Maximum Sustainable Yield of Populations with Continuous Age-Structure

ROGER ARDITI  
Institute of Zoology and Animal Ecology, University of Lausanne,  
CH-1015 Lausanne, Switzerland

AND

BERNARD DACOROGNA  
Department of Mathematics, Federal Polytechnic School,  
CH-1015 Lausanne, Switzerland

Received 26 July 1991

ABSTRACT

We address the problem of finding the harvesting policy that will maximize the yield and maintain a population in a steady state. The population is characterized by continuous age classes and therefore follows differential equations. Here, we assume that the equations are linear (no density dependence).

Two possible constraints are considered: either recruitment or total population are fixed to a constant. Under these conditions, the optimal policy is to harvest the fraction of a younger age class and to harvest totally an older age class. The optimal solution can be calculated explicitly if the fecundity and mortality schedules are given.

The solution is compared to the simpler strategy of harvesting all individuals beyond a single age class. It is shown that the latter strategy can be much less profitable than harvesting two age classes because it cannot take account of the different values of individuals according to their age.

INTRODUCTION

Harvesting theory is concerned with the rational exploitation of biological resources, like populations of game, fish, and trees. The theory has developed in the context of models where the population is represented by a simple scalar variable as well as in the context of models with age structure or stage structure. Both types of models have considered the simple linear case (density-independent) as well as nonlinear cases obtained when survival or fecundity rates are assumed to depend on the population density. Getz and Haight have developed...
oped structured models as generalizations of the classical Leslie matrix approach. This discrete-time model applies to populations exhibiting distinct reproduction seasons, like many biological resources in temperate environments. However, many species of animals and plants reproduce continuously throughout the year, especially in the tropics. In such cases, the discrete approach is not appropriate. Continuous models lead naturally to differential equations rather than difference equations. Because humans reproduce all year round, classical human demography has generated a large body of literature on continuous models [8,9,10] that has also been extended to animal and plant problems [2].

In this paper, we present such an application by considering a relatively simple question: the determination of the harvesting policy that ensures the maximum sustainable yield (MSY) from a population in a steady state. For unstructured populations, it was one of the oldest problems of fisheries theory [7,11]. For structured populations, Beddington and Taylor [1] have solved the linear problem in the case of discrete age structure. The optimal policy consists in harvesting two age classes only: a younger age class is partially harvested and an older age class is completely harvested (implying that no individual will exist beyond this age). Getz and Haight [5] have made a general treatment for nonlinear models with discrete age or stage structures. In the present paper, we address the linear case for a population with continuous age structure.

1. CONTINUOUS LINEAR AGE-STRUCTURED MODEL WITH HARVESTING

For the most part, we follow the notations of Getz and Haight [5], modified as appropriate to the continuous setting of the model. Let \( t \) denote the time and \( a \) the age of an individual. The state variable \( x(a,t) \) is the number of individuals aged \( a \) present at time \( t \). The life history of this population is characterized by two schedules: \( b(a) \), the rate at which individuals aged \( a \) reproduce and \( \mu(a) \), the rate at which individuals aged \( a \) die. These rates are assumed to be intrinsic to the species and not to depend on the time \( t \). The linear, density-independent assumption asserts that these rates do not vary with population abundance.

The action of harvesting is represented by the function \( u(a,t) \), the number of individuals aged \( a \) that are removed at time \( t \). We also define the function \( h(a,t) = u(a,t)/x(a,t) \) as the removal rate. It is the additional mortality rate due to harvesting. Thereafter, it will also be called the control. The population dynamics is described by a differen-
tial equation with longitudinal differentiation along the cohort:
\[ \frac{\partial x(a,t)}{\partial a} + \frac{\partial x(a,t)}{\partial t} = -\left[ \mu(a) + h(a,t) \right] x(a,t). \] (1.1)

Therefore, the proportion of the cohort born at time \( t - a \) that has survived until time \( t \) is:
\[ s(a,t) = \exp \left\{ - \int_a^t \left[ \mu(a - t + \alpha) + h(a - t + \alpha, \alpha) \right] d\alpha \right\}. \] (1.2)

Note that, if this cohort has suffered no harvesting until age \( a \), \( s(a,t) \) reduces to the intrinsic survival schedule \( k(a) \):
\[ k(a) = \exp \left\{ - \int_0^a \mu(\alpha) d\alpha. \right\} \] (1.3)

At any time, the number of newborns is given by
\[ x(0,t) = \int_a^\infty b(a) x(a,t) da. \] (1.4)

At time \( t = 0 \), the population is assumed to have the initial structure \( x(a,0) = x_0(a) \). At time \( t \), the individuals older than \( a = t \) are the survivors of the cohorts already present at \( t = 0 \). Younger individuals are the survivors of cohorts born after \( t = 0 \). The solution of the differential equation (1.1) is therefore
\[ x(a,t) = \begin{cases} x(a-t,0) s(a,t) & \text{if } a \geq t, \\ x(0,t-a) s(a,t) & \text{if } a < t. \end{cases} \] (1.5)

Let \( w(a,t) \) denote the value (e.g., in dollars) at time \( t \) of an individual aged \( a \). The performance \( J \) of a harvesting strategy \( h \) applied from time \( t = 0 \) to time \( t = T \) is
\[ J(h) = \int_0^T \int_0^\infty w(a,t) h(a,t) x(a,t) da dt + \int_0^\infty L(w(a,T) x(a,T)) da, \] (1.6)

where the first term is the total value of the harvested individuals and the second term is the residual value of the population remaining at time \( T \).
2. THE MAXIMUM SUSTAINABLE YIELD PROBLEM

This problem consists in finding the harvesting strategy $h$ that maximizes the value of the harvest under steady-state conditions. We assume therefore that the whole population structure is constant in time: $x(a, t) = x_0(a)$ for all $t$. For simplicity, we can drop the subscript $0$ and write $x(a)$. We also assume that the value function $w$ does not vary with time and we disregard any value attached to the standing stock per se: $L = 0$. Therefore, we can eliminate the time from the problem and simplify Equations (1.1)-(1.5).

We also assume that $h \geq 0$ (no negative harvesting, i.e., no stocking) and write the maximization problem as

$$\max_{h \geq 0} \left\{ J(h) = \int_0^\infty w(a) h(a) x(a) da \right\}. \quad (2.1)$$

Since the mortality and fecundity schedules are assumed to be independent of the population density, we further need a constraint on the size of the population. We consider two possible constraints, which give rise to two different problems: either (P) recruitment $x(0)$ is fixed to a constant, or (Q) the total population is fixed to a constant. With no loss of generality, these constants can be set to 1. [The problem solved by Beddington and Taylor [1] was the discrete analogue of (Q)].

The result is as follows. In both cases (constant recruitment or constant total population) the optimal strategy is to harvest partially at some age $a$, then to harvest completely at a final age $b$ and not to harvest at any other age. The optimal population will then be, in both cases, of the form

$$x(a) = \begin{cases} l(a) x(0) & \text{if } a < \tilde{a} \\ \theta l(a) x(0) & \text{if } \tilde{a} \leq a < \tilde{b} \\ 0 & \text{if } a \geq \tilde{b} \end{cases} \quad (2.2)$$

where $\theta$ is a constant smaller than one.

In both cases, we will be able to give explicitly the values $\tilde{a}, \tilde{b}, \theta$ [and $x(0)$ in case of problem (Q)] in terms of $w$, $b$ and $\mu$.

To be more precise, we make the following hypotheses:

$$\begin{cases} \mu, b, w : \mathbb{R}_+ \rightarrow \mathbb{R}_+ = [0, \infty) \text{ are piecewise continuous}, \\ w \text{ is } C^1 \text{ and } \mu(a) \geq \mu_0 > 0 \text{ for } a \text{ large enough}. \end{cases} \quad (H)$$
We now need to establish the following conventions, which will be explained in the discussion that follows.

**CONVENTIONS**

(i) We agree to replace the expression

\[ J(h) = \int_0^\infty w(a) x(a) h(a) \, da \quad (2.3) \]

by

\[ J(x) = w(0) x(0) + \int_0^\infty \left[ w'(a) - \mu(a) w(a) \right] x(a) \, da. \quad (2.4) \]

These two definitions are strictly equivalent (see following) if \( h \in C^0 \) and \( x \in C^1 \). The advantage of (2.4) over (2.3) is that (2.4) is well defined even when \( x \) is discontinuous (e.g., if the control \( h \) consists in removing individuals at a precisely given age only), which is obviously our case.

(ii) In the same spirit, we agree to replace (1.1), which with our hypothesis \( x(a,t) = x(a) \) is just

\[ x'(a) = - \left[ \mu(a) + h(a) \right] x(a), \quad (2.5) \]

by

\[ \frac{x(a)}{l(a)} = \frac{x(0)}{l(0)} e^{-H(a)} \quad (2.6) \]

where

\[ H(a) = \int_0^a h(\alpha) \, d\alpha. \quad (2.7) \]

As before, (2.5) and (2.6) are obviously strictly equivalent if \( h \) is continuous. The advantage of (2.6) over (2.5) is again that (2.6) is meaningful even if \( H \) (and thus \( x \)) is discontinuous.

**PROOF OF THE EQUIVALENCE OF (2.3) AND (2.4)**

Using (2.5) we have

\[ x(a) h(a) = - \mu(a) x(a) - x'(a). \]
Therefore

\[ J(h) = \int_0^\infty w(a) x(a) h(a) \, da \]

\[ = \int_0^\infty [ - w(a) x'(a) - w(a) \mu(a) x(a) ] \, da. \]

Integrating by parts, bearing in mind that \( x(\infty) = 0 \) since \( \mu \) satisfies (H), we find that

\[ J(h) = w(0) x(0) + \int_0^\infty [ w'(a) - w(a) \mu(a) ] x(a) \, da = J(x). \]

**Remarks**

(i) The conventions just presented are needed only to avoid the use of distribution theory. In our case it is natural to allow discontinuous \( x \) and "Dirac masses" \( h \), and the meaning of (2.3) and (2.5) has to be reinterpreted. This is what we have done in (2.4) and (2.6).

(ii) Note also that the hypothesis \( h \geq 0 \) (i.e., no stocking) is then equivalent to \( H \) [in (2.7)] increasing and, therefore, is equivalent (see (2.6)) to

\[ \frac{x(a)}{l(a)} \leq \frac{x(b)}{l(b)} \text{ for every } a \leq b. \quad (2.8) \]

We can now state the theorem that gives the solution of the problems (P) and (Q) mentioned earlier.

We first define the set of admissible functions:

\[ X = \{ x: \mathbb{R}_+ \to \mathbb{R}_+ \text{ piecewise continuous and continuous from the right satisfying (2.8)} \}. \]

The fact that \( x \) is assumed to be continuous from the right is just a convention that says that harvesting is performed before the point of discontinuity.

The constraints corresponding to problems (P) and (Q), respectively, further restrict the set \( X \) to

\[ X_P = \left\{ x: x \in X \text{ and } \int_0^\infty x(a) b(a) \, da = x(0) = 1 \right\}. \]
and

\[ X_0 = \left\{ x : x \in X \text{ and } \int_0^x x(a) b(a) \, da = x(0) \text{ and } \int_0^x x(a) \, da = 1 \right\}. \]

Problems (P) and (Q) can therefore be stated as

\[ \max \{ J(x) : x \in X_p \} \quad (P) \]

and

\[ \times \]

\[ \max \{ J(x) : x \in X_0 \}. \quad (Q) \]

**THEOREM. PART I.**

Problems (P) and (Q) admit solutions that always have (in either case) the following form:

\[ x(a) = \begin{cases} 
  l(a) x(0) & \text{if } a < \tilde{a} \\
  \theta l(a) x(0) & \text{if } \tilde{a} \leq a < \tilde{b} \\
  0 & \text{if } a \geq \tilde{b} 
\end{cases} \quad (2.9) \]

where \( \tilde{a} < \tilde{b} \) are constants and \( 0 \leq \theta \leq 1 \). Furthermore, \( h(a) = 0 \) if \( a \neq \tilde{a} \) and \( a \neq \tilde{b} \), and \( h \) consists in reducing the population \( l(a) \) to \( \theta l(a) \) at age \( \tilde{a} \) and removing the whole population at age \( \tilde{b} \).

Part II. Moreover, \( \tilde{a} \), \( \tilde{b} \) and \( \theta \) are given explicitly as follows.

Let \( \tilde{a} \) be a solution of

\[ \int_0^{\tilde{a}} b(\alpha) l(\alpha) \, d\alpha = 1. \quad (2.10) \]

Let, for \( a \leq \tilde{a} \leq b \),

\[ J(a, b) = \frac{\left[ \int_a^{\tilde{a}} b(\alpha) l(\alpha) \, d\alpha \right] w(b) l(b) + \left[ \int_{\tilde{a}}^b \frac{b(\alpha)}{l(\alpha)} l(\alpha) \, d\alpha \right] w(a) l(a)}{\int_{\tilde{a}}^b \frac{b(\alpha)}{l(\alpha)} l(\alpha) \, d\alpha}. \quad (2.11) \]

Then, in the case of problem (P), \( \tilde{a} \) and \( \tilde{b} \) are given by the maximal condition

\[ J(\tilde{a}, \tilde{b}) \geq J(a, b) \text{ for every } a \in [0, \tilde{a}] \text{ and } b \in [\tilde{a}, \infty). \quad (2.12) \]
Furthermore, $\theta$ is given by

$$
\theta = \frac{\int_{\tilde{a}}^{\tilde{b}} b(\alpha) l(\alpha) \, d\alpha}{\int_{\tilde{a}}^{\tilde{b}} b(\alpha) l(\alpha) \, d\alpha}.
$$

(2.13)

In the case of problem (Q), $\tilde{a} \leq \tilde{a} \leq \tilde{b}$ are given by

$$
x(0) \int(\tilde{a}, \tilde{b}) \geq x(0) \int(\alpha, \beta) \quad \text{for every } \alpha, \beta \text{ with } \alpha \leq \tilde{a} \leq \beta,
$$

(2.14)

where $\int(\alpha, \beta)$ and $\theta$ are given by (2.11) and (2.13), respectively, and

$$
x(0) = \frac{\int_{\alpha}^{\beta} b(\alpha) l(\alpha) \, d\alpha}{\int_{0}^{\alpha} l(\alpha) \, d\alpha \int_{\alpha}^{\beta} b(\alpha) l(\alpha) \, d\alpha + \int_{\alpha}^{\beta} b(\alpha) l(\alpha) \, d\alpha \int_{\alpha}^{\beta} l(\alpha) \, d\alpha}.
$$

(2.15)

Remarks

(i) Observe first that in (2.10) $\tilde{a}$ needs not to be unique (this could happen if $\beta = 0$ in some parts). However, once assumed that $\tilde{a}$ satisfies (2.10), the definition (2.11) of $\int$ is independent of the value of $\tilde{a}$.

(ii) Note that in both cases uniqueness of solutions $x$ is related to uniqueness of $\tilde{a}$ and $\tilde{b}$. For example, in the case of fixed recruitment (P), the uniqueness of solutions depends only on the uniqueness of $(\tilde{a}, \tilde{b})$ satisfying (2.12).

(iii) The same methods can be applied to the discrete cases studied by Getz and Haight [5] and Beddington and Taylor [1]. The difference is that the fecundity schedule $b$ is then a sum of Dirac masses rather than a piecewise continuous function. Note that in the following proof, $b$ only appears as $|b(\alpha)|l(\alpha) \, d\alpha$ and $l$ is continuous by hypothesis. The mortality $\mu$ must be assumed to be continuous, as usual.

(iv) Pontryagin's Maximum Principle could have been used to show that impulse controls were a necessary condition. However, our approach has the advantage of giving the control explicitly, showing that the PMP's condition is also sufficient.

(v) Problems (P) and (Q) are obviously intimately related. Indeed, the solutions exhibit the same features. However, they are not equivalent. If the solution $x_p$ of (P) is renormalized to $x'_p$ such that $\int x'_p(\alpha) \, d\alpha = 1$, it becomes a function admissible for (Q), but it does not
necessarily give the maximum yield under such constraint. Similarly, if the solution $x_0$ of (Q) is renormalized to $x'$ such that $x'(0) = 1$, it becomes admissible for (P) but is not necessarily the maximizing solution of (P).

2.1 DISCUSSION

As in the discrete case [1], the optimal strategy consists in harvesting individuals at exactly two ages: partial harvest at a younger age $\tilde{a}$ and total harvest at an older age $\tilde{b}$. With some organisms like trees or deer, it will be possible to apply this policy in an approximate way. However, it will not be feasible in many other cases. With fish, for example, the common practice consists in setting the minimal mesh size of fishing nets. Assuming that body size is perfectly correlated with age, and assuming also that fishing is perfectly efficient, this policy consists therefore in harvesting all individuals beyond a critical age $\tilde{c}$. It is then interesting to compare this practice with the ideal strategy of harvesting at two distinct ages.

Harvesting at a single age $\tilde{c}$ gives the following population distribution instead of (2.2):

$$x(a) = \begin{cases} \mathbb{I}(a) x(0) & \text{if } a < \tilde{c} \\ 0 & \text{if } a \geq \tilde{c}. \end{cases}$$

Whether constraints (P) or (Q) are considered, i.e., fixed recruitment or fixed population size, the conditions $x \in X_P$ or $x \in X_Q$ both necessarily give

$$\int_0^{\tilde{c}} b(a) l(a) \, d\alpha = 1,$$

which, by definition (2.10), implies that $\tilde{c} = \tilde{a}$.

Under the constraint of single harvesting, the important consequence is that the harvesting age $\tilde{a}$, the age distribution $x$ and the control $h$ all become fixed. No maximization of $J(x)$ taking account of the value function $w(a)$ is possible. In particular, it is possible that $w(\tilde{a}) = 0$, in which case harvesting will bring no profit at all!

We compared numerically the performance of the two strategies in a few examples. The first example is based on the data for Red Deer in Scotland given by Beddington and Taylor [1]. They give a life table with survival, fecundity, and body weight as functions of age. To make the age structure continuous, we added the age zero and we used linear interpolations for noninteger ages. The age for single harvesting is
\( \hat{a} = 6.18 \). In the case of problem P (recruitment = 1), the performance in terms of numbers of individuals removed \((w = 1)\) is \( J_1 = 0.853 \). The double-harvest optimal strategy is \( \theta = 0.793, \hat{a} = 0.0, \) and \( \hat{b} = 7.0; \) it improves the performance very little to the value \( J_2 = 0.857 \). In the case of problem Q (total population size = 1), the performance of single harvesting at age \( \hat{a} = 6.18 \) is \( J_1 = 0.155 \). The double-harvest optimal strategy is \( \theta = 0.545, \hat{a} = 0.0 \) and \( \hat{b} = 9.0; \) the performance is significantly improved to the value \( J_2 = 0.196 \).

The last results are very close to those obtained by Beddington and Taylor [1] in the discrete framework. This justifies the discrete approach in retrospect. Mathematically, the discrete and continuous formalizations are very different. It is not at all obvious that the continuous solution (that we have found directly) can be obtained as the limit of the discrete solution.

In another example, we took a hypothetical population with the constant mortality \( \mu = 0.1 \), the fecundity schedule \( b(a) \) represented on Figure 1, and a bell-shaped value function \( w \) that we varied as indicated on Table 1. We report in Table 1 the strategies and performances that we found. Apparently, the performance attained with two harvests \( (J_2) \) is not very different from that obtained with a single harvest \( (J_1) \) as long as the value \( w(\hat{a}) \) is not too small. In some cases, the double-harvest strategy represents considerable improvement.

Density dependence has been ignored in the present model: survival and fecundity rates depend only on age, not on population size. It might be interesting to examine the consequence of introducing density de-

\[ \text{FIG. 1. Example of fecundity schedule for a hypothetical species.} \]

It reaches maximum fecundity at age 7 and does not reproduce after age 10. See text and Table 1 for further details.
TABLE I

Comparison of the Single-Harvest Performance $J_1$ with the Double-Harvest Performance $J_2$ in Problem (P) (Fixed Recruitment).

<table>
<thead>
<tr>
<th>$w$</th>
<th>$a(6-a)$</th>
<th>$a(5-a)$</th>
<th>$a(4-a)$</th>
<th>$a(3-a)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{a}$</td>
<td>3.11</td>
<td>3.11</td>
<td>3.11</td>
<td>3.11</td>
</tr>
<tr>
<td>$\hat{a}$</td>
<td>2.88</td>
<td>2.35</td>
<td>1.87</td>
<td>1.41</td>
</tr>
<tr>
<td>$\hat{b}$</td>
<td>10.05</td>
<td>10.02</td>
<td>10.00</td>
<td>10.08</td>
</tr>
<tr>
<td>$J_1$</td>
<td>6.58</td>
<td>4.30</td>
<td>2.02</td>
<td>0</td>
</tr>
<tr>
<td>$J_2$</td>
<td>6.61</td>
<td>4.67</td>
<td>3.07</td>
<td>1.79</td>
</tr>
</tbody>
</table>

In the first strategy, harvesting is performed at age $\hat{a}$. In the second strategy, harvesting is performed at ages $\hat{a}$ and $\hat{b}$. Four examples of bell-shaped value functions $w$ are given. Mortality is the constant $\mu = 0.1$ and the fecundity schedule $b(a)$ is given in Figure 1. Problem (Q) (fixed population size) gives similar results.

pendence such as

$$x(0, t) = \Phi \left( \int_{0}^{\infty} b(a) x(a, t) da \right)$$

where $\Phi$ is some nonlinear function. However, it has been shown in other contexts [6] that ignoring density dependence is a conservative assumption. Therefore, the present linear model should not lead to dangerous strategies.

3. PROOF OF THE THEOREM

3.1 PART I.

We prove the result for problem (P). The case of (Q) is closely analogous and will only be mentioned as required. We divide the proof into two steps.

Step 1. We will show that, if $x \in X_p$ is such that $\frac{x}{T}$ is piecewise constant, then

$$J(x) \leq \lambda_p$$

where

$$\lambda_p = \max \left\{ J(x) : x \in X_p, \frac{x}{T} \text{ is piecewise constant and has at most two discontinuities} \right\}.$$
Step 2. Assume that Step 1 has been established. We can then conclude the proof of Part I.

Let \( x \in X_p \). Given \( \varepsilon > 0 \), it is easy to find \( x_\varepsilon \in X_p \) with \( \frac{x_\varepsilon}{T} \) piecewise constant and such that \( J(x) \leq J(x_\varepsilon) + \varepsilon \). Applying then Step 1 to \( x_\varepsilon \), we get

\[
J(x) \leq \lambda_p + \varepsilon.
\]

Letting \( \varepsilon \to 0 \) and taking the maximum over all \( x \in X_p \), we obtain

\[
\max(P) \leq \lambda_p.
\]

Since the other inequality \( \lambda_p \leq \max(P) \) is trivial, we have indeed established Part 1 of the theorem, provided Step 1 has been proved. The case of problem (Q) is treated identically by replacing \( X_p \) and \( \lambda_p \) with \( X_Q \) and \( \lambda_Q \).

**Proof of Step 1.** Assume that \( x \in X_p \) is piecewise constant and has \( N \) discontinuities with \( N \geq 3 \). To obtain Step 1, it is sufficient to show that we can find \( y \in X_p \), \( \frac{y}{T} \) piecewise constant with \( N - 1 \) discontinuities such that

\[
J(x) \leq J(y).
\]

To simplify the notations, we let

\[
f(a) = w'(a) - w(a) \mu(a) \tag{3.2}
\]

\[
g(a) = f(a) + w(0) b(a). \tag{3.3}
\]

Since \( x \) is continuous from the right, we let \( x(a_-) = \lim_{a \to a^-} x(a) \) be the limit from the left. We let \( 0 < \alpha < \beta < \gamma \) be the first three consecutive points of discontinuities of \( \frac{x}{T} \) (see Figure 2a). We claim that we can find \( \varepsilon, \delta \in \mathbb{R} \) such that at least one of the following inequalities is an equality

\[
\frac{x(\gamma)}{l(\gamma)} \leq \frac{x(\beta)}{l(\beta)} + \delta \leq \frac{x(\alpha)}{l(\alpha)} + \varepsilon \leq \frac{x(\alpha_-)}{l(\alpha)} = \frac{x(0)}{l(0)} = 1 \tag{3.4}
\]

and, furthermore, with \( \varepsilon, \delta \) satisfying the two following conditions:

\[
\varepsilon \int^\beta_a b(a) l(a) \, da + \delta \int^\gamma_\beta b(a) l(a) \, da = 0, \tag{3.5}
\]
FIG. 2. The plain line is the given function $\frac{x}{f}$ with $N$ discontinuities. The function $\frac{y}{f}$ has $N - 1$ discontinuities, $J(x) < J(y)$. $\frac{x}{f}$ is represented by a dotted line in the interval where it differs from $\frac{x}{f}$: (a) for $a \in [\alpha, \gamma]$ in problem (P); (b) for $a \in [0, \gamma]$ in problem (Q).
and
\[ \varepsilon \int_\alpha^\beta f(a) l(a) \, da + \delta \int_\beta^\gamma f(a) l(a) \, da \geq 0. \] (3.6)

This can always be done. For example, if
\[ \int_\alpha^\beta f(a) l(a) \, da \int_\beta^\gamma b(a) l(a) \, da - \int_\alpha^\beta b(a) l(a) \, da \int_\alpha^\gamma f(a) l(a) \, da \leq 0, \]
(3.7)
choose \( \varepsilon \) such that
\[ \varepsilon = \left[ \frac{x(\beta)}{l(\beta)} - \frac{x(\alpha)}{l(\alpha)} \right] \frac{\int_\beta^\gamma b(a) l(a) \, da}{\int_\alpha^\gamma b(a) l(a) \, da} < 0 \] (3.8)
and \( \delta \) is given by
\[ \delta = \frac{x(\alpha)}{l(\alpha)} - \frac{x(\beta)}{l(\beta)} + \varepsilon. \] (3.9)

An elementary calculation shows that these choices of \( \varepsilon \) and \( \delta \) satisfy (3.4), (3.5), and (3.6). A similar construction can be done if the sign in (3.7) is the opposite.

We then define \( y \) in the following form
\[ y(a) \frac{l(a)}{l(\alpha)} = \begin{cases} 
\frac{x(0)}{l(0)} = 1 & \text{if } a \in [0, \alpha) \\
\frac{x(a)}{l(a)} + \varepsilon = \frac{x(\alpha)}{l(\alpha)} + \varepsilon & \text{if } a \in [\alpha, \beta) \\
\frac{x(a)}{l(a)} + \delta = \frac{x(\beta)}{l(\beta)} + \delta & \text{if } a \in [\beta, \gamma) \\
\frac{x(\gamma)}{l(\gamma)} & \text{if } a \in [\gamma, +\infty) 
\end{cases} \] (3.10)
and thus, in view of (3.4), (3.5), and (3.6), \( y \) has all the claimed properties. Figure 2a gives the picture of \( \frac{x}{\gamma} \) and \( \frac{y}{\gamma} \) in case (3.7) is satisfied.
We now briefly explain how to deal with the case (Q). An analogous construction can be done. Let $0 < \alpha < \beta < \gamma$ be the first three discontinuities of $x$. We claim that one can find $\varepsilon_1, \varepsilon_2, \varepsilon_3 \in \mathbb{R}$ such that at least one of the following inequalities is an equality

$$\frac{x(\gamma)}{l(\gamma)} \leq \frac{x(\beta)}{l(\beta)} + \varepsilon_3 \leq \frac{x(\alpha)}{l(\alpha)} + \varepsilon_2 \leq \frac{x(0)}{l(0)} + \varepsilon_1$$

and

$$\varepsilon_1 \int_0^\alpha \frac{1}{l(a)} \, da + \varepsilon_2 \int_{\alpha}^\beta \frac{1}{l(a)} \, da + \varepsilon_3 \int_{\beta}^\gamma \frac{1}{l(a)} \, da = 0$$

$$\varepsilon_1 \int_0^\alpha g(a) \, l(a) \, da + \varepsilon_2 \int_{\alpha}^\beta g(a) \, l(a) \, da + \varepsilon_3 \int_{\beta}^\gamma g(a) \, l(a) \, da \geq 0,$$}

where $g$ is defined in (3.3). We then define $y$ in the following way

$$y(a) = \begin{cases} 
\frac{x(a)}{l(a)} + \varepsilon_1 = \frac{x(0)}{l(0)} + \varepsilon_1 & \text{if } a \in [0, \alpha) \\
\frac{x(a)}{l(a)} + \varepsilon_2 = \frac{x(\alpha)}{l(\alpha)} + \varepsilon_2 & \text{if } a \in [\alpha, \beta) \\
\frac{x(a)}{l(a)} + \varepsilon_3 = \frac{x(\beta)}{l(\beta)} + \varepsilon_3 & \text{if } a \in [\beta, \gamma) \\
\frac{x(a)}{l(a)} & \text{if } a \in [\gamma, +\infty).}
\end{cases}$$

Then $y$ has all the claimed properties, i.e., $J(x) \leq J(y)$, $y \in X_Q$ and $y$ has $N - 1$ discontinuities. Figure 2b gives the picture of $x$ and $y$. This achieves Part I.

### 3.2 PART II.

By Part I we may assume that there exists $a$ and $b$ such that a solution of (P) or of (Q) is of the form

$$\frac{x(\alpha)}{l(\alpha)} = \begin{cases} 
x(0) & \text{if } 0 \leq \alpha < a \\
0 & \text{if } a \leq \alpha < b \\
0 & \text{if } \alpha > b.
\end{cases}$$

For such a function \( x \), \( J \) takes the following form:

\[
J(x) = x(0) \left[ w(0) + \int_a^b \left( w'(\alpha) - \mu(\alpha) w(\alpha) \right) l(\alpha) \, d\alpha \\
+ \theta \int_a^b \left( w'(\alpha) - \mu(\alpha) w(\alpha) \right) l(\alpha) \, d\alpha \right].
\]

Observing that \( l'(\alpha) = -\mu(\alpha) l(\alpha) \) and that \( l(0) = 1 \), we immediately get that

\[
J(x) = x(0) \left[ w(0) + w(a) l(a) - w(0) + \theta \left( w(b) l(b) - w(a) l(a) \right) \right] = x(0) \left[ \theta w(b) l(b) + (1 - \theta) w(a) l(a) \right].
\]

To obtain the value of \( \theta \), we use the fact that

\[
\int x(\alpha) b(\alpha) d\alpha = x(0),
\]

which in our case gives [simplifying by \( x(0) \)]

\[
\int_a^b b(\alpha) l(\alpha) \, d\alpha + \theta \int_a^b b(\alpha) l(\alpha) \, d\alpha = 1.
\]

Bearing in mind (2.10), we get immediately that

\[
\theta = \frac{\int_a^b b(\alpha) l(\alpha) \, d\alpha}{\int_a^b b(\alpha) l(\alpha) \, d\alpha}.
\]

Combining (3.11) and (3.12), we deduce that

\[
J(x) = x(0) \left\{ \frac{\left( \int_a^b b(\alpha) l(\alpha) \, d\alpha \right) w(b) l(b)}{\int_a^b b(\alpha) l(\alpha) \, d\alpha} + \frac{\left( \int_a^b b(\alpha) l(\alpha) \, d\alpha \right) w(a) l(a)}{\int_a^b b(\alpha) l(\alpha) \, d\alpha} \right\}
\]

\[
= x(0) J(a, b).
\]
The conclusion for (P) follows at once from the fact that $x(0) = 1$. In the case (Q) we have still to use the fact that

$$\int_0^\infty x(\alpha) \, d\alpha = 1,$$

i.e.,

$$x(0) \left( \int_0^a l(\alpha) \, d\alpha + \theta \int_a^b l(\alpha) \, d\alpha \right) = 1,$$

which leads to

$$x(0) = \frac{\int_a^b b(\alpha) l(\alpha) \, d\alpha}{\int_0^a l(\alpha) \, d\alpha \int_a^b b(\alpha) l(\alpha) \, d\alpha + \int_a^b b(\alpha) l(\alpha) \, d\alpha \int_a^b l(\alpha) \, d\alpha}.$$

(3.14)

The combination of (3.13) and (3.14) leads to the claimed result. ■

We thank François Weissbaum for performing the numerical calculations of the examples; Olivier Glaizot and Henni Saïah for helping with the preparation of the manuscript and the figures; and Yuri Dombrovsky, Wayne Getz, and Lev Ginzburg for interesting comments.

REFERENCES
