

EXACT COMPUTATION OF OPTIMAL INVENTORY POLICIES OVER AN UNBOUNDED HORIZON*

REFAEL HASSIN AND NIMROD MEGIDDO

An inventory scheduling model with forbidden time intervals is analyzed. The objective is to minimize the long-term average cost per time unit. Unlike most of the literature on inventory theory, no restrictive assumptions are made a priori about the nature of optimal solutions. Rather it is proved that optimal policies exist, and that some of them are cyclic with cycles of a particular structure. It is then shown that such optimal policies can be computed and an algorithm is given.

1. Introduction. In the classical Economic Order Quantity (EOQ) model one seeks a policy of inventory scheduling so as to minimize the long-term average cost per time unit. Inventory is depleted at a known constant rate R and can be replenished instantaneously at any time. There is a fixed cost of C_f per order, and the inventory holding cost is proportional to the amount of time and the amount of commodity held in stock: C_h per time unit for each unit of commodity. Given that an order is placed at time t , it is straightforward to find the time of the next order $t + x$ so as to minimize the average cost per time unit over $[t, t + x)$:

$$\frac{C_f}{x} + \frac{1}{2}C_hRx.$$

This turns out to be $x^* = \sqrt{2C_f/(C_hR)}$. A nice feature of this model is that the policy of ordering every x^* time units is optimal in the following strong sense: it minimizes the *limsup* (as T tends to infinity) of the average cost per unit over the interval $[0, T)$ among all possible schedules over an unbounded horizon.

There is a vast literature on inventory and production models generalizing the EOQ in numerous directions. Usually the models are stated precisely but the policies found are only approximately optimal or optimal only within a restricted class of policies which is more amenable for analysis (see, for example, [2], [3], [4], [5], [6], [8]). For example, many authors simply assume that there exists a cyclic optimal policy and set out to find a best cyclic one. Another common assumption is that orders are placed (or production is resumed) only when the inventory level reaches zero. In many cases such an assumption turns out to be wrong. There have been cases where authors made restrictive assumptions that were hard to justify and several mathematical pitfalls have been pointed out [1], [7], [9].

When we started the work presented in this paper our goal was to carry out an exact analysis of a model more general than the classical EOQ, where it is not a priori clear whether optimal policies exist, or whether optimality of a given policy can be decided. We were quite surprised to discover the complexities involved in the simple model we were looking at. It turned out, however, to be decidable in a nontrivial way.

*Received November 2, 1988; revised November 9, 1989.

AMS 1980 subject classification. Primary: 90B05. Secondary: 90A16.

LAOR 1973 subject classification. Main: Inventory. Cross references: optimization.

OR/MS Index 1978 subject classification. Primary: 346 Inventory/production/policies.

Key words. Optimal inventory policies, infinite horizon, deterministic.

We consider a deterministic inventory problem over an unbounded horizon. The objective is to minimize cost. More precisely, setup costs and inventory holding costs are averaged over time and a policy is evaluated by either the liminf or the limsup (as T tends to infinity) of the average total cost per time unit incurred during the first T time units. We actually show that there exist policies where this average tends to a limit which is equal to the infimum of the liminfs over all policies and hence also to the infimum of the limsup.

A precise statement of the problem is as follows.

PROBLEM 1.1. Given a real number α , $0 < \alpha < 1$, a commodity can be ordered at any time except during intervals of the form $(I, I + \alpha)$, where I is an integer. There is a fixed positive cost associated with every order and there is an inventory holding cost proportional to the amount of time and the amount of inventory. Inventory is depleted at a constant rate, which is called the demand rate. The order is filled instantaneously. A feasible solution is a policy consisting of feasible times of orders and amounts such that the demand is always satisfied. A policy Ξ is called *optimal* if the limsup of the average cost per time over the interval $[0, T)$, denoted $\bar{c}(\Xi)$, is minimum among all policies.

2. Preliminaries. The existence of feasible policies is obvious but we can actually prove:

PROPOSITION 2.1. *There exist optimal policies.*

PROOF. Suppose c^* is the infimum of $\bar{c}(\Xi)$ taken over all the feasible policies Ξ . Thus, for any $k > 0$ there exists a policy Ξ^k such that for every sufficiently large T , the average cost per time unit of Ξ^k over the interval $[0, T)$ is less than $c^* + 1/k$. The idea of the proof is to generate a policy Ξ which uses portions of the policies Ξ^k successively, for $k = 1, 2, \dots$. We construct below a sequence of positive integers $\{\tau_k\}$ which determines an optimal policy Ξ as follows. Over the interval $[0, \tau_1)$, the policy Ξ works essentially like Ξ^1 over the same interval, except that if the inventory level at τ_1 is positive, then the sizes of one or more orders before τ_1 are reduced to obtain a new feasible solution where the inventory level at τ_1 is zero. Denote $\sigma_j = \sum_{i=1}^j \tau_i$. Over the interval $[\sigma_{k-1}, \sigma_k)$, Ξ is essentially the same as Ξ^k during the interval $[0, \tau_k)$. Without loss of generality, assume that for every k the average cost of Ξ^k over the interval $[0, T)$ is bounded over all values of T . Now, there exists τ_1 such that for any $T \geq \tau_1$, the average cost over $[0, T)$ of running Ξ^1 over $[0, \tau_1)$, and then the first $T - \tau_1$ time units of Ξ^2 over $[\tau_1, T)$, is less than $c^* + 2$. Inductively, suppose we have determined $\tau_1, \dots, \tau_{k-1}$, and the policy Ξ has already been specified up to σ_{k-1} (and then Ξ^k is started) as explained above so that the following is true:

(i) For every $1 \leq j \leq k - 2$ and for any T , $\sigma_j \leq T < \sigma_{j+1}$, the average cost over $[0, T)$ is less than $c^* + 2/j$.

(ii) For any $T \geq \sigma_{k-1}$, the average cost over $[0, T)$ is less than $c^* + 2/(k - 1)$.

Since the average cost of Ξ^k over $[0, T)$ is less than $c^* + 1/k$ for any sufficiently large T , it follows that there exists τ_k such that if we start Ξ^{k+1} at $\sigma_k = \sigma_{k-1} + \tau_k$, then for every $T \geq \sigma_k$, the average cost over $[0, T)$ is less than $c^* + 2/k$. The complete construction gives a policy where the average cost over $[0, T)$, for $\sigma_j \leq T < \sigma_{j+1}$, is less than $c^* + 2/j$, and hence the limsup of the average cost is c^* . ■

Knowing that an optimal policy exists, the next question is how to find it. It is very useful to know whether a cyclic optimal policy exists, i.e., a policy which consists of repetitions of the same behavior over some finite interval. We note that in some staged decision problems there exist optimal policies but no cyclic optimal ones, as in the following example.

EXAMPLE 2.2. Suppose we have to choose at times $i = 0, 1, \dots$ numbers x_i , $0 \leq x_i \leq 1$. If $x_i < 1$, the payoff at time i is x_i . If $x_i = 1$, the payoff is zero. An optimal policy is to choose a sequence x_0, x_1, \dots that tends to 1. However, no cyclic sequence achieves 1 as the long-term average.

3. Existence of cyclic optimal policies. To simplify the presentation, we scale the data so that the following are true: (i) the demand rate is of one unit per time unit, (ii) the unit holding cost is one per time unit, and (iii) the fixed order cost is K per order.

Denote by $\{x\} = x - \lfloor x \rfloor$ the fractional part of x . A feasible policy Ξ involves an infinite sequence t_0, t_1, \dots , such that if t_i is not integral then $\{t_i\} \geq \alpha$. Without ambiguity, denote both the interval $[t_{i-1}, t_i)$ (the i th cycle) and its length by Δ_i .

REMARK 3.1. In the following propositions we prove the existence of optimal policies with certain properties. These propositions justify our focusing on optimal policies with these properties. In particular, subsequent to any of these propositions, we consider only policies with the properties proven so far.

PROPOSITION 3.2. *There exists an optimal policy where orders are placed only at times where the inventory level reaches zero.*

PROOF. Consider any optimal policy Ξ where orders of sizes Q_0^0, Q_1^0, \dots are placed at times $t_0 = 0, t_1, t_2, \dots$, respectively. Suppose some of the orders are placed when the inventory level is positive. Denote $\Xi^0 = \Xi$. Inductively, assume we have defined the policy Ξ^j where orders of sizes Q_0^j, Q_1^j, \dots are placed at the same times t_0, t_1, \dots , respectively, and suppose Ξ^j still has some orders placed when the inventory level is positive. We now define a policy Ξ^{j+1} as follows. Orders of Ξ^{j+1} are yet placed at the same times t_0, t_1, \dots . Let t_i be the first time in Ξ^j where an order is placed such that the inventory level X is positive, so $i \geq 1$. Define $Q_{i-1}^{j+1} = Q_{i-1}^j - X$ and $Q_i^{j+1} = Q_i^j + X$, i.e., reduce the amount ordered at t_{i-1} by X , and increase the amount ordered at t_i by X . Otherwise, the quantities of Ξ^{j+1} are the same as in Ξ^j . It follows by induction that Ξ^{j+1} is optimal and at t_i the inventory level is zero. This construction yields an infinite sequence of optimal policies where in each policy orders are placed at the same sequence of times t_0, t_1, \dots . For every t_i , consider the sequence of quantities Q_i^0, Q_i^1, \dots ordered at t_i in the policies Ξ^0, Ξ^1, \dots , respectively. Obviously, the sequence $\{Q_i^j\}_{j=0}^\infty$ increases at most once and decreases at most once. Thus, it eventually becomes constant. Denote these constants corresponding to the times t_0, t_1, \dots by Q_0^*, Q_1^*, \dots . Consider the policy Ξ^* defined by ordering these quantities at times t_0, t_1, \dots , respectively. Obviously, orders in Ξ^* are placed only when the level of inventory is zero. The policy Ξ^* is optimal since its cost is less than or equal to the cost of Ξ over any interval $[0, T)$. ■

PROPOSITION 3.3. *There exists an optimal solution such that*

- (i) if t_i is integral then $\Delta_{i+1} \geq \Delta_i$,
- (ii) if $\{t_i\} = \alpha$ then $\Delta_i \geq \Delta_{i+1}$, and
- (iii) if t_i is interior to a feasible interval, then $\Delta_i = \Delta_{i+1}$.

PROOF. The proof follows from the fact that the average holding cost per time unit over the interval $\Delta_i \cup \Delta_{i+1}$ is proportional to

$$\Delta_i^2 + \Delta_{i+1}^2 = (t_i - t_{i-1})^2 + (t_{i+1} - t_i)^2$$

which is strictly convex in t_i , and the minimum is attained when $2t_i = t_{i-1} + t_{i+1}$ (equivalently, $\Delta_i = \Delta_{i+1}$). ■

We now have to prove a certain property of shifts on the circle which helps when proving the existence of a cyclic optimal solution.

LEMMA 3.4. *If a process is defined on a circle by $f(\theta) = \theta + \sigma \pmod{2\pi}$ for some fixed σ (and any starting point), then*

(i) *the process is cyclic if and only if σ/π is rational,*

(ii) *if σ/π is irrational, then for any starting point, every point on the circle is an accumulation point.*

PROOF. The proof of (i) is obvious. For the proof of (ii), suppose θ is not an accumulation point. Thus there exists an interval I_1 of length ϵ , centered at θ , which is not visited infinitely many times. This can happen only if the intervals $I_{j+1} = I_j + \sigma$ ($j = 1, 2, \dots$) are not visited infinitely many times. For some i and j , $I_i \cap I_{i+j} \neq \emptyset$ and therefore $I_1 \cap I_j \neq \emptyset$. If $I_1 = I_j$ then σ/π must be rational. Otherwise, it follows that for some k the union of the intervals I_1, \dots, I_k covers the entire circle, which is a contradiction. ■

DEFINITION 3.5. A value T is called *feasible* if the policy of $\Delta_i = T$ for all i (i.e., $t_i = iT$) is feasible.

PROPOSITION 3.6. *A value T is feasible if and only if T is a rational number that can be represented with a denominator less than or equal to $1/\alpha$.*

PROOF. Consider a mapping from the real line onto a circle so that a point t is mapped to the point at the angle $2\pi t \pmod{2\pi}$. If T is irrational then by Lemma 3.4 every point on the circle is an accumulation point. But if the policy is feasible then there exists an arc of positive length which is not visited and hence T is rational. Suppose $T = J/N$, where J and N are relatively prime. Obviously, for any j , $\{jT\}$ is a multiple of $1/N$. Since there exist positive integers k and l such that $kJ - lN = 1$, it follows that

$$kT = k \frac{J}{N} = l + \frac{1}{N}.$$

Thus, for feasibility it is both necessary and sufficient that $1/N \geq \alpha$. ■

THEOREM 3.7. *There exists an optimal solution where for infinitely many i 's, $\{t_i\}$ is either 0 or α .*

PROOF. Consider any optimal solution (satisfying the properties discussed in Proposition 3.3). Suppose there is also only a finite number of i 's such that $\{t_i\}$ is either 0 or α . Consider the tail of the sequence where for all i , $\{t_i\} \notin \{0, \alpha\}$. By Proposition 3.3 we now have all the Δ_i 's equal. We argue that in this case we can shift the tail of the solution so that all the t_i 's become integral. The proof is as follows. Since all the Δ_i 's are equal to some T , it follows by Proposition 3.6 that T is rational with denominator $N \leq 1/\alpha$, and for infinitely many values of j , $\{jT\} = 1/N$. It follows that we can shift the solution by the amount $1/N$ so that all the order times are feasible and an infinite number of them are integers. ■

Below we consider the following four "finite" problems which we denote by P_{00} , $P_{\alpha\alpha}$, $P_{0\alpha}$, and $P_{\alpha 0}$. These problems are useful in analyzing the optimal solutions for the whole problem.

DEFINITION 3.8. (i) The problem P_{00} is the following: find integers J and N such that the policy of N equally spaced orders over the interval $[0, J]$ (i.e., order every J/N time units) minimizes the average cost per time unit over such intervals, subject to the conditions that the initial inventory level is zero and orders are not placed during intervals of the form $(I, I + \alpha)$, where I is integer.

(ii) The problem $P_{\alpha\alpha}$ is essentially the same as P_{00} except that the interval $[0, J]$ is replaced by $[\alpha, J + \alpha)$.

(iii) The problem $P_{0\alpha}$ is essentially the same as P_{00} except that the interval $[0, J]$ is replaced by $[0, J + \alpha)$.

(iv) The problem $P_{\alpha 0}$ is essentially the same as P_{00} except that the interval $[0, J]$ is replaced by $[\alpha, J)$.

Note that in the above problems we minimize over an infinite domain of values of T , so the existence of an optimal solution is not obvious. Furthermore, we show that the problems $P_{0\alpha}$ and $P_{\alpha 0}$ may not have optimal solutions. A value of T is said to be P_{00} -feasible if there exist J and N for which the policy is feasible such that $T = J/N$.

PROPOSITION 3.9. *Problem P_{00} has an optimal solution.*

PROOF. Let T denote the length of the interval between consecutive orders, i.e., $T = J/N$. The average cost per time unit is therefore

$$f(T) = \frac{K}{T} + \frac{1}{2}T.$$

The latter is a convex function of T and has a minimum at $T_0 = \sqrt{2K}$. It follows from Proposition 3.6 that T is P_{00} -feasible if and only if it is feasible (see Definition 3.5). In general, T_0 may not be feasible. We are interested in the infimum of $f(T)$ over all feasible values of T , i.e., $T = J/N$, where $N \leq 1/\alpha$. Obviously, for each N , there is a minimum of $f(J/N)$ as a function of J and hence there exists a global minimum. ■

REMARK 3.10. The analysis of the problem $P_{\alpha\alpha}$ is essentially the same as P_{00} , namely, T is $P_{\alpha\alpha}$ -feasible if and only if it is feasible, and hence it has an optimal solution. Moreover, the optimal solutions of P_{00} and $P_{\alpha\alpha}$ have the same value.

REMARK 3.11. The problems $P_{0\alpha}$ and $P_{\alpha 0}$ may not have optimal solutions. For example, consider $P_{0\alpha}$ where $\alpha = 1/3$ and the minimum of $f(T)$ occurs at $T_0 = 1/2$. The latter is not $P_{0\alpha}$ -feasible since there do not exist integers J and N such that $N/2 = J + 1/3$. However, by choosing J arbitrarily large and $N = 2J + 1$, we obtain $P_{0\alpha}$ -feasible policies with

$$T = \frac{J + 1/3}{2J + 1}$$

arbitrarily close to T_0 .

We now analyze the sets of $P_{0\alpha}$ -feasible and $P_{\alpha 0}$ -feasible values. Denote by $0 = \tau_0 < \tau_1 < \tau_2 < \dots$ the feasible values of T (we include $\tau = 0$ for the ease of notation below). The properties of the $P_{0\alpha}$ -feasible values, to be proven later (see Proposition 3.12), imply that these values can be enumerated in increasing order.

PROPOSITION 3.12. *For every $i \geq 1$, if τ_i is not $P_{0\alpha}$ -feasible, then there exist infinitely many $P_{0\alpha}$ -feasible values between τ_{i-1} and τ_i and the only accumulation point of this infinite set is τ_i .*

PROOF. Suppose τ_i is not $P_{0\alpha}$ -feasible. We first show that τ_i is an accumulation point of $P_{0\alpha}$ -feasible values less than τ_i . Let $\epsilon > 0$ be any number such that $\epsilon < \tau_i - \tau_{i-1}$. We argue that there exists a $P_{0\alpha}$ -feasible value T such that $\tau_i - \epsilon \leq T < \tau_i$. Suppose $\tau_i - \epsilon$ is not $P_{0\alpha}$ -feasible. Let N be the smallest integer such that $0 < \{N(\tau_i - \epsilon)\} < \alpha$, and let $J = \lfloor N(\tau_i - \epsilon) \rfloor$. Consider any value T such that $\tau_i - \epsilon < T < (J + \alpha)/N$; then $0 < \{NT\} < \alpha$, and hence T is not feasible. It follows that

$$\tau_i - \epsilon < T < \tau_i.$$

Since there are no feasible values between $\tau_i - \epsilon$ and τ_i , it follows that for every $j < N$, jT is not an integer. Therefore, $(J + \alpha)/N$ is $P_{0\alpha}$ -feasible.

Next, we show that there is no accumulation point of this set other than τ_i . Consider any value T' , $\tau_{i-1} < T' < \tau_i$. Thus, T' is infeasible. We claim that there are only finitely many $P_{0\alpha}$ -feasible values T such that $\tau_{i-1} < T < T'$. The proof is as follows. Let N' be the smallest integer such $0 < \{N'T'\} < \alpha$, and let $J' = \lfloor N'T' \rfloor$. Imagine decreasing the value of T from T' . Denote $T^- = J'/N'$. If $T^- < T < T'$ then $0 < \{N'T\} < \alpha$. Thus, for such a T to be $P_{0\alpha}$ -feasible, it is necessary that the corresponding value of N be less than N' . This implies that there are a finite number of $P_{0\alpha}$ -feasible values between T^- and T' . If T^- is feasible then $T^- = \tau_{i-1}$ and we are done. Otherwise, there exist N^- and J^- less than N' and J' , respectively, such that $J^- < N^-T^- < J^- + \alpha$. The above argument shows that there exist only finitely many $P_{0\alpha}$ -feasible values of T between J^-/N^- and T^- . The same argument can now be repeated. Every time we apply this argument the value of N' decreases. Thus, the process terminates in a finite number of steps, and the total number of $P_{0\alpha}$ -feasible values between τ_{i-1} and T' is finite. ■

REMARK 3.13. It follows from the second part of the above proof that if τ_i is $P_{0\alpha}$ -feasible, then there exist only a finite number of $P_{0\alpha}$ -feasible values between τ_{i-1} and τ_i .

PROPOSITION 3.14. *For every $i \geq 1$, if τ_{i-1} is not $P_{\alpha 0}$ -feasible, then there are infinitely many $P_{\alpha 0}$ -feasible values between τ_{i-1} and τ_i and the only accumulation point of this infinite set is τ_{i-1} . If τ_{i-1} is $P_{\alpha 0}$ -feasible, then there exist only a finite number of $P_{\alpha 0}$ -feasible values between τ_{i-1} and τ_i .*

PROOF. This claim is analogous to those of Proposition 3.12 and Remark 3.13. ■

We now show that the sequences of $P_{0\alpha}$ -feasible values between two consecutive feasible values τ_{i-1} and τ_i have special structures:

PROPOSITION 3.15. *If $\tau_i = J/N$ ($i \geq 1$; $\gcd(J, N) = 1$) is not $P_{0\alpha}$ -feasible, then there exist numbers $a = a(i)$ and $b = b(i)$ such that the set of $P_{0\alpha}$ -feasible values between τ_{i-1} and τ_i is the union of some finite set S_i and the set of all the values of the form*

$$\frac{a + Jx}{b + Nx}, \text{ where } x = 0, 1, 2, \dots$$

PROOF. We know that if τ_i is not $P_{0\alpha}$ -feasible, then it is the only accumulation point of $P_{0\alpha}$ -feasible values between τ_{i-1} and τ_i (Proposition 3.12). Let $\epsilon > 0$ be such that $\epsilon < \min(\tau_i - \tau_{i-1}, \alpha/N)$ and let T' be a $P_{0\alpha}$ -feasible value such that $\tau_i - \epsilon < T' < \tau_i$. Denote by S_i the set of all $P_{0\alpha}$ -feasible values T'' such that $\tau_{i-1} < T'' < T'$. This set is finite. Let J' and N' be such that $T' = (J' + \alpha)/N'$ (thus, note that $J' + \alpha = N'T'$). Since $\tau_i = J/N$,

$$\{J' + \alpha + N\tau_i\} = \alpha,$$

and for any j , $0 < j < N$, $\{j\tau_i\} \geq 1/N$, so that

$$\{J' + \alpha + j\tau_i\} \geq \alpha + 1/N.$$

Since $\epsilon < \alpha/N$, we have $N\tau_i > NT' > N\tau_i - N\epsilon > N\tau_i - \alpha$, so that

$$0 < \{(N + N')T'\} = \{J' + \alpha + NT'\} < \alpha$$

and for $j < N$,

$$\{(N' + j)T'\} = \{J' + \alpha + jT'\} \geq \alpha + 1/N - N\epsilon > 1/N \geq \alpha.$$

Thus, the least $P_{0\alpha}$ -feasible value which is greater than T' (and yet smaller than τ_i) is precisely

$$\frac{J' + J + \alpha}{N' + N}.$$

It is easy to verify that this implies the special structure we claim. ■

PROPOSITION 3.16. *If $\tau_i = J/N$ ($i \geq 1; \gcd(J, N) = 1$) is not $P_{\alpha 0}$ -feasible, then there exist numbers $c = c(i)$ and $d = d(i)$ such that the set of $P_{\alpha 0}$ -feasible values between τ_i and τ_{i+1} is the union of some finite set R_i and the set of all the values of the form*

$$\frac{c + Jy}{d + Ny}, \quad \text{where } y = 0, 1, 2, \dots$$

PROOF. The proof is analogous to that of Proposition 3.15. ■

REMARK 3.17. It follows from the proof of Proposition 3.15 that, given the rationals τ_{i-1}, τ_i , one can construct all the $P_{0\alpha}$ -feasible values between them, in the sense that the elements of a set S_i can be listed and a concise description of the rest of the values can be given in terms of a fractional linear function. The same is true for solutions of $P_{\alpha 0}$.

We now consider a more complicated problem $P_{0\alpha 0}$:

DEFINITION 3.18. The problem $P_{0\alpha 0}$ is as follows. Find integers J_1, J_2 and N_1, N_2 such that the policy consisting of N_1 equally spaced orders over the interval $[0, J_1 + \alpha)$, followed by N_2 equally spaced orders over the interval $[J_1 + \alpha, J_1 + J_2)$, minimizes the average cost per time unit over intervals of the form $[0, J_1 + J_2)$, subject to the following conditions:

- (i) the initial inventory level is zero,
- (ii) orders are equally spaced over each of the intervals $[0, J_1 + \alpha)$ and $[J_1 + \alpha, J_1 + J_2)$ (including one order at $J_1 + \alpha$), and
- (iii) orders are not placed during intervals of the form $(I, I + \alpha)$, where I is an integer.

REMARK 3.19. A feasible solution of $P_{0\alpha 0}$ can be described simply as a pair of solutions of $P_{0\alpha}$ and $P_{\alpha 0}$. Problem $P_{0\alpha 0}$ may have no optimal solution. This may happen, for example, when the minimizer $T_0 = \sqrt{2K}$ of $f(T)$ is an integer, e.g., if $K = 0.5$. Here for any $\alpha < 1$, the infimum of $P_{0\alpha 0}$ is $f(1)$ but the latter is not attained.

PROPOSITION 3.20. *If the problem $P_{0\alpha 0}$ does not have an optimal solution then the infimum of the average costs of feasible solutions of $P_{0\alpha 0}$ is equal to the minimum of the problem P_{00} .*

PROOF. It follows from Propositions 3.12 and 3.14 that the infimum is less than or equal to the minimum of P_{00} , since any value τ_i is either feasible or a limit of $P_{0\alpha}$ -feasible and $P_{\alpha 0}$ -feasible values. We now prove the other direction. Consider the problem $P_{0\alpha 0}$. For given J_i, N_i ($i = 1, 2$), denote

$$T_1 = \frac{J_1 + \alpha}{N_1}, \quad T_2 = \frac{J_2 - \alpha}{N_2}$$

(i.e., T_i , $i = 1, 2$, is the length of the interval between consecutive orders over the first and the second intervals, respectively). The average cost over the interval $[0, J_1 + J_2)$ is

$$\frac{(J_1 + \alpha)f(T_1) + (J_2 - \alpha)f(T_2)}{J_1 + J_2}.$$

Let λ denote the infimum of the latter taken over all feasible choices of J_i, N_i ($i = 1, 2$). Note that the claim applies to the case where the infimum is not attained. There exist sequences of values J_i^k, N_i^k tending to infinity with $T_i^k = J_i^k/N_i^k$ such that

- (i) T_i^k tends to T_i^* ,
- (ii) $f(T_i^k)$ tends to some limit $f(T_i^*)$, and
- (iii) the average cost

$$\frac{(J_1^k + \alpha)f(T_1^k) + (J_2^k - \alpha)f(T_2^k)}{J_1^k + J_2^k}$$

tends to λ .

By Propositions 3.12 and 3.14, the quantities T_i^* are feasible (in the sense of Definition 3.5) and hence the average costs $f(T_i^*)$ are greater than or equal to the minimum of the average cost taken over any policy with equally spaced orders. This minimum is precisely the minimum of P_{00} . ■

We can now prove the desired result:

THEOREM 3.21. *There exists a cyclic optimal policy Ξ where the defining cycle is an optimal solution of either P_{00} or $P_{0\alpha 0}$.*

PROOF. By Theorem 3.7 there exists an optimal solution where for an infinite number of i 's $\{t_i\} \in \{0, \alpha\}$. It suffices to consider the following three cases:

(i) For an infinite number of i 's, t_i is integral and for all i , $\{t_i\} \neq \alpha$. In this case the optimal value (of the long-term average cost) is a limit of weighted average of feasible values of problems of the form P_{00} . By Proposition 3.9, the latter has an optimal solution. Thus, the cyclic solution consisting of repetitions of an optimal solution of P_{00} must be optimal.

(ii) For an infinite number of i 's, $\{t_i\} = \alpha$ and for all i (except for $t_i = 0$), t_i is not integral. In this case the solution starting with an order of α at time 0, and continuing with repetitions of an optimal solution of $P_{\alpha\alpha}$, must be optimal. Moreover, by shifting this solution we actually get the same cyclic solution as in (i).

(iii) For an infinite number of i 's, t_i is integral and for an infinite number of i 's, $\{t_i\} = \alpha$. Consider an interval between two consecutive integral values t_i, t_j . Thus, there exist m values $t_k, t_i < t_k < t_j$ ($m \geq 0$), such that $\{t_k\} = \alpha$ and for any t_k such that $t_i < t_k < t_j$, $\{t_k\} \neq 0$. Suppose $m \geq 2$, and denote the points t_k in $[t_i, t_j)$ with $\{t_k\} = \alpha$ by t_{k_1}, \dots, t_{k_m} . The average cost per time unit over the interval $[t_i, t_j)$ is a weighted average of two averages: c_1 over $[t_{k_1}, t_{k_m})$ and c_2 over the union of the intervals $[t_i, t_{k_1})$ and $[t_{k_m}, t_j)$. Obviously, c_1 is greater than or equal to the optimum of the problem $P_{\alpha\alpha}$ (which is in turn equal to the optimum of P_{00}), and c_2 is greater than or equal to the optimum of $P_{0\alpha 0}$. It follows that the cyclic solution consisting of repetitions of the best of the optimal solutions (either of P_{00} or of $P_{0\alpha 0}$) must be optimal. ■

4. Computing an optimal policy. Having shown that a cyclic optimal solution exists, the natural question is whether an optimal cycle can be computed. The answer is not obvious since it requires a search over an infinite domain. For example, we

need to decide whether there exist feasible values

$$T_1 = \frac{a + Jx}{b + Nx} \quad \text{and} \quad T_2 = \frac{c + Jy}{d + Ny}$$

(x and y integral) for $P_{\alpha 0}$ and $P_{0\alpha}$, respectively, which together yield a solution of $P_{0\alpha 0}$ better than J/N , i.e.,

$$\frac{(a + Jx)f(T_1) + (c + Jy)f(T_2)}{(a + Jx) + (c + Jy)} < f\left(\frac{J}{N}\right).$$

This is the content of the following:

PROBLEM 4.1. Given $a, b, c, d, K, J, N > 0$, recognize whether there exist nonnegative integers x and y such that

$$\begin{aligned} \psi(x, y) &= \frac{K(b + Nx) + \frac{1}{2} \frac{(a + Jx)^2}{b + Nx} + K(d + Ny) + \frac{1}{2} \frac{(c + Jy)^2}{d + Ny}}{(a + Jx) + (c + Jy)} \\ &< K \cdot \frac{N}{J} + \frac{1}{2} \cdot \frac{J}{N}. \end{aligned}$$

PROPOSITION 4.2. *Problem 4.1 is decidable.*

PROOF. A pair $x, y \geq 0$ solves the inequality of Problem 4.1 if and only if the following quantity is negative:

$$\begin{aligned} &K\left(b - \frac{aN}{J}\right) + K\left(d - \frac{cN}{J}\right) + \frac{1}{2}(a + Jx)\left(\frac{a + Jx}{b + Nx} - \frac{J}{N}\right) \\ &+ \frac{1}{2}(c + Jy)\left(\frac{c + Jy}{d + Ny} - \frac{J}{N}\right) \end{aligned}$$

or, equivalently,

$$\begin{aligned} (*) \quad &\frac{1}{2} \frac{a + Jx}{b + Nx} \cdot \frac{aN - bJ}{N} + \frac{1}{2} \frac{c + Jy}{d + Ny} \cdot \frac{cN - dJ}{N} \\ &< K\left(\frac{aN}{J} - b\right) + K\left(\frac{cN}{J} - d\right). \end{aligned}$$

The key observation here is that the function

$$g(x) = \frac{a + Jx}{b + Nx}$$

is always monotone. It is increasing if $J/N > a/b$, constant if $J/N = a/b$, and decreasing if $J/N < a/b$. As x tends to infinity, $g(x)$ tends to J/N . A similar observation holds for $(c + Jy)/(d + Ny)$. All these imply that the left-hand side of (*) amounts to the sum of two monotone nonincreasing functions of x and y , respectively. Thus, the problem can be solved by letting x and y tend to infinity.

More precisely, the inequality of the problem has a solution if and only if

$$\frac{1}{2} \cdot \frac{J(aN - bJ)}{N^2} + \frac{1}{2} \cdot \frac{J(cN - dJ)}{N^2} < K\left(\frac{aN}{J} - b\right) + K\left(\frac{cN}{J} - d\right). \quad \blacksquare$$

If one wants to minimize $\psi(x, y)$ then the following procedure can be used. First, note the infimum is always less than or equal to the right-hand side, and we know how to decide whether it is strictly less than the latter. Assuming it is, for every value of x it suffices to check a finite number of values of y . As either x or y tends to infinity, the value of the left-hand side tends to that of the right-hand side, so after a finite number of steps the minimum is reached. However, we still have to develop the tools for recognizing the minimum when it is reached.

PROPOSITION 4.3. *The infimum of $\psi(x, y)$ (see Problem 4.1) over all nonnegative integers x, y is computable.*

PROOF. It follows from Proposition 4.2 that if the minimum does not exist then the infimum is equal to $f(J/N) = KN/J + \frac{1}{2}J/N$. Thus we now assume the minimum exists and is less than $f(J/N)$.

We first argue that for any $\epsilon > 0$ it can be decided whether there exist nonnegative integers x, y such that

$$\psi(x, y) < f(J/N) - \epsilon.$$

The idea is essentially the same as that in the proof of Proposition 4.2. The problem is equivalent to deciding the existence of x, y such that

$$\begin{aligned} (**) \quad & \frac{1}{2} \frac{a + Jx}{b + Nx} \cdot \frac{aN - bJ}{N} + \epsilon(a + Jx) + \frac{1}{2} \frac{c + Jy}{d + Ny} \cdot \frac{cN - dJ}{N} + \epsilon(c + Jy) \\ & < K\left(\frac{aN}{J} - b\right) + K\left(\frac{cN}{J} - d\right). \end{aligned}$$

Here the left-hand side is a sum of functions $g_1(x) = g_1(x; \epsilon)$ and $g_2(y) = g_2(y; \epsilon)$. Consider the function $g_1(x)$ (the function $g_2(y)$ can be analyzed in the same way). If $a/b = J/N$ the minimum of $g_1(x)$ is at $x = 0$. Otherwise, $g_1(x)$ is strictly convex with a unique minimum (over the reals) which can be derived analytically. The minimum of $g_1(x)$ over the nonnegative integers can then be found using a convexity argument. Thus, given any x and y we can decide whether they minimize $\psi(x, y)$. So, assuming the minimum exists, by enumeration we reach it and recognize it. \blacksquare

REMARK 4.4. It is possible to develop a more efficient procedure for minimizing $\psi(x, y)$. The question amounts to finding the maximum value ϵ^* of ϵ for which the value of

$$G(\epsilon) = \min_x g_1(x; \epsilon) + \min_y g_2(y; \epsilon)$$

is less than or equal to the right-hand side of (**). As the minimum of countably many linear functions, $G(\epsilon)$ is monotone increasing and concave. When the minimum of $\psi(x, y)$ is strictly less than $f(J/N)$, we not only know that but we also find x, y such that $\psi(x, y) < f(J/N)$. It follows that we can construct an interval containing ϵ^* over which $G(\epsilon)$ has only finitely many pieces and hence using binary search we can locate ϵ^* exactly.

REMARK 4.5. A simpler problem can be handled essentially in the same way: given $a, b, C, D, K, J, N > 0$, recognize whether there exists a nonnegative integer x

such that

$$\frac{K(b + Nx) + \frac{1}{2} \frac{(a + Jx)^2}{b + Nx} + C}{(a + Jx) + D} < K \frac{N}{J} + \frac{1}{2} \frac{J}{N},$$

and if so, find the minimum of the left-hand side.

Given feasible values $T_1 = (J_1 + \alpha)/N_1$ and $T_2 = (J_2 - \alpha)/N_2$ of $P_{0\alpha}$ and $P_{\alpha 0}$, respectively, denote by

$$\phi(J_1, N_1; J_2, N_2) = \frac{(J_1 + \alpha)f(T_1) + (J_2 - \alpha)f(T_2)}{J_1 + J_2}$$

the value of the corresponding solution of $P_{0\alpha 0}$. Recall that $T_0 = \sqrt{2K}$ minimizes $K/T + \frac{1}{2}T$ over the reals and $0 = \tau_0 < \tau_1 < \dots$ are the feasible values, i.e., rationals with denominators less than or equal to $1/\alpha$.

PROPOSITION 4.6. *Suppose $\tau_{i-1} < T_0 < \tau_i$ and let $T_1 = (J_1 + \alpha)/N_1$ be any $P_{0\alpha}$ -feasible value in the interval (T_0, τ_i) . Let $T_2 = (J_2 - \alpha)/N_2$ denote the maximal $P_{\alpha 0}$ -feasible value in the interval $[\tau_i, \tau_{i+1})$. Under these conditions,*

$$\phi(J_1, N_1; J'_2, N'_2) > \phi(J_1, N_1; J_2, N_2)$$

for any $P_{\alpha 0}$ -feasible value $T'_2 = (J'_2 - \alpha)/N'_2$ such that either (i) $T'_2 > J_2 - \alpha$ or (ii) $J'_2 > J_2$ and $T'_2 \geq \tau_{i+1}$.

PROOF. The assumptions on T_1 and T_2 imply $f(T_1) \leq f(T_2)$. Thus ϕ will increase if J_2 and N_2 are replaced by J'_2 and N'_2 such that $J'_2 > J_2$ and $f(J'_2/N'_2) \geq f(J_2/N_2)$. In case (i) the claim follows from the inequalities.

$$J'_2 - \alpha \geq T'_2 > J_2 - \alpha \quad \text{and} \quad T'_2 > J_2 - \alpha \geq T_2.$$

The proof of case (ii) is similar. ■

COROLLARY 4.7. *Assume the conditions of Proposition 4.6, and $f(\tau_i) < f(\tau_{i-1})$. Given any $P_{0\alpha}$ -feasible value $T_1 = (J_1 + \alpha)/N_1$ such that $f(T_1) < F_{00}$, in order to decide whether $\phi(J_1, N_1; J_2, N_2) < F_{00}$ for some $P_{\alpha 0}$ -feasible pair (J_2, N_2) , it suffices to consider candidates $T_2 = (J_2 - \alpha)/N_2$ only from a union of a finite number of intervals of the form $[\tau_{k-1}, \tau_k)$, and except for the case $k = i + 1$, in each interval only a finite number of values need to be considered.*

PROOF. The sequence of $P_{\alpha 0}$ -feasible values which lie between two consecutive values τ_{k-1} and τ_k (see Proposition 3.16) accumulates into τ_{k-1} . There are only a finite number of k 's such that $k < i$. If $k < i$ then the value $\phi(J_1, N_1; J_2, N_2)$ increases as the value $T_2 = (J_2 - \alpha)/N_2$ varies along the sequence. So, it suffices to check the first member of the sequence (as well as the finite number of values not in the sequence which lie in the same interval). Similarly, if $k = i$, it suffices to check among the members of the structured sequence only the values of T_2 which are greater than or equal to T_0 (there are only finitely many of these) and the largest one which is less than T_0 . If $k > i + 1$, then the claim follows from Proposition 4.6. ■

REMARK 4.8. The number of values considered in Corollary 4.7 can be further reduced by noting that $T = J/N$ is $P_{0\alpha}$ -feasible if and only if $T + 1 = (N + J)/N$ is. Now, if $T'_2 > T_0 + 1$ then $\phi(J_1, N_1; J'_2 - N'_2, N'_2) < \phi(J_1, N_1; J'_2, N'_2)$. Thus only values less than $T_0 + 1$ need to be checked. A similar argument shows that it suffices to consider only values greater than $T_0 - 1$.

5. Summary of the algorithm. We now sketch an algorithm for an optimal policy. Most of the effort in the algorithm goes into the determination of whether

$F_{0\alpha 0} < F_{00}$. Indeed, if this inequality holds and a feasible solution with $\phi(J_1, N_1; J_2, N_2) < F_{00}$ is known then, as we argue below, there remain only a finite number of solutions of $P_{0\alpha 0}$ that have to be checked before the optimal solution is found. The main steps of the algorithm are as follows.

(i) The first step is to calculate the minimum $T_0 = \sqrt{2K}$. If T_0 is rational with denominator $N \leq 1/\alpha$ then the policy of ordering every T_0 time units is optimal. Otherwise, find two nonnegative rationals τ_{i-1}, τ_i (see the definition preceding Proposition 3.12) with denominators not greater than $1/\alpha$, such that $\tau_{i-1} < T_0 < \tau_i$. The optimal solution of P_{00} (see Definition 3.8) is determined by either τ_{i-1} or τ_i , i.e., $F_{00} = \min\{f(\tau_{i-1}), f(\tau_i)\}$.

(ii) In this step, we construct (in the sense of Remark 3.17) the sets of feasible solutions of $P_{0\alpha}$ and $P_{\alpha 0}$ which lie strictly between τ_{i-1} and τ_i . If there are no feasible T 's neither for $P_{0\alpha}$ nor for $P_{\alpha 0}$ such that $f(T) < F_{00}$, then the optimal solution for the whole problem is the optimal solution of P_{00} found in (i), so we stop. Otherwise, we continue with step (iii).

(iii) We now determine whether $F_{0\alpha 0} < F_{00}$. We do this by analyzing the neighborhood of τ_{i-1} or τ_i , or both (depending on how P_{00} is attained as the minimum of $f(\tau_{i-1})$ and $f(\tau_i)$). Consider the case where $f(\tau_i) = F_{00} < f(\tau_{i-1})$; the other cases are handled analogously. Suppose there exists a feasible value $T_1 = (J_1 + \alpha)/N_1$ of $P_{0\alpha}$ such that $f(T_1) < F_{00}$, so T_1 lies strictly between τ_{i-1} and τ_i . Given T_1 we search for feasible values $T_2 = (J_2 - \alpha)/N_2$ of $P_{\alpha 0}$ so as to minimize the value $\phi(J_1, N_1; J_2, N_2)$ of the corresponding solution of $P_{0\alpha 0}$. This is done as follows. We are interested only in $T_2 = (J_2 - \alpha)/N_2$ such that $\phi(J_1, N_1; J_2, N_2) < F_{00}$. Proposition 4.6 (case (i)) explains how to bound the number of intervals $[\tau_{k-1}, \tau_k)$ that have to be checked, so we can restrict attention to a finite set of such intervals. Suppose first that $k \neq i + 1$; the case $k = i + 1$ is discussed in (iv). By Corollary 4.7, in each interval there is only a finite number of points that have to be considered. For each possible value of T_2 , we consider all the values of T_1 between τ_{i-1} and τ_i which are $P_{0\alpha}$ -feasible. Here we rely on Remark 4.5 for finding the optimal T_1 or concluding that the infimum is not less than F_{00} .

(iv) Now we consider the case $k = i + 1$. Here we have to consider an infinite number of values of both T_1 and T_2 . The two sequences tend to τ_{i-1} from different sides. We rely on Proposition 4.3 and either find an optimal pair T_1, T_2 or conclude that no such pair gives a value less than F_{00} .

6. Finding a good approximate solution. If one is satisfied with approximately optimal policies then the following proposition is useful. We know that the optimal value is the minimum of F_{00} and $F_{0\alpha 0}$. If F_{00} (which is easy to compute) is taken as an approximately optimal solution, the error can be estimated as follows.

PROPOSITION 6.1. *If J_i, N_i ($i = 1, 2$) define an optimal solution for $P_{0\alpha 0}$ such that*

$$f\left(\frac{J_1 + \alpha}{N_1}\right) \leq f\left(\frac{J_2 - \alpha}{N_2}\right),$$

then the average cost per time unit in this solution, $F_{0\alpha 0}$, and the optimal average cost per time unit, F_{00} , in P_{00} satisfy

$$F_{00} \leq F_{0\alpha 0} + \frac{1}{2N_1}.$$

PROOF. Denote by \bar{T} the smallest value greater than $T_1 = (J_1 + \alpha)/N_1$ which is feasible (for P_{00}), i.e., it is rational with denominator not greater than $1/\alpha$. Note that

$$\bar{T} \leq T_1 + \frac{(1 - \alpha)}{N_1}$$

since N_1 times the right-hand side is an integer. Actually, if $\bar{T} = j/n$ ($\gcd(j, n) = 1$), then by the classical Dirichlet's theorem,

$$\bar{T} - T_1 < \frac{1}{nN_1}.$$

Since $\bar{T} > T_1$,

$$\begin{aligned} F_{00} - F_{0\alpha 0} &\leq f(\bar{T}) - F_{0\alpha 0} \leq f(\bar{T}) - f(T_1) \\ &= \left(\frac{K}{\bar{T}} + \frac{1}{2}\bar{T} \right) - \left(\frac{K}{T_1} + \frac{1}{2}T_1 \right) \leq \frac{1}{2}(\bar{T} - T_1) \leq \frac{1}{2nN_1}. \quad \blacksquare \end{aligned}$$

Acknowledgements. We are grateful to Dafna Sheinwald for pointing out a delicate flaw in one of the proofs in an earlier version.

References

- [1] Andres, F. M. and Emmons, H. (1976). On the Optimal Packaging Frequency of Products Jointly Replenished. *Management Sci.* **22** 1165–1166.
- [2] Buzacott, J. A. and Ozkarahan, I. A. (1983). One and Two-Stage Scheduling of Two Products with Distributed Inserted Idle Time: The Benefits of a Controllable Production Rate. *Naval Res. Logist. Quart.* **30** 675–696.
- [3] Dobson, G. (1987). The Economic Lot-Scheduling Problem: Achieving Feasibility using Time-Varying Lot Sizes. *Oper. Res.* **35** 764–771.
- [4] Elmaghraby, S. L. (1978). The Economic Lot Scheduling Problem (ELSP): Review and Extensions. *Management Sci.* **24** 587–598.
- [5] Graves, S. C. (1981). A Review of Production Scheduling. *Oper. Res.* **29** 646–675.
- [6] Maxwell, W. L. and Singh, H. (1983). The Effect of Restricting Cycle Times in the Economic Lot Scheduling Problem. *IIE Trans.* **15** 235–241.
- [7] Muckstadt, J. A. and Singer, H. M. (1978). Comments on Single Cycle Continuous Review Policies for Arborescent Production/Inventory Systems. *Management Sci.* **24** 1766–1768.
- [8] Roundy, R. (1986). A 98-Percent-Effective Lot-Sizing Rule for a Multi-Product, Multi-Stage Production Inventory System. *Math. Oper. Res.* **11** 699–727.
- [9] Schweitzer, P. J. and Silver, E. A. (1983). Mathematical Pitfalls in the One Machine Multi-Product Economic Lot Scheduling Problem. *Oper. Res.* **31** 401–405.

HASSIN: SCHOOL OF MATHEMATICS, TEL AVIV UNIVERSITY, TEL AVIV 69978, ISRAEL

MEGIDDO: IBM RESEARCH DIVISION, ALMADEN RESEARCH CENTER, SAN JOSE, CALIFORNIA 95120-6099

SCHOOL OF MATHEMATICAL SCIENCES, TEL AVIV UNIVERSITY, TEL AVIV 69978, ISRAEL