IIE Transactions (2004) **36**, 611–628

ISSN: 0740-817X print / 1545-8830 online
DOI: 10.1080/07408170490278229

Parallelism of continuous- and discrete-time production planning problems

EUGENE KHMELNITSKY[1] and MICHAL TZUR[2,*]

[1]*Department of Industrial Engineering, Tel Aviv University, Tel Aviv 69978, Israel*
E-mail: xmel@eng.tau.ac.il
[2]*Department of Industrial Engineering, Tel Aviv University, Tel Aviv 69978, Israel*
E-mail: tzur@eng.tau.ac.il

Received February 2002 and accepted September 2003

We consider a capacitated production-inventory problem in both discrete- and continuous-time, stationary settings. In the discrete-time setting we analyze the infinite horizon capacitated dynamic lot-sizing problem, find the optimal solution and characterize its properties. For the continuous-time setting we formulate a new problem, which we claim to be an appropriate counterpart of the above discrete-time problem. No other counterpart model was found in the literature, including the vast literature on optimal control, which presumably deals with similar problems but in a continuous-time framework. The new problem formulation is the basis for a new class of models, which forms an alternative way of analyzing certain dynamic lot-sizing problems. This new alternative could sometimes be simpler than the analysis in the discrete case.

1. Introduction

When formulating an optimization problem, a modeler may choose either a discrete or continuous framework. In some cases, the derivation of the results may be simplified under one framework, compared to the other. The alternative method may reach similar results and insights, but in a more tedious and unstructured way. Hence, the ability to choose the "appropriate" or "convenient" framework is useful, provided that switching between the alternatives is easy.

In particular, discretization of time is a well-known idea. In the production-inventory field, one of the earliest examples is the seminal paper by Wagner and Whitin (1958), which aimed at solving a dynamic version of the continuous Economic Order Quantity (EOQ) problem. Their approach for solving the dynamic problem was by discretizing time, which resulted in the well-known dynamic lot-sizing model, then solved by dynamic programming.

The dynamic lot-sizing problem and its extensions have become one of the most investigated production-inventory problems. The motivation for analyzing these problems has evolved from approximating continuous problems to describing situations in which a decision may actually be taken only once in a period, thus leading naturally to the use of discrete time periods. As far as we know, no other works have used a variation of the dynamic lot-sizing problem in order to analyze its continuous counterpart, or *vice versa*.

The latter observation is somewhat surprising, given that many of the extensions of the basic dynamic lot-sizing problems are difficult problems, in particular some of those that involve capacity constraints; see for example, Florian and Klein (1971) and Florian *et al*. (1980). Moreover, although the broad area of *optimal control* deals with presumably similar problems but in a continuous framework, we found that none of the models considered there can serve as a continuous counterpart of lot-sizing problems.

In this paper we first offer an explanation for the above observation, which is related to the meaning of the production cost component, referred to as the *setup* or the *fixed* cost. We claim that the real meaning of this cost component has implications for the type of model that appropriately represents the problem, and therefore, in general, the two terms are not interchangeable. Then, we identify a class of models that has not been suggested previously in the literature, which forms a new continuous counterpart to the class of dynamic lot-sizing problems. The new class of models brings an alternative path to analyzing certain dynamic lot-sizing problems, which could sometimes be simpler than the analysis in the discrete case.

2. Modeling framework

Our focus in this section is on the meaning and modeling implications of the cost component, which is sometimes

*Currently visiting the Transportation Center and IE/MS Department, Northwestern University.

referred to as a setup cost, sometimes as a fixed cost, and is denoted in this paper as K in both cases.

A setup cost is a cost component, which is incurred whenever an initiation of production occurs. It represents a one-time cost of preparing the machine or the entire facility for production, for example, the cost of the workforce that makes the required preparations on the machine. In models where a batch is ordered from an outside supplier, the setup cost may be the amount charged by the supplier, which could represent, for example, the transportation cost associated with the delivery. The notion of a setup cost is well-known from the EOQ framework, where a cost is incurred whenever production of a batch is initiated. The production can be instantaneous, implying an infinite production rate, or it may take some time, implying a finite production rate. If the production rate is finite, the resulting model is sometimes referred to as the *Economic Production Quantity* (*EPQ*). In either case we refer to this family of models as EOQ. Regardless of the production rate, the setup cost is incurred exactly once, for a batch of any size.

On the other hand, a fixed cost is a cost component, which is incurred during production, that is, at any time when production occurs. It represents the cost of keeping the machine operating, for example, the electricity cost incurred when the machine operates, or the cost of the workforce that operates the machine. That is, the fixed cost depends on the time in which the machine is on, regardless of the quantity produced. Note that typically in production planning problems, there exists another cost component, the variable production cost, which is also incurred whenever production occurs. However, as opposed to a fixed cost, a variable production cost is quantity dependent, regardless of the length of time in which production occurs. (Only if the production rate is *a-priori* constant, then do the fixed and variable costs coincide.)

Thus, the common feature of both the setup and the fixed cost is their independence of the quantity produced. This is also the reason why they are sometimes referred to as interchangeable costs. However, we concur that the real meaning of this cost component is important for an appropriate modeling of the problem. It turns out that the difference in modeling is much more notable in the continuous framework than in the discrete framework, as discussed below.

In order to focus on the new modeling framework and to obtain clear insights, we focus in this paper on a stationary parameter environment, that is, on problems in which all the parameters, and in particular the demand, are constant over time. We consider a capacitated setting, that is, a setting in which the amount produced at any time (or in a given period) is limited.

Consider, for example, the EOQ problem. In this problem, the demand rate is constant over an infinite horizon, a setup cost is incurred whenever an initiation of production occurs, and holding costs are incurred per unit of inventory held, per unit of time. The closest discrete counterpart of the EOQ is the dynamic lot-sizing problem, in which demand in every period is constant, and the same cost components as in the EOQ are present. Even though typically in the dynamic lot-sizing problem the demand varies from period to period, we still refer to it as the discrete counterpart of the EOQ, since it is merely a special case of the dynamic demand case. With a capacity constraint, we obtain the capacitated EOQ and the capacitated dynamic lot-sizing problems, respectively.

Now consider how the above two problems (EOQ and dynamic lot sizing) change, when the setup cost is replaced by a fixed cost. For the discrete case, the dynamic lot-sizing problem does not change, only the meanings of some of its components change. In particular, a period represents the length of time available for production and K represents the fixed cost of producing during this length of time. The production in two consecutive periods is in fact continuous and therefore the fixed cost is proportional to the number of periods in which production occurs and is not related to the initiation of production as in the case of a setup cost. However, for the continuous case, there is no variation of the EOQ (or EPQ) model, which properly accounts for a fixed cost component without also including a setup cost component. The setup cost component in these models is related to the initiation of production, as in the previous context. In Silver (1990) both the setup and fixed cost components are considered, but if the setup cost and time are set to zero, the model is no longer realistic since it results in a chattering regime, see below. (See Wolsey (1989) for a discrete model that considers both setup and fixed costs.) Instead, a natural framework to analyze the continuous case with a fixed cost and without a setup cost appears to be the optimal control framework. Moreover, in the optimal control framework the production rate is allowed to vary. The above discussion is summarized in Table 1.

Focusing now on the optimal control cell of Table 1, we note that production-inventory planning problems, analyzed by the optimal control methodology, have been studied extensively in the literature. We mention here the pioneering work by Hwang *et al.* (1967) who modeled a simple problem of aggregate production planning in a continuous-time form. Bensoussan *et al.* (1983) considered both discrete- and continuous-time production planning problems, and within the continuous-time framework they considered both continuous and impulse control formulations. More recent works for example Sethi and Zhang (1995), Maimon *et al.* (1998) and Dauzere-Peres *et al.* (2000), have extensively studied the continuous-time production control models in deterministic and stochastic environments. The solution methodology is usually based on

Table 1. Modeling framework

Meaning of K/Time	*Continuous*	*Discrete*
Setup cost	EOQ	Lot sizing
Fixed cost	Optimal control	Lot sizing

either Hamilton-Jacobi-Bellman dynamic programming or the Pontryagin maximum principle. For linear costs and simple demand functions (constant, cyclic, etc.) the optimal production can be obtained in a closed form. For more complicated cases, development of specific numerical procedures is required.

However, even with the existing extensive optimal control literature on production-inventory problems, we claim that an appropriate counterpart of the lot-sizing problem with fixed cost, has not yet been analyzed. One important reason is that a straightforward "translation" of the discrete model to a continuous one typically results in a production regime known as *chattering*. Under a chattering regime, the production rate jumps from zero to its maximum value (back and forth) an infinite number of times at any (even a very small) time interval. Clearly a chattering regime is not practical, and cannot be implemented without modifications. Moreover, the ability to change the production rate at an infinite speed at any time is also not a realistic assumption. We conclude that a new continuous-time model is required, which overcomes both difficulties. Such a model is suggested in this paper.

The rest of the paper is organized as follows. In Section 3 we describe the discrete-time model, discuss known and new results and present the optimal solution for the problem. In Section 4 we introduce the new continuous-time model, analyze its properties and obtain its solution. Finally, in Section 5 we discuss analogies between the results of the previous two sections. We also define general guidelines for a modeling analogy, based on which we further justify the suggested new continuous-time model.

3. The discrete-time model

We consider a discrete-time production inventory problem over an infinite planning horizon. Our analysis is based, in part, on results from the literature for the finite horizon version of the problem; those results are reviewed in Section 3.1. In Section 3.2 we provide a new analysis for the infinite horizon problem. We derive structural properties and obtain the optimal solution for this problem.

3.1. *The finite horizon problem*

The finite horizon problem, known as the capacitated dynamic lot-sizing problem, is defined by the following parameters:

T = number of periods;
d = demand in each period;
h = holding cost per unit per period;
p = variable production cost per unit;
K = fixed cost, incurred in any period in which production occurs;
C = production capacity per period.

To solve this problem, one needs to determine the production quantity in each period. A feasible production plan has to satisfy demand on time without violating the capacity constraints. The objective is to minimize the total cost (this is equivalent to minimizing the average total cost per period), which consists of fixed and variable production costs, and holding costs. We refer to this problem as (F-CDLSP) (Finite Horizon, Capacitated Dynamic Lot-Sizing Problem) and a Mixed Integer Linear Programming (MILP) formulation of it is given in Appendix A.

Note that the total variable production cost, p per unit, is contributing a constant to the objective value under any feasible policy, and therefore may be ignored. (F-CDLSP) with non-stationary demand and capacity parameters was shown to be NP-hard by Florian *et al*. (1980). In fact, it is NP-hard even in many special cases, as was shown in Bitran and Yanasse (1982). However, for stationary capacity parameters, (F-CDLSP) has a polynomial-time solution, see Florian and Klein (1971) and more recently Van Hoesel and Wagelmans (1996). Some results from the above mentioned papers are reviewed next, and will be used in the analysis of the infinite horizon problem.

As in the MILP formulation, we denote Q_t as the production quantity in period t and S_t as the inventory level at the end of period t. A production in period t is called *full* if it equals C, i.e., $Q_t = C$. A production in period t is called *fractional* if it is strictly between zero and C, i.e., $0 < Q_t < C$.

Lemma 1. (*Florian and Klein*, 1971): *There exists an optimal schedule such that between any pair of fractional production periods there is at least one period with a zero inventory. This property is often referred to as the fractional production property.*

The lemma is a direct result of the theory on concave cost network flows, applied to the network describing our problem. See Denardo (1982) for additional dynamic lot-sizing applications. As a result of Lemma 1, we have:

Corollary 1. *There exists an optimal schedule which satisfies the following*; *between any two consecutive periods with zero inventory*, *there exists at most one fractional production period.*

Due to the corollary, the following definition is used.

A *subplan* $(t_1, t_2)(1 \le t_1 \le t_2 \le T)$ is a set of consecutive periods, starting with period t_1 and ending with period t_2, such that $S_{t_1-1} = S_{t_2} = 0$, $S_t > 0$ for all $t_1 - 1 < t < t_2$ and at most one period within the subplan has fractional production period.

Thus, a subplan (t_1, t_2) consists of at most one fractional production period and as many full production periods as required to cover all demand within the subplan. We refer to the rest of the periods as *zero production periods*.

The following lemma is due to Baker *et al*. (1978).

Lemma 2.

(a) *For any subplan* (t_1, t_2), *the fractional production period* (*if one exists*) *is the first period of the subplan.*

(b) *The locations of the full production periods within a subplan are as late as possible (without incurring a backlog).*

Intuitive arguments for this lemma are as follows:

(a) Suppose that in the optimal solution, the statement does not hold. Then at the beginning of the fractional period there is positive inventory, and a reduction in cost can occur by reducing the amount of beginning inventory by ε, and increasing the production amount in the fractional period by ε. This is a contradiction to the optimality of the suggested solution.
(b) This part is true due to the fact that the fixed costs in a cycle are determined by the number of periods in the subplan, and the above policy minimizes the holding costs.

3.2. *The infinite horizon problem*

We now consider the infinite horizon version of problem (F-CDLSP), i.e., when $T \to \infty$, and the objective function is the limit (when $T \to \infty$) on the average cost. We refer to this problem as (CDLSP).

Theorem 1 proves the periodicity property of the solution, for the case of rational demand and capacity parameters.

Theorem 1. *If the d and C parameters are rational, then there exists an optimal solution for the discrete-time problem which is cyclic and consists of subplans of identical length. The length of those subplans is bounded by:* $\bar{n} \equiv \min\{n: nd/C \text{ is an integer}\}$.

Proof. In Appendix B.

Define *cycle* (n) of length n, as a subplan (t_1, t_2) in which $t_2 - t_1 + 1 = n$. Since all parameters are stationary, all subplans with the same number of periods are identical. Note that since the horizon of a subplan is finite, Lemmas 1 and 2 hold.

For ease of notation, and without loss of generality, we slightly change the definition of a fractional period. Since from now on we consider only solutions which consist of subplans, we define a fractional production period to be as before (strictly between zero and C) when $nd \bmod(C) \neq 0$ and to be exactly C units (in the first period of the subplan), when $nd \bmod(C) = 0$. As a result, all cycles have exactly one fractional period and we have the following corollary:

Corollary 2.

(a) *The optimal schedule of cycle* (n) *consists of one fractional production period of size*:

$$f(n) = \begin{cases} nd \bmod(C) & \text{if } nd \bmod(C) \neq 0, \\ C & \text{if } nd \bmod(C) = 0, \end{cases}$$

$c(n)$ full production periods where $c(n) = (nd - f(n))/C$, and $z(n)$ zero production periods where $z(n) = n - c(n) - 1$.

(b) *Only values of n for which $f(n) \geq d$ are possible cycle length values. (Otherwise, the fractional production in the first period of the cycle would not cover the demand of that period.)*

Next we would like to explore some properties of the optimal solution, which are of interest and are useful for the desired comparison with the continuous version of the problem. As a result of Corollary 1 and Corollary 2 above, the optimal solution could consist of one of the following three types of cycles:

Cycle type 1: The production in every period is d. That is: $Q_t = d \ \forall t$, $n = 1$. As a result, no inventory is ever carried. This cycle type is feasible for every $0 < d \leq C$.
Cycle type 2: There is one fractional production period, at least one full production period, and at least one zero production period. This cycle type is feasible for every $0 < d < C$.
Cycle type 3: There is one fractional production period, one or several zero production periods and no full production periods. This cycle type is feasible only for $0 < d \leq C/2$.

We refer to a solution that is a representative of cycle type i as solution type i, $i = 1,2,3$. Note that given n, the values of $f(n)$, $c(n)$ and $z(n)$ are determined according to Corollary 2. Given those values, the cycle type is determined according to the above cycle type definitions.

We conclude that solving the discrete-time problem consists in finding the best cycle length $1 \leq n^* \leq \bar{n}$. Once n^* is determined, the optimal schedule within the cycle is determined by Lemma 2.

The next lemma presents a closed-form formula for the average cost per period as a function of the cycle length, n. Note that this formula is independent of the cycle type, as for each value of n only one cycle type is well-defined. The lemma is based on the following condition, which is also used in other places in the paper: *cycle*(n) is feasible if $f(n)$ satisfies: $d \leq f(n) \leq C$ (Corollary 2(b)). Recalling that $f(n) = nd - c(n)C$, using it in the above inequalities, dividing by d and using the integrality of n, we get:

$$\left\lceil \frac{c(n)C}{d} \right\rceil + 1 \leq n \leq \left\lfloor \frac{(c(n)+1)C}{d} \right\rfloor. \tag{1}$$

In Equation (1) and in the following, the term $\lceil x \rceil$ corresponds to the smallest integer which is greater than or equal to x, and the term $\lfloor x \rfloor$ corresponds to the largest integer which is smaller than or equal to x.

Lemma 3. *The average cost per period of a cyclic solution as a function of n is*:

$$j(n) = \frac{c(n)+1}{n} K + \frac{1}{2} hd(n-1) - Chc(n) + \frac{Ch}{n} \sum_{i=1}^{c(n)} \left\lceil \frac{iC}{d} \right\rceil. \tag{2}$$

Table 2. The observed production and inventory quantities

Period	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Q_t	10	20	0	20	0	20	0	20	0	0
S_t	1	12	3	14	5	16	7	18	9	0

Proof. In Appendix C.

To demonstrate the behavior of the holding costs in the above cost function, consider the following example:

Example 1: $d = 9$; $C = 20$; $h = 1$; $K = 15$; $n = 10 \Rightarrow c(n) = 4, f(n) = 10$.

The production and inventory quantities listed in Table 2 are observed during the cycle. There are five production periods and a total of 85 units carried during the cycle, therefore the average cost per period is: $(5 \times 15 + 85)/10 = 16$. One can verify that this is also obtained by Equation (2).

For each point in the (C, d)-plane, the result of minimizing $j(n)$ as a function of n determines the optimal cycle type for that point. Therefore, according to the resulting optimal cycle type, the (C, d)-plane may be partitioned into three regions. In Fig. 1(a) we show our conjecture (which is proved for some cases) regarding how these regions look. The broken line $d = C$ in the figure distinguishes between the feasible area ($C \geq d$) and the non-feasible area ($C < d$). Bold digits 1, 2 or 3 denote the regions where cycle types 1, 2 or 3 are optimal, respectively. Note that region 2 of the (C, d)-plane (where cycle type 2 is optimal) is not continuous. It consists of an infinite number of closed "wings", each located right above the lines $d = C/r$, $r = 2, 3, \ldots, \infty$, as depicted in bold in the figure. To explain the partition of Fig. 1(a), we consider the following five cases, corresponding to the lines in the (C, d)-plane that distinguish between the regions 1, 2 and 3. We conjecture the exact form of the lines in cases 1–3 and prove it in cases 4–5. The conjectures in cases 1–3 are based on and confirmed by an extensive simulation study that we performed, with a wide variety of parameter combinations.

Case 1: For $C > 2K/h$, the line $d = K/h$ distinguishes between optimal cycle types 1 and 3.

Indeed, solution type 3 has the cost function:

$$j(n) = \frac{K}{n} + \frac{1}{2}hd(n-1), \tag{3}$$

which is obtained by substituting $c(n) = 0$ in Equation (2). Comparing Equation (3) as a function of n with the cost of solution type 1, K, we obtain that solution type 1 is better when:

$$d > \max_{n \geq 2} \frac{2K}{nh} = \frac{K}{h}.$$

Since solution type 2 appears to have a higher cost in this area of the (C,d)-plane, we conjecture that it does not influence the comparison; see an example and an intuitive explanation in Appendix D.

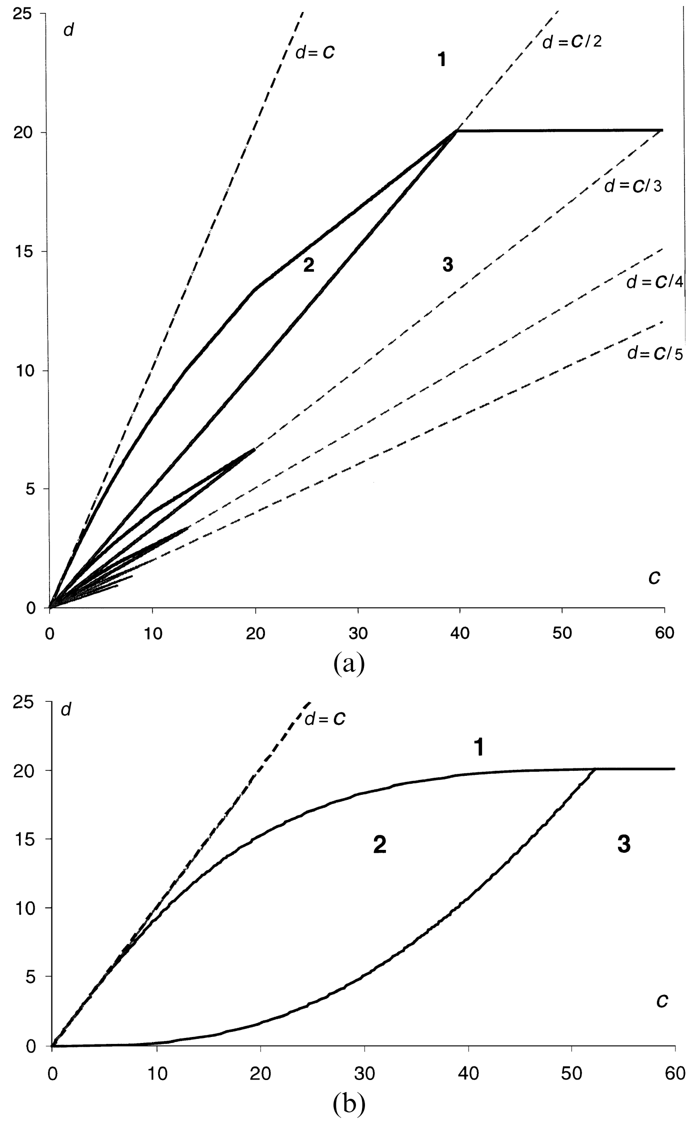

Fig. 1. Partition of the (C, d)-plane for: (a) the discrete-time problem; and (b) the continuous-time problem.

Case 2: For each $m = 1, 2, \ldots, \infty$, the line:

$$d = C\frac{2K/Ch + m(m+1)}{(m+2)(m+1)},$$

distinguishes between optimal cycle types 1 and 2 in the interval:

$$\frac{1}{m+1}\frac{2K}{h} < C < \frac{1}{m}\frac{2K}{h}.$$

Indeed, cycle type 3 does not exist in this area since $d > C/2$ for all m, and cycle type 2 exists for only $c(n) \geq m$; otherwise, for $c(n) < m$ no feasible n exists, see Equation (1). We conjecture that the best solution of this type is characterized by $c(n) = m$ and $n = m + 2$. By comparing this solution with the solution of cycle type 1, i.e., $c(n) = 0$ and $n = 1$, we obtain the line distinguishing between cycle types 1 and 2 for each m.

Case 3: For each $r = 2, 3, \ldots, \infty$ and $m = 1, 2, \ldots, \infty$, the line:

$$d = C\frac{2K/Ch + m(m+1)r^2}{r(mr + r + 1)(mr + 1)},$$

distinguishes between optimal cycle types 2 and 3 in the interval:

$$\frac{1}{r(m+1)}\frac{2K}{h} < C < \frac{1}{rm}\frac{2K}{h}.$$

Here we conjecture that the best solution of cycle type 2 is characterized by $c(n) = r - 1$ with $n = mr + 1$ and that the best solution of cycle type 3 is characterized by $c(n) = 0$ and $n = m$. By comparing the costs of these specific solutions, we obtain the line distinguishing between optimal cycle types 2 and 3. Since solution type 1 appears to have a higher cost in this area of the (C,d)-plane, it does not influence the comparison.

Unlike the previous three cases where some conjectures have been adopted, in the following two cases we rigorously prove that for each $r = 2, 3, \ldots, \infty$ the line $d = C/r$, distinguishes between optimal cycle types 2 and 3 in the interval $0 \le C < 2K/((r-1)h)$.

Case 4: For each $r = 2, 3, \ldots, \infty$, a small $\varepsilon > 0$ and $d = (C(1-\varepsilon)/r)$, cycle type 3 is optimal in the interval $0 \le C < 2K/((r-1)h)$. The optimal cycle length n^* is r, and $c(n^*) = 0$.

We can prove this case by considering that for a sufficiently small ε, and for each $i = 1, \ldots, c(n)$, we have:

$$\left\lceil \frac{iC}{d} \right\rceil = \left\lceil \frac{ir}{1-\varepsilon} \right\rceil = ir + 1.$$

By substituting this expression into Equation (2), we find:

$$j(n) = \frac{c(n)+1}{n}K + \frac{1}{2}hd(n-1) - Chc(n) + \frac{Ch}{n}c(n) + \frac{Ch}{2n}r\,c(n)(c(n)+1). \quad (4)$$

Let $c(n) = 0$. Then, from Equation (1), $1 \le n \le r$. From Equation (4), and by noting that $j(n)$ is convex, the optimal n is:

$$n = \sqrt{\frac{2Kr}{Ch(1-\varepsilon)}},$$

rounded up or down, whichever has the lower $j(n)$ value. Since:

$$\sqrt{\frac{2Kr}{Ch(1-\varepsilon)}} > \sqrt{\frac{(r-1)r}{(1-\varepsilon)}},$$

the right hand side of this inequality equals either r or $r - 1$, and $n \le r$, we conclude that n itself would equal either r or $r - 1$, whichever has the lower $j(n)$ value. By comparing $j(r)$ with $j(r-1)$, we find that for the case under consideration $n = r$.

Let $c(n) > 0$. For a fixed $c(n)$, and by plugging d into Equation (2), we observe that $j(n)$ is convex and that:

$$\left.\frac{\partial j(n)}{\partial n}\right|_{n=(c(n)+1)r} < 0.$$

Since $n = (c(n)+1)r$ is the upper limit of n, we conclude that n takes its maximum possible value, i.e., $n = (c(n)+1)r$. Now, by minimizing $j(n)$ for $n = (c(n)+1)r$ as a function of $c(n)$, we find that the minimum is achieved for $c(n) = 1$, i.e., $n = 2r$. From Equation (4) we calculate $j(r)$ and $j(2r)$, which are optimal for the cases $c(n) = 0$ and $c(n) > 0$ respectively, and conclude that $j(r) < j(2r)$. Thus, $n^* = r$, $c(n^*) = 0$, which, by definition, means that cycle type 3 is optimal.

Case 5: For each $r = 2, 3, \ldots, \infty$, a small $\varepsilon > 0$ and $d = (C(1+\varepsilon)/r)$, cycle type 2 is optimal in the interval $0 \le C < 2K/((r-1)h)$.

We can prove this case by considering that for a sufficiently small ε and for each $i = 1, \ldots, c(n)$, we have:

$$\left\lceil \frac{iC}{d} \right\rceil = \left\lceil \frac{ir}{1+\varepsilon} \right\rceil = ir.$$

By substituting this expression into Equations (2) and (1), we find:

$$j(n) = \frac{c(n)+1}{n}K + \frac{1}{2}hd(n-1) - Chc(n) + \frac{Ch}{2n}rc(n)(c(n)+1),$$
$$c(n)r + 1 \le n \le (c(n)+1)r - 1.$$

Since for a fixed $c(n)$, $j(n)$ is convex and:

$$\left.\frac{\partial j(n)}{\partial n}\right|_{n=(c(n)+1)r-1} < 0,$$

we conclude that n takes its maximum possible value, i.e., $n = (c(n)+1)r - 1$. Now, by minimizing $j(n)$ for $n = (c(n)+1)r - 1$ as a function of $c(n)$, we find that the minimum is achieved for:

$$n = \sqrt{\frac{1}{\varepsilon}\left(\frac{2K}{Ch} - r + 1\right)},$$

rounded up or down, which is a particular solution of cycle type 2.

4. A new continuous-time model

In this section we present a new continuous-time production control model. We build the model so that it represents the problem scenario of (CDLSP) of the previous section as closely as possible but within a reasonable model complexity.

In contrast to the previous section, continuous-time production planning models assume that the decision about the production quantity is made at each point in time, rather than once per period. As a result, some of the notions introduced in the previous section change, although their physical meaning is similar. The fixed cost, K, and the holding cost per unit, h, become cost rates (measured in dollars per time unit); the production and demand quantities $Q(t)$ and d, respectively, become rates (measured in units per time unit) which satisfy the following material balance equation:

$$\dot{S}(t) = Q(t) - d \quad \text{for } 0 \leq t < \infty, \quad S(0) = 0, \tag{5}$$

where $\dot{S}(t)$ represents the rate of change in the inventory level at time t.

Backlogs are disallowed and the production rate is bounded from above by a finite capacity, C:

$$S(t) \geq 0, \quad 0 \leq t < \infty, \tag{6}$$

$$0 \leq Q(t) \leq C, \quad 0 \leq t < \infty. \tag{7}$$

The problem is to minimize the total cost of inventory and production per time unit:

$$\min J = \lim_{T\to\infty} \frac{1}{T} \int_0^T [hS(t) + \text{Cost}(Q(t))]\text{d}t, \tag{8}$$

where

$$\text{Cost}(Q(t)) = \begin{cases} K + pQ(t) & \text{if } Q(t) > 0, \\ 0 & \text{if } Q(t) = 0, \end{cases}$$

reflects both the fixed cost K which is incurred at each interval of time $(t, t + \text{d}t)$ where $Q(t) > 0$, and the variable production cost which is proportional to the production rate $Q(t)$ at time t(with a coefficient p); h is the cost of holding a unit in inventory for a unit of time. Therefore, the cost in Equation (8) has to be minimized subject to constraints (5)–(7). We refer the reader to Remark 1 at the end of this model description for a further discussion on the choice of the above cost function.

The optimal production schedule of the described system leads to the "chattering" regime at which $Q(t)$ undergoes an infinite number of jumps from $Q(t) = 0$ to $Q(t) = C$ and back to $Q(t) = 0$ at any (even very small) time interval, see Steindl (2001). That is:

$$Q(t) = \lim_{\Delta\to 0} \begin{cases} C, & (k-1)\Delta \leq t < (k-1)\Delta + (d/C)\Delta \\ 0, & (k-1)\Delta + (d/C)\Delta \leq t < k\Delta \\ & \qquad k = 1, \ldots, \infty. \end{cases},$$

The chattering solution is the best one, having a zero inventory and a minimum possible fixed cost, $(d/C)K$ (in \$ per time unit). If, for example, $C = 10$, $K = 20$, $h = 1$ and $d = 8$, then in $8/10 = 80\%$ of the time we produce at the rate $C = 10$ and in the remaining 20% of the time we do not produce. Thus, the cost is \$16 per time unit. The stationary solution for this example, $Q(t) = 8$, results in costs of $K = 20$ (\$ per time unit).

To avoid chattering while properly modeling the problem, we restrict the change of the production rate as follows:

$$\dot{Q}(t) = v(t), \quad 0 \leq t < \infty, \tag{9}$$

$$|v(t)| \leq M, \quad 0 \leq t < \infty. \tag{10}$$

Here, an additional decision variable $v(t)$ controls the change in the production rate $\dot{Q}(t)$ at time t. The parameter M denotes the maximum allowed rate of change in $Q(t)$. Constraints (9) and (10) reflect the "inertia" effect that is closely related to a setup which re-configures a production system to produce at a different rate. For the above example with $M = 111$, the optimal solution turns out to be a cyclic one with a cycle length of 1.5 time units and a cost of \$18.33 per time unit (see details of the solution below in this section). We refer the reader to Remark 2 for a further discussion on the meaning of this constraint and to guideline 2 in Section 5 for a discussion on the choice of the parameter M. The condition $d \leq C$ is required to ensure feasibility. The variable production cost p is now dropped, since the cost associated with it is a constant under any feasible solution.

Before we proceed to discuss the optimal solution of the model just presented, let us make two important remarks that are related to our modeling choices. In particular, we address here the analogy between the new model and the (CDLSP) of the previous section.

Remark 1. It may seem at first that since K in the continuous model is a rate, incurred whenever producing and not only at the initiation of production, it is not an analogous situation to the setup cost in the discrete problem. Note, however, that in the discrete case K is incurred, in fact, in a similar way. It is incurred in every time period when there is production, regardless of whether or not there was production in the previous period. Therefore, also in the discrete case K is incurred whenever producing, not whenever initiating production. (In some models, an additional cost is considered, referred to as a *startup cost*, which is incurred when a setup cost is incurred in a given period but not in the previous period, see for example, Wolsey (1989). We do not consider here such costs.) In other words, the analogy exists since the fixed cost in the discrete case is proportional to the number of periods in which production occurs, while in the continuous case it is proportional to the length of time in which production occurs. When the period length in the discrete model is short, the analogy is even more apparent. See also the discussion in Section 2 with respect to the model associated with the lower right corner of Table 1.

Remark 2. The constraint on the change in the production rate in the continuous model does not have an apparent counterpart in the discrete model. However, when considering carefully the meaning of such a restriction in the discrete case, one realizes that a similar restriction, in fact does exist, due to the fact that a period in this model has a duration. In other words, going from no production to full production (or *vice versa*) in two consecutive periods takes an amount of time, which is equal to the length of the period (or part of it). You can imagine that within a period, part of the time may be used to accelerate the production, part of it to production, and so on. Since this is a discrete model we do not observe the activities within a period, only their total. Therefore, the change in production is not instantaneous. If the period length in the discrete-time problem is very short, this can be modeled in the continuous case by choosing an appropriately large value of M. The constraint in the continuous case merely helps to avoid the chattering regime which is not realistic and which is not feasible in the discrete case.

To conclude the above discussion we argue based on the above arguments that the new continuous model is analogous to the (CDLSP). In Section 5 we somewhat formalize the notion of "analogous." While we are not aware of any other model that can be claimed to be analogous, it is possible that in the future a new and different model will be developed, that will also be claimed to be analogous to (CDLSP). Our claim does not preclude this possibility, and therefore we refer to our model as *an* analogous, not necessarily *the* only analogous model that can be envisioned. The reason for possible multiple models claiming to be analogous to (CDLSP) is that the discrete and continuous models will never be exactly the same. See further discussion in Section 5.

We now turn to analyzing the optimal solution of the above new continuous model. The following lemma characterizes the optimal production schedule. The proof is based on the maximum principle and is placed in Appendix E (see a comprehensive survey of the maximum principles in Hartl *et al*. (1995)).

Lemma 4. *An optimal solution to problem* (5)–(10) *possesses the following three properties*.

1. *It consists of a sequence of four types of regimes*:
   full production (*FP*), *that is*: $Q(t) = C$;
   no production (*NP*), *that is*: $Q(t) = 0$;
   adequate production (*AP*), *that is*: $Q(t) = d$;
   changeover regime (*CO*), *that is*: $v(t) = M$ *or* $v(t) = -M$.
2. $S(t) = 0$ *when* t *is at the AP regime*.
3. *A CO regime*, *which changes the production rate from value* Q_1 *to value* Q_2, *can occur only at an interval of time* $[t, t + \tau]$ *for which*:

$$\frac{1}{\tau}\int_t^{t+\tau} \psi(t')\mathrm{d}t' = \frac{Cost(Q_2) - Cost(Q_1)}{Q_2 - Q_1}, \qquad (11)$$

where the function $\psi(t)$ *satisfies the co-state equation:*

$$\mathrm{d}\psi(t) = h\mathrm{d}t - \mathrm{d}\mu(t),$$

with unspecified boundary conditions. Here, $\mu(t)$ *is a non-decreasing function*, $\mathrm{d}\mu(t) \geq 0$, *satisfying the complementary slackness condition*:

$$\mathrm{d}\mu(t) = 0 \quad \textit{if } S(t) > 0.$$

Proof. In Appendix E.

In Theorem 2 we define five types of cycles that characterize the solution. For cycle types 2–5 we depict the solutions schematically in Figs. 2–5 (for cycle type 1 the figure is trivial). In these figures T_{per} represents the cycle length and y is the length of the first regime in each cycle. Plots of $Q(t)$ and $S(t)$ are depicted in part (a) of each figure, where $S(t)$ is calculated by integrating Equation (6) over t. The co-state variables $\psi(t)$, which are depicted in part (b) of each figure, are discussed in the proof of Lemma 4 and Theorem 2.

Theorem 2. *The following cyclic solutions* (*cycle types*) *satisfy the properties of optimality stated in Lemma* 4 (*see Figures* 2–5):

Cycle type 1: $Q(t) = d$ *for* $t \in (0, \infty)$. *The solution is constant and therefore can be regarded as a cyclic one with any cycle length* $T_{\text{per}} > 0$. *It exists for all values of* d, *i.e.*, *for* $0 < d \leq C$.

Cycle type 2:

$$Q(t) = \begin{cases} 0 & \textit{if } t \in [0, y], \\ M(t - y) & \textit{if } t \in [y, y + C/M], \\ C & \textit{if } t \in [y + C/M, T_{\text{per}} - C/M]. \\ -M(t - T_{\text{per}}) & \textit{if } t \in [T_{\text{per}} - C/M, T_{\text{per}}], \end{cases}$$

where:

$$y = T_{\text{per}}\left(1 - \frac{d}{C}\right) - \frac{C}{M}.$$

The solution is cyclic with:

$$T_{\text{per}} \geq \frac{C^2}{Md(C - d)} \max\{d, C - d\}.$$

It exists for $0 < d < C$.

Cycle type 3:

$$Q(t) = \begin{cases} Mt & \textit{if } t \in [0, y] \\ -M(t - 2y) & \textit{if } t \in [y, 2y], \\ 0 & \textit{if } t \in [T_{\text{per}} - 2y, T_{\text{per}}] \end{cases}$$

where:

$$y = \sqrt{\frac{dT_{\text{per}}}{M}}.$$

The solution is cyclic with:

$$\frac{4d}{M} \leq T_{\text{per}} \leq \frac{C^2}{Md}.$$

It exists for $0 < d \leq C/2$.

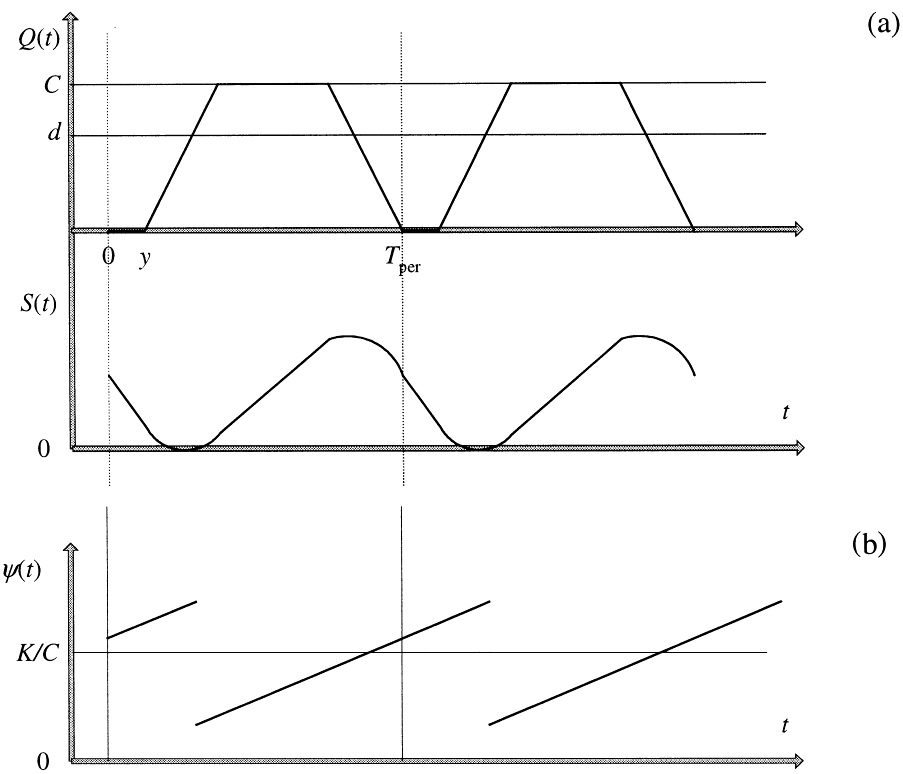

Fig. 2. A solution of cycle type 2: (a) state variables; (b) co-state variables.

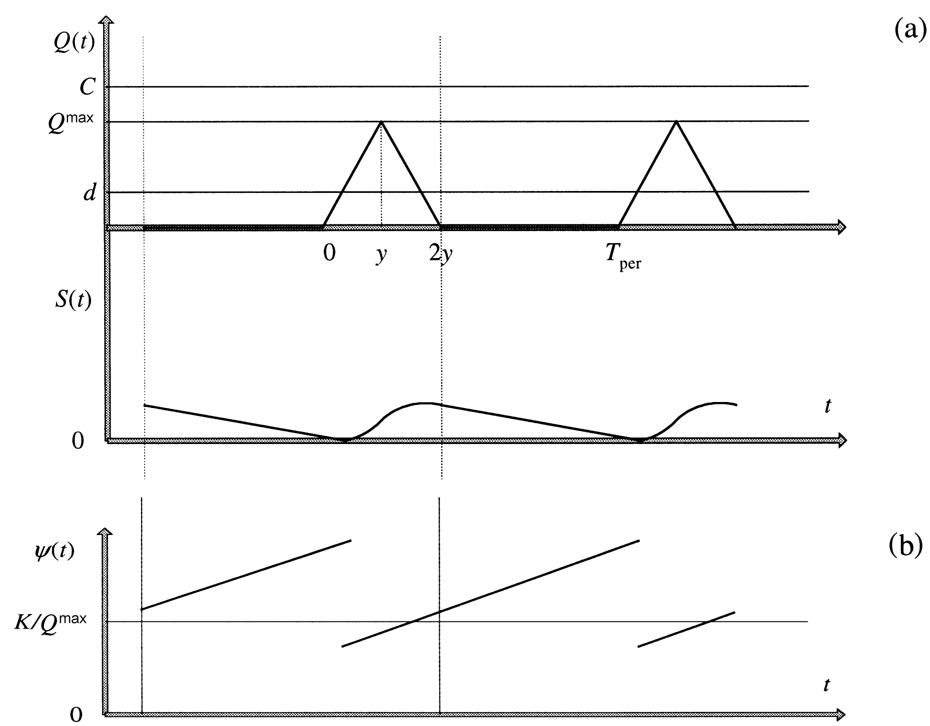

Fig. 3. A solution of cycle type 3: (a) state variables; (b) co-state variables.

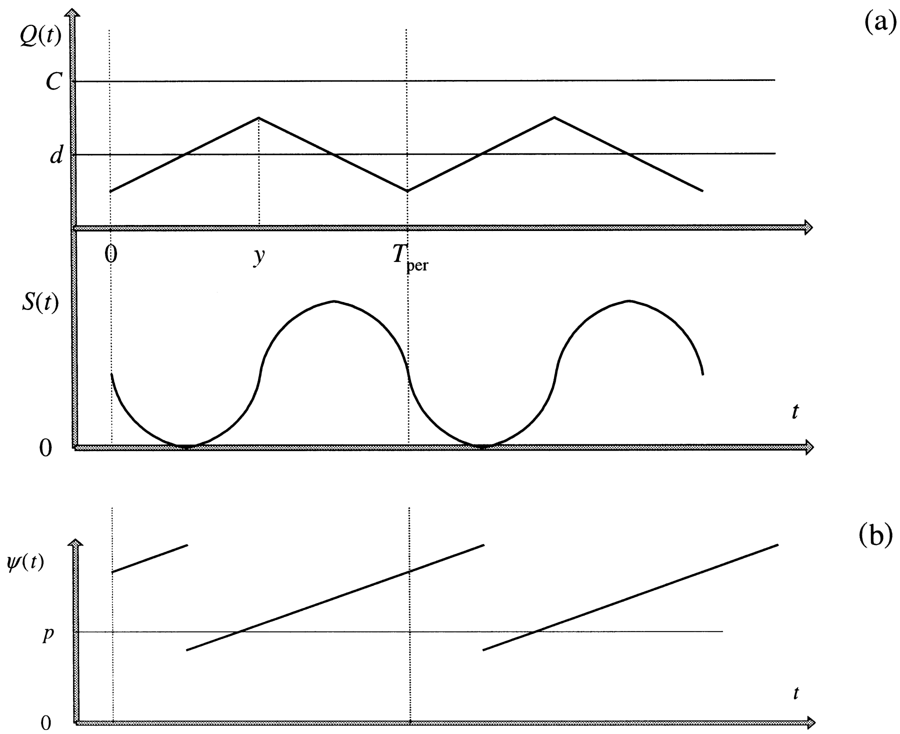

Fig. 4. A solution to cycle type 4: (a) state variables; (b) co-state variables.

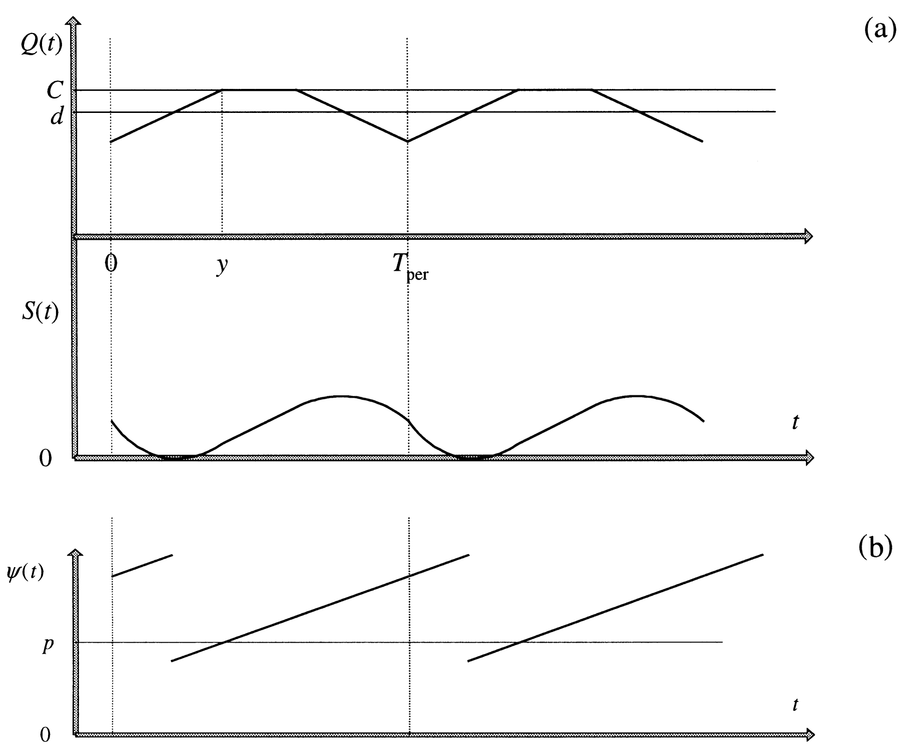

Fig. 5. A solution of cycle type 5: (a) state variables; (b) co-state variables.

Cycle type 4: $Q(t) = \begin{cases} d + M(t - y/2) & \text{if } t \in [0, y], \\ d - M\left(t - \frac{3y}{2}\right) & \text{if } t \in [y, T_{\text{per}}], \end{cases}$

where:

$$y = \frac{T_{\text{per}}}{2}.$$

The solution is cyclic with:

$$0 < T_{\text{per}} \le \frac{4}{M} \min\{d, C - d\}.$$

It exists for $0 < d < C$.

Cycle type 5:

$$Q(t) = \begin{cases} C + M(t - y) & \text{if } t \in [0, y], \\ C & \text{if } t \in [y, T_{\text{per}} - y], \\ C - M(t + y - T_{\text{per}}) & \text{if } t \in [T_{\text{per}} - y, T_{\text{per}}], \end{cases}$$

where:

$$y = \sqrt{\frac{(C - d)T_{\text{per}}}{M}}.$$

The solution is cyclic with:

$$\frac{4(C - d)}{M} \le T_{\text{per}} \le \frac{C^2}{M(C - d)}.$$

It exists for $d \ge C/2$.

Proof. In Appendix F.

To locate the best solution among the five considered in Theorem 2, we calculate the value of the objective function (8) for each cycle type. Denote by j the cost of a cyclic solution per time unit, i.e.:

$$j = \frac{1}{T_{\text{per}}} \int_0^{T_{\text{per}}} (hS(t) + \text{Cost}(Q(t)))\text{d}t \tag{12}$$

and by j_i the cost of a solution of cycle type i, $i = 1, \ldots, 5$. Then, analytical calculation of Equation (12) results in a closed-form expression for each of the cycle types:

$$j_1 = K;$$
$$j_2 = K\left(\frac{d}{C} + \frac{C}{MT_{\text{per}}}\right) + \frac{hd}{2}\left(T_{\text{per}} - \frac{C}{M}\right)\left(1 - \frac{d}{C}\right);$$
$$j_3 = 2K\sqrt{\frac{d}{MT_{\text{per}}}} + \frac{hd}{2M}(\sqrt{MT_{\text{per}}} - \sqrt{d}^2);$$
$$j_4 = K + \frac{hMT_{\text{per}}^2}{32};$$
$$j_5 = K + \frac{h(C - d)}{2M}(\sqrt{MT_{\text{per}}} - \sqrt{C - d})^2.$$

From these expressions, one can conclude that cycle types 4 and 5 can never be better than the others because j_4 and j_5 monotonously increase as functions of T_{per} within the scope of the solution existence.

The minimum of j_3 as a function of T_{per} is achieved for the root of the equation $(A + dMT_{\text{per}})^2 = d(MT_{\text{per}})^3$, where $A = 2KM/h$. The minimum of j_2 as a function of T_{per} is achieved for:

$$T_{\text{per}} = \frac{C}{M}\sqrt{\frac{A}{d(C - d)}}.$$

The minimum of j_1 is achieved for any $T_{\text{per}} > 0$.

By comparing the values of j_1, j_2 and j_3, we find that: Cycle type 3 is better than the others when:

$$0 < d \le \sqrt{A}\sqrt{\frac{2}{11 + 5\sqrt{5}}} \quad \text{and} \quad A \le C^2\left(\frac{C}{d} - 1\right).$$

Cycle type 2 is better than the others when:

$$A \ge C^2\left(\frac{C}{d} - 1\right) \quad \text{and} \quad A + Cd \le 2C\sqrt{\frac{Ad}{C - d}}.$$

Cycle type 1 is better than the others for all other combinations of parameters.

The partition of the (C,d) parameter plane for an example with $K = 20$ and $h = 1$ with respect to the optimal cycle types is presented in Fig. 1(b). The broken line in the figure represents the $C = d$ line which distinguishes between the feasible area ($C \ge d$) and the non-feasible area ($C < d$). The bold digits 1, 2 and 3 mark the regions of the (C, d)-plane where respective cycle types 1, 2 and 3 are optimal.

Theorem 3. *No other solution satisfies the optimality properties stated in Lemma 4.*

Proof. In Appendix G.

5. Discussion

The model presented in Section 4 fits into the continuous and fixed cost section of Table 1. We would also like to claim that it represents an analogy of the lot-sizing model with a fixed cost. However, this raises the question of how to determine if a particular continuous model is an appropriate analogy to its discrete counterpart (or *vice versa*). While determining an analogy between models is a broad and interesting topic that is beyond the scope of this paper, we nevertheless define minimal guidelines for such a determination. The guidelines we list below were motivated by our modeling experience of the model presented in Section 4. Note that they imply that an analogous model may not be unique, so there may be more than one way for an appropriate modeling. For this reason the guidelines are related mostly to the solution of the problem, rather than to the formulation of the problem. In addition, our paper and our guidelines do not consider the issue of solution "translation" from one type of model (e.g., continuous) to the other (e.g., discrete). We believe that here too there are many different ways in which this can be done, and we do not want our problem definition or analogy guidelines to depend on a particular mechanism. For this reason, our guidelines are often somewhat vague, however, they are useful in excluding formulations that may be considered as analogous, but do not satisfy those minimal requirements. While in defining those guidelines we had in mind the analogy for the (CDLSP) problem, we believe that they are also appropriate for other discrete-continuous problems.

Guidelines for a modeling analogy between discrete and continuous time problems:

1. *Feasibility*: The solution of each problem should be feasible to implement in the environment in which it is defined.
2. *Solution similarity*: The optimal solutions of both the discrete and the continuous problems should have similar properties.
3. *Cost similarity*: The optimal solutions of both the discrete and the continuous problems should incur close cost values and the maximum difference (over all possible parameter values) in the cost should either be finite or go to infinity at a rate of a smaller order than do the parameter values.

The feasibility requirement is easy to justify, and is motivated by the need to exclude situations in which the solution cannot be implemented, for example, a chattering solution.

Point 2 is stated in an informal way. However, we will demonstrate the similarity of the solutions to the discrete and continuous problems that we have considered. In general, this point is motivated by the necessity of having a mechanism that can "translate" the solution in one environment (in which a solution can be found more easily) to the other. Based on similar properties, such a mechanism is expected to ensure near-optimality of the translated solution.

In point 3, note that in most cases the cost values indeed cannot be identical, however, we expect them to be close, as defined above. In addition, it would be reasonable to determine that when two alternative modeling analogies are examined, the smaller the difference in the cost value between the discrete and the continuous solutions, the better the analogy between the models, given that the first two points of the guidelines hold.

Next we discuss the analogy between the infinite-horizon discrete-capacitated dynamic lot-sizing model (Section 3) and its new suggested continuous-time counterpart (Section 4). We examine the analogy according to the above guidelines and demonstrate that all three points of the guidelines are satisfied.

5.1. *Guideline 1: Feasibility*

Clearly the discrete-time solution is feasible in its environment. The solution of the new continuous-time problem is feasible to implement, since the production rate change is controlled by the user. In particular, the chattering regime is no longer feasible.

5.2. *Guideline 2: Solution similarity*

We demonstrate the similarity between the solutions of the above two problems with respect to several characteristics.

5.2.1. *Cycle types*

In both problems we found that only three cycle types may constitute the solution. Moreover, these cycle types are very similar in both problems. In cycle type 1 the production follows the demand exactly, so that the inventory level is always zero. In cycle type 2, the solution exploits the full production capacity in some periods while in other periods/times it reaches zero. The transition between full production and zero production in both problems is done as fast as possible; in the discrete problem this means that consecutive periods observe a change from zero to full capacity (or *vice versa*), whereas in the continuous problem the maximal change in the production rate is used for an interval of time of length C/M to switch from one extreme to the other. In cycle type 3 the full capacity level is not *regularly* used within the cycle. Rather, zero and intermediate production levels are sufficient in an optimal solution to track the demand. We note that with cycle type 3 the full production capacity may be reached at most once within the cycle, in the first period in the discrete problem, and in an isolated time point in the continuous problem. Cycle types 1 and 2 exist over the entire (C,d)-plane, while cycle type 3 does not exist when $d > C/2$, for both problems.

5.2.2. *Partition of the (C,d)-plane*

For specific C and d values, one of the three cycle types is better than the others. Therefore, the (C,d)-plane is partitioned into three regions in which one cycle type dominates the others. Figure 1(a and b) presents such a partition for both problems for the values of $K = 20$ and $h = 1$. The general location of the three cycle types in the (C,d)-plane is similar for both problems. For example, when d is close to C, cycle type 1 is optimal. Another example is the line distinguishing between cycle types 1 and 2. It looks similar in both problems, but the exact location is different.

The line distinguishing between cycle type 1 and cycle type 3 is:

$$d = \sqrt{\frac{2}{11 + 5\sqrt{5}}}\sqrt{\frac{2KM}{h}},$$

in the continuous case and $d = K/h$ in the discrete case. This is due to the additional time-related parameter that exists in the continuous problem, M. We can view this time-related parameter as a measure for the "closeness" of the discrete and continuous problems in the following way; the location of the line in the (C,d)-plane shapes the entire partition and establishes the scale of the plane [see Fig. 1(a and b)]. Therefore, to make the partition of the (C,d)-plane in the two problems as close as possible, the line must coincide for the two cases. This occurs if the parameter M is chosen as:

$$M = \frac{K}{h} \cdot \frac{(11 + 5\sqrt{5})}{4}.$$

In our opinion, for such M the partitioning of the (C,d)-plane according to the discrete and the continuous models is the closest, see also guideline 3 for additional support of this choice. Therefore, if, for example, the continuous-time problem is formulated for the purpose of approximating the discrete-time problem, we recommend that M be chosen in this way. In cases where the problem describes a specific environment, M should be chosen according to practical considerations.

5.3. *Guideline 3: Cost similarity*

The objective of the discrete model is to minimize the average cost per period (see the objective function of the formulation in Appendix A), while the objective of the continuous model is to minimize the average cost per time unit, see Equation (8). To correctly compare between these costs, we set the time unit used in the formulation of the continuous model to be equal to the period length used in the discrete model. Then, the following lemma proves the cost similarity of the two problems.

Lemma 5. *Let $j_d(K, h, C, d)$ be the optimal cost per period in the discrete-time problem and $j_c(K, h, C, d)$ be the optimal cost per time unit in the continuous-time problem for a given value of the parameter M. Then, $|j_c(K, h, C, d) - j_d(K, h, C, d)| \le o(K)$.*

Proof. Consider the case when $K \to \infty$. For a specific point on the (C,d)-plane in the continuous-time problem (see Fig. 1(b)), there exists a sufficiently large K such that the point belongs to the second cycle type area. The cycle length:

$$T_{\text{per}} = \sqrt{K}\sqrt{\frac{2C^2}{d(C-d)Mh}},$$

goes to infinity and

$$j_c(K, h, C, d) = j_2 = \frac{d}{C}K + \sqrt{2d(C-d)}\sqrt{\frac{Kh}{M}} - \frac{hd(C-d)}{2M}. \quad (13)$$

When $K \to \infty$ in the discrete-time problem, there again exists a sufficiently large K such that a point on the (C,d)-plane is within the area of the second cycle type (see Fig. 1(a)). The number of full production periods within the cycle $c(n) \to (nd/C) - 1$. Now, from Equation (2) it follows that for large K:

$$\frac{d}{C}K + \frac{1}{2}h(C-d) < j_d(K, h, C, d) < \frac{d}{C}K + \frac{1}{2}h(C+d). \quad (14)$$

The lemma immediately follows from Equations (13) and (14) for the considered case.

The case of $h \to 0$ is very similar to that considered above. When $K \to 0$ and/or $h \to \infty$, the optimal cycle type is one in both continuous- and discrete-time problems. As a result, $|j_c(K, h, C, d) - j_d(K, h, C, d)| = 0$. When $C \to \infty$, either cycle type 3 or cycle type 1 is optimal. In the former case the value of C is not relevant, i.e., the difference in the lemma remains constant. In the latter case, the difference in the lemma is zero.

The case when $C \to 0$ (therefore $d \to 0$ as well) is equivalent to the case when $h \to 0$. This can be shown by rescaling the part units. Finally, in both cases when $d \to 0$ and $d \to C$ the difference in the lemma goes to zero. ■

The following corollary additionally motivates the choice of the parameter M as discussed in guideline 2 above. The corollary strengthens the previous lemma and claims that under such a choice the difference between $j_c(K, h, C, d)$ and $j_d(K, h, C, d)$ is limited throughout the entire space of the parameters (K, h, C, d).

Corollary 3. *Let, in the conditions of Lemma 5, the parameter M be chosen as $M = Kq/h$, where q is an arbitrary positive constant, then for each q, $|j_c(K, h, C, d) - j_d(K, h, C, d)| \le O(1)$.*

Proof. The proof of Corollary 3 is similar to that of Lemma 5. ■

Figure 6 shows the objective values obtained for both problems when using the optimal policy for $K = 20$, $h = 1$, $C = 10$ and

$$M = \frac{K}{h} \cdot \frac{(11 + 5\sqrt{5})}{4} = 110.9.$$

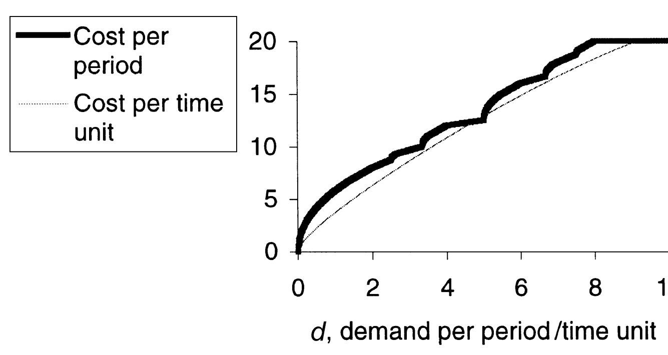

Fig. 6. The optimal cost per period (for the discrete-time problem) and the optimal cost per time unit (for the continuous-time problem) as a function of demand for $C = 10$, $K = 20$ and $h = 1$.

The graph shows the optimal per period/time unit objective value as a function of the demand per period/time unit, which varies between zero and C. In general, the values are similar, however, the objective value of the discrete problem is almost always higher than that of the continuous problem. We believe that this is due to the ability of the continuous solution to change at any point in time, as opposed to the discrete solution, which can change only at the end of a period. On the other hand, the inertia of the production rate in the continuous-time case can result in a higher cost than in the similar discrete-time case (see, for example, point $d = 5.1$ in Fig. 6).

Thus, we have shown that the three guidelines for the analogy, hold. Therefore, the new proposed continuous-time problem can be considered as an appropriate counterpart for the capacitated lot-sizing model. Future research is needed to better understand this parallelism for other versions of production planning models, but we believe that the foundations for that have been established in this paper.

Acknowledgement

The authors wish to thank the participants of the Research Forum in the Department of Industrial Engineering at Tel Aviv University for their comments on an earlier draft of this paper.

References

Baker, K.R., Dixon, P., Magazine, M.J. and Silver, E.A. (1978) An algorithm for the dynamic lot-size problem with time-varying production capacity constraints. *Management Science*, **24**, 1710–1720.

Bensoussan, A., Crouhy, M. and Proth, J.M. (1983) *Mathematical Theory of Production Planning*, North Holland, Amsterdam, The Netherlands.

Bitran, G.B. and Yanasse, H.H. (1982) Computational complexity of the capacitated lot size problem. *Management Science*, **28**, 1174–1185.

Dauzere-Peres, S., Gershwin, S.B. and Sevaux, M. (2000) Models and solving procedures for continuous-time production planning, *IIE Transactions*, **32**, 93–103.

Denardo, E.V. (1982) *Dynamic Programming: Models and Applications*, Prentice-Hall, Englewood, Cliffs, NJ.

Florian, M. and Klein, M. (1971) Deterministic production planning with concave costs and capacity constraints. *Management Science*, **18**, 12–20.

Florian, M., Lenstra, J.K. and Rinnooy Kan, A.H.G. (1980) Deterministic production planning: algorithms and complexity. *Management Science*, **26**, 669–679.

Hartl, R.F., Sethi, S.P. and Vickson R.G. (1995) A survey of the maximum principles for optimal control problems with state constraints. *SIAM Review*, **37**, 181–218.

Hwang, C.L., Fan, L.T. and Erickson, L.I. (1967) Optimal production planning by the maximum principle. *Management Science*, **13**, 751–755.

Maimon, O., Khmelnitsky, E. and Kogan, K. (1998) *Optimal Flow Control in Manufacturing Systems: Production Planning and Scheduling*, Kluwer, Boston, MA.

Sethi, S.P. and Zhang, Q. (1995) *Hierarchical Decision Making in Stochastic Manufacturing Systems*, Birkhauser, Boston, MA.

Silver, E.A. (1990) Deliberately slowing down output in a family production context. *International Journal of Production Research*, **28**, 17–27.

Steindl, A. (2001) Fast oscillations in a singularly perturbed production model, *Mathematical Models and Methods in Applied Sciences*, **11**(1), 149–162.

Van Hoesel, S. and Wagelmans, A. (1996) An $O(T^3)$ algorithm for the economic lot-sizing problem with constant capacities. *Management Science*, **42**, 142–150.

Wagner, H.M. and Whitin, T.M. (1958) Dynamic version of the economic lot size model. *Management Science*, **5**, 89–96.

Wolsey, L.A. (1989) Uncapacitated lot-sizing problems with start-up costs. *Operations Research*, **37**(5), 741–747.

Appendices

Appendix A: MILP formulation of (F-CDLSP)

Let Q_t = production quantity in period t;

$$Y_t = \begin{cases} 1 & \text{if } Q_t > 0, \\ 0 & \text{otherwise;} \end{cases}$$

S_t = inventory level at the end of period t;

$$\text{Min} \frac{1}{T} \sum_{t=1}^{T} (pQ_t + hS_t + KY_t),$$

subject to

$$\begin{aligned} & S_0 = 0, \\ & S_t = S_{t-1} + Q_t - d && t = 1, \ldots, T, \\ & Q_t \le CY_t && t = 1, \ldots, T, \\ & Q_t \ge 0, S_t \ge 0 && t = 1, \ldots, T, \\ & Y_t = 0, 1 && t = 1, \ldots, T. \end{aligned}$$

Appendix B: Proof of Theorem 1

We first prove that there exists an optimal solution which consists of subplans of bounded lengths. This is accomplished by showing first that there exists a finite period with a zero ending inventory. To that end, assume by contradiction that in all periods the ending inventory is strictly positive. Note that when the starting inventory in a given period is larger than or equal to d, no production occurs in that period, as postponing it will result in inventory cost savings. Note also that if C and d are rational, then $\bar{n}$ is finite. We consider the following cases:

Case 1: All periods in which production occurs are full production periods. In this case, after $\bar{n}$ periods in which $\bar{n}d/C$ (which is an integer, by assumption) full production periods take place, the ending inventory is zero, a contradiction.

Case 2: There exists one fractional production period. In this case, consider the first $\bar{n}d/C$ consecutive periods in which full production occurs, after the fractional production period. Those $\bar{n}d/C$ full production periods cover the demand of exactly $\bar{n}$ periods, therefore before and after their occurrence there exists an equal amount of positive inventory. As a result, and since all other parameters are stationary, the next $\bar{n}$ periods will observe the same production and inventory levels (otherwise one of these schedules can be improved), that is, the solution becomes cyclic after the fractional production has occurred. In that case, reducing the size of the fractional production period by ε results in a cost saving of $h\varepsilon$ for every period, starting with the fractional period. This can be repeated until the size of the fractional production period reaches zero (which brings us back to case 1), or until one of the periods observes a zero ending inventory level.

Case 3: There exist at least two fractional production periods. Consider the first and second fractional production periods, denoted by f_1 and f_2, respectively (and recall that all periods in between have a positive ending inventory). Then, consider an alternative schedule in which the production quantity in period f_1 is reduced by ε, and the production quantity in period f_2 is increased by ε. As a result, in all periods between f_1 and f_2 a cost saving of $h\varepsilon$ occurs, thus the new schedule is better than the previous one. This can be repeated until either the size of f_1 drops to zero, or the size of f_2 reaches its capacity, or until the ending inventory of one of the periods in between reaches zero. In the first two cases, there is one less fractional production period and the process may be repeated with the next fractional production period (if it exists). In the third case, a period with a zero ending inventory has been reached.

Second, since a period with a zero ending inventory has been reached, and since all parameters are stationary, we are at exactly the same situation as at the beginning of the horizon. Therefore, the same solution repeats itself.

To prove that the length of a subplan is bounded by $\bar{n}$, we now consider one subplan. Note that the total number of units produced in a subplan of length n is nd, therefore when nd/C is an integer, no fractional period exists and the last period in the cycle has a zero ending inventory. If n is chosen to be larger than $\bar{n}$, then at the end of $n-\bar{n}$ periods, the ending inventory will be zero, since the amount produced after that (through full production periods only) exactly covers the demand in the remaining $\bar{n}$ periods. ■

Appendix C: Proof of Lemma 3

According to Equation (8):

$$\left\lceil \frac{c(n)C}{d} \right\rceil + 1 \le n \le \left\lfloor \frac{(c(n)+1)\,C}{d} \right\rfloor.$$

Next, note that the total production during the full production periods covers the demand of:

$$\left\lceil \frac{c(n)C}{d} \right\rceil,$$

periods (one of them may be covered only partially), and therefore given n and $c(n)$ that satisfy Equation (8), the number of periods between the first period (a fractional production period) and the first full production period is:

$$n - \left\lceil \frac{c(n)C}{d} \right\rceil.$$

At this time interval the inventory level S_t decreases by d units each period, starting at $S = f(n) - d = (n-1)d - c(n)C$. Therefore, the number of units held in inventory during this time interval forms an algebraic series, and the total inventory cost associated with this interval is the unit holding cost h, multiplied by the sum of this series which, after some algebra, can be shown to be equal to:

$$\begin{aligned}
&\frac{h}{2}\left(d\left(n + \left\lceil \frac{c(n)C}{d} \right\rceil - 1\right) - 2Cc(n)\right)\left(n - \left\lceil \frac{c(n)C}{d} \right\rceil\right) \\
&\quad = \frac{h}{2}\left[d\left(n(n-1) - \left\lceil \frac{c(n)C}{d} \right\rceil\left(\left\lceil \frac{c(n)C}{d} \right\rceil - 1\right)\right)\right. \\
&\qquad \left. - 2nc(n)C + 2c(n)C\left\lceil \frac{c(n)C}{d} \right\rceil\right].
\end{aligned}$$

Similarly, the number of periods between the first and the second full production periods is:

$$\left\lceil \frac{c(n)C}{d} \right\rceil - \left\lceil \frac{(c(n)-1)C}{d} \right\rceil.$$

Again, at this interval the inventory level decreases by d units each period, starting at:

$$\begin{aligned}
S &= f(n) + C - d\left(n - \left\lceil \frac{c(n)C}{d} \right\rceil + 1\right) \\
&= d\left(\left\lceil \frac{c(n)C}{d} \right\rceil - 1\right) - (c(n)-1)C,
\end{aligned}$$

therefore the total inventory during this interval is the sum of an algebraic series, and the inventory cost associated with this interval is:

$$\begin{aligned}
&\frac{h}{2}\left(d\left(\left\lceil \frac{c(n)C}{d} \right\rceil + \left\lceil \frac{(c(n)-1)C}{d} \right\rceil - 1\right) - 2C(c(n)-1)\right) \\
&\quad \times\left(\left\lceil \frac{c(n)C}{d} \right\rceil - \left\lceil \frac{(c(n)-1)C}{d} \right\rceil\right) \\
&= \frac{h}{2}\left[d\left(\left\lceil \frac{c(n)C}{d} \right\rceil\left(\left\lceil \frac{c(n)C}{d} \right\rceil - 1\right) - \left\lceil \frac{(c(n)-1)C}{d} \right\rceil\right.\right. \\
&\quad \left.\left. \times\left(\left\lceil \frac{(c(n)-1)C}{d} \right\rceil - 1\right)\right)\right] \\
&= \frac{h}{2}\left[-2(c(n)-1)C\left\lceil \frac{c(n)C}{d} \right\rceil + \left\lceil \frac{(c(n)-1)C}{d} \right\rceil\right].
\end{aligned}$$

Subsequent intervals follow similar calculations. Finally, the number of periods between the last full production period and the end of the cycle is:

$$\left\lceil \frac{C}{d} \right\rceil.$$

The inventory in the first period of this interval is:

$$d\left(\left\lceil \frac{C}{d} \right\rceil - 1\right),$$

and it decreases by d units each period until it reaches zero in the last period. Therefore, the inventory cost associated with this interval is:

$$\frac{hd}{2}\left\lceil \frac{C}{d} \right\rceil\left(\left\lceil \frac{C}{d} \right\rceil - 1\right).$$

Summing up the inventory cost over the cycle, we obtain (after some algebra):

$$\frac{1}{2}hd(n-1) - Chc(n) + \frac{Ch}{n}\sum_{i=1}^{c(n)}\left\lceil \frac{iC}{d} \right\rceil,$$

from which the lemma immediately follows. ∎

Appendix D: Example for case 1 of the (C,d)-plane partition of the discrete problem

The following examples illustrate the conjecture of case 1 by considering the points (56,22) and (56,18) of the (C,d)-plane, which belong to regions 1 and 3 respectively (see Fig. 1(a)). Tables A1 and A2 show that the costs $j(n)$ of the solutions of cycle type 2 are significantly greater and therefore do not influence the competition between the solutions of cycle types 1 and 3. The comparison is presented for small cycles, $n \le 10$; for the larger cycle lengths, all of which belong to cycle type 2, $j(n)$ is even greater.

In this example, and generally in case 1, where the C value is relatively large, solutions of type 2 cannot be optimal since this would require producing in large lots of $Q(t) = C$ in some periods t, thereby increasing the inventory holding costs. With a large C and a small d a better option is to produce in lots smaller than C (EOQ-like solutions belonging to cycle type 3, finding the best trade-off between the inventory holding and the fixed costs). With a large C and a large d a better option is to exactly track the demand, having zero inventory in each period (solution type 1). Thus, we conjectured that solutions of type 2 have higher costs in this area of the (C,d)-plane and therefore they do not influence the partition, as in the previous examples.

Table A1. Cost as a function of the cycle type for $C = 56$ and $d = 22$

$c(n)$	n	*Cycle type*	$j(n)$
0	1	1	20.0
	2	3	21.0
1	4	2	29.0
	5	2	29.6
2	7	2	34.57
3	9	2	34.67
	10	2	34.2

Table A2. Cost as a function of the cycle type for $C = 56$ and $d = 18$

$c(n)$	n	*Cycle type*	$j(n)$
0	1	1	20.0
	2	3	18.0
	3	3	24.67
1	5	2	32.8
	6	2	33.0
2	8	2	35.5
	9	2	35.11

Appendix E: Proof of Lemma 4

To characterize the optimal solution of problem (5)–(10), we apply the maximum principle that declares that:

- The optimal control $v(t)$ maximizes the Hamiltonian function at each point of time, i.e.:

$$v(t) = \arg\max_{0 \le |v| \le M} H(S(t), Q(t), \psi(t), \psi^Q(t), v, t). \quad \text{(A1)}$$

- The Hamiltonian function is:

$$H(t) = -hS(t) - \text{Cost}(Q(t)) + \psi(t)(Q(t) - d) + \psi^Q(t)v(t). \quad \text{(A2)}$$

- The co-state variables $\psi(t)$ and $\psi^Q(t)$ are left-continuous function satisfying the co-state equations:

$$d\psi(t) = h\,dt - d\mu(t), \quad \text{(A3)}$$

$$d\psi^Q(t) = -\psi(t)dt + \text{Cost}'_Q(Q(t))dt - d\mu_1^Q(t) + d\mu_2^Q(t),\ \psi^Q(0) = \psi^Q(T+0) = 0. \quad \text{(A4)}$$

- Lagrange multipliers of the state constraints, Equations (6) and (7), $d\mu(t)$, $d\mu_1^Q(t)$ and $d\mu_2^Q(t)$, are measure functions satisfying the non-negativity and complementary slackness conditions:

$$d\mu(t) \ge 0, \quad d\mu_1^Q(t) \ge 0, \quad d\mu_2^Q(t) \ge 0, \quad \text{(A5)}$$

$$\int_0^\infty S(t)d\mu(t) = 0, \quad \int_0^\infty Q(t)d\mu_1^Q(t) = 0, \quad \int_0^\infty (C - Q(t))\,d\mu_2^Q(t) = 0. \quad \text{(A6)}$$

Maximization of the Hamiltonian as a function of $v(t)$ results in:

$$v(t) = \begin{cases} M, & \text{if } \psi^Q(t) > 0, \\ -M, & \text{if } \psi^Q(t) < 0, \\ \alpha, \alpha \in [-M, M], & \text{if } \psi^Q(t) = 0. \end{cases} \quad \text{(A7)}$$

The CO regime immediately follows from the first two lines of Equation (A7). The other three regimes stated in the lemma follow from the third line of Equation (A7). Indeed, let $\psi^Q(t) = 0$ at an interval of time $Y \subseteq [0, T_c]$. Consider four possible cases for the Lagrange multipliers.

1. $d\mu(t) > 0, d\mu_1^Q(t) = 0, d\mu_2^Q(t) = 0$ at Y. From the complementary slackness conditions (A6), co-state equations (A3) and (A4) and state equations (5) and (9), it follows that at Y:

$$S(t) = 0, \quad Q(t) = d, \quad v(t) = 0,$$
$$\psi(t) = \mathrm{Cost}'_Q(Q(t)) = p, \quad d\mu(t) = h\mathrm{d}t.$$

   This is the AP regime. It can potentially occur everywhere. Note that AP is the unique regime for which $d\mu(t) > 0$. Indeed, assuming that either $d\mu_1^Q(t) > 0$ or $d\mu_2^Q(t) > 0$, we obtain particular cases of the AP regime. The strict positiveness of both $d\mu_1^Q(t)$ and $d\mu_2^Q(t)$ is impossible because $Q(t)$ cannot be simultaneously equal to zero and to C (see Equation (A6)).

2. $d\mu(t) = 0, d\mu_1^Q(t) = 0, d\mu_2^Q(t) = 0$ at Y. From the complementary slackness conditions (A6), co-state equations (A3) and (A4) and state equations (5) and (9), it follows that at Y:

$$\psi(t) = p \quad \text{and} \quad \dot{\psi}(t) = h.$$

   These contradict each other.

3. $d\mu(t) = 0, d\mu_1^Q(t) = 0, d\mu_2^Q(t) > 0$ at Y. From the complementary slackness conditions (A6), co-state equations (A3) and (A4) and state equations (5) and (9), it follows that at Y:

$$Q(t) = C, \quad v(t) = 0.$$

   This is the FP regime. The necessary conditions for the regime to occur, which follow from Equations (A4) and (A5), are:

$$\psi(t) \geq p.$$

4. $d\mu(t) = 0, d\mu_1^Q(t) > 0, d\mu_2^Q(t) = 0$ at Y. From the complementary slackness conditions (A6), co-state equations (A3) and (A4) and state equations (5) and (9), it follows that at Y:

$$Q(t) = 0, \quad v(t) = 0.$$

   This is the NP regime.

Since we have enumerated all possible cases of the Lagrange multipliers, no other production regime can occur along the optimal trajectory. That has proved the first two properties stated in the lemma. To prove the third property, we first note that if a CO regime occurs on the interval $[t, t+\tau] \subseteq [0, \infty)$, then $d\mu_1^Q(t') = d\mu_2^Q(t') = 0$ for $t' \in [t, t+\tau]$. This follows from the complementary slackness Equation (A6). Next, at the beginning and at the end of the CO interval, $\psi^Q(t)$ is zero. Thus, by integrating the co-state equation (A4) on this interval, we obtain:

$$\begin{aligned} \frac{1}{\tau}\int_t^{t+\tau} \psi(t')\mathrm{d}t' &= \frac{1}{\tau}\int_t^{t+\tau} \mathrm{Cost}'_Q(Q(t'))\mathrm{d}t' \\ &= \frac{1}{\tau}\int_t^{t+\tau} \frac{\mathrm{Cost}'_Q(Q(t'))}{\dot{Q}(t')}\mathrm{d}Q(t') \\ \frac{|\mathrm{Cost}(Q_2) - \mathrm{Cost}(Q_1)|}{\tau M} &= \frac{\mathrm{Cost}(Q_2) - \mathrm{Cost}(Q_1)}{Q_2 - Q_1}. \end{aligned} \tag{A8}$$

■

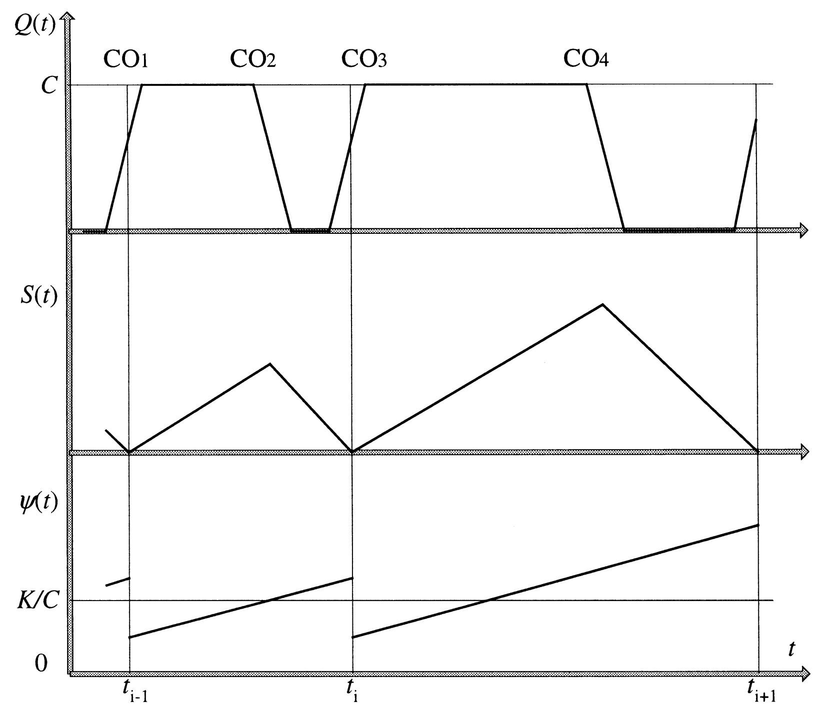

Fig. A1. Non-periodic solution.

Appendix F: Proof of Theorem 2

First we note that the solutions presented in the theorem contain only the four regimes defined in Lemma 4, i.e., optimality property 1 holds. Property 2 stated in the same lemma is meaningful only for cycle type 1, for which it is trivially satisfied. Therefore, to prove the theorem, it remains to show that optimality property 3 holds for each cycle type, i.e., that there exists a co-state variable $\psi(t)$ that satisfies Equations (A3) and (A8). For cycle type 1, such a co-state variable is simply $\psi(t) \equiv 0$ for $t \in (0, \infty)$. The co-state variable for cycle types 2–5 is depicted in Figs. 2(b)–5(b). ■

Appendix G: Proof of Theorem 3

Let Z denote the set of time points at which $S(t) = 0$. Consider the following three cases.

1. $Z = \cup\{t_i\}$, $i = 1, \ldots, I$, $I \gg 0$, i.e., Z is a set of a finite number of isolated points. Then, $S(t) > 0$ for $t > t_I$ and $\psi(t)$ increases with rate h always after t_I. Therefore, there exists a point $\hat{t} > t_I$ after which no change of regimes occurs. If the last regime is:
   - NP or CO, then the trajectory is not feasible;
   - FP or AP, then the trajectory can be easily improved.
   These contradict the optimality of the solution.
2. $Z = \cup\{t_i\}$ is a cyclic set of an infinite number of isolated points, i.e., for all i, $t_{i+1} - t_i = t_i - t_{i-1}$. These are solutions of cycle types 2, 3, 4 and 5.
3. $Z = \cup\{t_i\}$ is a non-cyclic set of an infinite number of isolated points, i.e., for some i, $t_{i+1} - t_i \neq t_i - t_{i-1}$ (see Fig. A1). If condition (A8) holds for CO_1, CO_2 and CO_3, then it cannot hold for CO_4.
4. $Z = (0, \infty)$, then $Q(t) \equiv d$ which coincides with cycle type 1.
5. Z contains a finite interval. Then, after this interval, the trajectory of the solution is obtained unambiguously which is dictated by the state and co-state dynamics. This trajectory will finally come to a contradiction with the optimality condition, Equation (A8), unless the parameters of the problem are chosen such that $j_1 = j_3$ or $j_1 = j_2$. ■

Biographies

Eugene Khmelnitsky received a M.S. degree in Applied Mathematics from the Moscow Institute of Physics and Technology and a Ph.D. degree in Industrial Engineering from Tel-Aviv University. He is a Senior Lecturer in the Department of Industrial Engineering, Tel-Aviv University, Israel. His research interests are in the areas of control and optimization of complex dynamic systems, production planning and scheduling. He is the author of the book *Optimal Flow Control in Manufacturing Systems: Production Planning and Scheduling* published by Kluwer in 1998.

Michal Tzur is a Senior Lecturer in the Department of Industrial Engineering at Tel-Aviv University, Israel, and is currently a visiting Associate Professor at the Transportation Center and the IEMS Department at Northwestern University. She joined Tel-Aviv University in 1994 after spending 3 years at the Operations and Information Management Department at the Wharton School, University of Pennsylvania. She received her B.A. from Tel-Aviv University and her M. Phil and Ph.D. from Columbia University. Her research interests are in the areas of supply chain management, multi-echelon inventory management, production planning and operations scheduling.

Contributed by the Supply Chains/Production-Inventory Systems Department