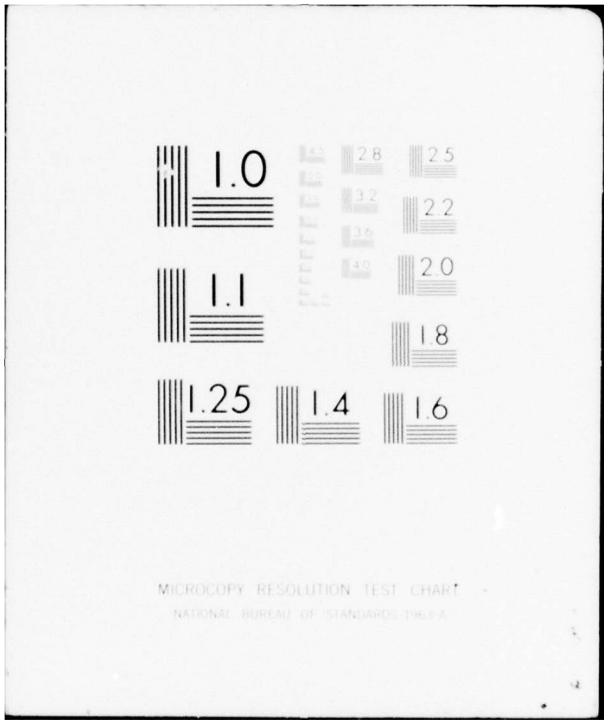
AD-A042 559 MASSACHUSETTS INST OF TECH CAMBRIDGE OPERATIONS RESE--ETC F/G 5/3
ISSUES OF HIERARCHICAL PLANNING IN MULTI-STAGE PRODUCTION SYSTE--ETC(U)
JUL 77 D CANDEA N00014-75-C-0556
UNCLASSIFIED TR-134 NL

1 OF 3
AD A042559




1.0
1.1
1.25
1.4
1.6
2.8
2.5
3.2
2.2
3.6
2.0
4.0
1.8
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

AD A042559

ISSUES OF HIERARCHICAL PLANNING IN MULTI-STAGE PRODUCTION SYSTEMS

by

DAN CANDEA

Technical Report No. 134

OPERATIONS RESEARCH CENTER

DDC
RECEIVED
AUG 9 1977
C

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited



MASSACHUSETTS INSTITUTE
OF
TECHNOLOGY

July 1977

AD NO.
DDC FILE COPY

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Date Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report No. 134

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE (and Subtitle): ISSUES OF HIERARCHICAL PLANNING IN MULTI-STAGE PRODUCTION SYSTEMS.

5. TYPE OF REPORT & PERIOD COVERED: Technical Report, July 1977

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s): Dan Candea

8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-0556

9. PERFORMING ORGANIZATION NAME AND ADDRESS
M.I.T. Operations Research Center
77 Massachusetts Avenue
Cambridge, MA 02139

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 347-027

11. CONTROLLING OFFICE NAME AND ADDRESS
O.R. Branch, ONR Navy Dept.
800 North Quincy Street
Arlington, VA 22217

12. REPORT DATE: July 1977

13. NUMBER OF PAGES: 266 pages

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): TR-134

15. SECURITY CLASS. (of this report): 264p.

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Releasable without limitation on dissemination.

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

DDC
AUG 9 1977
RECEIVED
C

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)
Multi-Stage Production Systems
Hierarchical Production Planning
Dynamic Programming

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

See page 4.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

270 720

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Date Entered)

ISSUES OF HIERARCHICAL PLANNING IN MULTI-STAGE PRODUCTION SYSTEMS

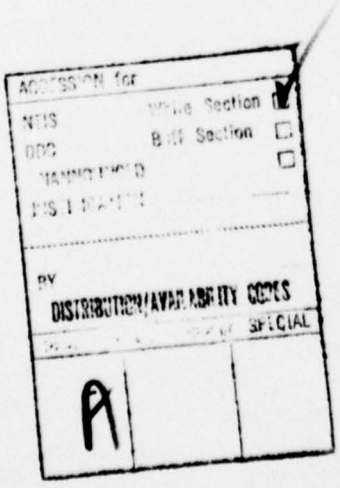
by

DAN CANDEA

Technical Report No. 134

Work Performed Under
Contract N00014-75-C-0556, Office of Naval Research
Multilevel Logistics Organization Models
NR 347-027 M.I.T. OSP 82491

Operations Research Center
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139



July 1977

FOREWORD

The Operations Research Center at the Massachusetts Institute of Technology is an interdepartmental activity devoted to graduate education and research in the field of operations research. The work of the Center is supported, in part, by government contracts and industrial grants-in-aid. The work reported herein was supported (in part) by the Office of Naval Research under Contract N00014-75-C-0556.

Jeremy F. Shapiro
Richard C. Larson
Co-Directors

ABSTRACT

The multi-stage production system is viewed as a production process in which component parts have to be obtained by manufacturing or by purchasing, then assembled into subassemblies, assemblies, and finally into the finished good.

The problem and its characteristics are described, and the principal difficulties associated with this kind of problems are identified. Both computational and managerial aspects are discussed and used to motivate the adoption of the hierarchical approach to production planning. According to this approach, the overall problem is partitioned into two levels: the aggregate level and the detailed (or disaggregation) level.

At the aggregate level a new formulation for the aggregate planning model is given, in order to bring computational feasibility to situations in which older formulations went beyond the capabilities of current linear programming codes. The resulting aggregate model is a large scale system that lends itself to solution by column generation. A dynamic programming algorithm for the generation of columns is developed.

Next, at the disaggregation level, the problem of computing optimal lot sizes in multi-stage systems is addressed and solved. Since exact solution procedures are found to be either very expensive or computationally infeasible, a heuristic approach is adopted and results are reported. For more complex situations, in which parts are common to several end products or where there is independent demand for parts, even the heuristics become infeasible; therefore it is suggested that myopic lot sizing policies be used.

The issue of safety stocks for aggregate planning is treated in another chapter. The impact of the rolling horizon policy upon safety stocks is first examined in the simpler setting of single stage production. Then, the sources of uncertainties in multi-stage systems are identified and ways to build protection against them are developed.

In the concluding chapter a number of research topics, which are worth investigating in ghe author's opinion, are indicated.

TABLE OF CONTENTS

	Page
Abstract	2
Acknowledgements	4
Table of contents	5
List of figures	9
List of tables	12
Chapter 1 - INTRODUCTION	13
1.1 General Statement of the Problem	13
1.2 Global vs. Hierarchical Approach	20
1.3 Outline of the Thesis	23
Chapter 2 - A SURVEY OF THE LITERATURE	25
2.1 Typical Issues and Modelling Assumptions	25
2.2 A Survey of Multi-Stage Models	27
2.2.1 Infinite Horizon Models	31
2.2.2 Finite Horizon Models	34
Chapter 3 - A HIERARCHICAL APPROACH	41
3.1 The Aggregate Level	41
3.1.1 Modelling Backorders	43
3.1.2 Issues of Aggregation	45
3.1.2.1 Aggregation of Items	45
3.1.2.2 An Example of Aggregation	51
3.1.2.3 Aggregation of Individual Facilities (Machines and Work Places) into	

	Page
Aggregate Facilities or Stages	66
3.1.3 Handling Initial Inventories of Component Parts in the Aggregate Planning Model	68
3.1.4 A Two-Level Aggregation for Very Large Operations	72
3.2 Issues of Disaggregation	75
3.2.1 A Conceptual Approach to Disaggregation	77
3.2.2 Composition Factors in Family Structures	84
Chapter 4 - THE AGGREGATE PLANNING MODEL	88
4.1 A Simple Example and The Elementary Problem	88
4.2 Elementary vs. Non-Elementary Production Schedules	98
4.3 More Complex F2 Formulations for Uncapacitated Multi-Stage Systems	105
4.3.1 The Case with Initial Inventories	106
4.3.2 The Case with Initial Inventories of Parts Common to Two or More End Products	120
4.3.3 The Case with Non-Zero Lead Times and Initial Inventories	125
4.3.4 The Case of Assembly Structures with Diverging Arcs	134
4.3.5 The Case with Independent Demand for Component Parts	137
4.3.6 Further Refinements for the Uncapacitated Problem	137
4.4 F2 Formulations for Capacitated Multi-Stage Systems	145

Page

4.4.1 The Dominant Character of the Elementary Extreme Flow Schedules 149

4.4.2 Column Generation by Dynamic Programming 152

4.4.3 Other Ways of Reducing the Computational Effort 158

4.4.3.1 Multiple Pricing 158

4.4.3.2 Revised F2 Formulations 159

4.4.4 Concluding Comments 164

Chapter 5 - ECONOMIC LOT SIZE DETERMINATION IN MULTI-STAGE SYSTEMS 166

5.1 The Case with Independent Demand for the End Product Only 167

5.2 The Case with Independent Demand for the End Product and Component Parts 198

5.3 Solving the Lot Sizing Problem 203

Chapter 6 - ISSUES OF SAFETY STOCKS CALCULATIONS 213

6.1 Forecast Fluctuations and Safety Stocks for Aggregate Planning with a Rolling Horizon in Single-Stage Systems 220

6.2 Safety Stocks in Multi-Stage Systems 225

6.2.1 Safety Stocks in a Multi-Stage One Time Production - The Case of Uncertainties in Requirements 228

6.2.2 Safety Stocks in Multi-Stage Systems Planned on a Rolling Horizon Basis - The Case of Uncertainties in Requirements 235

6.2.3 Safety Stocks in Multi-Stage Systems - The Case

	Page
of Uncertainties in Supply	247
Chapter 7 - REVIEW AND FUTURE RESEARCH	250
Glossary	255
Reference list	258
Biographical note	267

LIST OF FIGURES

		Page
1.1	A typical multi-stage production process	14
1.2	The multi-stage production process of fig. 1.1 in simplified representation	15
2.1	Assembly-line structure	36
3.1	A two stage system	43
3.2	Example of parts that cannot be aggregated	48
3.3	Example of parts that can be aggregated	50
3.4	Detailed structure of product 17	52
3.5	Aggregate structure 1	55
3.6	Aggregate structure 2	56
3.7	Aggregate structure 3	57
3.8	Aggregate structure 4	62
3.9	Final aggregate structure	64
3.10	Structure in which parts 1 and 2 can be aggregated together	69
3.11	Structure at the higher aggregate level	74
3.12	Family structure of product type 17	79
3.13	Two stage structures	84
3.14	Family structure of products in fig. 3.13	85
3.15	Three stage structures	86
3.16	Family structure of products in fig. 3.15	87
4.1	A simple aggregate structure	89
4.2	Two end products sharing one common part	95
4.3	The two products of figure 4.2, decoupled	96

		Page
4.4	Feasible elementary production schedules for the products of fig. 4.3	96
4.5	Tabular representation of the feasible elementary production schedules of fig. 4.4	97
4.6	Product structure considered by the set J_t^{3*1} of elementary schedules	112
4.7	Two products with a common part	120
4.8	A simple structure with three stages	125
4.9	A case with initial inventories of component parts 1, 2, and 4	130
4.10	Product structure with diverging arcs	135
4.11	The structure of figure 4.10 as viewed by an elementary schedule of type (b)	137
4.12	A five stage product structure	139
4.13	A five stage system	153
4.14	Time-phasing for a complete elementary schedule j_t for the product of fig. 4.13	155
4.15	End product with two component parts	160
5.1	Basic cell for pure assembly systems (basic cell of type I)	169
5.2	Basic cell for assembly systems with diverging arcs (basic cell of type II)	170
5.3	The simplest basic cell of type II (only two immediate successors)	170
5.4	Fluctuations of the installation stock at stage 1	175

		Page
5.5	The j-th subsystem	184
5.6	Typical cost curve	195
5.7	A case with independent demand for parts 5 and 10	199
5.8	Multi-stage structure with dummy facilities	200
5.9	Multi-stage structure	204

LIST OF TABLES

	Page
3.1 Characteristics of the component parts	53
3.2 Relevant information after aggregation at the final assembly stage	60
3.3 Relevant information for aggregation at the plastics shop stage	61
3.4 Relevant information for aggregation at the casting shop stage	63
3.5 Relevant information for aggregation at the forging shop stage	65
5.1 Results of the heuristic lot sizing	210
5.2 Relevant data for the problems of table 5.1	211
5.3 Comparison between unconstrained optimum lot sizes, and lot sizes provided by the best heuristic solution available	212

CHAPTER 1 - INTRODUCTION

1.1. General Statement of the Problem

In this research, the setting we are referring to as a MULTI-STAGE PRODUCTION SYSTEM is thought of as being a production process in which component parts have to be obtained by manufacturing or by purchasing, then assembled into subassemblies, assemblies, and finally into the finished good.

In the schematic representation of a multi-stage production process there are three elements that are either shown explicitly or implied:

- production operations such as: fabrication, assembly, packaging, other kinds of processings;

- stocking points representing inventories of raw materials, parts, subassemblies, and finished goods;

- flows of materials from operations to stocking points, and from stocking points to other operations.

Figure 1.1 shows a typical process structure. For the sake of simplicity, in what follows we will not represent the stocking points, although it will be implied that they are present following every operation. Figure 1.2 represents the process of figure 1.1 in the simplified version.

It is to be emphasized that there are other types of multi-stage problems that are not intended to be treated here, such as distribution

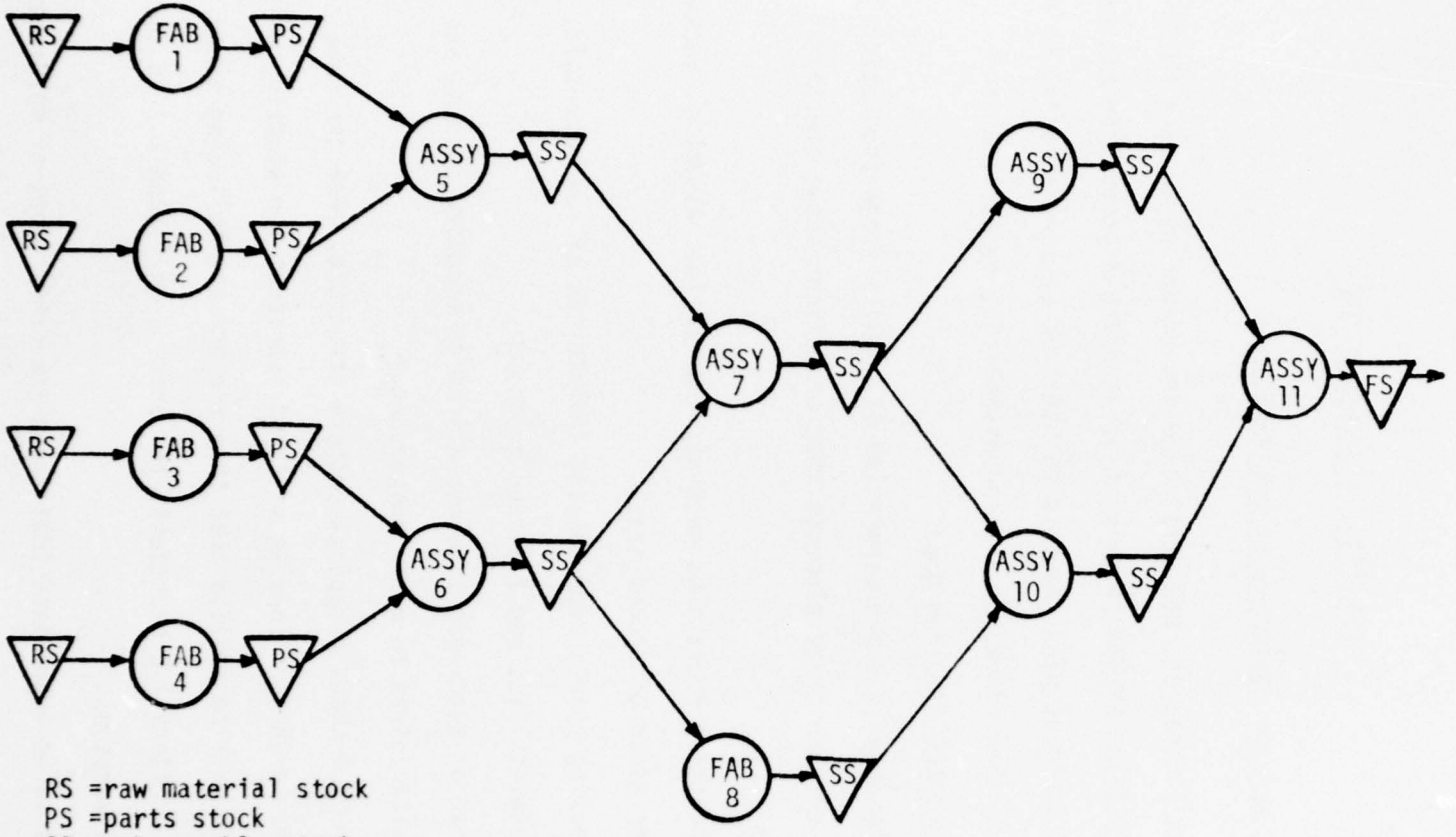


Fig. 1.1- A typical multi-stage production process.

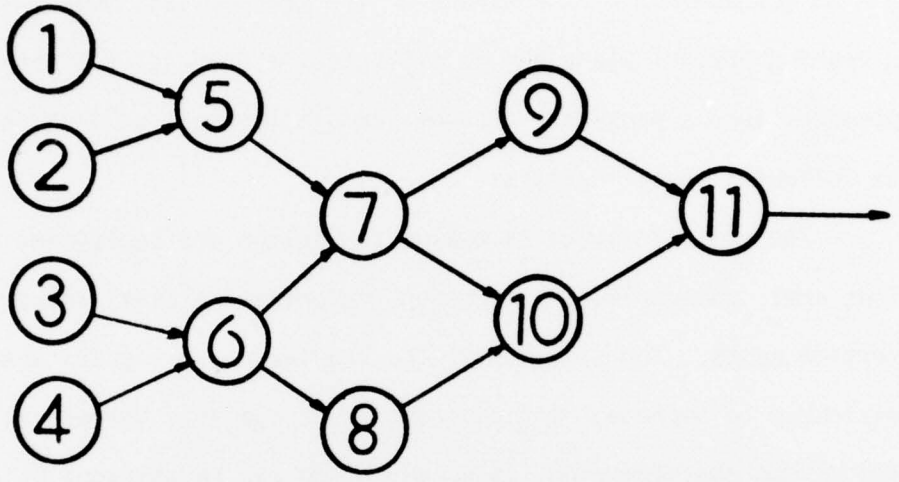


Fig. 1.2- The multi-stage production process of fig. 1.1 in simplified representation.

systems, and job shop scheduling and dispatching problems.

As an expository procedure, whenever a concept, for which there is no generally established definition in the literature, will be used for the first time, its contents, in the context of this research, will be fully specified. For further reference, a glossary of terms will be attached at the end.

The purpose of a planning system for the multi-stage process defined above, would be to determine the optimum production plan specifying when and how much to produce at every stage; if the size of the work force is also a decision variable the plan will have to contain

information with respect to the levels of hiring and layoff.

To get an idea of the nature of the mathematical models involved in drawing optimal plans and to motivate the need for the hierarchical approach, let us present a problem formulation that will abide by the following specifications:

- The costs involved in making production and employment decisions are: the variable production, inventory holding, setup, and overtime costs. For simplicity, the regular payroll costs are considered to represent a fixed commitment and are, therefore, excluded from consideration, as no advantage can be obtained by economizing on the use of regular time. The formulation will be later expanded to include additional costs and decisions (good discussions on production costs are provided by Holt et al., ch. 3 [49], and McGarrah, ch. 1.4, 5.4 [70]).

- The objective of the optimization is the minimization of the total costs over the planning horizon.

- The objective function will have linear cost components, with the exception of the setup cost. We have chosen this structure because it will lead to an aggregate planning model of the linear programming type, that presents a number of important advantages (Hax [47]).

The formulation of the multi-stage production planning model is then:

$$\text{Min } z = \sum_{s=1}^{S} \sum_{i_s=1}^{N_s} \sum_{t=1+\ell_s}^{T+\ell_s} [B_{i_s,t-\ell_{i_s}} \delta(x_{i_s t}) + v_{i_s,t-\ell_{i_s}} x_{i_s t} + h_{i_s t} I_{i_s t}] + \sum_{s=1}^{S} \sum_{t=1}^{T} c_t^s o_t^s \tag{1.1}$$

s.t.

$$I_{i_s,t-1} + x_{i_s t} - I_{i_s t} = r_{i_s t} \qquad \left.\begin{array}{l} t=1,\ldots,T+\ell_{i_s} \\ i_s=1,\ldots,N_s \\ s=1,\ldots,S \end{array}\right\} \tag{1.2}$$

$$r_{i_s t} = d_{i_s t} + \sum_{q=1}^{S} \sum_{i_q=1}^{N_q} a_{i_s i_q} x_{i_q,t+\ell_{i_q}} \tag{1.3}$$

$$\sum_{i_s=1}^{N_s} [m_{i_s}' \delta(x_{i_s,t+\ell_{i_s}}) + m_{i_s} x_{i_s,t+\ell_{i_s}}] - o_t^s \le R_t^s \qquad \left.\begin{array}{l} t=1,\ldots,T \\ s=1,\ldots,S \end{array}\right\} \tag{1.4}$$

$$o_t^s \le O_t^s \tag{1.5}$$

Nonnegativity constraints (1.6)

$$\delta(x_{i_s t}) = \begin{cases} 1 & \text{if } x_{i_s t} > 0 \\ 0 & \text{otherwise} \end{cases} \tag{1.7}$$

where $a_{i_s i_q}$ = composition factor; shows how many units of i_s are required per unit of i_q; $a_{i_s i_q}$ is defined only for pairs (s,q) of stages that are in an immediate predecessor-immediate successor relationship to each other; in all other cases $a_{i_s i_q}$ is zero.

$B_{i_s t}$ = cost of a setup for item i_s in period t.

c_t^s = cost of one hour of overtime at stage s in period t.

$d_{i_s t}$ = independent demand (i.e., requirements originating from

outside the model) for item i_s in period t.

$h_{i_s t}$ = inventory carrying cost per unit of item i_s held in stock from period t to t+1.

i_s = the i-th item produced at stage s; there is a total of N_s items produced by stage s.

$I_{i_s t}$ = units of item i_s to be left over as inventory at the end of period t.

l_{i_s} = production lead time for item i_s.

m_{i_s} = resource (in this case,labor) consumption for the production of one unit of item i_s.

m'_{i_s} = resource (in this case,labor) consumption by a setup for item i_s.

R_t^s, O_t^s = upper bounds on available regular and overtime labor, respectively, in period t at stage s. Here, we are using manhours as the limiting resource; other limitations upon the productive capacity can also be modelled such as limited equipment, energy, or raw materials availability. In fact, the issue of capacity can get very involved when it comes to assessing the productive capacity of a shop, plant or industry (Zabel [97], [98]).

$r_{i_s t}$ = total requirements for item i_s in period t.

s = the s-th stage; there are a total of S stages.

A stage is a pool of productive resource(s) that can be utilized to perform one or a number of operations upon one

or several items; a stage can include one or more machines or/and work places. The terms stage and facility will be used interchangeably.

t = denotes the t-th time period. The planning horizon contains T periods; this means that T production decisions have to be made for every item i_s.

$v_{i_s t}$ = the cost of producing one unit of item i_s when production starts at the beginning of period t.

$x_{i_s t}$ = amount of production of item i_s delivered to stock by stage s at the beginning of period t.

Problem (1.1) - (1.7) is a capacitated multi-state lot size model. The objective function (1.1) represents the sum of the above mentioned cost elements. (1.2) are the regular inventory balance equations, while (1.3) defines the total requirements for item i_s in period t as having two components : an independent demand, coming from customers, and a dependent demand, that is generated inside the planning model by stage s upon its immediate predecessor stages. Constraints (1.4) put an upper ceiling to the total amount of labor that can be planned for consumption at stage s in period t, and constraints (1.5) show that the amount of overtime is limited. The Kronecker delta $\delta(x_{i_s t})$ helps model the fact that a setup is incurred, with its associated cost and resource consumption, only when some positive x is produced.

(1.1) - (1.7) is a general formulation in that it allows a stage s to have <u>any</u> number of immediate predecessors and successors. The only restriction is that the network representing the process structure

(see figure 1.2) should be acyclic. A network is acyclic when it contains no cycles, that is, it is not possible for a material flow to leave a stage and then return to the same stage.

As a matter of convention, variable $x_{i_s t}$ represents the amount of production of item i_s delivered (or scheduled to be delivered) to stock at the beginning of period t; by necessity, that batch has to be started in production l_{i_s} periods in advance, i.e., at the beginning of period $t-l_{i_s}$. The inventory $I_{i_s t}$ is always counted at the end of period t. Therefore, by equation (1.2) requirement $r_{i_s t}$ can occur and be served at any time during period t. Also notice that, when costs are functions of time, the variable production cost is charged in the period when production starts.

1.2. Global vs. Hierarchical Approach

If someone wanted and could solve problem (1.1) - (1.7) we would say that a global approach is adopted, because all cost elements and all decisions are considered. It is immediately clear that for any real life mid-term (1 - 1-1/2 year) problem the global approach is not a viable alternative for two good reasons:

- The computational aspect - problem (1.1) - (1.7) is a large-scale minimization of a non-linear cost function over a set of non-linear constraints. If, for example, there were 20 facilities, 12 time periods (i.e. 1 year), and an average of 30 items per facility (giving a total of 600 component parts and su bassemblies, which is a small number for a manufacturing environment) the problem would have 7200 integer (0-1) variables and 7200 inventory balance equations plus

several hundreds of other constraints. Even a linear program with that many constraints, with no special structure, would be computationally infeasible. (About 4000 to 5000 is the upper bound on the number of constraints, without special structure, that can be handled by current linear programming codes; a problem with 4000-5000 rows would require, for solution, a computer time of the order of hours).

- The managerial aspect - in the first place the management decision process does not have a global structure; it is, rather, hierarchical with upper level decisions providing constraints upon lower level decisions (Anthony [4]). Secondly, providing all the detailed pieces of information, required by the model, for every single item over the entire planning horizon is most likely an impossible task.

For these reasons, a partitioning of some sort has to be performed upon problem (1.1) - (1.7). One approach (Hax and Meal [44]) starts with the assumption that setups are of secondary importance and, therefore, the setup cost is dominated by the costs of: producing, holding inventory, overtime, and changing the work force level. Consequently, as the first step in the production planning process, rather than solving a capacitated lot size model, setups are ignored and thus a linear programming model results. As mentioned above, the resulting linear program might still be too large to be solved; to cut its size down, items that share certain characteristics (to be seen later) are aggregated together into finished product types and aggregate parts (collectively called aggregate types). Notice that the product types

and aggregate parts have no physical correspondents; therefore, if they are used in the planning model setups have to be ignored anyway as there can be no setup for an abstract planning category. Due to this aggregate process and also because no detailed decisions (lot sizing, sequencing) are considered, this problem is called the aggregate planning model.

Since, as already mentioned, the end product types and aggregate parts are abstract notions, the aggregate plan cannot be directly implemented. It has to be disaggregated into quantities to be produced of each item, at each stage, in each time period. It is at this level of the planning process that setups, ignored at the aggregate level, are taken into account.

The approach outlined above has been termed hierarchical planning (by Hax and Meal [44]). It has been applied to one stage problems, and a number of disaggregation schemes have been experimented with (Hax et al. [46]). Results are encouraging in that, although the approach is obviously a suboptimization compared to the global approach, the partitioning into an aggregate level and a disaggregation level has not substantially increased total cost (Bitran and Hax [9]). The hierarchical approach not only makes the problem computationally feasible, but it is also managerially appealing: it parallels the hierarchical structure of the management decision process, reduces the amount and the degree of detail of information needed for planning purposes, and increases the accuracy of forecasts (broader, aggregate finished product types can be forecast with less errors than detailed, individual

items). There are, however, a number of issues that have to be resolved, related to the feasibility and consistency of the disaggregation process; they will not be brought up here, but good discussions can be found in Gabbay [34] and Golovin [36].

1.3 Outline of the Thesis

This thesis examines the multi-stage production planning in the context of the hierarchical approach. Rather than being a study experimenting with various methodologies of aggregation - disaggregation, this research has put the major effort in developing new results in areas relevant to the hierarchical planning in multi-stage systems. The necessity of this emphasis on new theoretical work has become evident after we have conducted a literature survey. The survey showed that a number of much needed developments are still missing from the area of multi-stage production, and the research tools for one stage system cannot be directly applied or are insufficient.

In Chapter 2 an attempt is made to identify major modelling issues and assumptions, on the basis of which a critical literature survey is conducted.

Chapter 3 presents an extension of the hierarchical approach from one stage system to multi-stage production; necessary and sufficient conditions for the aggregation of individual items into aggregate categories (product types and aggregate parts) are derived. An example is provided, and suggestions are offered for a two level aggregation for very large operations

The next two chapters deal with optimization problems at the two

levels of the hierarchical scheme. Chapter 4 regards the aggregate planning model as a large scale system. After establishing the difficulties associated with solving it, a new formulation of the problem is developed and the solution procedure by column generation is outlined. Chapter 5 goes to the other end (the detailed level) of the hierarchy, and the problem of the optimum lot sizes in multi-stage systems is approached and solved.

The issue of safety stocks for aggregate planning is tackled in chapter 6. As a byproduct of the results obtained for the multi-stage case, new findings are also reported for planning with rolling horizon in the one stage case.

Finally, chapter 7 puts the problem in perspective and offers suggestions for future areas of research.

CHAPTER 2 - A SURVEY OF THE LITERATURE

2.1. Typical Issues and Modelling Assumptions

In the area of aggregate planning in one stage production systems excellent and intensive surveys are given by Hax [47], and Johnson and Montgomery [52]; Eilon [32], although less comprehensive, also offers a coverage of the problem. For the multi-stage case no comparable survey could be located, with the exception, maybe, of chapter 3-8, 4-7 in Johnson and Montgomery [52].

In order to provide a systematic review of the literature we found that a number of issues and assumptions, discussed below, are central to determining the approach and the generality of the model.

The Planning Horizon assumption leads to two categories of models: models with finite horizon, and infinite horizon. The finite planning horizon assumption normally leads to planning models of some mathematical programming type (depending on the cost structure). The infinite horizon models are, in general, EOQ type models (where a trade-off between inventory holding and setup cost is sought).

The Production Capacity can be regarded as limited (capacitated models) or infinite (uncapacitated models).

The Cost Structure covers a wide spectrum from linear (von Lanzenauer [61]) to unrestricted functions (Williams [95]); the following cost structures have been identified:

- linear costs,
- setup cost plus linear production and inventory costs,
- general concave costs,
- piecewise concave costs, (i.e., concave on $(-\infty, 0]$. and on $[0, +\infty)$ but not on $(-\infty, +\infty)$,
- unrestricted cost functions.

Number of Products - there could be one product, or several products considered. Let us note that if a multi-product model has linear costs, linear constraints, and is uncapacitated, it immediately decomposes into a number of single product models equal to the number of products.

The Type of Demand can be deterministic or stochastic, and in either case it can be stationary (time invariant parameters) or non-stationary (the parameters change from period to period).

Structure of the production process:

- facilities (or stages) in series - each stage can have no more than one successor and one predecessor;
- facilities in parallel - each stage serves only market demands (i.e., independent demands) and no other facility;
- pure assembly systems - each stage can have any number of predecessor stages, but at most one successor stage;
- assembly networks with diverging arcs, or general structure systems - each stage can have any number of predecessor or successor stages (e.g., Figure 1.2).

Production Rates - there can be assumed infinite production rates (or instantaneous production), or finite production rates.

How the work-in-process moves between stages:

- For finite horizon models, which normally have a mathematical programming formulation, this assumption is not an issue since the material flow is governed by the inventory balance equations. According to these equations any excess amount at stage s is stored at the stocking point immediately following that stage, and a cost is charged according to the inventory level determined at the end of the time period. The stages that are successors of s withdraw from the inventory available at the stocking point as needed.
- For infinite horizon models this assumption determines the level of in-process inventories as a function of time. Work-in-process can move: in series (a batch cannot leave stage s before it is completely processed by that stage), in parallel (sub-batches of one or several pieces can be taken to the following operation as soon as they are finished at stage s but before the entire batch is finished), and mixed.

2.2. A Survey of Multi-Stage Models

Before briefly discussing the results reported in some representative papers we will mention some features of a more general nature.

Depending on the assumption about the planning horizon there are two different approaches to the problem. The infinite horizon models, where

development resembles the one stage order point-order quanitity models, start by evaluating the in-process inventories, and then minimize the sum of setup and inventory holding costs. Usually, demand is deterministic and constant, costs are time - invariant; in this static enviroment, the assumption of time - invariant lot sizes is reasonable. If demand is also occuring at a constant rate (i.e., continuous demand), time - invariant lot sizes are optimal (Schwarz and Schrage [81]). If production is instantaneous Crowston and Wagner [23], and Crowston et.al. [25] have shown that in pure assembly systems it is optimal to have the lot size at stage s an integral multiple of the lot size at the stage immediately succeeding s . This is not necessarily true, however, when production is non-instantaneous (see Jensen and Khan [51]), or when the structure of the process is an assembly network with diverging arcs (to be discussed later in Chapter 5). Whenever lot sizing in integral multiples is assumed or is optimal, the multiples are computed by dynamic programming procedures, by branch and bound algorithms, or even by considering all possible combinations of lot sizes if the problem is small enough (Taha and Skeith [87]). To reduce the search space, many procedures start by developing heuristic bounds on the range of lot size multiples to be considered.

Under the finite horizon assumption, normally leading to a mathematical programming model, demand is assumed discrete, varying or constant, deterministic or stochastic. Clearly, the infinite horizon can be viewed as a limiting case of the finite horizon, with the periods being made infinitely small and the number of periods infinitely large. The equivalent of what used to be the production rate is now the lead time. A

lead time of zero is equivalent to the former instantaneous production (production started and ended within one time period), while a positive lead time is similar to a finite production rate (when the processing of a batch will take up a number of time periods).

If the production process structure under study is facilities in series, the problem can be represented as a network flow problem [1] (of the transshipment type). If the cost structure is linear network algorithms, of the dynamic programming type, exist (Zangwill [100] for the single product case; Veinott [92] extended Zangwill's approach to arborescent and assembly line structures).

When an assembly structure is treated, the resulting model is no longer a network flow problem; it becomes a combinatorial problem, and either dynamic programming or branch and bound algorithms can be used to obtain a solution. The authors who studied this kind of problems put considerable effort into:

- exploiting the characteristics of the optimal solution to give good formulations for the recursive equations (for the dynamic programming algorithms);
- finding good bounds (for the branch and bound algorithms);
- finding "reasonable" restrictive assumptions upon the cost functions, which would enable them to develop better bounds and more efficient algorithms.

---

[1] An interesting network interpretation for an uncapacitated single stage problem with variable production costs plus inventory holding, and production rate change costs is given by Hu and Prager [50].

In general, the models we found in the literature are uncapacitated, with some exceptions:

- some models provide the constraint that during a time period only one product may be processed on any one facility (Dorsey et al. [30], [31], von Lanzenauer [62]);
- von Lanzenauer [61] provides capacity constraints in his linear programming variable workforce multi-stage model;
- Gabbay [33], [34] treats the capacitated case for the facilities in series structure, under rather restrictive cost assumptions;
- Taha and Skeith [87] allow limits on storage space;
- Klingman et. al. [55] treat the capacitated problem for facilities in parallel with one period planning horizon;
- Mangiameli [68] formulates a mixed integer capacitated problem and attempts solution by Benders method; his formulation is restrictive in that no inventory of component parts is allowed between pairs of adjacent stages.

In the above discussion we have not included Material Requirements Planning (MRP) research, although there are papers that try to deal with the lot sizing problem in MRP systems under capacity constraints (e.g., Biggs and Hahn [8], Collier [19]). The reason for this is that MRP, although it grew to become a separate area of production management, is an incomplete approach to planning production in multi-stage systems. It totally relies on the existence of a master schedule (Orlicky [77]) that is assumed feasible with respect to the capacity limitations at all stages involved in the production process. Unfortunately, the literature search

showed that good procedures to draw such master schedules are yet to be developed (for a more extensive critique of MRP see Kanodia [54]).

In our survey of the literature no model was found to be hierarchical and apparently the problem is not percieved to have this nature by most authors, with the exception of Goodman [37] who classifies the decisions to be made in a multi-stage system into: the aggregate planning process, the derivation of the master schedule, and the detailed planning of the manufacturing operations. Also, some hints about the hierarchical structure are offered by Plossl and Wight [79], and Magee and Boodman [67].

Since the assumption on the finiteness of the planning horizon proved to be of great importance, leading to two different types of formulations and solutions to the problem, we will divide the papers reviewed below into two categories. Brief remarks will be provided with each reference.

2.2.1. Infinite Horizon Models

CROWSTON and WAGNER [23], and CROWSTON et. al. [25] considered pure assembly structures, single product, uncapacitated environment, time invariant costs. They have proven the optimality of what had since long become a well known policy in production processes, namely: the lot size at some stage s has to be a positive integral multiple of the lot size at the succeeding stage. For this policy to remain optimal under non-instantaneous production rate, it is necessary to assume something which the two papers do not explicitly state: processing of a batch at a given stage s cannot start before all the needed material input (from predecessor stages) for that batch is

available to stage s. Unless this assumption is made, it is easy to produce examples that would invalidate the optimality of the integral multiples policy (Jensen and Kahn [51] p. 131, Thomas [88]). For the derivation of the optimal set of integral multiples a dynamic programming (DP) algorithm is developed.

CROWSTON, WAGNER, and HENSHAW [22] develop heuristic procedures for the determination of the lot size multiples. It is suggested to use the DP algorithm of [23], [25] for problems of moderate size, and the heuristics for larger problems.

JENSEN and KAHN [51] - for uncapacitated serial facilities, dynamic programming is used to determine the optimal cycle time (length of time between two subsequent start-ups) for each stage, so as to minimize setup plus inventory holding costs. Also, the minimum delays between the initial start-up times for successive stages are computed, in order to forestall shortages of component parts. "Reasonable" minimum and maximum number of production runs per year are established before the DP algorithm is applied.

SCHWARZ and SCHRAGE [81] look at ways to find the optimal lot sizes values in a pure assembly system under the integral multiples policy. First a branch and bound algorithm is used; then, a system myopic policy is developed, by which the lot size multiple at stage s is determined by isolating stage s and its immediate successor from the production process network. 500 test problem produced good results: in 50% of the cases system myopic policies were optimal, and in the other cases the average error did not exceed 5% of the optimal cost.

TAHA and SKEITH [87] - for uncapacitated facilities in series it is suggested ([87] p. 160) that the optimal lot size multipliers be found by trying different combinations of values. If the limits on storage space become constraining a constrained lot size model should be considered. In such a case the authors assert ([87] p. 161) that "no general procedure can be suggested for this problem except probably by the use of trial and error."

THOMAS [88] provides an interative algorithm to compute the lot sizes successively stage by stage (starting with either the first or the last stage of the series structure considered).

DARUKHANAVALA [29] deals with assembly structures with diverging arcs. Demand for the finished product is taken to be stochastic with constant mean. An algorithm is presented, which derives the dependent stochastic demand distributions at every stage, and then computes optimum lot sizes with the objective of minimizing the total setup, inventory holding, and safety stock costs.

SZENDROVITZ [85], and SZENDROVITZ and WESOLOWSKY [86] present two models for the serial facilities setting. One model assumes the lot size to be uniform at all stages, and equal sized sub-batches can be transported from a given stage to the next production stage before the production of the lot is finished. The second model assumes that the lot size at a stage is an integer multiple of the lot size at the succeeding stage, and the product is transferred between stages in lots. An optimal solution is found for the first model, and a good approximation to the optimal solution for the second model.

GOYAL [38] considers the production of one item, which is then put in different packages. There is a setup every time a batch is manufactured; also a setup is incurred whenever the packaging operation is prepared for a certain type of package. Under the objective of minimizing setup plus inventory holding costs, an algorithm is developed to determine the optimum manufacturing and packaging frequencies. Criticism of the assumption of equally spaced packaging runs has been brought by Andres and Emmons [3].

2.2.2. Finite Horizon Models

ZANGWILL [99] - there are N facilities forming an acyclic network; each facility produces a single item, that can be delivered to the market or can be input to another facility; backorders are permitted. The setting is uncapacitated, demand is deterministic; production cost is concave, and inventory cost is piecewise concave (that is, concave on $(-\infty, 0]$ and on $[0, +\infty)$ but not on $(-\infty, +\infty)$). The optimum solution is shown to be part of a dominant set of schedules. Algorithms, of the dynamic programming type, are developed for the parallel facilities, and for the serial facilities cases. More general, assembly type structures, are not covered because of the insurmountable difficulties involved in developing an efficient algorithm.

ZANGWILL [100] - the second part of the paper deals with a single product, uncapacitated, series facilities production process. Demand for the finished product is deterministic, and production and inventory

holding costs are concave on the nonnegative orthant (zero at the origin); no backorders are allowed. A network interpretation of the inventory flow is provided, in which a node is attached to every stage-time period pair. The optimum solution is of the extreme flow type, that is, at every node at most one arc entering the node will have a positive flow:

$$x_{st} I_{s,t-1} = 0, \qquad \text{all } s, t$$

where s is the stage, t the period, x_{st} production at stage s in period t, and $I_{s,t-1}$ inventory at stage s at the end of period $t-1$ and beginning of period t. A dynamic programming althorithm, that starts with the final stage and the last time period, and works backwards, was developed.

VEINOTT [92] has extended the work by Zangwill [100] to cover structures more complex than serial facilities. Maintaining the uncapacitated environment, with a single product, deterministic non-stationary demand for both the final product and the component parts, and no backorders, Veinott considers assembly-line structures (Figure 2.1).

The proposed solution algorithm is based on an enumeration of all extreme flow production schedules for the final stage (stage 9 in Figure 2.1). Every such production schedule generates a set of requirements upon the stages that are immediate predecessors to the final stages, and thus the problem decomposes into a number of pro-

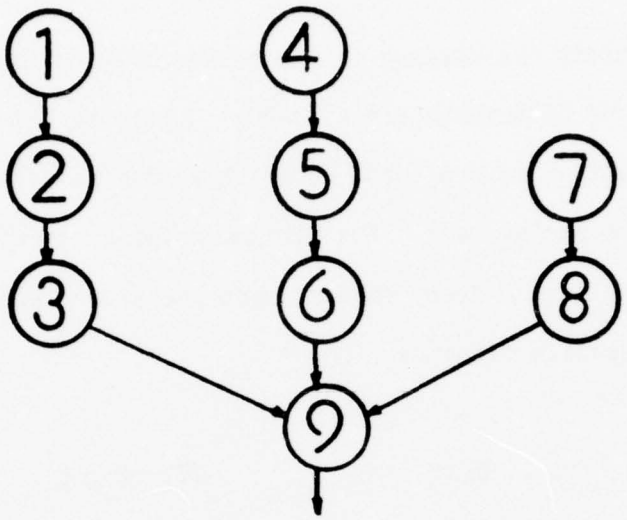


Fig. 2.1 - Assembly-line structure.

blems with facilities in series. These, then, can be solved by Zangwill's [100] algorithm. The computational effort is linear in the number of facilities in the system.

Other extensions to distribution systems have been provided.

CROWSTON and WAGNER [23] proposed a 0-1 formulation for uncapacitated serial or pure assembly system, to be solved by Benders method; no computational experience is reported.

CROWSTON and WAGNER [26] studied the uncapacitated pure assembly systems with concave production costs and linear inventory holding costs. Under the additional assumptions that production costs are nonincreasing in time, and inventory holding costs are nondecreasing over stages the optimal solution is an extreme flow and has the nested

property. The nested property implies that if stage s produces in period t its immediate successor will also produce in period t. A dynamic programming algorithm is proposed, whose solution time increases exponentially with the number of time periods, and linearly with the number of stages. For the case where the number of periods is large, but the structure is close to serial, a branch and bound algorithm is suggested.

CROWSTON et. al. [24] develop heuristic decision rules to solve a "newsboy problem" applied to a multi-stage production process, when the final product has a limited sales season.

DORSEY et. al. [30], [31] look at the problem of facilities in parallel, with multiple products, and capacity constraints in the sense that during a time period only one product can be assigned to any one facility. The problem is formulated as a network flow problem, for which an efficient solution procedure is given. A number of extensions are suggested, such as: warehouse space limitations, variable number of facilities available in each period; these can be solved by more general minimum cost flow algorithms.

von LANZENAUER [61] formulates the aggregate planning problem in multi-stage systems as a linear program with variable work force (no setups are considered). The objective is to determine the production and employment levels so as to maximize total profits over the planning horizon. The problem we see with this formulation is that it can easily become computationally infeasible (see Chapter 4) because of the very large number of inventory balance equations.

von LANZENAUER [62] - this paper considers a general multi-stage system that produces a number of different products. No assembly operations can take place; every product goes through a serial sequence of operations, which sequence can be different from product to product but known according to technological requirements. The problem is to determine the sequence of products at every stage, as well as the lot size for each product with the objective of minimizing the total cost (over the planning horizon) of the setups, holding inventory, and shortages. The formulation is a 0-1 mixed integer program, with the capacity constraint that during a time period only one product may be processed at any one stage. On the computational side, however, the author himself says [62] p. 111: "With the presently available integer programming codes it is impossible to solve problems of fairly realistic size."

GABBAY [33], [34] addresses the capacitated serial structure, with single and multiple products, under linear production and holding costs. For the single product case, under the assumption that production costs are non-increasing in time and inventory holding costs are non-decreasing over stages, it is shown that the problem may be solved by treating each stage separately, starting with the last stage and last period. For the multi-product case algorithms are developed under the quite restrictive assumption that the same percentage of their final value is added, at some stage, to <u>all</u> products processed in the system.

LOVE [66] - an uncapacitated, serial facilities, single product system, with deterministic demand is considered. Production and holding costs are concave on the positive orthant and zero at the origin. The paper establishes that under production costs that are non-increasing in time, and inventory holding costs that are non-decreasing over stages the optimal production schedule is nested and extreme flow. Using these properties, Zangwill's [100] dynamic programming algorithm is specialized.

WILLIAMS [95] treats the single product, uncapacitated facilities in series, with stochastic stationary demand for the end product; the objective function includes the setup cost and production and inventory costs that are unrestricted in form. A dynamic programming algorithm is formulated, in which the functional equation expresses the expected total cost from period t $(1 \leq t \leq T)$ to the last period T in the planning horizon.

BRYAN et. al. [14] - the problem is that of stock distribution in a four-stage manufacturing process in the case where demand for the finished good is seasonal. The two costs that have to be traded off are: losses on liquidation of stock left over at the end of the season, and profits lost as a result of running out of stock or as a result of failure to fill customers' orders immediately. The optimum quantity of inventory to be held at each of the four stages, as of the beginning of the sales season, is determined.

A number of other papers treat multi-stage systems; thus, KALYMON [53] deals with an uncapacitated arborescent structure, KLINGMAN et. al.

[55] addresses a single period, capacitated parallel facilities problem, where lot sizes have to be selected from a predetermined finite set of possible lot sizes, and a paper by MANGIAMELI [68] has been already mentioned earlier in section 2.2.

If we are now to compare the amount of research that has been produced in the area of multi-stage production vs. single stage, a substantial disparity will be immediately evident, in the sense that the multi-stage field is much less developed, although the vast majority of production processes are of a multi-stage nature. This is explainable in part by the added degree of difficulty, and by the still limited capability to conduct large scale optimizations.

Researchers who have adopted a global approach to the problem, as are most papers under the finite horizon heading, spent their efforts in finding properties of the optimum solutions and, then, developing DP, branch and bound, or network algorithms. Although many results of highly theoretical value have been found, the algorithms are of limited practicality, applicable only to small problems or to problems with special cost structures.

As for the infinite horizon models, they only address part of the planning problem, with obvious limitations brought about by the restrictive assumptions made.

CHAPTER 3 - A HIERARCHICAL APPROACH

The survey of the literature shows that a conceptually new way of tackling the multi-stage production planning problem is needed. This author's opinion is that the hierarchical decomposition of the global problem is a reasonable approach, that will make the solution to the multi-stage production planning problem both managerially appealing and amenable to computational feasibility.

Our proposal here is to extend the hierarchical methodology, presented in section 1.2 for one stage systems, to the multi-stage case. Thus, the overall problem will be decomposed into:

- the aggregate problem, or the aggregate level, and
- the disaggregation problem, or the detailed level.

The basic assumption we are making is that setup costs (which are costs of a detail nature) are dominated by the aggregate costs or, put otherwise, setup costs are of secondary importance. The hierarchy of costs leads to the hierarchy of goals: optimize the aggregate problem first, and, then, perform the disaggregation optimally. The disaggregation will have to observe the constraints imposed by the aggregate plan.[1]

3.1. The Aggregate Level

The aggregate planning model, as already mentioned, will be a linear program. Formally it looks like the global model (1.1) - (1.7), without

---

[1] This hierarchy of goals is somewhat similar to goal programming; see Lee [65].

including setups:

$$\text{Min } z = \sum_s \sum_{i_s} \sum_{t=1+\ell_{i_s}}^{T+\ell_{i_s}} (v_{i_s,t-\ell_{i_s}} x_{i_s t} + h_{i_s t} I_{i_s t}) + \sum_s \sum_t c_t^s o_t^s \qquad (3.1)$$

<u>s.t.</u>

$$I_{i_s,t-1} + x_{i_s t} - I_{i_s t} = r_{i_s t} \qquad (3.2)$$

$$r_{i_s t} = d_{i_s t} = \sum_{q=1}^{S} \sum_{i_q=1}^{N_q} a_{i_s i_q} x_{i_q,t+\ell_{i_q}} \qquad (3.3)$$

$$t = 1,\ldots,T+\ell_{i_s}; \quad i_s = 1,\ldots,N_s; \quad s = 1,\ldots,S$$

$$\sum_{i_s} m_{i_s} x_{i_s,t+\ell_{i_s}} - o_t^s \leq R_t^s \qquad (3.4)$$

$$o_t^s \leq O_t^s \qquad (3.5)$$

$$t = 1,\ldots,T; \quad s = 1,\ldots,S$$

Nonnegativity constraints (3.6)

The notations stayed unchanged; wherever the summation limits are not specified, it is assumed that: $s = 1,\ldots,S$; $i_s = 1,\ldots,N_s$; $t = 1,\ldots,T$.

It is not a difficult matter to expand the model formulation so as to include: hiring and layoff decisions, multiple resources, backorders, and independent demand for parts.

3.1.1. Modelling Backorders

A note of caution about backorder modelling is in order here, before we proceed further: if the aggregate model does not contain independent demand for parts, backorders can only be planned at the end product level but not at the component parts level. If the model includes independent demand for parts, backorders can be planned at the parts level but only to the extent that independent demand can be backlogged. The explanation is simple: if one allows backordering of the dependent demand for parts, a paradoxical situation could develop in which the needed parts are backordered (i.e., produced in a later period) while assembly will take place on schedule.

To illustrate consider a two stage model (Figure 3.1), in which stage A represents part production, stage B assembly of the finished product, lead times are zero, and the composition factor $a_{AB} = 1$.

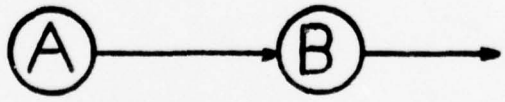


Fig. 3.1- A two stage system.

If there is no independent demand upon stage A, and if we allow backordering at A and B , the inventory balance equations are:

$$I^{+}_{A,t-1} + x_{At} + I^{-}_{At} = x_{Bt} + I^{+}_{At} + I^{-}_{A,t-1} \qquad (3.7)$$

$$I^{+}_{B,t-1} + x_{Bt} + I^{-}_{Bt} = d_{Bt} + I^{+}_{Bt} + I^{-}_{B,t-1} \qquad (3.8)$$

where I^{+} is positive inventory, and I^{-} represents backorders.

Suppose that the amount x_{Bt} scheduled to be assembled is too large, so that the quantity of parts available, i.e., $I^{+}_{A,t-1} + x_{At}$, cannot cover the amount required by x_{Bt}. In this case there is nothing to prevent the LP from backordering production of parts; technically this would be done by increasing I^{-}_{At} until equality is achieved in (3.7) without having to lower the amount x_{Bt} to be assembled. Evidently, such a solution can not be accepted.

If there is independent demand d_{At} for component parts A, backorders have to be limited to independent demand only. This can be done by re-writing equations (3.2) - (3.3) in a suitable form, and by imposing an additional constraint upon the component part producing stage:

$$I^{+}_{A,t-1} + x_{At} - I^{+}_{At} = r'_{At} + r''_{At} \qquad (3.9)$$

$$r'_{At} = x_{Bt} \qquad (3.10)$$

$$r''_{At} = d_{At} + I^{-}_{A,t-1} - I^{-}_{At} \qquad (3.11)$$

$$r''_{At} \geq 0 \qquad (3.12)$$

$$I^{+}_{B,t-1} + x_{Bt} - I^{+}_{Bt} = d_{Bt} + I^{-}_{B,t-1} - I^{-}_{Bt} \qquad (3.13)$$

The inventory balance equation for the final stage did not change (it is the same as (3.8)). But, for the component parts stage the requirements term is expressed as the sum of two components: r'_{At} and r''_{At}, corresponding to the dependent and independent demand. By the nonnegativity constraint (3.12), it is ensured that only independent demand can be backordered.

3.1.2. Issues of Aggregation

From the example given in section 1.2, it is clear that the LP formulation (3.1) - (3.6) can easily become too large to be computationally feasible. Therefore, an aggregation of individual items (either end products or component parts) into aggregate types, which will reduce both the number of variables and the number of rows, will have to be performed; also resources will have to be pooled into aggregate stages.

3.1.2.1. Aggregation of Items

There is a condition for the grouping of individual items into types: the solution obtained from the aggregate model has to not differ from the aggregated solution that would be obtained if we solved the detailed model (i.e., before aggregation) and, then, aggregated the result. Given that we are using a LP at the aggregate level, the specific conditions are not hard to find. Let c be a $(1xn)$ vector, x a $(nx1)$ vector, A a (mxn) matrix, and b a $(mx1)$ vector. The aggregate model is:

$$\min z = cx \tag{3.14}$$

s.t.

$$Ax = b \tag{3.15}$$

$$x \geq 0 \tag{3.16}$$

If $A = (A_1, A_2, \ldots, A_n)$, where $A_1, A_2, \ldots, A_n$ are the columns of A , let vector V_j , $j = 1, \ldots, n$, be defined as $V_j = \begin{pmatrix} c_j \\ A_j \end{pmatrix}$.

Then, two variables x_i , x_j can be collaspsed into one aggregate variable X_{ij} , without affecting the optimum value of the objective function, if $V_i = V_j$.

If x_i , x_j represent two items satisfying the $V_i = V_j$ condition, and if we aggregate them together as X_{ij} and solve the LP , we will then be indifferent with respect to what mix of x_i , x_j will be produced in reality as long as the sum $(x_i + x_j)$ will be equal to the solution X_{ij} . The mixing decision will not and cannot affect the optimum objective value. Since we would be indifferent (in the LP context) whether x_i , or x_j , or both are produced, we will call the group represented by X_{ij} an "EITHER ... OR ..." aggregate type.

Besides the conditions implied by $V_i = V_j$, there are also other aspects that have to be considered when creating the aggregate types, aspects which are important for disaggregation purposes.

Let us now summarize the requirements that have to be satisfied by two or more items in order to be included into the same aggregate type:

1) The items in question have to have the same routing through the set of production facilities; the same resources have to be consumed, and

lead times have to be similar for all member items of the aggregate type.

2) The items have to have similar resource consumption at any given stage (facility).

3) The items have to have similar variable production and inventory holding costs at any given facility involved in their production.

4) The items should enter into the same next level assembly(ies) in the same proportions. If the items in question are end products this requirement is equivalent to having similar demand patterns[1], that is, similar trend and seasonalities, although not necessarily the same level of overall demand.

Requirements (1) through (3) above are a direct consequence of the $V_i = V_j$ condition, with the sole exception that the mathematically rigorous equality sign of $V_i = V_j$ has been worded as "similar" rather than "identical", in order to accommodate realism and some flexibility. We will say that two or more items (either end products or component parts) that satisfy requirements (1), (2) and (3) have similar production profiles.

---

[1] These conditions are valid for a linear programming aggregate planning model. If other types of models are used, the conditions may change accordingly. It is interesting to note, for instance, that in the case of a quadratic cost model Krajewski et. al. [58] reported that, under conditions of varying demand, the parameters of the aggregate cost function, contrarily to the general belief, are highly related to demand.

If the aggregated items are finished products, by requirement (4) they have to have similar demand patterns. This is so because the aggregate plan cannot distinguish among items within a product type, so that it will assume that all member items' demands behave the same way the aggregate demand does. If in reality this is not true, the disaggregation might lead to holding inventory in the wrong member item, which,under tight capacity, can lead to backorders.

For the component parts, the following example will justify requirement (4): suppose there are two inidividual parts, PART 1 and PART 2, processed on the same machine, and satisfying requirements (1) - (3). PART 1 goes into ASSEMBLY 1, and PART 2 goes into both ASSEMBLY 1 and ASSEMBLY 2 (Figure 3.2).

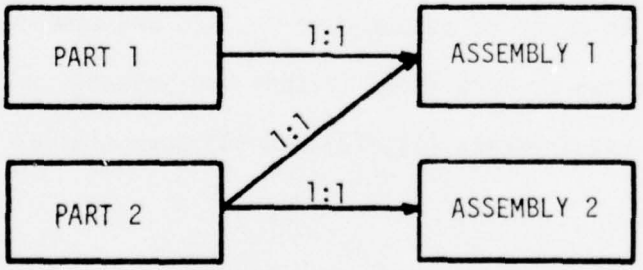


Fig. 3.2- Example of parts that cannot be aggregated

If one tries to aggregate the two parts together, there will be insurmountable difficulties when trying to handle initial, or starting, inventories. Indeed, assume that there is an initial stock of 2000 units of PART 1 and nothing of PART 2, and the 2000 units will be input to the aggregate model as starting inventory (i.e., the familiar I_0) of the

aggregate part defined as: EITHER PART 1 OR PART 2 . Suppose that the aggregate plan schedules in production 1000 units of ASSEMBLY 1 and 2000 units of ASSEMBLY 2 for the upcoming period; clearly, 1000 units of PART 1 and 3000 units of PART 2 are required, or a total of 4000 units of aggregate part. The model sees 2000 units already available and will plan to produce 2000 more. But this is wrong since we knew that 3000 units of PART 2 will be needed; this happened because the model did not distinguish between PART 1 and PART 2 and used stock of PART 1 to satisfy requirements of PART 2.

The way to avoid this misallocation of capacity is to not let the model handle the initial inventories but rather to allocate them outside the model (something similar is done for the finished products to obtain what are called effective or net demands; see Bitran and Hax [97]). The difficulty is that, unlike the end product case, when dealing with component parts the requirements for them are unknown prior to solving the model, and thus it is impossible to allocate initial inventory against forthcoming requirements. In this situation, what condition (4) does is to ensure that the requirements for the two individual parts always occur in the same proportion; therefore, before running the aggregate planning model we will re-balance the inventories of parts to ensure the feasibility of the disaggregation (this will be discussed in more detail later). To illustrate, let us change figure 3.2 to figure 3.3.

PART 2 and PART 1 are required for ASSEMBLY 1 in proportion of 3:1 , and for ASSEMBLY 2 in the same proportion of 3:1 . No matter of what production levels are set for the two assemblies, PART 2 and PART 1 will

always be needed in proportion of 3:1 .

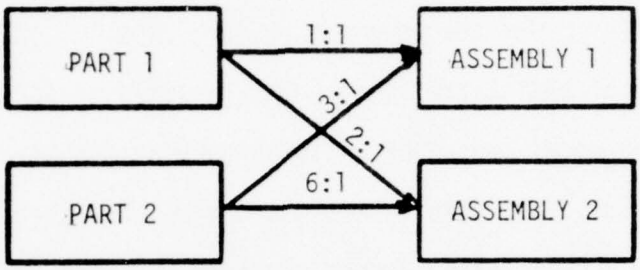


Fig. 3.3- Example of parts that can be aggregated

This is the meaning of condition (4) above, and although it does not answer directly the question of: "how much of the initial inventory of PART 1 should be considered as starting aggregate inventory," it will be of help when handling the initial stocks of parts.

If there is independent demand for parts, there are several ways in which this can be dealt with:

- If the volume of spare parts required is very large, and especially when parts are still needed for older models of end product that are no longer made, production of spare parts can be regarded as an activity separated from the manufacturing of the end products. Special plants for spare parts can be operated, or subcontracting can be used.
- If spare part production has to be included in the "regular" aggregate plan, drawn for the production of end products, the creation

of aggregate parts would require that the member parts should also have similar independent demand patterns, in additon to requirements (1) through (4). Since all these conditions would most likely be extremely hard to meet, the problem can be tackled in one of the following ways:

a) If the volume of spare parts is small compared to regular production, their effect will be minimal; therefore the creation of aggregate parts will still be conducted by conditions (1) - (4).

b) If the volume of spare parts is significant, separate production and inventory variables will be assigned to the regular production and to the production for service purposes of the same component part. The aggregation will be performed according to conditions (1) - (4), separately for spare parts (regarded as end products) and separately for component parts.

In both cases (a) and (b), outside the aggregate model and before running it, independent demand for parts, over the corresponding lead time, will be filled from initial inventory and, then, the remaining stock of parts will be re-balanced as it will be shown later.

3.1.2.2. An Example of Aggregation

To illustrate the way in which the requirements (1) - (4) govern the grouping of individual items into aggregate types consider the following simple example: a product is assembled from 16 component parts as shown in figure 3.4. Four shops or facilities are involved in the production of parts: casting, forging, fabrication and plastics. Normally lead times are about a month, except at casting where the lead time is two months.

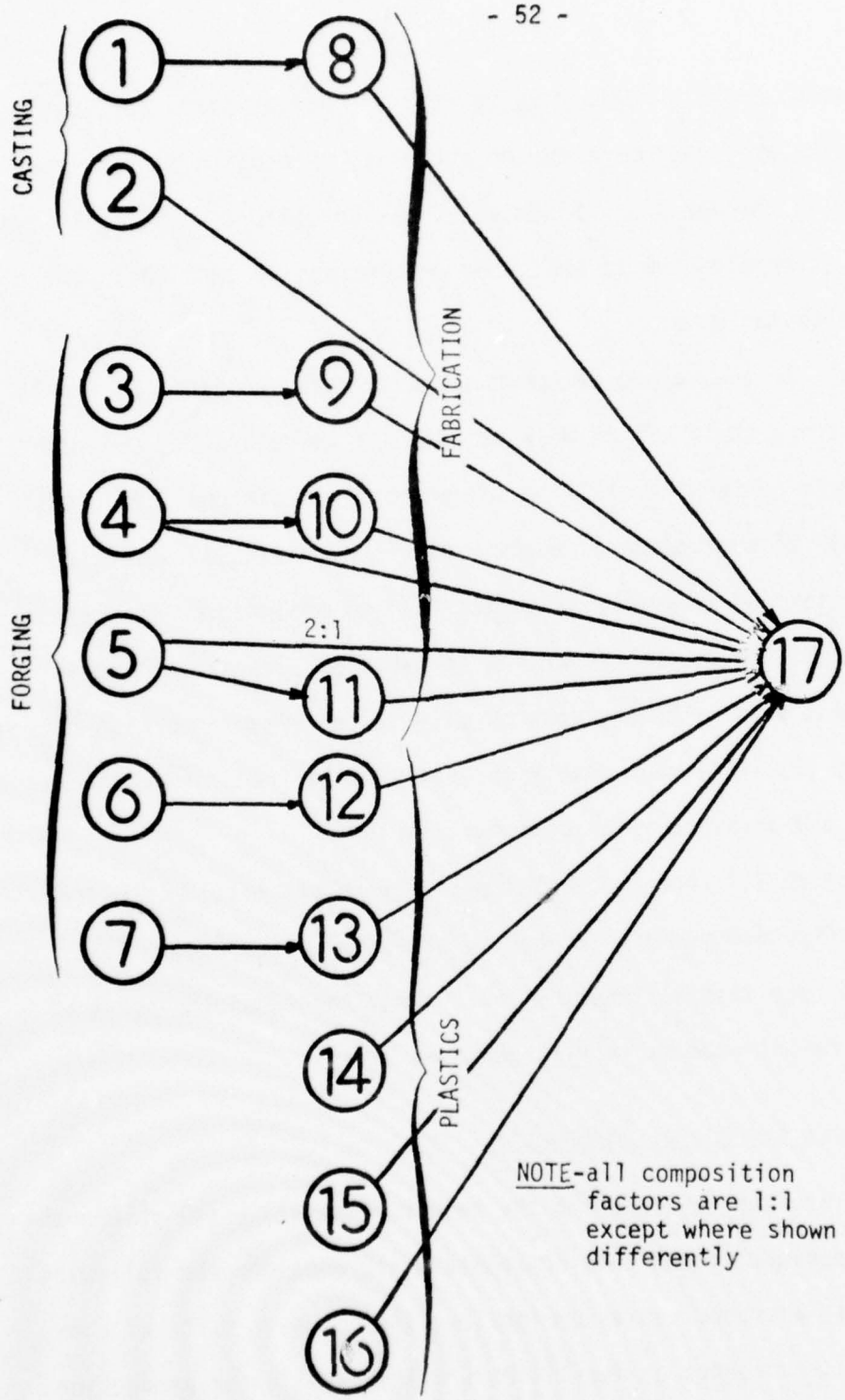


Fig. 3.4- Detailed structure of product 17.

Table 3.1 summarizes the characteristics of all 16 component parts. We are considering one productive resource (e.g., labor) in every shop, and we are assuming that the percentage holding cost charge (with respect to the value of the stock item) is uniform over all stages and all inventory items. There are no constraints with respect to raw materials availability.

Table 3.1 - Characteristics of the Component Parts

Part No.	Shop	Lead Time (months)	Productivity (Units/Hour)	Value of Completed Part ($/Unit)
1	Casting	2	10	3
2	Casting	2	10	3
3	Forging	1	5	7
4	Forging	1	5	7
5	Forging	1	5	7
6	Forging	1	5	15
7	Forging	1	5	7
8	Fabrication	1	20	12
9	Fabrication	1	20	12
10	Fabrication	1	20	12
11	Fabrication	1	20	12
12	Plastics	1	60	17
13	Plastics	1	60	17
14	Plastics	1	60	17
15	Plastics	1	60	17
16	Plastics	1	60	17

The production process goes as follows: parts 1 and 2 are cast, but before assembly part 1 has a machining operation (after which it becomes part 8). Parts 3, 4, 5, 6 and 7 are forged; after machining,parts 3, 4, 5 become parts 9, 10 and 11 respectively; 6 and 7 are plastic coated. Components 14, 15, 16 are all mold injected.

The end product 17 comes in two models: one with all plastic parts and plastic coated parts red, and the other one with all blue. The difference in color has no impact on production costs and productivities, and there is no significant cost associated with switching production from red to blue and vice versa. The two (red, and blue) end products present identical demand patterns; no component part exhibits independent demand.

Our task now is to perform the aggregation in order to be able to run an aggregate planning model. We will work thorugh the example in detail, in order to make the aggregation conditions clear. The example has been set in such a way as to emphasize the requirements imposed upon the aggregation by multi-stage structure considerations rather than by the costs and other parameters.

The aggregation will start with stage 17 and will proceed by production facilities.

FINAL ASSEMBLY (end product 17)

As we have already mentioned, the two end products are identical except for the difference in the color of the plastic and plastic coated parts. Consequently, they have identical production profiles and since, by assumption, they exhibit similar demand patterns, they can be aggregated to form a product type, say, PRODUCT TYPE 1 .

If we identify the product with red plastic parts by 17r and the product with blue parts by 17b , we would write

$$1 \text{ unit of PRODUCT TYPE } 1 = (\underset{1 \text{ unit of } 17r}{\text{EITHER}} \ , \ \underset{1 \text{ unit of } 17b}{\text{OR}})$$

After this step has been completed we are left with the aggregate structure shown in figure 3.5 (again, "r" depicts component parts belonging to 17r, and "b" component parts belonging to 17b).

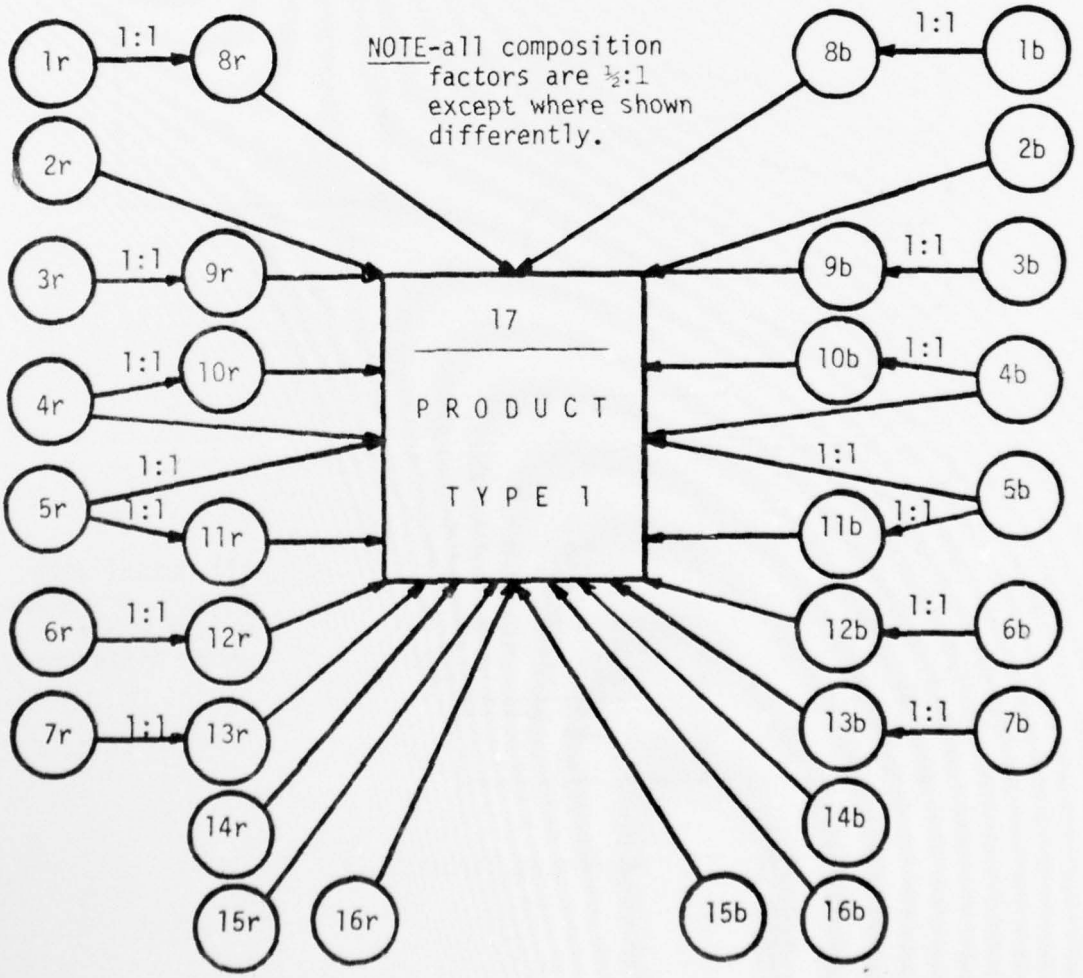


Fig. 3.5- Aggregate structure 1.

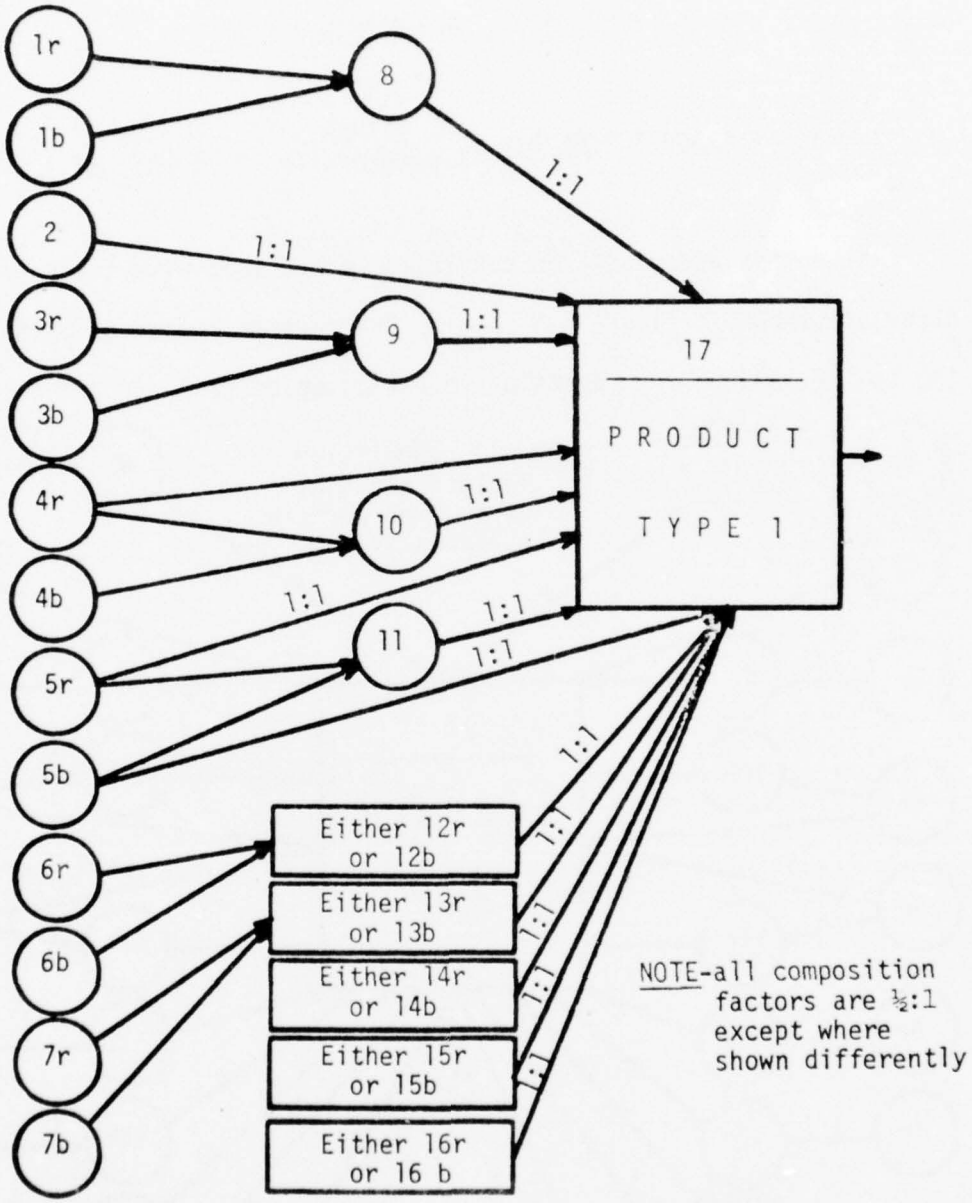


Fig. 3.6- Aggregate structure 2.

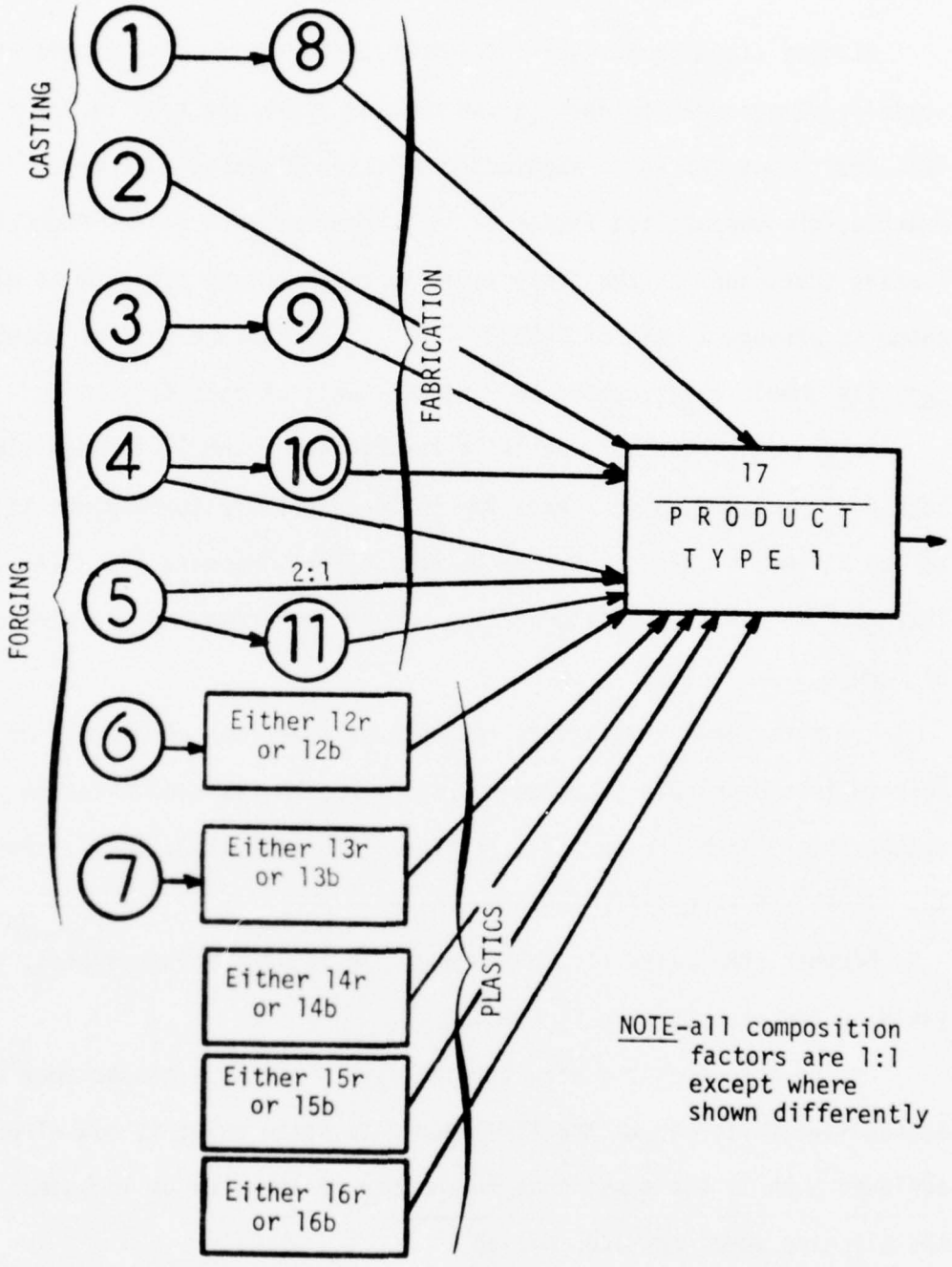


Fig. 3.7- Aggregate structure 3.

At this time, because of the aggregation, the meaning of the composition factors has changed, given that an aggregate type is not a physical entity but rather an abstract production planning category. For example, the composition factor of 1:1 between part 5a and PRODUCT TYPE 1 means that, when at the final assembly stage enough capacity is allocated to produce 1 unit of PRODUCT TYPE 1, in the forging shop enough capacity should be allocated to produce 1 unit of part 5r .

As described earlier, parts 1r through 11r , and 1b through 11b are identical, respectively. Parts 12r to 16r and their correspondents 12b to 16b differ in color, but they do satisfy requirements (1) - (4) . Therefore, a new step of aggregation can be performed and the structure of figure 3.6 results.

For notational simplicity, and because there was absolutely no difference (not even color) between parts 8r and 8b, we used notation (8) rather than "Either 8r or 8b" ; similarly for (9), (10), (11) . Part 12, 13, 14, 15, 16 show differences in color.

Further, the pairs (1r,1b) through (7r,7b) can be aggregated, thus yielding the structure in figure 3.7 .

Let us note that the structure of figure 3.7 is a consequence of the aggregation performed at the final assembly stage only; it has already achieved much in the sense that the number of items to be included into the planning model has been halved.

At this point we will try to find other opportunities for aggregation at other stages.

FABRICATION SHOP (Parts 8, 9, 10, 11)

To systematically approach the inspection of items for aggregation purposes we will summarize all relevant information in table 3.2.

Lines 1, and 3 through 7 in table 3.2 provide the necessary information to determine whether requirements (1) - (4) are fulfilled. The conclusion is that parts 9, 10, 11 do satisfy conditions (1) - (4), while part 8 does not (e.g., the routing of part 8 and its component part 1 differs from the routing of parts 9, 10, 11). Hence, an aggregate part, call it FAB 1 can be defined at the fabrication stage:

$$\begin{matrix}\text{1 unit of} \\ \text{FAB 1}\end{matrix} = \left(\begin{matrix}\text{EITHER} \\ \text{1 unit of 9}\end{matrix} \text{ , } \begin{matrix}\text{OR} \\ \text{1 unit of 10}\end{matrix} \text{ , } \begin{matrix}\text{OR} \\ \text{1 unit of 11}\end{matrix}\right)$$

Part 8 stands by itself.

PLASTICS SHOP (Parts 12, 13, 14, 15, 16)[1]

Table 3.3 concentrates the information necessary to perform the aggregation at the plastics facility. By analyzing the table we find out that:

- if it had not been for different values added, parts 12 and 13 could have been aggregated together;
- Parts 14, 15 and 16 comply with requirements (1) - (4), thus forming an aggregate part called PLAST 1:

---

[1] For simplicity we are saying part 12, but we understand that this is aggregate part (either 12r or 12b); similarly for parts 13 - 16 .

Table 3.2 - Relevant Information After Aggregation at the Final Assembly Stage

line #	Part to be analyzed	Part 8		Part 9		Part 10		Part 11	
1	Shops involved in the production of the analyzed part and its components	Casting	Fabrication	Forging	Fabrication	Forging	Fabrication	Forging	Fabrication
2	What are the parts related to the analyzed part	1	8	3	9	4	10	5	11
3	Lead time (mos.)	2	1	1	1	1	1	1	1
4	Productivity (units/hour)	10	20	5	20	5	20	5	20
5	Value when completely processed ($/unit)	3	12	7	12	7	12	7	12
6	Value added ($/unit)	3	9	7	5	7	5	7	5
7	Assembly or part immediately succeeding the analyzed part; composition factors	17 $a_{8,17} = 1$		17 $a_{9,17} = 1$		17 $a_{10,17} = 1$		17 $a_{11,17} = 1$	

Table 3.3 - Relevant Information for Aggregation at the Plastics Shop Stage

line #	Part to be analyzed	Part 12		Part 13		Part 14	Part 15	Part 16
1	Shops involved in the production of the analyzed part and its components	Forging	Plastics	Forging	Plastics	Plastics	Plastics	Plastics
2	What are the parts related to the analyzed part	6	12	7	13	14	15	16
3	Lead time (mos.)	1	1	1	1	1	1	1
4	Productivity (units/hour)	5	60	5	60	60	60	60
5	Value when completely processed ($/unit)	15	17	7	17	17	17	17
6	Value added ($/unit)	15	2	7	10	17	17	17
7	Assembly or part immediately succeeding the analyzed part; composition factors	17 $a_{12,17} = 1$		17 $a_{13,17} = 1$		17 $a_{14,17} = 1$	17 $a_{15,17} = 1$	17 $a_{16,17} = 1$

$$\text{1 unit of PLAST 1} = \left(\begin{array}{c} \text{EITHER} \\ \text{1 unit of} \\ \text{14r or 14b} \end{array} , \begin{array}{c} \text{OR} \\ \text{1 unit of} \\ \text{15r or 15b} \end{array} , \begin{array}{c} \text{OR} \\ \text{1 unit of} \\ \text{16r or 16b} \end{array} \right)$$

As this point the structure of PRODUCT TYPE 1 looks like in figure 3.8.

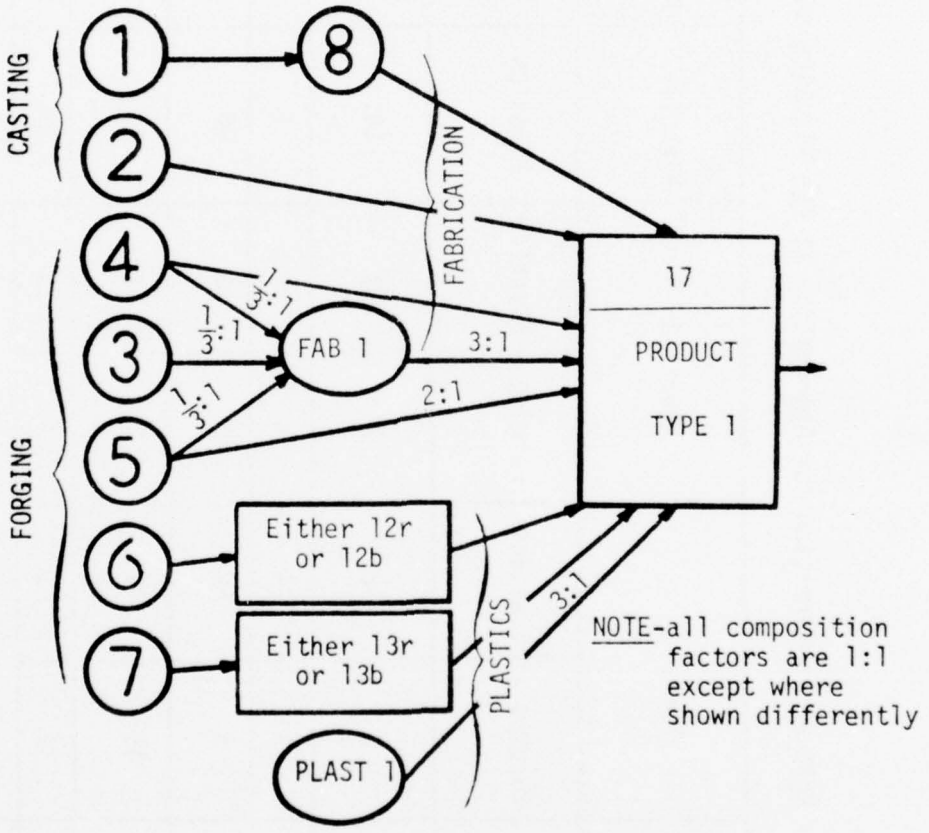


Fig. 3.8- Aggregate structure 4.

composition factors of 1/3:1 have the meaning that, when at the fabrication stage enough capacity is allocated to produce one unit of FAB 1 , in the forging shop 1/3 of the capacity needed to produce one unit of part 3, plus 1/3 of the capacity for one unit of part 4, plus 1/3 of the capacity needed for one unit of part 5 will have to be allocated.

We will look now into possibilities of further aggregating at the casting and forging shop stages.

CASTING SHOP (Parts 1 and 2)

Table 3.4 - Relevant Information

for Aggregation at the Casting Shop Stage

line #	Part to be analyzed	Part 1	Part 2
1	Shops involved in the production of the analyzed part and its components	Casting	Casting
2	What are the parts related to be analyzed part	1	2
3	Lead time (months)	2	2
4	Productivity (units/hour)	10	10
5	Value when completely processed ($/unit)	3	3
6	Value added ($/unit)	3	3
7	Assembly or part immediately succeeding the analyzed part; composition factors	8 $a_{1,8} = 1$	17 $a_{2,17} = 1$

Clearly, parts 1 and 2 have identical production profiles since requirements (1) - (3) are complied with. However, because they go into different next level assemblies (or parts) requirement (4) is violated; therefore, parts 1 and 2 cannot be aggregated together.

FORGING SHOP (Parts 3, 4, 5, 6, 7; relevant information in table 3.5)

With the exception of part 6, all other parts (i.e., 3, 4, 5 and 7) have identical production profiles. However, they do not satisfy requirement (4), that is, they either do not go into the same next level assemblies, or if they do (like parts 4 and 5) they do not have composition factors satisfying the proportionality condition of requirement (4). Consequently, no aggregation can be performed at the forging stage.

The final aggregate structure is presented in figure 3.9 along with the definitions of all aggregate types created.

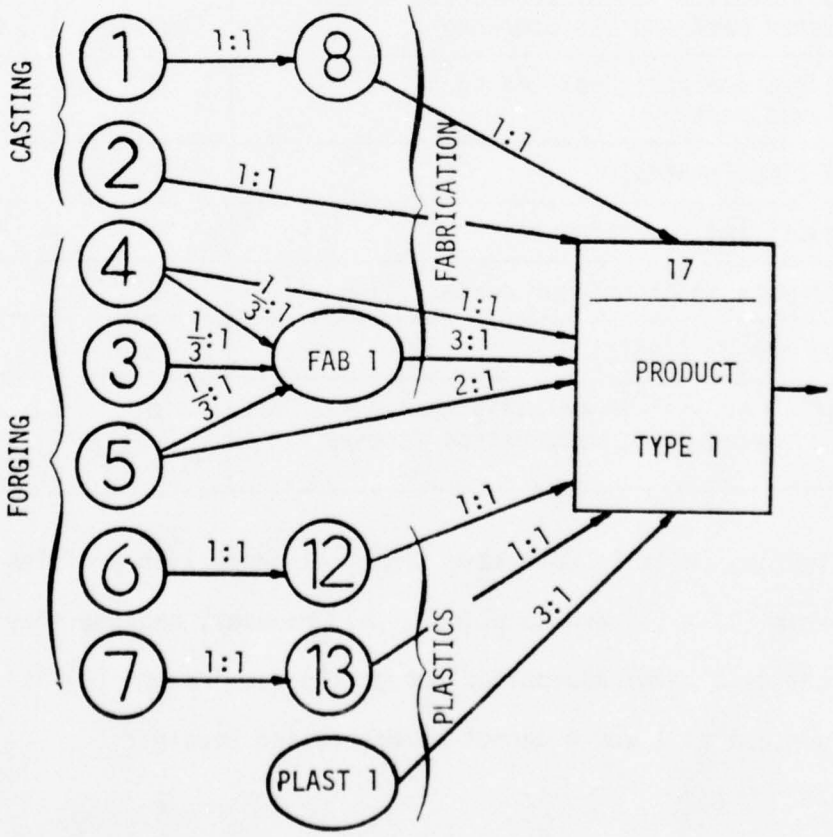


Fig. 3.9- Final aggregate structure.

Table 3.5 - Relevant Information for Aggregation at the Forging Shop Stage

Line #	Part to be analyzed	Part 3	Part 4	Part 5	Part 6	Part 7
1	Shops involved in the production of the analyzed part and its components	Forging	Forging	Forging	Forging	Forging
2	What are the parts related to the analyzed part	3	4	5	6	7
3	Lead time (mos.)	1	1	1	1	1
4	Productivity (unit/hour)	5	5	5	5	5
5	Value when completely processed ($/unit)	7	7	7	15	7
6	Value added ($/unit)	7	7	7	15	7
7	Assembly or part immediately succeeding the analyzed part; composition factors	FAB 1 $a_{3,FAB\ 1} = \frac{1}{3}$	FAB 1, 17 $a_{4,FAB\ 1} = \frac{1}{3}$ $a_{4,17} = 1$	FAB 1, 17 $a_{5,FAB\ 1} = \frac{1}{3}$ $a_{5,17} = 2$	12 $a_{6,12} = 1$	13 $a_{7,13} = 1$

$$\text{FAB 1} = \left(\begin{array}{c} \text{EITHER} \\ \text{1 unit of 9} \end{array} , \begin{array}{c} \text{OR} \\ \text{1 unit of 10} \end{array} , \begin{array}{c} \text{OR} \\ \text{1 unit of 11} \end{array} \right)$$

$$12 = \left(\begin{array}{c} \text{EITHER} \\ \text{1 unit of 12r} \end{array} , \begin{array}{c} \text{OR} \\ \text{1 unit of 12b} \end{array} \right)$$

$$13 = \left(\begin{array}{c} \text{EITHER} \\ \text{1 unit of 13r} \end{array} , \begin{array}{c} \text{OR} \\ \text{1 unit of 13b} \end{array} \right)$$

$$\text{PLAST 1} = \left(\begin{array}{c} \text{EITHER} \\ \text{1 unit of 14r} \\ \text{OR} \\ \text{1 unit of 14b} \end{array} , \begin{array}{c} \text{OR} \\ \text{1 unit of 15r} \\ \text{OR} \\ \text{1 unit of 15b} \end{array} , \begin{array}{c} \text{OR} \\ \text{1 unit of 16r} \\ \text{OR} \\ \text{1 unit of 16b} \end{array} \right)$$

$$\text{PRODUCT TYPE 1} = \left(\begin{array}{c} \text{EITHER} \\ \text{1 unit of 17r} \end{array} , \begin{array}{c} \text{OR} \\ \text{1 unit of 17b} \end{array} \right)$$

Parts 1 through 8 are identical for both products 17r and 17b, thus making unnecessary the definition of aggregate parts.

In terms of the problem size, let us point out that the initial total number of 34 individual items has been reduced, by aggregation, to 13 aggregate types.

3.1.2.3. Aggregation of Individual Facilities (Machines and Work Places) into Aggregate Facilities or Stages.

In the example of section 3.1.2.2 the aggregation has been performed by facilities, starting with the final assembly stage. The facilities we have been referring to were aggregate facilities: the fabrication shop, casting shop, etc., rather than individual machines or work places.

The aggregate stages, once defined, provide the capacity constraints (3.4), (3.5) of the aggregate planning model; they also have an effect

upon the inventory balance equations (3.2), (3.3) in that the number of aggregate types defined in the LP cannot be smaller than the number of facilities (i.e., there has to be at least one aggregate part or product type per facility identified in the aggregate model).

There are both advantages and disadvantages in having many or few facilities to consider in the planning model. If many facilities are defined, they permit a lower degree of aggregation, hence a smaller loss of detailed information; there would also be a higher degree of flexibility in the LP model. On the other hand, the more facilities defined, the shorter the lead times (e.g., the lead time at the grinding machines group might be one week, while the lead time at the fabrication shop level could well be one month). Since the aggregate planning model only handles lead times equal to a time period or a multiple of it, a low degree of aggregation would require many periods in order to cover a given planning horizon. This, in turn, increases the size of the model and the numbers of parameters to be estimated.

In creating aggregate facilities, the managerial aspect has to play a central role. Indeed, the aggregate plan will specify for every facility a set of production and employment levels, that will have to be implemented by managers. Therefore, the aggregate facilities identified in the model have to represent "manageable" units, related to some already existing, "naturally" established facilities, determined by the nature of the production activity and by the existing organizational structure. Also, it should not be forgotten that a number of parameters (e.g., capacity constraints, productivity factors, lead times) have to be estimated

and this task can be rendered easier or more difficult depending on how facilities are defined. As an example, the manager of a fabrication shop might have no problem in estimating that it takes about 4 weeks to get a lot of long arbors through the shop, and about 8 weeks for a batch of small drums for textile machines. However, he might find it difficult to estimate lead times for the group of milling machines if, normally, jobs are released as production orders to the fabrication shop rather than to specific machine centers.

3.1.3 Handling Initial Inventories of Component Parts in the Aggregate Planning Model.

The discussion on the difficulties associated with handling initial inventories of component parts at the aggregate planning level has already been opened in section 3.1.2.1 The fourth aggregation condition stated there, is intended to help us with this problem.

To continue the discussion and propose a solution, consider the following example: two component parts 1 and 2 satisfy the aggregation requirements (1) through (4), and therefore are grouped into an "either 1 or 2" aggregate part, call it A (figure 3.10).

Suppose that the initial inventories are $I_{10} = 10$ units for part 1, and $I_{20} = 0$ for part 2; for simplicity assume zero lead times.

The aggregate planning model recognizes neither part 1 nor part 2; only the aggregate part A is identified.

WHAT SHOULD BE THE INITIAL INVENTORY I_{A0} INPUT TO THE AGGREGATE PLANNING MODEL?

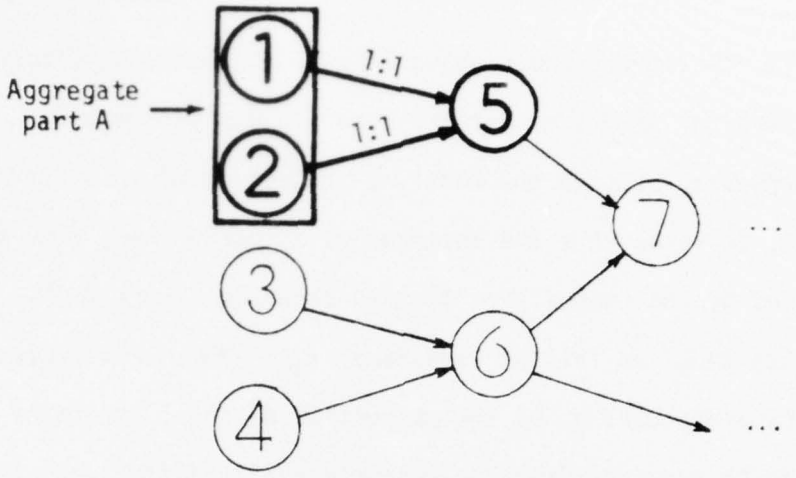


Fig. 3.10- Structure in which parts 1 and 2 can be aggregated together.

We will examine, critically, four strategies that can be used:

a) Simply input to the aggregate model $I_{A0} = I_{10} + I_{20} = 10$.

Depending on the value of x_{51} (production of subassembly 5 in period 1) in the LP optimal solution, the following situations could develop:

- if $1 \leq x_{51} \leq 5$ the model will order no production of A (i.e., $x_{A1} = 0$), hence, nothing will have to be produced of part 2. This is wrong because between 1 and 5 units of part 2 will be needed, while none will be available.
- if $6 \leq x_{51} \leq 9$ a production of 2 to 8 units of A will be decided, which will enable us to produce 2 to 8 units of part 2. This is wrong again, because the amount of parts 2 needed will between 6 (corresponding to $x_{51} = 6$) and 9(when $x_{51} = 9$) .
- If $x_{31} \geq 10$ the model will decide correctly.

Clearly, the errors shown above resulted because the model has been using inventory of part 1 to satisfy requirements for part 2.

b) Schedule, outside the model, a production of 10 units of part 2 in period 1, to re-balance the inventories of parts 1 and 2 to the 1:1 ratio implied by the composition factors shown in figure 3.10. Input to the aggregate model an initial inventory $I_{A0} = 20$. Obviously, the amount of capacity used up by the production of the 10 units of part 2 will have to be deducted from the resource availability input to the LP. With this strategy, the model will produce the correct decisions; however, our decision might have been not too good. Indeed, suppose that in the optimal solution $x_{51} < 10$; in this case less than 10 sets of parts are needed right away, and our decision to make 10 sets available diverted a certain capacity from other needs and unnecessarily increased the inventory held in stock.

c) Look at the previous aggregate solution (obtained when the aggregate planning model was solved one period earlier) and act based on the belief that the new solution will differ only marginally from the old one; thus, the old requirements for parts 1 and 2 will be considered as tentatively valid (the old solution will be denoted by the superscript "old").

Suppose:

$$x_{51}^{old} = 5 \; ; \quad x_{52}^{old} = 4 \; ; \quad x_{53}^{old} = 5$$

where period 1 is the upcoming period, period 2 is the next period, etc. Then schedule outside the model a production of 5 units of part 2 in

period 1, and 4 units in period 2, and input to the model the re-balanced inventories of parts: 10 units of aggregate part A available in period 1, 8 units available in period 2, and 1 unit in period 3. Technically this can be achieved by making $I_{A0} = 10$, and by modifying the inventory balance equations for part A in periods 2 and 3 as follows:

$$I_{A1} + 8 + x_{A2} - I_{A2} = 2x_{52}$$

$$I_{A2} + 1 + x_{A3} - I_{A3} = 2x_{53}$$

Let us remark that neither (b) nor (c) can lead to an infeasible disaggregation as (a) could have done. Notice, also, that approach (c) provides a lower inventory holding cost and a more reasonable utilization of resources than (b).

d) Proceed similarly to strategy (c); after it has been decided to produce 5 units of part 2 in period 1, and 4 units in period 2, analyze the possibility of moving to period 1 the 4 units scheduled for period 2. The analysis is based on the old optimal shadow prices obtained when the aggregate model was solved a period earlier. If the sum of the shadow price of the resource constraint for period 1 plus the cost of holding part 2 for one period is lower than the shadow price associated with the resource constraint for period 2, it pays to produce the 4 units in period 1.

This approach represents an attempt to make, outside the model, decision as close to LP optimality as possible.[1] Notice that by moving production of part 2 to an earlier period, the feasibility of disaggregation is in no way affected.

3.1.4. A Two-Level Aggregation for Very Large Operations

In very large operations, such as automobile manufacturing, major subassemblies or groups of parts are supplied by quasi-independent facilities or vendors, whose capacities have only minor or no interactions with each other. In such a case, the aggregate planning phase would consist of two hierarchical levels.

The Higher Aggregate Level is driven by the demand forecasts for the end products. Its purpose is to coordinate production and supply of major subassemblies and sets of parts. The output of this aggregate level would be a feasible plan of requirements for the major subassemblies and groups of parts.

For instance, suppose that in an automobile manufacturing operation, the major subassemblies and sets of parts are:

1. engine (mechanical components)
2. components of the electric system

---

(1) We are saying "as close as possible" because the shadow prices are valid only within a limited range of the right hand side values; therefore, tying up some of the available capacity might lead to a change in shadow prices.

3. body parts

4. transmission, braking, and steering systems

5. rubber parts 1

6. plastic and rubber parts 2

7. tires.

For every group of parts above, a lead time[1] and cost parameters will have to be assessed.

The multi-stage structure corresponding to this level of aggregation is shown in figure 3.11.

The result of this level will be an assembly plan (or master schedule) for the end product,and a feasible schedule of requirements for all major subassemblies and sets of parts.

The Lower Aggregate Level would be a collection of multi-stage aggregate planning models of the type (3.1) - (3.6). There would be such a model for every major subassembly and set of parts. Every model will be driven by the schedule of requirements generated by the higher aggregate level. Of course, no backorders may be incorporated in the planning models at this lower level because they would jeorpadize the entire coordination of production and supply achieved at the higher aggregate level.

---

[1] Lead times estimates will be assessed on statistical grounds, as an answer to the question: "What has been in the past the normal time elapsed between the placement of an order for, say, a batch of tires and the delivery of that order ?"

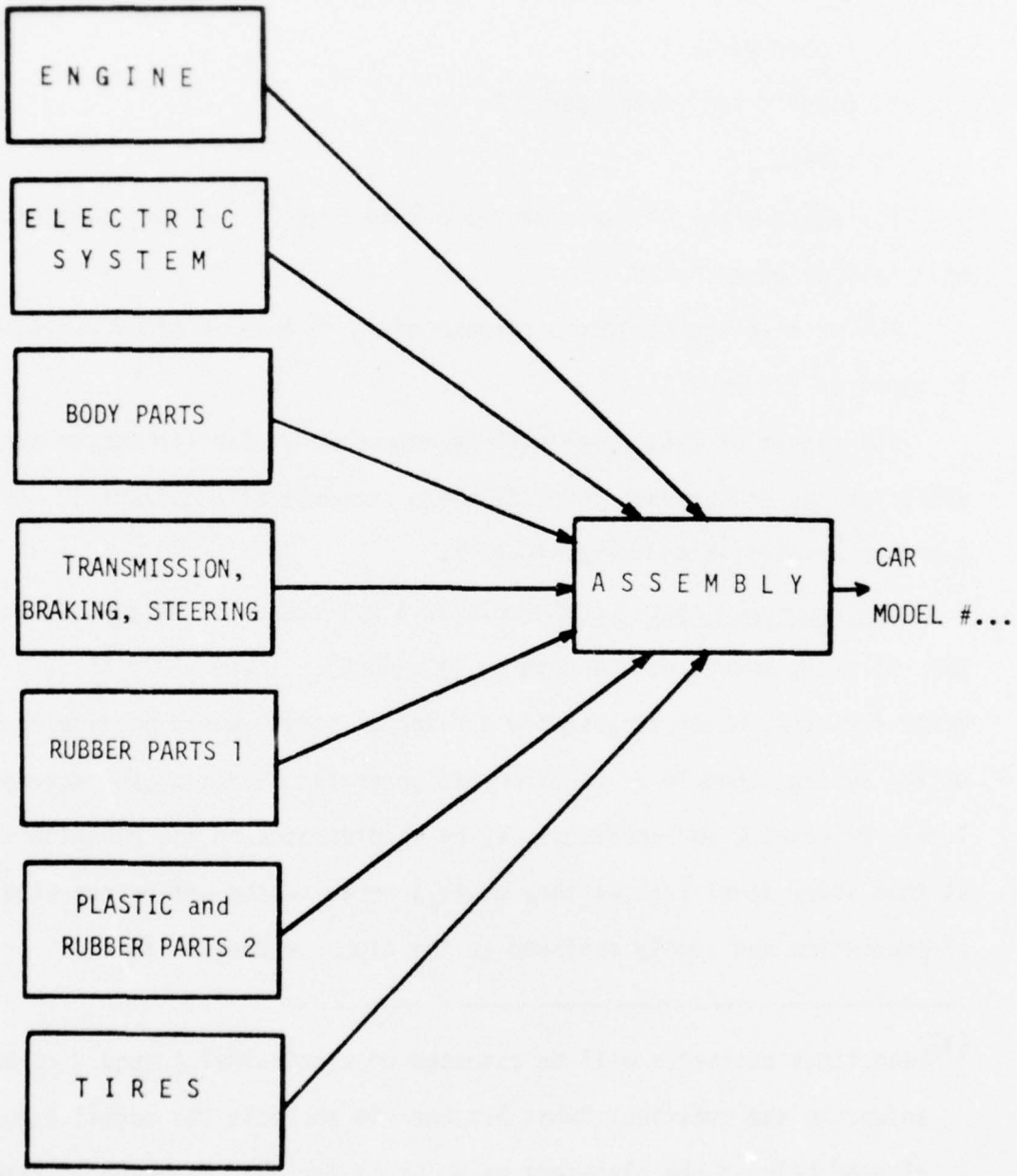


Fig. 3.11- Structure at the higher aggregate level.

It is apparent from this two-level scheme that the purpose of the higher aggregate level is to decompose what would be an enormous multi-stage problem into several smaller multi-stage problems.

3.2. Issues of Disaggregation

Once the solution of the aggregate planning model (call it the aggregate plan) has been obtained, there are still two problems to be resolved before a production plan that can be implemented is obtained:

- the aggregate plan specifies production levels for the aggregate types, and these have no physical correspondents;

- the cost structure, under which the optimal aggregate plan was computed, was incomplete since it did not contain setup costs.

It is the role of the disaggregation to look into these issues and to come up with implementable production and inventory levels. The disaggregation will have to be performed in such a way as to observe the capacity allocation specified in the aggregate plan. This means that

the sum of the production levels, for any given period t , of all the individual items included in an aggregate type must be equal to the amount of production of that aggregate type, called for by the aggregate plan for period t . In this way the accumulation of seasonal stocks of parts and finished goods is ensured, and the detailed plan (obtained by disaggregation) will not perturb the optimality of the aggregate solution.

From what we have just said it follows that the disaggregation can take place for each aggregate type independently, since the capacity and time phasing interactions have been all considered in drawing the aggregate plan. This is very important because of two reasons:

a) The independent disaggregation for aggregate types also implies that the disaggregation takes place within each stage (or facility) without too much interaction with other stages; this is managerially appealing since it eliminates a lot of communication and correlation problems. Ideally, if the managers of all the facilities involved are provided with knowledge of the aggregate plan and a set of feasible disaggregation rules, and if the aggregate production levels are implemented as specified, the production process should be able to proceed ryhthmically[1] and smoothly.

---

[1] Of course, because of uncertain and/or unexpected factors perturbations will occur all the time. However, this does not invalidate the the drive towards optimality; rather, it requires that the planning technique be robust enough to avoid destabilization. For the single

(continued on next page)

b) The basic conceptual approach to disaggregation can, then, be borrowed from single stage systems (Hax et. al. [46]).

3.2.1. A Conceptual Approach to Disaggregation

Consider an aggregate part for instance, that includes a number of individual parts. It is possible that when some piece of equipment is set up to produce one of these parts, other parts in the type can also be produced with only a minor additional setup or with no additional set-up at all. We will, therefore, define a family to be a group of items which share a common manufacturing setup (Hax and Meal [44]). The items in the group are part of the same aggregate type. Economies of scale are accomplished by replenishing jointly items belonging to the same family.

Thus, if we are to refer back to the example of section 3.1.2.2, the aggregate structure of figure 3.9 will have to be investigated for the possible creation of families within each aggregate type as follows:

- PRODUCT TYPE 1 - since the two end products have no structural differences, both the red and the blue product can be assembled with the same setup of the assembly facility; hence, PRODUCT TYPE 1 contains only one family, call it FAM 17.

---

stage case the hierarchical approach has proven to be very robust (Bitran and Hax [9]) even under forecast errors of up to ±30% . It is based on these encouraging results that we advocate as worthwhile to investigate the extension of this approach to multi-stage systems.

- Aggregate parts 1 through 8 - each contains two identical parts, distinguished only by the fact that they go into two different end products: 17r and 17b . Each aggregate part 1 through 8 will consist of one family; hence, we will have FAM 1 through FAM 8 .
- FAB 1 contains three component parts: 9, 10 and 11 . Suppose part 9 and 10 share the same setup. Two families will be, then, formed within FAB 1: FABFAM 1 and FABFAM 2 .
- Since the mold injection and the plastic coating machines need a setup every time the color of plastic is changed, only component parts that have the same color can possibly be included into the same family. Therefore, we will have FAM 12r , FAM 12b , FAM 13r, FAM 13b , PLASTFAMr (including 14r, 15r, 16r), and PLASTFAMb (including 14b, 15b, 16b) .

After this analysis, the family structure looks like in figure 3.12. we will discuss, in the next section, the issue of how to determine the composition factors for the family structure.

The disaggregation will proceed in two steps:

- the family disaggregation, by which the amount of production planned for the aggregate type will be broken down into run quantities for the member families;
- the item disaggregation, by which the amount of production set for the family will be distributed to individual member items. As it will be seen later, if the lead time is non-zero the disaggregation will have to take place for a number of periods ahead, although

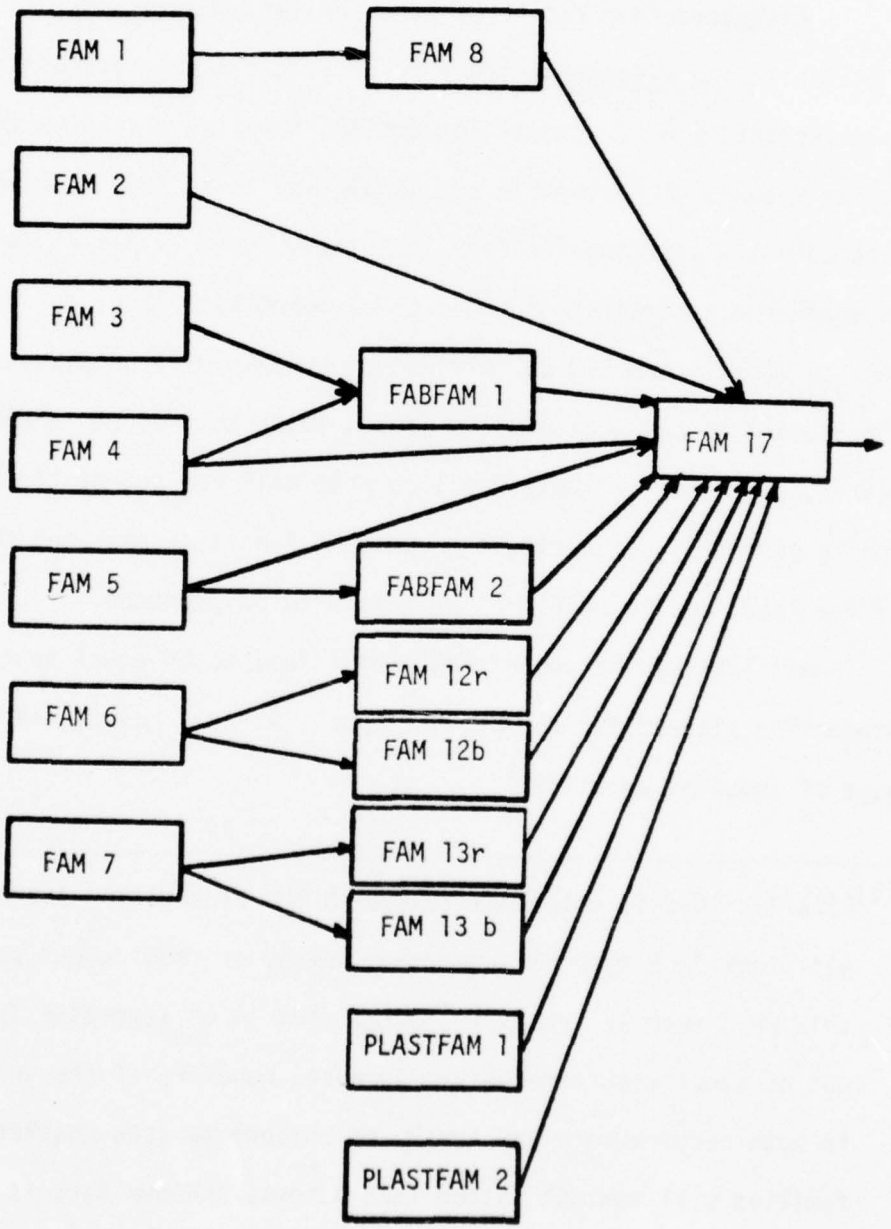


Fig. 3.12- Family structure of product type 17.

only the production plans for the upcoming period will be implemented. This is different from the single stage case, where the disaggregation had to be done for the upcoming period only.

The family disaggregation - in order for the setup costs to be taken into account, the run quantities for the families will have to be related to the setups. A reasonable and simple way to do this, and which has also been the starting strategy in the work done by Hax et. al. [46], is to computeunconstrained economic order quantities (EOQ) for every family (how to compute the EOQ in multi-stage systems will be discussed in Chapter 5), and to see what happens if one tries to produce in EOQ's. Of course, only that or those families that will run out of stock in the upcoming period[1] (or during the upcoming lead time plus one review period, if the lead time is non-zero) will have to be produced.

Most likely, the sum of EOQ's will fail to be equal to the amount of production planned for the product type. At this point, there are two ways of inducing equality:

---

[1] Clearly, this is not consistent with the aggregation assumption that all items in a type have the same demand or requirement patterns; if this were exactly true all the families in an aggregate type would run out of stock simultaneously. Because, however, of the uncertainties in both requirements and supply of components (see Chapter 6) some families will run out faster than others, and our task is to try to prevent stockout situations from developing.

- the Hax-Meal [44] disaggregation technique by which the unconstrained EOQ's are adjusted (increased or decreased) so as to ensure that the capacity allocated to the aggregate type is neither underutilized, nor exceeded;
- the Winters [96] approach by which the unconstrained EOQ's are left unchanged; all the families in the aggregate type are ranked by the times they will run out of stock. Starting with the family which will run out at the earliest time, EOQ's are accumulated until the type's allotted capacity is exhausted.

We will not go into further details here, since the two techniques are fully documented in the indicated references. However, let us mention that Winter's method may lead to exhausting the capacity allocated to the type before we can schedule in production all the families that will run out of stock in the upcoming period (or lead time plus review period). We can see three ways to resolve this situation:

- allocate for each family which runs out a minimum amount of production rather than the EOQ (the minimum production can be the quantity just sufficient to make the inventory last over the lead time plus the review period);
- switch to the Hax-Meal procedure;
- re-solve the aggregate plan under the constraint that a certain amount of capacity be transferred to the aggregate type in need from other types that ulitize the same resources and have capacity in excess of the immediate requirements.

Other family disaggregation techniques, for instance the equalization of run out times (Hax et. al. [46]), can be found. In order to evaluate their efficiency, they have to be tested by simulation under a variety of conditions involving design parameters such as: demand patterns for the end products and for spare parts, forecast accuracy, structure of the end products, setup costs, capacity restrictions, length of the planning horizon. All these are likely to influence the performance of the planning system, and only a thorough statistical analysis of the experimental data can answer questions about the effectiveness and comparative efficiency of the various disaggregation techniques.

The item disaggregation seems to be less involved because, for the upcoming period, all the costs have already been taken into account in the previous steps, so that any feasible disaggregation of a family run quantity will have the same total cost. However, it is not to be forgotten that all items in a family share the major setup for the family; therefore, even if only one item of the family needs to be produced, the major setup has to take place. Consequently it is reasonable to distribute the family run quantity among its member items such as to have them all trigger simultaneously; in this way, every time the family setup is incurred all member items will be produced. For obvious reasons, this technique is called the equalization of runout times.

The calculation of the item run quantities is simple, that is, they will have to be set in such a way that the items runout time equal the runout time of the family (Bitran and Hax [9], Hax and Meal [44]). One aspect is, however, specific to the multi-stage structure: while in the

single stage case the runout times can be computed using the demand forecasts made for the items in question, in the multi-stage case only the end product disaggregation will use demand forecasts. The disaggregation at the component levels will have to be based on the requirements for parts generated by exploding[1] the production schedules of higher level parts and assemblies.

For instance, if some component part has a lead time of 2 months, we will have to know the projected requirements for that part for at least 3 periods in advance (lead time plus review period); hence, at the immediate successor stages, the disaggregation of the aggregate plan will have to be carried out for 3 months ahead. If the initial inventory of the component part in our example is large, and thus expected to cover more than 3 periods, we might still want to know for how many more months it is expected to last; to answer this question, the disaggregation at the successor stages would have to go beyond the initial 3 months.

Thus, as already mentioned before, when lead times are non-zero, plans have to be disaggregated for a number of periods ahead, although only the plans for the upcoming periods will be implemented before the model is updated and re-run.

There is another issue we should pay attention to in the course of disaggregation, namely, bounds on the EOQ's. Sometimes unconstrained lot

---

[1] The term "explosion" has been borrowed from MRP nomenclature (Orlicky [77], p. 56).

sizing yields either too small or too large EOQ's. For instance we might not want to order in such small sizes so that two or more setups would take place for the same item during one month. Also, as requirements projections cannot be trusted over very long periods of time, we should limit our production commitments to a maximum order quantity of, say, at most one year. Of course, these limits are management decisions and can be changed whenever deemed necessary. Also, when items with limited sales seasons are planned, overstock limits have to be computed and enforced. A more detailed exposition of the bounds issue is presented by Hax [42], and for overstock limits Crowston et. al. [24], Bryan et. al. [14].

3.2.2. Composition Factors in Family Structures

To make the issue clear consider a very simple example: there are two products (figure 3.13) one assembled from red plastic parts, the other assembled from blue plastic parts. Except for the color, the two products are otherwise identical, and display similar demand patterns.

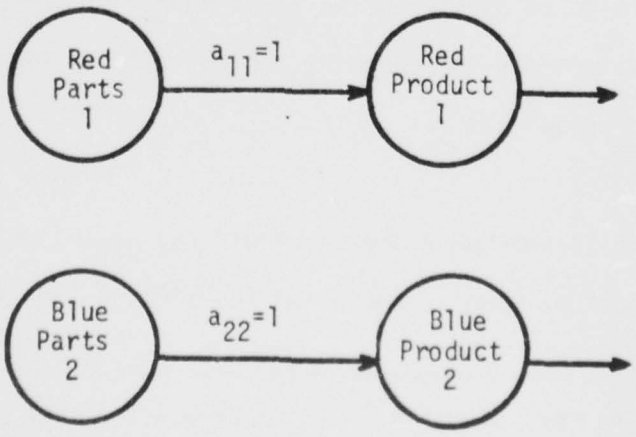


Fig. 3.13- Two stage structures.

Assume that, although the parts have similar production profiles, the machine that makes the parts has to be set up every time the color is changed; however, when the assembly facility is set up either of the end products can be assembled. Hence, the family structure will be as shown in figure 3.14.

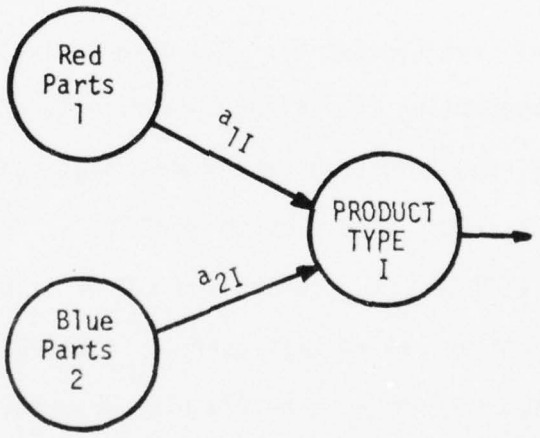


Fig. 3.14- Family structure of products in fig. 3.13.

What should be the composition factors a_{1I} , a_{2I} ?

Clearly, they cannot be 1:1 (as in the detailed structure of figure 3.13) because, suppose the demand for Red Product 1 is $d_1 = 1000$ and the demand for Red Product 2 is $d_2 = 2000$; the resulting demand for the Product Type I is $d_I = 3000$. Then, composition factors of 1:1 would imply that both the Red Parts 1 and the Blue Parts 2 are required in amounts of 3000 each, which is obviously wrong.

Therefore, for the purpose of computing EOQ'a for the multi-stage family structure we will determine the composition factors by weighting the original factors by the demands. Thus we obtain:

$$a_{1I} = a_{11} \frac{d_1}{d_1+d_2} = \frac{1}{3} \quad ; \quad a_{2I} = a_{22} \frac{d_2}{d_1+d_2} = \frac{2}{3}$$

There are several reasons to justify this approach:

- the composition factors computed as shown above will yield the correct demand requirements for the component parts;
- the EOQ computations are conducted under deterministic demand assumptions (see Chapter 5). The deterministic demand, coupled with the assumption of similar demand patterns and with the equalization of runout times at the item disaggregation level, will ensure that, every time a batch of PRODUCT TYPE 1 is assembled, 1/3 of it will be red products and 2/3 of it blue products. This is equivalent to saying that every unit of PRODUCT TYPE 1 is composed of 1/3 red product and 2/3 blue product, which is in agreement with the composition factors computed above.

If the products in our example had three stages instead of two (figure 3.15), and if Castings 1 and Castings 2 had identical production

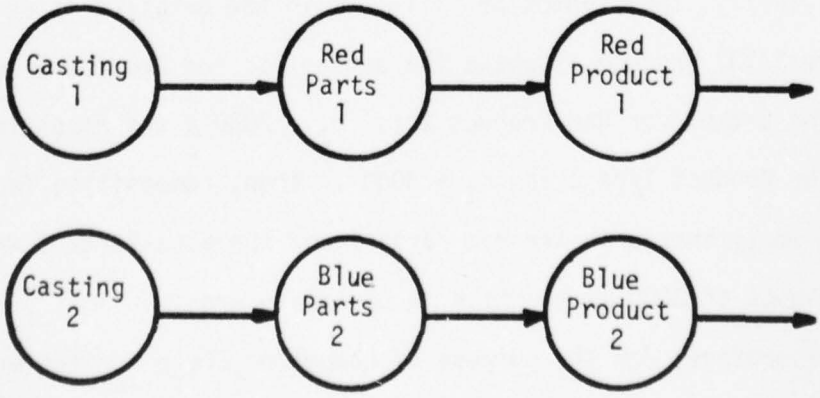


Fig. 3.15- Three stage structures.

profiles and used the same setup, the family structure would be the one depicted in figure 3.16. The composition of factors shown have been obtained by a similar line of reasoning as discussed above.

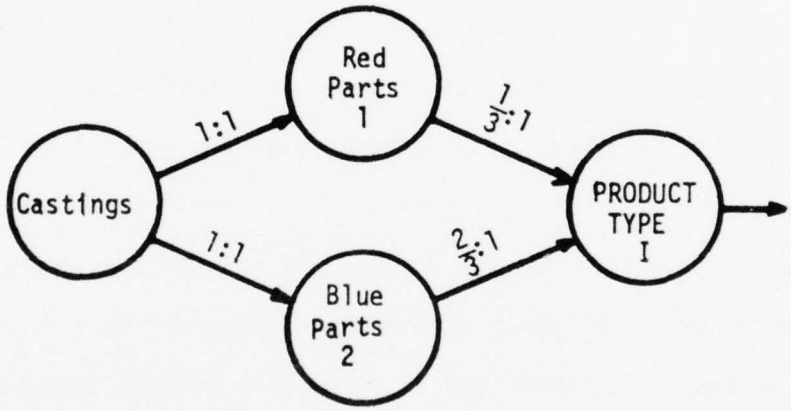


Fig. 3.16- Family structure of products in fig. 3.15.

CHAPTER 4 - THE AGGREGATE PLANNING MODEL

In section 3.1 the aggregate planning model, of the linear programming type, has been introduced, and we have already seen that the need to reduce its size was one reason behind the aggregation process. However, as there are limits on how far the aggregation can go without introducing unacceptable distortions into the solution, the aggregate model might still be too large to solve because of the excessively large number of rows.

Therefore, we are considering in this chapter a re-formulation of the problem, by which the LP with a large number of rows is converted into an equivalent LP with less rows but many columns instead. The advantage is that, although both LP's involve large scale optimization, in this case we can better handle many columns rather than many rows. For simplicity of expression, we will use throughout chapter 4 the terms "end product" and "part" interchangeably with "product type" and "aggregate part", respectively. No confusion can arise since only aggregate planning issues will be discussed.

4.1 A Simple Example and the Elementary Problem

For ease of further reference let us call model (3.1) - (3.6) of section 3.1, with possible extensions to accommodate variable work force, multiple resources, etc., formulation F1 . The new type of formulation, that we are going to derive, will be referred to as F2.

Consider a simple example of an aggregate structure: a product type 3, assembled from two aggregate parts, 1 and 2 (figure 4.1).

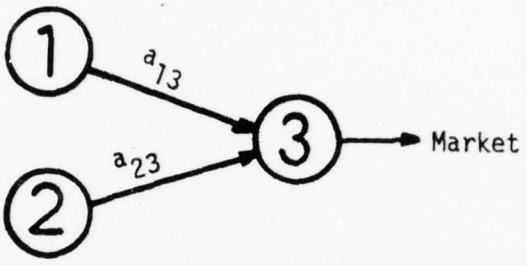


Fig. 4.1- A simple aggregate structure.

In addition, assume that there is a demand of one unit in period p for product 3, and no demand at all in any other period of the planning horizon. No setups are considered; there are no initial inventories, no lead times, and capacity is unlimited (notice that, because we are considering the uncapacitated case, with no setups, there is no eed to bother with multiple product types, for the moment).

We will call this example the elementary problem due to the demand of only one unit in only one period. The corresponding Fl formulation is:

$$\min z = \sum_{s=1}^{3} \sum_{t=1}^{p} (v_{st}x_{st} + h_{st}I_{st}) \tag{4.1}$$

s.t.

$$\left.\begin{aligned} I_{1,t-1} + x_{1t} - I_{1t} - a_{13}x_{3t} &= 0 \qquad (4.2)\\ I_{2,t-1} + x_{2t} - I_{2t} - a_{23}x_{3t} &= 0 \qquad (4.3) \end{aligned}\right\} t=1,\ldots,p$$

$$I_{3,t-1} + x_{3t} - I_{3t} = \begin{cases} 0 \text{ if } 1 \leq t < p \\ 1 \text{ if } t = p \end{cases} \tag{4.4}$$

$$\text{Nonnegativity} \tag{4.5}$$

where the notations of section 1.1 have been used.

We have, by assumption, $I_{10} = I_{20} = I_{30} = 0$; also, we will not include the variables I_{1p} , I_{2p} , I_{3p} in the formulation since the minimization would any way drive them to zero.

Formulation (4.1) - (4.5) is, in matrix notation, of the following form:

$$\min z = cX \tag{4.6}$$

s.t.

$$AX = b \tag{4.7}$$

$$X \geq 0 \tag{4.8}$$

Matrix A is Leontief (Dantzig [27], [28]; Veinott [91]) since it has exactly one positive element in each column and there is a nonnegative X for which Ax is strictly positive. Indeed, if in (4.1) - (4.5) there had been positive independent demands at all stages in all periods 1 through p, a nonnegative vector of productions and inventories could have been certainly found, such as to satisfy all demands.

The optimal solution to the Leontief substitution system (4.1) - (4.5) wil have the following important property (Veinott [92], Corollary 2):

$$I_{s,t-1}\, x_{st} = 0 \tag{4.9}$$

(4.9) defines what is known as the extreme flow property, and we will call the solution to (4.1) - (4.5) an extreme flow elementary production schedule. The extreme points of $Ax = b$ are extreme flows. Hence, an alternative way of solving the elementary problem, rather than using the simplex method, clearly emerges: provide a complete list of extreme flow elementary schedules, compute their costs, and choose the cheapest one as the optimal solution. Formally we write this problem as follows:

$$\min z = \sum_{h_p \in J_p} C_{j_p} \delta_{j_p} \tag{4.10}$$

s.t.

$$\sum_{j_p \in J_p} \delta_{j_p} = 1 \tag{4.11}$$

$$\delta_{j_p} = \begin{cases} 1 \text{ if schedule } j_p \text{ is chosen} \\ 0 \text{ otherwise} \end{cases} \tag{4.12}$$

where J_p is the set of extreme flow elementary schedules for the given problem, and C_{j_p} is the cost of the j_p - th schedule. What an elementary schedule j_p does, is to specify the production plan for the manufacturing of the one unit of end product demanded in period p ; thus, it will show when the parts should be made and when the assembly should taken place. The cost C_{j_p} of such a plan can be immediately calculated, given that the variable production costs and holding costs are all known.

Notice that although (4.10) - (4.12) is an integer program, it can be solved as a LP by the simplex method. But what is more important is that, by the new formulation, we have achieved our goal of turning the F1 formulation of the elementary problem, that had $3p$ rows, into an equivalent program with fewer rows (one row in this case) but with a larger number of columns.

If the demand in period p is $d_p > 1$, and $d_t = 0$ for $1 \leq t < p$ the problem decomposes into d_t elementary problems; this is true because, given the uncapacitated and linear nature of the problem, a production plan is optimal when every unit demanded is produced optimally.

When we "consolidate" the d_t elementary solutions into the optimal solution to (4.10) - (4.12) we need a new variable Θ_{j_p} that will indicate the number of times the extreme flow elementary schedule j_p is being used in the optimal solution. For any given $j_p \in J_p$, Θ_{j_p} represents the sum of the corresponding δ_{j_p}'s over the solutions to the d_p elementary problems. The formulation becomes then:

$$\min z = \sum_{j_p \in J_p} C_{j_p} \Theta_{j_p} \tag{4.13}$$

s.t.

$$\sum_{j_p \in J_p} \Theta_{j_p} = d_p \tag{4.14}$$

$$\Theta_{j_p} \geq 0 \quad , \quad \text{all } j_p \in J_p \tag{4.15}$$

Clearly, as the problem is uncapacitated only one of the schedules in J_p will be used to produce all d_p units required in period p; thus, all Θ_{j_p}'s in the solution of (4.13) - (4.15) will be zero except for one Θ_{j_p} which will take on a value of d_p.

The next "complication" is when we have N end products or product types and positive demands in every period over a T period planning

horizon. Again, based on the linearity of the objective function and of the constraints set, the uncapacitated multi-product problem decomposes into $\sum_i \sum_t d_{it}$ elementary problems, and, thus, the F2 formulation becomes:

$$\min z = \sum_{i=1}^{} \sum_{t=1}^{} \sum_{j_t \in J_t^i} C_{ij_t} \Theta_{ij_t} \tag{4.16}$$

s.t.

$$\sum_{j_t \in J_t^i} \Theta_{ij_t} = d_{it} , \quad \begin{cases} i = 1,\ldots,N \\ t = 1,\ldots,T \end{cases} \tag{4.17}$$

$$\Theta_{ij_t} \geq 0 , \quad \begin{cases} i = 1,\ldots,N \\ t = 1,\ldots,T \\ j_t \in J_t \end{cases} \tag{4.18}$$

where J_t^i is the set of elementary extreme flow schedules capable of serving one unit of demand for product i $(i = 1,\ldots,N)$ in period t $(t = 1,\ldots,T)$. The pair (ij_t) denotes the j-th schedule in J_t^i.

Let us recall that there are, in total, S facilities (or stages); stage s produces N_s items. Two or more end products can have common parts, and thus compete for resources at various facilities (so far, capacity has been assumed unlimited, but we will approach the capacitated case later). The new formulations, F2, do not show facilities and items explicitly; all time-phasing relationships and technological constraints are implicit in the definitions of elementary production schedules. For

this reason only end products have to be identified; they have been denoted by $i = 1,\ldots,N$, without subscripting for stage.

Example - suppose there are 7 facilities, where each of them manufactures or assembles one item (figure 4.2).

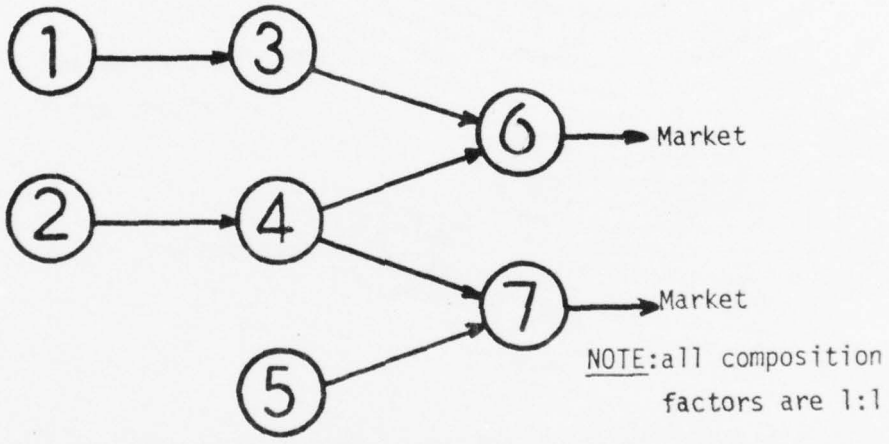


Fig. 4.2- Two end products sharing one common part.

Given that the problem is uncapacitated and linear, each end product can be treated independently (figure 4.3).

Suppose the lead times are all equal to 1 time period, and there is only one unit of demand for each of the end products in period 5 . Figure 4.4 shows one possible elementary production schedule for each product, in the form of Gantt charts. The heavy segments indicate the period in which the activity (production or assembly) takes place.

There is still an alternative(tabular)way to represent these schedules, which will also be used in later sections (figure 4.5). In this representation, number 1 put in a box shows the period in which production starts at that particular stage. An empty box shows that no

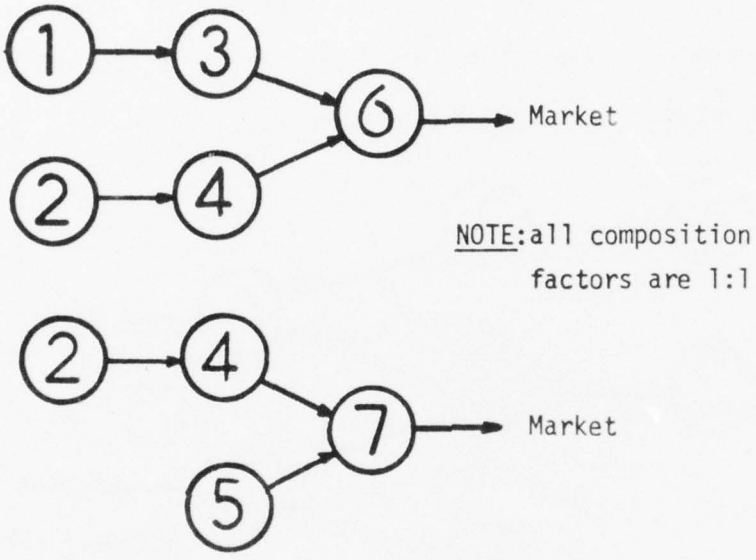


Fig. 4.3- The two products of fig. 4.2, decoupled.

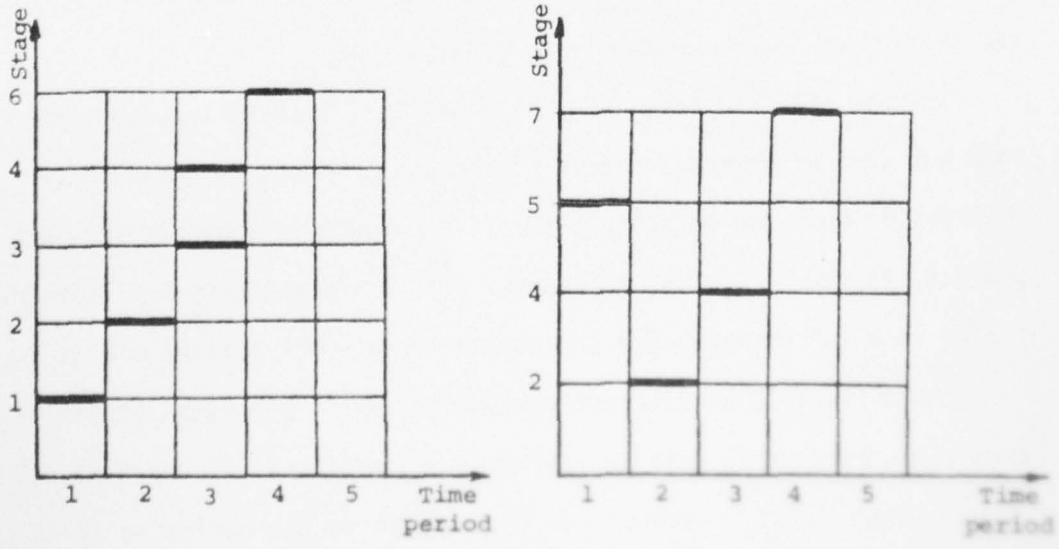
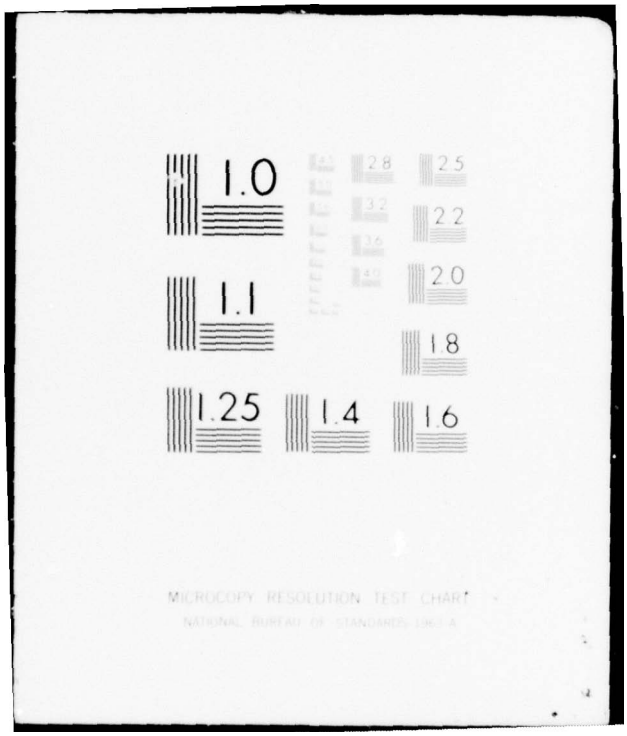


Fig. 4.4- Feasible elementary production schedules for the products of fig. 4.3.

AD-A042 559 MASSACHUSETTS INST OF TECH CAMBRIDGE OPERATIONS RESE--ETC F/G 5/3
ISSUES OF HIERARCHICAL PLANNING IN MULTI-STAGE PRODUCTION SYSTE--ETC(U)
JUL 77 D CANDEA N00014-75-C-0556
UNCLASSIFIED TR-134 NL

2 OF 3
AD A042559




1.0
1.1
1.25
1.4
1.6
2.8
2.5
3.2
2.2
3.6
2.0
4.0
1.8
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

Stage \ Per.	1	2	3	4	5
1	1				
2		1			
3			1		
4			1		
6				1	
Demand	0	0	0	0	1

a)

Stage \ Per.	1	2	3	4	5
2		1			
4			1		
5	1				
7				1	
Demand	0	0	0	0	1

b)

Fig. 4.5- Tabular representation of the feasible elementary production schedules of fig. 4.4.

production takes place.

It is obvious that in our example there are a number of other feasible elementary production schedules. For the schedules shown in figure 4.4 and 4.5 it is easy to compute the costs. Suppose both schedules are number one in their respective sets; the two end products will be called product 6 and 7; the notations will be those used in problem (4.1) - (4.5). The costs, then, are:

$$C_{61_5} = v_{11} + h_{11} + v_{22} + v_{33} + v_{43} + v_{64} \tag{4.19}$$

$$C_{71_5} = v_{22} + v_{43} + v_{51} + h_{51} + h_{52} + h_{53} + v_{74} \tag{4.20}$$

* * * *

Certainly, the number of columns in the LP formulation (4.16) - (4.18) is extremely large. Therefore, we suggest using a column generation technique (Bradley et. al. [10], chapter 12) in dealing with this problem; details on this issue will be given in a later section.

4.2. Elementary vs. Non-Elementary Production Schedules

We realize that problem (4.1) - (4.5) can be re-formulated in a manner similar to (4.16) - (4.18), but using another type of schedules different from the elementary schedules. To see the point, consider again the product of figure 4.1, with composition factors $a_{13} = 1$, $a_{23} = 1$. Assume, for simplicity, zero lead times and a planning horizon of 3 periods, with one unit of the end product demanded in each period. Instead of utilizing three sets of elementary schedules, having each set serve demand in only one period, we can define a set of non-elementary schedules aimed at satisfying demand over the entire planning horizon. To limit the number of the non-elementary schedules (their total number would be 65 for this problem) we will only consider those that satisfy the extreme flow property (4.9). There are 28 such schedules and they are all listed below; the numbers in the boxes show how many units of the part or assembly, made at that stage, are produced in that particular period.

Schedule 1

Per. / Stage	1	2	3
1	1	1	1
2	1	1	1
3	1	1	1
Demand	1	1	1

Schedule 2

Per. / Stage	1	2	3
1	1	2	
2	1	1	1
3	1	1	1
Demand	1	1	1

Schedule 3

Per. / Stage	1	2	3
1	1	1	1
2	1	2	
3	1	1	1
Demand	1	1	1

Schedule 4

Per. / Stage	1	2	3
1	1	2	
2	1	2	
3	1	1	1
Demand	1	1	1

Schedule 5

Per. / Stage	1	2	3
1	1	2	
2	1	2	
3	1	2	
Demand	1	1	1

Schedule 6

Per. / Stage	1	2	3
1	2		1
2	1	1	1
3	1	1	1
Demand	1	1	1

Schedule 7

Per. / Stage	1	2	3
1	3		
2	1	1	1
3	1	1	1
Demand	1	1	1

Schedule 8

Per. / Stage	1	2	3
1	2		1
2	1	2	
3	1	1	1
Demand	1	1	1

Schedule 9

Per. / Stage	1	2	3
1	3		
2	1	2	
3	1	1	1
Demand	1	1	1

Schedule 10

Per. / Stage	1	2	3
1	3		
2	1	2	
3	1	2	
Demand	1	1	1

Schedule 11

Per. / Stage	1	2	3
1	1	1	1
2	2		1
3	1	1	1
Demand	1	1	1

Schedule 12

Per. / Stage	1	2	3
1	1	2	
2	2		1
3	1	1	1
Demand	1	1	1

Schedule 13

Per. / Stage	1	2	3
1	1	1	1
2	3		
3	1	1	1
Demand	1	1	1

Schedule 14

Per. / Stage	1	2	3
1	1	2	
2	3		
3	1	1	1
Demand	1	1	1

Schedule 15

Per. / Stage	1	2	3
1	1	2	
2	3		
3	1	2	
Demand	1	1	1

Schedule 16

Per. / Stage	1	2	3
1	2		1
2	2		1
3	1	1	1
Demand	1	1	1

Schedule 17

Per. / Stage	1	2	3
1	3		
2	2		1
3	1	1	1
Demand	1	1	1

Schedule 18

Per / Stage	1	2	3
1	2		1
2	3		
3	1	1	1
Demand	1	1	1

Schedule 19

Per. / Stage	1	2	3
1	3		
2	3		
3	1	2	
Demand	1	1	1

Schedule 20

Per. / Stage	1	2	3
1	2		1
2	2		1
3	2		1
Demand	1	1	1

Schedule 21

Per. / Stage	1	2	3
1	2	1	
2	2		1
3	2		1
Demand	1	1	1

Schedule 22

Per. / Stage	1	2	3
1	3		
2	2		1
3	2		1
Demand	1	1	1

Schedule 23

Per. / Stage	1	2	3
1	2		1
2	2	1	
3	2		1
Demand	1	1	1

Schedule 24

Per. / Stage	1	2	3
1	2		1
2	3		
3	2		1
Demand	1	1	1

Schedule 25

Per. / Stage	1	2	3
1	2	1	
2	2	1	
3	2		1
Demand	1	1	1

Schedule 26

Per. / Stage	1	2	3
1	3		
2	2	1	
3	2		1
Demand	1	1	1

Schedule 27

Per. / Stage	1	2	3
1	2	1	
2	3		
3	2		1
Demand	1	1	1

Schedule 28

Per. / Stage	1	2	3
1	3		
2	3		
3	3		
Demand	1	1	1

As opposed to the non-elementary schedules, there is a total of 19 elementary extreme flow schedules, shown below; they are numbered by the j_t notation (e.g., 10_3 denotes the 10-th elementary schedule that can be used to satisfy one unit of demand in period 3).

Schedule 1_1

Per. / Stage	1	2	3
1	1		
2	1		
3	1		
Demand	1	0	0

Schedule 1_2

Per. / Stage	1	2	3
1		1	
2		1	
3		1	
Demand	0	1	0

Schedule 2_2

Per. / Stage	1	2	3
1	1		
2		1	
3		1	
Demand	0	1	0

Schedule 3_2

Per. / Stage	1	2	3
1		1	
2	1		
3		1	
Demand	0	1	0

Schedule 4_2

Per. / Stage	1	2	3
1	1		
2	1		
3		1	
Demand	0	1	0

Schedule 5_2

Per. / Stage	1	2	3
1	1		
2	1		
3	1		
Demand	0	1	0

Schedule 1_3

Per. / Stage	1	2	3
1			1
2			1
3			1
Demand	0	0	1

Schedule 2_3

Per. / Stage	1	2	3
1		1	
2			1
3			1
Demand	0	0	1

Schedule 3_3

Per. / Stage	1	2	3
1	1		
2			1
3			1
Demand	0	0	1

Schedule 4_3

Per. / Stage	1	2	3
1			1
2		1	
3			1
Demand	0	0	1

Schedule 5_3

Per. / Stage	1	2	3
1			1
2	1		
3			1
Demand	0	0	1

Schedule 6_3

Per. / Stage	1	2	3
1		1	
2		1	
3			1
Demand	0	0	1

Schedule 7_3

Per. / Stage	1	2	3
1	1		
2		1	
3			1
Demand	0	0	1

Schedule 8_3

Per. / Stage	1	2	3
1		1	
2	1		
3			1
Demand	0	0	1

Schedule 9_3

Per. / Stage	1	2	3
1		1	
2		1	
		1	
Demand	0	0	1

Schedule 10_3

Per. / Stage	1	2	3
1	1		
2		1	
3		1	
Demand	0	0	1

Schedule 11_3

Per. / Stage	1	2	3
1		1	
2	1		
3		1	
Demand	0	0	1

Schedule 12_3

Per. / Stage	1	2	3
1	1		
2	1		
3		1	
Demand	0	0	1

Schedule 13_3

Per. / Stage	1	2	3
1	1		
2	1		
3	1		
Demand	0	0	1

The formulation with non-elementary extreme flow schedules, for the multi-product case, would be similar to (4.16) - (4.18), except that the set of equations (4.17) would be replaced by N convexity constraints of the following type:

$$\sum_{j \in J^i} \theta_{ij} = 1 \quad , \qquad i = 1,\ldots,N \qquad (4.21)$$

where J^i is the set of non-elementary schedules for product i, and the pair (ij) denotes the j-th schedule in J^i. The meaning of (4.21) is that, for each end product i, demand must be exactly met regardless by what combination of production schedules from J^i.

The advantage of this formulation is that it has T times less constraints than (4.16) - (4.18). The price that would be paid consists of two disadvantages:

- From the example presented above we can conclude that, in general, the number of elementary schedules is substantially smaller than the number of non-elementary schedules. Therefore, when a column generation procedure is used for solution, formulation (4.16) - (4.18) presents the potential of a faster rate of convergence.
- The non-elementary schedules depend on the values of the demand forecasts over the planning horizon. Thus, since the demand forecasts normally change after every updating, the entire set of schedules used to initialize the aggregate planning model would have to be generated from scratch every time (again, assuming that a column generation technique is used to solve the model).

We conclude by deciding to discard the non-elementary schedules from further consideration, and by noting that any such schedule can be obtained as a combination of elementary schedules. For instance, schedule 1 is a consolidation of $1_1 - 1_2 - 1_3$, schedule 2 is obtained from $1_1 - 1_2 - 2_3$, etc.

4.3. More Complex F2 Formulations for Uncapacitated Multi-Stage Systems

As we have shown in section 4.1, the equivalence between F2 and F1 is based entirely on the fact that the linear programming aggregate planning model F1 is a Leontief substitution system.

Four additional issues will be investigated in this section:

- initial inventories,
- non-zero lead times,
- assembly networks with diverging arcs,
- independent demand for component parts.

We will show that even when these features are considered, F1 is still Leontief or can be decomposed into Leontief subsystems. We will, thus, be able to show that a F2 formulation is equivalent to F1 under the most general conditions.

4.3.1. The Case with Initial Inventories

We still maintain the assumptions of zero lead times, pure assembly structures, and independent demand for parts, but relax the assumption of no starting stock. Notice that we are not concerned with initial inventories of the end product because they can be easily exhausted by allocating them, outside the model, against the demand forecasts.

Consider the product shown in figure 4.1, and assume that some initial inventory I_{10} of part 1 is available, but none of part 2.

What the model will do will be to meet part of the demand: d'_{31} , d'_{32} , d'_{33} , d'_{34} ,..., d'_{3T} with products using part 1 from available stock; the other part of the demand: d''_{31} ,..., d''_{3T} will be met with finished units for which part 1 has to be newly produced.

Hence, the system of equations (4.2) - (4.4) can be viewed as being composed of two subsystems S_1 and S_2 as follows (the equations will be written for just $T = 4$):

$$
\begin{array}{llllllll}
 & (S_1) & & & & & & \\
I_{10}+\cancel{x'_{11}}-I'_{11} & & & -a_{13}x'_{31} & & & = 0 & \\
I'_{11}+\cancel{x'_{12}}-I'_{12} & & & & -a_{13}x'_{32} & & = 0 & \\
I'_{12}+\cancel{x'_{13}}-I'_{13} & & & & & -a_{13}x'_{33} & = 0 & \\
I'_{13}+\cancel{x'_{14}} & & & & & & -a_{13}x'_{34} = 0 & \\
x'_{21}-I'_{21} & & -a_{23}x'_{31} & & & & = 0 & \\
I'_{21}+x'_{22}-I'_{22} & & & -a_{23}x'_{32} & & & = 0 & \\
I'_{22}+x'_{23}-I'_{23} & & & & -a_{23}x'_{33} & & = 0 & (4.22) \\
I'_{23}+x'_{24} & & & & & -a_{23}x'_{34} & = 0 & \\
x'_{31}-I'_{31} & & & & & & = d'_{31} & \\
I'_{31}+x'_{32}-I'_{32} & & & & & & = d'_{32} & \\
I'_{32}+x'_{33}-I'_{33} & & & & & & = d'_{33} & \\
I'_{33}+x'_{34} & & & & & & = d'_{34} &
\end{array}
$$

NOTE - $x'_{11},\ldots,x'_{14}$ have been crosed out because requirements for part 1 derived from demand $d'_{31},\ldots,d'_{34}$ are all satisfied only from the initial stock I_{10}, without any new production of part 1, i.e., $x'_{11} = x'_{12} = x'_{13} = x'_{14} = 0$.

- I'_{14}, I'_{24}, I'_{34} have not been included since they would be, any way, set to zero by the minimization. The implicit assumption is that initial inventories are not as large as to overflow beyond the last period of the planning horizon. If that were to happen (which is unlikely in real life situations) the amount of initial inventory should be reduced to what would just cover total requirements over the planning horizon; the model would not be able to use more, any way.

$$
\begin{array}{cccccccccccc}
 & & & & & & & & & & & \\
 & & & & & & & & & & & \\
 & & & & & & & & & & & \\
 & & & & & & & & & & & \\
 & & & & & & & & & & & \\
 & & & & & & & & & & & \\
 & & & & & & & & & & & \\
 & & & & & & & & & & & \\
 & & & & & & & & & & & \\
 & & & & & & & & & & & \\
 & & & & & & & & & & & \\
 & & & & & & & & & & & \\
 & & & & & & & & & & & \\
 & & & & & & & & & & & \\
 & & & & & & & & & & &
\end{array}
$$

(S_2)

$$
\begin{array}{cccccccccccc}
x''_{11}-I''_{11} & & & & & & -a_{13}x''_{31} & & & & & = 0 \\
I''_{11}+x''_{12}-I''_{12} & & & & & & & -a_{13}x''_{32} & & & & = 0 \\
 & I''_{12}+x''_{13}-I''_{13} & & & & & & & -a_{13}x''_{33} & & & = 0 \\
 & & I''_{13}+x''_{14} & & & & & & & -a_{13}x''_{34} & & = 0 \\
 & & & x''_{21}-I''_{21} & & & -a_{23}x''_{31} & & & & & = 0 \\
 & & & I''_{21}+x''_{22}-I''_{22} & & & & -a_{23}x''_{32} & & & & = 0 \\
 & & & & I''_{22}+x''_{23}-I''_{23} & & & & -a_{23}x''_{33} & & & = 0 \\
 & & & & & I''_{23}+x''_{24} & & & & -a_{23}x''_{34} & & = 0 \\
 & & & & & & x''_{31}-I''_{31} & & & & & = d''_{31} \\
 & & & & & & & I''_{31}+x''_{32}-I''_{32} & & & & = d''_{32} \\
 & & & & & & & & I''_{32}+x''_{33}-I''_{33} & & = d''_{33} \\
 & & & & & & & & & I''_{33}+x''_{34} & & = d''_{34}
\end{array}
\qquad (4.23)
$$

NOTE - I''_{14}, I''_{24}, I''_{34} have not been included since they would be, any way, set to zero by the minimization.

By construction of S_1 , S_2 we have (for $s = 1,2,3$):

$$d'_{31} + d''_{31} = d_{31} \qquad x'_{s1} + x''_{s1} = x_{s1} \qquad I'_{s1} + I''_{s1} = I_{s1}$$

$$d'_{32} + d''_{32} = d_{32} \qquad x'_{s2} + x''_{s2} = x_{s2} \qquad I'_{s2} + I''_{s2} = I_{s2}$$

$$d'_{33} + d''_{33} = d_{33} \qquad x'_{s3} + x''_{s3} = x_{s3} \qquad I'_{s3} + I''_{s3} = I_{s3}$$

$$d'_{34} + d''_{34} = d_{34} \qquad x'_{s4} + x''_{s4} = x_{s4} \qquad I'_{s4} + I''_{s4} = I_{s4}$$

Both (S_1) and (S_2) are Leontief substitution systems, with the optimum solutions satisfying

$$I'_{s,t-1}\, x'_{st} = 0$$

$$I''_{s,t-1}\, x''_{st} = 0$$

(Notice that period (1) is forced by construction of (S_1) and (S_2) to have the extreme flow property met.)

The initial system of equations (4.2) - (4.4), call it (S) , can be replaced by (S_1) and (S_2) since:

$$(S) = (S_1) \cup (S_2)$$

and

$$(S_1) \cap (S_2) = \phi$$

Here, we understand by (S) , (S_1) , (S_2) the corresponding feasible regions.

The two systems (S_1) and (S_2) can be easily set up:

- If $I_{10} < a_{13} d_{31} \Rightarrow a_{13}d'_{31} = I_{10}$,

- If $a_{13}d_{31} \leq I_{10} < a_{13}(d_{31}+d_{32}) \Rightarrow \begin{cases} d'_{31} = d_{31} \\ a_{13}d'_{32} = I_{10} - a_{13}d_{31} \end{cases}$, etc.

However, as it will turn out, we are not interested in the particular values of d'_{3t} and d''_{3t} ; they are only helping us with carrying out the argument, without being actually needed in solving the model. Then:

$$d''_{31} = d_{31} - d'_{31}$$

$$d''_{32} = d_{32} - d'_{32} \text{ , etc.}$$

Thus, the initial problem:

min

(OBJ. FCT.)

s.t.

(S)

nonnegativity

can be replaced by the equivalent problem:

min

(OBJ. FCT.)' + (OBJ. FCT.)"

s.t.

(S_1)

(S_2) (4.25)

nonnegativity

which, given $S_1 \cap S_2 = \phi$, can be decoupled into two independent uncapacitated problems with Leontief technology matrices. (OBJ. FCT.)' contains only variables x'_{st} , I'_{st} , and (OBJ. FCT.)" only variables x''_{st} , I''_{st} .

Let us introduce some new definitions:

- The set J_t^3 of elementary extreme flow schedules will be called the set of complete elementary schedules, because, similarly to section 4.1, such a schedule assumes no initial inventories and, therefore, has to produce all the parts and assemblies completely.
- The set J_t^{3*1} contains special elementary extreme flow schedules that assume that component part 1 is in stock and, therefore, all that has to be done is to produce part 2 and assemble end product 3 . In fact, the product considered by J_t^{3*1} looks like in figure 4.6; no production cost is charged for part 1, only holding cost from period 1 until assembly into end product 3 takes place.

In line with the reasoning in section 4.1, the two subsystems obtained from (4.25) can be given equivalent F2 formulations as follows:

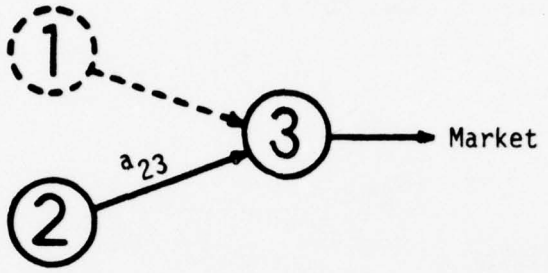


Fig. 4.6- Product structure considered by the set J_t^{3*1} of elementary schedules.

(4.26)

F1

min
(OBJ. FCT.)"
s.t.
(S_2)
nonnegativity

<⟶ equivalent ⟶>

F2

$$\min \sum_{t=1}^{T} \sum_{j_t \in J_t^3} C_{3j_t} \Theta_{3j_t}$$

s.t.

$$\sum_{j_t \in J_t^3} \Theta_{3j_t} = d''_{3t}, \quad \text{all } t \tag{4.27}$$

nonnegativity

(4.28)

F1

min
(OBJ. FCT.)'
s.t.
(S_1)
nonnegativity

<⟶ equivalent ⟶>

F2

$$\min \sum_{t=1}^{T} \sum_{j_t^* \in J_t^{3*1}} C_{3j_t^*} \Theta_{3j_t^*}$$

s.t.

$$\sum_{j_t^* \in J_t^{3*1}} \Theta_{3j_t^*} = d'_{3t}, \quad \text{all } t \tag{4.29}$$

nonnegativity

where j_t^* denotes a special elementary extreme flow schedule serving demand in period t.

It is obvious that in problem (4.29):

$$\sum_{t=1}^{T} \sum_{j_t^* \in J_t^{3*1}} \Theta_{3j_t^*} = \sum_{t=1}^{T} d'_{3t} = I_{10} \tag{4.30}$$

If (4.26) and (4.28) are consolidated, the original F1 formulation (4.24) results. If (4.27) and (4.29) are consolidated, a F2 formulation equivalent to the original F1 formulation will result:

$$\left.\begin{aligned} \min \quad & \sum_{t=1}^{T} \sum_{j_t \in J_t^3} C_{3j_t} \Theta_{3j_t} + \sum_{t=1}^{T} \sum_{j_t^* \in J_t^{3*1}} C_{3j^*} \Theta_{3j_t^*} \\ & \underline{\text{s.t.}} \\ & \sum_{j_t \in J_t^3} \Theta_{3j_t} + \sum_{j_t^* \in J_t^{3*1}} \Theta_{3j_t^*} = d_{3t} \text{ , all } t \\ & \sum_{t=1}^{T} \sum_{j_t^* \in J_t^{3*1}} \Theta_{3j_t^*} = I_{10} \\ & \text{nonnegativity} \end{aligned}\right\} \tag{4.31}$$

Two important things should be pointed out:

- When (4.31) is obtained from consolidating (4.27) and (4.29), constraint (4.30) has to be appended because the identities of d'_{3t} and d''_{3t} are lost in the consolidation process. What constraint

(4.30) is saying is that precisely as many units of end product 3 will be assembled using part 1 from stock, as there are units of part 1 in the initial inventory.

- Although the two uncapacitated systems (4.26) and (4.28), or (4.27) and (4.29) are Leontief (and, therefore, have extreme flow solutions), the overall problem (4.31) is not Leontief, and its solution can be non-extreme flow.

* * * *

Suppose now that there is initial stock of both part 1 and 2. Along the same lines followed above, the overall system (S) can be decomposed into:

- subsystem (S_1) that meets demands $d_{31}^1,\ldots,d_{3T}^1$ with end products using only part 1 from initial stock (part 2 is newly manufactured);
- subsystem (S_2) that meets demands $d_{31}^2,\ldots,d_{3T}^2$ with end products using only part 2 from initial stock (part 1 is newly manufactured);
- subsystem (S_{12}) that meets demands $d_{31}^{12},\ldots,d_{3T}^{12}$ with products using both part 1 and part 2 from available stock;
- subsystem (S_3) that meets demands $d_{31}^3,\ldots,d_{3T}^3$ with all newly manufactured component parts.

Similarly to the previous case, we can write:

$$d_{3t}^1 + d_{3t}^2 + d_{3t}^{12} + d_{3t}^3 = d_{3t}\ , \quad \text{all } t \tag{4.32}$$

Every subsystem identified above is Leontief and, thus, has an extreme-flow optimum. With each subsystem a set of elementary extreme flow schedules can be associated:

- $J_t^{3,1}$ contains all possible schedules that use only part 1 from stock, and supply one unit of finished product 3 at the beginning of period t ;
- $J_t^{3,2}$ contains schedules that use only part 2 from stock;
- $J_t^{3,12}$ contains schedules that use both part 1 and part 2 from stock;
- J_t^{3} is the set of complete schedules by which all parts have to be newly manufactured.

Also define the set J_t^{3*1} whose extreme-flow elementary schedules use part 1 from stock, either alone or in combination with part 2 from stock. A similar definition applies to J_t^{3*2}. Thus:

$$J_t^{3*1} = J_t^{3,1} \cup J_t^{3,12} \tag{4.33}$$

$$J_t^{3*2} = J_t^{3,2} \cup J_t^{3,12} \tag{4.34}$$

Let us note that it might be tempting to say that a schedule in $J_t^{3,12}$ is cheaper than any schedule in $J_t^{3,1}$ or $J_t^{3,2}$, but this is not necessarily true; it would be easy to show the contrary.

By the same argument as in the earlier example, we can provide the F1 and F2 formulations corresponding to the four subsystems:

(4.35) F1

$$\begin{cases} \min \\ \quad (\text{OBJ. FCT.})^1 \\ \text{s.t.} \\ \quad (S_1) \\ \text{nonnegativity} \end{cases}$$

equivalent $\longleftrightarrow$

F2

$$\begin{cases} \min \quad \sum_{t=1}^{T} \sum_{j_t^* \in J_t^{3,1}} C_{3j_t^*} \Theta_{3j_t^*} \\ \text{s.t.} \\ \quad \sum_{j_t^* \in J_t^{3,1}} \Theta_{3j_t^*} = d_{3t}^1 \text{ , all } t \\ \text{nonnegativity} \end{cases} \tag{4.36}$$

(4.37) F1

$$\begin{cases} \min \\ \quad (\text{OBJ. FCT.})^2 \\ \text{s.t.} \\ \quad (S_2) \\ \text{nonnegativity} \end{cases}$$

equivalent $\longleftrightarrow$

F2

$$\begin{cases} \min \quad \sum_{t=1}^{T} \sum_{j_t^* \in J_t^{3,2}} C_{3j_t^*} \Theta_{3j_t^*} \\ \text{s.t.} \\ \quad \sum_{j_t^* \in J_t^{3,2}} \Theta_{3j_t^*} = d_{3t}^2 \text{ , all } t \\ \text{nonnegativity} \end{cases} \tag{4.38}$$

(4.39)

F1

min

$(\text{OBJ. FCT.})^{12}$

s.t.

(S_{12})

nonnegativity

equivalent

F2

$$\min \quad \sum_{t=1}^{T} \sum_{j_t^* \in J_t^{3,12}} C_{3j_t^*} \Theta_{3j_t^*} \qquad (4.40)$$

$$\text{s.t.} \quad \sum_{j_t^* \in J_t^{3,2}} \Theta_{3j_t^*} = d_{3t}^2 \text{ , all } t$$

nonnegativity

(4.41)

F1

min

$(\text{OBJ. FCT.})^{3}$

s.t.

(S_3)

nonnegativity

equivalent

F2

$$\min \quad \sum_{t=1}^{T} \sum_{j_t \in J_t^3} C_{3j_t} \Theta_{3j_t} \qquad (4.42)$$

$$\text{s.t.} \quad \sum_{j_t \in J_t^3} \Theta_{3j_t} = d_{3t}^3 \text{ , all } t$$

nonnegativity

where $(\text{OBJ. FCT})^1$, $(\text{OBJ. FCT})^2$, etc. are the objective functions corresponding to the subsystems.

Evidently, the following equations hold:

$$\sum_{t=1}^{T} (d_{3t}^1 + d_{3t}^2) = I_{10} \qquad (4.43)$$

$$\sum_{t=1}^{T} (d_{3t}^{2} + d_{3t}^{12}) = I_{20} \qquad (4.44)$$

By consolidating (4.36), (4.38), (4.40), and (4.42) a F2 formulation equivalent to the original problem (system S) is obtained. Again, we have to be careful to append (4.43) and (4.44) to the new formulation since, in the process of consolidation, the identities of d_{3t}^{1}, d_{3t}^{2}, d_{3t}^{12}, and d_{3t}^{3} are lost. We would like to mention that, under the reasonable assumption that $I_{10} \leq \sum_{t} a_{13} d_{3t}$ and $I_{20} \leq \sum_{t} a_{23} d_{3t}$, (4.43) and (4.44) do hold at equality because, once there is some initial inventory, it is cheaper to use it than holding it and producing other parts instead. Thus, the complete F2 formulation for our problem is:

$$\min \sum_{t=1}^{T} \sum_{j_t^* \in J_t^{3,1}} C_{3j_t^*}\,\Theta_{3j_t^*} + \sum_{j_t^* \in J_t^{3,2}} C_{3j_t^*}\,\Theta_{3j_t^*} + \sum_{j_t^* \in J_t^{3,12}} C_{3j_t^*}\,\Theta_{3j_t^*} + \sum_{j_t \in J_t^{3}} C_{3j_t}\,\Theta_{3j_t} \tag{4.45}$$

s.t.

$$\sum_{j_t^* \in J_t^{3,1}} \Theta_{3j_t^*} + \sum_{j_t^* \in J_t^{3,2}} \Theta_{3j_t^*} + \sum_{j_t^* \in J_t^{3,12}} \Theta_{3j_t^*} + \sum_{j_t \in J_t^{3}} \Theta_{3j_t} = d_{3t}, \text{ all } t$$

$$\sum_{t=1}^{T} \sum_{j_t^* \in J_t^{3,1}} \Theta_{3j_t^*} + \sum_{t=1}^{T} \sum_{j_t^* \in J_t^{3,12}} \Theta_{3j_t^*} = I_{10}$$

$$\sum_{t=1}^{T} \sum_{j_t^* \in J_t^{3,2}} \Theta_{3j_t^*} + \sum_{t=1}^{T} \sum_{j_t^* \in J_t^{3,12}} \Theta_{3j_t^*} = I_{20}$$

nonnegativity

The first constraint ensures that demand in period t is met. The second constraint does not permit more or less special schedules using part 1 from initial stock than there is available initial inventory of part 1. The third constraint has a similar interpretation, but with respect to part 2.

Given the uncapacitated nature of the problem, the solution to (4.45) is the reunion of the solutions to the four subsystems.

The extension of the above analysis to other uncapacitated pure assembly systems, regardless of the number of stages, is straightforward.

4.3.2. The Case with Initial Inventories of Parts Common to Two or More End Products

Assumptions: zero lead times, every product has a pure assembly structure, and there is no independent demand for parts. The two products shown in figure 4.7 share component part 2. Suppose there is initial inventory I_{10} , I_{20} , I_{40} of component parts 1, 2, and 4 , respectively.

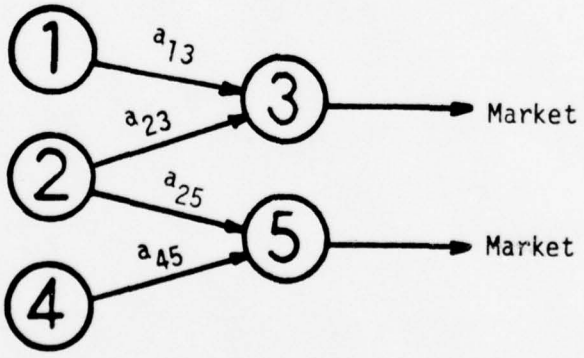


Fig. 4.7- Two products with a common part.

Although the overall problem is not Leontief, it can be decomposed into a number of subsystems; they will be shown below together with the corresponding sets of elementary extreme flow production schedules:

For END PRODUCT 3 (assembled at stage 3):

- (S_1^3) uses only part 1 from stock and serves $d_{31}^1,\ldots,d_{3T}^1$; $J_t^{3,1}$.
- (S_2^3) uses only part 2 from stock to serve $d_{31}^2,\ldots,d_{3T}^2$; $J_t^{3,2}$.
- (S_{12}^3) uses both part 1 and 2 from stock to serve $d_{31}^{12},\ldots,d_{3T}^{12}$; $J_t^{3,12}$.
- (S_3^3) serves $d_{31}^3,\ldots,d_{3T}^3$ with all newly manufactured units; J_t^3 .

For END PRODUCT 5 (assembled at stage 5):

- (S_2^5) uses only part 2 from stock to serve $d_{51}^2,\ldots,d_{5T}^2$; $J_t^{5,2}$.
- (S_4^5) uses only part 4 from stock to serve $d_{51}^4,\ldots,d_{5T}^4$; $J_t^{5,4}$.
- (S_{24}^5) uses both part 2 and 4 from stock to serve $d_{51}^{24},\ldots,d_{5T}^{24}$; $J_t^{5,24}$.
- (S_5^5) serves $d_{51}^5,\ldots,d_{5T}^5$ with all newly made finished units; J_t^5.

The following relations hold:

$$\Sigma_t(d_{31}^1 + d_{3t}^{12}) = I_{10} \qquad (4.46)$$

$$\Sigma_t(d_{3t}^{12} + d_{3t}^2 + d_{5t}^2 + d_{5t}^{24}) = I_{20} \qquad (4.47)$$

$$\Sigma_t(d_{5t}^{24} + d_{5t}^4) = I_{40} \qquad (4.48)$$

$$d^{2}_{5t} + d^{4}_{5t} + d^{24}_{5t} + d^{5}_{5t} = d_{5t} \text{ , all } t \qquad (4.49)$$

$$d^{1}_{3t} + d^{2}_{3t} + d^{12}_{3t} + d^{3}_{3t} + d_{3t} \text{ , all } t \qquad (4.50)$$

(4.46) - (4.48) will have to be incorporated in the new F2 formulation, similarly to the way it has been done in the previous cases.

The follwoing notations will be useful in simplifying the writing of the model:

J_t^{3*1} = the set of special elementary extreme flow schedules that supply one unit of PRODUCT 3 to period t , by using part 1 from stock (part 2 can be either from stock or newly made);

J_t^{3*2} = similar to the above, except that it refers to part 2;

J_t^{5*2} = the set of special elementary extreme flow schedules that supply one unit of PRODUCT 5 to period t , by using part 2 from stock (part 4 can be either from stock or newly made);

J_t^{5*4} = similar to the above, but referring to part 4 .

J_t^{3*} = the totality of all special elementary extreme flow schedules for PRODUCT 3 and period t .

J_t^{5*} = the totality of all special elementary extreme flow schedules for PRODUCT 5 and period t .

The formal definitions are:

$$J_t^{3*1} = J_t^{3,1} \cup J_t^{3,12} \tag{4.51}$$

$$J_t^{3*2} = J_t^{3,2} \cup J_t^{3,12} \tag{4.52}$$

$$J_t^{5*2} = J_t^{5,2} \cup J_t^{5,24} \tag{4.53}$$

$$J_t^{5*4} = J_t^{5,4} \cup J_t^{5,24} \tag{4.54}$$

$$J_t^{3*} = J_t^{3*1} \cup J_t^{3*2} \tag{4.55}$$

$$J_t^{5*} = J_t^{5*2} \cup J_t^{5*4} \tag{4.56}$$

and the F2 formulation is:

$$\min \sum_t \left[\sum_{j_t \in J_t^3} C_{3j_t} \Theta_{3j_t} + \sum_{j_t^* \in J_t^{3*}} C_{3j_t^*} \Theta_{3j_t^*} + \sum_{j_t \in J_t^5} C_{5j_t} \Theta_{5j_t} + \right.$$

$$\left. + \sum_{j_t^* \in J_t^{5*}} C_{5j_t} \Theta_{5j_t} \right] \tag{4.57}$$

s.t.

$$\sum_{j_t \in J_t^3} \Theta_{3j_t} + \sum_{j_t^* \in J_t^{3*}} \Theta_{3j_t^*} = d_{3t} \text{ , all } t \tag{4.58}$$

$$\sum_{j_t \in J_t^5} \Theta_{5j_t} + \sum_{j_t^* \in J_t^{5*}} \Theta_{5j_t^*} = d_{5t} \text{ , all } t \tag{4.59}$$

$$\sum_t \left[\sum_{j_t^* \in J_t^{3,1}} \Theta_{3j_t^*} + \sum_{j_t^* \in J_t^{3,12}} \Theta_{3j_t^*} \right] = I_{10} \tag{4.60}$$

$$\sum_t \left[\sum_{j_t^* \in J_t^{3,12}} \Theta_{3j_t^*} + \sum_{j_t^* \in J_t^{3,2}} \Theta_{3j_t^*} + \sum_{j_t^* \in J_t^{5,2}} \Theta_{5j_t^*} + \sum_{j_t^* \in J_t^{5,24}} \Theta_{5j_t^*} \right] =$$

$$= I_{20} \tag{4.61}$$

$$\sum_t \left[\sum_{j_t^* \in J_t^{5,4}} \Theta_{5j_t^*} + \sum_{j_t^* \in J_t^{5,24}} \Theta_{5j_t^*} \right] = I_{40} \tag{4.62}$$

nonnegativity

Equations (4.58), (4.59) require that demand be met in all time periods by any combination of complete and special elementary extreme flow schedules. Constraints (4.60) - (4.62) ensure that the total number of special schedules that use a certain part from stock will be exactly equal to the amount of initial inventory of that part.

4.3.3. The Case with Non-Zero Lead Times and Initial Inventories

Assumptions: every product has a pure assembly structure, and there is no independent demand for parts.

In this case, aside from initial inventories, there are deliveries of parts throughout the lead times from production started in periods prior to the current period 1 . The system does no longer have an extreme-flow solution since in the early stages property $I_{s,t-1}\, x_{st} = 0$ may be violated. Therefore, we will again decompose the problem into Leontief subsystems in order to be able to work with extreme flow schedules.

Consider, first, a simple example with three stages only (figure 4.8). I_{10} , I_{20} are initial inventories of parts 1 and 2 . The lead

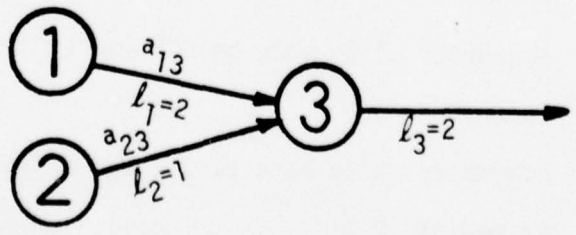


Fig. 4.8- A simple structure with three stages.

times are shown in the picture as ℓ_1 = 2 periods, ℓ_2 = 1 period , ℓ_3 = 2 periods. Since the decomposition into Leontief subsystems has been described in detail in the previous sections, we will only show the resulting sets of elementary extreme flow schedules, for which the following notations will be used:

$J_{t,v}^{i,p_1p_2}$ = the set of special elementary schedules that produce one unit of end product i for period t, by using part p_1 and part p_2 from the initial inventories available in period v ;

J_t^i = the set of complete elementary schedules for end product i and period t .

The sets of schedules are then:

$J_{t,1}^{3,1}$ = the set of special elemenatry extreme flow schedules that use part 1 from the stock available in period 1 ;

$J_{t,2}^{3,1}$ = the set of special elementary extreme flow schedules that use part 1 from the stock available in period 2 ;

$J_{t,1}^{3,2}$ = schedules that use part 2 from the stock available in period 1 ;

$J_{t,1}^{3,12}$ = schedules using both part 1 and 2 from stocks available in period 1 ;

$J_{t,2}^{3,12}$ = schedules using both part 1 and 2 from stocks available in period 2 ;

J_t^3 = complete schedules.

The F2 formulation is presented below as (4.63) - (4.68):

min

$$\sum_{t=3}^{T+2} \sum_{j_t \in J_t^3} C_{3j_t} \Theta_{3j_t} + \sum_{j_t^* \in J_t^{3*}} C_{3j_t^*} \Theta_{3j_t^*} + \sum_{s=1}^{2} h_{s1} I_{s1} \qquad (4.63)$$

s.t.

$$\sum_{j_t \in J_t^3} \Theta_{3j_t} + \sum_{v=1}^{2} \sum_{j_t^* \in J_{t,v}^{3*}} \Theta_{3j_t^*} = d_{3t}, \quad t=3,\ldots,T+2 \qquad (4.64)$$

$$\sum_{t=3}^{T+2} \sum_{j_t^* \in J_{t,1}^{3*1}} \Theta_{3j_t^*} + I_{11} = I_{10} + DEL_1^1 \qquad (4.65)$$

$$\sum_{t=3}^{T+2} \sum_{j_t^* \in J_{t,1}^{3*2}} \Theta_{3j_t^*} + I_{21} = I_{20} + DEL_1^2 \qquad (4.66)$$

$$\sum_{t=4}^{T+2} \sum_{j_t^* \in J_{t,2}^{3*1}} \Theta_{3j_t^*} = I_{11} + DEL_2^1 \qquad (4.67)$$

$$\sum_{t=4}^{T+2} \sum_{j_t^* \in J_{t,2}^{3,12}} \Theta_{3j_t^*} = I_{21} \qquad (4.68)$$

nonnegativity

A number of things require explanations:

- J_t^{3*} has been defined in (4.55);
- $J_{t,v}^{3*}$ is, in fact, $J_{t,v}^{3*1} \cup J_{t,v}^{3*2}$, where $J_{t,v}^{3*1} = J_{t,v}^{3,1} \cup J_{t,v}^{3,12}$ and $J_{t,v}^{3*2} = J_{t,v}^{3,2} \cup J_{t,v}^{3,12}$.
- Since the lead time of stage 3 is two periods, the earliest demand that can be "reached" by any production scheduled is that of period 3; this is the reason why the first summation in the objective function and the t index in (4.64) run from $t=3$. Also, since the planning horizon is T period long, T production decisions are called for; this requires the time index t to run up to period $T+2$. However, $J_t^3 = \phi$ for $t=3$, since the cumulated lead time for making an end product with no parts from stock is three periods.
- Notation DEL_1^1 representes scheduled deliveries of part 1 in period 1, from production started two periods ahead; similarly, DEL_2^1 are scheduled deliveries of part 1 in period 2. As part 2 has a lead time of one period only, only deliveries DEL_1^2 in period 1 are expected.
- Since there are deliveries of part 1 over the first two periods, and because both one unit of part 1 and one unit of part 2 are needed to assemble the end product, it is necessary to provide the possibility for the excess stock of part 1 or 2 to "move" to period 2 and be used then. The excess can occur because of the imbalances between what is available of part 1 and part 2 in period 1. There-

fore, the inventory variables I_{11} and I_{21} have been incorporated into the formulation. Notice that there is no need to allow the transmission of stock of parts to period 3 or later, because the elementary schedules would account for the holding cost of any component part available in period 1 or 2 and used for assemblying the end product 3 in a later period.

- There is no need for a set of special schedules of the type $J^{3,2}_{t,2}$, that would produce end product 3 for period t with component parts 2 available in stock in period 2, because all inventory of part 2 becomes available in period 1 (initial stock plus deliveries). Although some inventory I_{21} will "migrate" to period 2, any schedule in $J^{3,2}_{t,1}$ will be able to reflect in its cost the holding of one unit of part 2 from period 1 to a later period. It is important, then, to emphasize that <u>variables I_{11} and I_{21} handle only inventories that are not included in any of the schedules which use parts available in period 1</u> (e.g., $J^{3,1}_{t,1}$, $J^{3,2}_{t,1}$, $J^{3,12}_{t,1}$).

Now it is easy to interpret model (4.63) - (4.68). Constraints (4.64) ensure that demand for the end product is met, regardless of the combination of schedules by which this is accomplished. Expression (4.65) represents an inventory balance equation for period 1:

- the left hand side shows the amount of part 1 consumed, by the special schedules in $J^{3*1}_{t,1}$, from inventory available in period 1, plus the amount of part 1 put in stock at the end of period 1;

- the right hand side represents the availabilities of part 1 in period 1 , namely, initial inventory I_{10} and scheduled deliveries DEL_1^1 .

Equations (4.66) - (4.68) have similar interpretations.

* * * *

Consider now an example (figure 4.9) that brings another, and the last, complication in terms of initial inventories. For simplicity, assume that only initial stocks and scheduled deliveries of parts 1, 2, and 4 exist, and none of assemblies 3 and 5 .

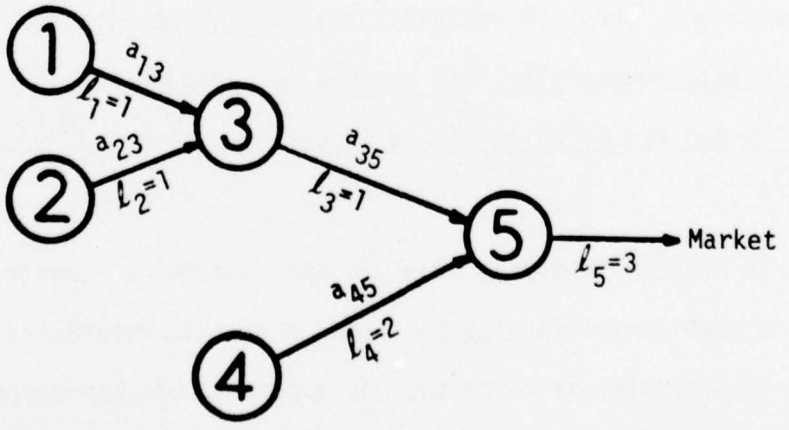


Fig. 4.9- A case with initial inventories of component parts 1, 2 and 4.

Similarly to the previous examples, the overall problem can be decomposed into as many Leontief subsystems as there are ways of producing one unit of end product 5 . Thus, the following sets of elementary extreme flow schedules can be identified:

$J^{5,1}_{t,1}$ - only part 1 used from stock, available in period 1;

$J^{5,2}_{t,1}$ - only part 2 used from stock, available in period 1;

$J^{5,4}_{t,1}$ - only part 4 used from stock, available in period 1;

$J^{5,4}_{t,2}$ - only part 4 used from stock, available in period 2;

$J^{5,12}_{t,1}$ - both part 1 and 2 used from stock, available in period 1;

$J^{5,1\text{-}4}_{t,1\text{-}2}$ - part 1 used from stock, available in period 1 , and part 4 used from stock available in period 2; all other components newly manufactured;

$J^{5,2\text{-}4}_{t,1\text{-}2}$ - part 2 used from stock, available in period 1, and part 4 used from stock available in period 2; all other components newly manufactured;

$J^{5,12\text{-}4}_{t,\ 1\text{-}2}$ - parts 1 and 2 used from stock, available in period 1; part 4 used from stock available in period 2;

J^{5}_{t} - nothing used from stock; all newly made components.

The F2 formulation is straightforward; therefore, we will only present the inventory balance equations (4.69) - (4.72) (see next page) .

$$\sum_{t=6}^{16} \sum_{j_t^* \varepsilon J_{t,1}^{5,1}} \theta_{5j_t^*} + \sum_{t=6}^{16} \sum_{j_t^* \varepsilon J_{t,1}^{5,12}} \theta_{5j_t^*} + \sum_{t=6}^{16} \sum_{j_t^* \varepsilon J_{t,1-2}^{5,1-4}} \theta_{5j_t^*} + \sum_{t=5}^{16} \sum_{j_t^* \varepsilon J_{t,1-2}^{5,12-4}} \theta_{5j_t^*} = I_{10} + DEL_1^1 \tag{4.69}$$

$$\sum_{t=6}^{16} \sum_{j_t^* \varepsilon J_{t,1}^{5,2}} \theta_{5j_t^*} + \sum_{t=6}^{16} \sum_{j_t^* \varepsilon J_{t,1}^{5,12}} \theta_{5j_t^*} + \sum_{t=6}^{16} \sum_{j_t^* \varepsilon J_{t,1-2}^{5,2-4}} \theta_{5j_t^*} + \sum_{t=5}^{16} \sum_{j_t^* \varepsilon J_{t,1-2}^{5,12-4}} \theta_{5j_t^*} = I_{20} + DEL_1^2 \tag{4.70}$$

$$\sum_{t=6}^{16} \sum_{j_t^* \varepsilon J_{t,1}^{5,4}} \theta_{5j_t^*} + I_{41} = I_{40} + DEL_1^4 \tag{4.71}$$

$$\sum_{t=6}^{16} \sum_{j_t^* \varepsilon J_{t,2}^{5,4}} \theta_{5j_t^*} + \sum_{t=6}^{16} \sum_{j_t^* \varepsilon J_{t,1-2}^{5,1-4}} \theta_{5j_t^*} + \sum_{t=6}^{16} \sum_{j_t^* \varepsilon J_{t,1-2}^{5,2-4}} \theta_{5j_t^*} + \sum_{t=5}^{16} \sum_{j_t^* \varepsilon J_{t,1-2}^{5,12-4}} \theta_{5j_t^*} = I_{41} + DEL_2^4 \tag{4.72}$$

Assumed horizon: T=13 periods. Because of the similarities with equations (4.65) - (4.68) it is not necessary to insist with further details. One remark, however: equations (4.71) is redundant for this particular problem, and has been included only to have a more complete representation. Indeed, given that no initial inventories and no scheduled deliveries of assembly 3 are assumed, part 4 alone cannot be used in period 1 because the assembly of 5 requires both part 3 and 4 . Thus $J^{5,4}_{t,1} = \phi$.

* * * *

The two examples analyzed above will permit us to draw several conclusions:

- When two or more component parts go into the same (sub)assembly, like parts 1 and 2 in figure 4.8 , their inventory balance equations will have to span the same number of periods, number which has to equal the longest lead time (see equations (4.65) - (4.68) that span 2 periods, although $\ell_1 = 2$ and $\ell_2 = 1$) . If however, the parts enter into different assemblies (e.g., parts 1 and 4 in figure 4.9) their inventory balance equations can cover different numbers of time intervals because they are needed at different points in time in the production process.
- Technological constraints have to be observed when deciding what sets of schedules should be incorporated into the model, in an effort to keep their number as small as possible. For instance, for the product of figure 4.9 it would not make sense to consider

a set $J^{5,124}_{t,1}$ because, although parts 1, 2, and 4 can be available as initial stock in period 1 , only parts 1 and 2 can be immediately used. Part 4 can be used only in period 2 at the earliest; hence, the inclusion of set $J^{5,12-4}_{t,\ 1-2}$ is correct.

- By reasons of technological feasibility, in our latest example we concluded that $J^{5,4}_{t,1} = \phi$. Why not also $J^{5,4}_{t,2} = \phi$, since it would take two periods to make assembly 3 from "scratch"? It is true that part 4 could not be used before period 3 by a schedule of type $J^{5,4}_{t,2}$, and therefore we might consider discarding $J^{5,4}_{t,2}$ and using a set $J^{5,4}_{t,3}$ instead. But, this would require an additional inventory balance equation and a new variable I_{42} to permit the flow of stock of part 4 into period 3 . This approach, then, unnecessarily increases the size of the problem. Notice that the final result is the same; the difference is that, rather than explicitly representing the flow of inventory of part 4 out of period 2 via variable I_{42} , the same stock will flow inside schedules $J^{5,4}_{t,2}$ whose costs will reflect the holding of inventory (see, for instance, the cost computing equations (4.19) and (4.20)) .

4.3.4. The case of Assembly Structures with Diverging Arcs

Assumptions: zero lead times, zero initial inventories, no independent demand for component parts.

The product we are now considering has part 2 going into both component 4 and subassembly 3 (figure 4.10).

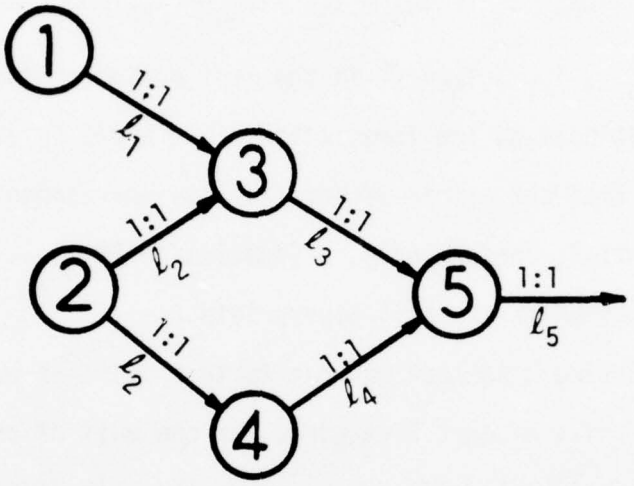


Fig. 4.10- Product structure with diverging arcs.

For period t, the inventory balance equations of type (3.2) - (3.3) (see section 3.1) for this case are:

$$I_{1,t-1}+x_{1t}-I_{1t} \quad -x_{3t} = 0 \qquad (4.73)$$

$$I_{2,t-1}+x_{2t}-I_{2t} \quad -x_{3t} \quad -x_{4t} = 0 \qquad (4.74)$$

$$I_{3,t-1}+x_{3t}-I_{3t} \quad -x_{5t} = 0 \qquad (4.75)$$

$$I_{4,t-1}+x_{4t}-I_{4t} \quad -x_{5t} = 0 \qquad (4.76)$$

$$I_{5,t-1}+x_{5t}-I_{5t} = d_{5t} \qquad (4.77)$$

By assumption, $I_{1Q} = I_{2Q} = I_{3Q} = I_{4Q} = I_{5Q} = 0$, and variables I_{1T} , I_{2T} , I_{3T} , I_{4T} , I_{5T} (T is the last period of the planning horizon) will not be included in the formulation since would be zero any way. It is then clear that the matrix of coefficients corresponding to (4.73) - (4.77) is Leontief; consequently, a formulation F2 in terms of elementary extreme flow scheudles is still appropriate.

Before closing this section, one further point is worth mentioning. There are two units of part 2 required for one unit of end product 5; two kinds of elementary schedules can be generated with respect to component part 2:

a) an elementary schedule that produces both units of 2 at the same time;

b) an elementary schedule that allows the two units of 2 to be produced in different time periods, as required by the production schedule of subassemblies 3 and 4 . In fact, a schedule of type (b) implies that the product of figure 4.10 has, for aggregate planning purposes, the structure shown in figure 4.11 Parts 2' and 2" represent the same component part 2, except for their destinations.

If we analyze equations (4.73) - (4.77), pertaining to the F1 formulation, it becomes clear that approach (b) is implied rather than (a); indeed, part 3 and 4 might very well be scheduled for production in two different periods. Hence, the two units of 2 , 2' and 2" , might as well be produced in different time periods (recall that no setups are involved).

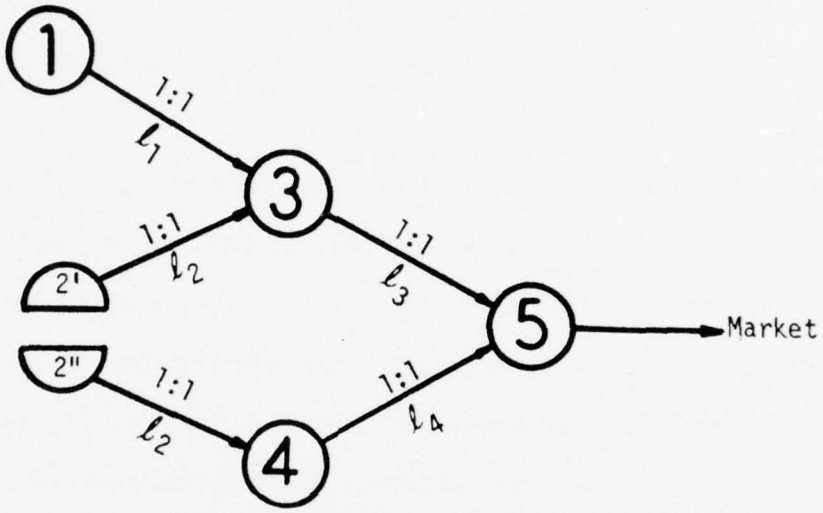


Fig. 4.11- The structure of fig. 4.10 as viewed by an elementary schedule of type (b).

Therefore, to guarantee that the resulting F2 formulation is equivalent to the original F1 , elementary extreme flow schedules of type (b) have to be used. We can also make the remark that approach (a) would tend to yield a costlier solution because it is more constrained than (b).

4.3.5. The Case with Independent Demand for Component Parts

A component part used for service purposes, thus also exhibiting independent demand, will be considered as just another end product; the resulting case of multiple end products has been treated in section 4.1 .

4.3.6. Further Refinements for the Uncapacitated Problem

It should be emphasized that, although the features introduced in sections 4.3.1 through 4.3.5 have been treated independently from each other, they can be superimposed due to the linear nature of the aggregate planning model. For this reason, the general multi-stage problem with

initial inventories, non-zero lead times, assembly structures with diverging arcs, multiple products, and independent demand for component parts can still be given a F2 formulation in terms of elementary extreme flow schedules.

As it has already been mentioned, the F2 formulations achieved the desired reduction in the number of rows, at the cost of increased number of columns;also (see discussion in section 4.4.4) most of the constrains have special structure (Generalized Upper Bounding Constraints). We intend to use the column generation technique to solve F2; an algorithm will be developed in a later section. The problem encountered here, however, is that at every iteration one elementary schedule has to be generated for every set of complete and special schedules. If there are initial inventories, the number of sets can become very large: there have been 4 sets in the second example of section 4.3.1 , 8 sets in the example in section 4.3.2 , 9 sets in the second example of section 4.3.3. The longer the lead times, and the more complex the product structures, the larger the number of sets of schedules that have to be considered whenever initial inventories exist. In what follows we will derive some results that will lead to a reduction in the computational effort for uncapacitated problems, under mild restrictions imposed upon the cost structure. Later, the capacitated problem will be treated, for which two approaches, for cutting down the computations, will be proposed:

- multiple pricing (Lasdon and Terjung [64]), and
- conversion of initial inventories into an equivalent amount of productive capacity.

Consider the product shown in figure 4.12(same product as in section 4.3.3, figure 4.9). There are five stages, and every stage produces one part. Assume that the cost structure satisfies the following conditions (for notations refer to problem (1.1) - (1.7) of section 1.1):

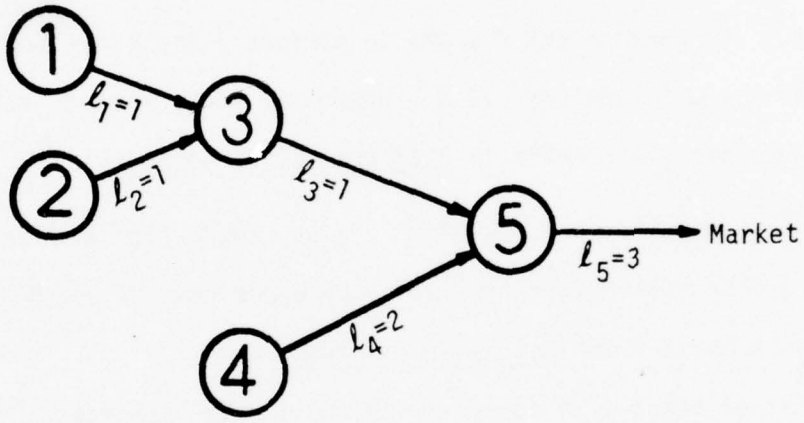


Fig. 4.12- A five stage product structure.

$$v_{st} \geq v_{s,t+1} \tag{4.78}$$

$$h_{3t} \geq h_{1t} + h_{2t} \tag{4.79}$$

$$h_{5t} \geq h_{3t} + h_{4t} \tag{4.80}$$

These assumptions, although somewhat restrictive, are reasonable. In most cases we would expect to use a constant variable production cost, i.e., $v_{st} = v_{s,t+1}$, with the possibility of discounting over time, in which case $v_{st} > v_{s,t+1}$. The restriction on holding cost simply states that the value added at every stage is non-negative; in most cases we expect this restriction to hold at inequality.

The above assumptions imply that in an uncapacitated environment one would produce as late as possible in time, and would hold inventories in earlier stages rather than in later stages.

For simplicity, assume that there are initial inventories only of parts 1, 2 and 4 . Also, there are scheduled deliveries of parts: in period 1 for parts 1 and 2 , and in periods 1 and 2 for part 4 .

As shown in section 4.3.3 , there are three sets of special schedules which use parts available in <u>different time periods:</u> $J^{5,1-4}_{t,1-2}$, $J^{5,2-4}_{t,1-2}$, $J^{5,12-4}_{t,\ 1-2}$. We will show that for $t \geq 6$, these sets are unnecessary because they are costwise dominated by other sets of special schedules which use parts from <u>the same time period</u>.

Let us start with comparing the following two sets:

$$J1 = J^{5,1-4}_{t,1-2} \qquad \text{vs.} \qquad J2 = J^{5,1}_{t,1}$$

Let:

π_{5t} = the dual variable associated with the demand constraint for period t (see, for instance, equation (4.64));

π^1_1 = the dual variable associated with the inventory balance equation (4.69);

π^2_1 = the dual variable associated with (4.70);

π^4_1 = the dual variable associated with (4.71);

π^4_2 = the dual variable associated with (4.72);

b' , b" = period when part 2 is scheduled in production by a schedule in J1 , and in J2 , respectively;

c' , c" = similarly, production periods for part 3 ;

d" = period when part 4 is started in production by a schedule in J2 .

The reduced costs for one schedule in J1 and for one in J2 are:

$$\overline{C}_{5,j_t^* \varepsilon J1} = \sum_{p=1}^{c'-1} h_{1p} + v_{2b'} + \sum_{p=b'+1}^{c'-1} h_{2p} + v_{3c'} +$$

$$+ \sum_{p=c'+1}^{e'-1} h_{3p} + \underline{\sum_{p=2}^{e'-1} h_{4p}} + v_{5e'} +$$

$$+ \sum_{p=e'+3}^{t-1} h_{5p} - \pi_{5t} - \pi_1^1 - \underline{\pi_2^4}$$

$$\overline{C}_{5,j_t^* \varepsilon J2} = \sum_{p=1}^{c"-1} h_{1p} + v_{2b"} + \sum_{p=b"+1}^{c"-1} h_{2p} + v_{3c"} +$$

$$+ \sum_{p=c"+1}^{e"-1} h_{3p} + \underline{v_{4d"} + \sum_{p=d"+2}^{e"-1} h_{4p}} + v_{5e"} +$$

$$+ \sum_{p=e"+3}^{t-1} h_{5p} - \pi_{5t} - \pi_1^1$$

Given assumptions (4.78) - (4.80) we have:

$$b' = b'' = b \; ; \qquad c' = c'' = c \; ; \qquad e' = e'' = e$$

Moreover:

$$b = t - 5 \; ; \qquad c = t - 4 \; ; \qquad e = t - 3$$

Thus, to compare the two reduced costs one has to look only at the underlined terms:

$$\underline{\sum_{p=2}^{e-1} h_{4p} - \pi_2^4} \qquad \text{vs.} \qquad v_{4d} + \underline{\sum_{p=d+2}^{e-1} h_{4p}} .$$

π_2^4 is negative because a cost saving is achieved when more parts are available as initial inventory. Thus, the comparison becomes:

$$\underline{\sum_{p=2}^{e-1} h_{4p} + |\pi_2^4|} \qquad \text{vs.} \qquad v_{4d} + \underline{\sum_{p=d+2}^{e-1} h_{4p}} .$$

Since $d \geq 1$, clearly:

$$\sum_{p=2}^{e-1} h_{4p} \geq \sum_{p=d+2}^{e-1} h_{4p} \tag{4.81}$$

$|\pi_2^4|$ is the marginal cost saving associated with one unit of part 4 being available in stock in period 2 . If we did not have that unit in inventory we would have to make it as early as possible (or more precisely, we should have started it in production before period 1); the earliest time is period 1 , and the cost is v_{41} . Therefore, the saving $|\pi_2^4|$ is at least as large as the cost of making one unit of part 4 if it were not available in stock:

$$|\pi_2^4| \geq v_{41}$$

Since $d \geq 1$, by (4.78) we have:

$$|\pi_2^4| \geq v_{41} \geq v_{4d} \tag{4.82}$$

By (4.81) and (4.82) we can conclude that:

$$\overline{C}_{5,j_t^* \varepsilon J_{t,1}^{5,1}} \leq \overline{C}_{5,j_t^* \varepsilon J_{t,1-2}^{5,1-4}} \quad , \qquad t \geq 6$$

By similar developments, the following results can be obtained:

$$\overline{C}_{5,j_t^* \varepsilon J_{t,2}^{5,4}} \leq \overline{C}_{5,j_t^* \varepsilon J_{t,1-2}^{5,1-4}} \quad , \qquad t \geq 6$$

$$\overline{C}_{5,j_t^* \varepsilon J_{t,1}^{5,2}} \leq \overline{C}_{5,j_t^* \varepsilon J_{t,1-2}^{5,2-4}} \quad , \qquad t \geq 6$$

$$\bar{C}_{5,j_t^* \in J_{t,2}^{5,4}} \leq \bar{C}_{5,j_t^* \in J_{t,1-2}^{5,2-4}} , \qquad t \geq 6$$

$$\bar{C}_{5,j_t^* \in J_{t,1}^{5,12}} \leq \bar{C}_{5,j_t^* \in J_{t,\,1-2}^{5,12-4}} , \qquad t \geq 6$$

$$\bar{C}_{5,j_t^* \in J_{t,2}^{5,4}} \leq \bar{C}_{5,j_t^* \in J_{t,\,1-2}^{5,12-4}} , \qquad t \geq 6 .$$

These results suggest that, under the stated cost assumptions and for t sufficiently large, we can exclude from the uncapacitated aggregate model the sets of special schedules which use component parts from initial inventories available in different time periods.

Two comments, however:

- for small t, demand can only be satisfied with schedules which use parts from stocks available in different time periods (in our example, demand in period $t = 5$ can only be satisfied with a schedule in $J_{t,\,1-2}^{5,12-4}$). But, even in this case sets $J_{t,1-2}^{5,1-4}$ and $J_{t,1-2}^{5,2-4}$ can be discarded.
- if total demand over the planning horizon is small relative to the initial inventories and scheduled deliveries of parts, it might be impossible to use up the available stocks of parts (as required by constraints (4.69) - (4.72)) unless the special schedules which use

parts from different time periods are also included in the formulation. In practice, however, we do not expect cases of this sort.

4.4. F2 Formulations for Capacitated Multi-Stage Systems

The uncapacitated cases studied so far have constituted the vehicle for introducing concepts and basic results related to the new formulation for the multi-stage aggregate production problem. Real situations, however, involve limited resources; therefore, extensions to cover the capacitated case have to be developed.

The F1 formulation of the capacitated problem has been presented as (3.2) - (3.6) in section 3.1 . It is an easy matter to also impose capacity restrictions in the F2 formulation. For illustration consider the very simple case where there are no initial inventories, and no scheduled deliveries (from production started prior to period 1) over the lead time; there are S stages, each with one productive resource, and every stage only produces one part or assembly. A total of N end products (among which there can also be components used for service purposes) have to be scheduled; the planning horizon has T time periods.

The uncapacitated F2 formulation is only slightly different from model (4.16) - (4.18) of section 4.1. The following additional notations will be required to write the capacitated version:

L_i = cumulative lead time for end product i , i.e., the shortest time necessary to manufacture a complete unit of end product when no components are available as initial inventory;

$\ell^{s}_{ij_tp}$ = amount of resource consumed by the elementary schedule j_t for end product i in period p at stage s.

If m_s is the resource consumption per unit of component produced at stage s, then:

$$\ell^{s}_{ij_tp} = \begin{cases} m_s & \text{if schedule } j_t \text{ requires, in period } p, \\ & \text{the production of the component part made} \\ & \text{by stage } s \text{ for product } i. \\ 0 & \text{otherwise} \end{cases} \qquad (4.83)$$

Example - for both end products in figure 4.3 of section 4.1 the cumulative lead time are 3 time periods. In figure 4.5 two elementary production schedules are shown; both produce end products to satisfy demand in period $t = 5$. Let the two schedules be number 1 in their respective sets; the two end products are number 6 and number 7, as shown in figure 4.3. The resource consumption factors are then:

- for product $i = 6$:

$\ell^{1}_{61_5 1} = m_1$; $\ell^{2}_{61_5 2} = m_2$;

$\ell^{3}_{61_5 3} = m_3$; $\ell^{4}_{61_5 3} = m_4$;

$\ell^{6}_{61_5 4} = m_6$; all other factors are zero .

- for product $i = 7$:

$$\ell^{2}_{71_5 2} = m_2 \; ; \qquad \ell^{4}_{71_5 3} = m_4 \; ;$$

$$\ell^{5}_{71_5 1} = m_5 \; ; \qquad \ell^{7}_{71_5 4} = m_7 \; ;$$

all other factors are zero.

* * * *

The capacitated multi-stage aggregate planning model is:

$$\min z = \sum_{i=1}^{N} \sum_{t=L_i+1}^{L_i+T} \sum_{j_t \in J_t^i} C_{ij_t} \Theta_{ij_t} \tag{4.84}$$

<u>s.t.</u>

$$\sum_{j_t \in J_t^i} \Theta_{ij_t} = d_{it} \; , \qquad \begin{cases} i = 1,\ldots,N \\ t = L_i+1,\ldots,L_i+T \end{cases} \tag{4.85}$$

$$\sum_{i=1}^{N} \sum_{t=L_i+1}^{L_i+T} \sum_{j_t \in J_t^i} \ell^{s}_{ij_t p} \Theta_{ij_t} \leq R^{s}_{p} \; , \quad \begin{cases} s = 1,\ldots, S \\ p = 1,\ldots, T \end{cases} \tag{4.86}$$

$$\Theta_{ij_t} \geq 0 \; , \qquad \text{all } ij_t \tag{4.87}$$

where R^{s}_{p} is the amount of productive resource available at stage s in period p .

Any extensions of this model to cover features mentioned for the F1 formulation are straightforward, and many of them have been shown in detail for the uncapacitated case. However, the issue of modelling backorders has not been brought up yet, in the F2 context, because no backorders are normally considered in the uncapacitated case. As mentioned in section 3.1.1, no backorders should be shown at any stage except at finished product level. Consequently, the demand equations have to be modified accordingly, and the corresponding cost of backordering has to be included in the objective function.

$$\min z = \sum_{i=1}^{N} \sum_{t=L_i+1}^{L_i+T} \sum_{j_t \in J_t^i} C_{ij_t}\, \theta_{ij_t} + \sum_{i=1}^{N} \sum_{t=L_i+1}^{L_i+T} b_{it}\, I_{it}^{-}$$

s.t. (4.88)

$$\sum_{j_t \in J_t^i} \theta_{ij_t} + I_{it}^{-} - I_{i,t-1}^{-} = d_{it}\,, \quad \begin{cases} i = 1,\ldots,N \\ t = L_i+1,\ldots,L_i+T \end{cases}$$

constraints (4.86); *nonnegativity*

where I_{it}^{-} is the level of backorders of finished product i at the end of period t and b_{it} is the cost of backordering one unit of i in period t. Notice that no positive inventories are shown, because they are implicitly considered by the production schedules.

When backorders are allowed in the formulation, and when also independent demand for component parts exist, some difficulties arise with respect to handling initial inventories of parts. Thus, suppose that the initial stock of some part is insufficient to cover both independent and

dependent demand for that part over its lead time; in such a case back-orders will occur by necessity. At this point of the planning process managerial input will be required to decide whether independent or dependent demand has to go unsatisfied, and to what extent.

Notice that in the capacitated formulations we kept using only elementary extreme flow schedules, although in the general case the aggregate solution will not be an extreme flow.
It is also clear that the capacitated model cannot be decomposed into independent Leontief subsystems because of the capacity constraints. these circumstances, it is not immediately obvious that it is sufficient to consider only extreme flow schedules, unless we can prove this to be true.

4.4.1. The Dominant Character of the Elementary Extreme Flow Schedules

The way to prove that it is sufficient to consider only extreme flow elementary schedules is to show that no improvement in the aggregate objective function can be achieved by also considering elementary non-extreme flow schedules.

Thus, let the aggregate planning problem be similar to (4.84) - (4.87), with the exception that a set $J_t^{i,inf}$ will be used instead of J_t^i J_t^i . $J_t^{i,inf}$ is the infinitely large set of elementary schedules, both extreme and non-extreme flow, that can supply one unit of product i to the market in period t .

$$\min z = \sum_{i=1}^{N} \sum_{t=L_i+1}^{L_i+T} \sum_{j_t \in J_t^{i,inf}} C_{ij_t} \theta_{ij_t} \tag{4.89}$$

s.t.

$$\sum_{j_t \in J_t^{i,inf}} \theta_{ij_t} = d_{it} , \quad \begin{cases} i = 1,\ldots,N \\ t = L_i+1,\ldots,L_i+T \end{cases} \tag{4.90}$$

$$\sum_{i=1}^{N} \sum_{t=L_i+1}^{L_i+T} \sum_{j_t \in J_t^{i,inf}} \ell^s_{ij_tp} \theta_{ij_t} \leq R^s_p , \tag{4.91}$$

$$s = 1,\ldots,S$$

$$p = 1,\ldots,T$$

$$\theta_{ij_t} \geq 0 , \qquad \text{all } ij_t \tag{4.92}$$

We have to show that, for any vector of dual variables π associated with constraints (4.90) - (4.91), the "most promising" column of all the columns in (4.89) - (4.92) is always an elementary extreme flow schedule. Formally, our problem is:

- find column $(ij_t)^o$ having minimum reduced cost:

$$(ij_t)^o: \quad \bar{C}_{(ij_t)^o} = \min_{\substack{\text{all pairs } (i,t) \\ i=1,\ldots,N \\ t=1,\ldots,T \\ j_t \in J_t^{i,inf}}} [\bar{C}_{ij_t}] \tag{4.93}$$

where

$$\bar{C}_{ij_t} = C_{ij_t} - \sum_s \sum_p \ell^s_{ij_t p} \pi_{sp} - \pi'_{it}$$

π_{sp} is associated with (4.91), and π'_{it} with (4.90); both are given.

- show that schedule j_t of column $(ij_t)^o$ is an extreme flow schedule, i.e.:

$$j_t \epsilon J^i_t \quad \text{, where } (i,t) \text{ are associated with column } (ij_t)^o \text{ in (4.93).}$$

We will conduct the search in (4.93) by pairs of (i,t):

$$(ij_t)^o : \bar{C}_{(ij_t)^o} = \min_{\left\{\begin{array}{l}\text{all pairs } (i,t) \\ i=1,\ldots,N \\ t=1,\ldots,T\end{array}\right.} \left[\min_{j_t \epsilon J^{i,\inf}_t} \bar{C}_{ij_t} \right]$$

The inner minimization consists of choosing the best schedule by which d_{it} units of product i will be made available to the market in period t; production cost in period t, at stage s, is $(v_{st} - m_s \pi_{st})$ per unit, and the holding cost is h_{st}.

Since the inner minimization is unconstrained, the result is an extreme flow schedule. There are NT inner minimizations, leading to

NT extreme flow schedules, of which one will have the lowest reduced cost; its corresponding column will be $(ij_t)^o$ of (4.93) . Thus, at any point during the solution of (4.89) - (4.92) , the "best" column is some elementary extreme flow schedule, and therefore there is no need to keep the non-extreme flow schedules around.

This result was to be expected for the following simple reason: the extreme flow schedules constitute the extreme points of the feasible region $\mathcal{F}_u$ of the uncapacitated problem; inside and on the boundaries of $\mathcal{F}_u$ there are non-extreme flow schedules. When capacity constraints are imposed, these will cut off parts of $\mathcal{F}_u$, creating a smaller feasible region $\mathcal{F}_c$. Since $\mathcal{F}_c \subset \mathcal{F}_u$, clearly any point in $\mathcal{F}_c$ can be represented as a convex combination of two or more extreme points of $\mathcal{F}_u$.

4.4.2. Column Generation by Dynamic Programming

As we have already mentioned, because of the very large number of columns in the F2 formulations, we suggest using a column generation technique for solution (Lasdon [63], Bradley et. al. [10], chapter 12). Thus, instead of considering all the columns in the capacitated aggregate model (4.84) - (4.87) , we start out with only a small subset of columns and solve this reduced LP . π_{it} and π_p^s are the optimal dual variables associated with constraints (4.85) and (4.86), respectively, in the reduced model.

The reduced costs are given by:

$$\bar{C}_{ij_t} = C_{ij_t} - \sum_s \sum_p \pi^s_p \ell^s_{ij_t p} - \pi_{it} \qquad (4.94)$$

The entering column is the one which has the smallest reduced cost, or

$$\min_i \min_t \min_{j_t} \bar{C}_{ij_t} \qquad (4.95)$$

By solving (4.95), a new elementary extreme flow schedule can be generated. (4.95) can be easily solved by dynamic programming.

We will describe the algorithm in the context of a simple example (figure 4.13). Let a be the period in which production of part 1 is scheduled to start, b the period in which part 2 is scheduled in production, c the production period for part 3 , d for part 4 , and e for the final assembly.

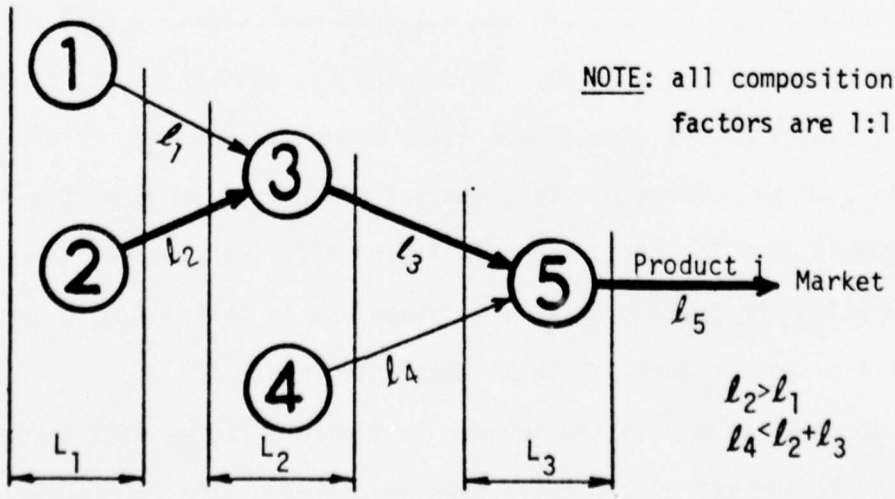


Fig. 4.13- A five stage system.

Suppose the end product assembled at stage 5 is the i-th product; the reduced cost (4.94) can then be written in more detail as:

$$\bar{C}_{ij_t} = (v_{1a} - \pi_a^1 m_1) + \sum_{p=a+\ell_1}^{c-1} h_{1p} + (v_{2b} - \pi_b^2 m_2) + \sum_{p=b+\ell_2}^{c-1} h_{2p} +$$

$$+ (v_{3c} - \pi_c^3 m_3) + \sum_{p=c+\ell_3}^{e-1} h_{3p} + (v_{4d} - \pi_d^4 m_4) + \sum_{p=d+\ell_4}^{e-1} h_{4p} +$$

$$+ (v_{5e} - \pi_e^5 m_5) + \sum_{p=e+\ell_5}^{t-1} h_{5p} - \pi_{it}$$

Notice that all $\pi_p^s \leq 0$, so the effect will be that variable production costs will be increased in those periods in which capacity is tight.

The task associated with generating a schedule is to decide upon the values of a , b , c , d , and e ; this can be easily done with a dynamic programming algorithm. To develop it, consider first the time-phasing relationships (see figure 4.14) among the 5 stages of product i . The critical path for completing the end product is shown with solid lines in figure 4.13. The solid lines in figure 4.14 indicate the time limits within which the processing at each stage has to take place, given that one unit of end product has to be available in period t .

In connection with the network in figure 4.13, we will define a level to be the set of all facilities separated by the same number of nodes

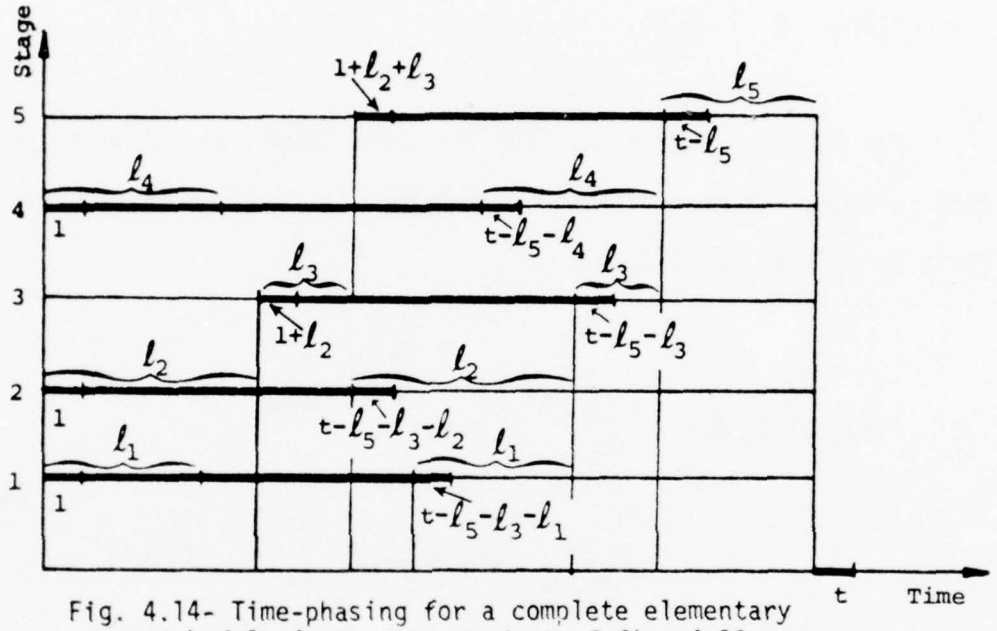


Fig. 4.14- Time-phasing for a complete elementary schedule j_t for the product of fig. 4.13.

from the finished good inventory (recall that a stocking point is assumed after each aggregate facility). Thus, level L_1 contains facilities 1 and 2 , level L_2 facilities 3 and 4 , and level L_3 the final assembly facility 5 (levels are numbered starting from raw materials towards the finished good).

The dyanmic programming will start from level L_1 and will proceed towards the last level L_3 .

Let $f_s(\cdot)$ be the minimal cost program for periods 1 to the beginning of period $(\cdot)$ at facility s ; also, $f_F(t)$ is the minimal cost program for periods 1 to the beginning of period t at the finished product stocking point. The state variables will be the numbers of the time periods in which production is started. Also let:

$$f_1(a) = 0 \quad ; \quad f_2(b) = 0 \quad ; \quad f_4(d) = 0$$

The functional equations for the three levels are shown below (in what follows, a summation whose lower limit is larger than the upper limit is defined to be zero):

$$\begin{aligned} f_3(c) \quad (1+\ell_2 \le c \le t-\ell_5-\ell_3) \;=\; & \min_{1 \le a \le c-\ell_1} \left[(v_{1a} - \pi_a^1 m_1) + \sum_{p=a+\ell_1}^{c-1} h_{1p} + f_1(a) \right] + \\ & + \min_{1 \le b \le c-\ell_2} \left[(v_{2b} - \pi_b^2 m_3) + \sum_{p=b+\ell_2}^{c-1} h_{2p} + f_2(b) \right] \end{aligned} \tag{4.96}$$

$$\begin{aligned} f_5(e) \quad (1+\ell_2+\ell_3 \le e \le t-\ell_5) \;=\; & \min_{1+\ell_2 \le c \le e-\ell_3} \left[(v_{3c} - \pi_c^3 m_3) + \sum_{p=c+\ell_3}^{e-1} h_{3p} + f_3(c) \right] + \\ & + \min_{1 \le d \le e-\ell_4} \left[(v_{4d} - \pi_d^4 m_4) + \sum_{p=d+\ell_4}^{e-1} h_{4p} + f_4(d) \right] \end{aligned} \tag{4.97}$$

$$f_F(t) = \min_{1+\ell_2+\ell_3 \le e \le t-\ell_5} \left[(v_{5e} - \pi_e^5 m_5) + \sum_{p=e+\ell_5}^{t-1} h_{5p} + f_5(e) \right] \tag{4.98}$$

The innermost minimization in (4.95) has i and t fixed; therefore, π_{it} is irrelevant to it. Consequently, the solution to (4.96) - (4.98) will be precisely the solution to the innermost minimization in (4.95).

The optimum a , b , c , d and e will define the new elementary extreme flow schedule for product i and period t . After repeating the algorithm for every i and t , the best schedule can be chosen as required by (4.95); if its reduced cost is negative it will be appended to the aggregate planning model. If no negative reduced cost can be found, the optimal aggregate solution has been obtained.

Under mild cost assumptions the search space can be reduced and, thus, the efficiency of the algorithm can be improved.

For some product i , let a_1 , b_1 , c_1 , d_1 , e_1 be the optimal solution for $t=t_1$, and a_2 , b_2 , c_2 , d_2 , e_2 the optimal solution for $t=t_2$. It is easy to show that, if:

$$h_{sp} \geq h_{s,p+1} \qquad \text{all } s,p \qquad (4.99)$$

then, in our example:

$$t_2 > t_1 \Rightarrow e_2 \geq e_1 \Rightarrow \begin{Bmatrix} d_2 \geq d_1 \\ c_2 \geq c_1 \end{Bmatrix} \Rightarrow \begin{Bmatrix} b_2 \geq b_1 \\ a_2 \geq a_1 \end{Bmatrix} \qquad (4.100)$$

Condition (4.99) is met in many practical cases; indeed, normally the holding cost is a percentage of the cost of that item, and in most situations the production cost is constant over the planning horizon or, if discounting is applied, it will decrease with time.

The proof can be done by contradiction and will be omitted. The

implication is that, when a new schedule is generated, we should start out with the smallest t so that, with increasing t, we can take advantage of (4.100).

When a special schedule (one which uses one or several parts from stock), rather than a complete schedule, has to be generated, the only thing that changes is the time-phasing diagram, because the parts in stock do not have to be produced. The dynamic programming algorithm remains, in principle, the same. Notice that the minimization of type (4.95) has to be carried over for every class (or set) of complete and special schedules.

In concluding, let us mention that, in case the composition factors are not all 1:1, the unit production cost and holding cost in the functional equations will have to be adjusted in proportion to the number of parts required per unit of end product.

4.4.3. Other Ways of Reducing the Computational Effort

4.4.3.1. Multiple Pricing

One way by which one can try to cut down the number of computations involved in the solution by column generation is multiple pricing. In solving (4.95), a total of NT schedules are generated, T schedules for each end product. Of all these, the minimum cost schedule is supposed to be chosen and appended to the aggregate planning model, and the others will be discarded; this is a single pricing strategy.

It is clear that this is not the most economical procedure. Lasdon and Terjung [64], who solved Manne's [69] problem, reported good computa-

tional results and prospects of faster convergence with the use of multiple pricing. In our case, multiple pricing will be applied as follows:

- Generate NT new columns, given a vector π of dual variables.
- The column with the most negative reduced cost will be appended to the reduced aggregate model, and will enter the basis; the other columns will not be discarded.
- Solve the new aggregate planning model; a new vector π' of dual variables will be obtained.
- Price out the columns not used so far, and have the one with the most negative reduced cost enter the basis. The reason for this step is given by Lasdon and Terjung [64], p. 951: "Often π and π' will not differ radically, so that some of columns...may again price out negative".
- Repeat the procedure until no column (of the initial NT generated) will price out negative.
- Generate another set of NT columns and start out from the beginning. In order to keep the LP from growing too large, the non-basic columns can be dropped. Should there be later prospects for them to become candidates to enter the basis, they will be re-generated automatically.

4.4.3.2. Revised F2 Formulations

We have already mentioned (section 4.3.6), and have been concerned with the large number of classes (or sets) of elementary extreme flow schedules in problems with initial inventories and deliveries over the

lead times. If the number of sets is large, the computational effort involved in generating the best column can become substantial. A reinterpretation of the initial inventories, however, will help us reduce the number of sets to one only, without any increase in the number of constraints.

Consider, again, the simple product shown in figure 4.8 and reillustrated here for convenience (figure 4.15). There are initial inven-

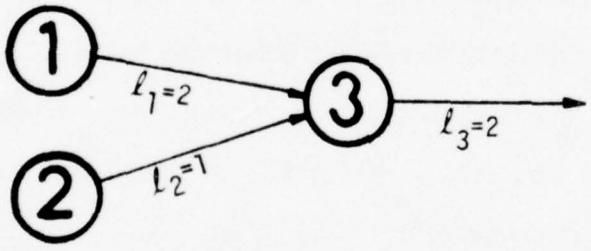


Fig. 4.15- End product with two component parts.

tories I_{10} and I_{20} ; there are also scheduled deliveries of part 1 over its 2 period lead time, DEL_1^1 and DEL_2^1 , and of part 2 over its one period lead time, DEL_1^2 . In section 4.3.3, it has been shown that, besides the complete elementary schedules, 5 sets of special schedules have to be considered if the resulting F2 formulation is to be equivalent to the original F1 formulation. The sets of special schedules represent, in fact, all the possible combinations of part 1 and 2 from stock for assemblying product 3 .

Let us, however, think of inventories as productive resources; for instance, one unit of part 1 in stock at the beginning of period 1 is

equivalent to considering that the amount of resource required to produce that unit of part 1 was available and was consumed one lead time before, i.e., in period -2 . Thus, an appropriate number of past time periods is brought into the picture: two periods for part 1 and one period for part 2 . Their resource availabilities are easily computed by the use of the m_1 and m_2 resource consumption factors, see model (1.1) - (1.7):

$$R^1_{-2} = m_1(I_{10} + DEL^1_1)$$

$$R^1_{-1} = m_1\ DEL^1_2$$

$$R^2_{-1} = m_2\ (I_{20} + DEL^2_1)$$

The resource availabilities R^s_t in all other periods remain unchanged so that the revised capacitated F2 formulation is:

$$\min z = \sum_{t=3}^{T+2} \sum_{j_t \in J_t^3} C_{3j_t}\, \theta_{3j_t} \tag{4.101}$$

s.t.

$$\sum_{j_t \in J_t^3} \theta_{3j_t} = d_{3t}\,, \quad t=3,\ldots,T+2 \tag{4.102}$$

$$\sum_{t=3}^{T+2} \sum_{j_t \in J_t^3} \ell^s_{3j_t p}\, \theta_{3j_t} \leq R^s_p\,, \quad \begin{cases} s=1,\ldots,3 \\ p=1,\ldots,T \end{cases} \tag{4.103}$$

$$\sum_{t=3}^{T+2} \sum_{j_t \in J_t^3} \ell^s_{3j_t,(-2)}\, \theta_{3j_t} = R^1_{-2} \tag{4.104}$$

$$\sum_{t=3}^{T+2} \sum_{j_t \in J_t^3} \ell^1_{3j_t,(-1)}\, \theta_{3j_t} = R^1_{-1} \tag{4.105}$$

$$\sum_{t=3}^{T+2} \sum_{j_t \in J_t^3} \ell^1_{3j_t,(-1)}\, \theta_{3j_t} = R^2_{-1} \tag{4.106}$$

$$\text{nonnegativity} \tag{4.107}$$

It is important to note that of all sets of schedules, only the complete schedules have still to be considered in the above F2 formulation. Thus, a substantial reduction in the number of columns, and in the computational effort for the generation of columns, is achieved.

Several remarks are necessary:

- The initial inventories and deliveries of aggregate parts have to be re-balanced, as discussed in section 3.1.3, before they can be converted into equivalent productive capacities.
- No increase in the number of constraints takes place with the revised formulation compared to the "older" F2 formulation; rather, a reduction might be achieved (compare the capacitated version of (4.63) - (4.68) with (4.101) - (4.107), both representing the same problem).
- Since no production cost is charged for a part used from stock, attention should be paid to the cost computations for those schedules that "produce" parts in time periods prior to $t=1$, i.e., in periods with (-) sign. A zero variable production cost v_{st} has to be charged for production in those periods; however, the regular inventory holding costs would apply after the part is finished.
- Notice that, equivalently to constraints (4.65) - (4.68), constraints (4.104) - (4.106) are equalities. This ensures that all initial stock is completely consumed. Moreover, the imposed equality is reasonable and there should be no difficulties in achieving it. Indeed, since no production cost is charged for using the resources in "past" periods, the column generation procedure will tend to favor those schedules "producing" parts in those periods.
- Because of the increase in the number of time periods, an increase in the computational effort per column generated by dynamic programming is expected.

However, this increase is likely to be quite limited, as in most cases condition (4.99) is fulfilled. Also, the increase will be more than offset by the reduction in the number of classes of schedules to be generated.

- Extensions of (4.101) - (4.107) to include multiple products, multiple resources, etc. are straightforward.

We would like to mention that we have not introduced this simplification earlier, in order to allow a more complete treatment of the nature of the F2 formulations and all the issues involved.

4.4.4. Concluding Comments

It is apparent that in the F2 formulations inventories "flow" within the elementary production schedules rather than being represented by explicit variables in the model. However, it is a trivial matter to retrieve all the production and inventory data from the solution to a F2 aggregate model.

In terms of computational requirements, the F2 formulation has normally fewer constraints than the F1 formulation, and in any case the constraints have special structure. Thus, if we re-consider the case with 20 facilities (S=20) having one limited resource each, 12 time periods (T=12), and an average of 30 items produced per facility, the number of constraints involved in a fixed work force model with overtime allowance would be:

- F1 formulations (see (3.1) - (3.6) in section 3.1):

7200 inventory balance equations,

240 capacity constraints,

240 overtime constraints.

- F2 formulations (consider an extension of (4.101) - (4.107)); assume there is independent demand for every aggregate part:

7200 demand equations of the Generalized Upper Bounding type,

240 overtime constraints of the Generalized Upper Bounding type,

240 capacity constraints,

Constraints of the (4.104) - (4.106) type whose number depends on how long the lead times are.

Thus, the F2 formulation appears to offer a computational advantage, when compared to F1, because of two reasons:

- F1 has many constraints with no special structure and the computational effort , when solving LP's, grows with the third power of the number of constraints;
- F2 has many constraints of the Generalized Upper Bounding type, that can be better handled (up to 50,000) than constraints with no special structure.

Moreover, in this example F1 is obviously infeasible, while F2 is not.

CHAPTER 5 - ECONOMIC LOT SIZE DETERMINATION IN MULTI-STAGE SYSTEMS

The computation of economic lot sizes is part of the disaggregation process, by which implementable production plans are obtained and the cost structure is completed by bringing the setup costs into the picture.

A well established policy in multi-stage lot sizing has been to consider the economic order quantity (EOQ) at some stage s as a positive integral multiple of the EOQ at the immediate successor stage. Crowston and Wagner [23], and Crowston et. al. [25] have shown that this policy is optimal in the case of pure assembly systems.

In this thesis we have adopted a broader view by considering a more general case of multi-stage system, namely assembly systems with diverging arcs. We will address the computation of the EOQ's in such system for two cases:

- the case where there is independent demand for the end product only;
- the case where there is also independent demand for some component parts.

We should mention that, in fact, in multi-stage systems there are two general classes of lot size models (Krajewski [59]): one class that does not recognize dependencies between stages, and another class in which these dependencies constitute the heart of the matter. The former category is currently used in MRP systems, Berry [6]; we will discard it because it ignores the very nature of the system under study, that is the multi-stage aspect.

5.1. The Case with Independent Demand for the End Product Only

The assumptions under which we will conduct the derivation of the optimal lot sizing policy are the "normal" assumptions made in EOQ research:

- Single product.
- Uncapacitated production process.
- Infinite planning horizon.[1]
- Constant demand for the end product; the end product is withdrawn from finished good inventory one unit at a time, at a constant rate.
- Instantaneous production.
- Zero lead times.
- Time-invariant cost functions: setup cost and inventory holding cost.
- Time-invariant lot sizes at all stages.
- No shortages allowed.
- At moment zero all stages produce their respective lot sizes.
- A stage can start processing only when all input materials required for the production of an EOQ are available at that stage. Raw materials required by the first level stages are assumed to be available in unlimited supply.

---

[1] Schwarz [80], p. 555 justifies why an infinite planning horizon is desirable.

THE LOT SIZING PROBLEM is: find the EOQ's that minimize the average cost per unit produced.

Several terms and notations have to be defined:

The echelon stock introduced by Clark [15]) of stage s is the number of units of the part produced at stage s which have passed through stage s but have not been sold yet (i.e., they are still in the system).

The installation stock of stage s is the amount of inventory stored at the stocking point immediately following stage s.

Corresponding to the two concepts just introduced, we will have for stage s two kinds of inventory holding costs:

- the installation inventory holding cost, h_s, which is proportional to the total value of the part made at stage s;
- the echelon inventory holding cost, H_s, which is proportional to the value added at stage s; the value added is assumed to be nonnegative at each stage.

$a(s)$ = the set of immediate successors of stage s.

$A(s)$ = the set of all successors of stage s.

$b(s)$ = the set of immediate predecessors of stage s.

$B(s)$ = the set of all predecessors of stage s.

a_{sq} = composition factor.

$A_{sS} = \sum_{q \in a(s)} a_{sq} \cdot A_{qS}$ = number of units of stage s output per unit of end product.

Q_s = lot size at stage s .

In a pure assembly system every facility responds to requirements from one successor only; hence, the "basic cell" of the pure assembly systems is:

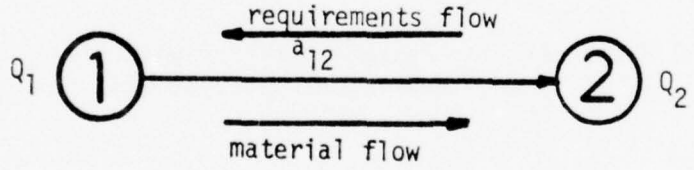


Fig. 5.1- Basic cell for pure assembly systems (basic cell of type I).

In structures with diverging arcs another basic cell has to be considered (see figure 5.2); notice that the cell of type I can be viewed as a special case of the basic cell of type II. Any assembly system with diverging arcs can be structured from cells of type II.

For ease of later reference, let us show here (figure 5.3) the simplest example of cell of type II, i.e., one that has only two immediate successors.

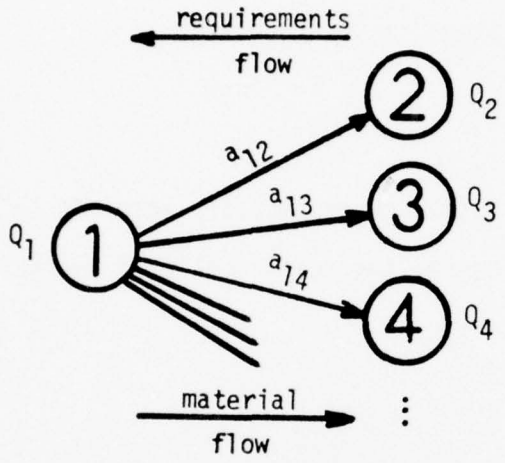


Fig. 5.2- Basic cell for assembly systems with diverging arcs (basic cell of type II).

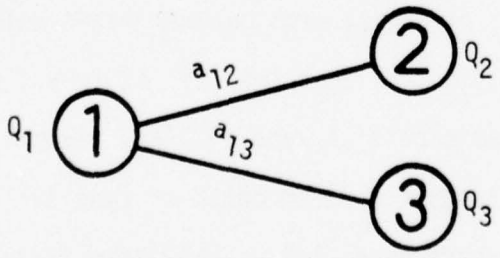


Fig. 5.3- The simplest basic cell of type II (only two immediate successors).

We would like to point out that we have made no assumption about having zero initial inventories; as it will be shown later, there might arise situations in which a permanent stock has to be held at some stage in order to prevent a stockout situation. Then, this permanent stock will have to be there from the very beginning, as initial inventory. We chose this approach because it is richer and would permit us to gain more insight into the structure of the cost functions involved. However, at the end we will also specialize our results for the case of no initial inventories.

PROPOSITION 5.1 - Let stage s have one or several immediate successors. Then, an optimal solution to the lot sizing problem, with lot size rational numbers, has the property that there are points in time, occurring with some periodicity P, at which production takes place simultaneously at stage s and at all its immediate successors.

Proof. For simplicity assume stage s has just two immediate successors (see figure 5.3). By assumption, in time period 1 (moment zero) all stages have to produce.

In an optimal solution, at any stage successive batches of production cover non-overlapping time intervals, that is,one lot is produced after the previous one has been depleted.

By assumption, the lot size at a stage is a rational number, constant over time; the composition factors are also rational (see section 3.2.2); therefore, we can find three positive integers q_1, q_2, q_3,

with no common divisor, satisfying the following equalities:

$$t = 1 + q_1 \cdot \frac{Q_1}{A_{1S}D} = 1 + q_2 \cdot \frac{Q_2}{A_{2S}D} = 1 + q_3 \cdot \frac{Q_3}{A_{3S}D}$$

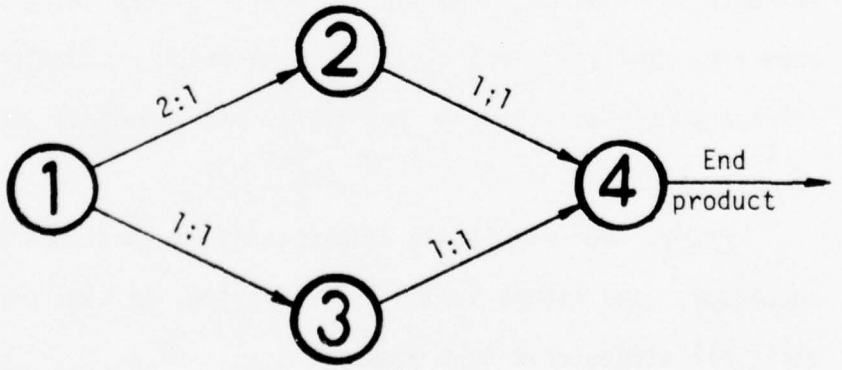
Then, t is the point in time at which production takes place at all three stages simultaneously.

The periodicity of these points is given by:

$$P = q_1 \cdot \frac{Q_1}{A_{1S}D} = q_2 \cdot \frac{Q_2}{A_{2S}D} = q_3 \cdot \frac{Q_3}{A_{3S}D}$$

Clearly, P is also the period with which the installation stock at stage 1 cycles.

Example



D = 10 units/period;

$Q_1 = 72$; $Q_2 = 30$; $Q_3 = 20$;

$A_{1S} = 3$; $A_{2S} = 1$; $A_{3S} = 1$.

For $q_1 = 5$, $q_2 = 4$, $q_3 = 6$ one obtains:

$$P = 5 \cdot \left(\frac{72}{30}\right) = 4 \cdot \left(\frac{30}{10}\right) = 6 \cdot \left(\frac{20}{10}\right) = 12$$

Thus, in periods 1 , 13 , 25 , 37,... production occurs simultaneously at stages 1, 2 , 3 ; in the general case, production takes place simultaneously at stage s and all its immediate successors in periods 1 , P+1 , 2P+1,... ‖

The extension of the above proof to the case with just one successor or with more than two successors is straightforward.

PROPOSITION 5.2 - In a two level system composed of a stage s with several immediate successors s_1 , s_2 , s_3 ,..., s_n , with the successor lot sizes held constant at values Q_{s_1} , Q_{s_2} ,..., Q_{s_n} , respectively (where Q_{s_1} , Q_{s_2} ,... are positive rational numbers), the total cost per unit time associated with a lot size Q_s (positive rational number) at stage s is:

$$Z(Q_s) = \frac{B_s D A_{sS}}{Q_s} + H_s \cdot \frac{Q_s - A_{sS}}{2} + h_s \cdot \max_t(\bar{I}_t)$$

where $\bar{I}(t) = -I(t)$ and $I(t)$ is the installation inventory at stage s as a function of time.

Proof. For simplicity assume only two successors and refer to figure 5.3. The total cost at stage 1 has three components:

- The setup cost $\frac{B_1 DA_{1S}}{Q_1}$ where B_1 is the setup cost per lot at stage 1 .
- The echelon inventory holding cost arising from the well known "saw tooth" pattern of stock replenished periodically: $H_1 \cdot \frac{Q_1 - A_{1S}}{2}$; A_{1S} is subtracted because, by assumption, the end product is withdrawn from stock one unit at a time, which results in A_{1S} units of echelon stock of stage 1 being withdrawn at a time.
- The cost of holding an initial permanent inventory at stage 1 to prevent running out of stock. Indeed, if the lot sizes Q_1 , Q_2 , Q_3 are not well coordinated a stock-out situation may develop.

Example

Consider again the example of Proposition 5.1, and let us follow the fluctuations of the installation inventory level at stage 1 during a cycle of P=12 periods. In this particular example we would have to continuously keep in inventory 52 units of part 1 to prevent the stock-out at the beginning of period 10. Recall that lot sizes are assumed constant; therefore it is not possible to increase, preventively, the batch size just before a stockout would occur (see fig. 5.4).

The inventory at stage 1 , at any point t in time, can be computed with the following equation:

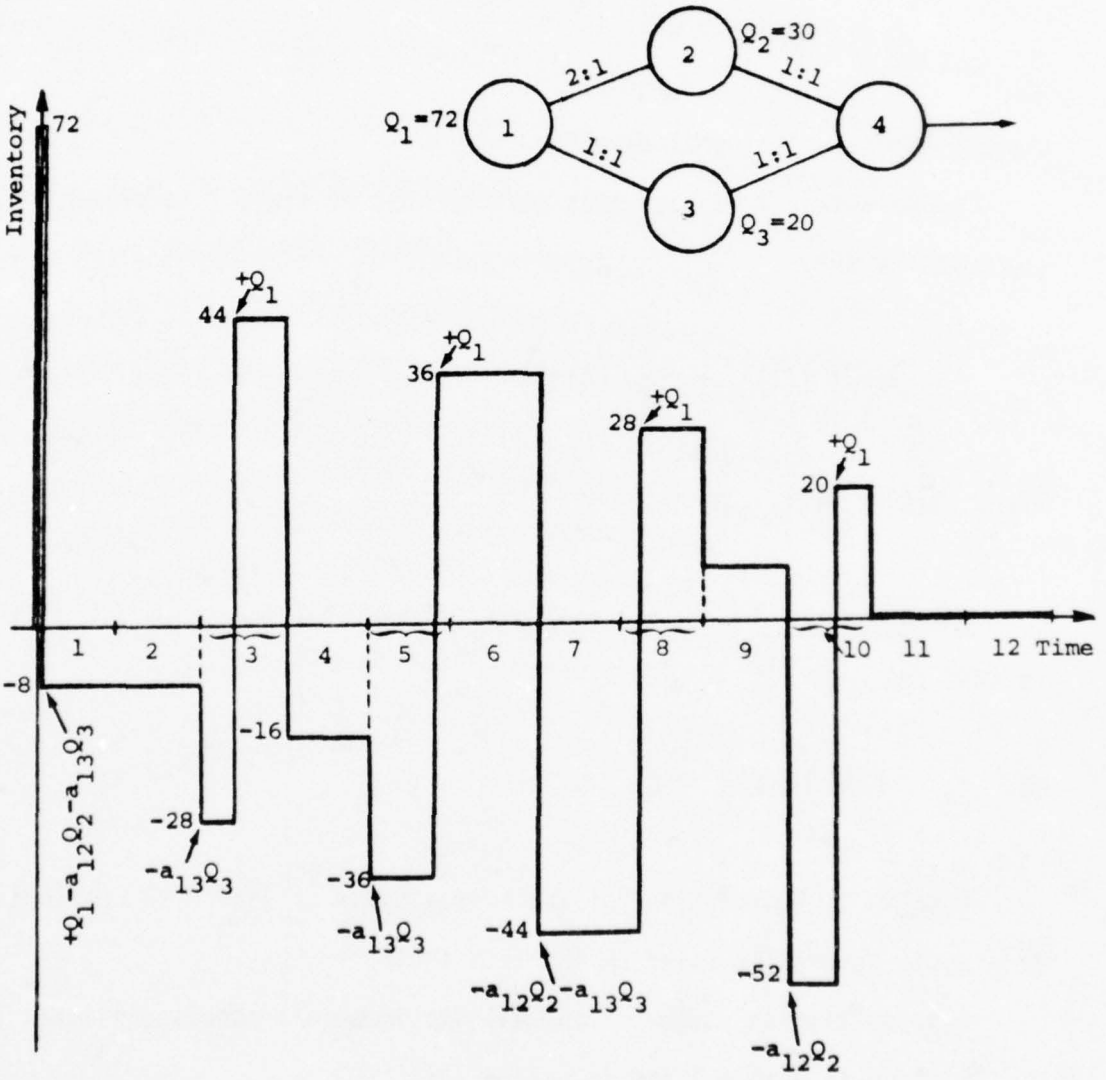


Fig. 5.4- Fluctuations of the installation stock at stage 1.

$$I(t) = \left[\frac{tA_{1S}D}{Q_1}\right] \cdot Q_1 - \left[\frac{tA_{2S}D}{Q_2}\right] \cdot Q_2 a_{12} - \left[\frac{tA_{3S}D}{Q_3}\right] \cdot Q_3 a_{13} +$$

$$+ Q_1 - Q_2 a_{12} - Q_3 a_{13}$$

where [] denotes "integral part" .

The permanent inventory that must be kept at stage 1 to prevent stockouts is then:

$$-\min_t \left(I(t) \right) = \max_t \left(\bar{I}(t) \right)$$

where $\bar{I}(t) = -I(t)$. ‖

PROPOSITION 5.3 - In the two level system of Proposition 5.2 , it is true that

$$\min_t \left(I(t) \right) \leq 0 .$$

Proof. By Proposition 5.1 there is a cycle of length P periods. This cycle is characterized by the fact that:

- production at stage s and all its immediate successors takes place in period 1 and in period P+1 ;
- the total amount produced by stage s during the cycle exactly equals the sum of the amount required by all immediate successors during the cycle.

The first statement follows directly from the definition of the cycle.

The second statement is derived as follows: the periodicity of the cycle is:

$$P = q_s \cdot \frac{Q_s}{A_{ss}D} = q_{s1} \cdot \frac{Q_{s1}}{A_{s_1 s}D} = \ldots = q_{s_n} \cdot \frac{Q_{s_n}}{A_{s_n s}D}$$

where $s_1, \ldots, s_n$ is the set of all immediate successors of s, and $q_s, q_{s_1}, \ldots, q_{s_n}$ have no common divisor. Let us note that q_s, $q_{s_1}, \ldots, q_{s_n}$ represent the number of batches produced during the cycle at stage s, $s_1, \ldots, s_n$ respectively.

$$q_s \cdot Q_s = q_{s_1} \cdot Q_{s_1} \cdot \frac{A_{ss}}{A_{s_1 s}}$$

But $A_{ss} = a_{ss_1} \cdot A_{s_1 s} + a_{ss_2} \cdot A_{s_2 s} + \ldots + a_{ss_n} \cdot A_{s_n s}$

Thus:

$$q_s \cdot Q_s = q_{s_1} \cdot Q_{s_1} \cdot a_{ss_1} + q_{s_1} \cdot Q_{s_1} \cdot \frac{A_{s_2 s}}{A_{s_1 s}} \cdot a_{ss_2} + \ldots + q_{s_1} \cdot Q_{s_1} \cdot \frac{A_{s_n s}}{A_{s_1 s}} \cdot a_{ss_n}$$

From the set of equalities defining P we have that:

$$q_{s_1} Q_{s_1} \cdot \frac{A_{s_2 S}}{A_{s_1 S}} a_{ss_2} = q_{s_2} \cdot Q_{s_2} \cdot a_{ss_2}$$

$$\vdots \qquad\qquad \vdots$$

$$q_{s_1} \cdot Q_{s_1} \cdot \frac{A_{s_n S}}{A_{s_1 S}} \cdot a_{ss_n} = q_{s_n} \cdot Q_{s_n} \cdot a_{ss_n}$$

Therefore:

$$q_s \cdot Q_s = q_{s_1} \cdot Q_{s_1} \cdot a_{ss_1} + q_{s_2} \cdot Q_{s_2} \cdot a_{ss_2} + \ldots + q_{s_n} \cdot Q_{s_n} \cdot a_{ss_n} ,$$

so the second statement above is justified.

Consequently, at the end of period P the installation stock at stage s is zero. Hence, the minimum value of the installation stock $I(t)$ at stage s cannot possibly be larger than zero during the cycle; however, it can be negative;

$$\Rightarrow \min (I(t)) \leq 0 . \|$$

PROPOSITION 5.4 - For the purpose of describing the time behavior of the installation stock $I(t)$ at stage s, the two level system of Proposition 5.2 can be decomposed into n two stage subsystems. The j^{th} subsystem $(1 \leq j \leq n)$ will be composed of stages s and s_j; the number of parts s required per unit of subassembly s_j is considered to be $a'_{ss_j} = \frac{A_{sS}}{A_{s_j S}}$. If $I_j(t)$ is the installation stock function at stage s in the j^{th} subsystem then:

$$I(t) = \frac{a_{ss_1}}{a'_{ss_1}} \cdot I_1(t) + \frac{a_{ss_2}}{a'_{ss_2}} \cdot I_2(t) + \cdots + \frac{a_{ss_n}}{a'_{ss_n}} \cdot I_n(t)$$

Proof. The installation inventory at stage s in the system of Proposition (5.2) is:

$$I(t) = \left[\frac{tA_{sS}D}{Q_s}\right] \cdot Q_s - \left[\frac{tA_{s_1S}D}{Q_{s_1}}\right] \cdot Q_{s_1} \cdot a_{ss_1} - \cdots - \left[\frac{tA_{s_nS}D}{Q_{s_n}}\right] \cdot Q_{s_n} \cdot a_{ss_n} +$$

$$+ Q_s - Q_{s_1} \cdot a_{ss_1} - \cdots - Q_{s_n} \cdot a_{ss_n}$$

The installation inventory at stage s in the $j^{\underline{th}}$ subsystem is:

$$I_j(t) = \left[\frac{tA_{sS}D}{Q_s}\right] \cdot Q_s - \left[\frac{tA_{s_jS}D}{Q_{s_j}}\right] \cdot Q_{s_j} \cdot a'_{ss_j} + Q_s - Q_{s_j} \cdot a'_{ss_j}$$

Then:

$$\frac{a_{ss_1}}{a'_{ss_1}} \cdot I_1(t) + \frac{a_{ss_2}}{a'_{ss_2}} \cdot I_2(t) + \cdots + \frac{a_{ss_n}}{a'_{ss_n}} \cdot I_n(t) =$$

$$= a_{ss_1} \cdot \frac{A_{s_1S}}{A_{sS}} \cdot I_1(t) + a_{ss_2} \cdot \frac{A_{s_2S}}{A_{sS}} \cdot I_2(t) + \cdots + a_{ss_n} \cdot \frac{A_{s_nS}}{A_{sS}} \cdot I_n(t) =$$

$$= \frac{a_{ss_1}A_{s_1S} + \cdots + a_{ss_n}A_{s_nS}}{A_{sS}} \cdot \left[\frac{tA_{sS}D}{Q_s}\right] \cdot Q_s -$$

$$- \sum_{j=1}^{n} \left[\frac{tA_{s_j S^D}}{Q_{s_j}} \right] \cdot Q_{s_j} \cdot a'_{ss_j} \cdot \frac{a_{ss_j}}{a'_{ss_j}} + \frac{a_{ss_1} \cdot A_{s_1 S} + \cdots + a_{ss_n} \cdot A_{s_n S}}{A_{sS}} \cdot Q_s -$$

$$- \sum_{j=1}^{n} Q_{s_j} \cdot a'_{ss_j} \cdot \frac{a_{ss_j}}{a'_{ss_j}} = I(t) \quad \|$$

PROPOSITION 5.5 - In the system of Proposition 5.2 there are subcycles of length $P' \leq P$ (P defined in Proposition 5.1); in the first period of every such subcycle production takes place simultaneously at all immediate successors of stage s.

Proof. By reason of $Q_{s_1}, Q_{s_2}, \ldots, Q_{s_n}$ being rational numbers we can find n positive integers $\alpha_1, \alpha_2, \ldots, \alpha_n$, with no common divisor, defining the periodicity P' of the subcycle:

$$P' = \alpha_1 \frac{Q_{s_1}}{A_{s_1 S^D}} = \alpha_2 \frac{Q_{s_2}}{A_{s_2 S^D}} = \ldots = \alpha_n \frac{Q_{s_n}}{A_{s_n S^D}}$$

Production takes place simultaneously at all immediate successors $s_1, s_2, \ldots, s_n$ in and only in periods $1, P'+1, 2P'+1, 3P'+1$, etc.

Example

Consider the example of proposition 5.1. We have already seen that there is a cycle $P=12$ periods. The subcycle P' is obtained for $\alpha_1 = 2$, $\alpha_2 = 3$:

$$P' = 2\cdot\left(\frac{30}{10}\right) = 3\cdot\left(\frac{20}{10}\right) = 6$$

In periods 1 , 7 , 13 , 19 , ... production occurs simultaneously at stages 2 and 3 . ‖

LEMMA 5.6 - The cost function $Z(Q_s)$ associated with the choice of Q_s in the two level system of Proposition 5.2 has its minimum approximated by:

$$Q_s^* = k(\alpha_j Q_{s_j} \cdot \frac{A_{sS}}{A_{s_j S}}) , \qquad \text{any } j\varepsilon\{1,\ldots,n\}$$

where k is a positive integer, and α_j is defined in the proof of Proposition 5.5.

Proof. We will develop two bounding functions for $Z(Q_s)$; then, we will show that for $Q_s = Q_s^*$ function $Z(Q_s)$ is equal to its bounds. The proof contains five parts.

Part a - The Upper Bounding Function.

Let us decompose the two level system of Proposition 5.2 into n subsystems as shown in Proposition 5.4 . For the $j^{\underline{th}}$ subsystem, the total cost per unit time is:

$$Z_j(Q_s) = \frac{B_s DA_{sS}}{Q_s} + H_s \cdot \frac{Q_s - A_{sS}}{2} + h_s \cdot \max_t (\bar{I}_j(t))$$

where $I_j(t)$ was defined in Proposition 5.4 and $\bar{I}_j(t) = -I_j(t)$.

If we define $U(Q_s) = \sum_{j=1}^{n} \frac{a_{ss_j}}{a'_{ss_j}} Z_j(Q_s)$ we get:

$$U(Q_s) = \frac{\sum_{j=1}^{n} a_{ss_j} A_{s_j S}}{A_{sS}} \cdot \frac{B_s DA_{sS}}{Q_s} + \frac{\sum_{j=1}^{n} a_{ss_j} A_{s_j S}}{A_{sS}} \cdot H_s \cdot \frac{Q_s - A_{sS}}{2} +$$

$$+ h_s \sum_{j=1}^{n} \frac{a_{ss_j}}{a'_{ss_j}} \max_t (\bar{I}_j(t))$$

$$U(Q_s) = \frac{B_s DA_{sS}}{Q_s} + H_s \cdot \frac{Q_s - A_{sS}}{2} + h_s \cdot \sum_{j=1}^{n} \max_t \left(\frac{a_{ss_j}}{a'_{ss_j}} \bar{I}_j(t) \right)$$

Now recall that by proposition 5.4:

$$\bar{I}(t) = \sum_{j=1}^{n} \frac{a_{ss_j}}{a'_{ss_j}} \bar{I}_j(t)$$

Therefore, the $Z(Q_s)$ cost function can be written as:

$$Z(Q_s) = \frac{B_s DA_{sS}}{Q_s} + H_s \cdot \frac{Q_s - A_{sS}}{2} + h_s \cdot \max_t (\bar{I}_t) =$$

$$= \frac{B_s DA_{sS}}{Q_s} + H_s \cdot \frac{Q_s - A_{sS}}{2} + h_s \cdot \max_t \left(\sum_{j=1}^{n} \frac{a_{ss_j}}{a'_{ss_j}} \bar{I}_j(t) \right)$$

As the sum of the maximums is at least as large as the maximum of the sum, a comparison between $U(Q_s)$ and $Z(Q_s)$ leads to the following result:

$$U(Q_s) \geq Z(Q_s)$$

Part b - The Lower Bounding Function.

By Proposition 5.3 $\min_t (I(t)) \leq 0$; hence:

$$\max_t (\bar{I}(t)) = -\min_t (I(t)) \geq 0 .$$

As all h_s , a_{ss_j} , a'_{ss_j} are positive, we can write:

$$Z(Q_s) \geq \frac{B_s DA_{sS}}{Q_s} + H_s \cdot \frac{Q_s - A_{sS}}{2}$$

or, if we let:

$$W(Q_s) = \frac{B_s D A_{sS}}{Q_s} + H_s \cdot \frac{Q_s - A_{sS}}{2}$$

then:

$$Z(Q_s) \geq W(Q_s)$$

Part c - Consider the $j^{\underline{th}}$ subsystem, as defined by Proposition 5.4:

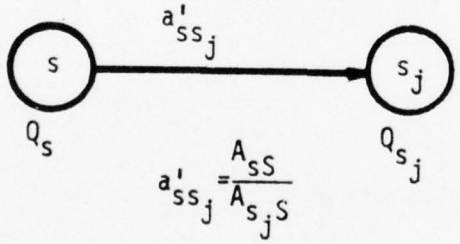


Fig. 5.5- The j-th subsystem.

In this subsystem there is a cyclical variation of the installation stock between stages s and s_j, with periodicity:

$$P_j = \beta_s \cdot \frac{Q_s}{A_{sS} D} = \beta_{s_j} \cdot \frac{Q_{s_j}}{A_{s_j S} D} \leq P$$

where β_s, β_{s_j} are relatively prime positive integers. Production takes place simultaneously at stages s and s_j in periods 1, P_j+1, $2P_j+1$, $3P_j+1$, etc. Let us note that if $P_j < P$ then P is an integral multiple of P_j.

The permanent installation stock that has to be held, between stage s and s_j , in order to prevent a stockout is (adapted from Crowston et. al. [25], p. 525):

$$\max_t \left(\bar{I}_j(t)\right) = Q_{s_j} a'_{ss_j} \cdot \frac{\beta_s - 1}{\beta_s}$$

If Q_s is chosen to be an integral multiple of $Q_{s_j} \cdot \frac{A_{sS}}{A_{s_jS}}$, then obviously $\beta_s = 1$ and $\max_t \left(\bar{I}_j(t)\right) = o$.

Suppose now that Q_s^* is some multiple of $\alpha_j Q_{s_j} \cdot \frac{A_{sS}}{A_{s_jS}}$:

$$Q_s^* = k\left(\alpha_j Q_{s_j} \cdot \frac{A_{sS}}{A_{s_jS}}\right), \quad \text{any } j \in \{1, 2, \ldots, n\},$$
$$\text{and } k = \text{positive integer.}$$

In this case, as k and α_j are integers, Q_s^* is also a multiple of $Q_{s_j} \cdot \frac{A_{sS}}{A_{s_jS}}$ so that , by the result mentioned above we have:[1]

---

[1] Recall that by Proposition 5.5 an integral multiple of $\alpha_j Q_{s_j} \cdot \frac{A_{sS}}{A_{s_jS}}$ is also an integral multiple of $\alpha_1 Q_{s_1} \cdot \frac{A_{sS}}{A_{s_1S}}$, $\alpha_2 Q_{s_2} \cdot \frac{A_{sS}}{A_{s_2S}}$, etc.

$$Z_j(Q_s^*) = \frac{B_s DA_{sS}}{Q_s^*} + H_s \cdot \frac{Q_s^* - A_{sS}}{2}, \qquad \text{all } j \in \{1, 2, \ldots, n\}.$$

Also, for $Q_s = Q_s^*$ the two bounding functions become:

$$U(Q_s^*) = \sum_{j=1}^{n} \frac{a_{ss_j}}{a'_{ss_j}} Z_j(Q_s^*) = \frac{B_s DA_{sS}}{Q_s^*} + H \cdot \frac{Q_s^* - A_{sS}}{2}$$

and

$$W(Q_s^*) = \frac{B_s DA_{sS}}{Q_s^*} + H_s \cdot \frac{Q_s^* - A_{sS}}{2}$$

so that, evidently $U(Q_s^*) = W(Q_s^*)$.

Since by part a and b:

$$U(Q_s) \geq Z(Q_s) \geq W(Q_s)$$

it immediately follows that

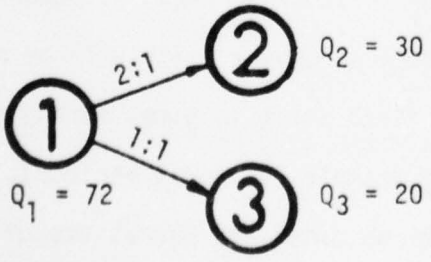
$$Z(Q_s^*) = U(Q_s^*) = W(Q_s^*) = \frac{B_s DA_{sS}}{Q_s^*} + H_s \cdot \frac{Q_s^* - A_{sS}}{2}$$

Part d - Here we will show that for any Q_s other than Q_s^* of part c function $Z(Q_s)$ lies strictly above its lower bounding function:

$$Z(Q_s) > W(Q_s) .$$

By Proposition 5.5 periods 1 , P'+1 , 2P'+1,... are the only points in time when production occurs at all immediate successors of s simultaneously. $\alpha_j Q_{s_j} \cdot \frac{A_{sS}}{A_{s_j S}}$, $j \in \{1,2,\ldots,n\}$, gives us precisely the number of units that must be produced by stage s to cover the requirements for parts s during one subcycle P' . Consequently, if Q_s is equal to $\alpha_j Q_{s_j} \cdot \frac{A_{sS}}{A_{s_j S}}$, $j \in \{1,2,\ldots,n\}$, or to some integral multiple of $\alpha_j Q_{s_j} \frac{A_{sS}}{A_{s_j S}}$, every time stage s produces, all its immediate successors will also be producing. If, however, $Q_s \neq k\left(\alpha_j Q_{s_j} \cdot \frac{A_{sS}}{A_{s_j S}}\right)$, $j \in \{1,2,\ldots,n\}$, k = positive integer, stage s will produce at least once during the cycle P , and at that point in time at least one of its immediate successors will not be producing.

Example



D = 10 units/period

A_{1S} =3; A_{2s} = 1 ; A_{3S} = 1

P = 12 periods; P' = 6 periods

In the production schedule given in the table below production at stages 2, and 3 always occurs in the first period of the subcycle P' , while production at stage 1 (symbolized by an arrow) can take place anywhere inside a period so as to make the production points at stage 1 equidistant in time.

	Per.1	2	3	4	5	6	7	8	9	10	11	12
Stage 1	72 ↓		72 ↓		72 ↓			72 ↓		72 ↓		
Stage 2	Q_2=30			30			30			30		
Stage 3	Q_3=20		20		20		20		20		20	
	P'						P'					
	P											

This example clearly illustrates that, except for the first period in a cycle P , production does not take place simultaneously at s and all its immediate successors.

We want to show that, if $Q_s \neq Q_s^*$ a stockout tends to occur at some time during a cycle P , and therefore a permanent stock will be required.

Let t_o , $1 < t_o < P+1$, be the first point in time, during cycle P, when stage s produces (in the above example t_o = 2.4 periods). Suppose that a number i , $1 \leq i \leq n$, of the n immediate successors of stage s do not produce at moment t_o , while the remaining (n-i) immediate successors do produce at t_o . Re-number the immediate successors so

that $s_1, \ldots, s_i$ do not produce at t_o, and $s_{i+1}, \ldots, s_n$ do produce.

Obviously:

$$t_o = \frac{Q_s}{A_{sS}D}$$

During the interval $[1, t_o - \delta t]$, where δt is infinitely small, stage s_1 produces a number of batches (including the batch produced at the beginning of cycle P) equal to:

$$t_o\left[\frac{A_{s_1S}D}{Q_{s_1}} + 1\right] = \left[\frac{Q_s}{A_{sS}D} \cdot \frac{A_{s_1S}D}{Q_{s_1}} + 1\right] = \left[\frac{Q_s A_{s_1S}}{A_{sS}Q_{s_1}}\right] + 1$$

where [ ] denotes "integral part of".

Since, by assumption, stages s and s_1 do not produce simultaneously at moment t_o, Q_s is not an integral multiple of

$Q_{s_1} \cdot \frac{A_{sS}}{A_{s_1S}}$. It then follows directly that:

$$\left[\frac{Q_s A_{s_1S}}{A_{sS}Q_{s_1}}\right] + 1 > \frac{Q_s A_{s_1S}}{A_{sS}Q_{s_1}}$$

or

$$\left\{\left[\frac{Q_s A_{s_1 S}}{A_{sS} Q_{s_1}}\right] + 1\right\} \cdot Q_{s_1} \cdot a_{ss_1} > Q_s \cdot \frac{A_{s_1 S}}{A_{sS}} \cdot a_{ss_1} \qquad (5.1)$$

Inequalities similar to (5.1) can be written for all i immediate successors that do not produce at time t_o.

With stage s_{i+1} things change. Since stage s_{i+1} does produce at $t_o \Rightarrow Q_s$ is an integral multiple of

$$Q_{s_{i+1}} \cdot \frac{A_{sS}}{A_{s_{i+1} S}} :$$

$$Q_s = m_{i+1} \cdot Q_{s_{i+1}} \cdot \frac{A_{sS}}{A_{s_{i+1} S}}$$

where m_{i+1} is a positive integer. This implies:

$$m_{i+1} \cdot Q_{s_{i+1}} \cdot a_{ss_{i+1}} = Q_s \cdot \frac{A_{s_{i+1} S}}{A_{sS}} \cdot a_{ss_{i+1}} \qquad (5.2)$$

Equalities similar to (5.2) can be written for all $(n-i)$ immediate successors that do produce at time t_o.

When (5.1) and (5.2) are put together, we obtain:

$$\left.\begin{array}{l}\left\{\left[\dfrac{Q_s A_{s_1 S}}{A_{sS} Q_{s_1}}\right] + 1\right\} \cdot Q_{s_1} \cdot a_{ss_1} \\ \vdots \\ \left\{\left[\dfrac{Q_s A_{s_i S}}{A_{sS} Q_{s_i}}\right] + 1\right\} \cdot Q_{s_i} \cdot a_{ss_i}\end{array}\right\} > \left.\begin{array}{l} Q_s \cdot \dfrac{A_{s_1 S}}{A_{sS}} \cdot a_{ss_1} \\ \vdots \\ Q_s \dfrac{A_{s_i S}}{A_{sS}} \cdot a_{ss_i}\end{array}\right\} \text{ i inequalities}$$

$$\left.\begin{array}{l} m_{i+1} Q_{s_{i+1}} a_{ss_{i+1}} \\ \vdots \\ m_n Q_{s_n} a_{ss_n}\end{array}\right\} = \left.\begin{array}{l} Q_s \cdot \dfrac{A_{s_{i+1} S}}{A_{sS}} a_{ss_{i+1}} \\ \vdots \\ Q_s \cdot \dfrac{A_{s_n}}{A_{sS}} a_{ss_n}\end{array}\right\} (n-i) \text{equalities}$$

$$\sum_{j=1}^{i} \left\{\left[\frac{Q_s A_{s_j S}}{A_{sS} Q_{s_j}}\right] + 1\right\} \cdot Q_{s_j} \cdot a_{ss_j} + \sum_{j=i+1}^{n} m_j Q_{s_j} a_{ss_j} > \sum_{j=1}^{n} Q_s \cdot \frac{A_{s_j S} a_{ss_j}}{A_{sS}}$$

Since by definition:

$$\sum_{j=1}^{n} A_{s_j S} a_{ss_j} = A_{sS}$$

the final relation is:

$$\sum_{j=1}^{i} \left\{ \left[\frac{Q_s A_{s_j S}}{A_{sS} Q_{s_j}} \right] + 1 \right\} \cdot Q_{s_j} \cdot a_{ss_j} + \sum_{j=i+1}^{n} m_j Q_{s_j} a_{ss_j} > Q_s \qquad (5.3)$$

The left hand side of (5.3) represents amount of parts required from stage s during the interval $[1 , t_o - \delta t]$, while the right hand side represents amount of parts made available by stage s during the same interval. If no initial stock is provided, a stockout would occur because the installation stock at moment $(t_o - \delta t)$ is negative:

$$I(t_o - \delta t) = Q_s - \sum_{j=1}^{i} \left\{ \left[\frac{Q_s A_{s_j S}}{A_{sS} Q_{s_j}} \right] + 1 \right\} \cdot Q_{s_j} \cdot a_{ss_j} - \sum_{j=i+1}^{n} m_j Q_{s_j} a_{ss_j} < 0$$

This implies that for any Q_s which is not an integral multiple of

$\alpha_j Q_{s_j} \frac{A_{sS}}{A_{s_j S}}$ (i.e., $Q_s \neq Q_s^*$) :

$$\min_t I_{Q_s \neq Q_s^*}(t) < 0$$

or

$$\max_t (\bar{I}_{Q_s \neq Q_s^*}(t)) = -\min_t (I_{Q_s \neq Q_s^*}(t)) > 0$$

and, consequently:

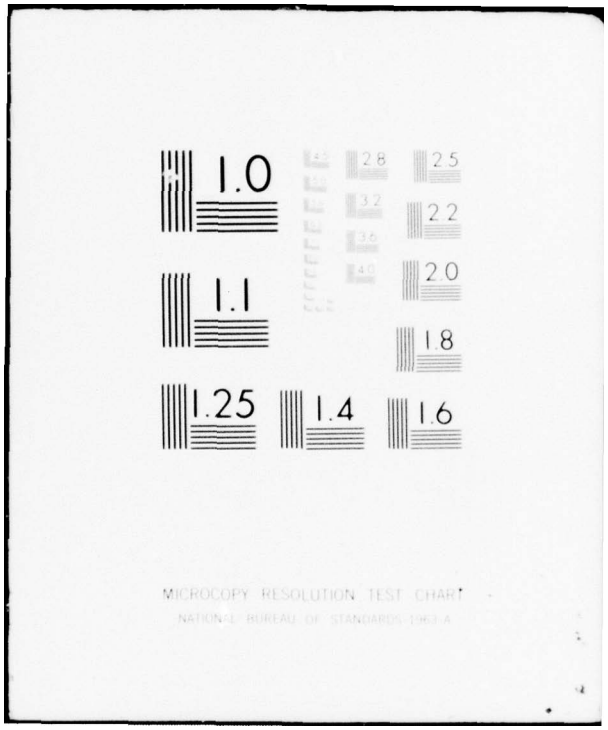
AD-A042 559 MASSACHUSETTS INST OF TECH CAMBRIDGE OPERATIONS RESE--ETC F/G 5/3
ISSUES OF HIERARCHICAL PLANNING IN MULTI-STAGE PRODUCTION SYSTE--ETC(U)
JUL 77 D CANDEA N00014-75-C-0556
UNCLASSIFIED TR-134 NL

3 OF 3
AD
A042559



END

DATE
FILMED
8-77


1.0
1.1
1.25
1.4
1.6
2.8
2.5
3.2
2.2
3.6
2.0
4.0
1.8
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

$$Z(Q_s \neq Q_s^*) = \frac{B_s D A_{sS}}{Q_s} + H_s \cdot \frac{Q_s - A_{sS}}{2} + h_s \cdot \max_t \left(\bar{I}_{Q_s \neq Q_s^*}(t) \right) >$$

$$> \frac{B_s D A_{sS}}{Q_s} + H_s \cdot \frac{Q_s - A_{sS}}{2} = W(Q_s \neq Q_s^*)$$

or

$$Z(Q_s \neq Q_s^*) > W(Q_s \neq Q_s^*)$$

Part e - Analysis of the Results.

We have found that

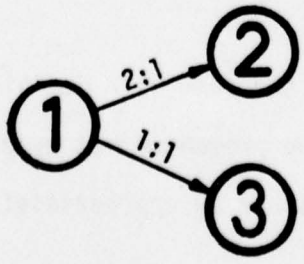
$$U(Q_s) \geq Z(Q_s) \geq W(Q_s) ,$$

with

$$U(Q_s^*) = Z(Q_s^*) = W(Q_s^*)$$

$$Z(Q_s \neq Q_s^*) > W(Q_s \neq Q_s^*)$$

The interpretation becomes straightforward after we consider the following example:



B_1 = 80 (setup cost at stage 1)

D = 1000 units/period

A_{1S} = 3 ; A_{2S} = 1 ; A_{3S} = 1

h_1 = 2 ; H_1 = 2

Q_2 = 30 ; Q_3 = 20; $Q^* = k \cdot 180$

In figure 5.6 cost curve $Z(Q_1)$ and the two bounding functions are shown for our case. $Q_{1\ opt}$ is the lot size for which $W(Q_1)$ achieves its minimum (i.e., the unconstrained EOQ, computed with echelon inventory holding cost). It can be seen how $Z(Q_1)$ is "pinched" between the two bounds for Q_1 an integral multiple of 180. For $Q_1 \neq k \cdot 180$ the cost curve is strictly above $W(Q_1)$.

In the process of searching for the optimal lot sizes, curve $W(Q_1)$ will stay fixed, while $U(Q_1)$ and $Z(Q_1)$ will "jump" around as Q_2 and Q_3 are changed; of course, the contact points between the three curves will be different for different Q_2 and Q_3.

If we concentrate upon the interval bracketing $Q_{1\ opt}$, it is clear that, inside the interval, $Z(Q_1)$ can take on a value lower than $Z(Q_1 = 360)$ or $Z(Q_1 = 540)$. Therefore, we can only say that $Q_1 = Q_1^*$ (in this case $Q_1^* = 540$) represents only an approximation to the optimum of the cost function $Z(Q_1)$. This approximation is likely to be good, given the flatness of $W(Q_1)$ around its minimum (Solomon [84]); there are, also, chances that the minimum of $Z(Q_1)$ will be exactly given by some point of contact between the three curves (i.e., a multiple of

$$\alpha_j Q_j \cdot \frac{A_{sS}}{A_{s_j S}}) \ . \ \|$$

* * * *

We will now specialize our results for the case with no initial inventories, in which case stockouts have to be prevented not by holding a certain amount of permanent inventory but rather by appropriately choosing the lot sizes.

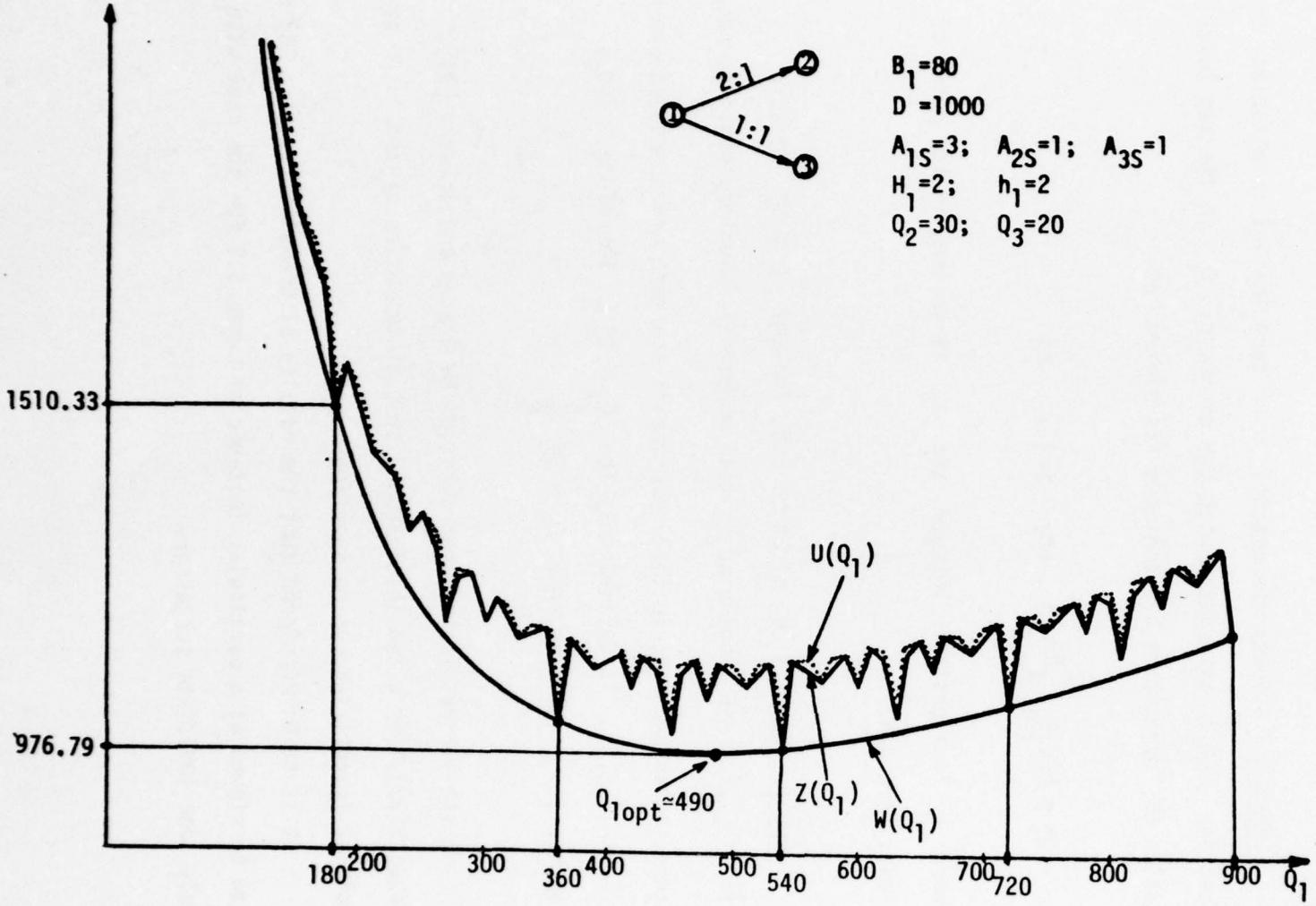


Fig. 5.6- Typical cost curve.

LEMMA 5.7 - Under the assumption of zero initial inventories, cost function $Z(Q_s)$ associated with the choice of Q_s in the two level system of Proposition 5.2 achieves its minimum at

$$Q_s^* = k(\alpha_j Q_{s_j} \cdot \frac{A_{sS}}{A_{s_jS}}) \text{ , any } j\in\{1,\ldots,n\}$$

where k is a positive integer, and α_j is defined in the proof of Proposition 5.5.

Proof. By part d of Lemma 5.6, for any $Q_s \neq Q_s^*$ we have $\min_t I_{Q_s\neq Q_s^*}(t) < 0$, hence an initial permanent inventory would be required. Since, by assumption, initial inventories are zero and shortages are not allowed, $Z(Q_s)$ is defined only for $Q_s = Q_s^*$. Therefore, $Z(Q_s)$ reaches its minimum at some $Q_s = Q_s^*$. ∎

A similar result has been obtained by Graves and Schwarz [39], Schwarz[80], for a two level arborescent distribution system, with zero initial inventories and no shortages.

It is to be mentioned that the results of Crowston et. al. [25] can now be viewed as a particular instance of Lemma 5.7 for the case with only one immediate successor.

THEOREM 5.8 - Under the assumption of zero initial inventories, the optimal solution to THE LOT SIZING PROBLEM in an assembly system with diverging arcs has the lot size Q_s at stage $s(s=1,\ldots,S-1)$ with the following properties:

- if stage s has only one immediate successor $a(s)$ then:

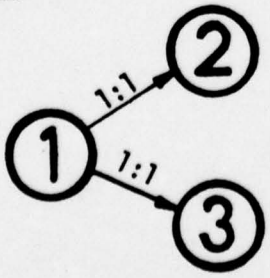
$$Q_s = k \cdot (a_{sa(s)} Q_{a(s)}) , \qquad k = \text{positive integer;}$$

- if stage s has multiple immediate successors, then Q_s is given by Q_s^* of Lemma 5.7 .

Proof. The proof is by induction over the levels of the multi-stage systems and will be omitted.

In closing this section we would like to point out that the optimal lot size Q_s^* , when stage s has multiple successors, is not necessarily a multiple of the lot size at all successor stages, but rather a multiple of the requirements over the subcycle P' .

Example



$Q_2 = 60$; $Q_3 = 90$

$A_{1S} = 3$; $A_{2S} = 2$; $A_{3S} = 1$

$$P' = \alpha_2 \cdot \frac{Q_2}{A_{2S}} = \alpha_3 \cdot \frac{Q_3}{A_{3S}} \Rightarrow \alpha_2 = 3 \ , \ \alpha_3 = 1$$

$$Q_1^* = k \cdot \left(\alpha_2 \cdot \frac{Q_2 A_{1S}}{A_{2S}} \right) = k \cdot 270$$

For k=1 we have $Q_1^* = 270$, which is not an integral multiple of $Q_2 = 60$.

5.2. The Case with Independent Demand for the End Product and Component Parts

Assumptions: all assumptions of section 5.1 stay unchanged; additionally, we assume that the independent demand for component parts is constant and occurring one unit at a time.

This is the case where some or all component parts are needed not only for end product assemblying but also for service purposes; the lot size calculations will have to reflect the existence of multiple sources of demand. (figure 5.7)

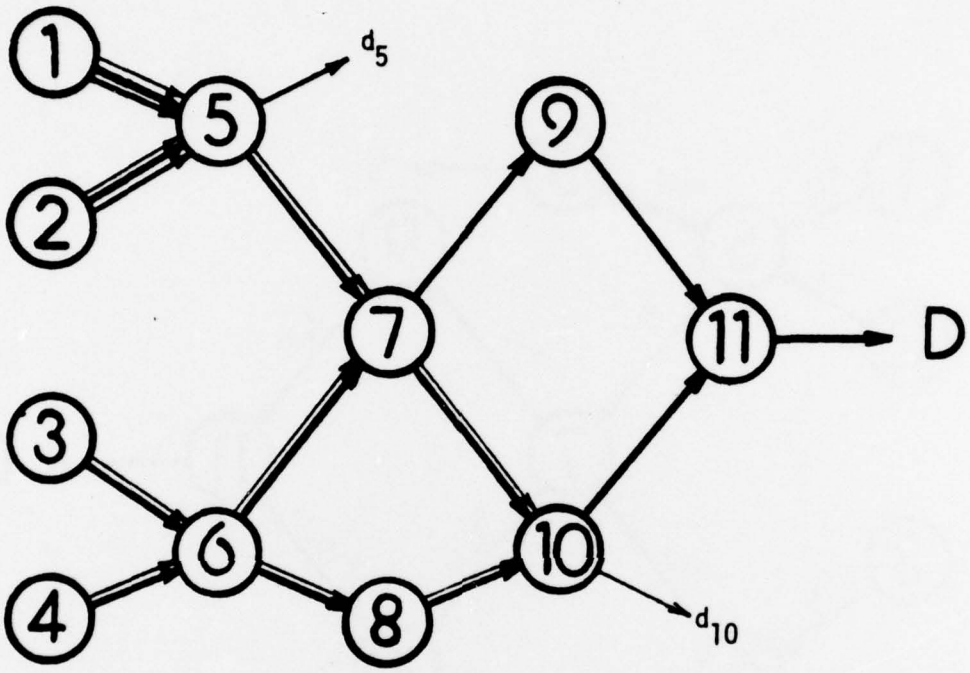


Fig. 5.7- A case with independent demand for parts 5 and 10.

Clearly, then:

- demand upon stages 1 , 2 , 5 originates from three sources,
- demand upon stages 3 , 4 , 6 , 7 , 8 , 10 comes from two sources, and
- demand upon stages 9 , 11 has only one source of origin.

In order to be able to analyze the problem in the broader framework of structures with diverging arcs, we will introduce one dummy facility for every part required for service purposes; the independent demand will then be associated with the dummy facility (see facilities 5' and 10' in figure (5.8).

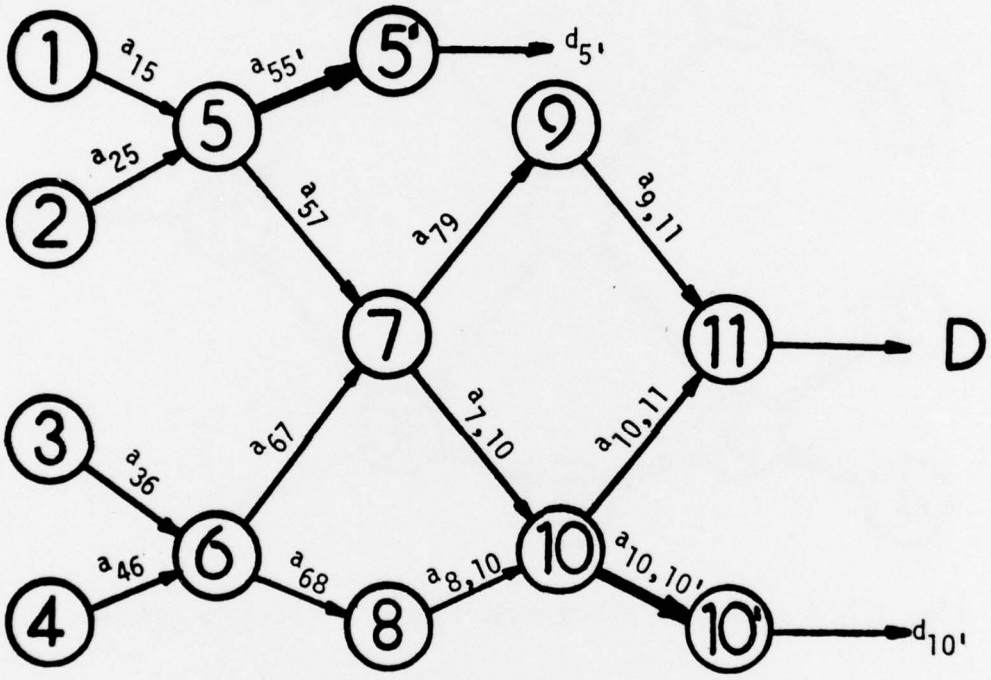


Fig. 5.8- Multi-stage structure with dummy facilities.

A dummy facility performs the following activity: as soon as one unit of spare part is demanded, it immediately withdraws that unit from inventory and ships it. The setup cost associated with a dummy facility is zero, and the lot size is equal to 1 ; this implies that the average inventory is zero. Also, since no real productive activity takes place

at a dummy facility, the value added is zero.

Under these circumstances the theory developed for the case without spare part demand can easily be extended. We will show below the changes that are required to account for the multiplicity of demand sources.

To the notations of section 5.1 we will add:

A_{sq} = number of units of stage s output per unit of stage q output, where q is not an immediate successor of s, but rather a more distant successor. If q is an immediate successor s, $A_{sq} \equiv a_{sq}$.

Let D_s be the demand rate perceived by stage s. D_s is the cumulative effect of the demand for the end product and of the demand for spare parts. Thus, for the structure of figure 5.8 we have:

$$D_1 = A_{1,11}\, D + A_{15}\, d_5 + A_{1,10}\, d_{10}$$

$$D_2 = A_{2,11}\, D + A_{25}\, d_5 + A_{2,10}\, d_{10}$$

$$D_3 = A_{3,11}\, D + A_{3,10}\, d_{10}$$

$$D_4 = A_{4,11}\, D + A_{4,10}\, d_{10}$$

$$D_5 = A_{5,11}\, D + a_{55}\, d_5 + A_{5,10}\, d_{10} \quad ; \quad D_5 = d_5$$

$$D_6 = A_{6,11}\, D + A_{6,10}\, d_{10}$$

etc.

In general, we can write:

$$D_s = \sum_{q=1}^{s-1} A_{sq} d_q + A_{sS} D$$

where S includes the original and the dummy facilities

d_q = demand for service purposes for the output of facility q (in our example $d_{5'} \neq 0$, and $d_{10'} \neq 0$; all other $d_q = 0$).

Obviously, some A_{sq} are zero.

All the results developed in section 5.1 can be easily adapted to reflect the new situation. We will now re-state Lemma 5.7 to show that, at optimality, lot sizes should still be integral multiples of the sub-cycle requirements at the successor stages.

LEMMA 5.7-A - Under the assumption of zero initial inventories, cost function $Z(Q_s)$ associated with the choice of Q_s in the two level system of Proposition 5.2 achieves its minimum at

$$Q_s^* = k \cdot \left(\alpha_j Q_{s_j} \cdot \frac{D_s}{D_{s_j}} \right), \quad \text{any } j \in \{1,\ldots,n\}$$

where k is a positive integer, and α_j defines the subcycle P' as follows:

$$P' = \alpha_1 \frac{Q_{s_1}}{D_{s_1}} = \cdots = \alpha_n \frac{Q_{s_n}}{D_{s_n}}$$

5.3. Solving the Lot Sizing Problem

The solution to THE LOT SIZING PROBLEM is completely defined if the EOQ at the finished good stage and the lot size multipliers k are specified. Two approaches can be followed in order to develop a solution procedure: an exact approach, or a heuristic approach.

We have developed an exact procedure of the dynamic programming type, similar to that by Crowston et. al. [25]. However, insurmountable computational difficulties developed. Because of the diverging arcs in the product structure, the dynamic programming involved more than one stage variables and, although bounds on the search space have been developed, the computer memory requirements made small problem impractical and larger problems impossible to solve. For this reason we have discarded dynamic programming as a viable alternative and have investigated the heuristic approach.

* * * *

Let us consider now the case with independent demand for the end product only.

Under an integral multiple policy, the costs at all stages consist only of setup and echelon holding costs. Since the EOQ's at various stages are related to each other by the integral multipliers, the cost function for the entire multi-stage system can be expressed in terms of the EOQ at the end product stage and the multipliers.

Consider, as an example, the structure in figure 5.9. The notation we will use for multipliers will be, for instance:

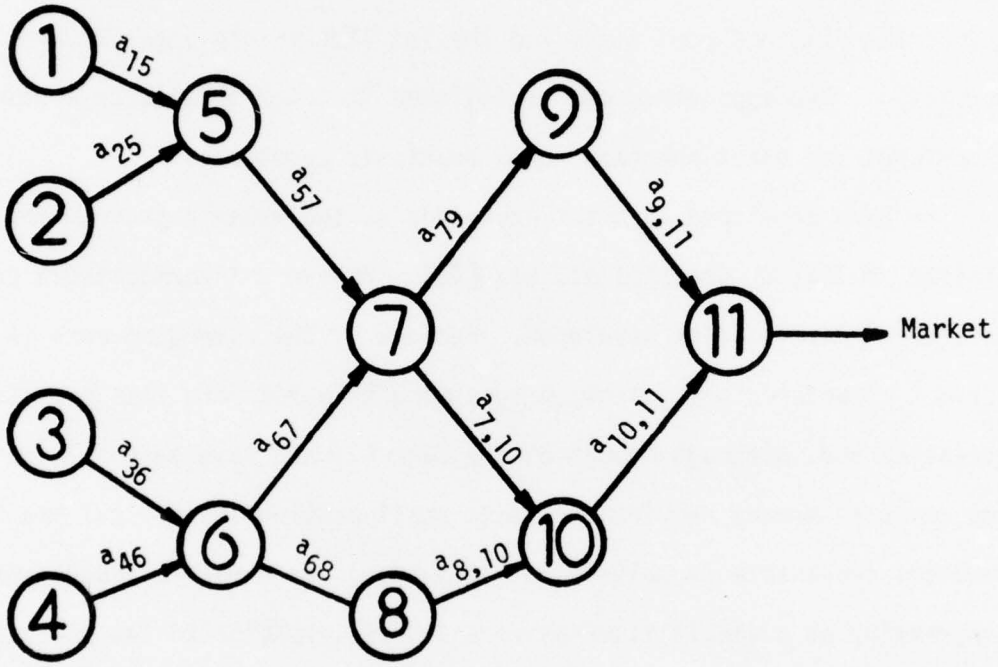


Fig. 5.9- Multi-stage structure.

$k_{9,11}$ = positive integer that yields Q_9 as a function of Q_{11} ,

$k^{9}_{7\,10}$ = positive integer that yields Q_7 as a function of either Q_9 or Q_{10} .

Thus, for our example we can write the following

$Q_9 = K_9 Q_{11}$, $\quad K_9 = k_{9,11}\, a_{9,11}$

$Q_{10} = K_{10} Q_{11}$, $\quad K_{10} = k_{10,11}\, a_{10,11}$

$Q_7 = K_7 Q_{11}$, $\quad K_7 = k_{\substack{9\\7\\10}}\, \alpha_9 K_9 \frac{A_{7S}}{A_{9S}}$

$\frac{\alpha_9 Q_9}{A_{9S}} = \frac{\alpha_{10} Q_{10}}{A_{10,S}}$, where α_9 , α_{10} are relatively prime positive integers.

$Q_8 = K_8 Q_{11}$, $\quad K_8 = k_{8,10}\, a_{8,10}\, K_{10}$

$Q_6 = K_6 Q_{11}$, $\quad K_{\substack{6\\8}}^{7} = k_{\substack{6\\8}}^{7} \alpha_7 K_7 \frac{A_{6S}}{A_{7S}}$

$\frac{\alpha_7 Q_7}{A_{7S}} = \frac{\alpha_8 Q_8}{A_{8S}}$, where α_7 , α_8 are relatively prime positive integers

$Q_5 = K_5 Q_{11}$, $\quad K_5 = k_{57}\, a_{57}\, K_7$

$Q_4 = K_4 Q_{11}$, $\quad K_4 = k_{46}\, a_{46}\, K_6$

$Q_3 = K_3 Q_{11}$, $\quad K_3 = k_{36}\, a_{36}\, K_6$

$Q_2 = K_2 Q_{11}$, $\quad K_2 = k_{25}\, a_{25}\, K_5$

$Q_1 = K_1 Q_{11}$, $\quad K_1 = k_{15}\, a_{15}\, K_5$

The TOTAL COST for the whole multi-stage system is the sum of the costs at all stages:

$$\text{TOTAL COST} = \sum_{s=1}^{S} \left(\frac{DA_{sS}B_s}{K_s Q_S} + H_s \cdot \frac{K_s Q_S - A_{sS}}{2} \right) \tag{5.3}$$

For a given vector M^j of lot size multiples

$$M^j = \left\{ K_1^j , K_2^j , \ldots , K_{S-1}^j \right\} ,$$

the optimal $Q_{S,j,opt}$ is obtained by setting to zero the derivative of TOTAL COST with respect to Q_S .

$$Q_{S,j,opt} = \sqrt{\frac{2D \cdot \sum_{s=1}^{S} \frac{A_{sS}B_s}{K_s^j}}{\sum_{s=1}^{S} H_s K_s^j}} \tag{5.4}$$

Crowston et. al. [22] suggested three heuristics for searching the sets of integral multiples. Two of them will be adapted to our problem, to search for the optimum vector M* which will minimize the TOTAL COST function.

Let us note here that the introduction of lot size independent lead times would not affect our results. Such lead times would only increase the inventory holding cost by a constant amount.

The two heuristics we have already mentioned have been programmed and run. The basic search module is the same for both of them; the difference is the initialization procedure:

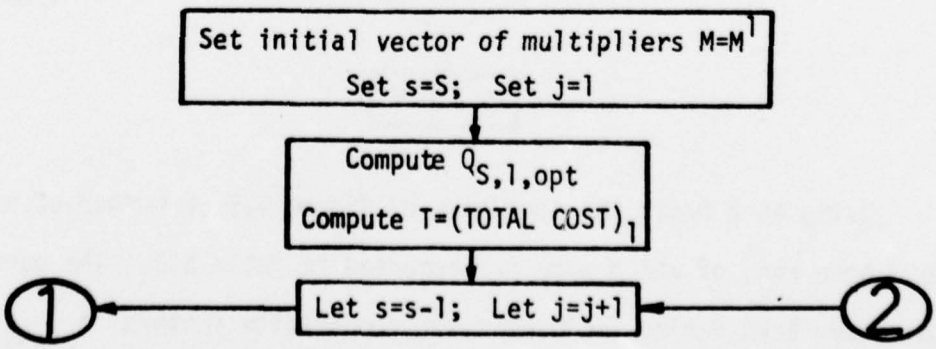
Heuristic 1

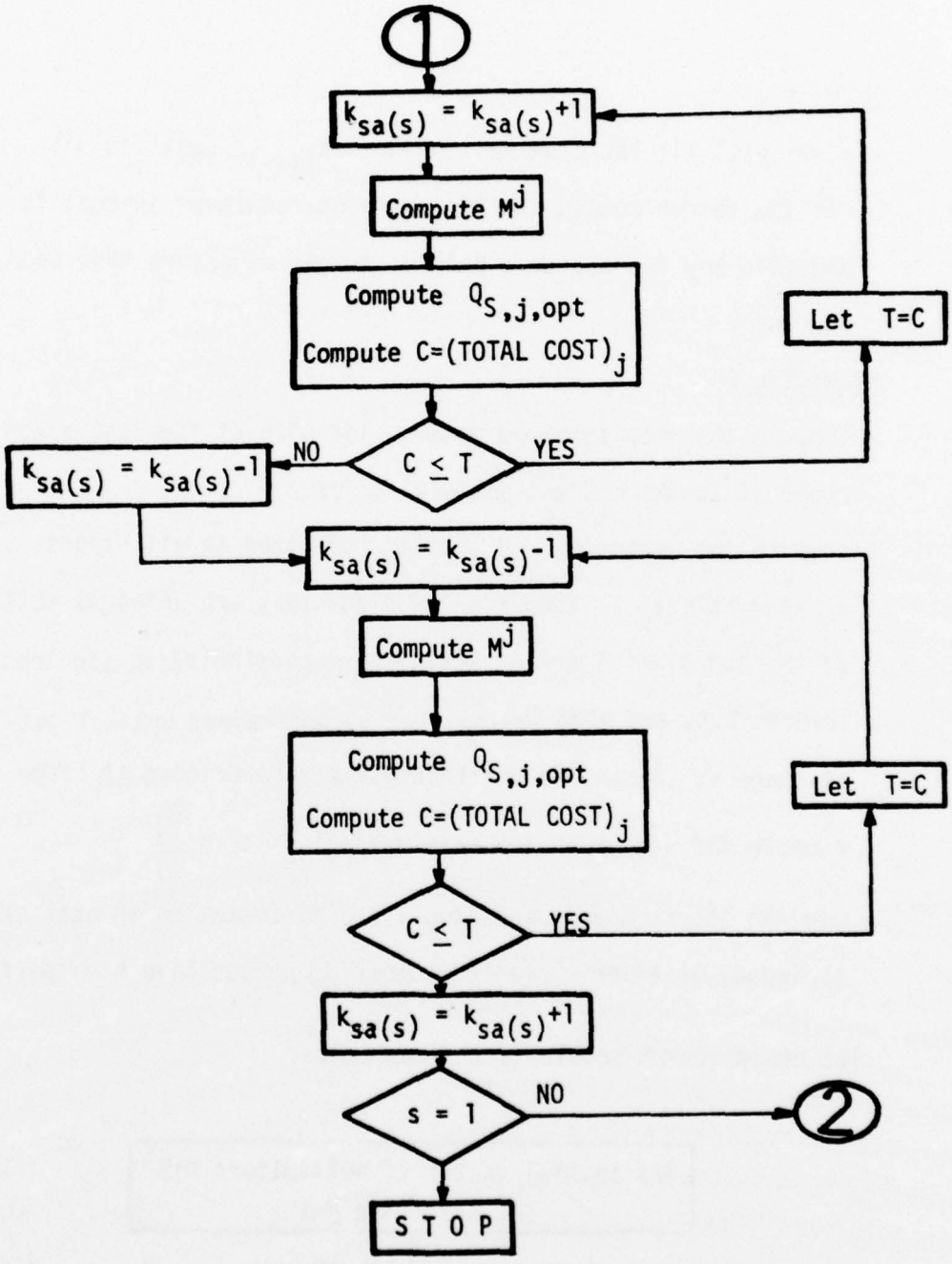
- start with all lot size multipliers $k_{sa(s)}$ equal to 1 ;
- use the search module and stop when no reduction in cost is possible any longer, or a pre-set number of passes have been performed.

Heuristic 2

- compute the unconstrained optimum lot size at the last stage and round it to the nearest integral value;
- compute the unconstrained optimum lot sizes at all stages;
- at every stage s find two lot sizes that are integral multiples of the lot size(s) at the immediate successor(s), in the sense of Theorem 5.8, and also bracket the unconstrained optimal lot size at stage s; choose the one that gives a lower cost at stage s.
- compute the vector of lot size multiples $K_s = \frac{Q_s}{Q_S}$;
- use the search module and stop after no reduction in cost can be achieved, or after a pre-set number of passes have been performed.

The basic search module is charted below:





Using as a basis the structure of figure 5.9, a number of examples have been run, of which some are reported in table 5.1. The parameters that have been varied included: the composition factors $a_{sa(s)}$, the

setup costs, echelon inventory holding costs, and demand. Table 5.2 summarizes the relevant data for the five problems reported in table 5.1. Table 5.3 presents both the unconstrained optimum lot sizes and the lot sizes produced by the best heuristic of the two tested.

In all runs both heuristics terminated in a few iterations. Apparently the level of demand influences neither the number of iterations until termination, nor the values of the lot size multipliers.

* * * *

We have extended the two heuristics to cover the case where there is also independent demand for component parts; the basic structure considered was that of figure 5.7, with independent demand for the end product and for parts 5 and 10. Because the search had to be conducted in more than one dimension, the computer time involved was very large, and we concluded that even heuristics cannot efficiently solve these kinds of problems. It is conceivable that for the case where parts will be common to several end products even the heuristic lot sizing will become computationally infeasible. Therefore, it would be worth investigating what has come to be known as "myopic lot sizing policies", in which one optimizes the objective function with respect to any two stages and ignoring multistage interactions (Schwarz and Schrage [81], Graves and Schwarz [39]).

Table 5.1 - Results of the heuristic lot sizing

Example #	HEURISTIC 1						HEURISTIC 2						Lower bound L	% over lower bound of best known solution
	Initialization run		Final solution				Initialization run		Final solution					
	Cost	Q_{11}	Cost	Q_{11}	Lot size multiples vector $\{K_s\}$	# of iter.	Cost	Q_{11}	Cost	Q_{11}	Lot size multiples vector $\{K_s\}$	# of iter.		
1	17203	83	15673	42	{8,8,12, 12,8,12, 4,4,2,2,1}	2	25883	9	15340	22	{12,12,36, 36,12,36, 12,12,6,2,1}	3	14281	7.4
2	38595	186	35173	93	{8,8,12, 12,8,12, 4,4,2,2,1}	2	58004	20	34429	50	{12,12,36, 36,12,36, 12,12,6,2,1}	3	32063	7.4
3	13155	216	11688	152	{6,2,3,3, 2,3,2,1, 1,1,1}	3	14839	18	11957	27	{36,12,18, 18,12,18, 12,6,3,6,1}	3	11131	5
4	25227	57	23938	52	{2,10,6, 18,2,6, 2,2,1,1,1}	2	36377	8	23938	52	{2,10,6, 18,2,6,2, 2,1,1,1}	2	17829	34.3
5	39844	72	39844	72	{30,20,33, 33,10,33, 5,2,1,2,1}	1	42954	39	41769	38	{60,40,66, 66,20,66, 10,4,1,2,1}	2	25773	54.6

NOTE: -All data used in the above example runs are listed in table 5.2.
-The lower bound L is the cost of using unconstrained lot sizes at all stages with the permanent installation inventories ignored.

Table 5.2- Relevant data for the problems of table 5.1

	Example 1				Example 2				Example 3				Example 4				Example 5			
Stage #	Composition factors	Setup costs	Echelon holding costs	Demand	Composition factors	Setup costs	Echelon holding costs	Demand	Composition factors	Setup costs	Echelon holding costs	Demand	Composition factors	Setup costs	Echelon holding costs	Demand	Composition factors	Setup costs	Echelon holding costs	Demand
1	$a_{15}=1$	60	4	D = 1000	$a_{15}=1$	60	4	D = 5000	$a_{15}=1$	800	4	D = 1000	$a_{15}=1$	60	40	D = 1000	$a_{15}=3$	60	4	D = 1000
2	$a_{25}=1$	80	5		$a_{25}=1$	80	5		$a_{25}=1$	80	5		$a_{25}=1$	80	1		$a_{25}=2$	80	5	
3	$a_{36}=1$	50	4		$a_{36}=1$	50	4		$a_{36}=1$	50	4		$a_{36}=1$	50	40		$a_{36}=1$	50	4	
4	$a_{46}=1$	40	2		$a_{46}=1$	40	2		$a_{46}=1$	40	2		$a_{46}=1$	40	0.5		$a_{46}=1$	40	2	
5	$a_{57}=1$	70	2		$a_{57}=1$	70	2		$a_{57}=1$	70	2		$a_{57}=1$	70	20		$a_{57}=2$	70	2	
6	$a_{67}=2$ $a_{68}=1$	90	3		$a_{67}=2$ $a_{68}=1$	90	3		$a_{67}=1$ $a_{68}=1$	90	3		$a_{67}=2$ $a_{68}=1$	90	0.6		$a_{67}=5$ $a_{68}=4$	90	3	
7	$a_{79}=1$ $a_{7,10}=1$	80	2		$a_{79}=1$ $a_{7,10}=1$	80	2		$a_{79}=1$ $a_{7,10}=1$	80	2		$a_{79}=1$ $a_{7,10}=1$	80	20		$a_{79}=1$ $a_{7,10}=2$	800	2	
8	$a_{8,10}=2$	100	3		$a_{8,10}=2$	100	3		$a_{8,10}=1$	100	3		$a_{8,10}=2$	100	0.5		$a_{8,10}=1$	100	3	
9	$a_{9,11}=1$	50	2		$a_{9,11}=1$	50	2		$a_{9,11}=1$	50	2		$a_{9,11}=1$	50	20		$a_{9,11}=1$	50	2	
10	$a_{10,11}=1$	60	40		$a_{10,11}=1$	60	40		$a_{10,11}=1$	60	1		$a_{10,11}=1$	60	0.25		$a_{10,11}=2$	60	1	
11		40	80			40	80			5	2			40	20			40	2	

Table 5.3- Comparison between unconstrained optimum lot sizes, and lot sizes provided by the best heuristic solution available

	Example 1		Example 2		Example 3		Example 4		Example 5	
Stage #	Unconstrained optimum lot sizes	Lot sizes by heuristic# 2	Unconstrained optimum lot sizes	Lot sizes by heuristic# 2	Unconstrained optimum lot sizes	Lot sizes by heuristic# 1	Unconstrained optimum lot sizes	Lot sizes by heuristic# 1 or #2	Unconstrained optimum lot sizes	Lot sizes by heuristic# 1
1	245	264	548	600	894	912	77	104	949	2160
2	253	264	566	600	253	304	566	520	800	1440
3	387	792	866	1800	274	456	122	312	908	2376
4	490	792	1095	1800	346	456	980	936	1149	2376
5	374	264	837	600	374	304	118	104	837	720
6	600	792	1342	1800	424	456	1342	312	1407	2376
7	400	264	894	600	400	304	126	104	2000	360
8	365	264	816	600	258	152	894	104	365	144
9	224	132	500	300	224	152	71	52	224	72
10	55	44	122	100	346	152	693	52	490	144
11	32	22	71	50	71	152	63	52	200	72

NOTE: All lot sizes have been rounded off to the nearest integer.

CHAPTER 6 - ISSUES OF SAFETY STOCKS CALCULATIONS

The aggregate planning models of the linear programming type have the disadvantage that they cannot handle demand uncertainties directly. Therefore, one way to account for the uncertain production environment is by imposing the restriction that inventories in the aggregate model should not be depleted below a safety stock level; technically, this can be done either by formulating constraints as such (Hax [47]), or by artificially increasing the demand forecasts input to the model so as to reflect the safety stock requirements.

A large assortment of formulas is available for setting safety stocks in inventory systems (Brown [13], Hax [42]). Based on some appropriate definition for the desired service level, safety stocks are computed and held to protect againgst demand uncertainties.

When planning production on a rolling horizon basis, forecasts are made and revised continuously. It is clear from the nature of the planning process that, for a given future time period, several successive forecasts will be made before the production plans for that period are actually implemented. We will analyze the process of demand forecast fluctuations, and will show how it affects the planning process and the setting of safety stocks.

A good treatment of forecasting techniques based on smoothing, and related analyses of forecast errors, can be found in Montgomery and Johnson [73], Chapter 4, 6, 7 and Brown [11], [12]. Several issues, adapted

from these references, will be presented below without proof, in order to serve us later.

Assume that the time series $\{ D_t \}$ we are analyzing can be represented by:

$$D_t = \sum_{i=1}^{k} b_i f_i(t) + \varepsilon_t \ , \qquad t=1,\ldots,T \qquad (6.1)$$

where b_i are coefficients and $f_i(t)$ are mathematical functions of time. For instance, the constant demand model $D_t = b_1 + \varepsilon_t$, the linear trend model $D_t = b_1 + b_2 t + \varepsilon_t$, and the 12 period seasonal model

$D_t = b_1 + b_2 \sin \frac{2\pi t}{12} + b_3 \cos \frac{2\pi t}{12} + \varepsilon_t$ are particular instances of (6.1).

ε_t is the random demand error; it is assumed that $E(\varepsilon_t) = 0$, σ_ε^2 is time-invariant, and demands in nonoverlapping time periods are independent.

The $f_i(t)$ time functions will be assumed to be so that their values at time period $(t+1)$ are linear combinations of the same functions evaluated at time t::

$$f_i(t+1) = L_{i1} f_1(t) + \ldots + L_{ik} f_k(t) \qquad i = 1 ,\ldots, k \qquad (6.2)$$

The $k \times k$ matrix of the $L_{ij} (i=1 ,\ldots, k \ ; \ j=1 ,\ldots, k)$ is called the transition matrix L . Hence,

$$f(t+1) = Lf(t) \tag{6.3}$$

where $f(t)$ is the $k \times 1$ vector of the $f_i(t)$. A large number of functions satisfy (6.2); these include polynomial, exponential, and trigonometric functions. Montgomery and Johnson [73], chapter 4-3, 4-4 developed the L matrix for various forecasting models.

The estimators $\hat{b}_1, \ldots, \hat{b}_k$ for the model parameters will be computed by the discounted least squares criterion:

$$\min SS_E = \sum_{t=1}^{T} \beta^{T-t} [D_t - f'(t)\hat{b}(T)]^2 \tag{6.4}$$

where β is the discounting factor chosen so that $0 < \beta < 1$; T is the current period, and $b(T)$ is the vector of estimates computed at the end of the current period.

Notice that the assumption is that the time period is fixed; therefore, every time the forecast is updated the model parameters have to be recomputed from scratch in order to incorporate the latest observations. As opposed to this, the technique called <u>direct smoothing</u> assumes the time origin to be at the end of the current period T, and updates the model coefficients simply by using the old estimates and the forecast error recorded in the current period T. Obviously, the time origin itself will have to be updated every time we update the forecast.

To distinguish between the two approaches we will use coefficients a_i for direct smoothing, rather than b_i. The time series model is then:

$$D_{T-j} = \sum_{i=1}^{k} a_i(T)\, f_i(-j) + \varepsilon_{T-j} \qquad j = 0, 1, \ldots, T-1 \qquad (6.5)$$

or in matrix notation:

$$D_{T-j} = f'(-j)\, a(T) + \varepsilon_{T-j} \qquad j = 0, 1, \ldots, T-1 \qquad (6.6)$$

where T = the current time period, considered the time origin,

D_{T-j} = actual demand that occurred in period T-j,

f(-j) = vector of time functions evaluated at time -j , i.e., j periods before moment zero,

a(T) = vector of model parameters, whose i-th element is $a_i(T)$, evaluated at time T . $\hat{a}$ will denote the k x 1 vector of estimates of a .

By minimizing the sum of squared errors one obtains:

$$\hat{a}(T) = L'\hat{a}(T-1) + h e_T(T-1) \qquad (6.7)$$

where L is the transition matrix mentioned earlier, h is the smoothing vector (dimensionality k x 1) , and $e_T(T-1)$ is the single-period forecast error. A τ-period forecast is made when at the end of period T the demand of period $T+\tau$ is forecasted; for $\tau=1$ the single or one period forecast is made, i.e. for one period ahead.

The smoothing vector is given by

$$h = G^{-1} f(0) \qquad (6.8)$$

where the k x k matrix G is:

$$G = \lim_{T\to\infty} G(T) = \lim_{T\to\infty} \sum_{j=0}^{T-1} \beta^j f(-j)f'(-j) \tag{6.9}$$

The one period forcast error recorded in period T is:

$$e_T(T-1) = D_T - F_T(T-1) \tag{6.10}$$

where $F_T(T-1)$ is the one period forecast for period T, made at the end of period T-1, or:

$$F_T(T-1) = f'(1)\,\hat{a}(T-1) \tag{6.11}$$

The variability of the $\hat{a}(T)$ is measured by the k x k covariance matrix V:

$$V = G^{-1} M G^{-1} \sigma_\varepsilon^2 \tag{6.12}$$

where M can be expressed as follows:

$$M = \sum_{j=0}^{\infty} \beta^{2j} f(-j)f'(-j) \tag{6.13}$$

By definition, the τ-period forecast error is:

$$e_{T+\tau}(T) = D_{T+\tau} - F_{T+\tau}(T) \tag{6.14}$$

in which $D_{T+\tau}$ is the actual demand in $T+\tau$, and $F_{T+\tau}(T)$ is the forecast made in period T for period $T+\tau$. $D_{T+\tau}$ and $F_{T+\tau}(T)$ are independent, hence, the variance of $e_{T+\tau}(T)$ is:

$$\text{Var}\,[e_{T+\tau}(T)] = \text{Var}\,[D_{T+\tau}] + \text{Var}\,[F_{T+\tau}(T)]$$

or

$$\boxed{\text{Var}\,[e_{T+\tau}(T)] = \sigma_\varepsilon^2 + \text{Var}\,[F_{T+\tau}(T)]} \qquad (6.15)$$

Since $F_{T+\tau}(T)$ is given by an expression similar to (6.11), i.e.:

$$F_{T+\tau}(T) = f'(\tau)\,\hat{a}(T)\,, \qquad \tau = 1,2,\ldots \qquad (6.16)$$

we can write:

$$\boxed{\text{Var}\,[F_{T+\tau}(T)] = f'(\tau)\,V\,f(\tau)} \qquad (6.17)$$

We can specialize equations (6.15) and (6.17) for various forecasting models. Thus, for the constant demand model $D_t = a_1 + \varepsilon_t$, we have:

$$\text{Var}\,[F_{T+\tau}(T)] = \frac{\alpha_1}{2-\alpha_1}\,\sigma_\varepsilon^2 \qquad (6.18)$$

$$\text{Var}\,[e_{T+\tau}(T)] = \frac{2}{2-\alpha_1}\,\sigma_\varepsilon^2 \qquad (6.19)$$

where α_1 is the smoothing constant for the constant demand model; $\alpha_1 = 1-\beta_1$, where β_1 is the discounting factor in the constant demand model.

- for the linear trend model $D_t = a_1+a_2t+\varepsilon_t$

$$\text{Var}[F_{T+\tau}(T)] = \frac{\alpha_2}{(2-\alpha_2)^3}\left[(10-14\alpha_2+5\alpha_2^2) + 2\alpha_2^2\tau^2 + 2\alpha_2(4-3\alpha_2)\right]\sigma_\varepsilon^2 \qquad (6.20)$$

and $\text{Var}[e_{T+\tau}(T)]$ can be computed by (6.15).

Also, a large number of seasonal models can be found in Brown [12], pp. 187-193 with all their statistical properties listed in tabular form.

We should note that the smoothing constant differs from model to model. For the constant demand model we denoted it by α_1 and, as a general rule, it takes on values between 0.01 and 0.3 . As the number of parameters in the model increases, the smoothing constant should decrease (Montgomery and Johnson [73], pp. 67-69, 90-96 offer a good discussion of this topic). Let α_k be the smoothing constant in a k-parameter model; then, a general guide to how the two constants relate to each other is the following equality:

$$1 - \alpha_k = (1-\alpha_1)^k \qquad (6.21)$$

6.1. Forecast Fluctuations and Safety Stocks for Aggregate Planning with a Rolling Horizon in Single - Stage Systems

As we have already mentioned, when using a rolling horizon policy the demand forecast for some given future time period will undergo a number of changes before a production plan for that period is actually implemented. Thus, if we are at the end of the current period T and the future period is $T+\tau$, there will be $\tau-1$ forecast changes before we arrive at period $T+\tau$.

Let:

$\Delta_{T+\tau}$ = cumulated change, i.e., the result of the $\tau-1$ updates;

$\Delta_{T+\tau}(T,T+1)$ = change in the forecast for period $T+\tau$, when the forecast is done once at the end of period T and then at the end of period $T+1$.

$$\Delta_{T+\tau} = \Delta_{T+\tau}(T,T+1) + \Delta_{T+\tau}(T+1, T+2) +\ldots+ \Delta_{T+\tau}(T+\tau-2, T+\tau-1) =$$

$$= [F_{T+\tau}(T) - F_{T+\tau}(T+1)] + [F_{T+\tau}(T+1) - F_{T+\tau}(T+2)] +\ldots$$

$$\ldots+ [F_{T+\tau}(T+\tau-2) - F_{T+\tau}(T+\tau-1)] = F_{T+\tau}(T) - F_{T+\tau}(T+\tau-1) \quad (6.22)$$

With an unbiased forecasting system:

$$E[\Delta_{T+\tau}] = 0 \quad (6.23)$$

If demand is normally distributed, $\Delta_{T+\tau}$ is also normally distributed with variance $\text{Var}[\Delta_{T+\tau}]$. Thus, if at the end of period T the forecast for period $T+\tau$ is $F_{T+\tau}(T)$, we will be, say, 95% confident that forecast $F_{T+\tau}(T+\tau-1)$ will be within $\pm 1.96\sqrt{\text{Var}[\Delta_{T+\tau}]}$ of the initial forecast $F_{T+\tau}(T)$.

The implications of these fluctuations can be made clear by means of a simple example. Consider that we have a production capacity of 10115 units per period; demand is believed to be constant, forecasting is done by smoothing with $\alpha_1 = 0.25$, and the current forecast is 10000 ±100 per period (notice that relation (6.19) implies for the constant demand model that the forecast error does not depend on how far into the future is the period for which the forecast is made). Production is single stage, with time-invariant costs, and with zero lead times; however, the production level is decided upon and implemented before actual demand is known. Assume that we want to be able to meet the maximum forecasted demand; there are no initial inventories.

When the planning model is set up the demand forecasts input to it are 10000 in every period, and a safety stock of 100 will have to be provided. Consequently, the production plan will call for a production of 10100 units in the first period and 10000 units per period thereafter.

Suppose that one month passes and actual demand turns out to be 10100. No stockout occurred since we had enough safety stock; also, there is no inventory left over. The revised forecast will now be 10025 ±100 per period, so we would have to produce 10125 units in the first period

to meet the demand forecast of 10025 plus to provide the safety stock of 100 . But, this plan is infeasible since the available capacity only permits us to produce 10115 units/period.

Let us, then, see what went wrong. There were two facts that have been overlooked from the very beginning:

- an actual demand higher than the forecast would result in an increased updated forecast;
- an actual demand higher than the forecast would deplete the safety stock, thus creating the need to replenish the safety stock.

Both effects would call for increased production levels and, therefore, can lead to infeasibility (i.e., planning backorders) under tight capacity conditions.

After this analysis we see the safety stock as being composed of three parts:

a) one part to protect against the discrepancy between the forecast and the actual demand;

b) a second part to protect against upward forecast fluctuations after updating;

c) a third part to protect against requirements for safety stock replenishment.

We will show how to compute all three components:

Component (a) is based on the variance of the forecast error. If T is the current period (i.e., the last period for which we have observed data), the variance of the forecast error for period $T+\tau$ is given by

(6.15). Component (a) of safety stock will consist of an appropriate number k of standard deviations of the forecast error (k depends on the desired service level). In case the lead time is not zero, the cumulative forecast error over the lead time plus a review period will serve as the basis for calculations.

Component (b) is based on the variance of the cumulated change $\Delta_{T+\tau}$ and will consist of a number k of standard deviations of the cumulated change. $\Delta_{T+\tau}$ and its variance can be computed starting from the definition (6.22).

Component (c) is based on the cumulative forecast error. Given the forecasting model, it is possible to determine what is the "maximum reasonable" (as determined by a number of k standard deviations) cumulative forecast error that can occur over a time span of 1 period, 2 periods, 3 periods, etc., up to and including the entire planning horizon. These cumulative forecast errors will give us directly the "maximum reasonable" safety stock replenishment requirements that could occur over the corresponding time spans.

Let:

$CF_L(T)$ = cumulative forecast over a L period time interval, when the forecasting process takes place at the end of the current period T ;

$E_L(T)$ = cumulative forecast error over a L period time span when the forecasts are made in period T .

By definition:

$$E_L(T) = \sum_{t=1}^{L} D_t - CF_L(T)$$

The variances are (Montgomery and Johnson [73], p. 140):

$$\text{Var}\,[CF_L(T)] = [\sum_{t=1}^{L} f(t)]' \; V[\sum_{t=1}^{L} f(t)] \qquad (6.24)$$

$$\boxed{\text{Var}\,[E_L(T)] = L\sigma_\varepsilon^2 + \text{Var}\,[CF_L(T)]} \qquad (6.25)$$

Thus, component (c) for period 1 is $k\sqrt{\text{Var}\,[E_1(t)]}$, for period 2 is $k\sqrt{\text{Var}\,[E_2(t)]}$, etc.

In concluding this section we should emphasize that the need for components (b) and (c) is a direct consequence of using the rolling horizon policy. It is also important to realize that component (a) offers protection at the time when demand materializes, while components (b) and (c) offer protection at the time when the planning takes place. Therefore, when capacity is loose, and thus probably able to absorb increased requirements coming from upward forecast fluctuations and from safety stock replenishment, components (b) and (c) might not be needed.

6.2. Safety Stocks in Multi-Stage Systems

Most studies of the safety stocks in multi stage production come from the area of Material Requirement Planning (MRP). While researchs tend to agree that the traditional techniques of determining safety stocks in inventory systems have to be changed when put to work in an MRP environment, there is not that much harmony when the problem is to decide where to keep safety stocks and how much. Answers to this problem range within a wide spectrum, going from the opinion that "safety stock is properly applied only to inventory items subject to independent demand" (Orlicky [77], p. 79)[1] to the more general view by which safety stocks should be held at all levels of a multi-stage production system since all are affected by some sort of uncertainties (Meal [71], New [74], Whybark and Williams [94], Miller [72]).

There has not been much theoretical work done with respect to this topic. Miller [72] has tried to formalize the computation of safety stocks in an MRP system; simulation studies (Whybark and Williams [94]) have investigated the circumstances under which using safety stocks or safety lead times is more appropriate. Most theoretical results with regard to inventories fluctuations under uncertain demand have been produced by research in multi-stage inventory systems, and in distribution

---

[1] Orlicky [77], p. 80 admits, however, that on an exception basis safety stocks could also be planned at purchased items level.

systems rather than in production (Simpson [83], Bryan et. al. [14], Clark and Scarf [17], [18], Hanssmann [40] and [41], Part IV; extensive research papers have been written by Clark [16], Veinott [90], Aggarwal [2]).

Another field of research dealing with inter-stage inventories is the design of production lines; there, questions about the location of bunkers for in-process inventory, and the size of these bunkers, have to be answered. These problems are, however, different from our concern since they are at a detailed level, where such issues as the distribution of processing times at adjacent stages, and the sequencing of jobs through stages have to be considered. Also, the buffer stocks, computed as a trade off between inventory cost and station down-time, are not intended to protect against demand uncertainties; see Young [101], Kraemer and Love [57], and for a review Koenigsberg [56].

In multi-stage production systems, the problems are somewhat different in nature because there are in fact two types of demand we can identify (Orlicky et. al. [75]):

- independent demand generated outside the planning model; this is essentially demand coming from customers in general;
- dependent demand generated inside the planning model by a stage s upon its immediate prodecessor stages.

The two types of demand are quite different; for instance even if the independent demand for the end product is constant and continuous, because of the production in batches the dependent demand (or require-

ments) upon stages producing component parts will be lumpier and lumpier the farther away we get from the final assembly stage.

In this research we hold the view that wherever there are uncertainties that could produce adverse effects, appropriate protection should be built in the form of safety stocks or safety lead times or both. The function of safety stocks is "to absorb all the normal fluctuations in supply and usage which may occur during lead time, and thus to avoid disruption of the production process" (New [74], p. 2). Therefore it would be helpful to identify first the nature of fluctuations, or uncertainties, at some arbitrary stage s of the production process; afterwards, we can look into ways to buffer against them.

There are two categories of uncertainties on one hand: quantity and timing uncertainties; if we look from another viewpoint other two categories can be identified: demand and supply uncertainties. Thus, we can talk about:

- demand quantity uncertainty - the gross requirements upon stage s fluctuate because of changes at higher level items on which the part is used;
- demand timing uncertainty;
- supply quantity uncertainty because of scrap losses or overruns;
- supply timing uncertainty that arises either in the form of uncertain lead times in one's own parts production shops, or in the form of uncertain vendor delivery times.

It is important to point out that we will look at the safety stocks in the context of our aggregate planning model. Although it is true that safety stocks are held at the stock keeping unit level and are considered when computing run out times in the disaggregation process, it is at the aggregate level that they are first used and actually planned into production. The aggregate level of our hierarchical planning scheme is meant to achieve what is probably one of the major criticism of the MRP approach: the integration, in the sense of capacity planning, of what is called in MRP the master schedule (i.e., the production plan at the end-product level), and the parts production.(1)

We will treat first the uncertainties coming from "above": demand quantity and timing uncertainities. We will start out with the case of a one time production and, after the nature of the problem is understood, we will try to shed some light upon the dynamic case, where production for stock is driven by a set of demand forecasts and planning takes place on a rolling horizon basis.

6.2.1. Safety Stocks in a Multi-Stage One Time Production - The Case of Uncertainties in Requirements

Suppose that an item is assembled from a set of machined parts; it takes one month to produce the parts, and one month to assemble the final

---

(1) The criticism is that, in principle, MRP regards the master schedule as a given input, without any concern for how it is obtained and for how should capacity constraints, present at parts production levels, be incorporated in the master schedule (Vollman [93]).

product; capacity is unlimited at all stages. Assume that all we are concerned with, relative to this product, is to be able to serve at least, say, 98% of the demand in period 3 .

We have the following data available to forecast our item's demand:

- the demand is uncertain but has a constant mean (unknown);
- the forecast for month 1 shows a demand of 1000 units,
- the variance of the demand is $\sigma_\varepsilon^2 = 100$ units;
- the forecasting system uses exponential smoothing with $\alpha = 0.25$.

By (6.15) Var $[e_\tau(0)] = 1.143\sigma_\varepsilon^2 = 114.3$; the forecasting interval is ± 2.054 standard deviations, or ± 22 .

Thus, at time zero our forecast is:

$$F_1(0) = 1000 \pm 22$$
$$F_2(0) = 1000 \pm 22$$
$$F_3(0) = 1000 \pm 22$$

Consider that we do the following: based on the information currently available, we plan to produce 1022 units of end product to be available in period 3 (and, thus, by current estimates at least 98% of the demand in period 3 will be served). This requires that 1022 sets of parts be started in production in period 1 and completed within the one month lead time.

Now suppose that the first month went by and the actual demand that occurred in period 1 was 1016. The updated forecast intervals will be:

$$F_2(1) = 1004 \pm 22$$

$$F_3(1) = 1004 \pm 22$$

The 1022 sets of parts are due in shortly, at the beginning of the upcoming (i.e., second) month. But, at this point trouble starts to develop. Indeed, to meet the goal of serving at least 98% of the demand in the third month we would have to schedule in production a batch of 1026 end products; however, this is impossible since only 1022 sets of components will be available. Therefore, the production plan we are just about to draw at the end of period 1 and the beginning of period 2 is infeasible.

The above example was meant to emphasize the need for a safety stock at the component parts level; the purpose of this safety stock would be to protect against the forecast fluctuations (+4 units in our example). Therefore, if we had taken the need for this protection into account we would have proceeded as follows: at the beginning of period 1 (moment zero) we would have recognized the likelihood of the difference:

$$\Delta_{T+3}(T,T+1) = F_{T+3}(T) - F_{T+3}(T+1)$$

Since the forecasting model is a constand demand process we have:

$$F_{T+3}(T+1) \equiv F_{T+2}(T+1)$$

$$F_{T+3}(T) \equiv F_{T+2}(T)$$

so that

$$\Delta_{T+3}(T,T+1) = F_{T+2}(T) - F_{T+2}(T+1)$$

or

$$\Delta_{T+3}(T,T+1) = \alpha F_{T+2}(T) - \alpha D_{T+1}$$

Clearly

$$E[\Delta_{T+3}(T,T+1)] = 0$$

and, it is easy to show that:

$$\text{Var}\,[\Delta_{T+3}(T,T+1)] = \frac{2\alpha^2}{2-\alpha}\ \sigma_\varepsilon^2 = \frac{0.125}{1.75}\,\alpha_\varepsilon^2 = 7.14$$

$$\sqrt{\text{Var}\,[\Delta_{T+3}(T,T+1)]} = 2.67$$

Since our goal is a 98% service level, a safety stock of 2.054 standard deviations is appropriate, or $2.054 \cdot 2.67 \simeq 6$ units.

Thus, the decision which would have protected us against forecast fluctuations would have been to produce 1022 + 6 = 1028 sets of component parts for period 2 . Then, even if in the first period the maximum "reasonable" forecasting error of +22 would have occurred, thus leading to an updated forecast of 1006 ±22, we would have still been able to start in production 1028 end products since the necessary parts would have been there.

We would like to point out that this approach is fundamentally different from the case where safety stocks are provided only at the end product level. In the latter case a safety stock to guard against the forecase error over the cumulative lead time of 2 periods should have been provided. The standard deviation of the cumulative forecast error over 2 periods can be shown to be (formula (6.25)):

$$\sqrt{\text{Var}\,[E_2(T)]} = 1.6\,\sigma_\varepsilon = 16\ .$$

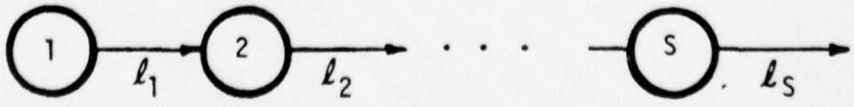
For 98% protection the safety stock should be:

$$2.054 \cdot 16 \simeq 33 \text{ units.}$$

Two remarks can be made about this safety stock of 33 units:

- it is inappropriate because it ignores that planning is done with a rolling horizon, and therefore the forecast is revised before the batch of end product is actually started in production;
- it is more expensive than the safety stock described in the above approach (which had 6 units at the component parts level, and 22 units at end product level).

To generalize, consider a product whose structure shows S stages (considered to be serial, for simplicity). A lead time ℓ_s is associated with stage s . Period T denotes the current period (in a rolling horizon policy).



The cumulative lead time is:

$$L = \ell_1 + \ell_2 + \ldots + \ell_S \ .$$

Our goal is to serve at least a certain percentage of demand in period $T+L+1$.

When production is started at stage 1 we are aware of the fact that before the batch of end products will be released for production a number of revisions of the forecast will take place. Therefore, a safety stock SS_1 should be provided at stage 1 to protect against the forecast fluctuations:

$$SS_1 \xrightarrow[\text{to}]{\text{related}} \Delta_{T+L+1}(T, T+L-\ell_S) = F_{T+L+1}(T) - F_{T+L+1}(T+L-\ell_S)$$

Similarly, when we arrive at the point in time where production at stage 2 has to be started, a safety stock SS_2 has to be included:

$$SS_2 \xrightarrow[\text{to}]{\text{related}} \Delta_{T+L-\ell_1+1}(T, T+L-\ell_1-\ell_S) = F_{T+L-\ell_1+1}(T) -$$

$$- F_{T+L-\ell_1+1}(T+L-\ell_1-\ell_S)$$

For an arbitrary stage s:

$$SS_s \xrightarrow[\text{to}]{\text{related}} \Delta_{T+L-\ell_1-\ldots-\ell_{s-1}+1}(T, T+L-\ell_1-\ldots-\ell_{s-1}-\ell_s) =$$

$$= F_{T+L-\ell_1-\ldots-\ell_{s-1}+1}(T) -$$

$$- F_{T+L-\ell_1-\ldots-\ell_{s-1}+1}(T+L-\ell_1-\ldots-\ell_{s-1}-\ell_s)$$

It is easy to see that when the current period T is just ℓ_s+1 periods away from the target period, the forecast $F_{T+\ell_s+1}(T)$ is made and at the beginning of the upcoming period T+1 the batch of end products is started in production. No safety stock to protect against forecast fluctuations is needed for the finished product:

$$SS_S \xrightarrow[\text{to}]{\text{related}} \Delta \begin{matrix} T, T+L-\ell_1-\ell_2-\ldots-\ell_{n-1}-\ell_n \\ T+L+1-\ell_1-\ell_2-\ldots-\ell_{S-1} \end{matrix} = \Delta \begin{matrix} T, T \\ T+\ell_S+1 \end{matrix} =$$

$$= F_{T+L+1-\ell_1-\ell_2-\ldots-\ell_{S-1}}(T) -$$

$$- F_{T+L+1-\ell_1-\ell_2-\ldots-\ell_{S-1}}(T+L-\ell_1-\ell_2-\ldots-\ell_{S-1}-\ell_S) =$$

$$= F_{T+\ell_S+1}(T) - F_{T+\ell_S+1}(T) = 0$$

The safety stock of finished product will be set such as to provide protection only against the forecast error $e_{T+\ell_S+1}(T)$.

To set safety stocks SS_s as shown above, fluctuations of the $\Delta_{T+\tau}(T,T+\tau-1-\ell_S)$ type and their variances will have to be calculated, given the forecasting model in use. Since, normally, unbiased forecasting systems are utilized

$$E[\Delta_{T+\tau}(T,T+\tau-1-\ell_S)] = 0$$

and it is sufficient to set the safety stock SS_s equal to some appropriate number of standard deviations of the fluctuation.

In closing this section, one final note: in developing the procedure for setting safety stocks, S serial stages have been considered. It is obvious that we could have as well an assembly structure. The central issue is to identify the difference, in time, between the moment when production starts at stage s and the moment when production starts at the last stage S ; call this difference δ_s . Then:

$$\Delta_{T+\delta_s+1+\ell_S}(T,T+\delta_s) = F_{T+\delta_s+1+\ell_S}(T) - F_{T+\delta_s+1+\ell_S}(T+\delta_s) \qquad (6.26)$$

and the corresponding variance can be obtained afterwards.

6.2.2. Safety Stocks in Multi-Stage Systems Planned on a Rolling Horizon Basis--The Case of Uncertainties in Requirements

In this section we will go one step further and will consider the dynamic case, where production is driven by a set of demand forecasts that extend over the planning horizon.Forecasts and inventories are reviewed at

the end of every period, and every time a new production plan is drawn.

As pointed out by Miller [72], p. 20, the issue of setting safety stocks in a dynamic environment of this sort is little understood. Therefore, we will make the following assumptions which, although somewhat restrictive, will give us a start and help us shed some light upon this issue:

- the costs (variable costs, inventory holding, and overtime) in the aggregate planning model are time invariant;
- there is ample regular time capacity, so that capacity restrictions are never constraining; thus, there is no need for overtime.

Under these circumstances, the aggregate model will plan to deliver, at the beginning of period t, an amount of production precisely equal to the requirement in period t (i.e., there will be no seasonal stock buildup). This situation offers the advantage that it will be easier for us to establish relationships between the changes in demand and the resulting changes in the production levels.

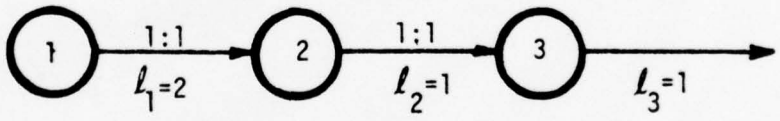
We start by noting that the amount of <u>safety stock of component parts required as protection against forecast fluctuations</u> is the same as in the one time production case (section 6.2.2.)

An additional phenomenon, however, occurs in the dynamic case; because of the discrepancies between actual demands and forecasts, the safety stock of finished product might need to be replenished. This would unexpectedly increase the requirements for component parts, and in case these are not available the aggregate plan will be infeasible although the productive capacity is assumed to be in ample supply. Therefore,

the safety stock at the components level should include a certain amount to support the replenishment of end product safety stock.

The two effects shown above, against which we need protection at the component parts level, constitute what we understood to be the uncertainties in both demand quantity and demand timing.

For the moment let us forget about the fluctuations in forecast, and concentrate on the safety stock replenishment problem. Consider the case of a product with three stages:



ℓ_1, ℓ_2, ℓ_3 = lead times

The inventory balance equations, relating to this product, in an aggregate planning model are as follows (assume no independent demand for component parts):

$$I_{10}+x_{11}-I_{11} = x_{22} \quad (6.27)$$

$$I_{11}+x_{12}-I_{12} = x_{23} \quad (6.28)$$

$$I_{12}+x_{13}-I_{13} = x_{24} \quad (6.29)$$

$$\vdots$$

$$I_{20}+x_{21}-I_{21} = x_{32} \quad (6.30)$$

$$I_{21}+x_{22}-I_{22} = x_{33} \quad (6.31)$$

$$I_{22}+x_{23}-I_{23} = x_{34} \quad (6.32)$$

$$\vdots$$

$$I_{30}+x_{31}-I_{31} = F_1(0) \quad (6.33)$$

$$I_{31}+x_{32}-I_{32} = F_2(0) \quad (6.34)$$

$$I_{32}+x_{33}-I_{33} = F_3(0) \quad (6.35)$$

$$\vdots$$

$$I_{st} \geq \underline{b}_{st}, \quad s=1,2,3; \quad t=1,2,\ldots \quad (6.36)$$

Recall that, for instance, x_{23} denotes production delivered at the beginning of period 3 (started at the beginning of period 2) by stage 2; $\underline{b}_{st}$ is the lower bound on inventory at stage s in period t. Because of the lead times, we have no control over x_{11}, x_{12}, x_{21}, and x_{31}; these variables represent production started before the current planning horizon.

Suppose now that the aggregate model is solved and the optimal values I_{st}, x_{st} of the decision variables are found. Due to the initial assumptions, constraints (6.36) will solve at equality. Next, the

production plan for the upcoming (i.e., the first) period will be implemented. Assume that the actual demand D_1 turns out to be by 1 unit larger than forecast $F_1(0)$. Consequently, the inventory I_{31} at the end of the first period will be 1 unit short of target. Consider now that one period has gone by and that we are in a position to run the aggregate model again; clearly, the new I_{30} (which will be equal to the old I_{31} that we have just talked about) will be 1 unit lower than desired. Since x_{31} is beyond our control, the new I_{31} will be also lower by 1 unit.(1) Unless corrective action is initiated, it is easy to see that I_{32}, I_{33}, etc., will all be 1 unit short of target. To avoid this, the model will increase x_{32} by 1 unit, i.e., the batch of end product to be started in period 1 will be increased by 1 unit.

Thus, a forecast error of +1 in the first period was followed by an increase of +1 in the level of production of finished product in the immediately following period. As we were carrying safety stocks, the larger than expected demand did not produce a stockout, but has consumed from the safety stock. This is why the safety stock has to be replenished.

In turn, a larger x_{32} will require a larger input of component parts (see (6.30)). Reasoning as above we will find that I_{21} will be below the target by one unit, and the model will decide to increase x_{22} by 1 unit. Again, supposedly we were carrying a safety stock at stage 2

---

(1) Since the discrepancy between actual demand and forecast was small, we assume that the forecast did not, practically, change.

so that the surge in x_{32} did not cause a stockout, but the safety stock has to be re-built.

Similarly, we will find that at stage 1 the value of x_{13} will have to go up by 1 unit too, while both I_{11} and I_{12} fall by 1 unit. Notice here that, as the lead time is two periods, the inventory at stage 1 will "feel" the cumulated effect of the independent demand forecast errors over two periods rather than just one.

Thus, a forecast error at the end product level in period t will reverberate through the system, and will produce a 1 to 1 increase or decrease in the level of productions started in period t+1 at all stages, as well as a fall in inventories.

In analyzing the corrective action taken at stage s at moment zero we realize that the next opportunity to make another correction will be only after one review period, and its effect will not be felt until after the corresponding lead time ℓ_s. Therefore, the current corrective action should take care of a time span equal to the lead time ℓ_s plus one review period.

This is nothing new, because in general when setting safety stocks, one has to look ahead one lead time plus one review period. However, the following difference between the independent demand items and the dependent demand items should be noted:

a) The independent demand item is confronted with uncertain demands in all periods of the one lead time plus one review period time span. The safety stock will have, then, to respond to the cumulated forecast error over periods $1,2,3,\ldots,\ell_s$, ℓ_s+1 .

b) The dependent demand item produced by stage s will have to satisfy a set of requirements obtained by exploding back the values of the production levels already set for the immediate successor stages. (see equations (6.27) - (6.32)) . Thus, in our example, the value of x_{22} tells us that so many sets of component parts 2 will be started in production in period 1, and therefore precisely x_{22} units of part 1 will be withdrawn from the inventory of stage 1 . In other words, the requirement (or dependent demand) upon stage 1 in period 1 is exactly known once you have the solution of the planning model. This is in contrast with the requirements for stage 1 in later periods, i. e., x_{23} , x_{24} ,..., which are not deterministic, since the production decisions will be most likely changed during subsequent revisions of the production plan. Hence, the safety stock at stage s(s=1,2,...,S-1) will have to include a certain amount as protection against the cumulative error of the independent demand forecast over periods $1,2,\ldots,\ell_s$.(1) This result is similar to that obtained by Hanssmann [40], p. 497 in treating the optimal inventory location and control in multi-level production and distribution systems (if in his results we set the average shortage time to zero, and if we make the demand independent of the delivery time, our findings are recovered).

The total safety stock at stage s(s=1,2,...,S-1) will consist, then, of two parts:

---

(1) Recall that the forecast error in the independent demand affects inventories at stages 1, 2,..., S-1 with a lag of one period.

$$SS_{s\neq S} = \left\{\begin{array}{l}\text{Protection against}\\ \text{forecast fluctuations}\end{array}\right\} + \left\{\begin{array}{l}\text{Protection against requirements}\\ \text{derived from safety stock}\\ \text{replenishment at higher level}\\ \text{stages}\end{array}\right\} \quad (6.37)$$

Assume the intervening random variables to be normally distributed; if the service level at stage s is defined by the goal of serving at least a certain percentage of the requirements upon stage s in any given period, let us associate a number of k standard deviations with the stated service level. Then, (6.37) becomes:

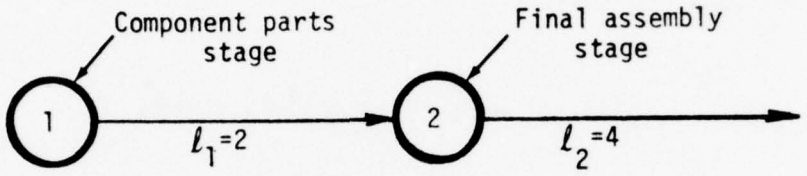
$$SS_{s\neq S} = k\sqrt{\text{Var}\,[\Delta_{T+\delta_s+1+\ell_S}(T,T+\delta_s)]} + k\sqrt{\text{Var}\,[E_{\ell_S}(T)]} \quad (6.38)$$

where $\Delta_{T+\delta_s+1+\ell_S}(T,T+\delta_s)$ is defined by (6.26) and $E_{\ell_S}(T)$ is the cumulative forecast error over the lead time ℓ_s when the forecasting is done at the end of the current period T. $\text{Var}\,[E_{\ell_S}(T)]$ can be computed by (6.25).

The safety stock for the end product will be the regular safety stock, computed based on the cumulative forecast error over the lead time plus the review period; a number of k standard deviations will be considered to correspond to the desired service level:

$$SS_S = k\sqrt{\text{Var}\,[E_{\ell_S+1}(T)]} \quad (6.39)$$

Example



Assume the demand has a linear trend, and the smoothing constant is $\alpha = 0.051$. At least 95% of the requirements upon both stage 1 and 2 have to be served at any moment. The variance of the demand is $\sigma_\varepsilon^2 = 2500$. Determine the safety stocks at both stages 1 and 2 .

Solution - let the current period be moment T=0 . Before applying (6.37) and (6.38) we will identify the following quantities:

$\ell_1 = 2$; $\ell_2 = 4$; $\delta_1 = 2$

$S = 2$; $k = 1.645$

The change in the forecast for period $T+\delta_1+1+\ell_2 = 7$, when the forecasting is done once in the current period T=0 and then in period T=2 , is $\Delta_7(0,2)$; its standard deviation for the given linear trend model can be shown to be:

$$\sqrt{\text{Var}\,[\Delta_7(0,2)]} = 0.16278624\ \sigma_\varepsilon = 8.13931$$

We need the cumulative forecast errors over $\ell_1 = 2$ periods, and over $\ell_2+1 = 5$ periods. First compute the cumulative forecasts (for notations see equation (6.5)):

$$CF_2(0) = F_1(0)+F_2(0) = [\hat{a}_1(0)+\hat{a}_2(0)]+[\hat{a}_1(0)+2\hat{a}_2(0)]=$$

$$= 2\hat{a}_1(0)+3\hat{a}_2(0)$$

$$CF_5(0) = \sum_{i=1}^{5} F_i(0) = 5\hat{a}_1(0)+15\hat{a}_2(0)$$

For the estimators we have (Montgomery and Johnson [73], p. 138):

$$\text{Var}\,[\hat{a}_1(0)] = 0.06405767\ \sigma_\varepsilon^2$$

$$\text{Var}\,[\hat{a}_2(0)] = 0.00003583\ \sigma_\varepsilon^2$$

$$\text{Cov}\,[\hat{a}_1(0), \hat{a}_2(0)] = 0.00135153\ \sigma_\varepsilon^2$$

Then:

$$\text{Var}\,[CF_2(0)] = 4\ \text{Var}\,[\hat{a}_1(0)] + 9\ \text{Var}\,[\hat{a}_2(0)] +$$

$$+ 12\ \text{Cov}\,[\hat{a}_1(0), \hat{a}_2(0)] = 0.2727706\ \sigma_\varepsilon^2$$

$$\text{Var}\,[CF_5(0)] = 25\ \text{Var}\,[\hat{a}_1(0)] + 225\ \text{Var}\,[\hat{a}_2(0)] +$$

$$+ 150\ \text{Cov}\,[\hat{a}_1(0), \hat{a}_2(0)] = 1.81222\ \sigma_\varepsilon^2$$

The standard deviations of the cumulative forecast errors can be computed with (6.25):

$$\sqrt{\text{Var}\,[E_2(0)]} = \sqrt{2\sigma_\varepsilon^2 + 0.2727706\ \sigma_\varepsilon^2} = 75.3785$$

$$\sqrt{\text{Var}\,[E_5(0)]} = \sqrt{5\sigma_\varepsilon^2 + 1.81222\ \sigma_\varepsilon^2} = 130.5011$$

The safety stocks are:

$$SS_1 = 1.645 \cdot (8.13931 + 75.3785) = 137 \text{ units}$$

$$SS_2 = 1.645 \cdot 130.5011 = 215 \text{ units.}$$

* * * *

Several remarks are in order at this point:

1) In the above developments and illustrations we have assumed that one unit of output from stage s goes into every unit of end product. In case this is not true we will re-define the unit of output from stage s so as to include all components produced by stage s that go into one unit of end product.

2) If the output from some stage s is used in several different end products it should be straightforward to use equation (6.38). Indeed, if those end products are independent the variance of the forecast fluctuations and the variance of the cumulative forecast errors will be obtained as sums of variances. If the end products are not independent the relevant covariances will have to be considered.

3) Due to the assumption that capacity was ample we did not have to consider the issue of protecting against capacity shortages because of fluctuating forecasts (as in section 6.1). But we are aware of the possibility of having all necessary component parts without being able to produce the desired amount of end product because of lack of adequate capacity at the final assembly stage. In such a case of tight capacity the results of section 6.1 would be applicable and the safety stock for the end product will have to be supplemented.

4) While the assumption of time invariant aggregate costs is reasonable, the assumption of ample capacity is restrictive. As already mentioned, these assumption resulted in a "lot for lot" production plan, by which the amount of production delivered at each stage was equal to the net requirement upon that stage for the period in question. This feature made it possible to study the effect, upon earlier stages, of the forecast error at the finished product level, and thus permitted us to decide how much safety stock was necessary at every stage s. In the context of tight capacity, however, the effects of perturbations at the end product level (in the form of fluctuations in forecasts, or in the form of forecast errors) upon the inventory at some arbitrary stage $s \neq S$ are not easily and accurately predictable. Therefore, we see no straightforward extension of the results of this section to the more general capacity constrained aggregate planning model, although we believe that the basic approach should be similar.

6.2.3. Safety Stocks in Multi-Stage Systems--The Case of Uncertainties in Supply

In the previous section we have addressed the problem of uncertainties coming from "above"; now we will look at the problem of uncertainties in the supply quantity and timing.

Two relatively recent studies, one by New [74] and the other by Whybark and Williams [94], deal with the problem of uncertainties in MRP systems, and their recommendations for the case of uncertain supply can be adapted to planning with a multi-stage aggregate model.

We should recall that stage s is "connected" with its predecessors $p \in b(s)$ by two links: composition factors a_{ps}, and lead time ℓ_p. Supply uncertainties can affect these two links as it will be shown below:

1) If an amount of x_s units have to be started in production at stage s and $a_{ps}\, x_s$ units of stage p output are ordered, it is possible that the actual receipts will differ from the amount ordered because of scrap losses or overruns at stage p, or because the stage p itself was confronted with a shortage of lower level material. Of course, we are more concerned with receipts lower than the order rather than with the reverse situation. To provide protection, the problem can be approached in two ways: build a safety stock at stage s, or build a safety stock at the predecessor stage p. By the former approach, when the receipts of components are lower than ordered an amount of production smaller than x_s will have to be started at stage s, and the difference will be made up from the available safety stock (thus making it possible for stage s

to deliver the entire amount x_s). The latter approach would be implemented via the so called "scrap allowances", in which case the composition factors a_{ps} are augmented by "yield factors" to protect stage s from possible scrap losses at predecessor stages. In this way, the extra inventory will be in the output from stage p rather than from stage s .

How large a safety stock or "yield factors" should be used is a problem that can be solved by statistical analysis of the history of receipts vs. orders. After assessing a probability distribution of receipt vs. orders, the situation can be formulated as a newsboy problem and solved accordingly (Vachani [89] considers the reject allowances problem in the context of serial structures).

A question that arises, however, is when to use safety stocks at stage s and when to use "yield factors". There are several aspects that would influence this issue: if, for instance, components are made in the company's own fabrication shop probably the yield factors will be used. However, New [74], p. 14 considers that: "The use of overall scrap allowance rates is not to be recommended and in any case inventory due solely to batching often renders most such allowances unnecessary except for high loss items." For components or raw materials that are purchased and present uncertainties in both quantity and timing (e.g., imports of scarce raw materials) one would want to build up some safety stocks to prevent production disruption and at the same time orders larger than the expected requirements can be placed to counteract receipts short of target.

2) If lead times are uncertain, a safety lead time will be added to the "regular" lead time ℓ_p , thus making allowance for possible late

arrivals. Again, there should be no problem with setting appropriate safety lead times after a statistical analysis of a history of lead time variability is conducted.

In concluding, we mention that the simulation study of Whybark and Williams (performed in a MRP environment) led to their recommending [94], p. 605: "Under conditions of uncertainty in timing, safety lead time is the preferred technique, while safety stock is preferred under conditions of quantity uncertainty. These conclusions did not change with the source of the uncertainity (demand or supply), lot sizing technique, lead time, average demand level, uncertainty level...."

CHAPTER 7 - REVIEW AND FUTURE RESEARCH

This thesis has presented a study of the production planning process in multi-stage systems, in the framework of a hierarchical approach. Every effort was made to treat the problem in a setting as general as possible, in order to obtain results applicable to a large range of situations.

The introduction of the thesis was intended to place the problem in the appropriate context, and to identify the principal difficulties associated with this kind of problems. The literature on multi-stage systems was reviewed critically, and the finding was that a new conceptual way of tackling the problem was needed.

After showing how to partition the overall problem into two levels: the aggregate level and the detailed level, work has been done at both ends of the spectrum. The treatment was kept somewhat informal by the inclusion of a large number of simple examples intended to clarify the issues.

At the aggregate planning level a new formulation for the planning model was given in order to bring computational feasibility to situations in which older formulations went beyond the capabilities of current linear programming computer codes. For solution, the problem was tackled using the tools of large scale optimization, namely, column generation. For the generation of columns a simple dynamic programming algorithm was developed. Refinements to improve the efficiency of both the formulation and the column generation procedure have been presented.

Next, the detailed level of the hierarchy was concentrated upon, and the lot sizing problem in multi-stage systems was addressed and solved. Since exact solution algorithms proved to be prohibitively expensive for small problems and completely infeasible for more complex cases, the heuristic approach was adopted. Two heuristics were programmed and run, and the results were reported. For more complex situations, allowing independent demand for component parts or considering parts common to several end products, even the heuristics become computationally infeasible; therefore, it was suggested that myopic lot sizing policies be used.

Chapter six of this research examines the problem of safety stocks for aggregate planning. Starting with the simpler case of one stage production, the implications of using a rolling horizon policy were investigated; it was found that the rolling horizon policy affects the way safety stocks have to be computed. Then, the safety stocks in multi-stage structures were considered, and ways to compute them have been suggested.

When this research was started, it was intended to be mostly experimental work, aimed at investigating the efficiency of various aggregation and disaggregation schemes. However, because of the lack of adequate modelling support in the multi-stage area, this direction had to be temporarily abandoned.

* * * *

In what follows we will suggest research topics which are worth investigating in this author's opinion:

- In order to evaluate the efficiency of the proposed approach and of the different methods of disaggregation, experimentations have to be conducted under various conditions involving design parameters such as: varying demand patterns for the end products and for spare parts, accuracy of the demand forecasts, structure of the finished products, length of the planning horizon, capacity restrictions (from tight to loose), setup costs.
- The issue of feedback in hierarchical systems has not been satisfactorily approached yet. While it is obvious that the aggregate level passes down information in the form of capacity allocations, the only feedback from bottom up are the initial starting conditions. Some work has been done by Shwimer [82], but it was limited to a job shop environment.
- A substantial amount of work has been done in the area of job shop scheduling and dispatching (Conway et. al. [20]); however, for realistic situations theoretical analysis remains combinatorially locked, and this is why today we are confronted with an extremely large number of dispatching rules (see Panwalkar and Iskander [78]) whose efficiency could only be tested by simulation. It is hoped, and this would be interesting to investigate, that "good" aggregate planning and disaggregation would greatly alleviate the many problems related to machine loading (Holstein [48]), job sequencing and scheduling.

- In the section on safety stocks we have analyzed cases where we wanted most of the requirements, say 90-95-98%, served off the shelf. If this restriction can be relaxed we will be led to investigate where to locate inventories in the system and how large should they be as a function of the desired delivery time to the customer. Some results have been obtained: Hanssmann [27] (for stationary demand), Orlicky et. al. [76]. Some more research suggestions for MRP systems can be found in Berry and Whybark [7].
- Traditionally, production and distribution have been studied separately, but this is not the best approach (see critiques by Hax [43] and Aggarwal [1]). How should they interact with each other, and at what level should they be coordinated? A model to tackle this problem was presented by Krauss [60], who extended Geoffrion and Graves' [35] approach (by Benders decomposition) to allow for the simultaneous determination of a production schedule for each product class, by production process, at the various plants. Benli and Nanda [5] used decomposition to solve a distribution-production problem in which the production cost had two components: the fixed cost of constructing a plant to produce a certain product, and the variable production cost.
- How would multi-stage hierarchical planning apply to environments other than manufacturing (e.g., service organizations)?
- It is worth exploring the role MRP can play in the context of the multi-stage hierarchical approach, because there are many manufacturing organizations with MRP systems already in use.

- The hierarchical approach requires the investigation of a number of issues related to the design and implementation of such systems (for a presentation of the problem in general see Hax [45]); these issues go beyond the modelling aspects, and regard things like the distribution of decision making at various levels of management.

GLOSSARY

Aggregate part = a group of component parts manufactured or purchased at the same stage; abstract category created for aggregate planning purposes (see chapter 3).

Aggregate type = either a finished product type or an aggregate part.

Arborescent structure = each stage can have any number of successor stages but at most one predecessor stage.

Assembly line structure = all stages have at most one predecessor and at most one successor, with the exception of the final stage that can have more than one predecessor stage.

Assembly network with diverging arcs = each stage can have any number of predecessor or successor stages; the general type of production structure (see figure 1.2).

* * * *

Complete elementary production schedule = an elementary production schedule that produces all component parts, i.e. uses no parts from initial stock (as opposed to a special schedule).

Composition factor a_{sq} = how many units of s are required per unit of q; a_{sq} is defined only for pairs (s,q) that are in an immediate predecessor - immediate successor relationship to each other; in all other cases a_{sq} is zero.

* * * *

Dependent demand = demand for component parts, generated inside the planning model by a stage s upon its immediate predecessor stages.

DP = dynamic programming.

* * * *

Echelon stock = the number of units which have passed through stage s but have not been sold yet, i.e. they are still in the system.

Elementary production schedule for period t = a production schedule that can serve one unit of demand for the end product in period t.

* * * *

Facility = same as Stage.

* * * *

General structure system = same as Assembly network with diverging arcs.

* * * *

Independent demand = demand originating from outside the planning model; demand from customers, in general.

Installation stock = the amount of inventory stored at the stocking point immediately following a stage.

* * * *

Level = the set of all facilities separated, in the production structure network, by the same number of nodes from the finished good inventory point.

LP = linear program(ming).

* * * *

Network with diverging arcs = same as Assembly network with diverging arcs.

* * * *

Product type = a group of end products; abstract category created for aggregate planning purposes (see chapter 3).

Pure assembly system = each stage can have any number of predecessor stages, but only one successor stage.

* * * *

Special elementary production schedule = an elementary production schedule that uses some or all component parts from initial inventories, and produces the rest or none (as opposed to a complete schedule).

Stage = a pool of productive resource(s) that can be utilized to perform one or a number of operations upon the product; a stage can include one or more machines or/and work places. Stage and facility are used interchangeably.

REFERENCES

1. Aggarwal, S. - A Critique of "The Distribution System Simulator", by M.M. Connors, et. al., Management Science, Vol. 20, No. 4 December, Part I, 1973, pp. 482-486.

2. Aggarwal, S.C. - A Review of Current Inventory Theory and Its Applications. International Journal of Production Research, No. 12, 1974, pp. 443-482.

3. Andres, F.M. and Emmons, H. - On the Optimal Packaging Frequency of Products Jointly Replenished, Management Science, Vol. 22, No. 10, June 1976, pp. 1165-1166.

4. Anthony, R.N. - Planning and Control Systems: A Framework for Analysis, Harvard University, Boston, Mass. 1965.

5. Benli, O. and Nanda, P. - A Solution Procedure for Location-Allocation-Production Problems, Syracuse, New York, Department of Industrial Engineering and Operations Research, College of Engineering, Syracuse University, 1977.

6. Berry, W.L. - Lot Sizing Procedures for Requirement Planning System: A Framework for Analysis, Production and Inventory Management, Vol. 13, No. 2, 1972, pp. 19-34.

7. Berry, W.L. and Whybark, D.C. - Research Perspectives for Material Requirements Planning Systems, Production and Inventory Management, Vol. 16, No. 2, 1975, pp. 19-25.

8. Biggs, J.R., and Hahn, C.K. - The Impact of Capacity on Lot Sizing Decisions, in The Problems of Disaggregation in Manufacturing and Service Organizations - Conference Proceedings. The Ohio State University, March 10-11, 1977, Columbus, Ohio.

9. Bitran, G.R., and Hax, A.C. - On the Design of Hierarchical Production Planning Systems, Decision Sciences, Vol. 8, No. 1, January, 1977, pp. 28-55.

10. Bradley, S.P., Hax, A.C., and Magnanti, T.L. - Applied Mathematical Programming, Addison Wesley, Reading, Massachusetts, 1977.

11. Brown, R.G. - Statistical Forecasting for Inventory Control, McGraw-Hill, New York, 1959.

12. Brown, R.G. - Smoothing, Forecasting and Prediction, Prentice Hall, Englewood Cliffs, New Jersey, 1963.

13. Brown, R.G. - Decision Rules for Inventory Management, Holt, Rinehart and Winston, New York, 1967.

14. Bryan, J.G., Wadsworth, G.P. and Whitin, T.M. - A Multi-Stage Inventory Model, Naval Research Logistics Quarterly, Vol. 2, 1955, pp. 25-37.

15. Clark, A.J. - A Dynamic, Single-Item, Multi-Echelon Inventory Model, RM-2297, Santa Monica, California, The RAND Corporation, December 1958.

16. Clark, A.J. - An Informal Survey of Multi-Echelon Inventory Theory, Naval Research Logistics Quarterly, Vol. 19, 1972, pp. 621-650.

17. Clark, A.J. and Scarf, H. - Optimal Policies for a Multi-Echelon Inventory Problem, Management Science, Vol. 6, No. 4, July, 1960, pp. 475-490.

18. Clark, A.J. and Scarf, H. - Approxiate Solutions to a Simple Multi-Echelon Inventory Problem, in Arrow, Karlin, and Scarf: "Studies in Applied Probability and Management Science", Chapter 5, Stanford University Press, Stanford, Calif., 1962.

19. Collier, D.A. - The Effects of Joint Lot Size Strategies on Capacity Related Performance in a Multi-Stage Production-Inventory System, in The Problems of Disaggregation in Manufacturing and Service Organizations - Conference Proceedings. The Ohio State University, March 10-11, 1977, Columbus, Ohio.

20. Conway, R.W., Maxwell, W.L. and Miller, L.W. - Theory of Scheduling, Addison Wesley Publishing Co., Reading, Mass., 1967.

21. Crowe, W.R. - The Language of Lot Size, Operations Research, Vol. 19, No. 4, July-August 1971, pp. 992-997.

22. Crowston, W.B., Wagner, M. and Henshaw, A. - A Comparison of Exact and Heuristic Routines for Lot Size Determination in Multi-Stage Assembly Systems. Working Paper No. 475-70, Cambridge, Mass., Alfred P. Sloan School of Management, M.I.T., September 1970.

23. Crowston, W.B. and Wagner, M. - Lot Size Determination in Multi-Stage Assembly Systems, Working Paper No. 508-71, Cambridge, Mass., Alfred P. Sloan School of Management, M.I.T., January, 1971.

24. Crowston, W.B., Hausman, W.H. and Kampe, W.R. II.- Multi-Stage Production for Stochastic Seasonal Demand, Working Paper No. 587-72, Cambridge, Mass., Alfred P. Sloan School of Management, M.I.T., April 1972.

25. Crowston, W.B., Wagner, M. and Williams, J.F. - Economic Lot Size Determination in Multi-Stage Assembly Systems, Management Science, Vol. 19, No. 5, January 1973, pp. 517-527.

26. Crowston, W.B. and Wagner, M.H. - Dynamic Lot Size Models for Multi-Stage Assembly Systems. Management Science, Vol. 20, No. 1, September 1973, pp. 14-21.

27. Dantzig, G.B. - Optimal Solution to a Dynamic Leontief Model with Substitution, Econometrica, No. 23, 1955, pp. 295-302.

28. Dantzig, G.B. - On the Status of Multi-Stage Linear Programming Problems, Management Science, Vol. 6, October, 1959, pp. 53-72.

29. Darukhanavala, P.P. - An Economic Lot Size Algorithm for Multi-Level Inventory Systems with Acyclic Network Structures, in The Problems of Disaggregation in Manufacturing and Service Organizations - Conference Proceedings. The Ohio State University, March 10-11, 1977, Columbus, Ohio.

30. Dorsey, R.C., Hodgson, T.J. and Ratliff, H.D. - A Production-Scheduling Problem with Batch Processing, Operations Research, Vol. 22, No. 6, November-December 1974, pp. 1271-1279.

31. Dorsey, R.C., Hodgson, T.J., and Ratliff, D.H. - A Network Approach to a Multi-Facility, Multi-Product Production Scheduling Problem without Backordering, Management Science, Vol. 21, No. 7, March 1975, pp. 813-822.

32. Eilon, S. - Five Approaches to Aggregate Production Planning, AIIE Transactions, Vol. 7, No. 2, June 1975, pp. 118-131.

33. Gabbay, H. - Some Issues in Hierarchical Planning, Technical Report No. 112, Cambridge, Mass., Operations Research Center, M.I.T., June 1975.

34. Gabbay, H. - A Hierarchical Approach to Production Planning, Technical Report No. 120, Cambridge, Mass., Operations Research Center, M.I.T., December 1975.

35. Geoffrion, A.M. and Graves, G.W. - Multi-Commodity Distribution System Design by Benders Decomposition, Management Science, Vol. 20, No. 5, January 1974, pp. 822-844.

36. Golovin, J.J. - Hierarchical Integration of Planning and Control, Technical Report No. 116, Cambridge, Mass., Operations Research Center, M.I.T., September 1975.

37. Goodman, S.H. - Impact of a Product Structure Measure on the Effectiveness of Lot Sizing Techniques, in The Problems of Disaggregation in Manufacturing and Service Organizations - Conference Proceedings, The Ohio State University, March 10-11, 1977, Columbus, Ohio.

38. Goyal, S.K. - Determination of Optimum Packaging Frequency for Items Jointly Replenished, Management Science, Vol. 21, No. 4, December 1974, pp. 436-443.

39. Graves, S.C. and Schwarz, L.B. - Single Cycle Continuous Review Policies for Arborescent Production/Inventory Systems, Management Science, Vol. 23, No. 5, January 1977, pp. 529-540.

40. Hanssmann, F. - Optimal Inventory Location and Control in Production and Distribution Networks, Operations Research, Vol. 7, No. 4, July-August 1959, pp. 483-498.

41. Hanssmann, F. - Operations Research in Production and Inventory Control, Wiley, New York, 1962.

42. Hax, A.C. - Technical Note on Inventory Control, Case No. 4-172-173, Boston, Mass., Harvard University, 1971.

43, Hax, A.C. - A Comment on the "Distribution System Simulator", Management Science, Vol. 21, No. 2, October 1974, pp. 233-236.

44. Hax, A.C. and Meal, H.C. - Hierarchical Integration of Production Planning and Scheduling, in "Studies in Management Sciences", M.A. Geisler (editor), Vol. I, Logistics, North Holland - American Elsevier, 1975, pp. 53-59.

45. Hax, A.C. - The Design of Large Scale Logistics Systems: A Survey and an Approach in "Modern Trends in Logistics Research", W. Marlow (editor), M.I.T. Press, Cambridge, Mass., 1976, pp. 59-96.

46. Hax, A.C., Golovin, J.J., Bosyj, M. and Victor, T. - COMS: A Computer Based Operations Management System, Technical Report No. 121, Cambridge, Mass., Operations Research Center, M.I.T., January 1976.

47. Hax, A.C. - Aggregate Production Planning, in "Handbook of Operations Research", J. Moder and S.E. Elmaghraby (editors), N.Y. Van Nostrand Reinhold, New York, to appear.

48. Holstein, W.K. - Production Planning and Control Integrated, Harvard Business Review, Vol. 46, No. 3, May-June, 1968, pp. 121-140.

49. Holt, C.C., Modigliani, F., Muth, J.F., and Simon, H.A. - Planning Production, Inventories, and Work Force. Prentice-Hall, Englewood Cliffs, New Jersey, 1960.

50. Hu, C.T. and Prager, W. - Network Analysis of Production Smoothing, Naval Research Logistics Quarterly, Vol. 6, No. 1, March 1959, pp. 17-23.

51. Jensen, P.A. and Khan, H.A. - Scheduling in a Multi-Stage Production System with Setup and Inventory Costs, AIIE Transactions, Vol. 4, No. 2, 1972, pp. 126-133.

52. Johnson, L.A. and Montgomery, D.C. - Operations Research in Production Planning, Scheduling, and Inventory Control, Wiley, New York, 1974.

53. Kalymon, B.A. - A Decomposition Algorithm for Arborescence Inventory Systems, Operations Research, Vol. 20, No. 4, July-August 1972, pp. 860-874.

54. Kanodia, A.S. - Material Requirements Planning: A Study, Technical Report No. 128, Cambridge, Mass., Operations Research Center, M.I.T., November 1976.

55. Klingman, D.D., Soland, R.M., and Ross, G.T. - Optimal Lot-Sizing and Machine Loading for Multiple Products, in The Problems of Disaggregation in Manufacturing and Service Organizations - Conference Proceedings, The Ohio State University, March 10-11, 1977, Columnbus, Ohio.

56. Koenigsberg, E. - Production Lines and Internal Storage - A review, Management Science, Vol. 5, No. 4, July 1959, pp. 410-433.

57. Kraemer, S.A. and Love R.F. - A Model for Optimizing the Buffer Inventory Storage Size in a Sequential Production System, AIIE Transactions, Vol. 2, No. 1, May 1970, pp. 64-69.

58. Krajewski, L.J., Mabert, V.A., and Thompson, H.E. - Quadratic Inventory Cost Approximations and the Aggregation of Individual Products, Management Science, Vol. 19, No. 11, July 1973, pp. 1229-1240.

59. Krajewski, L.J. - Detailed Planning in Multi-Stage Systems - A Review of the Literature with Recommendations for Future Research. Presented at the Annual Midwest Meeting of the American Institute for Decision Sciences, San Francisco, California, November 1976.

60. Krauss, G.H. - An Approach to the Analysis of Integrated Production-Distribution Systems, in The Problems of Disaggregation in Manufacturing and Service Organizations - Conference Proceedings, The Ohio State University, March 10-11, 1977, Columbus, Ohio.

61. von Lanzenauer, C.H. - Production and Employment Scheduling in Multi-Stage Production Systems, Naval Research Logistics Quarterly, Vol. 17, No. 2, June 1970, pp. 193-198.

62. von Lanzenauer, C.H. - A Production Scheduling Model by Bivalent Linear Programming, Management Science, Vol. 17, No. 1, September 1970, pp. 105-111.

63. Lasdon, L.S. - Optimization Theory for Large Systems. The Macmillan Company, New York, 1970.

64. Lasdon, L.S. and Terjung, R.C. - An Efficient Algorithm for Multi-Item Scheduling, Operations Research, Vol. 19, No. 4, 1971, pp. 946-969.

65. Lee, S.M. - Goal Programming for Decision Analysis, Auerbach Publishers, Inc., Philadelphia, Pennsylvania, 1972.

66. Love, S.F. - A Facilities in Series Inventory Model with Nested Schedules, Management Science, Vol. 18, No. 5, Part I, January 1972, pp. 327-338.

67. Magee, J.F. and Boodman, D.M. - Production Planning and Inventory Control, McGraw-Hill, New York, 1967.

68. Mangiameli, P.M. - Large Scale Optimization for Master Production Scheduling, in The Problems of Disaggregation in Manufacturing and Service Organizations - Conference Proceedings, The Ohio State University, March 10-11, 1977, Columbus, Ohio.

69. Manne, A.S. - Programming of Economic Lot Sizes, Management Science, Vol. 4, No. 2, January 1958, pp. 115-135.

70. McGarrah, R.E. - Production and Logistics Management: Text and Cases, John Wiley and Sons, New York, 1963.

71. Meal, H.C. - Manufacturing Control and Material Requirements Planning, Paper presented at the Seventh Annual Meeting of the American Institute for Decision Sciences, Cincinnati, Ohio, November 1975.

72. Miller, J.G. - Hedging the Master Schedule, Working Paper, Boston, Mass., Graduate School of Business Administration, Harvard University, March 1977.

73. Montgomery, D.C. and Johnson, L.A. - Forecasting and Time Series Analysis, McGraw-Hill, New York, 1976.

74. New, C. - Safety Stocks for Requirements Planning, Production and Inventory Management, Vol. 16, No. 2, 1975, pp. 1-18.

75. Orlicky, J.A., Plossl, G.W. and Wight, O.W. - Material Requirements Planning Systems. Paper Presented at the 13th International Conference of the American Production and Inventory Control Society, Cincinnati, Ohio, October 1970.

76. Orlicky, J.A., Plossl, G.W. and Wight, O.W. - Structuring the Bill of Material for MRP, Production and Inventory Management, Vol. 13, No. 4, 1972, pp. 19-42.

77. Orlick, J.A. - Material Requirements Planning: The New Way of Life in Production and Inventory Management, McGraw-Hill, New York, 1975.

78. Panwalkar, S.S. and Iskander, W. - A Survey of Scheduling Rules, Operations Research, Vol. 25, No. 1, January-February, 1977, pp. 45-61.

79. Plossl, G.W. and Wight, O.W. - Production and Inventory Control, Prentice-Hall, Englewood Cliffs, New Jersey, 1967.

80. Schwarz, L.B. - A Simple Continuous Review Deterministic One-Warehouse N-Retailer Inventory Problem, Management Science, Vol. 19, No. 5, January 1973, pp. 555-566.

81. Schwarz, L.B. and Schrage, L. - Optimal and System Myopic Policies for Multi-Echelon Production/Inventory Assembly Systems, Management Science, Vol. 21, No. 11, July 1975, pp. 1285-1294.

82. Shwimer, J. - Interaction Between Aggregate and Detailed Scheduling in a Job Shop, Unpublished Ph.D. Thesis, Cambridge, Mass., Alfred P. Sloan School of Management, M.I.T., 1972.

83. Simpson, K.F., Jr. - In-Process Inventories, Operations Research, Vol. 6, No. 6, November-December 1958, pp. 863-873.

84. Solomon, M.J. - The Use of an Economic Lot Range in Scheduling Production, Management Science, Vol. 5, No. 4, July 1959, pp. 434-442.

85. Szendrovits, A.Z. - Manufacturing Cycle Time Determination for a Multi-Stage Economic Production Quantity Model, Management Science, Vol. 22, No. 3, November 1975, pp. 298-308.

86. Szendrovits, A.Z. and Wesolowsky, G.O. - Variations in Optimizing Serial Multi-Stage Production-Inventory Systems, in The Problems of Disaggregation in Manufacturing and Service Organizations - Conference Proceedings, The Ohio State University, March 10-11, 1977, Columbus, Ohio.

87. Taha, H.A. and Skeith, R.W. - The Economic Lot Sizes in Multi-Stage Production Systems, AIIE Transactions, June 1970, pp. 157-162.

88. Thomas, A.B. - Optimizing a Multi-Stage Production Process, Operational Research Quarterly, Vol. 14, No. 2, June 1963, pp. 201-213.

89. Vachani, M. - Determining Optimum Reject Allowances for Multi-Stage Job-Lot Manufacturing, AIIE Transactions, Vol. II, No. 1, March 1970, pp. 70-77.

90. Veinott, A.F., Jr. - The Status of Mathematical Inventory Theory, Management Science, Vol. 12, No. 11, July 1966, pp. 745-777.

91. Veinott, A.F., Jr. - Extreme Points of Leontief Substitution Systems. Linear Algebra and Its Applications, Vol. 1, 1968, pp. 181-194.

92. Veinott, A.F., Jr. Minimum Concave-Cost Solution of Leontief Substitution Models of Multi-Facility Inventory Systems, Operations Research, Vol. 17, March-April 1969, pp. 262-291.

93. Yollmann, T.E. - Capacity Planning: The Missing Link, Production and Inventory Management, Vol. 14, No. 1, 1973, pp. 61-74.

94. Whybark, D.C. and Williams, J.G.- Material Requirements Planning Under Uncertainty, Decision Sciences, Vol. 7, No. 4, October 1976, pp. 595-606.

95. Williams, J.F. - Multi-Echelon Production Scheduling when Demand is Stochastic, Management Science, Vol. 20, No. 9, May 1974, pp. 1253-1263.

96. Winters, P.R. - Constrained Inventory Rules for Production Smoothing, Management Science, Vol. 8, No. 4, July 1962, pp. 470-481.

97. Zabel, E. - On the Meaning and Use of a Capacity Concept, Naval Research Logistics Quarterly, Vol. 2, 1955, pp. 237-249.

98. Zabel, E. - Measures of Industry Capacity, Naval Research Logistics Quarterly, Vol. 3, 1956, pp. 229-244.

99. Zangwill, W.I. - A Deterministic Multi-Product, Multi-Facility Production and Inventory Model, Operations Research, Vol. 14, No. 3, May-June 1966, pp. 486-506.

100. Zangwill, W.I. - A Backlogging Model and a Multi-Echelon Model of a Dynamic Economic Lot Size Production System - A Network Approach, Management Science, Vol. 15, No. 9, May 1969, pp. 506-527.

101. Young, H.H. - Optimization Methods for Production Lines, Journal of Industrial Engineering, Vol. 18, No. 1, January 1967, pp. 70-78.

BIOGRAPHICAL NOTE

Dan Candea was born in Cluj, Romania, on November 29, 1946. He graduated from the Polytechnic Institute of Cluj in 1969 with the degree of Master of Mechanical Engineering, with Honors. He was an Assistant Professor at the same institute between 1969 - 1973.

In 1972 - 1973 he took part in an European contest and won a Ford Foundation fellowship in the European Doctoral Fellowship Program in Management Education. He entered the Doctoral Program in Management Science at the Alfred P. Sloan School of Management at the Massachusetts Institute of Technology, in 1973.

Mr. Candea received the Ph.D. from M.I.T. in 1977 and returned to the Polytechnic Institute of Cluj as an Assistant Professor.