)

Decoupled policies: This is the same as Definition 2.0.2 in the preceding chapter. We call a policy decoupled, if it can be represented in terms of mappings \( \tilde{\mu}_t : S \times Z \times Y \rightarrow \tilde{U} \), so that

\[
\mu_i^t(x) = \tilde{\mu}_t(s_t, z_t^i, y_t^i), \quad \forall \ i, \ t.
\]

In words, a decoupled policy is a policy where the decision of whether or not to release a unit from its current location can be written as a function of the state of the modulating Markov chain, the location of the unit, and the distance of the corresponding customer. Moreover, the function is the same for every unit.

State-dependent echelon base stock policies: A policy is a state-dependent echelon base stock policy if for every \( t \), every state \( x \), every location \( z \in Z \) that corresponds to a stage of the original system, and every \( s \in S \), there exists a nonnegative integer \( S_t^{z-1}(s) \) such that

\[
\begin{align*}
\text{\# units released from } z & \quad \{i \mid z^i = z, \ \mu_i^t(x) = 1\} \\
= \min & \quad \left\{ \begin{array}{l}
\text{basestock level} \\
\left( S_t^{z-1}(s) \right) - \left( \left| \left\{ i \mid 1 \leq z^i \leq z - 1 \right\} \right| - \left| \left\{ i \mid y_t^i = 1 \right\} \right| \right)
\end{array} \right\} \\
\text{echelon inventory at } z-1 & \quad \{i \mid z^i = z\} \\
\text{\# units at } z & \quad \end{align*}
\]

We are using here the notation \( |A| \) to denote the cardinality of a set \( A \), and the notation \( (a)^+ \) to denote \( \max\{0, a\} \). In words, such a policy operates as follows: For
every location \( z \), it calculates the echelon inventory at \( z - 1 \) and releases enough units (to the extent that they are available) to raise this number to a target value \( S^{z-1}_t(s) \). The echelon inventory at \( z - 1 \) is the number of available units at locations 1, \ldots, z - 1 minus the backlogged demand.

Note that if the initial state is a monotonic state (which we always assume to be the case) and one uses a monotonic policy, the state of the system at any time in the planning horizon will be monotonic.

We say that a set of policies \( \Pi_0 \) is optimal if

\[
\inf_{\pi \in \Pi_0} J^\pi_{t,T}(x) = J^*_t(x)
\]

for all \( t \) and all monotonic states \( x \). Non-monotonic states are not a concern, because we can reindex the units and obtain an equivalent monotonic state. Once we have a policy that is optimal starting from any monotonic state, an optimal policy starting from non-monotonic states is readily available.

In the next section, we show that the intersection of monotonic, committed, and decoupled policies is optimal. The following proposition will then imply that state-dependent echelon base stock policies are optimal, which is the main result of this chapter.

**Proposition 3.2.1.** A monotonic, committed, and decoupled policy is a state dependent echelon base stock policy.

**Proof.** Suppose that \( \pi = \{\mu^i_t\} \) is a monotonic, committed, and decoupled policy. By the definition of decoupled policies, there exist mappings \( \hat{\mu}_t : S \times Z \times Y \rightarrow \hat{U} \), so that \( \mu^i_t(x^i_t) = \hat{\mu}_t(s, z^i_t, y^i_t) \) for all \( i \) and \( t \). Since \( \pi \) is monotonic, \( \hat{\mu}_t(s, z, y) \) has to be non-increasing in \( y \) over the set \( \{y \mid y \geq 1\} \), for every \( t, s, \) and \( z \). Indeed, suppose that this is not the case. Then, there exists some \( (t, s, z, y) \) such that for some positive \( y' < y \), \( \hat{\mu}_t(s, z, y) = 1 \) but \( \hat{\mu}_t(s, z, y') = 0 \). Consider a monotonic state \( x_t \) for the overall problem such that \( x^i_t = (z, y), \ \ x^i_t = (z, y') \), and \( J^{\pi}_{t,T}(x_t) < \infty \). (Finiteness of \( J^{\pi}_{t,T}(x_t) \) is the same as requiring that only finitely many units have been released from the outside supplier, i.e., location \( N + 1 \). Note
also that since \( y' < y \) and the state \( x_t \) is monotonic, we must have \( j < i \). In this case, under policy \( \pi \), the higher-indexed unit \( i \) will move ahead of unit \( j \), and the new state will not be monotonic. Therefore, \( \pi \) cannot be monotonic, which is a contradiction.

Fix some \( s \in S, t, \) and some location \( z \) that corresponds to a stage (non-artificial location) of the system. Since \( \hat{\mu}_t(s, z, y) \) is non-increasing in \( y \) (for \( y \geq 1 \)), there exists some \( y^* \in \mathbb{N}_0 \cup \{\infty\} \) such that, for positive \( y \) we have \( \hat{\mu}_t(s, z, y) = 1 \) if and only if \( y \leq y^* \). Out of all units at location \( z \), the policy \( \pi \) releases those units that are paired with customers that have already arrived (\( y \leq 1 \)) and with the next \( y^* - 1 \) customers that have not yet arrived (\( y = 2, \ldots, y^* \)). Equivalently, it tries to set the echelon inventory level at \( z - 1 \) (number of units in locations 0, 1, \ldots, \( z - 1 \), minus the number of customers that have arrived) to \( y^* - 1 \) (to the extent that the required units are available at location \( z \)). But this is identical to the decision under a state-dependent echelon base stock policy with \( S_t^{z-1}(s) = y^* - 1 \). \( \square \)

### 3.3 Finite Horizon Analysis

In this section, the finite horizon model is analyzed and the optimality of state dependent echelon base stock policies is established. We establish the result by proving a series of propositions, each of which shows that a certain set of policies is optimal.

**Proposition 3.3.1.** The set of monotonic policies is optimal.

*Proof.* By our assumptions on the leadtimes, units do not overtake each other while in transit. In addition, whenever some units are released from a particular stage, it does not matter which particular units are released, since all units are identical. Hence, given any policy \( \pi \) and monotonic initial state \( x \), there exists another policy \( \pi' \), that at any time and for each location, releases the exact same number of units, but always chooses among the available units those that have the smallest indices. The new policy \( \pi' \) has the same expected cost, and is also monotonic. (For example, in Figure 3-1, instead of setting \( u_t^4 \) and \( u_t^7 \) to 1, we could have set \( u_t^3 \) and \( u_t^6 \) to 1. This would correspond to a monotonic policy that results in an equivalent – in terms of cost – next state.) \( \square \)
Note that a monotonic policy is not necessarily committed. This is because there are no restrictions on the actions of a monotonic policy when the initial state is not monotonic. Even when the initial state is monotonic, a monotonic policy is not necessarily committed. For an example, suppose that units 1 and 2 are available at location 1, no customer has arrived, and the policy chooses \( u^1 = 0, u^2 = 1 \). Neither unit moves to location 0, and monotonicity is preserved. Still, the movement of units is the same as under the alternative choices \( u^1 = 0 \) and \( u^2 = 0 \), which is consistent with the condition required by the definition of committed policies. More generally, it is seen that whenever a monotonic policy is not committed, the result of the chosen decisions is the same as if a monotonic and committed policy is followed. This argument is formalized in the next result.

**Proposition 3.3.2.** The set of monotonic and committed policies is optimal.

*Proof.* Consider a monotonic initial state. Under a monotonic policy, units are released from stage 1 to arrived customers in the order of the units' indices. Thus, for unit \( i \) to be given to a customer, the units \( 1, \ldots, i - 1 \) must have been given to other customers, now or in the past. (Otherwise unit \( i \) would move to location 0, while a lower-indexed unit would stay at location 1, which would contradict monotonicity.) Therefore, unit \( i \) is given to some customer only if customer \( i \) has already arrived. Hence the state evolution is the same as under an additional restriction that a unit \( i \) can be released from location 1 only when customer \( i \) has arrived, which corresponds to a monotonic and committed policy. Therefore, we can modify a monotonic policy so that it is also committed, without increasing the costs. \( \square \)

**Proposition 3.3.3.** The set of committed policies is optimal.

*Proof.* We have just shown that monotonic and committed policies are optimal. Therefore, the class of committed policies (which is no smaller) is also optimal. \( \square \)

The original form of the problem, with all possible policies allowed is not decomposable, because of the coupling that arises when units are delivered (from location 1) to customers. Indeed, if there are two units at location 1, and if only one customer has arrived, we may set
\( u^i_t = 1 \) for each unit, but only one of them will move to location 0. Thus, the movement (that is, the dynamics) of the unit that was not delivered is affected by the presence of the unit that was delivered, contradicting assumption A4 for decomposable systems in Definition 2.0.1.

On the other hand, once we restrict to the set of committed policies, this coupling is eliminated, and the system becomes decomposable. A unit \( i \) can be released from location 1 only if customer \( i \) has arrived. In that case, whether the unit will move or not to location 0 is not affected by the state or decision variables corresponding to other unit-customer pairs.

**Proposition 3.3.4.** The set of committed and decoupled policies is optimal and

\[
J^*_T(x) = \sum_{i=1}^{\infty} J^*_T(s, z^i, y^i)
\]

for every monotonic state \( x \).

**Proof.** Consider the overall problem, together with an additional restriction to committed policies. Such a restriction can be represented as a change in the dynamics of the system, instead of a restriction on the policy space. In particular, suppose that an unit \( i \) is at location 1 at some time \( t \). Then, restricting \( \mu^i_t(x_t) \) to be 0 when \( y^i_t > 1 \) has the same effect as defining the dynamics of the system so that \( z^i_{t+1} = 0 \) if and only if either \( z^i_t = 0 \) or \( y^i_t = z^i_t = 1 \) and \( \mu^i_t(x_t) = 1 \). Note that after such a modification of the dynamics, the system is decomposable and Lemma 2.0.1 implies the optimality of decoupled policies.

Let us also note at this point that for a decomposable problem the existence of an optimal policy is guaranteed, because there exists a decoupled policy which is optimal for the overall problem.

Since the system with the redefined dynamics as in the previous proof is decomposable, it consists of a series of subproblems as in Section 2. From now on, whenever we refer to a "subproblem," we will mean the single-unit single-customer problem in which the dynamics have been redefined; that is, \( z^i_{t+1} = 0 \) if and only if either \( z^i_t = 0 \) or \( y^i_t = z^i_t = 1 \) and \( \mu^i_t(x_t) = 1 \). Thus, a restriction to committed policies will be implied on the corresponding overall problem.

37
**Definition 3.3.1.** For any $k, s \in S, z \in Z, y \in Y$, let $\hat{U}_k^*(s, z, y) \subset \{0, 1\}$ be the set of all decisions that are optimal if a subproblem is found at state $(s, z, y)$ at time $t = T - k$, that is $k$ time steps before the end of the horizon.

The next lemma establishes that if an optimal policy dictates that a certain unit must released when the distance of the corresponding customer is $y$, then it will also be optimal to release the unit when the distance of the corresponding customer is smaller than $y$. This is intuitive because as the customer comes closer, there is more urgency to move the unit towards stage 1.

**Lemma 3.3.1.** For every $(k, s, z, y)$, if $\hat{U}_k^*(s, z, y) = \{1\}$, then $1 \in \hat{U}_k^*(s, z, y')$, for every $y' < y$.

**Proof.** If $z = 0$, then $U_k^*(s, z, y) = \{0, 1\}$, and there is nothing to prove. We therefore assume that $z > 0$. Suppose that there exist some $(k, s, z, y)$ and $(k, s, z, y')$, with $y' < y$, such that $\hat{U}_k^*(s, z, y) = \{1\}$ and $\hat{U}_k^*(s, z, y') = \{0\}$. Let $t = T - k$. Consider a monotonic state $x_t$ for the overall problem such that $x_t^i = (z, y)$, $x_t^j = (z, y')$, and $J_{\hat{U}_T}(x_t) < \infty$. Note that since $y' < y$, we must have $j < i$. Then according to Lemma 2.0.1(3), the decision $u_t$ under any optimal policy must satisfy $u_t^i = 1$ and $u_t^j = 0$. This means that the higher-indexed unit $i$ will move ahead of unit $j$, and the new state will not be monotonic. Therefore, no monotonic policy can be optimal, which contradicts Proposition 3.3.1. \qed

**Proposition 3.3.5.** The set of monotonic, committed, and decoupled policies is optimal.

**Proof.** Let us fix $t, s$, and $z$. Let $k = T - t$ be the number of remaining time steps. We consider three cases.

(a) If there are infinitely many $y$ for which $\hat{U}_k^*(s, z, y) = \{1\}$, then by Lemma 3.3.1, $1 \in \hat{U}_k^*(s, z, y)$ for every $y$, and we let $\mu_t(s, z, y) = 1$ for all $y$.

(b) If there is no $y$ for which $\hat{U}_k^*(s, z, y) = \{1\}$, we let $\mu_t(s, z, y) = 0$ for all $y$.  

38
(c) If there is a largest $y$ for which $\hat{U}_k^*(s, z, y) = \{1\}$, call it $y^*$, then we have $0 \in \hat{U}_k^*(s, z, y)$, for every $y > y^*$, and by Lemma 3.3.1, $1 \in \hat{U}_k^*(s, z, y)$, for every $y < y^*$. We then let $\hat{\mu}_t(s, z, y) = 1$ if and only if $y \leq y^*$.

The above described procedure is repeated for every $t$, $s$, and $z$. This results in functions $\hat{\mu}_t$ that satisfy $\hat{\mu}_t(s, z, y) \in \hat{U}_k^*(s, z, y)$ for all $(t, s, z, y)$. According to Lemma 2.0.1(3), choosing the decision according to $\hat{\mu}_t$ for each unit at each time step constitutes an optimal (and also decoupled and committed) policy. Furthermore, by our construction, $\hat{\mu}_t(s, z, y)$ is a monotonically nonincreasing function of $y$. That is, if two different units are at the same location, and if a lower-indexed unit (with smaller $y$) is not released (decision equal to 0), then any higher-indexed unit (with a larger value of $y$) is not released either. It follows that this is a monotonic policy, thus establishing the existence of a monotonic, committed, and decoupled policy which is optimal.

**Remark:** The fact that an optimal policy $\{\hat{\mu}_t\}$ for the subproblem can be chosen so that it is monotone in $y$ can be established using a traditional inductive argument, based on the dynamic programming recursion. For example, [1] studies the recursion for the infinite horizon single-unit single-customer problem with deterministic leadtimes and i.i.d. demands, and provides a lengthy algebraic derivation. In contrast, the proof we gave only relies on qualitative arguments.

**Theorem 3.3.1.** The set of state-dependent echelon base stock policies is optimal.

**Proof.** Follows from Propositions 3.2.1 and 3.3.5.

Note that as a corollary of the preceding proof, the basestock levels $S_t^{y-1}(s) = y^* - 1$ are readily determined once an optimal policy and the corresponding sets $\hat{U}_k^*(s, z, y)$ for the subproblem are available.
3.4 Infinite Horizon Analysis

This section deals with the case where the planning horizon is infinite. We study both the expected total discounted cost and expected average cost per unit time criteria. We start with the part of the analysis that is common to both criteria.

In the infinite horizon setting, we consider stationary policies. A stationary policy is one of the form \((\mu, \mu, \ldots)\), with \(\mu: X \mapsto U\), so that the decision at each time is a function of the current state but not of the current time. In the infinite horizon context, we refer to a stationary policy of this type as policy \(\mu\). Let \(\Omega\) denote the set of all stationary infinite horizon policies.

Similarly, for the subproblems, we refer to a stationary policy of the form \((\hat{\mu}, \hat{\mu}, \ldots)\) with \(\hat{\mu}: S \times Z \times Y \mapsto \hat{U}\) as policy \(\hat{\mu}\). Given a fixed discount factor \(\alpha \in [0, 1]\), let \(\hat{J}_\infty^*(s, z, y)\) and \(\hat{J}_\infty^*(s, z, y)\) be the infinite horizon expected total discounted cost of policy \(\hat{\mu}\), and the corresponding optimal cost, respectively. Let \(\hat{J}_T^*(s, z, y)\) be the expected total discounted cost of using the stationary policy \(\hat{\mu}\) in a subproblem over a finite horizon of length \(T\), given that the initial state of the subproblem is \((s, z, y)\). We will still use the definitions introduced in Section 2 and Subsection 3.2.2, which have obvious extensions to the infinite horizon case.

**Definition 3.4.1.** For any \(s \in S, z \in Z, y \in Y\), let \(\hat{U}_\infty^*(s, z, y) \subset \{0, 1\}\) be the set of all decisions that are optimal if a subproblem is found at state \((s, z, y)\).

The next lemma relates the finite and infinite horizon versions of the subproblem.

**Lemma 3.4.1.** For any fixed \(\alpha \in [0, 1]\), and any \(s, z, y\), we have

\[
\lim_{T \to \infty} \hat{J}_T^*(s, z, y) = \hat{J}_\infty^*(s, z, y).
\]

Furthermore, if \(u \in \hat{U}_T^*(s, z, y)\) for infinitely many choices of \(T\), then \(u \in \hat{U}_\infty^*(s, z, y)\).

**Proof.** For any given initial state \((s, z, y)\) for a subproblem, the number of possible future states is finite; this is because \(y\) and \(z\) cannot increase. Therefore, general results for finite-state Markov decision problems apply. When \(\alpha < 1\), the convergence of \(\hat{J}_T^*(s, z, y)\)
to \( J^*_\infty(s, z, y) \) is immediate. When \( \alpha = 1 \), we have a "stochastic shortest path problem," ([7]) and any "improper" policy (that is, any policy that is not guaranteed to eventually deliver the unit to the customer) incurs infinite cost, due to eternal backlogging. Under this condition, the claimed convergence is again guaranteed.

The last part of the lemma is elementary and we only sketch the argument. The optimality of a particular decision for a certain time horizon \( t \) can be expressed in terms of an associated Bellman equation. By taking the limit in that Bellman equation as \( t \) goes to infinity, we recover a condition that asserts optimality of the same decision for an infinite horizon problem. \( \square \)

**Lemma 3.4.2.** For every \((s, z, y)\), if \( \hat{U}^*_\infty(s, z, y) = \{1\} \), then \( 1 \in \hat{U}^*_\infty(s, z, y') \), for every \( y' < y \).

**Proof.** If \( \hat{U}^*_\infty(s, z, y) = \{1\} \), Lemma 3.4.1 implies that there is a \( t' > 0 \) such that \( \hat{U}^*_t(s, z, y) = \{1\} \) for all \( t > t' \). Then, by Lemma 3.3.1, we have \( 1 \in \hat{U}^*_t(s, z, y') \) for every \( t > t' \) and \( y' < y \). Hence, by Lemma 3.4.1, \( 1 \in \hat{U}^*_\infty(s, z, y') \) for every \( y' < y \). \( \square \)

**Proposition 3.4.1.** There exists an optimal policy \( \hat{\mu}^* \) for the infinite horizon subproblem such that \( \hat{\mu}^*(s, z, y) \) is a monotonically nonincreasing function of \( y \).

**Proof.** The argument is identical to the one in the proof of Proposition 3.3.5, using \( \hat{U}^*_\infty(s, z, y) \) in place of \( \hat{U}^*_t(s, z, y) \), and by invoking Lemma 3.4.2 in place of Lemma 3.3.1. \( \square \)

**Proposition 3.4.2.** Let \( \mu^* \) be the stationary, decoupled, and committed policy for the overall problem that uses the optimal subproblem policy \( \hat{\mu}^* \) of Prop. 3.4.1 for each unit-customer pair. Then, \( \mu^* \) is a stationary state dependent echelon base stock policy.

**Proof.** By definition, \( \mu^* \) is decoupled and committed. By Proposition 3.4.1, \( \hat{\mu}^* \) is a monotonically nonincreasing function of \( y \). If two different units are at the same location, and if a lower-indexed unit (with smaller \( y \)) is not released (decision equal to 0), then any higher-indexed unit (with a larger value of \( y \)) is not released either. It follows that \( \mu^* \) is a monotonic policy as well. The rest of the argument is almost identical to Proposition 3.2.1. As in the
finite horizon case (Proposition 3.2.1), a monotonic, committed and decoupled policy is a state dependent echelon base stock policy and the result follows. 

We have so far constructed a stationary state dependent echelon base stock policy \( \mu^* \). This policy is constructed as a limit of optimal policies for the corresponding finite horizon problems. It should then be no surprise that \( \mu^* \) is optimal for the infinite horizon problem. However, some careful limiting arguments are needed to make this rigorous. This is the subject of the next two subsections.

3.4.1 Discounted Cost Criterion

In this subsection, we focus on the infinite horizon expected total discounted cost. In particular, the cost of a stationary policy \( \mu \), starting from an initial state \( x = \{ s, (z^1, y^1), (z^2, y^2), \ldots \} \), is defined as

\[
J^\mu_\infty(x) = \lim_{T \to \infty} E \left\{ \sum_{t=0}^{T-1} \alpha^t \cdot g(x_t, \mu(x_t)) \right\} \left| x_0 = x \right.
\]

where \( 0 < \alpha < 1 \). The infinite horizon optimal cost is defined by

\[
J^*_\infty(x) = \inf_{\mu \in \Omega} J^\mu_\infty(x).
\]

A policy \( \mu \) is said to be \textit{optimal} if

\[
J^\mu_\infty(x) = J^*_\infty(x),
\]

for every monotonic state \( x \).

A stationary policy can be used over any time horizon, finite or infinite. Let \( J^\mu_T(x) \) be the expected total discounted cost of using the stationary policy \( \mu \) during a finite planning horizon of length \( T \), starting with the initial state \( x_0 = x \). We then have

\[
J^\mu_\infty(x) = \lim_{T \to \infty} J^\mu_T(x).
\]

Recall that \( J^\mu_T(x) \) is defined as the optimal expected cost with a planning horizon from time 0 until time \( T \), given that \( x_0 = x \). Note that the finite horizon optimization is not
restricted to stationary policies. Hence, $J^*_T (x)$, the optimal cost among all policies within a finite planning horizon of $T$ periods, satisfies $J^*_T (x) \leq J^*_T (x)$ for any stationary policy $\mu$.

By Proposition 3.3.4, we have,

$$J^*_T (x) = \sum_{i=1}^{\infty} J^*_T (s, z^i, y^i),$$

for any monotonic state $x$. Hence, for every monotonic state $x$, we have

$$J^*_\infty (x) = \inf_{\mu \in \Omega} \lim_{T \to \infty} J^*_T (x)$$

$$\geq \lim_{T \to \infty} \inf_{\mu \in \Omega} J^*_T (x)$$

$$\geq \lim_{T \to \infty} J^*_T (x)$$

$$= \lim_{T \to \infty} \sum_{i=1}^{\infty} \hat{J}^*_T (s, z^i, y^i)$$

$$= \sum_{i=1}^{\infty} \lim_{T \to \infty} \hat{J}^*_T (s, z^i, y^i)$$

$$= \sum_{i=1}^{\infty} \hat{J}^*_\infty (s, z^i, y^i),$$

where the exchange of the limit and the summation is warranted by the monotone convergence theorem, since the functions $\hat{J}^*_T$ are monotonically increasing in $T$.

The above inequality provides a lower bound for the optimal cost. Consider now the decoupled policy $\mu^*$ from Proposition 3.4.2, which uses an optimal subproblem policy $\hat{\mu}^*$ for each unit-customer pair. The cost of $\mu^*$ is

$$J^*_{\infty} (x) = \sum_{i=1}^{\infty} \hat{J}^*_{\infty} (s, z^i, y^i) = \sum_{i=1}^{\infty} \hat{J}^*_\infty (s, z^i, y^i).$$

For a monotonic state $x$, this is equal to the lower bound, hence $\mu^*$ is optimal for the overall problem.

**Theorem 3.4.1.** The set of state dependent echelon base stock policies is optimal under the infinite horizon discounted cost criterion.

**Proof.** The policy $\mu^*$ is a state dependent echelon base stock policy, by Proposition 3.4.2. For a monotonic state $x$, it attains the lower bound and is optimal. \qed
3.4.2 Average Cost Criterion

In this subsection, we consider the infinite horizon average cost criterion.

The average cost of a policy $\mu$, starting from an initial state $x = \{s, (z^1, y^1), (z^2, y^2), \ldots\}$, is defined as

$$
\lambda^\mu(x) = \limsup_{T \to \infty} \frac{1}{T} \mathbb{E} \left\{ \sum_{t=0}^{T-1} g(x_t, \mu(x_t)) \mid x_0 = x \right\} = \limsup_{T \to \infty} \frac{1}{T} J^\mu_T(x).
$$

The optimal average cost is defined by

$$
\lambda^*(x) = \inf_{\mu \in \mathcal{P}} \lambda^\mu(x).
$$

A policy $\mu$ is said to be optimal if

$$
\lambda^\mu(x) = \lambda^*(x),
$$

for every monotonic state $x$.

For any monotonic state $x$, we have

$$
\lambda^*(x) = \inf_{\mu \in \mathcal{P}} \limsup_{T \to \infty} \frac{1}{T} J^\mu_T(x)
\geq \limsup_{T \to \infty} \frac{1}{T} \inf_{\mu \in \mathcal{P}} J^\mu_T(x)
\geq \limsup_{T \to \infty} \frac{1}{T} J^*_T(x)
= \limsup_{T \to \infty} \frac{1}{T} \sum_{i=1}^{\infty} \tilde{J}^*_T(s, z^i, y^i). \tag{3.1}
$$

The right hand side of the above inequality is a lower bound on the optimal infinite horizon average cost. We will show that the state dependent echelon base stock policy $\mu^*$ from Proposition 3.4.2 achieves this lower bound, and is therefore optimal.

The lemmas that follow analyze the structure of the optimal cost function for the subproblem. The first lemma establishes a uniform upper bound for the cost incurred by a unit that lies at the outside supplier, that is, location $N + 1$. 

44
Lemma 3.4.3. There exists a scalar $C_{\text{max}}$ such that

\[ \hat{J}_{T}^{*}(s, N+1, y) \leq \hat{J}_{T}^{\text{opt}}(s, N+1, y) \leq \hat{J}_{\infty}^{*}(s, N+1, y) \leq C_{\text{max}}, \]

for every $(T, s, y)$, where $\mu^{*}$ is the optimal subproblem policy from Proposition 3.4.1.

Proof. The first two inequalities are obvious, so we concentrate on the third. Note that the cost for any unit-customer pair in any given single period is bounded by $b + h_{\text{max}}$, where $h_{\text{max}} = \max_{i} h_{i}$. Consider a policy where an unit is kept at location $N+1$ until the customer arrives and then the unit is pushed through the system as quickly as possible (wait-push policy). Such a policy will incur a positive cost only while the unit is in transit in the system, which is at most $(N+1)$ periods. Thus, the infinite horizon expected cost of the wait-push policy is bounded by $C_{\text{max}} = (N+1) \cdot (b + h_{\text{max}})$.

Lemma 3.4.4. There exists an integer $Y_{\text{max}}$ such that any optimal policy for a finite horizon subproblem with large enough horizon, or any optimal policy for the infinite horizon subproblem, will not release the unit from stage $N+1$ if the distance of the customer is larger than $Y_{\text{max}}$. Formally, there exist some $t_{\text{max}} > 0$ and $Y_{\text{max}}$ such that for every $s$ and for every $y > Y_{\text{max}}$, $\hat{U}_{t}^{*}(s, N+1, y) = \{0\}$ for $t > t_{\text{max}}$, and $\hat{U}_{\infty}^{*}(s, N+1, y) = \{0\}$.

Proof. First consider the infinite horizon subproblem. Once the unit leaves stage $N+1$, a positive holding cost has to be incurred, at least until the customer arrives. Let $h_{\text{min}} > 0$ be the minimum holding cost rate at locations 1, \ldots, $N$. Suppose that the customer distance is $y$. Let $\tau(s, y)$ be the number of periods until the customer's arrival, given that the state of the Markov chain is currently $s$, and let $e(s, y) = E[\tau(s, y)]$. If the unit is released from stage $N+1$, the expected remaining cost is at least $e(s, y) \cdot h_{\text{min}}$. Clearly, for every $s$, $e(s, y)$ is nondecreasing in $y$ and diverges as $y$ goes to infinity. Hence, there is an integer $Y_{\text{max}}$ such that $e(s, y) \cdot h_{\text{min}} > C_{\text{max}}$ for every $y \geq Y_{\text{max}}$ and every $s$. Therefore, any policy that releases an unit from stage $N+1$ while the distance of the customer is greater than $Y_{\text{max}}$ will have an expected cost larger than $C_{\text{max}}$ and cannot be optimal. This proves the result for the infinite horizon case.
We now consider the case of a finite horizon $T$. If the unit is released from location $N+1$, the expected cost is at least $E[\min\{\tau(s, y), T\}] \cdot h_{\min}$. For any $y > Y_{\max}$, this is larger than or equal to $E[\min\{\tau(s, Y_{\max}), T\}] \cdot h_{\min}$. As $T$ increases to infinity, $E[\min\{\tau(s, Y_{\max}), T\}] \cdot h_{\min}$ converges to $E[\tau(s, Y_{\max})] \cdot h_{\min} = e(s, Y_{\max}) \cdot h_{\min} > C_{\max}$. Therefore, there exists some $t_{\max}$ such that if $y > Y_{\max}$ and $T > t_{\max}$, then the cost of a released unit is larger than $C_{\max}$, and therefore larger than the optimal cost. Hence, it is not optimal to release the unit, and the finite horizon proof is complete. 

\[ \square \]

**Lemma 3.4.5.** Consider the finite horizon subproblem. There exists an integer $K$ such that if the customer arrives $K$ periods or more before the end of the horizon, then under any optimal policy the unit will be given to the customer before the end of the horizon.

**Proof.** First, consider the infinite horizon subproblem where the customer has already arrived but the unit is not given to the customer yet. In this case, moving the unit as quickly as possible is the unique optimal control since waiting at a certain location for one period does nothing but add one extra period of holding and penalty cost. Hence, $\hat{U}^*_t(s, z, 1) = \{1\}$ for every $s$ and $z > 0$. Now fix some $s$ and $z > 0$. By Lemma 3.4.1, we cannot have $0 \in \hat{U}^*_t(s, z, 1)$ for infinitely many choices of $t$ and therefore there exists some $t'(s, z)$ such that $\hat{U}^*_t(s, z, 1) = \{1\}$ for every $t > t'(s, z)$. Let $K = N + 2 + \max_{s, z} t'(s, z)$. Now, consider the $T$-horizon subproblem and suppose that a customer arrives at time $k$ where $T - k \geq K$. In this case, the optimal decision is to move the unit for periods $t = k, k + 1, \ldots, k + N + 1$. Thus, the unit will move through the system as quickly as possible, and will be given to the customer before the end of the horizon. \[ \square \]

Let $\bar{d}$ be the expected demand per unit time, in steady state. In particular,

$$\bar{d} = \sum_{s \in S} \lim_{t \to \infty} \text{Prob}(s_t = s) E[d_t \mid s_t = s].$$

Note that the limits defining the steady state probabilities $\lim_{t \to \infty} P(s_t = s)$ exist and are independent of $s_0$ because we have assumed that $s_t$ is irreducible and aperiodic.
Our next result establishes that if the horizon $T$ is large and a customer $y$ is fairly certain to arrive well before time $T$, the cost-to-go $\hat{J}_T^* (s, N + 1, y)$ is very close to the infinite horizon cost $\hat{J}_\infty^* (s, N + 1, y)$, in a uniform sense.

**Lemma 3.4.6.** Fix some $\varepsilon > 0$. For every $s$ and for every $y$ such that $y \leq (\bar{d} - \varepsilon) \cdot T$, we have

$$\left| \hat{J}_T^* (s, N + 1, y) - \hat{J}_\infty^* (s, N + 1, y) \right| \leq f (T, s)$$

for some $f : \mathbb{N} \times S \rightarrow \mathbb{R}$ such that $\lim_{T \rightarrow \infty} f (T, s) = 0$.

**Proof.** Note that $\hat{J}_T^* (s, N + 1, y) \leq \hat{J}_\infty^* (s, N + 1, y)$, since having more time periods can only increase the costs. We will next establish an inequality in the reverse direction.

Given a time horizon $T > t_{\text{max}}$, consider the following non-stationary policy for the infinite horizon subproblem with initial state $(s, N + 1, y)$: in the first $T$ periods, employ an optimal policy for the $T$-horizon problem, and then employ the wait-push policy. At the end of period $T$, the unit is either at location 0 (given to the customer), or at a location greater than 0 but less than $N + 1$ (in the system), or at location $N + 1$ (at the supplier). If the unit is given to the customer, there is no more cost. If the unit is in the system, then the distance of the customer is at most $Y_{\text{max}}$, by Lemma 3.4.4. Hence, the remaining cost can be bounded by a value that does not depend on $y$. Let $v$ be this value. If the unit is at the supplier, then the cost of the wait-push policy is bounded by $C_{\text{max}}$, by Lemma 3.4.3. Hence, the cost of this combined policy is bounded by

$$\hat{J}_T^* (s, N + 1, y) + \text{Prob} \{ \text{Unit not given to the customer by time } T \} \cdot (C_{\text{max}} + v).$$

Moreover, this cost has to be at least as large as the optimal infinite horizon cost. So, we have:

$$\hat{J}_\infty^* (s, N + 1, y) - \hat{J}_T^* (s, N + 1, y) \leq \text{Prob} \{ \text{Unit not given to the customer by time } T \} \cdot (C_{\text{max}} + v).$$

Let $d^{T-K} (s)$ be a random variable denoting the sum of the demands in $T - K$ periods, starting with a period where the Markov chain is in state $s$, where $K$ is the constant of
Lemma 3.4.5. If the unit is not given to the customer by time $T$, then by Lemma 3.4.5, the customer has not arrived by time $T - K$, that is, $d^{T-K}(s) < y - 1$. Hence,

$$
\tilde{J}_\infty^r (s, N + 1, y) - \tilde{J}_0^r (s, N + 1, y) \leq \text{Prob} \left\{ d^{T-K}(s) < (\bar{d} - \epsilon)T \right\} \cdot (C_{\max} + v).
$$

For $T > K$, let $f(T, s)$ be the right hand side of the above inequality. It remains to show that $\lim_{T \to \infty} f(T, s) = 0$.

Indeed,

$$
\text{Prob} \left\{ d^{T-K}(s) < (\bar{d} - \epsilon)T \right\} = \text{Prob} \left\{ \frac{d^{T-K}(s)}{T} < \bar{d} - \epsilon \right\}.
$$

As $T \to \infty$, by the law of large numbers for Markov reward processes, $d^{T-K}(s)/T$ converges to $\bar{d}$ almost surely, and therefore, in probability. Therefore, the probability we are considering converges to zero, and so does $f(T, s)$.

We are now ready for the main part of our argument, which is to show that the state dependent echelon base stock policy $\mu^*$ from Proposition 3.4.2 achieves the lower bound in Eq. (3.1), and is therefore optimal.

**Proposition 3.4.3.** The policy $\mu^*$ from Proposition 3.4.2 is optimal.

**Proof.** Since $\mu^*$ is a committed and decoupled policy, we have for every finite horizon $T$,

$$
\tilde{J}_{T}^{\mu^*} (x) = \sum_{i=1}^{\infty} \tilde{J}^{\mu^*}_{T} (s, z^i, y^i).
$$

Then, by the definition of the infinite horizon average cost,

$$
\lambda^{\mu^*} (x) = \limsup_{T \to \infty} \frac{1}{T} \sum_{i=1}^{\infty} \tilde{J}^{\mu^*}_{T} (s, z^i, y^i).
$$

Let us fix a monotonic initial state $x = \{s, (z^1, y^1), (z^2, y^2), \ldots \}$. If there is an infinite number of units in locations other than $N + 1$, the optimal cost-to-go is infinite and there is nothing to prove. We can therefore assume there is a finite number $k$ of units in locations $1, \ldots, N$. Let $\ell$ be the number of units whose corresponding customers have already
arrived to the system, so that \( y^i = i - \ell + 1 \) for \( i > \ell \). We have

\[
\lambda^\mu(x) = \limsup_{T \to \infty} \frac{1}{T} \sum_{i=1}^{\infty} \tilde{j}_T^\mu(s, z^i, y^i)
\leq \limsup_{T \to \infty} \frac{1}{T} \sum_{i=1}^{k+\ell} \tilde{j}_T^\mu(s, z^i, y^i) + \limsup_{T \to \infty} \frac{1}{T} \sum_{i=k+\ell+1}^{\infty} \tilde{j}_T^\mu(s, N+1, i - \ell + 1)
= \limsup_{T \to \infty} \frac{1}{T} \sum_{i=1}^{\infty} \tilde{j}_T^\mu(s, N+1, i).
\]

Let \( \epsilon > 0 \) be a constant less than \( \bar{d}/2 \). We decompose the above expression as follows:

\[
\limsup_{T \to \infty} \frac{1}{T} \sum_{i=1}^{\infty} \tilde{j}_T^\mu(s, N+1, i)
\leq \limsup_{T \to \infty} \frac{1}{T} \sum_{i=1}^{[\epsilon T]} \tilde{j}_T^\mu(s, N+1, i) + \limsup_{T \to \infty} \frac{1}{T} \sum_{i=[\epsilon T]+1}^{[(\bar{d}-\epsilon)T]} \tilde{j}_T^\mu(s, N+1, i)
+ \limsup_{T \to \infty} \frac{1}{T} \sum_{i=([\epsilon T]+1)}^{[(\bar{d}-\epsilon)T]} \tilde{j}_T^\mu(s, N+1, i) + \limsup_{T \to \infty} \frac{1}{T} \sum_{i=([\bar{d}-\epsilon)T]+1}^{[\bar{d}T]} \tilde{j}_T^\mu(s, N+1, i).
\]

We will show that the first and the third terms in the above sum go to zero as \( \epsilon \to 0 \), and that the fourth term is equal to zero. Using Lemma 3.4.3, the first term satisfies

\[
\limsup_{T \to \infty} \frac{1}{T} \sum_{i=1}^{[\epsilon T]} \tilde{j}_T^\mu(s, N+1, i) \leq \limsup_{T \to \infty} \frac{1}{T} \epsilon TC_{max} \to 0.
\]

Similarly, for the third term,

\[
\limsup_{T \to \infty} \frac{1}{T} \sum_{i=([\epsilon T]+1)}^{[(\bar{d}-\epsilon)T]} \tilde{j}_T^\mu(s, N+1, i) \leq \limsup_{T \to \infty} \frac{1}{T} 2\epsilon TC_{max} \to 0.
\]

To get the result for the fourth term, consider a unit \( i \) and its corresponding customer. After an interval of \( T \) periods, the distance of this customer will be \( (i - d^T(s))^+ + 1 \), where \( d^T(s) \) is the random variable denoting the sum of demands in \( T \) periods, starting from a period with the Markov chain in state \( s \) (assuming that the customer has not received the unit within the interval). By Lemma 3.4.4, if \( (i - d^T(s))^+ + 1 > Y_{max} \), unit \( i \) will not be
released from location $N + 1$ and this unit customer pair will have a cost of 0 during the $T$-step horizon. For any unit that is released from location $N + 1$, the expected cost can be at most $C_{\text{max}}$. Therefore,

$$
\limsup_{T \to \infty} \frac{1}{T} \sum_{i = \lfloor (\bar{d} + \epsilon) T \rfloor + 1}^{\infty} \tilde{J}^\mu_T (s, N + 1, i)
\leq \limsup_{T \to \infty} \frac{1}{T} \cdot C_{\text{max}} \cdot E \left\{ (d^T (s) + Y_{\text{max}} - \lfloor (\bar{d} + \epsilon) \cdot T \rfloor - 1)^+ \right\} = 0,
$$

using the law of large numbers. This means that only the second term in the right hand side of Eq. (3.2) remains positive as $\epsilon \downarrow 0$, and

$$
\lambda^\mu (x) = \limsup_{T \to \infty} \frac{1}{T} \sum_{i = \lfloor \epsilon T \rfloor + 1}^{\lfloor (\bar{d} - \epsilon) T \rfloor} \tilde{J}^\mu_T (s, N + 1, i) + f(\epsilon)
\leq \limsup_{T \to \infty} \frac{1}{T} \sum_{i = \lfloor \epsilon T \rfloor + 1}^{\lfloor (\bar{d} - \epsilon) T \rfloor} \tilde{J}^\mu_\infty (s, N + 1, i) + f(\epsilon),
$$

for some function $f$ that satisfies $\lim_{\epsilon \downarrow 0} f(\epsilon) = 0$.

We now use Eq. (3.1), to obtain

$$
\lambda^* (x) \geq \limsup_{T \to \infty} \frac{1}{T} \sum_{i=1}^{\infty} \tilde{J}^*_T (s, z^i, y^i)
\geq \limsup_{T \to \infty} \frac{1}{T} \sum_{i = \lfloor \epsilon T \rfloor + 1}^{\lfloor (\bar{d} - \epsilon) T \rfloor} \tilde{J}^*_T (s, N + 1, i)
\geq \limsup_{T \to \infty} \frac{1}{T} \sum_{i = \lfloor \epsilon T \rfloor + 1}^{\lfloor (\bar{d} - \epsilon) T \rfloor} \tilde{J}^*_\infty (s, N + 1, i) - \limsup_{T \to \infty} \frac{1}{T} \bar{d} T f(s, T)
= \limsup_{T \to \infty} \frac{1}{T} \sum_{i = \lfloor \epsilon T \rfloor + 1}^{\lfloor (\bar{d} - \epsilon) T \rfloor} \tilde{J}^*_\infty (s, N + 1, i)
$$

where the last inequality uses Lemma 3.4.6. By comparing the above two inequalities, and using the fact $\tilde{J}^*_\infty = \tilde{J}^\mu_\infty$ (optimality of $\mu^*$ for the infinite horizon subproblem), we obtain

$$
\lambda^\mu (x) \leq \lambda^* (x) + f(\epsilon).
$$
By taking the limit as $\epsilon$ decreases to zero, we obtain $\lambda^*(x) \leq \lambda^*(x)$, which establishes the optimality of $\mu^*$.

By Proposition 3.4.2, $\mu^*$ is a state dependent echelon base stock policy, which leads to our main result.

**Theorem 3.4.2.** The set of state dependent echelon base stock policies is optimal under the infinite horizon average cost criterion.

We close this section, by providing a characterization of the infinite horizon average cost. The expected cost incurred by successive customers is of the form $\hat{J}_\infty^*(s, N + 1, y)$, for ever increasing values of $y$. Over a time interval of length $T$, about $\bar{d}T$ customers are expected to arrive suggesting that the average cost per unit time is of the form $\bar{d} \lim_{y \to \infty} \hat{J}_\infty^*(s, N + 1, y)$.

The proposition that follows shows that the above limit exists and that the above intuition is correct. We will need, however, a minor assumption on the nature of the demand process. We say that the demand process $\{d_t\}$ is of the lattice type if there exists an integer $\ell > 1$ such that for every $s$, the conditional distribution of $d_t$, given $s_t = s$, is concentrated on the integer multiples of $\ell$. Otherwise, we say that $\{d_t\}$ is of the non-lattice type.

**Lemma 3.4.7.** Suppose that the demand process is of the non-lattice type. Then,

(a) There exists a constant $C$, such that

$$\lim_{y \to \infty} \hat{J}_\infty^*(s, N + 1, y) = C, \quad \forall \ s \in S.$$ 

(b) For every state $x$ such that the number of units in locations other than $N + 1$ is finite, we have

$$\lambda^*(x) = C\bar{d}.$$ 

In particular, the optimal average cost is the same for all initial states.

**Proof.** (a) Let $F_s(r)$ be the probability mass function of the distribution of the demand when the Markov chain is in state $s$. Let $P_{i,j} = \text{Prob}(s_{t+1} = j \mid s_t = i)$ be the transition
probabilities of the Markov chain \( s_t \). For every \( y > Y_{\text{max}} \), Lemma 3.4.4 states that the optimal decision is to not release the unit, i.e., \( u = 0 \), and the dynamic programming equation for the subproblem takes the form

\[
\hat{J}_{\infty}^*(s, N + 1, y) = \sum_{j \in \mathcal{S}} \sum_{r=0}^{\infty} P_{s,j} \cdot F_s(r) \cdot \hat{J}_{\infty}^*(j, N + 1, (y - 1 - r)^+) + 1).
\]

This equation is in the form of a Markov renewal equation. Since the Markov chain \( s_t \) is irreducible and aperiodic, and since the demand process \( d_t \) is of the non-lattice type, Proposition 4.17 in Chapter 10 of Çinlar [10] applies and shows that the solution of the Markov renewal equation converges as \( y \) goes to infinity, to a constant that does not depend on \( s \).

(b) The proof of Proposition 3.4.3 shows that

\[
\lambda^*(x) = \lim_{\epsilon \downarrow 0} \left[ \lim_{T \to \infty} \sup_{\tau \geq \epsilon T} \left\{ \frac{(\bar{d} - \epsilon)T}{\epsilon T + 1} \sum_{i=\epsilon T + 1}^{(\bar{d} - \epsilon)T} \hat{J}_{\infty}^*(s, N + 1, i) + f(\epsilon) \right\} \right]
\]

\[
= \lim_{\epsilon \downarrow 0} [(\bar{d} - 2\epsilon)C + f(\epsilon)]
\]

\[
= \bar{d}C,
\]

where the second equality follows from part (a).

The nonlattice assumption is hardly restrictive. In the lattice case, customers always arrive in batches of size \( \ell > 1 \). We can then treat each batch as a single customer and associate with it a batch of \( \ell \) units, which is treated as a single unit, which brings us back to the non-lattice case.

### 3.5 Algorithmic Issues

The subject of this section is a new algorithmic approach for determining the optimal echelon base stock levels in a serial inventory system.
While proving the optimality of state dependent echelon base stock policies, we established that decoupled and committed policies are optimal. Under such policies, the different unit-customer pairs are controlled separately and independent of each other. Besides providing a simple proof technique, such a decomposition also gives rise to efficient algorithms for computing the base stock levels. Instead of applying a dynamic programming algorithm on the larger problem involving all units and customers, we find an optimal policy for a subproblem involving a single unit-customer pair. That policy can be taken monotonic, by the argument used in the proof of Prop. 3.3.5. As in the proof of Proposition 3.2.1, the resulting base stock levels are readily obtained.

3.5.1 Subproblem Formulation

The subproblem to be solved is as follows. Given a single unit and a single customer, the goal is to move the unit through the serial system in a way that minimizes the expected inventory costs and backorder penalties. We assume that the leadtime model is the extension of Kaplan's model given in Section 3.2. Recall that in that model, at each time period \( t \), there is a random variable \( \rho^i_t \) that determines which outstanding orders will arrive at stage \( i \). More precisely, an order will be delivered if and only if it was placed \( \rho^i_t \) or more time units ago. Let \( \rho_t = (\rho^1_t, \rho^2_t, \ldots, \rho^M_t) \) be the vector of leadtime variables associated with the various stages. We assume that the statistics of \( \rho_t \) are given in terms of a conditional probability distribution, given the state \( s_t \) of the modulating Markov chain. If the unit is in transit, the holding cost rate that is charged is the rate at the destination echelon. In addition, we assume that the sequence of events within a period is as follows. First, a previously released unit may arrive at its destination location (depending on the previous period's leadtime random variable). The resulting new state is determined and observed. Then, the decision of whether or not to release the unit in this period is made. Finally, the demand and other random variables are realized, and holding and/or backorder costs are charged.

Let \((s_t, z_t, y_t)\) be the state of the system (subproblem) at time \( t \), where \( s_t \) is the state of the modulating Markov chain at time \( t \), \( z_t \) is the location of the unit at the beginning of
period \( t \) (after the move of the previous period is completed), and \( y_t \) is the distance of the customer at the beginning of period \( t \).

Let \( u_t \in \{0, 1\} \) be the control variable at time \( t \), where

\[
u_t = \begin{cases} 
1, & \text{if the unit is released from its current location,} \\
0, & \text{if the unit is kept at its current location.}
\end{cases}
\]

Of course, this decision can have an effect only if the location corresponds to an actual stage other than stage 1. Otherwise, we still allow a choice of 0 or 1 for \( u_t \), however this choice has no bearing on the evolution of the system or the costs charged. The location of a unit was defined in Subsection 3.2.1 and indicates that the unit is at an actual stage, or in transit between two actual stages for a certain time, or has been given to a customer (location 0). Let \( A \subset Z \) be the set of locations corresponding to actual stages in the original system plus the location \( N + 1 \). In addition, let

\[
u(z) = \max_{i \in A \text{ and } i \leq z} i, \quad \forall z > 0,
\]

i.e., \( v(z) \) is the location that has a label less than or equal to \( z \) and corresponds to an actual stage that is closest to location \( z \).

The following random variables affect the evolution of the system:

d_\ell: The demand at time \( t \),

\( r_\ell: \) A random variable that takes on the value 1 if the unit will go to \( v(z_\ell - 1) \) if released, or 0 if it will go to \( z_\ell - 1 \). Its distribution depends on \( z_\ell \) and \( s_\ell \). This variable determines whether a unit that is in transit will arrive to the next actual stage or whether it will stay in transit. Given the probability distribution of the leadtime random vector \( \rho_t = (\rho_1^t, \rho_2^t, \ldots, \rho_M^t) \) and \( z_\ell \) and \( s_\ell \), the probability distribution of \( r_\ell \) can easily be calculated.

Finally, the Markov chain that modulates the demand and leadtime processes evolves according to known transition probabilities.
The dynamics of the system are stationary. Costs are incurred until the state becomes \((s, 0, 0)\) for some \(s\) or until the end of the horizon is reached. States of the form \((s, z, 0)\) with \(z > 0\), or \((s, 0, y)\) with \(y > 0\) are impossible ("degenerate") because if the unit is given to a customer, the customer should have received a unit and vice versa. We assume that the initial state is not degenerate and the following dynamics guarantee that no degenerate state will be reached. The evolution of the distance of the customer is as follows. For every \(t\),

\[
y_{t+1} = \begin{cases} 
  y_t - d_t, & \text{if } y_t - d_t \geq 1 \text{ and } y_t > 1, \\
  1, & \text{if } y_t - d_t < 1 \text{ and } y_t > 1, \\
  1 - u_t I_{z_t=1}, & \text{if } y_t = 1, \\
  0, & \text{if } y_t = 0.
\end{cases}
\]

The evolution of the location of the unit is a little more involved due to the stochastic nature of the leadtimes. For every \(t\),

\[
z_{t+1} = \begin{cases} 
  0, & \text{if } z_t = 0, \\
  1 - I_{y_t=1} \cdot u_t, & \text{if } z_t = 1, \\
  (z_t - 1) \cdot (1 - r_t) + v (z_t - 1) \cdot r_t, & \text{if } z_t \geq 2 \text{ and } z_t \not\in A, \\
  (z_t - u_t) \cdot (1 - r_t) + v (z_t - u_t) \cdot r_t, & \text{if } z_t \geq 2 \text{ and } z_t \in A,
\end{cases}
\]

where we use the notation \(I_B\) to denote the indicator function of a set \(B\), that is,

\[
I_B = \begin{cases} 
  1, & \text{If } B \text{ is true}, \\
  0, & \text{If } B \text{ is false}.
\end{cases}
\]

The one period costs are stationary and are defined by

\[
\hat{g}(s_t, z_t, y_t, u_t) = \hat{h}_{z_t} - \hat{h}_{z_t} \cdot I_{y_t=1} \cdot I_{y_t=1} \cdot u_t \\
+ b \cdot I_{y_t=1} \cdot (1 - I_{z_t=1} \cdot u_t) \\
+ \hat{c}_{z_t} \cdot u_t.
\]

We define \(\hat{h}_{N+1} = 0\), \(\hat{h}_0 = 0\), and \(\hat{c}_0 = 0\), and set \(\hat{h}_i\) and \(\hat{c}_i\) to be the appropriate holding cost rate and order cost rate, respectively, for location \(i\). The holding cost rate for units in
artificial locations is assumed to be the rate associated with the downstream actual stage. Order costs for locations other than the actual stages should be taken to be zero. Then, the first line of (3.3) gives the holding cost, the second line gives the backorder cost, and the third line gives the ordering cost.

Note that the state space of the subproblem is infinite because $y$ can be arbitrarily large. However, Lemma 3.4.4 states that an optimal policy will never release a unit from the external supplier (location $N + 1$) unless $y \leq Y_{\text{max}}$. We are therefore justified in seeking optimal policies only starting from states in which $y \leq Y_{\text{max}}$. This makes the state space of the subproblem finite, of cardinality $O(N \cdot Y_{\text{max}} \cdot |S|)$.

### 3.5.2 Finite Horizon Complexity

Now that the description of the subproblem is complete, we can use dynamic programming to determine optimal policies. Consider the standard backward recursion algorithm for the finite horizon version of the subproblem.

**Proposition 3.5.1.** The complexity of the standard backward recursion dynamic programming algorithm for the finite horizon version of the subproblem given in Subsection 3.5.1, with a planning horizon of $T$, is $O\left(N \cdot Y_{\text{max}}^2 \cdot |S|^2 \cdot T\right)$.

**Proof.** The cardinality of the state space is $O(N \cdot Y_{\text{max}} \cdot |S|)$. For each state, the number of possible controls is 2, and the number of possible next states is bounded above by $2 \cdot Y_{\text{max}} \cdot |S|$. Hence, the complexity of the standard backward recursion algorithm for the finite horizon version of the dynamic program with a planning horizon of $T$ is $O\left(N \cdot Y_{\text{max}}^2 \cdot |S|^2 \cdot T\right)$. □

### 3.5.3 Infinite Horizon Complexity

We now describe an algorithm for the infinite horizon dynamic program, which is more efficient than standard methods. To help us analyze the complexity of this algorithm, we need the following definition of an optimal stopping problem and a corresponding lemma.
**Definition 3.5.1.** Consider an $n$-state Markov chain. Suppose that there is a cost for being at a given state. In addition, suppose that there is a controller that has an option to stop the Markov chain at any time, and that for each state there is a cost associated with stopping the Markov chain at that state. After the Markov chain is stopped, no more costs are incurred. The optimal stopping problem is the infinite horizon problem of stopping this Markov chain at such a time so that the total expected cost is minimized.

**Lemma 3.5.1.** An optimal stopping problem with $n$ states can be solved in $O(n^3)$ time.

**Proof.** (Outline) Consider the policy iteration algorithm, starting with the policy that stops at every state. We claim that the policy iteration algorithm takes at most $n$ iterations. This is because the cost-to-go of a state cannot increase in the course of the policy iteration algorithm. Thus, if the cost-to-go of not stopping becomes smaller than the cost of stopping at a given state, it remains smaller in subsequent iterations. Thus, with each policy iteration, the policy is either the same (in which case, we have an optimal policy), or the number of states at which the policy stops increases.

Let $k_i$ be the number of states at which the policy changes at the $i^{th}$ iteration. At each policy iteration, we need to solve a new system of equations with $n$ unknowns. But the transition matrix differs from the previous one in only $k_i$ rows (the ones where the policy changed). This is a rank $k_i$ modification. We can use the following fact from numerical linear algebra. If $A$ is $n \times n$ and $A^{-1}$ is available, and if $B - A$ has rank $k$, then we can compute $B^{-1}$ in $O(n^2 \cdot k)$ time. (This is possible because of the Sherman-Morrison-Woodbury formula in [23]). Therefore, the total complexity is $O(\sum_i k_i \cdot n^2)$ and since $\sum_i k_i \cdot n^2 \leq n^3$, the result follows. \hfill \Box

**Lemma 3.5.2.** Fix a pair $(z, y)$ of a unit location $z$ and customer distance $y$. Suppose that $\hat{J}^*_\infty(s', z', y')$ is available for every $(s', z', y')$ such that $z' \leq z$, $y' \leq y$, and $(z', y') \neq (z, y)$. Then, the values of $\hat{J}^*_\infty(s, z, y)$, for all $s \in S$, can be found in time $O(Y_{\text{max}} \cdot |S|^2)$ if $z$ is an artificial stage ($z \notin A$), and in time $O(Y_{\text{max}} \cdot |S|^2 + |S|^3)$ if $z$ is an actual stage ($z \in A$).

**Proof.** Consider first the case where $z$ is an artificial stage. Then, the location $z'$ at the
next time is guaranteed to satisfy $z' < z$. The Bellman equation for $J^*_\infty(s, z, y)$ in involves the known values of $J^*_\infty(s', z', y')$ for the various possible next states $(s', z', y')$. For each $s \in S$, there are at most $2 \cdot Y_{\text{max}} \cdot |S|$ possible next states, and the complexity estimate $O \left( Y_{\text{max}} \cdot |S|^2 \right)$ follows.

Suppose now that $z$ corresponds to an actual stage ($z \in A$). Given a current state $(s, z, y)$, the successor state is of the form $(s', z, y)$ as long as the demand is zero and the decision is to not release the unit. We view a release decision as a stopping decision and a nonzero demand as a forced stopping. Indeed, when we write down the Bellman equation for the various states of the form $(s, z, y)$, it takes the form of the Bellman equation for a stopping problem for a Markov chain with $|S|$ states. The transition probabilities, stopping, and continuation costs for this optimal stopping problem can be computed in time $O \left( Y_{\text{max}} \cdot |S|^2 \right)$. (This is because we have $|S|$ states of the form $(s, z, y)$ and for each such state at most $2 \cdot Y_{\text{max}} \cdot |S|$ possible next states.) By Lemma 3.5.1, the corresponding optimal stopping problem can be solved in $O \left( |S|^3 \right)$ time. Hence, the claimed complexity estimate follows.

\[ \square \]

**Proposition 3.5.2.** An optimal policy for the infinite horizon version of the subproblem described in Subsection 3.5.1 can be found with $O \left( M \cdot Y_{\text{max}} \cdot |S|^3 + N \cdot Y_{\text{max}}^2 \cdot |S|^2 \right)$, where $M$ is the number of actual (non-artificial) stages.

**Proof.** We only consider states of the form $(s, z, y)$ with $y \leq Y_{\text{max}}$. We calculate the values $J^*_\infty(s, z, y)$ recursively, starting with $J^*_\infty(s, z, 0) = 0$ for every $s$ and $z$, and $J^*_\infty(s, 0, y) = 0$ for every $s$ and $y$. The recursion proceeds by considering progressively larger values of $z$ and for each $z$, by considering progressively larger values of $y$. For each $(z, y)$ pair, the computational complexity of calculating $J^*_\infty(s, z, y)$ for all $s$ is given by Lemma 3.5.2. There are $O(N \cdot Y_{\text{max}})$ pairs $(z, y)$ to be considered, and $O(M \cdot Y_{\text{max}})$ pairs for which $z$ corresponds to an actual stage. The result follows.

\[ \square \]

The algorithms reported in this section are fairly efficient in terms of complexity. There is no other work that presents optimal algorithms for a multi-echelon inventory control problem with stochastic leadtimes and Markov-modulated demand. However, the decomposition of
the problem into single unit-customer pairs can be applied to special cases of our model that correspond to problems that have been studied before. For example, the serial system with deterministic leadtimes and i.i.d. demands, i.e., the model of Clark & Scarf [16], is such a special case, as is the serial system with deterministic leadtimes and Markovian demands studied in Chen & Song [12]. The single unit approach can solve the finite horizon version of the problem in Clark & Scarf [16] in $O(N \cdot Y_{\text{max}}^2 + M \cdot Y_{\text{max}}^2 \cdot T)$, which is a little better than what was stated in Prop 3.5.1. To see this, note that no decisions are made at the artificial locations. During the analysis, it was easier to introduce them, for the purposes of the proof, however, one can define a state space that involves only the actual locations and location 0, and modify the system dynamics appropriately. To work with such a model, the statistics about the total demand within the leadtime need to be calculated. This can be done once at the beginning of the algorithm, and takes time $O(N \cdot Y_{\text{max}}^2)$. Then, the dynamic programming recursion takes $O(M \cdot Y_{\text{max}}^2 \cdot T)$. The model of Clark & Scarf has been studied very extensively in the literature. The algorithms generated by the stage by stage decomposition suggested by Clark & Scarf, have complexities of $O(N \cdot Y_{\text{max}}^2 + M \cdot Y_{\text{max}}^2 \cdot T)$ and $O(N \cdot Y_{\text{max}}^2)$, for finite horizon and infinite horizon versions of the problem, respectively. These are consistent with the complexity of the algorithms developed through the single unit decomposition.

For the infinite horizon version with the average cost criterion, Gallego and Zipkin [22] present a streamlined version of the Clark & Scarf algorithm. (This streamlined version was initially suggested by Chen & Zheng [14], but they did not state it in the compact form found in Gallego and Zipkin [22].) This streamlined version is quite simple, and can be stated in just a few lines. However, for the finite horizon and the infinite horizon with the discounted cost criterion versions, we are not aware of such streamlined recursions, and the original decomposition given by Clark & Scarf needs to be used.

Another special case of the problem described in this chapter corresponds to a serial system with Markov modulated demand but deterministic leadtimes. For the infinite horizon version of this problem with an average cost criterion, Chen & Song [12] give an alternative
algorithm to calculate the base stock levels. Even though this reference does not report any precise complexity results, after inspecting their algorithm, it appears that it can be implemented in time $O(N \cdot Y_{\text{max}}^2 \cdot |S|^3)$, if the Sherman-Morrison-Woodbury formula is used. While this is not far from the complexity of our algorithm (the complexity of our algorithm is slightly lower), it only applies to a special case of our formulation (deterministic leadtimes) and only for the infinite horizon average cost criterion.

Even though it seems like the decomposition suggested by Clark & Scarf and our unit by unit decomposition give rise to algorithms with similar complexity estimates, we believe that our approach possesses some additional advantages. For the canonical model of Clark & Scarf, there exist streamlined proofs and algorithms. However, as one adds variations to the model, such as Markov modulated demands, our method is easily adapted to handle these new variations of the problem. In addition, our algorithms are very easy to understand and to implement; they are basically problems involving a single unit and a single customer.
Chapter 4

Serial Systems with Expediting

4.1 Introduction

In this chapter, we analyze a serial system with expediting options. Recall the classical serial system of Clark and Scarf [16]. In that model, there is a number of stages, and the units need to go through all these stages in order to be ready to be given to a customer. In our model, we now allow for the option of shipping a unit from a stage to any other stage downstream, not just the next stage, “expediting”, at a certain cost. This option gives the controller an additional flexibility that can be utilized as the uncertain demand is revealed.

Such a model can be applied in several settings. For example, in a supply chain setting, suppose that the units go through a national warehouse, a regional warehouse, and then a local inventory location. In the classical model, all the units have to go through all these locations. Our model allows the flexibility to ship (expedite) units from the national warehouse directly to the local inventory location, bypassing the regional warehouse. This may be a faster way to replenish inventory, albeit possibly at a higher cost. However, it may be a worthwhile option to consider in the face of uncertain demand. Another setting where this model may be applicable is a manufacturing operation. Suppose that a unit needs to go through several processes sequentially. An expediting decision in this case can correspond to either renting additional machines or hiring temporary employees or outsourcing, in order
to make sure that these subsequent processes can be completed faster than usual. Such a situation can be handled using the model in this chapter as well.

In this chapter, we use the single unit approach to characterize the structure of optimal policies for the multi-echelon inventory problem with expediting. We call the resulting policies “extended echelon base stock policies”. These policies, which are a generalization of echelon base stock policies, can be described by \((M + 1) \cdot M/2\) threshold values, where \(M\) is the number of stages in the system. In particular, since shipments from every stage to every downstream stage is possible, now we need extended base stock levels for every origin stage \(i\) to every destination stage \(j < i\). These extended echelon base stock levels completely determine an optimal policy. Given these threshold levels and the echelon inventory positions in the system at a given time, the optimal ordering levels can be computed via a simple iterative calculation, as described at the end of Section 4.3. We achieve this result by using the single unit approach, i.e. by showing that the problem can be decomposed into a series of subproblems each of which involves a single unit and a single customer. This approach then automatically gives rise to an algorithm for computing an optimal policy by solving a single unit subproblem. This is enough to compute the optimal extended base stock levels.

The only other work that considers a multi-echelon systems with expediting is Lawson and Porteus [27]. They represent their model a little differently than we do, but the two models are closely related, as discussed at the end of this chapter in Subsection 4.5.1. They also consider a serial system. In their model, the units have to go through all the stages. The transition from a stage to another can be expedited to make it essentially instantaneous. Since a unit can be expedited through several stages within the same period, this model is very similar to ours. It involves a more specialized cost structure, in that the cost of expediting through several stages is equal to the sum of the one step expediting costs. Our model assumes a more general expediting cost structure, that we call supermodularity (See Section 4.3 for details). However, because of a subtle difference between how the two models treat units that are expedited to stage 1, the model of [27] is not a special case of our model automatically. A slight modification in their model, or a slight modification in our analysis
makes such a comparison possible though, as explained in Subsection 4.5.1.

Lawson and Porteus [27] characterize optimal policies for their problem, which they name “top-down echelon base stock policies”. Similar to our “extended echelon base stock policies”, “top-down echelon base stock policies” consist of a number of threshold values, such that given those values and the echelon inventory positions, the optimal ordering levels can be determined through a simple iteration. However, their policies require only $2M + 1$ threshold values, and hence are somewhat simpler. On the other hand, this is not very surprising, since their model has a more specialized expediting cost structure, i.e. the expediting costs are additive.

Lawson & Porteus [27] is the only work that deals with expediting in a multi-echelon setting. There is a number of contributions in the literature, for single stage systems. This literature is commonly referred to as the “dual supply mode problem”. In particular, there is a single stage and two supply modes, one of them faster than the other. Daniel [17] characterizes the structure of optimal policies when there are two suppliers and the leadtimes are 0 and 1 respectively. Fukuda [21] extends this result to the case when the leadtimes are $k$ and $k + 1$ respectively. If the leadtimes differ by more than one period, Whittemore and Saunders [37] give conditions under which it is optimal to use only one supply mode.

The rest of this chapter has four sections. In Section 4.2, the problem formulation is given. Sections 4.3 and 4.4 analyze the finite and infinite horizon versions of the problem, respectively. The chapter concludes with Section 4.5, where the relationship between our model and that of Lawson and Porteus [27] is investigated, and then some extensions to the model are discussed.

### 4.2 Problem Formulation

We consider a single-product serial inventory system consisting of $M$ stages, indexed by $1, \ldots, M$. Customer demand can only be satisfied by units at stage 1. Any demand that is not immediately satisfied is backlogged. Stages can place orders for units to be delivered
from any upstream stage in the supply chain. In other words, stage \( i (i = 1, \ldots, M) \) can place an order for units stored at any stage \( j > i \), by incurring a certain cost. Stage \( M \) receives replenishments from an outside supplier with unlimited stock. For notational simplicity, we label the outside supplier as stage \( M + 1 \). We assume that inventory is reviewed periodically, hence a discrete time model can be used. The order leadtime for every origin destination pair is assumed to be equal to one period. (see Section 4.5 for a discussion of other leadtime models).

We assume that demand is independent and identically distributed over time, however our approach allows the analysis to be easily extended to the case where demand is Markovian modulated, similar to Chapter 3. We assume that demand takes on nonnegative integer values and the expected demand per period is finite. Let \( d_t \) be the demand at time \( t \).

The system involves inventory holding, ordering and backorder costs. In particular, we assume:

(a) For each stage \( i \), there is an inventory holding cost rate \( h_i \) that gets charged at each time period to each unit at that stage. We assume that the holding cost rate \( h_{M+1} \) at the external supplier is zero.

(b) There is a backorder cost rate \( b \) which is charged at each time step for each unit of backlogged demand.

(c) For each order from stage \( i \) to stage \( j \), there is a cost of \( c_{i,j} \) per unit.

We assume that the holding cost and backorder cost parameters are positive and that the shipping cost is non-negative.

The detail-oriented reader may have noticed that the model has not been specified in full detail: we would still need to describe the relative timing of observing the demand, fulfilling the demand, placing orders, receiving orders, and charging the costs. Different choices with respect to these details result, in general, to slightly different optimal costs and policies. Whatever specific choices are made, the arguments used for our subsequent results remain unaffected. For specificity, however, we make one assumption about delivery of units to
customers. We assume that if a customer arrives during period \( t \), a decision to give a unit to the customer can only be made at time \( t + 1 \), or later.

4.3 Finite Horizon Analysis

In this section, the finite horizon model is analyzed. We show that if the shipping costs have a structure that we call supermodularity, then the overall problem can be decomposed into a series of single unit subproblems. This analysis allows us to characterize the structure of optimal policies for the overall problem as extended echelon base stock policies, that are defined later in this section.

Following the central theme of the thesis, the analysis is based on the idea of decomposing the overall problem into a series of single-unit single-customer subproblems. After showing that the decomposition goes through, we analyze the structure of the resulting policies. They are not as simple as echelon base stock policies, but are quite manageable, in the sense that they can be summarized by \((M + 1) \cdot M/2\) threshold values. These values can be seen as extended base stock levels, that depend on both the stage from which units are to be shipped as well as the stage that is the intended destination. Given these \((M + 1) \cdot M/2\) threshold levels and the echelon inventory positions for each stage, the optimal ordering levels can be computed via a simple iteration, that is given at the end of this section.

To facilitate the single-unit analysis, we use the following state space: For each unit-customer pair \( i, i \in \mathbb{N} \), we have a vector \((z^i_t, y^i_t)\), with \(z^i_t \in Z = \{0, 1, \ldots, M + 1\}\) and \(y^i_t \in Y = \mathbb{N}_0\), where \(z^i_t \) is the location of unit \( i \) at time \( t \), and \(y^i_t \) is the distance of customer \( i \) at time \( t \). The state of the system consists of a countably infinite number of such vectors, one for each unit-customer pair, i.e.,

\[ x_t = \{(z^1_t, y^1_t), (z^2_t, y^2_t), \ldots \}. \]

Note that since all the leadtimes are assumed to be equal to one period, we do not need to insert artificial locations to model units in transit. The only artificial location is location
0, that corresponds to units that are already delivered to customers. Any other location \( i \) corresponds to a stage \( i \) of the system.

The control vector is an infinite sequence, \( u_t = (u_1^t, u_2^t, \ldots) \), where the \( i^{th} \) component \( u_i^t \) corresponds to a holding or shipping decision for the \( i^{th} \) unit. If unit \( i \) is at location 0, it is already delivered to a customer and is unaffected by \( u_i^t \). If unit \( i \) is in location \( j > 1 \), then \( u_i^t \) can take on a value in the set \( \{1, 2, \ldots, j\} \). The decision \( u_i^t = j \) corresponds to holding unit \( i \) at location \( j \). The decision \( u_i^t = k \) for some \( k \) with \( 1 \leq k < j \) corresponds to shipping unit \( i \) from location \( j \) to location \( k \). Finally, if unit \( i \) is at stage 1, a decision \( u_i^t = 0 \) releases this unit so that it can be given to a customer. In case the number \( n \) of units released from stage 1 is larger than the number \( m \) of customers whose demand is backlogged, only \( m \) of these units are given to customers (i.e., move to location 0), and the remaining \( n - m \) units stay at location 1. Otherwise, all \( n \) units are given to customers. The rules about which units are given to customers and which customers receive a unit are the following: If unit \( i \) is released to be given to a customer and customer \( i \)'s demand is backlogged, then unit \( i \) is given to customer \( i \). After all such matchings are done, if there are extra units and customers, the units and customers with the lowest indices are chosen until one side is empty.

This setting is slightly different than the one in Chapter 3, but the monotonic, committed and decoupled policy classes have analogous definitions in this setting as well. The modifications that are needed for formal definitions are obvious and to avoid replication, we omit them. Suffice it to say that monotonic policies still refer to policies that maintain the monotonic order among the units, committed policies still refer to policies that release a unit only if the corresponding customer is there, and decoupled policies still refer to policies that control each unit-customer pair independently of other units and customers.

To show that the problem can be decomposed into single unit subproblems, recall that in Chapter 3 we first showed that monotonic policies are optimal, then that committed policies are optimal, and finally that the problem is decomposable. For the expediting problem, the part where this framework may have a problem is the first step, i.e., when we need to show that monotonic policies are optimal. In fact, monotonic policies are not optimal if we do not
Figure 4-1: The supermodularity assumption means that if two units are to be expedited from two given stages $i$ and $j$ to two other stages $k$ and $l$, then the option where the unit originating from the earlier stage ends up at the later destination is at least as costly as the other option. In this figure, this means that option a) is at least as costly as option b).

Having any restriction on the shipping costs. Next, we define a property for the shipping costs that we call supermodularity, and show that this property allows us to restrict our attention to monotonic policies:

**Definition 4.3.1.** We define the shipping costs to be supermodular, if for any stages $i > j \geq k > l$, the following holds (See Figure 4-1):

$$c_{i,j} + c_{j,k} \geq c_{i,k} + c_{j,l}$$

The supermodularity assumption may be satisfied in certain settings and it may not in others. For example, in cases where the cost of a shipment is a convex function of the number of echelons that the unit travels, the supermodularity assumption is satisfied. However, in certain other cases, a single shipment over a long distance may be less expensive than the sum of two small shipments, and in such a case, the supermodularity assumption is violated when $j = k$. Nevertheless, from this point on, we will assume that the shipping costs are supermodular. We are able to characterize the structure of optimal policies for this class of
shipping costs, which is a quite large class. Of course, it is desirable to extend the results to
shipping costs that are not supermodular, however we suspect that the optimal policies in
that case may potentially be much more complicated.

**Proposition 4.3.1.** *The set of monotonic policies is optimal.*

*Proof.* Whenever some units are shipped from a particular stage, it does not matter which
particular units are shipped, since all units are identical. Hence, one can always choose units
so that the lower indexed ones will end up in lower locations. Hence, we can easily prevent
overtaking of units that are in the same location at a certain time. There is still the issue
that a unit may overtake another unit that is in a lower echelon. However, since the shipping
costs are supermodular, any decision that results in such an overtake can be replaced with
another that has at most the same cost and prevents overtaking, and the resulting states
are identical for optimization purposes. For example, in Figure 4-1, option b) can be chosen
instead of option a). Hence, given any policy \( \pi \) and monotonic initial state \( x \), there exists
another policy \( \pi' \), that at any time and for each location, ships the exact same number of
units to the same destinations, but prevents overtake. The new policy \( \pi' \) has at most the
same expected cost, and is also monotonic. \( \square \)

As we mentioned above, showing that monotonic policies are optimal was the only step
in the framework of showing the optimality of monotonic, committed and then of decoupled
policies that needed extra caution. Now, that this step is complete, the rest of the framework
can be applied without any problems and the problem can be decomposed into subproblems
each of which has a single unit and a single customer. We next state this result.

**Proposition 4.3.2.** *The set of committed and decoupled policies is optimal and

\[ J^*_T(x) = \sum_{i=1}^{\infty} J^*_T(z^i, y^i) \]

for every monotonic state \( x = \{(z^1, y^1), (z^2, y^2), \ldots\} \).

*Proof.* The steps in the proof are essentially the same as the proofs of Propositions 3.3.2,
3.3.3 and 3.3.4 and we only provide a sketch. In particular, under the restriction to monotonic
policies, we can restrict our attention to committed policies, since one can pair units with customers and be sure that a customer's demand can be satisfied using the corresponding unit, without sacrificing performance. Then, we relax the restriction to monotonic policies and focus on the class of committed policies. We obtain an equivalent system that represents this committed restriction as a modification of the system dynamics, and this system is decomposable. Lemma 2.0.1 then implies that decoupled policies are optimal.

At this point, we have shown that if the shipping costs are supermodular, the overall problem can be decomposed into a series of identical single unit subproblems. Next, we analyze the structure of these subproblems in more detail in order to characterize the structure of the resulting optimal policies for the overall problem.

**Definition 4.3.2.** For any \( k, z \in Z, y \in Y \), let \( \hat{U}_k^*(z, y) \subset \{1, 2, \ldots, M + 1\} \) be the set of all decisions that are optimal if a subproblem is found at state \((z, y)\) at time \( t = T - k \), that is \( k \) time steps before the end of the horizon. Also, let \( \hat{U}_*^*(z, y) \) be defined similarly for an infinite horizon problem.

**Lemma 4.3.1.** There exists an integer \( Y_{\text{max}} \) such that any optimal policy for a finite or infinite horizon subproblem will not release the unit from stage \( M + 1 \) if the distance of the customer is larger than \( Y_{\text{max}} \). Formally, there exist some \( Y_{\text{max}} \) such that for every \( y > Y_{\text{max}} \) and every \( t \), \( \hat{U}_t^*(M + 1, y) = \{M + 1\} \) and \( \hat{U}_\infty^*(M + 1, y) = \{M + 1\} \).

**Proof.** First, let us note that we already have a proof when the horizon is infinite or when the horizon is large enough, since the proof of Lemma 3.4.4 applies with obvious modifications. This allows us to state that there exists an integer \( Y'_{\text{max}} \) such that any optimal policy for a finite horizon subproblem with large enough horizon, or any optimal policy for the infinite horizon subproblem, will not release the item from stage \( M + 1 \) if the distance of the customer is larger than \( Y'_{\text{max}} \). Formally, there exist some \( t_{\text{max}} > 0 \) and \( Y'_{\text{max}} \) such that for every \( y > Y'_{\text{max}} \), \( \hat{U}_t^*(M + 1, y) = \{M + 1\} \) for \( t > t_{\text{max}} \), and \( \hat{U}_\infty^*(M + 1, y) = \{M + 1\} \).

Now, let us consider a horizon length \( t \leq t_{\text{max}} \). The cost of the policy that keeps the unit at the outside supplier \( M + 1 \) will be 0 if the customer does not show up within the time
horizon and will be at most $b \cdot t$ if the customer does show up within the time horizon. Let $p_t(y)$ be the probability that a customer at distance $y$ will show up within $t$ time periods. Hence, the cost of the policy that keeps the unit at $M + 1$ is at most $p_t(y) \cdot b \cdot t$. Now, consider a policy that does not keep the unit at location $M + 1$. The cost of such a policy is at least $(1 - p_t(y)) \cdot t \cdot h_{\text{min}}$, where $h_{\text{min}} > 0$ is the minimum holding cost in the system. Since for every $t$, $p_t(y)$ is non-increasing in $y$ and goes to 0 as $y$ goes to infinity, there exists some $Y(t)$ such that $p_t(y) \cdot b \cdot t < (1 - p_t(y)) \cdot t \cdot h_{\text{min}}$ for every $y > Y(t)$. This means that for any $t$, if $y > Y(t)$, then $\hat{U}_t^*(M + 1, y) = \{M + 1\}$. Now, let $Y_{\text{max}} = \max \{Y'_t, \max_{t < t_{\text{max}}} Y(t)\}$.

The above Lemma tells us that under any optimal policy, a unit will not be ordered from the outside supplier, unless the distance of the corresponding customer is $Y_{\text{max}}$ or less. Assuming that the initial state of the system is such that no unit with a corresponding customer distance larger than $Y_{\text{max}}$ is in locations $1, \ldots, M$, this ensures that under any optimal policy, no unit with a corresponding customer distance greater than $Y_{\text{max}}$ will be anywhere except the outside supplier throughout the time horizon. Hence, when trying to find an optimal decoupled policy, we only need to focus on states with a $y$ component less than or equal to $Y_{\text{max}}$.

We next define a mapping $M_k(z, y)$ that is linked to optimal policies for the subproblem.

**Definition 4.3.3.** For any $k, z > 0$ and $y > 0$ and $y \leq Y_{\text{max}}$, we define $M_k(z, y)$ as follows:

- Let $M_k(z, y) = \max \left\{ u | u \in \hat{U}_k^*(z, Y_{\text{max}}) \right\}$ for $y = Y_{\text{max}}$.

- Starting with $z = M + 1$ and $y = Y_{\text{max}}$, compute the values of $M_k(z, y)$ recursively first for decreasing values of $y$ until $z = M + 1$ and $y = 1$ and then by decreasing the value of $z$ by 1 and going through values $y = Y_{\text{max}}, Y_{\text{max}} - 1, \ldots, 1$ for $z = M$, etc., using the following formula:

$$M_k(z, y) = \max \left\{ u | u \in \hat{U}_k^*(z, y) \text{ and } u \leq M_k(z', y + 1) \text{ for any } z' \geq z \right\}$$  \hspace{1cm} (4.1)
The idea here is to choose a particular decoupled and committed policy, that is also monotonic. The recursion is defined in such a way that out of all possible choices from the \( \hat{U} \) sets, we choose the one that moves a unit as little as possible, subject to not violating monotonicity. However, in order for this to work, the maximum in the recursion needs to be attained for every point. This is what we show next.

**Lemma 4.3.2.** For every \( (k, z, y) \), the maximum in equation (4.1) is attained by some \( u \in \{1, 2, \ldots, M + 1\} \).

**Proof.** Suppose the statement is not true. Then, there exists some \( (k, z, y) \), such that the maximum in equation (4.1) does not exist (i.e. the corresponding set of \( u \)'s is empty) and this is the first such \( (z, y) \) that is encountered in the recursion described in Definition 4.3.3. However, in this case, no monotonic policy can be optimal, as we will demonstrate next. To do this, we first describe a monotonic state of the system. Recall that by the third part of Lemma 2.0.1, any optimal policy has to consist of actions that belong to the \( \hat{U} \) sets. We then show that if the maximum is not attained, then any such policy (hence any optimal policy) will be forced to act in a non-monotonic fashion at the monotonic state that is described. Hence, no monotonic policy can be optimal, which is a contradiction.

To construct the monotonic state that demonstrates the contradiction, we describe a sequence of points in the \( (z, y) \) space. An example is given in Figure 4-2, and the figure may aid in following the construction. Let \( (z^y, \hat{y}) \) be the first point in the recursion where the maximum in equation (4.1) does not exist. Now, consider the following sequence of points \( \{(z^y, \hat{y}), (z_{y+1}, \hat{y} + 1), (z_{y+2}, \hat{y} + 2), \ldots, (z_{y_{\max}}, y_{\max})\} \) in the \( (z, y) \) space. Given \( (z_{y}, y) \), determine \( (z_{y+1}, y + 1) \) as follows: Look at the values of \( M_k(z', y + 1) \) for all \( z' \geq z_y \) and choose a \( z' \) that results in the smallest value. Note that this value has to be greater than or equal to \( M_k(z_{y}, y) \), by equation (4.1). Hence, in this sequence, the value of \( M_k(z_{y}, y) \) is a non-decreasing function of \( y \), with a number of jumps between which \( M_k(z_{y}, y) \) stays constant.

We want to concentrate on these jump points, since these points provide us with some concrete information about the \( \hat{U} \) sets. Let \( (z_{y_i}, \hat{y}_i) \) be the \( i^{th} \) jump point, i.e. the \( i^{th} \) point in the sequence for which \( M_k(z_{y_i+1}, y + 1) > M_k(z_{y}, y) \). Let \( n \) be the number of jump points.
Figure 4-2: The figure shows an example of how to construct a monotonic state that demonstrates a contradiction, if the maximum in equation 4.1 is not attained (See the proof of Lemma 4.1). The table has the $M_k(z, y)$ values. In the example, (12, 3) is the first point where the maximum in the equation is not attained. Then, a sequence of points is followed in the $(z, y)$ space, as described in the proof. On this path, there are three jump points, at (13, 6) (from 6 to 8), at (14, 10) (from 8 to 9), and at (14, 11) (from 9 to 13). These jump points, along with the beginning and end points of the path, determine the monotonic state that results in a contradiction.
For all the jump points, we have some information about the sets $\hat{U}_k^*(z_{yi}, y_i)$. In particular, we know that any value in the set \(\{M_k(z_{yi}, y_i) + 1, M_k(z_{yi}, y_i) + 2, \ldots, M_k(z_{yi} + 1, y_i + 1)\}\) cannot be in the set $\hat{U}_k^*(z_{yi}, y_i)$. This is clear by equation (4.1), since otherwise $M_k(z_{yi}, y_i)$ would be higher. Now, consider a monotonic state, that has the following units with the corresponding customer distances:

- A unit at location $z_{\hat{y}}$ with a corresponding customer distance of $\hat{y}$. Label this as unit $c_0$.

- $n$ units with unit location and customer distance values $(z_{y_1}, y_1), (z_{y_2}, y_2), \ldots, (z_{y_n}, y_n)$. Label these units as units $c_1$ through $c_n$.

- A unit at location $z_{Y_{\max}}$ with a corresponding customer distance of $Y_{\max}$. Label this as unit $c_{n+1}$.

Clearly, there are monotonic states that have these units as a part, since the customer distance of a unit at a lower indexed location is lower for all these units that are described. Now, consider an optimal policy that is also monotonic. We look at the behaviour of such a policy in particular about the units that are described:

- Since $M_k(z_{Y_{\max}}, Y_{\max}) = \text{max}\{u | u \in \hat{U}_k^*(z_{Y_{\max}}, Y_{\max})\}$, any optimal policy will ship unit $c_{n+1}$ to location $M_k(z_{Y_{\max}}, Y_{\max})$ or lower.

- Because $n$ is the last jump point and hence $M_k(z_{y_n + 1}, y_n + 1) = M_k(z_{Y_{\max}}, Y_{\max})$, we know that any action in the set \(\{M_k(z_{y_n}, y_n) + 1, M_k(z_{y_n}, y_n) + 2, \ldots, M_k(z_{Y_{\max}}, Y_{\max})\}\) cannot be in $\hat{U}_k^*(z_{y_n}, y_n)$. This means that any monotonic and optimal policy will ship unit $c_n$ to location $M_k(z_{y_n}, y_n)$ or lower.

- Now, any action in the set \(\{M_k(z_{y_{n-1}}, y_{n-1}) + 1, M_k(z_{y_{n-1}}, y_{n-1}) + 2, \ldots, M_k(z_{y_n}, y_n)\}\) cannot be in $\hat{U}_k^*(z_{y_{n-1}}, y_{n-1})$. This means that any monotonic and optimal policy will ship unit $c_{n-1}$ to location $M_k(z_{y_{n-1}}, y_{n-1})$ or lower.
• We can repeat the same argument for all the units $c_i$ and finally we get that any monotonic and optimal policy will ship unit $c_1$ to location $M_k(z_{y_1}, y_1)$ or lower. Note that $M_k(z_{y_1}, y_1) = M_k(z_{y+1}, y+1)$ since this is the first jump point.

• Any monotonic policy will have to ship unit $c_0$ to location $M_k(z_{y+1}, y+1)$ or lower to preserve monotonicity. However, since $(z_{y}, y)$ is a point for which the set in equation (4.1) is empty, this particular action $u$ cannot be in $\tilde{U}_k^*(z_{y}, y)$. This means that no monotonic policy can be optimal.

Let us label the set of subproblem states such that $z \leq M$ and $y > Y_{\text{max}}$ as prohibited states. Note that if the initial state of a subproblem is not of prohibited type, then the subproblem will not reach a prohibited state under any optimal subproblem policy, by Lemma 4.3.1. Let $\tilde{\mu}_t(z, y) = M_{T-t}(z, y)$ for all $t, z$ and $y \leq Y_{\text{max}}$ and $\tilde{\mu}_t(M + 1, y) = M + 1$ for all $t$ and all $y > Y_{\text{max}}$. This results in functions $\tilde{\mu}_t$ that satisfy $\tilde{\mu}_t(z, y) \in \tilde{U}_k^*(z, y)$ for all time periods and all subproblem states, except for the prohibited ones. By an argument virtually identical to Lemma 2.0.1(3), choosing the decision according to $\tilde{\mu}_t$ for each unit at each time step constitutes an optimal (and also decoupled and committed) policy, for all initial states such that the corresponding subproblem states are not of prohibited type.  

Furthermore, by our construction, $\tilde{\mu}_t(z, y) \leq \tilde{\mu}_t(z', y')$ for any $z \leq z'$ and $y \leq y'$, hence this policy is also monotonic. In Chapter 3, we had shown that a monotonic, committed and decoupled policy is an echelon base stock policy. In this case, the monotonic, committed and decoupled policy that we have is not as simple as an echelon base stock policy but is a policy that can be described by $(M + 1) \cdot M/2$ threshold values, which we call the extended base stock levels. We next define these values:

Definition 4.3.4. We define the extended base stock levels $S_i^{t,j}$ as follows: Fix some $i, j$ and $t$ such that $i > j \geq 1$. There are two alternatives:

\[ \text{[The difference here arises because Lemma 2.0.1(3) applies when } \tilde{\mu} \text{ is defined at all states, whereas we defined } \tilde{\mu} \text{ at all states except for a set of states that are impossible to reach under any optimal policy.} \]
Figure 4.3: The figure shows the structure of the optimal policy for the single unit subproblem that is obtained through the analysis in this section. Consider a certain stage \( i \). The customer distance axis is segmented into several intervals for this stage, where the end points of the intervals are determined by the extended echelon base stock levels. Given that a unit is at stage \( i \), the decision of whether and where to ship the unit depends on in which interval the distance of the corresponding customer falls. This figure depicts a situation, where the unit is at stage 4 and the distance of the corresponding customer is 5. Since 5 is between \( S_4^{4.3} = 5 \) and \( S_4^{4.2} + 1 = 4 \), the unit is shipped to stage 3.

(i) \( \max\{y | \mu_t(i, y) \leq j\} \) exists: Let \( S_t^{i,j} = \max\{y | \mu_t(i, y) \leq j\} \).

(ii) \( \mu_t(i, y) > j \) for all \( y \): Let \( S_t^{i,j} = -\infty \).

Note that by this construction, given the extended base stock levels \( S_t^{i,j} \), the whole function \( \mu_t \) is determined, since we know that \( \mu_t \) is non-decreasing in \( y \) and \( S_t^{i,j} \) are exactly the points where the function increases. Between these step points, the function is constant.

Let us now look at what kind of an ordering rule this decoupled and optimal policy corresponds to in the overall problem. Let \( I_t^i \) be the echelon inventory position of stage \( i \) at the beginning of period \( t \), (The echelon inventory position of a stage \( i \) is the total number of units at stages 1 through \( i \) minus the backlog.) Consider a particular stage \( i > 1 \) and assume that \( I_t^i > I_t^{i-1} \). Otherwise, there are no units at stage \( i \) and there is no decision to make. There are \( I_t^i - I_t^{i-1} \) units at stage \( i \). Since the state is monotonic, these units correspond to customers with distances \((I_t^{i-1} + 1)^+ + 1, (I_t^{i-1} + 2)^+ + 1, \ldots, (I_t^i)^+ + 1\). Given these customer
distances and the extended echelon base stock levels, the number of units that need to be shipped from stage \( i \) to other stages is determined. We now define this policy:

**Definition 4.3.5.** We define a policy to be of extended echelon base stock type, if there exist integers \( S_t^{i,j} \) for all pairs of stages \( (i, j) \) where \( i \geq j \geq 1 \) and for all \( t \) such that the quantity to be shipped from \( i \) to \( j \) is:

\[
\left( \min\{(I_t^i)^+, S_t^{i,j}\} - \max\{(I_t^{i-1})^+, S_t^{i,j-1}\} \right)^+
\]

Note that shipping from \( i \) to \( i \) actually means that the unit is kept at stage \( i \). Also, note that an echelon base stock policy is a special case of an extended echelon base stock policy where \( S_t^{i,j} = -\infty \) for every \( j < i - 1 \). In that case, the number of units shipped from stage \( i \) to stages other than the next stage \( i - 1 \) are 0. The number of units shipped from stage \( i \) to stage \( i - 1 \) are:

\[
\left( \min\{(I_t^i)^+, S_t^{i,i-1}\} - (I_t^{i-1})^+ \right)^+
\]

which is equivalent to the number of units shipped to stage \( i - 1 \) under an echelon base stock policy with base stock level \( S_t^{i,i-1} - 1 \) for stage \( i - 1 \).

### 4.4 Infinite Horizon Analysis

This section deals with the case where the planning horizon is infinite. The analysis of this problem is very similar to the analysis in Section 3.4. In particular, once the limiting infinite horizon policy is defined properly for this problem, the analysis corresponding to Subsections 3.4.1 and 3.4.2 where the optimality of such a policy is established goes through with virtually no change (except for obvious modifications). Thus, in this section, we provide only the first part of the infinite horizon analysis where the limiting infinite horizon policy is defined.

In the infinite horizon setting, we consider stationary policies. A stationary policy is one of the form \((\mu, \mu, \ldots)\), with \( \mu : X \mapsto U \), so that the decision at each time is a function of the current state but not of the current time. In the infinite horizon context, we refer to
a stationary policy of this type as policy \( \mu \). Let \( \Omega \) denote the set of all stationary infinite horizon policies.

Similarly, for the subproblems, we refer to a stationary policy of the form \((\hat{\mu}, \hat{\mu}, \ldots)\) with \( \hat{\mu} : Z \times Y \mapsto \hat{U} \) as policy \( \hat{\mu} \). Given a fixed discount factor \( \alpha \in [0, 1] \), let \( \hat{J}_\alpha^T(z, y) \) and \( \hat{J}_\alpha^*(z, y) \) be the infinite horizon expected total discounted cost of policy \( \hat{\mu} \), and the corresponding optimal cost, respectively. Let \( \hat{J}_T^T(z, y) \) be the expected total discounted cost of using the stationary policy \( \hat{\mu} \) in a subproblem over a finite horizon of length \( T \), given that the initial state of the subproblem is \((z, y)\).

**Definition 4.4.1.** For any \( z \in Z, y \in Y \), let \( \hat{U}_\infty^*(z, y) \subset \{1, 2, \ldots, M + 1\} \) be the set of all decisions that are optimal if a subproblem is found at state \((z, y)\).

At this point, in order to make the analysis more tractable, we are going to introduce an additional assumption. This assumption will allow us to extend the finite horizon results to infinite horizon without much difficulty.

**Assumption 4.4.1.** Assume that the infinite horizon subproblem has a unique optimal solution. In other words, assume that the sets \( \hat{U}_\infty^*(z, y) \) are singletons for every \( z \) and \( y \).

This assumption is clearly not true for many instances of the problem. However, the following argument explains our rationale for using this assumption. Even though the optimal solution to the infinite horizon subproblem may not be unique in many instances of the subproblem, one can perturb the data of the problem slightly to get another instance that has a unique optimal solution. We conjecture that one can do this in a way such that the perturbation is small enough so that the optimal policy for this perturbed subproblem will be optimal for the original subproblem as well. From this point on, the analysis is carried out using Assumption 4.4.1

**Lemma 4.4.1.** For any fixed \( \alpha \in [0, 1] \), and any \( z, y \), we have

\[
\lim_{T \to \infty} \hat{J}_T^*(z, y) = \hat{J}_\infty^*(z, y).
\]

Furthermore, there exists an integer \( \tilde{t} \) such that \( \hat{U}(z, y) = \hat{U}_\infty^*(z, y) \) for every \( z, y \) and every \( t > \tilde{t} \). 

77
Proof. For any given initial state \((z, y)\) for a subproblem, the number of possible future states is finite; this is because \(y\) and \(z\) cannot increase. Therefore, general results for finite-state Markov decision problems apply. When \(\alpha < 1\), the convergence of \(\hat{J}_t^*(z, y)\) to \(\hat{J}_\infty^*(z, y)\) is immediate. When \(\alpha = 1\), we have a “stochastic shortest path problem,” (See [7]) and any “improper” policy (that is, any policy that is not guaranteed to eventually deliver the unit to the customer) incurs infinite cost, due to eternal backlogging. Under this condition, the claimed convergence is again guaranteed.

To prove the last part of the lemma, note that the optimality of a particular decision can be expressed in terms of an associated Bellman equation both for finite and infinite horizon, since we have a finite state, finite control problem. In particular, let

\[
V_t(z, y, u) = g(z, y, u) + E_d \left[ \hat{J}_{t-1}^*(f(z, y, u, d)) \right]
\]

where \(d\) is the random demand per period and \(f\) is the state transition mapping. Let \(V_\infty(z, y, u)\) be the analogous term for the infinite horizon subproblem (using \(\hat{J}_\infty^*\) instead of \(\hat{J}_{t-1}^*\)). Then, a control \(u'\) is optimal when \(t\) periods are remaining (i.e. \(u' \in \hat{U}_t^*(z, y)\)) if and only if it minimizes \(V_t(z, y, u)\). Similarly, \(u'\) is optimal for the infinite horizon subproblem (i.e. \(u' \in \hat{U}_\infty^*(z, y)\)) if and only if it minimizes \(V_\infty(z, y, u)\). Now, since we know that \(\hat{J}_t \to \hat{J}_\infty\), clearly \(V_t \to V_\infty\). This means that for every \(z, y, u\) and every \(\epsilon > 0\), there exists some \(t(z, y, u, \epsilon)\) such that:

\[
V_\infty(z, y, u) - V_t(z, y, u) < \epsilon, \quad \forall t > t(z, y, u, \epsilon)
\]

Let \(t(c) = \max_{z, y, u} t(z, y, u, \epsilon)\). Now, fix some state \(z, y\) and let \(u'\) be the action that minimizes \(V_\infty(z, y, u)\). By definition, any action that is not in the optimal action set has a \(V\) value that is strictly larger. Then, there exists some \(\delta > 0\), such that any action \(u''\) that is not in \(\hat{U}_\infty^*(z, y)\) has \(V_\infty(z, y, u'') > V_\infty(z, y, u') + \delta\). For \(t > t(\delta/2)\) and for every \(u\) we have:

\[
V_\infty(z, y, u) - V_t(z, y, u) < \delta/2
\]

78
Now, assume that \( u'' \notin \hat{U}_\infty^*(z, y) \). Then, for \( t > t(\delta/2) \):

\[
V_t(z, y, u'') + \delta/2 > V_\infty(z, y, u'') \\
> V_\infty(z, y, u') + \delta \\
\geq \min_u V_t(z, y, u) + \delta
\]

where the last inequality follows since having more periods can only increase costs. Hence, we have:

\[
V_t(z, y, u'') > \min_u V_t(z, y, u) + \delta/2
\]

which means that \( u'' \) is not in \( \hat{U}_t^*(z, y) \). Therefore, we have \( \hat{U}_t^*(z, y) \subset \hat{U}_\infty^*(z, y) \) for every \( t > \bar{t} \), where \( \bar{t} = t(\delta/2) \). Since the sets \( \hat{U}_\infty^*(z, y) \) are singletons, and since the sets \( \hat{U}_t^*(z, y) \) are non-empty, this subset relation implies that for \( t \) large enough, the finite horizon optimal action sets are singletons as well and \( \hat{U}_t^*(z, y) = \hat{U}_\infty^*(z, y) \) for every \( t > \bar{t} \).

**Definition 4.4.2.** For any \( z > 0 \) and \( y > 0 \), and \( y \leq Y_{\max} \) we define \( M(z, y) \) as follows:

- Let \( M(z, y) = \max \{ u | u \in \hat{U}^*(z, Y_{\max}) \} \) for \( y = Y_{\max} \).

- Starting with \( z = M + 1 \) and \( y = Y_{\max} \), compute the values of \( M(z, y) \) recursively first for decreasing values of \( y \) until \( z = M + 1 \) and \( y = 1 \) and then by decreasing the value of \( z \) by 1 and going through values \( y = Y_{\max}, Y_{\max} - 1, \ldots, 1 \) for \( z = M \), etc., using the following formula:

\[
M(z, y) = \max \{ u | u \in \hat{U}_\infty^*(z, y) \text{ and } u \leq M(z', y + 1) \text{ for any } z' \geq z \} \quad (4.2)
\]

**Lemma 4.4.2.** For every \( (z, y) \), the maximum in equation (4.2) is attained by some \( u \in \{1, 2, \ldots, M + 1\} \) (i.e. the set on the right hand side of equation (4.2) is non-empty).

**Proof.** Suppose not. Note that the second part of Lemma 4.4.1 indicates that the \( \hat{U}_\infty^* \) sets and \( \hat{U}_t^* \) sets are the same for large enough \( t \). Since the recursions for the finite and infinite horizon versions have the same structure and rely only on the corresponding \( \hat{U} \) sets, this
means that the maximum is not attained for the some \((z, y)\) and some \(t\) large enough. This is a contradiction with Lemma 4.3.2. \(\square\)

**Proposition 4.4.1.** Let \(\hat{\mu}(z, y) = M(z, y)\) for all \(z\) and \(y \leq Y_{\text{max}}\) and \(\hat{\mu}(M + 1, y) = M + 1\) for all \(y > Y_{\text{max}}\). Then, the policy \(\hat{\mu}\) is optimal for the subproblem, given that the initial state is not prohibited.

**Proof.** This determines an action for all the states except for the ones that have \(z \leq M\) and \(y > Y_{\text{max}}\), i.e., the prohibited states. (See the discussion following the proof of Lemma 4.3.2.) This, in addition with the construction of \(M(z, y)\) ensure that the functions \(\hat{\mu}\) satisfy \(\hat{\mu}(z, y) \in \hat{U}^*_\infty(z, y)\) for all states that are not prohibited. Hence \(\hat{\mu}\) is optimal for the subproblem by the definition of \(\hat{U}^*_\infty(z, y)\), given that the initial state is not prohibited. \(\square\)

**Proposition 4.4.2.** Let \(\mu^*\) be the stationary, decoupled, and committed policy for the overall problem that uses the optimal subproblem policy \(\hat{\mu}\) of Prop. 4.4.1 for each unit-customer pair. Then, \(\mu^*\) is a stationary extended echelon base stock policy.

**Proof.** By definition, \(\mu^*\) is decoupled and committed. Furthermore, it has the same structure as its finite horizon counterparts, since the \(\hat{\mu}\) was constructed in the same way (through a recursion of the \(M_k\) function). In particular, this policy can be summarized by \((M + 1) \cdot M/2\) extended base stock levels, and is an extended base stock policy. \(\square\)

**Theorem 4.4.1.** If the shipping costs are supermodular, then the set of extended echelon base stock policies is optimal for both the discounted cost and the average cost criteria.

**Proof.** We have so far constructed a stationary extended echelon base stock policy \(\mu^*\). This policy is constructed as a limit of optimal policies for the corresponding finite horizon problems. It should then be no surprise that \(\mu^*\) is optimal for the infinite horizon problem. Some careful limiting arguments are needed to make this rigorous. However, these arguments are virtually identical to the ones in Subsections 3.4.1 and 3.4.2 and thus we omit them. \(\square\)
4.5 Discussion

In this section, we discuss the relationship between our model and the model of Lawson and Porteus [27], along with some possible extensions.

4.5.1 Comparison with Lawson and Porteus [27]

The paper by Lawson and Porteus [27] deals with a problem that is very similar to ours, i.e., a serial system with expediting. They present their model a little differently, though. In particular, they have a serial system and the goods have to go through all the stages in this serial system. However, the transition from one stage to the next can take one period (via the regular shipment method) or can be instantaneous, which is called expedited delivery. The expedited delivery and the regular shipment from a stage to the next has a particular cost associated with it. In addition, a unit can be expedited through several stages in the system within the same period. The cost of such a shipment will be the sum of all the expediting costs for the stages that the unit goes through. Even though this seems like a quite different model at first, the two models are similar. In particular, in both systems, goods can go through the system fast or slow, by incurring appropriate costs. In [27], the cost of expediting through several stages is additive, whereas we assume supermodularity on the cost structure. Even though their problem is not a special case of ours automatically, there are two ways in which the problems are closely related. First, a small modification in the model of [27] makes their problem a special case of ours. Second, a small modification in our analysis allows our method to analyze their problem. In the previous sections, we chose to present our model without this modification though, in order to make the exposition more simple. We now briefly explain these two modifications.

Figure 4-4 shows the sequence of events in [27]. In every period, expediting decisions and regular shipment decisions are made. The difference between these two is that the expedited orders arrive before the demand is observed and the regular orders arrive after. Hence, if a unit is expedited to stage 1, it can be used to satisfy customer demand in the current
period. However, note that this advantage is only relevant when the unit is shipped to stage 1. Suppose that we have a particular unit at a stage $i$ and we are considering shipping this unit to stage $j > 1$. Since $j$ is not the final stage, this unit will not be used to satisfy demand in the current period. One way of shipping the unit from $i$ to $j$ is to expedite it through all the stages between them. This has a cost that is the sum of all the one step expediting costs. Another way is to expedite the unit from $i$ to $j + 1$ and then to ship the unit via regular shipment from $j + 1$ to $j$. Since the regular shipment cost from $j + 1$ to $j$ is at most the expedited shipment cost from $j + 1$ to $j$, this way of shipping has a smaller cost. Since the unit will not be used to satisfy demand in the current period, the two ways are equivalent in terms of the location of the unit at the beginning of the next period. Hence, we can restrict our attention to policies that use regular shipment for the last step whenever shipping a unit from a stage $i$ to a stage $j > 1$. This means that for shipments other than the ones to stage 1, we can now merge the expedited and regular shipment decisions into a single decision, just like in our model. Now, the difference between our model and the one in [27] arises because of shipments to stage 1. In particular, in [27], a unit that is expedited into stage 1 can be used to satisfy the current period’s demand.

As we mentioned above, this difference can be eliminated in two different ways, either by slightly modifying their model or by slightly modifying our analysis. We start with the first option. Suppose that in their model, a unit that is expedited could not be used to satisfy demand in the current period. In this case, when we would like to ship a unit from a stage $i$
to stage 1, it is optimal to expedite the unit from $i$ to 2 and then to use regular shipment for
going from 2 to 1. Hence, regular and expediting order decisions can be merged, and now
the problem is a special case of our model that we presented in Section 4.2. However, there
is one more thing we need to check, i.e. the supermodularity of the shipping costs.

Consider the model in [27]. Let $c_i$ be the cost of expediting from stage $i$ to stage $i - 1$
and let $c'_i$ be the cost of regular shipment from $i$ to $i - 1$. With the small modification that
we made, we know that a shipment from $i$ to $j$ will involve expedited shipments from $i$ to
$j + 1$ and a regular shipment from $j + 1$ to $j$. Hence, the cost of shipping from $i$ to $j$ is:

$$ c_{i,j} = \sum_{n=i}^{j+2} c_n + c'_{j+1} $$

where we use the definition $\sum_{n=a}^{b} f(n) = 0$ if $b < a$. Now, let us see if this cost structure is
supermodular. For $i > j > k > l$:

$$ c_{i,l} + c_{j,k} = \sum_{n=i}^{l+2} c_n + c'_{l+1} + \sum_{n=j}^{k+2} c_n + c'_{k+1} $$

$$ = \sum_{n=i}^{k+2} c_n + c'_{k+1} + \sum_{n=j}^{l+2} c_n + c'_{l+1} $$

$$ = c_{i,k} + c_{j,l} $$

At this point, we have demonstrated that except for the case where $j = k$, the supermod-
ularity assumption is satisfied, and in particular, the required inequality in the definition of
supermodularity (Definition 4.3) is an equality in this case. Suppose that $j = k$. Then:

$$ c_{i,l} + c_{j,k} = \sum_{n=i}^{l+2} c_n + c'_{l+1} + 0 $$

$$ = \sum_{n=i}^{k+2} c_n + c'_{k+1} + \sum_{n=j}^{l+2} c_n + c'_{l+1} + (c_{k+1} - c'_{k+1}) $$

$$ = c_{i,k} + c_{j,l} + (c_{k+1} - c'_{k+1}) $$

$$ \geq c_{i,k} + c_{j,l} $$

since $c_{k+1} \geq c'_{k+1}$ (cost of expediting is at least as much as the cost of regular shipment).
Hence, the supermodularity assumption is satisfied, and the problem in [27] with the small modification is a special case of our problem.

Now, we explain the other option, namely modifying our analysis slightly in order to handle the model in [27]. As we explained before, the difference between the two models arises from the fact that a unit that is expedited to stage 1 becomes available to satisfy demand in the current period whereas a unit that is shipped regularly to stage 1 does not. Hence, we can take care of this subtlety by differentiating between these two actions. In particular, we introduce two artificial stages to the system, as depicted in Figure 4-5. We now have three locations corresponding to stage 1. Shipping to 1 will correspond to using regular shipment for the last step, hence the unit will not be available to be given to a customer within the current period. Shipping to 1A (1-Available) corresponds to expediting the last step, i.e., the unit is available to be given to any customer. Finally, shipping to 1CA (1-Committed-Available) corresponds to expediting the last step as well but the unit is available to be given only to the corresponding customer. (This last one 1-CA is introduced only in order to make the proof of our analysis possible. This can be introduced at the appropriate point in the proof).

Using such a modification and by setting the appropriate costs for the different situations, the model in [27] can be handled via the single unit approach as well.

The optimal policies in [27] are called “top-down base stock policies”. Similar to our extended base stock policies, they have a number of thresholds and the optimal ordering levels can be determined via a simple computation, given the echelon inventory positions and these threshold values. However, their policies require only $2M + 1$ threshold levels, whereas our policies require $(M + 1) \cdot M/2$ threshold levels. Since their model has additional structure (the cost structure is additive, i.e., the cost of expediting through several stages is the sum of the one step expediting costs), it is not surprising that a somewhat simpler policy can be shown to be optimal.
Figure 4-5: Example system with artificial locations to model availability to serve customer demand within the period. (Not all possible shipments are shown).

4.5.2 Extensions

In this subsection, we discuss some possible extensions to our model. In particular, we focus on the leadtime and the demand models.

In our model, the leadtimes for shipping from one stage to any other stage was assumed to be equal to one period. However, in many real world systems, longer leadtimes are required. Now we discuss the extent to which such leadtimes can be handled via our approach. If we have two stages $i + 1$ and $i$ and the leadtime between these two stages is greater than 1, we can allow this as long as the cost of expediting from a stage $j \geq i + 1$ to a stage $k \leq i$ is prohibitively expensive. The reason is, once a unit is in transit between $i + 1$ and $i$, we want to make sure that no other unit will overtake this unit. Another way to see this is as follows: To model the leadtime between $i + 1$ and $i$, we introduce artificial locations between these stages with unit leadtimes between each other. However, in order to accurately model this leadtime, we need to make sure that a unit in transit between $i + 1$ and $i$ goes through all
these artificial locations. This can be achieved by making the expediting cost from any of these artificial locations to a location other than the next one prohibitively high. However, then, in order to preserve the supermodularity of the cost structure, any shipment from a stage \( j \geq i + 1 \) to a stage \( k \leq i \) needs to be prohibitively expensive as well.

So, this analysis cannot handle situations where there is a large leadtime between two stages and it is also possible to expedite over these stages. This raises the question, whether another approach can yield the same result in such a situation. The answer is no, since in such a case, optimal policies may be potentially much more complicated. For example, consider the special case, where there is only a single location. Suppose that the regular order leadtime from the outside supplier and the expedited order leadtime differ by more than one period. Then optimal policies are more complicated than extended base stock policies, as can be seen in [37]. To see this, suppose that at some time, there are a number of units in transit between the outside supplier and the single stage. At this point, if a demand arrives, it may be optimal to expedite an extra unit from the outside supplier to serve the customer demand rather than waiting for the units in regular transit mode to arrive. Hence, it may be optimal to order a unit even though the echelon inventory position is not below the threshold level.

Another possible extension in leadtimes is to allow for stochastic leadtimes. Our method can handle stochastic leadtimes, as long as monotonicity is satisfied. However, even though there are many leadtime models that satisfy this criterion for this expediting model, at this point, we have not found any natural leadtime model that is interesting enough and that can be explained in a concise manner.

We can extend the demand model to Markov modulated demand easily, just like in Chapter 3. In this case, the extended base stock levels would of course depend on the state of the modulating Markov chain.

Finally, note that we have not given an explicit algorithm that determines the extended echelon base stock levels. However, since we have shown that the problem can be decomposed into single unit subproblems, the extended base stock levels can easily be determined by
solving a single unit subproblem. Such a problem is quite simple to solve using dynamic programming. Once the subproblem is solved and the optimal action sets $\hat{U}$ are determined, the extended echelon base stock levels are readily available.
Chapter 5

Serial Systems with Batch Ordering

5.1 Introduction

This chapter deals with a serial system with batch ordering constraints. In particular, the number of units shipped from a stage $k$ to another are required to be a multiple of a fixed lot size $Q_{k-1}$. Such a restriction is very common in practice and may correspond to cases where goods are transferred between echelons in fixed truckloads or full containers, or when such a restriction is placed in order to achieve economies of scale in ordering costs.

This problem is the extension of the model of Clark and Scarf [16] to include batch ordering constraints. Chen [11] has recently shown that in such a system an echelon $(R, nQ)$ policy is optimal, when the objective is to minimize the long run average cost. In an echelon $(R, nQ)$ policy, if the echelon inventory level of a given stage $k$ at time $t$ falls below a level $R^k_t$, a number of batches of size $Q_k$ are ordered to bring the echelon inventory level back above $R^k_t$. Such a policy can be viewed as an echelon base stock policy, modified to respect the batch ordering constraints. Chen’s lower bounding method is applies to infinite horizon problems with the average cost criteria and it is not clear how the results can be extended to other types of planning horizons. This chapter extends Chen’s results to finite horizon and infinite horizon with discounted cost criteria. An efficient algorithm to calculate an optimal policy is provided. In addition, the chapter shows the effectiveness of the decomposition
approach in this setting.

\((R, nQ)\) policies have received significant attention from researchers. For single stage systems, Veinott [36] shows that an \((R, nQ)\) policy is optimal when orders need to be multiples of \(Q\). For serial systems, De Bodt and Graves [18], Axsater and Rosling [5], and Chen and Zheng [13] have analyzed \((R, nQ)\) policies as reasonable heuristics.

The rest of this chapter has five sections. Section 5.2 gives the problem formulation, Sections 5.3 and 5.4 analyze the finite and infinite horizon versions of the problem, respectively. Section 5.5 describes an algorithm to compute optimal policies. The chapter concludes with Section 5.6, the appendix, that gives the details of the proofs of two propositions.

### 5.2 Problem Formulation

We consider a single-item serial inventory system consisting of \(M\) stages, indexed by \(1, \ldots, M\). Customer demand can only be satisfied by units at stage 1. Any demand that is not immediately satisfied is backlogged. The inventory at stage \(k\) \((k = 1, \ldots, M - 1)\) is replenished by placing an order for units stored at stage \(k + 1\). Stage \(M\) receives replenishments from an outside supplier with unlimited stock. For notational simplicity, we label the outside supplier as stage \(M + 1\). The ordering quantity between two stages \(k + 1\) and \(k\) is restricted to be a multiple of a given quantity \(Q_k\), for all \(k\). We assume that these batch sizes are nested, i.e., we have:

\[
Q_{k+1} = n_k Q_k, \quad k = 1, \ldots, M - 1,
\]

for some integer \(n_k\).

We assume that the system is periodically reviewed and, therefore, a discrete-time model can be employed. We employ a linear cost structure. In particular,

(a) For each stage \(k\), there is an inventory holding cost rate \(h_k\) that gets charged at each time period to each unit at that stage. We assume that the holding cost rate \(h_{M+1}\) at the external supplier is zero.

(b) For each stage \(k\), there is a cost \(c_k\) for shipping the unit from stage \(k + 1\) to stage \(k\).
(c) There is a backorder cost rate $b$, which is charged at each time step for each unit of backlogged demand.

We assume that the holding cost and backorder cost parameters are positive, and that the shipping cost is non-negative.

For notational convenience, we assume i.i.d. demand and unit leadtimes. However, all results can be easily extended to larger deterministic leadtimes, stochastic leadtimes, and/or Markovian demand scenarios, as in Chapter 3.

5.3 Finite Horizon Analysis

In this section, we characterize the structure of optimal policies for the finite horizon version of the problem. In particular, we show that the set of $(R, nQ)$ policies is optimal. This extends the result of Chen [11] to the finite horizon case.

The analysis is based on the decomposition idea that was utilized in the previous two chapters. In this case, we show that the problem decomposes into a series of subproblems, each of which involves a single batch and the customers that correspond to the units in the batch. (Note that since these customers are consecutive customers, the distance of the last one completely determines the distance of all the customers associated with a batch). The problem is formulated so that once a batch is shipped to a particular stage, it splits into several smaller batches. Hence, the decomposition result requires slightly more attention than the previous two chapters. In particular, we use a recursive argument to show that the problem can be decomposed into subproblems with large batches first and then given that, into subproblems with smaller batches. We then analyze the structure of the resulting policies by investigating the dynamic programming recursion for the corresponding subproblems. This shows that the resulting policies are also monotonic, which then implies the optimality of $(R, nQ)$ policies.

To facilitate the decomposition analysis, we use the following state space: For each unit-customer pair $i$, $i \in \mathbb{N}$, we have a vector $(z^i_t, y^i_t)$, with $z^i_t \in \mathbb{Z} = \{0, 1, \ldots, M + 1\}$ and
\( y_i^t \in Y = \mathbb{N}_0 \), where \( z_i^t \) is the location of unit \( i \) at time \( t \), and \( y_i^t \) is the distance of customer \( i \) at time \( t \). The state of the system consists of a countably infinite number of such vectors, one for each unit-customer pair, i.e.,

\[
    x_t = \left\{ (z_1^t, y_1^t), (z_2^t, y_2^t), \ldots \right\}.
\]

Note that since all the leadtimes are assumed to be equal to one period, we do not need to insert artificial locations to model units in transit. The only artificial location is location 0, that corresponds to units that are already delivered to customers. Any other location \( i \) corresponds to a stage \( i \) of the system.

The control vector is an infinite binary sequence, \( u_t = (u_1^t, u_2^t, \ldots) \), where the \( i^{th} \) component \( u_i^t \) corresponds to a "ship" or "hold" decision for the \( i^{th} \) unit. If unit \( i \) is at location 0, it is already delivered to a customer and is unaffected by \( u_i^t \). Otherwise, \( u_i^t = 0 \) corresponds to holding the unit at its current location, and \( u_i^t = 1 \) corresponds to shipping it to the next. The batch size restriction forces the total number of units shipped from a stage \( k + 1 \) to stage \( k \) (the number of units for which \( z_i^t = k + 1 \) and \( u_i^t = 1 \)) to be a multiple of \( Q_k \). Finally, if unit \( i \) is at stage 1, a decision \( u_i^t = 1 \) releases this unit so that it can be given to a customer. In case the number \( k \) of units released from stage 1 is larger than the number \( m \) of customers whose demand is backlogged, only \( m \) of these units are given to customers (i.e., move to location 0), and the remaining \( k - m \) units stay at location 1. Otherwise, all \( k \) units are given to customers. The rules about which units are given to customers and which customers receive an unit are the following: If unit \( i \) is released to be given to a customer and customer \( i \)'s demand is backlogged, then unit \( i \) is given to customer \( i \). After all such matchings are done, if there are extra units and customers, the units and customers with the lowest indices are chosen until one side is empty.

We again use analogous definitions for monotonic and committed policies as in Chapter 3. Monotonic policies refer to policies that maintain the monotonic order among the units and committed policies refer to policies that release a unit only if the corresponding customer has arrived.
Proposition 5.3.1. The set of monotonic and committed policies is optimal.

Proof. The proof is essentially the same as the proofs of Propositions 3.3.1 and 3.3.3 and we only provide a sketch. The set of monotonic policies is optimal, because whenever a number of units are to be shipped to the next stage, one can choose the ones with the lower index. Given a restriction to monotonic policies, one can pair units with customers and make sure that customer demands can be satisfied using the corresponding units, and hence the set of committed policies is optimal. □

Consider the overall problem, together with an additional restriction to committed policies. Such a restriction can be represented as a change in the dynamics of the system, instead of a restriction on the policy space. In particular, suppose that a unit $i$ is at location 1 at some time $t$. Then, restricting $u^i_t$ to be 0 when $y^i_t > 1$ has the same effect as defining the dynamics of the system so that $z^i_{t+1} = 0$ if and only if either $z^i_t = 0$ or $y^i_t = z^i_t = 1$ and $u^i_t = 1$. From now on, whenever we refer to the overall problem, we mean the problem with the modified dynamics as described above, so a restriction to committed policies will be implied.

Decoupled policies, as stated in Chapter 3 are clearly not optimal in this setting, due to the coupling that arises because of the batch sizing constraint. However, even though single unit problems are not decoupled, one can still use the idea of decomposability repeatedly to show that the overall problem consists of subproblems that involve not one but multiple units and multiple customers each. Before we proceed, some definitions are in order:

Definition 5.3.1. Let a $k$-unit subproblem be the problem that has the same dynamics as in the overall problem, except for the fact that the total number of units (including the ones at the outside supplier) and the total number of customers is $k$.

The state of the overall problem at time $t$ can be written as $(\tilde{z}_t, \tilde{y}_t)$, where $\tilde{z}_t = (z^1_t, z^2_t, \ldots)$ is the vector of all unit locations and $\tilde{y}_t = (y^1_t, y^2_t, \ldots)$ is the vector of all customer distances. For convenience, we use a similar notation for a $k$-unit subproblem. The state of a $k$-unit
subproblem can be written as $(\bar{z}_t, \bar{y}_t)$, where $\bar{z}_t = (z_{1t}^1, z_{1t}^2, \ldots, z_{kt}^k)$ is the vector of all $k$ unit locations and $\bar{y}_t = (y_{1t}^1, y_{1t}^2, \ldots, y_{kt}^k)$ is the vector of all $k$ customer distances.

A policy $\pi$ for the overall problem (for the $k$-unit subproblem) consists of a set of functions $\{\mu_t^i\}$ where each $\mu_t^i : Z^\infty \times Y^\infty \mapsto \hat{U} = \{0, 1\}$ ($\mu_t^i : Z^k \times Y^k \mapsto \hat{U} = \{0, 1\}$) prescribes the control for the $i$th unit at time $t$, according to $u_t^i = \mu_t^i(\bar{z}_t, \bar{y}_t)$. Let $\tilde{\mu}_t = (\mu_t^1, \mu_t^2, \ldots)$ ($\tilde{\mu}_t = (\mu_t^1, \mu_t^2, \ldots, \mu_t^k)$) be the part of this policy that applies at time $t$.

Whenever we use the vectors $\bar{z}_t$, $\bar{y}_t$ or $\tilde{\mu}_t$, the dimension will be apparent from the context.

**Definition 5.3.2.** (i) A state $(\bar{z}_t, \bar{y}_t)$ for the overall problem is *conforming*, if the total number of units in locations $0, 1, \ldots, i$ is an integer multiple of $Q_i$, for all $i = 1, \ldots, M$.

(ii) Now fix some $k$. A state $(\bar{z}_t, \bar{y}_t)$ for a $Q_k$-unit subproblem is *conforming* if the total number of units in locations $0$ through $i$ is an integer multiple of $Q_i$, for all $i = 1, \ldots, k$.

**Assumption 5.3.1.** We assume that the initial state of the overall problem is conforming.

Note that the above assumption, along with the fact that orders need to be multiples of batch sizes, guarantees that the state will be conforming throughout the horizon. Also note that in a conforming state, the number of units at stage $k$ is a multiple of $Q_{k-1}$, for all $k = 2, \ldots, M$. This is because of the assumption that $Q_k$ is an integer multiple of $Q_{k-1}$.

**Definition 5.3.3.** Let $a$ be a vector, finite or infinite dimensional, where $a^i$ is the $i$th component of $a$. We define $f_{j,b}(a)$ to be the $j$th subvector of size $b$ within $a$ (Assuming the dimension of $a$ is an integer multiple of $b$ if $a$ is finite dimensional). Formally:

$$f_{j,b}(a) = (a^{(j-1)b+1}, a^{(j-1)b+2}, \ldots, a^{jb})$$

Next, we make use of the result about decomposable systems from Chapter 2 repeatedly to show that the overall problem can be decomposed into subproblems that consist of multiple units and multiple customers, and then that these subproblems themselves consist of even smaller subproblems, etc.

Let $J^*_T(\bar{z}_0, \bar{y}_0)$ be the optimal cost of the overall problem with finite horizon $T$ and initial state $(\bar{z}_0, \bar{y}_0)$. Similarly, let $J^*_T(\bar{z}_0, \bar{y}_0)$ be the optimal cost of the $Q_k$-unit subproblem with finite horizon $T$ and initial state $(\bar{z}_0, \bar{y}_0)$. 
Proposition 5.3.2. The overall problem, with the modification of the dynamics introduced after the proof of Prop. 5.3.1, resulting in the restriction to committed policies, is decomposable into \( Q_M \)-unit subproblems. This means that:

(i) There exists an optimal policy \( \pi \) for the overall problem and functions \( \hat{\mu}_t^M : Z^{Q_M} \times Y^{Q_M} \mapsto \hat{U}^{Q_M} \) such that:

\[
f_{i,Q_M}(\hat{\mu}_t(z_i, y_i)) = \hat{\mu}_t^M \left( f_{i,Q_M}(z_i), f_{i,Q_M}(y_i) \right), \quad \forall i, t
\]

(ii) \n
\[
J_T^* (z_0, y_0) = \sum_{i=1}^{\infty} \hat{J}_T^M \left( f_{i,Q_M}(z_0), f_{i,Q_M}(y_0) \right)
\]

Proof. Since the initial state is conforming and since the batch sizes are nested, the overall system is decomposable, where the first subproblem consists of the first \( Q_M \) units and the first \( Q_M \) customers, the second subproblem consists of the second set of \( Q_M \) units and \( Q_M \) customers, and the \( i^{th} \) subproblem consists of the \( i^{th} \) group of \( Q_M \) units and customers. All the assumptions for decomposability are satisfied under this setting, hence Lemma 2.0.1 implies the result. \( \square \)

Proposition 5.3.3. The \( Q_k \)-unit subproblem, given that the initial state is conforming, and given that the \( Q_k \) units are not at stage \( k+1 \) or higher initially, is decomposable into \( Q_{k-1} \)-unit subproblems, for all \( k = 2, 3, \ldots, M \). This means that (when the initial state satisfies the conditions above):

(i) There exists an optimal policy \( \pi \) for the \( Q_k \)-unit subproblem and functions \( \hat{\mu}_t^{k-1} : Z^{Q_{k-1}} \times Y^{Q_{k-1}} \mapsto \hat{U}^{Q_{k-1}} \) such that:

\[
f_{i,Q_{k-1}}(\hat{\mu}_t(z_i, y_i)) = \hat{\mu}_t^{k-1} \left( f_{i,Q_{k-1}}(z_i), f_{i,Q_{k-1}}(y_i) \right), \quad \forall i, t
\]

(ii) \n
\[
\hat{J}_T^k(z_0, y_0) = \sum_{i=1}^{n_{k-1}} \hat{J}_T^{k-1} \left( f_{i,Q_{k-1}}(z_0), f_{i,Q_{k-1}}(y_0) \right)
\]

95
Proof. Given that the initial state is conforming, either all the $Q_k$ units will be at stage $k+1$ or higher, or all of them will be at stage $k$ or lower. The assumption is that they are at $k$ or lower. Hence, these units can be shipped in batch sizes of $Q_{k-1}$ or smaller. It then follows that the assumptions for decomposibility into $Q_{k-1}$-unit subproblems is satisfied. \qed

Let $c^k \cdot Q_{k-1}$ be the number of units at stage $k$ at the beginning of the horizon. (Assume $Q_0 = 1$ and since the initial state is conforming, $c^k$ are all non-negative integers). The two propositions above show that the overall problem can be decomposed into smaller subproblems. In particular, it can be decomposed into infinitely many subproblems, $c_k$ of which are $Q_{k-1}$-unit subproblems, for $i = 1, \ldots, M$ and the rest are $Q_M$-unit subproblems. In other words, the value of a particular state of the overall problem is the sum of the values of individual batches in the system. Moreover, there exists an optimal policy that ships or holds each batch by considering only the location of the batch and the distance of the corresponding customers.

Now that we have shown that the overall problem can be decomposed into smaller subproblems, we next analyze the structure of these subproblems in more detail. We show that there exists an optimal policy that has a certain monotonicity structure, and this monotonicity then leads to the optimality of $(R, nQ)$ policies.

Fix some $k$. Note that the value of a batch of size $Q_k$ at stage $k+1$ depends only on the remaining time and the distance of the final customer that corresponds to that batch. This follows, since the customers corresponding to a batch are consecutive customers, and given the distance of the last one, the distances of all of them are determined.

Suppose

$$\mathcal{Z}_t = (k+1, k+1, \ldots, k+1)_{Q_k \text{ components}}$$

and

$$\mathcal{y}_t = ((y - Q_k)^+ + 1, (y - Q_k)^+ + 2, \ldots, y)$$

is the state of a $Q_k$ unit subproblem, where the batch is at stage $k+1$ and the last customer corresponding to the batch has a distance of $y$. 

96
Then, we can write the value of this batch as:

\[ \tilde{J}^k_t(y) = \bar{J}^k_t(\tilde{z}_t, \tilde{y}_t). \]

In addition, we can describe the action of a \( Q_k \)-unit subproblem policy when the state is \((\tilde{z}_t, \tilde{y}_t)\) as described above as:

\[ \tilde{\mu}^k_t(y) = \begin{cases} 
0 & \text{if } \tilde{\mu}^k_t(\tilde{z}_t, \tilde{y}_t) = (0, 0, \ldots, 0) \\
1 & \text{if } \tilde{\mu}^k_t(\tilde{z}_t, \tilde{y}_t) = (1, 1, \ldots, 1) 
\end{cases} \]

i.e., \( \tilde{\mu}^k_t \) is the function that determines whether to ship this batch to stage \( k \) or not, depending on the distance of the last customer. (Note that shipping all or holding all \( Q_k \) units are the only two feasible controls).

**Proposition 5.3.4.** For every \( k \geq 1 \) and \( t \), there exists an optimal policy for the \( Q_k \)-unit subproblem that ships a batch of size \( Q_k \) to stage \( k \) if and only if the distance of the last customer corresponding to the batch is below a certain threshold \( \bar{y}^k_t \), where \( \bar{y}^k_t \in \mathbb{N} \cup \{\infty\} \cup \{-\infty\} \). In other words, there exists functions \( \bar{\mu}^k_t(y) \) that describe the action of this optimal policy when the \( Q_k \) units are at stage \( k + 1 \), such that \( \bar{\mu}^k_t(y) = 1 \) if and only if \( y \leq \bar{y}^k_t \).

**Proof.** See the appendix at the end of the chapter. \( \square \)

**Proposition 5.3.5.** The set of \((R, nQ)\) policies is optimal for the overall problem.

**Proof.** Using Propositions 5.3.2 and 5.3.3, a policy is optimal if it uses an action that is optimal for the \( Q_k \)-unit subproblem for every batch of size \( Q_k \) at stage \( k + 1 \), for all \( k \). One can choose the policies described in Proposition 5.3.4 at each stage, and since they are optimal for their respective subproblems, the resulting policy is optimal for the overall problem. Now, such a policy ships a batch if and only if the distance of the last customer is \( \bar{y}^k_t \) or less. Equivalently, it tries to set the echelon inventory position at stage \( k \) to a level above \( \bar{y}^k_t - 1 \). This is the same as an \((R, nQ)\) policy with \( R^k_t = \bar{y}^k_t - 1 \) for all \( k \) and \( t \). \( \square \)
5.4 Infinite Horizon Analysis

This section deals with the case where the planning horizon is infinite. We study the expected total discounted cost criterion.

In the infinite horizon setting, we consider stationary policies. A stationary policy is one of the form \((\mu, \mu, \ldots)\), with \(\mu : X \mapsto U\), so that the decision at each time is a function of the current state but not of the current time. In the infinite horizon context, we refer to a stationary policy of this type as policy \(\mu\). Let \(\Omega\) denote the set of all stationary infinite horizon policies.

For a \(Q_k\)-unit subproblem, given a fixed discount factor \(\alpha \in [0, 1]\), let \(\bar{J}_\infty^M(\bar{z}, \bar{y})\) be the optimal infinite horizon expected total discounted cost.

The next lemma relates the finite and infinite horizon versions of the \(Q_M\)-unit subproblem.

**Lemma 5.4.1.** For any \(k\), fixed \(\alpha \in [0, 1]\), and any state \((\bar{z}, \bar{y})\), we have

\[
\lim_{T \to \infty} 
J_T^k(\bar{z}, \bar{y}) = \bar{J}_\infty^k(\bar{z}, \bar{y}).
\]

**Proof.** For any given initial state for a \(Q_k\)-unit subproblem, the number of possible future states is finite, hence general results for finite-state Markov decision problems apply and the convergence is guaranteed. \(\square\)

Let

\[
\bar{J}_\infty^k(y) = \bar{J}_\infty^k(\bar{z}, \bar{y}),
\]

where

\[
\bar{z} = \left( k+1, k+1, \ldots, k+1 \right)_{Q_k \text{ components}}
\]

and

\[
\bar{y} = ((y - Q_k)^+ + 1, (y - Q_k)^+ + 2, \ldots, y)
\]

is the state of a \(Q_k\) unit subproblem where the batch is at stage \(k + 1\) and the last customer corresponding to the batch has a distance of \(y\).
In addition, we can describe the action of a $Q_k$-unit subproblem policy when the state is $(\bar{z}, \bar{y})$, as described above, as:

$$\bar{\mu}^k(y) = \begin{cases} 0 & \text{if } \mu^k(\bar{z}, \bar{y}) = (0, 0, \ldots, 0) \\ 1 & \text{if } \mu^k(\bar{z}, y) = (1, 1, \ldots, 1) \end{cases}$$

i.e., $\bar{\mu}^k$ is the function that determines whether to ship this batch to stage $k$ or not, depending on the distance of the last customer. (Note that shipping all or holding all $Q_k$ units are the only two feasible controls).

**Corollary 5.4.1.** For any $k$, fixed $\alpha \in [0, 1)$ and any $y$, we have

$$\lim_{T \to \infty} \bar{J}_T^k(y) = \bar{J}_\infty^k(y).$$

**Proposition 5.4.1.** For every $k \geq 1$, there exists an optimal policy for the $Q_k$-unit infinite horizon subproblem that ships a batch of size $Q_k$ to stage $k$ if and only if the distance of the last customer corresponding to the batch is below a certain threshold $\bar{y}^k$, where $\bar{y}^k \in \mathbb{N} \cup \infty \cup -\infty$. In other words, there exist functions $\bar{\mu}^k(y)$ that describe the action of this optimal policy when the $Q_k$ units are at stage $k + 1$, such that $\bar{\mu}^k(y) = 1$ if and only if $y \leq \bar{y}^k$.

**Proof.** See the appendix at the end of the chapter. \qed

**Proposition 5.4.2.** Let $\mu^*$ be a policy for the overall problem that chooses its actions at every stage as described in Proposition 5.4.1. Then, $\mu^*$ is a stationary $(R, nQ)$ policy.

**Proof.** Such a policy ships a batch if and only if the distance of the last customer is $\bar{y}^k$ or less. Equivalently, it tries to set the echelon inventory position at stage $k$ to a level above $\bar{y}^k - 1$. This is the same as an $(R, nQ)$ policy with $R^k = \bar{y}^k - 1$ for all $k$. \qed

We have so far constructed a stationary $(R, nQ)$ policy $\mu^*$. This policy is constructed as a limit of optimal policies for the corresponding finite horizon problems. It should then be no surprise that $\mu^*$ is optimal for the infinite horizon problem.

The infinite horizon optimal cost is defined by

$$J_\infty^k(\bar{z}, \bar{y}) = \inf_{\mu \in \Omega} J_\infty^\mu(\bar{z}, \bar{y}).$$
A stationary policy can be used over any time horizon, finite or infinite. Let $J_T^\mu (\bar{z}, \bar{y})$ be the expected total discounted cost of using the stationary policy $\mu$ during a finite planning horizon of length $T$, starting with the initial state $(\bar{z}, \bar{y})$. We then have

$$J_T^\mu (\bar{z}, \bar{y}) = \lim_{T \to \infty} J_T^\mu (\bar{z}, \bar{y}).$$

By Proposition 5.3.2, we have,

$$J_T^\star (\bar{z}, \bar{y}) = \sum_{i=1}^{\infty} \hat{J}_T^M (f_{i,Q_M}(z), f_{i,Q_M}(y)),$$

for any monotonic and conforming state $(\bar{z}, \bar{y})$. Hence, we have

$$J_T^\star (\bar{z}, \bar{y}) = \inf_{\mu \in \Omega} \lim_{T \to \infty} J_T^\mu (\bar{z}, \bar{y})$$

$$\geq \lim_{T \to \infty} \inf_{\mu \in \Omega} J_T^\mu (\bar{z}, \bar{y})$$

$$\geq \lim_{T \to \infty} J_T^\star (\bar{z}, \bar{y})$$

$$= \lim_{T \to \infty} \sum_{i=1}^{\infty} \hat{J}_T^M (f_{i,Q_M}(z), f_{i,Q_M}(y))$$

$$= \sum_{i=1}^{\infty} \lim_{T \to \infty} \hat{J}_T^M (f_{i,Q_M}(z), f_{i,Q_M}(y))$$

$$= \sum_{i=1}^{\infty} \hat{J}_\infty^M (f_{i,Q_M}(z), f_{i,Q_M}(y)),$$

where the exchange of the limit and the summation is warranted by the monotone convergence theorem, since the functions $\hat{J}_T^M$ are monotonically increasing in $T$.

The above inequality provides a lower bound for the optimal cost. Consider now the policy $\mu^\star$ from Proposition 5.4.1. The cost of $\mu^\star$ is

$$J_T^\mu^\star (\bar{z}, \bar{y}) = \sum_{i=1}^{\infty} \hat{J}_\infty^M (f_{i,Q_M}(z), f_{i,Q_M}(y)).$$

For a monotonic and conforming state, this is equal to the lower bound, hence $\mu^\star$ is optimal for the overall problem.

**Theorem 5.4.1.** The set of $(R, nQ)$ policies is optimal under the infinite horizon discounted cost criterion.
Proof. The policy $\mu^*$ is an $(R, nQ)$ policy, by Proposition 5.4.2. For a monotonic state, it attains the lower bound and is optimal. □

5.5 Algorithmic Issues

The decomposition idea is useful for finding the optimal $R^k$ levels as well. In particular, we can compute the optimal action for a $Q_k$-unit subproblem when the units are at $k + 1$ recursively, starting from $k = 1$ and going up.

Definition 5.5.1.

$$q(y, d) = (y - d - 1)^+ + 1$$

Thus, $q(y, d)$ is simply a function that takes the current distance of a customer and a realized demand $d$ and returns the distance of the customer after the demand.

Definition 5.5.2.

$$f^k(y, j) = (y - (n_{k-1} - j + 1)Q_{k-1} - 1)^+ + 1$$

Suppose that we have a batch of size $Q_k$ with a last customer distance $y$. $f^k(y, j)$ simply gives the last customer distance of the $j^{th}$ batch of size $Q_{k-1}$, if this batch of $Q_k$ is split into $n_{k-1}$ batches of size $Q_{k-1}$.

Since a $Q_k$-unit subproblem is a finite state finite control problem, an action is optimal if and only if it attains the minimum in an associated Bellman equation.

Let

$$V_t^k(y, 0) = g^k(y, 0) + E_d \left[ J_{k-1}^k(q(y, d)) \right]$$

and

$$V_t^k(y, 1) = g^k(y, 1) + E_d \left[ \sum_{j=1}^{n_k} J_{t-1}^{k-1}(f^k(q(y, d), j)) \right],$$

where $g^k(y, 0)$ ($g^k(y, 1)$) is the expected one period cost of keeping (shipping) the batch of $Q_k$ units with a last customer distance of $y$ at $k + 1$ (to stage $k$).

Using this relationship between the Bellman equation and optimal actions and using Proposition 5.3.3, we get:

101
Proposition 5.5.1. \( \bar{J}_t^k(y) = \min \{ V_t^k(y, 0), V_t^k(y, 1) \} \). In addition, it is optimal to ship a batch of size \( Q_k \) with a last customer distance of \( y \) from stage \( k + 1 \) to stage \( k \) at time \( t \) if and only if:

\[
V_t^k(y, 1) \leq V_t^k(y, 0)
\]

Solving this recursively for every \( k \), we can find the highest \( y \), for which it is optimal to ship a batch of size \( Q_k \) to stage \( k \), which in turn correspond to the \( R^k \) levels.

5.6 Appendix

In this appendix, we prove Propositions 5.3.4 and 5.4.1. Unlike other chapters, here we do not know how to prove that there exists a decoupled policy with the desired property (a threshold type of result on the customer distance), using only qualitative arguments. Hence, we now need to use the dynamic programming equations for the subproblem to verify this property directly.

Note that \( q(f^k(y, j), d) = f^k(q(y, d), j) \). This can be verified through simple algebra. The idea is, when we split a batch into smaller batches and then observe demand, this has the same result as if we view the demand first and then split the batch into smaller batches. This exchangability is used several times in the rest of the proof.

The one period costs for a \( Q_k \) unit subproblem are as follows:

\[
g^k(y, 1) = \begin{cases} 
    h_k Q_k + c_k Q_k + b(Q_k - y + 1)^+ & \text{if } k \geq 1 \\
    h_1 (1 - \mathbb{I}_{y=1}) & \text{if } k = 0 
\end{cases}
\]

\[
g^k(y, 0) = \begin{cases} 
    h_{k+1} Q_k + b(Q_k - y + 1)^+ & \text{if } k \geq 1 \\
    h_1 + \delta \mathbb{I}_{y=1} & \text{if } k = 0 
\end{cases}
\]

Let

\[
\Delta_t^k(y) = \sum_{j=1}^{n_k} \bar{J}_t^{k-1}(f^k(y, j)) - \bar{J}_t^k(y)
\]

We prove the proposition by first showing that \( \Delta_t^k(y) \) is non-decreasing in \( y \), and then given this monotonicity, we show that the proposition follows.
Lemma 5.6.1. If, for all \( d \), \( f(y,d) \) is non-decreasing in \( y \), then \( E_d \{ f(y,d) \} \) is non-decreasing in \( y \) as well.

Proof. By assumption, for any fixed value of \( d \), \( f(y,d) \) is non-decreasing in \( y \). Taking the expectation is simply taking a convex combination of such non-decreasing functions, which preserves the monotonicity. \( \square \)

Lemma 5.6.2. If \( f(y) \) and \( h(y) \) are both non-decreasing in \( y \), then \( k_1(y) = \min \{ f(y), h(y) \} \) and \( k_2(y) = \max \{ f(y), h(y) \} \) are non-decreasing in \( y \) as well.

Proof. First let us focus on \( k_1(y) \). Note that if over a region, one of the functions is consistently the minimum, then \( k_1(y) \) is clearly non-decreasing in this region since it is simply equal to one of the original functions. We only need to be concerned with transition regions, i.e. the regions where the minimum function changes.

Without loss of generality, suppose that \( y' \) is such a point, i.e. We then have \( f(y') \leq h(y') \) and \( h(y'+1) > h(y') \). Then, note that \( k_1(y'+1) = h(y'+1) \geq f(y') = k_1(y') \), so the property that \( k_1(y) \) is non-decreasing in \( y \) does not change when such a transition occurs either. This completes the proof of the first part of the lemma. A symmetric argument proves the same result for the maximization case. \( \square \)

Lemma 5.6.3. \( \Delta^k_t(y) \) is non-decreasing in \( y \) for all \( t \) and all \( k \geq 1 \).

Proof. We prove the result by induction. First, notice that the property is trivially true for \( t = 0 \) since all the values are 0. This is our base case. Now, assume that the property is true for \( t-1 \) and we next show that this implies that it is also true for \( t \).

\[
\Delta^k_t(y) = \sum_{j=1}^{n_y} \Pi^{k-1}_t(f^k(y,j)) - \Pi^k_t(y) \\
= \sum_{j=1}^{n_y} \min_{u_j} V^k_t(f^k(y,j), u_j) - \min_u V^k_t(y, u) \\
= \min_{u_1, \ldots, u_{n_y}} \left\{ \max_u \left\{ \sum_{j=1}^{n_y} V^{k-1}_t(f^k(y,j), u_j) - V^k_t(y, u) \right\} \right\}
\]
Now, if we can show that \( \sum_{j=1}^{n_k} V_t^{k-1}(f^k(y, j), u_j) - V_t^k(y, u) \) is non-decreasing in \( y \) for every possible choice of the values of \( (u, u_1, u_2, \ldots, u_{n_k}) \), then by Lemma 5.6.2, the result follows.

We have a number of choices about which \( u_j \)'s are going to be equal to 0 and which \( u_j \)'s are going to be equal to 1. Let \( S = \{ j | u_j = 0 \} \) and \( S' = \{ j | u_j = 1 \} \). In addition, there are two cases for \( u \), i.e. \( u = 1 \) and \( u = 0 \).

Case 1 : \( u = 1 \):

\[
\sum_{j=1}^{n_k} V_t^{k-1}(f^k(y, j), u_j) - V_t^k(y, u) \\
= \sum_{j \in S} V_t^{k-1}(f^k(y, j), 0) + \sum_{j \in S'} V_t^{k-1}(f^k(y, j), 1) - V_t^k(y, 1) \\
= \sum_{j \in S} [g^{k-1}(f^k(y, j), 0) + E_d \left[ \bar{j}^{k-1}(f^k(q(y, d), j)) \right]] \\
+ \sum_{j \in S'} \left[ g^{k-1}(f^k(y, j), 1) + E_d \left[ \sum_{l=1}^{n_k-1} \bar{j}^{k-2}(f^k(q(y, d), j), l) \right] \right] \\
- g^k(y, 1) - E_d \left[ \sum_{r=1}^{n_k} \bar{j}^{k-1}(f^k(q(y, d), r)) \right] \\
= \sum_{j \in S} g^{k-1}(f^k(y, j), 0) + \sum_{j \in S'} g^{k-1}(f^k(y, j), 1) - g^k(y, 1) \\
+ E_d \left[ \sum_{j \in S'} [\Delta_t^{k-1}(f^k(q(y, d), j))] \right]
\]

The last term in the above equation is non-decreasing in \( y \), using the induction hypothesis and Lemma 5.6.1. Hence, to complete the first case, it remains to show that show that

\( \sum_{j \in S} g^{k-1}(f^k(y, j), 0) + \sum_{j \in S'} g^{k-1}(f^k(y, j), 1) - g^k(y, 1) \) is non-decreasing in \( y \). Now, if
\( k \geq 2 \), we have:

\[
\sum_{j \in S} g^{k-1}(f^k(y, j), 0) + \sum_{j \in S'} g^{k-1}(f^k(y, j), 1) - g^{k}(y, 1)
\]

\[
= \sum_{j \in S} h_k Q_{k-1} + b(Q_{k-1} - f^k(y, j) + 1)^+
\]

\[
+ \sum_{j \in S'} h_{k-1} Q_{k-1} + c_{k-1} Q_{k-1} + b(Q_{k-1} - f^k(y, j) + 1)^+
\]

\[
- h_k Q_k - c_k Q_k - b(Q_k - y + 1)^+
\]

\[
= \sum_{j \in S} h_k Q_{k-1} + \sum_{j \in S'} h_{k-1} Q_{k-1} + c_{k-1} Q_{k-1}
\]

The above term does not depend on \( y \), so it is non-decreasing in \( y \). (This is the result of simple algebra, but the idea is as follows: The only term that is not obvious is the backlog cost term, i.e. the term that is multiplied by \( b \). However, this is the difference in backlog cost \( Q_k \) units at stage \( k \) versus \( Q_k \) units at stage \( k + 1 \) given that the distance of the last customer corresponding to the units is \( y \). However, since \( k \geq 1 \), none of these units can be given to customers in this period anyway, so the backlog costs are the same and cancel each other).

For \( k = 1 \), we have:

\[
\sum_{j \in S} g^0(f^0(y, j), 0) + \sum_{j \in S'} g^0(f^1(y, j), 1) - g^1(y, 1)
\]

\[
= \sum_{j \in S} h_1 + b \mathbb{1}_{f^1(y, j) = 1} + \sum_{j \in S'} h_1 (1 - \mathbb{1}_{f^1(y, j) = 1})
\]

\[
- h_1 Q_1 - c_1 Q_1 - b(Q_1 - y + 1)^+
\]

This is also non-decreasing in \( y \). (This also follows through simple algebra, but the idea is as follows: The only term that is not obvious is the backlog cost term, i.e. the term that is multiplied by \( b \). However, this is the difference in backlog cost of \( Q_1 \) units at stage 1 versus \( Q_1 \) units at stage 2, given that we choose to release some of the units at stage 1 to be given to customers, and that the distance of the last customer corresponding to the units is \( y \). If \( y \) is small enough, this means that some customers are waiting and hence being at 1 has
a smaller backlog cost, so the term is negative. As \( y \) increases, less and less customers are waiting, so the additional benefit of being at stage 1 versus stage 2 decreases, and eventually when no customers are waiting, the difference becomes 0). This concludes Case 1.

Case 2 : \( u = 0 \):

\[
\sum_{j=1}^{n_k} V_t^{k-1}(f^k(y, j), u_j) - V_t^k(y, u)
\]
\[
= \sum_{j \in S} V_t^{k-1}(f^k(y, j), 0) + \sum_{j \in S'} V_t^{k-1}(f^k(y, j), 1) - V_t^k(y, 0)
\]
\[
= \sum_{j \in S} \left[ g_t^{k-1}(f^k(y, j), 0) + E_d \left[ J_{t-1}^{k-1}(f^k(q(y, d), j)) \right] \right] 
\]
\[
+ \sum_{j \in S'} \left[ g_t^{k-1}(f^k(y, j), 1) + E_d \left[ \sum_{l=1}^{n_{k-1}} J_{t-1}^{k-2}(f^k(q(y, d), j), l) \right] \right] 
\]
\[
- g_t^k(y, 0) - E_d \left[ J_{t-1}^k(q(y, d)) \right]
\]
\[
= \sum_{j \in S} \left[ g_t^{k-1}(f^k(y, j), 0) + E_d \left[ J_{t-1}^{k-1}(f^k(q(y, d), j)) \right] \right] 
\]
\[
+ \sum_{j \in S'} \left[ g_t^{k-1}(f^k(y, j), 1) + E_d \left[ \sum_{l=1}^{n_{k-1}} J_{t-1}^{k-2}(f^k(q(y, d), j), l) \right] \right] 
\]
\[
- g_t^k(y, 0) + E_d \left[ \Delta_{t-1}^k(q(y, d)) - \sum_{r=1}^{n_k} J_{t-1}^{k-1}(f^k(q(y, d), r) \right]
\]
\[
= \sum_{j \in S} g_t^{k-1}(f^k(y, j), 0) + \sum_{j \in S'} g_t^{k-1}(f^k(y, j), 1) - g_t^k(y, 0)
\]
\[
+ E_d \left[ \sum_{j \in S'} \left[ \Delta_{t-1}^{k-1}(f^k(q(y, d), j)) \right] + \Delta_{t-1}^k(q(y, d)) \right]
\]

The last term in the above equation is non-decreasing in \( y \), using the induction hypothesis and Lemma 5.6.1. Hence, to complete the second case, it remains to show that show that the term \( \sum_{j \in S} g_t^{k-1}(f^k(y, j), 0) + \sum_{j \in S'} g_t^{k-1}(f^k(y, j), 1) - g_t^k(y, 0) \) is non-decreasing in \( y \). However, since \( k \geq 1 \), the part of \( g_t^k(y, 0) \) that does depend on \( y \) is the same as the corresponding part for \( g_t^k(y, 1) \) and hence just like in Case 1, this term is non-decreasing in \( y \).

\( \square \)

Let \( \hat{U}_t^k(y) \) be the set of optimal decisions for a batch of size \( Q_k \) at stage \( k + 1 \) with a last customer distance of \( y \) at time \( t \) for a \( Q_k \)-unit subproblem. Hence, \( 1 \in \hat{U}_t^k(y) \) means that it
is optimal to ship this batch to stage \( k \). We have:

\[
1 \in \tilde{\mathcal{U}}^k_t(y) \\
\iff V^k_t(y, 1) \leq V^k_t(y, 0) \\
\iff g^k(y, 1) + E_d \left[ \sum_{j=1}^{n_k} J^k_{t-1}(f^k(q(y, d), j)) \right] \leq g^k(y, 0) + E_d \left[ J^k_{t-1}(q(y, d)) \right] \\
\iff E_d \left[ \Delta^k_{t-1}(q(y, d)) \right] \leq g^k(y, 0) - g^k(y, 1) \tag{5.1}
\]

Fix some \( k \geq 1 \). Suppose that it is optimal to ship a batch with a last customer distance of \( y \), i.e. \( 1 \in \tilde{\mathcal{U}}^k_t(y) \). We need to verify that it is optimal to ship the batch even of the customer distance is \( y' < y \). Using equation (5.1), this is optimal if and only if

\[
E_d \left[ \Delta^k_{t-1}(q(y', d)) \right] \leq g^k(y', 0) - g^k(y', 1)
\]

Now,

\[
E_d \left[ \Delta^k_{t-1}(q(y', d)) \right] \\
\leq E_d \left[ \Delta^k_{t-1}(q(y', d)) \right] \\
\leq g^k(y, 0) - g^k(y, 1) \\
= g^k(y', 0) - g^k(y', 1)
\]

where the first inequality is due to Lemma 5.6.3, the second because it is optimal to ship the batch with a customer distance of \( y \), and the final equality follows simply from the definition of \( g^k \). (Note that \( k \geq 1 \)). Hence, using equation (5.1), we have:

\[
1 \in \tilde{\mathcal{U}}^k_t(y) \Rightarrow 1 \in \tilde{\mathcal{U}}^k_t(y'), \quad \forall y' < y \tag{5.2}
\]

Now, we are finally in a position to prove Proposition 5.3.4. We are going to choose a particular \( \tilde{\mu}_t^k(y) \) that prescribes the actions of an optimal \( Q_k \)-unit subproblem policy when the \( Q_k \) units are at stage \( k + 1 \) with a last customer distance \( y \) at time \( t \).

Fix some \( k \) and \( t \). There are three cases:

(i) There are infinitely many \( y \) for which \( 1 \in \tilde{\mathcal{U}}^k_t(y) \). Then, by equation 5.2, \( 1 \in \tilde{\mathcal{U}}^k_t(y) \) for all \( y \) and one can set \( \tilde{\mu}_t^k(y) = 1 \) for all \( y \). This corresponds to \( \tilde{y}_t^k = \infty \).
(ii) There is no $y$ for which $1 \in \hat{U}^k_t(y)$. Then, one can set $\bar{\mu}^k_t(y) = 0$ for all $y$. This corresponds to $\bar{y}^k_t = -\infty$.

(iii) There is a highest $y^*$ for which $1 \in \hat{U}^k_t(y)$. Then, using equation 5.2, one can set $\bar{\mu}^k_t(y) = 1$ if and only if $y \leq \bar{y}^k_t$.

The proof of Proposition 5.3.4 is now complete.

At this point, the proof for Proposition 5.4.1 is quite simple. Since we have shown through induction that $\Delta^k_t(y) = \sum_{j=1}^{n_k} J^k_{t-1} f^k(y, j) - J^k_t(y)$ is non-decreasing in $y$ for every finite $t$ and all $k \geq 1$ and since $\lim_{T \to \infty} J^k_T(y) = J^k_\infty(y)$ by Proposition 5.4.1, the induction also proves that $\Delta^k_\infty(y) = \sum_{j=1}^{n_k} J^k_\infty f^k(y, j) - J^k_\infty(y)$ is non-decreasing in $y$ for every $k \geq 1$. The rest of the proof is symmetric to the finite horizon case.
Chapter 6

Assembly Systems

6.1 Introduction

This chapter analyzes an assembly system. In such a system, a number of parts are acquired from outside suppliers and then assembled into subassemblies and eventually into a final product, which is then used to satisfy customer demand. We assume no capacities in ordering or assembly. The demand and the leadtimes are stochastic. The goal is to find an optimal ordering policy that minimizes the sum of holding, backorder and ordering costs.

This model is a generalization of the serial model in Chapter 3 to the assembly system case, except for the fact that here we assume stationary demand and leadtime statistics instead of the Markov modulated case. However, results here can easily be extended to include that case as well.

Schmidt & Nahmias [30] investigated a finite horizon model of two components assembled into a single final product, under deterministic leadtimes. They characterized the structure of optimal policies as belonging to a relatively complicated class, under general assumptions on initial inventory levels. Rosling [29] has shown that if the initial state of the system satisfies a certain condition, then the assembly system with deterministic leadtimes has an equivalent serial representation. Hence, results for serial systems can be used. Chen & Zheng [14] gave an alternative proof of this result. All these problems assumed linear ordering costs. De
Bodt & Graves [18] describe an approximate method to analyze an assembly system with fixed ordering costs, considering policies that use fixed lot sizes.

The main idea behind the fact that there is a serial representation of an assembly system under some conditions is that one would like to coordinate the orders of items that are to be assembled together. In particular, one would like to order an item with a short leadtime only if there will be a corresponding item with the long leadtime available by the time the former is delivered, so that the assembly can take place. Otherwise, unnecessary holding costs will be incurred. However, this argument does not hold in the case of stochastic leadtimes. We are not aware of any papers that either characterize optimal policies or provide exact computational methods for assembly systems with stochastic leadtimes.

In this chapter, we show that our assembly system problem can be decomposed into a series of subproblems each of which consists of a single kit of parts and a single customer. A kit of parts is a set of components, one from each outside supplier. Given this decomposition result, we can compute optimal policies and the corresponding costs by solving a single kit problem, rather than the overall problem. However, the optimal policies in this case are more complicated than echelon base stock policies, since the policies for a subproblem depend on the state of the whole kit. Still, the decomposition result, along with some results about the structure of the subproblems, gives us some information about the structure of optimal policies, even though we are unable to characterize them completely.

The rest of this chapter has four sections. Section 6.2 gives the problem formulation. Section 6.3 gives the analysis for the finite horizon problem. Section 6.4 discusses the algorithmic implications of the results and, finally, Section 6.5 discusses some extensions to the model.

6.2 Problem Formulation

We consider an uncapacitated assembly system with stochastic leadtimes. n parts are acquired from outside suppliers (which are assumed to have infinite stock), then possibly
assembled into subassemblies and finally into a final product. Let $M$ be the total number of different items, including parts acquired from outside suppliers, subassemblies and the final product. (For example, for the system in Figure 6-1, $M = 5$ and $n = 3$). Such an assembly system has a tree structure. Let $s(k)$ be the unique immediate successor of item $k$.

We assume that the units are chosen so that whenever a number of items are assembled together, exactly one unit from each item type is used. In addition, we assume that there are no capacity limits for the number of units that can be delivered or the number of assemblies that can be performed in a given period.

Without loss of generality, we assume that the assembly operation takes no time, once the required items are in hand. (Note that a non-zero time requirement can be modeled via the current model by introducing an additional item.) We view the network like a physical network, hence we use the following terminology: An item $k$ is shipped from a location $k$ to a location $s(k)$. Once all the required items are there, they are assembled into item $s(k)$. There is a leadtime for an item $k$ to be delivered from location $k$ to location $s(k)$. We assume that this time is stochastic, and has the structure of the stochastic leadtime model that was used in Chapter 3. In particular, we assume that the leadtime between location $k$ and $s(k)$ is upper bounded by some integer $l_k$. We assume that the probability that an outstanding
order arrives during the current period depends only on the amount of time since the order was placed and the item type \( k \) and, given these, it is conditionally independent of the history of the process until now. Finally, we assume that orders cannot overtake each other: an order cannot arrive at its destination before an earlier order does.

The leadtime model described above includes the obvious special case of deterministic leadtimes. As indicated in Chapter 3, it also includes a stochastic model which extends the model of Kaplan [26]. We now remind the reader about this model: At each time period \( t \), there is a random variable \( \rho_t^k \) that determines which outstanding orders of item \( k \) will be delivered to location \( s(k) \). More precisely, an outstanding order will be delivered to location \( s(k) \) if and only if it was placed \( \rho_t^k \) or more time units ago. Note that such a mechanism ensures that orders cannot overtake each other. Let \( \rho_t = (\rho_t^1, \rho_t^2, \ldots, \rho_t^M) \) be the vector of leadtime random variables associated with the various items. We assume that the statistics of \( \rho_t \) are independent and identically distributed over time. Notice that such a model allows for dependencies between the leadtime random variables corresponding to the same period but different items.

We assume that the demand is independent and identically distributed over time as well. Both demand and the leadtime models can easily be extended to the Markov modulated case as in Chapter 3, but we choose this particular model to simplify the exposition.

The cost structure that we use consists of linear holding, ordering, and backorder costs. In more detail, we assume:

(a) For each item \( k \), there is an inventory holding cost rate \( h_k \) that gets charged at each time period to each unit. We assume that the holding cost rate \( h_k \) for an item that is at an external supplier is zero. There are \( n \) items that are acquired from external suppliers. No holding cost is incurred for these items until they are ordered and are in transit. For concreteness, we also assume that after a unit is ordered and during its leadtime, the holding cost rate charged for this unit is the rate corresponding to the destination location.

(b) For each item \( k \), there is a cost \( c_k \) for initiating the shipment from location \( k \) to location
s(k).

(c) There is a backorder cost rate b which is charged at each time step for each unit of backlogged demand.

We assume that the holding cost parameters for locations other than the outside suppliers and the backorder cost parameter are positive and that the shipping cost is non-negative.

6.3 Finite Horizon Analysis

In this section, the finite horizon version of the problem is analyzed. We show that the problem can be decomposed into a series of subproblems, each of which consists of a single kit of parts and a single customer. A kit of parts is made up of n parts, each one from a different outside supplier. For example, if our final product is a computer system and we buy a computer, a monitor and a printer from different sources before we give the system to the customer, then a kit consists of these three components. Once we show that the problem can be decomposed, this gives us an efficient way to analyze this system. In addition, even though we do not characterize the structure of optimal policies completely, we give some results that shed light on the behaviour of an optimal policy.

Recall that in a serial system, the idea was that units and customers can be matched so that we know which unit is going to be given to which customer, and consequently we showed that the problem decouples into single unit single customer subproblems. Here, the main idea is that we can match a single kit with a customer and be sure that that customer will receive a final item that is assembled by putting together the parts in that kit. We can do this, because we have the no order crossing assumption on the stochastic leadtimes.

Before we proceed, we define two quantities, the location of a part (for the parts that are acquired from outside suppliers) and the location of a kit.

Definition 6.3.1. The location of a part: Note that there is a unique path from the outside supplier until the final item for each part acquired from an outside supplier. Now, concentrate
on a particular part. Let each stocking point on the path for this part be a location. In addition, insert $l_k - 1$ artificial locations between each stocking point $k$ and the next one $s(k)$, to model the delivery leadtime. Finally, the location 0 correspond to parts that were already assembled into a final item and given to a customer. Index all the other locations in increasing order of the maximum leadtime from that location until the end of the path.

This definition is very similar to the definition of the location of a unit that was given in Chapter 3. There, we had a serial system so each unit went through the same path. Here, we have an assembly system and each part goes through a unique path.

**Definition 6.3.2.** *The location of a kit:* The location of a kit is an $n$-dimensional vector whose components correspond to the locations of the $n$ parts that make up the kit.

Now, we form kits and label them according to their locations at time 0, so that we can treat them as individual objects. To do this, we observe the system at time 0. Kit 1 consists of the $n$ parts (one of each type) that have the smallest locations, breaking ties arbitrarily. Kit 2 consists of the $n$ parts that have the second smallest locations, etc. We form a countably infinite number of kits and label them in this way.

Just like in Chapter 3, we can now think of kit $i$ and customer $i$ as forming a pair.

We now define the state space for our problem. For each kit-customer pair we have a vector $(z_t^i, y_t^i)$. Here, $z_t^i = (z_{t_1}^i, z_{t_2}^i, \ldots, z_{t_n}^i)$ is the location of kit $i$ at time $t$, and $z_{t_k}^i$ is the location of part $k$ that is a component of kit $i$ at time $t$. $y_t^i$ is the distance of customer $i$ at time $t$. The state of the system consists of a countably infinite number of such vectors, one for each kit-customer pair:

$$x_t = \{(z_t^1, y_t^1), (z_t^2, y_t^2), \ldots\}$$

The control vector is an infinite sequence $u_t = (u_t^1, u_t^2, \ldots)$ where $u_t^i = (u_{t_1}^i, u_{t_2}^i, \ldots, u_{t_n}^i)$ and $u_{t_k}^i$ is 1 if part $k$ of kit $i$ is released from its current location and is 0 if it is held at its current location. Of course, a part is affected by this control only when it is at a non-artificial stage, i.e., if it is not in transit or given to the customer. In addition, note that there are certain restrictions on the shipment of parts. For example, if two parts need to be assembled
into a subassembly before they can be shipped further, they can only be released at the same
time (which actually corresponds to releasing the subassembly).

We define \( z^i_t \leq z^j_t \) if and only if \( z^{i,k}_t \leq z^{j,k}_t \) for all \( k \). Then, we modify the definition of a
monotonic state:

**Definition 6.3.3.** A state is monotonic if and only if \( z^i_t \leq z^j_t \) for every \( i < j \).

We use the familiar definition for monotonic policies, i.e., a monotonic policy is one that
keeps the state monotonic. Similarly, committed policies refer to policies that release a final
item (i.e., the parts that make it up) from location 1 only if the corresponding customer is
there. Decoupled policies refer to policies that make a decision about a particular kit by
only looking at that kit and the corresponding customer:

**Definition 6.3.4.** We call a policy decoupled, if it can be represented in terms of mappings
\( \mu_t \), so that the decision about a particular kit, \( \hat{u}^i_t \) depends only on the location of the kit and
the distance of the corresponding customer through

\[
\hat{u}^i_t = \mu_t(z^i_t, y^i_t), \quad \forall \ i, \ t.
\]

**Proposition 6.3.1.** The set of monotonic and committed policies is optimal.

*Proof.* The steps in the proof are essentially the same as the proofs of Propositions 3.3.1,
3.3.2, and we only provide a sketch. Monotonic policies are optimal, because whenever we
are at a monotonic state and we need to release some parts from a location, we can choose
to release the ones with the lowest index. Given that monotonic policies are optimal, we can
match kits with customers, and committed policies are optimal.

Since monotonic and committed policies are optimal, we can relax the restriction to
monotonic policies and, therefore committed policies are optimal.

**Proposition 6.3.2.** The set of committed and decoupled policies is optimal.

*Proof.* Similar to the proof of Proposition 3.3.4, the restriction to committed policies can be
represented as a change in the system dynamics as opposed to a restriction on the policy
space. With this modification, the system becomes decomposable into single kit single customer problems.

We have thus shown that the overall problem can be decomposed into much simpler subproblems, each of which consists of a single kit and a single customer. This is an important step towards the ultimate goal of completely characterizing the structure of optimal policies. In addition, it gives us a way to compute optimal policies relatively efficiently.

Even though we have not characterized the structure of optimal policies completely, we can still partially describe the behaviour of optimal policies. The following definition and lemma are geared towards this goal.

**Definition 6.3.5.** For any \( \tau, \bar{z}, y \) and \( y' \), let \( \bar{U}^*_\tau(\bar{z}, y) \) be the set of all decisions that are optimal if a subproblem is found at state \((\bar{z}, y), \tau\) time steps before the end of the horizon.

**Lemma 6.3.1.** For every \((\tau, k, \bar{z}, y)\), if the \(k\)th component of every \( u \in \bar{U}^*_\tau(\bar{z}, y) \) is equal to 1, then for every \( \bar{z}' \leq \bar{z} \) and \( y' < y \), there exists some \( u' \in \bar{U}^*_\tau(\bar{z}', y') \), whose \(k\)th component is 1.

**Proof.** Suppose that there exist some \((\tau, k, \bar{z}, y)\) and \((\tau, k, \bar{z}', y')\), with \( \bar{z}' \leq \bar{z} \) and \( y' < y \), such that the \(k\)th component of every \( u \in \bar{U}^*_\tau(\bar{z}, y) \) is 1 and the \(k\)th component of every \( u' \in \bar{U}^*_\tau(\bar{z}', y') \) is 0. Let \( t = T - \tau \). Consider a monotonic state for the overall problem such that the state of the \(i\)th kit-customer pair is \((\bar{z}, y)\) and the state of the \(j\)th kit-customer pair is \((\bar{z}', y')\). Note that since \( y' < y \), we must have \( j < i \). Then according to Lemma 2.0.1(3), the decision under any optimal policy must satisfy \( u^{i,k}_t = 1 \) and \( u^{j,k}_t = 0 \). This means that part \( k \) of the higher indexed kit \( i \) will move ahead of part \( k \) of kit \( j \), and the new state will not be monotonic. Therefore, no monotonic policy can be optimal, which contradicts Proposition 6.3.1.

Note that the above lemma is quite similar to Lemma 3.3.1, which in the serial system context allowed us to prove that the set of monotonic, committed and decoupled policy is optimal. However, the above Lemma is not enough for the analogous result to hold in
this setting. This is due to the interaction among the different parts. We conjecture that there is an analogous lemma that delas with a monotonicity property for the whole control vector, which in turn leads to the optimality of monotonic, committed and decoupled policies. However, even with such a lemma, the structure of optimal policies will not be as simple as echelon base stock policies, since the optimal control will depend on the location of the whole kit.

6.4 Algorithmic Issues

The result that the overall problem can be decomposed into a series of single kit subproblems provides us with a way to compute optimal policies relatively efficiently. In particular, one needs to solve a single dynamic program with a single kit and a single customer to compute the optimal policy and the corresponding cost.

Before we give the complexity result for this dynamic program, we need the following result that limits the customer distances that we need to consider:

**Lemma 6.4.1.** There exists an integer $Y_{\text{max}}$ such that any optimal policy for a single kit subproblem with horizon length $T$ will not release any of the parts of the kit from the outside suppliers if the distance of the customer is larger than $Y_{\text{max}}$. Formally, there exists some $Y_{\text{max}}$ such that for every $y > Y_{\text{max}}$ and every $T$, we have $\bar{U}^*_T(z, y) = \{(0, 0, \ldots, 0)\}$, if $z$ corresponds to the location of the kit where all the parts are at the outside suppliers.

**Proof.** The proof is along the lines of the proof of Lemma 4.3.1 and we only provide a sketch. For a large enough horizon length, releasing a part of the kit from the outside supplier before the corresponding customer is close enough incurs unnecessarily high cost, since once the part is released, it incurs a positive holding cost at least until the customer shows up, which can be a long time when $y$ is high. For horizons that are not that large, the probability that the customer will arrive before the end of the horizon decreases as $y$ goes to infinity. Hence, the cost of a policy that holds all the parts goes to 0 as $y$ gets larger, whereas the cost of a
policy that releases a part is positive. Hence, it is possible to find a large enough $y$ beyond which it is optimal not to release any of the parts of the kit.

Let $N$ be the number of locations (including the artificial ones corresponding to parts in transit) along the path with the maximum number of locations. Hence, $N$ is the maximum total leadtime to assemble a final item from raw materials if we try to put one together as quickly as possible.

**Proposition 6.4.1.** The complexity of the standard backward recursion dynamic programming algorithm for the finite horizon version of the single kit subproblem, with a planning horizon of $T$, is $O(N^n \cdot Y_{\text{max}}^2 \cdot T)$.

**Proof.** The cardinality of the state space is $O(N^n \cdot Y_{\text{max}})$. For each state, the number of possible controls is $2^n$, and given a state and a control, the number of possible next states is bounded above by $2 \cdot Y_{\text{max}}$. Hence, the complexity of the standard backward recursion algorithm for the finite horizon version of the dynamic program with a planning horizon of $T$ is $O(N^n \cdot Y_{\text{max}}^2 \cdot T)$.

We refer to this algorithm as relatively efficient, because it has a complexity that is exponential in $n$. Hence, solving the problem for large values of $n$ may not be possible. Still, it is much simpler than trying to solve the overall problem via dynamic programming, which is the only other exact approach for an assembly system with stochastic leadtimes that we are aware of.

### 6.5 Extensions

In the preceding sections, we provided the analysis for the finite horizon version of the problem. The results can be extended to infinite horizon as in Chapter 3. It is also possible to generalize the model to the case where the demands and the leadtimes are Markov modulated.

One interesting extension is the model where we have an assembly system along with batch ordering restrictions. It is possible to generalize the results in this chapter to this case.
In particular, one can show that the problem decomposes into a series of subproblems, each of which has a number of kits and the customers corresponding to those kits.

A similar problem was analyzed in Section 4 of Chen [11], where it was shown that if the lot sizes satisfy a certain condition, then the problem with deterministic leadtimes has an equivalent serial representation (similar to Rosling [29]). The required condition is as follows: Index the items in increasing order of their total leadtimes and let $Q_k$ be the lot size for item $k$. Then, the condition is that $Q_{k+1} = n_k Q_k$ for all $k$, where $n_k$ is a positive integer. For example, consider a system where two parts are assembled into a final product. The lot size for part 1 is 2 and the lot size for part 2 is 4. This system satisfies the condition only if the leadtime for part 2 is larger than the leadtime for part 1. However, for our method to work, no such restriction is required. Either leadtime can be larger and the problem decomposes into a series of subproblems each of which has 4 kits and 4 customers. Even when the lot sizes are not multiples for each other, the problem can be decomposed into a series of subproblems. The number of kits in each subproblem however, will then be the smallest common multiple of all the lot sizes.
Chapter 7

Conclusions

In this dissertation, we introduced a new perspective on multi-echelon inventory control, based on the idea of decomposing the overall problem into much simpler subproblems. This decomposition allowed us to characterize the structure of optimal policies for some important multi-echelon systems, and provides a method to compute optimal policies efficiently.

There are still many important problems in the multi-echelon inventory control realm that deserve further investigation. For example, serial systems with general cost structures, serial systems with capacities, and serial systems with lost sales are fundamental problems, whose optimal solutions are still unknown. For these problems, the decomposition does not go through, but the approach still provides a new perspective into the details of these notoriously difficult problems. For example, for serial systems with capacities and for serial systems with general non-linear cost structures, it is possible to show that monotonic and committed policies are optimal. However, even with the committed restriction, the problems are still coupled, in one case because of the capacity restriction and in the other because the marginal cost in one subsystem depends on other subsystems. For systems with lost sales, it is possible to show that monotonic policies are optimal, but committed policies are not. This is because once a demand is lost, the corresponding unit should be reassigned to another customer. Such observations, combined with others, may help us in identifying the aspects of a certain problem that make it difficult. This in turn, may suggest ideas about
how to overcome these difficulties.

Another fundamental system that was not studied in this dissertation is a distribution system. Consider a two echelon system, with a single warehouse and multiple retailers. For such a system, monotonic policies are optimal. However, one cannot match units with particular customers, because the retailer that the unit will be shipped to is not determined in advance. Hence, an exact decomposition is not possible. But, note that once the allocation of a certain unit to a certain retail site has been made, then the unit can be matched with a particular customer and the control of that unit can be made independently of other units, using the single unit approach. This suggests the following heuristic: Use an iterative method of allocating units to retail sites and then given the allocation, making shipment decisions followed by another allocation decision and another shipment decision and so on. We can solve the allocation problem by a simple dynamic program, where the costs of the dynamic program are determined using the value function of the single unit subproblem for the serial system.

The analysis here focused on systems with periodic review, but we believe that all the results can be extended to systems with continuous review with Poisson demand (or Poisson where the rate is Markov modulated) scenarios. Of course, the details would change slightly.
Bibliography


