Ca' Foscari University of Venice

Department of Economics

Working Paper

G. Fabbri, S. Faggian, and G. Freni

On the Mitra-Wan Forest Management Problem in Continuous Time

ISSN: 1827-3580
No. 28/WP/2013

Working Papers
Department of Economics
Ca' Foscari University of Venice
No. 28/WP/2013
ISSN 1827-3580

On the Mitra-Wan Forest Management Problem in Continuous Time

G. Fabbri[1], S. Faggian[2], G. Freni[3]

(1) *EPEE, Universite d'Evry-Val-d'Essonne (TEPP, FR-CNRS 3126), Departement d'Economie, 4 bd. Francois Mitterrand, 91025 Evry cedex, France. E-mail:* giorgio.fabbri@uniparthenope.it.
(2) *Department of Economics, University "Ca' Foscari" Venice, Italy. E-mail:* faggian@unive.it.
(3) *Department of Business and Economics, University of Naples Parthenope, Naples, Italy. E-mail:*giuseppe.freni@uniparthenope.it.

18th November 2013

Abstract

The paper provides a continuous time version of the well known discrete time Mitra-Wan model of optimal forest management, where a forest is harvested to maximize the utility of timber flow over an infinite time horizon. Besides varying with time, the state variable (describing available trees) and the other parameters of the problem vary continuously also with respect to the age of the trees. The evolution of the system is given in terms of a partial differential equation and later rephrased as an ordinary differential equation in an infinite dimensional space. The paper provides a classification of the behavior of optimal and maximal programs when the utility function is linear, convex, or strictly convex and the discount rate is positive or null. Formulas are provided for modified golden-rule configurations (uniform density functions with cutting at the ages that solve a Faustmann problem) and for *Faustmann policies*, and the optimality or maximality of such programs is discussed. In all different sets of data, it is shown that the optimal (or maximal) control is necessarily something more general than a function, i.e. a positive measure. In particular, in the case of strictly concave utility and null discount, when the Faustmann policy is not optimal, it is shown that optimal paths converges over time to the golden rule configuration, while in the case of strictly concave utility and positive discount the Faustmann policy is shown to be not optimal, contradicting the corresponding result in discrete time.

Keywords: Optimal harvesting problems, Forest Management, Measure-valued Control

JEL Codes: C61, C62, E22, D90, Q23.

Address for correspondence:
Silvia Faggian
Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta S.Giobbe
30121 Venezia - Italy
Phone: (++39) 041 2346913
e-mail: faggian@unive.it

The Working Paper Series
is available only on line
(http://www.unive.it/nqcontent.cfm?a_id=86302)
For editorial correspondence, please contact:
wp.dse@unive.it

Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta San Giobbe
30121 Venice Italy
Fax: ++39 041 2349210

On the Mitra–Wan Forest Management Problem in Continuous Time

G. Fabbri[*], S. Faggian[†] G. Freni[‡]

November 20, 2013

Abstract

The paper provides a continuous time version of the well known discrete time Mitra-Wan model of optimal forest management, where a forest is harvested to maximize the utility of timber flow over an infinite time horizon. Besides varying with time, the state variable (describing available trees) and the other parameters of the problem vary continuously also with respect to the age of the trees. The evolution of the system is given in terms of a partial differential equation and later rephrased as an ordinary differential equation in an infinite dimensional space. The paper provides a classification of the behavior of optimal and maximal programs when the utility function is linear, convex, or strictly convex and the discount rate is positive or null. Formulas are provided for modified golden-rule configurations (uniform density functions with cutting at the ages that solve a Faustmann problem) and for *Faustmann policies*, and the optimality or maximality of such programs is discussed. In all different sets of data, it is shown that the optimal (or maximal) control is necessarily something more general than a function, i.e. a positive measure. In particular, in the case of strictly concave utility and null discount, when the Faustmann policy is not optimal, it is shown that optimal paths converges over time to the golden rule configuration, while in the case of strictly concave utility and positive discount the Faustmann policy is shown to be not optimal, contradicting the corresponding result in discrete time.

Keywords: Optimal harvesting problems, Forest Management, Measure-valued Control.

JEL Classification: C61, C62, E22, D90, Q23.

1 Introduction

Although forest economics has a centuries-long history (see, *e.g.*, Samuelson (1995)), the first complete formulation of the forest management problem as a Ramsey-like optimal control model

[*]EPEE, Université d'Evry-Val-d'Essonne (TEPP, FR-CNRS 3126), Département d'Economie, 4 bd. François Mitterrand, 91025 Evry cedex, France. E-mail: giorgio.fabbri@uniparthenope.it. The work of Giorgio Fabbri was partially supported by the *Post-Doc Research Grant* of *Unicredit & Universities* and his research has been developed in the framework of the center of excellence LABEX MME-DII (ANR-11-LABX-0023-01)

[†]Department of Economics, University "*Ca' Foscari*" Venice, Italy. E-mail: `faggian@unive.it`.

[‡]Department of Business and Economics, University of Naples *Parthenope*, Naples, Italy. E-mail: `giuseppe.freni@uniparthenope.it`.

in discrete time is contained in two papers by T. Mitra and H. Wan of the early eighties of the last century (Mitra and Wan (1985, 1986)). In these two contributions, the focus is on the long run structure of the cutting/replanting strategy that maximizes over an infinite horizon the sum of utilities generated by the flows of timber that are obtained harvesting the trees of a forest. In the basic model, the forest comprises trees of different ages, up to a maximum age, that are cultivated on a unit piece of land that cannot be transferred to other uses, a productivity function gives the amount of wood that is obtained harvesting a tree of any given age, cutting and replanting costs are zero, and new saplings on the cleared land are the joint product of cutting. The main results are: 1) that the *Faustmann policy* (*i.e.*, cutting any tree that reach a critical age that maximizes the present value of bare forest land subject to an infinite sequence of planting cycles) is optimal when the utility function is linear, which implies a cycle in the configuration of the forest, 2) that optimal trajectories converge to the golden rule configuration (the uniform forest with the maximum sustainable yield) when the utility function is strictly concave and the discount factor is equal to 1, and 3) that cycles of the optimal trajectory reappear whenever future utility is discounted even if the utility function is strictly concave. Following this lead, almost the entire theory of optimal forest management has been developed in terms of discrete time (see Tahvonen (2004) and Khan and Piazza (2012) for recent lists of the extensions of the model) while, to our knowledge, a consistent continuous time version of the Mitra-Wan optimal forest management problem has never entered the literature.

The two usual justifications for the choice of discrete time – that, (i) transactions do not occur continuously, and that (ii) sometimes a "natural period" can be found for agricultural products (Foley (1975)) – have some appeal in forestry, but on balance there seem not to be compelling reasons to assume that transactions in timber markets are synchronized. Moreover, many forestry tasks (*e.g.*, tree felling, timber extraction, etc.) require time and it is likely therefore that these operations overlap not only for different agents but also for a single forestry firm. In addition, as it is the case with other natural resources, forests grow continuously and a single natural period for cutting and replanting activities can hardly be identified in general (*e.g.* at the tropics forests are cut continuously). Hence, no specific economic reason dictates that the discrete time should be preferred to the continuous time framework. If so, it is the technical complexity that continuous time introduces into the analysis of the model that explains the lack of contributions on continuous time control of a forest on a given area. Indeed, in elaborating Wicksell's classical model of natural aging process, Cass (1973) had already noted that formulating the model in continuous time leads to difficulties in the derivation of efficiency prices.

When the discrete time Mitra-Wan model is reformulated in continuous time, as it occurs here and in other vintage capital models, some difficulties arise: 1) the ages of capital goods (*i.e.* the ages of trees) vary continuously, so that the evolution in time of the state of the system is described by means of a partial differential equation; 2) the control appears also in the boundary condition (that is, we obtain a *boundary control system*); 3) candidate optimal controls are not functions (even if we allow almost everywhere defined measurable functions) but distributions, more precisely *measures*. We underline the fact that – contrary to what happens, as far as we know, in other vintage capital models – the control appearing in the evolutionary equation (the so called distributed control) *need* be a measure as a consequence of 3), so that the equation itself cannot be interpreted in $\mathbb{R}$ as usual (*i.e.*, pointwise) but calls for an extended formulation in order to make sense. All these features make indeed the problem mathematically challenging. Moreover, it is worth noticing that the sum of these technical difficulties lead to a problem that, differently from that in discrete time, is not a straightforward multi-sectoral generalization of the Ramsey model (for a similar conclusion see Khan and Piazza (2011)).

There have been various attempts to gain insight into the solution of the continuous time Mitra-Wan model either by studying stripped down versions of the full model or by adding further assumptions that somehow help simplify the analysis. Heaps (2006) (see also Heaps (1984)), for example, presented a model in which only the oldest trees can be harvested and proved that in his model the optimality conditions take the form of a delay differential equation [1]. Tahvonen and Salo (1999) (see also Tahvonen et al. (2001)), on the other hand, studied a model in which time is continuous, but trees are indivisible. This implies that the number of trees that can be grown on a given land is finite, but also that harvesting of the forest cannot provide a continuous flow of timber but only a sequence of mass points in connection with the jumps in the state variables. In Tahvonen and Salo (1999) storage of wood is then allowed so that the planner can smooth consumption during the time elapsed between two successive jumps. Finally, Salant (2013) analyzed the equilibrium price paths of different vintages of trees, in a simple model in which the forest land may be used in an alternative way and replanting is not allowed. In Salant (2013) the assumption that deforestation is irreversible allows to study optimal continuous harvesting/wood consumption without entering the complexities of distributed state variables.

In this paper we develop a new approach to handle the continuous time Mitra-Wan model that consists in reformulating the control problem for the partial differential equation as an equivalent problem for a ordinary differential equation but set in an infinite dimensional space, and in developing *ad hoc* techniques to perform the analysis. In our formulation we need neither to reduce the dimensionality of the problem (as in Tahvonen and Salo (1999) or Salant (2013)), nor to constrain the controls (as in Heaps (1984)).

The main purpose of developing a continuous time analog of the Mitra-Wan forestry model, and of giving jointly a classification of the behavior of optimal and maximal programs in the cases in which the utility is linear, convex, or strictly convex and the discount rate is positive or null, is to compare the properties of discrete-time optimal paths with the properties that emerge in the continuous-time framework. It turns out that there are significant differences. In particular in the present paper it is shown that:

a) the analog of golden-rule and *modified* golden-rule configurations is available for the continuous time model;

b) modified golden rules are optimal stationary solutions for the discounted model (with optimal cutting age M and timber stationary consumption level monotonically not increasing in the rate of discount), while in the undiscounted case they are maximal when the utility function is linear, and optimal when the utility is strictly concave (and we assume a unique golden rule configuration);

[1] In Heaps (2006) it is claimed that cutting only the oldest trees is a property of the optimal policy function of the full fledged model. However, in Heaps (2006) there is no formal proof of this result. For the discrete time two-age-classes model it is known that cutting the old trees *and* part of the young trees is optimal if the share of land on which old trees are planted is below a given threshold (see for example, Tahvonen (2004)). On the other hand, for the continuous time model discussed in this paper, constraining the control as suggested by Heaps would make unfeasible the shrinking in time of the length of the support of the state variable and thus a convergence result as that in Theorem 5.9 below would not hold.

c) if the golden-rule configuration is unique, then undiscounted maximal (or optimal) paths exist from any given initial configuration and, provided the utility function is strictly concave, converge over time to the golden rule configuration;

d) the *Faustmann policy* is optimal when the utility is linear and the discount positive, is maximal (and not optimal) when the utility is linear and the discount null, it is not optimal when the utility is strictly concave and the discount positive, for initial data in any neighborhood of the optimal steady state. In particular this result contradicts the analog in discrete time.

In the continuous time model we develop, although strategies are allowed to be measures, with the consequence that instantaneous cutting for the forest of any given age is allowed (the golden rule configuration is indeed measure-valued), the associated trajectories may be not, so to avoid mass points. To this extent, it is enough to consider initial distributions of the forest which are square integrable measurable functions and prove, as we do, that the whole trajectory preserves the same property.

Once the stage is set, we are able to extend to our continuous time model some of the classical results of the theory of optimal economic growth. Indeed, what we effectively show for the undiscounted case (see points c) and d)) is that Brock's (Brock (1970)) conclusions on the existence and "average" convergence of maximal paths hold in our continuous-time control framework and that the results can be strengthened to existence and asymptotic convergence of optimal paths as in Gale (1967) if the utility function is strictly concave ((Mitra and Wan (1986)), and (Khan and Piazza (2010)), had already shown that the same holds in discrete time). We in addition refine the above results by providing an example in the style of Brock (1970) and Peleg (1973), whose elaboration proves that optimal paths do not exist in the linear case.

Comparative statics results under b) are specific to the continuous time setting (hints for such results are in Samuelson (1995) see Figure 2 on pag. 133 and the first paragraph on pag. 134) and have not counterpart in the discrete time forestry literature. It is interesting to note that monotonicity of the modified golden rule consumption may not hold in models with several capital goods as the present one, while holds in the one sector Ramsey discounted model (see for example Mas-Colell et al. (1995) pp. 758-9). Regarding the *Faustmann policy* (point d)) we prove that, similarly to what occurs in the discrete time setting (see Mitra and Wan (1985), discounting does not affect the structure of the optimal policy when the utility function is linear. However, when the utility function is strictly concave we show that the periodic optimal solutions that is the distinguishing mark of the discounted discrete-time model (Mitra and Wan (1985) p.265, Salo and Tahvonen (2003) Proposition 1) is not optimal in our continuous-time framework.

We already mentioned that the mathematics underneath the problem is fine and challenging, due to the presence of the control in the boundary condition and to the fact that also the distributed control need be a distribution. We remark that, while the literature on distributed control systems is wide and well established also under the theoretical point of view, that on boundary control systems is not as much, even when described by a simple linear age structured equations as that in our problem. Indeed the presence of the control in the boundary condition yields discontinuities in the control operator and in the Hamiltonians which are difficult to handle. The techniques of functional analysis of rephrasing the problem into an abstract ODE in an infinite dimensional space of functions, and later suitably expand it to a space of distributions (see Section 3) in order to deal with discontinuities is performed in some theoretical and applied works. These techniques were first introduced in the economic literature by Barucci and Gozzi (1998, 2001), for a problem of optimal investment with *vintage capitals*, and then studied in various works, both under the theoretical point of view of Dynamic Programming (Faggian (2005, 2008), with finite horizon; Faggian and Gozzi (2010), with infinite horizon) and that of applications (Feichtinger et al. (2003, 2006); Faggian and Gozzi (2004)). We mention also the papers by Barucci and Gozzi (1999) and Faggian and Grosset (2013) with application of such techniques to problems of optimal advertising.

Nevertheless in none of these works the control space need to be a space of distributions (more often is the space of square integrable functions), thus the fact that our stationary optimal control is a distribution constitutes a new development in the literature. We strongly expect that some of the innovations we introduce will find application in vintage capital models and in other models comprising age distributed state variables (*e.g.*, demographic models). For example, having controls that are (not function but) positive measures may prove useful both in the analysis of the problem of endogenous scrapping of a machine and in the study of the investment profile (i. e., whether the investment is spread over a set of vintages or not, see Feichtinger et al. (2006)). Similarly, having controls that can be concentrated on single ages may help in developing theoretical models of the determinants of the optimal retirement age for not stationary populations endowed with a rich demographic structure (Chan and Zhu (1990), Heijdra and Romp (2009)). We also expect that our results and methods find useful applications in the analysis of other models with age distributed natural resources (for example, see Tahvonen (2009) for age distributed fisheries and Mitra et al. (1991) for the so called orchard model).

Our results will have a more direct bearing on the analysis of the continuous-time Ramsey version of the clay-clay vintage capital model of Solow, Tobin et al. (Solow et al. (1966)) developed by Boucekkine et al. (1997) (see also Boucekkine et al. (1998), Hritonenko and Yatsenko (2008)). In this model, indeed, a Faustmann-like age shows up as the steady state age at which old capital goods became obsolete and a Faustmann-like policy can be defined as the policy that replaces all stocks of capital goods that reach the critical age with the constraint that the new investment preserves full employment of labor. We expect advances on at least two fronts. First, as mentioned above, our control space will allow the handling of reversible investment, that is beyond the reach of current theory. Moreover, our strong convergence result for the strictly concave case, while providing a formal proof of the fact that the cycles induced by replacement echoes are dampened in the long run by the force of consumption smoothing only for the case in which the golden rule is optimal, it nevertheless should pave the way to the general turnpike result for the discounted model that the literature envisages (see Boucekkine et al. (1998) and Boucekkine et al. (2011)) (in the Ramsey vintage capital model with labor augmenting technical progress an exact golden rule path exists if the utility function is isoelastic. It is optimal - or maximal - when the discount rate is equal to the Mirrlees-Brock-Gale critical discount rate). In connection with this, since we have already established that the optimal cyclical paths that arises in the discrete-time discounted forestry model (with the radii of the cycles converging to zero with the discount rate (Tahvonen et al. (2001), Salo and Tahvonen (2003), Dasgupta and Mitra (2011))) is non optimal in continuous time, and since generically the optimal stationary state of the continuous-time model is unique, a follow up of our results should be an asymptotic turnpike result for the strictly concave continuous-time discounted Mitra-Wan model.

Finally it would be extremely interesting to develop formal necessary conditions in the form of the Maximum Principle, or perform Dynamic Programming to provide a finer characterization of the optimal solutions, in particular for the discounted utility setting for which the asymptotic properties are still unknown, and to characterize, at least in some interesting cases, the optimal policy function (Dasgupta and Mitra (2011)).

The paper is organized as follows. In section 2 we describe the model in continuous time, in section 3 we rephrase it into into abstract terms and introduce some useful notation, besides the formal definition of optimal and maximal strategies. In section 4 first we build the modified golden rules and Faustmann Policies. In section 5, we classify in four different subsections the behavior of optimal trajectories, according to the fact that the utility is linear, strictly concave (or general concave, in some cases), and that the discount rate is positive or null. We contextually establish whether Golden Rules and Faustmann policies constitute optimal programs, maximal programs, or none of the two. In section 6 we draw the conclusions. A rich Appendix, following the references section, completes the work with the proofs of the many theorem and all auxiliary technical results.

2 The continuous time model

A forest of unitary extension is described in terms of a density $x(t,s)$ which represents the part of the forest covered at time t by trees of a certain age s, with $t \geq 0$ and $s \geq 0$, and trees reaching a maximum (finite) age S. Starting from an initial density $x_0(s)$, trees grow in time and may be harvested with a certain cutting rate $c(t,s)$, performed at time t on trees of age s, and chosen by the optimizer. Harvested trees are instantaneously replaced by new saplings. The evolution of the system is then described by the following transport equation

$$\begin{cases} \frac{\partial x}{\partial t}(t,s) = -\frac{\partial x}{\partial s}(t,s) - c(t,s) & t > 0, \quad 0 < s \leq S \\ x(t,0) = \int_0^S c(t,s)\,\mathrm{d}s & t > 0 \\ x(0,s) = x_0(s) & 0 \leq s \leq S \end{cases} \tag{1}$$

where the variation of density $\frac{\partial x}{\partial t}(t,s)$ is due to ageing of trees $-\frac{\partial x}{\partial s}(t,s)$, and to harvesting $-c(t,s)$. Note also that in the boundary condition the quantity $x(t,0)$ of saplings of age zero at time t is assumed to coincide with the total amount of trees of different ages cut at time t, represented by $\int_0^S c(t,s)\mathrm{d}s$. [2]

In addition we require the strategy-trajectory couples (c,x) to satisfy some non negativity constraints, that is

$$c(t,s) \geq 0, \text{ and } x(t,s) \geq 0, \ \forall t \geq 0, \ 0 \leq s \leq S.$$

implying that only positive quantities are cut, and that the quantity of trees of all ages remains positive in time.

Some remarks are here due. First of all, note that $x(t,s)$ does not represent a spatial density. As a consequence, it may be imagined that trees grow far from one another, and not reciprocally interfering. Moreover, since the size of the forest is normalized to 1 at initial time, that is

[2] A dimensional interpretation of the boundary condition is the following. The quantity $x(t,0)ds$ represents the infinitesimal portion occupied at time t by trees of age 0, which has to be equal to the amount of trees (of all ages) cut at time t, represented by the quantity $\left[\int_0^S c(t,s)\mathrm{d}s\right] dt$. The condition then follows by observing that time and age vary jointly, that is $dt/ds = 1$.

$\int_0^S x_0(s)\,\mathrm{d}s = 1$, $x(t,s)$ may be interpreted as the percentage of the total forest which is covered at time t by trees of a certain age s. Thus it is expected that, as a consequence of the boundary condition, the total surface of the forest is covered in time by a constant amount (necessarily equal to 1) of trees of different ages, that is (see Proposition 3.4)

$$\int_0^S x(t,s)\,\mathrm{d}s \equiv \int_0^S x_0(s)\,\mathrm{d}s = 1, \quad \forall t \geq 0.$$

Now let $f(s)$ represent the *productivity* of a tree of age s. Summing all the contributions $f(s)c(t,s)ds$ of wood of different ages s harvested at time t, we obtain the total wood $w(c(t))$ harvested at time t, that is

$$w(c(t)) = \int_0^S f(s)c(t,s)ds. \tag{2}$$

Note that f is the continuous version of the timber-content function f defined by Mitra and Wan in (Mitra and Wan (1985, 1986)). We assume also f a regular function with support contained in $[0,\bar{s}]$, meaning in particular that trees of age greater than $\bar{s}$ are considered unproductive. Note also that it would make a little difference to assume trees eternal (that is, $S = \infty$), since the action takes place in $[0,\bar{s}]$ as a consequence of the choice of the productivity function.

Example 2.1 An example of a regular productivity function f is the following.

$$f(s) = \begin{cases} \exp\left[\left(x - \frac{M}{2}\right)^{-1}\left(x - \frac{3M}{2}\right)^{-1}\right] & s \in (M/2, 3M/2) \\ 0 & s \in [0, M/2] \cup [3M/2, +\infty) \end{cases}$$

Note that trees younger than $M/2$ and older than $3M/2$ are considered unproductive, while the productivity increases towards M and decreases afterwards. □

We anticipate that in the next section, where the problem is formalized, the analysis is restricted to a set of admissible controls $c(t,s)$ which are null and leave trajectories $x(t,s)$ null, for all $s \geq \bar{s}$. To simplify the mathematical work, we take into account initial data $x_0(s)$ which are null for all $s \geq \bar{s}$. Those assumptions are justified in view of the fact that the productivity f is null for $s \geq \bar{s}$, as one expects that an optimal trajectory yields null quantities of trees older than $\bar{s}$.

In addition, we introduce a *utility function* u, which is assumed bounded below, increasing and concave (possibly linear)[3], and an overall utility $U(c)$ defined in terms of u as

$$U(c) = \int_0^{+\infty} e^{-\rho t} u\left(w(c(t))\right)\,\mathrm{d}t.$$

The problem is maximizing in a suitable sense the overall utility $U(c)$, over the set of admissible strategies, with or without discount (that is, when $\rho > 0$ or $\rho = 0$, respectively). Note that when $\rho > 0$ the concavity of u implies the finiteness of $U(c)$, while when $\rho = 0$, $U(c)$ may indeed be infinite valued, and that need be taken into account when choosing a suitable definition of optimality. For U_T denoting the overall utility at a finite horizon T, that is

$$U_T(c) = \int_0^T e^{-\rho t} u\left(w(c(t))\right)\,\mathrm{d}t,$$

[3]For instance, $u(r)$ may be the identity function, or $\ln(r+1)$, or $r^{1-\sigma}$, with $0 < \sigma < 1$.

and for x_0 some initial stock, an admissible control strategy c^* is said to be *optimal at* x_0 if

$$\liminf_{T\to\infty} \big(U_T(c^*) - U_T(c)\big) \geq 0,$$

for every control strategy c, admissible at the same initial stock x_0. Note that the definition of optimality implies the property

$$\forall n \in \mathbb{N}, \exists T_n > 0 : \forall T \geq T_n \Rightarrow U_T(c^*) > U_T(c) - \frac{1}{n}.$$

saying that the control c^* yields definitively (namely, for a sufficiently large horizon T) a greater utility than any other control c starting at the same x_0 except for a (small) difference $\frac{1}{n}$. This idea of optimality is often referred to in literature as *overtaking* optimality, and the optimal control c^* is said to be (definitively) overtaking any other control c.

If the utility function fails to be strictly concave, then optimality defined this way will prove a too strong requirement, meaning that in some cases under study no control matching such definition will be available, and a weaker optimality property will have to be taken into account. Then, we say that an admissible control strategy is *maximal at* x_0 if

$$\limsup_{T\to\infty} \big(U_T(c^*) - U_T(c)\big) \geq 0$$

for every control strategy c admissible at the same initial stock x_0. Note that optimality of a control implies maximality, but the viceversa is false in general. An implication of the definition of maximality is that, given any other control c admissible at x_0 one has

$$\forall n \in \mathbb{N}, \exists T_n > 0 : U_{T_n}(c^*) > U_{T_n}(c) - \frac{1}{n}.$$

meaning that the control c^* is repeatedly overtaking c along a sequence of horizons $\{T_n\}$ or, from a reversed point of view, that no control c overtakes the control c^* definitively. Maximality indeed proves a more flexible idea of optimality, as it allows for controls engendering fluctuating behaviors of the overall utility U_T in time.

Comparison of non convergent infinite horizon integrals (or sums) occurs in the optimal growth literature since Ramsey (1928). As references on optimal/maximal controls, we recall that the notion of optimal (overtaking) controls was formalized by Von Weizsäcker (1965) in continuous time to study the existence of optimal programs in the aggregative model of growth, while Gale (1967) and McKenzie (2009) extensively studied optimal overtaking paths for the multisector optimal growth model in discrete time. Maximality of controls in the acceptation cited above is first found in Brock (1970) where existence of maximal programs for the n sectors optimal growth model in discrete time is established, while Halkin (1974) adapted the concept to continuous time and an infinite horizon. Extensions in continuous time of the results in Brock (1970) are also due to Brock and Haurie (1976), Carlson et al. (1987), and Zaslawski (2006). We should warn the reader that almost in every paper a different terminology is used (see Khan and Piazza (2012), f.n. 17) and that we follow the terminology used in McKenzie (1986), p.1286.

A final remark on the mathematics of the problem. The challenging issues go beyond the presence of the control in the boundary condition. Here is an example: the state equation has a solution which can be written easily by means of the characteristic method, as long as the control is an integrable function, given by

$$x(t,s) = \begin{cases} x_0\,(s-t) - \int_0^t c(t-\tau, s-\tau)\mathrm{d}\tau & s \geq t \\ \int_0^{\bar{s}} c(t-s, r)\mathrm{d}r - \int_0^s c(t-\tau, s-\tau)\,\mathrm{d}\tau & 0 \leq s < t. \end{cases} \tag{3}$$

Unfortunately, the space of admissible controls cannot be a space of functions, but need be a larger space – a space of measures, where optimal controls are shaped like Dirac's Deltas – in which a formula like the one above is not defined. Or needs to be redefined, as we explain in the next section.

3 The abstract problem

We here rephrase the problem in section 2 by means of semigroup theory, and formalize the assumptions accordingly. The original problem for the partial differential equation is rephrased as a problem for an ordinary differential equation, set in an infinite dimensional space. Roughly speaking, that means that rather than considering the state and the control as real functions of t and s, one sees them as functions of the only variable t, taking values in some space of functions of variable s (for instance the space of square integrable real functions), and $x(t)$ and $c(t)$ are interpreted as the functions of variable s defined by $x(t)[s] \equiv x(t,s)$ and $c(t)[s] \equiv c(t,s)$. The initial condition will be written as $x(0) = x_0$, the state equation as $x'(t) = Ax(t) + Bc(t)$ where x' indicates the derivative with respect to time, the differential operator $A = -\frac{\partial}{\partial s}$ is a linear operator between space of functions, while the control operator B represents the joint action of the control appearing into the PDE, and of the boundary condition, namely $Bc(t)[s] = -c(t)[s] + \left(\int_0^S c(t)[s]ds\right)\delta_0$ (where δ_0 is the Dirac's delta at 0).

We mention that usually the state/control spaces for such problems is $L^2(0,S)$, the space of real functions of the variable s which are square integrable in $[0,S]$, considered as a Hilbert space by means of the scalar product

$$\langle \phi, \psi \rangle_{L^2} = \int_0^S \phi(s)\psi(s)\,\mathrm{d}s, \quad \phi, \psi \in L^2(0,S)$$

while here we need to extend the problem to a larger space D' (the dual of some D space with $D \subset L^2 \subset D'$) which contains objects like the Dirac's deltas, and where the scalar product is replaced by the more general duality pairing

$$\langle \phi, \psi \rangle_{D',D}, \quad \phi \in D', \ \psi \in D$$

coinciding with the scalar product whenever $\phi \in L^2(0,S)$. It is very important to say that such formulation enables the possibility of choosing controls which are something more general than a function, that is, *positive measures*. That turns out to be fundamental, as optimal controls will be shaped like Dirac Delta's, to begin with the optimal control associated to stationary solutions (the so called *golden rule*, see section 4). We advise the reader that a full understanding of the following subsection 3.1 requires some familiarity with semigroup theory, and may be largely skipped at a first reading. For the general theory of strongly continuous semigroups we refer the reader to Engel and Nagel Engel and Nagel (1999) or Pazy (1983). For optimal control probem in infinite dimensional spaces, the reader is referred to Bensoussan et al. (2007)

3.1 The extended framework

We introduce some useful notation. Let X be a Banach space, X' its dual space, we denote with $\langle\cdot,\cdot\rangle_{X,X'}$ or simply by $\langle\cdot,\cdot\rangle$ the duality pairing. If $-\infty \leq \sigma_1 < \sigma_2 \leq \infty$, we denote with $L^p(\sigma_1, \sigma_2; X)$ the space functions of integrable p–norm defined in $[\sigma_1, \sigma_2]$ and taking values in X, or simply $L^2(0,\sigma)$ when $X = \mathbb{R}$. In particular we write $H^1(\sigma_1, \sigma_2)$ for the space of functions of $L^2(\sigma_1, \sigma_2)$ with (weak) derivative in $L^2(\sigma_1, \sigma_2)$, and $H^{-1}(\sigma_1, \sigma_2)$ for the dual of $H^1(\sigma_1, \sigma_2)$. We also denote with $L^2_{loc}(\sigma_1, \sigma_2; X)$ the set of functions from $[\sigma_1, \sigma_2]$ to X which are square integrable on every compact interval contained in $[\sigma_1, \sigma_2]$. With $C^k([\sigma_1, \sigma_2]; X)$ we denote the set of functions of class C^k from $[\sigma_1, \sigma_2]$ to X. Moreover we indicate with $\mathcal{L}(X, Y)$ ($\mathcal{L}(X)$, when $X = Y$) the space of continuous linear functions between the Banach spaces X and Y, with associated norm $|\cdot|_{\mathcal{L}(X,Y)}$ (respectively, $|\cdot|_{\mathcal{L}(X)}$).

Finally we denote by $\mathcal{R}$ the space of all Radon measures (positive and finite) on $[0, S]$. Given a measure $c \in \mathcal{R}$ we denote with $|\cdot|_R$ the mass of c (*i.e.* the measure of the set [0,S] with respect to c). One may show by Riesz Theorem (*e.g.* Rudin (1987) Theorem 2.14 page 40) that c can be represented as a positive linear functional on $C_0([0, S])$ and that $|c|_{\mathcal{R}} = \sup_{\phi \in C^0} \langle c, \phi\rangle / |\phi|_{C^0}$. Since $D \subseteq C_0([0, S])$ with continuous embedding then $\mathcal{R} \subseteq D'$ and, for any $c \in \mathcal{R}$, $|c|_{\mathcal{R}} \leq |c|_{D'}$.

We start by considering the space $L^2(0, S)$ and the definition of the translation semigroup on $L^2(0, S)$, $T(t) : L^2(0, S) \to L^2(0, S)$, $f \mapsto T(t)f$ where $T(t)f[s] = f(s - t)$, if $s \in [t, S]$, and $T(t)f[s] = 0$ otherwise. The generator of $T(t)$ is the operator $A : D(A) \to L^2(0, S)$ where $D(A) = \{f \in H^1(0, S) \ : \ f(0) = 0\}$, given by $A(f)[s] = -\frac{\partial f}{\partial s}(s)$. The adjoint of A^* is then $A^* : D(A^*) \to L^2(0, S)$ with $D(A^*) = \{f \in H^1(0, S) \ : \ f(S) = 0\}$ defined by $A^* f[s] = \frac{\partial f}{\partial s}(s)$, generating itself a translation semigroup $T^*(t) : L^2(0, S) \to L^2(0, S)$ given by $T^*(t)f[s] = f(s + t)$, if $s \in [0, S - t]$, and $T^*(t)f[s] = 0$ otherwise. Set

$$D \equiv D(A^*), \quad D' \equiv D(A^*)'.$$

By standard arguments $T(t)$ can be extended to a strongly continuous semigroup on D' and (with respect to the operator norm on D'), while $T^*(t)$ can be restricted to a strongly continuous semigroup on D. The generators of such semigroups are respectively an extension and a restriction of the ones in $L^2(0, S)$. For simplicity we keep denoting the semigroups and their generators by $T(t)$ and A, and by $T^*(t)$ and A^* respectively. For details we refer the reader to Faggian (2005); Faggian and Gozzi (2010).

We take D' both as control space and state space of the abstract problem.

Note that D' is a space of distributions and it contains the set $\mathcal{R}$ of all Radon measure on $[0, \bar{s}]$, and in particular all Dirac's measures δ_{s_0}, $s_0 \in [0, \bar{s}]$.
We define the cutoff function $\psi \in C^\infty([0, S]; \mathbb{R}^+)$ such that for fixed σ_1, σ_2, with $\bar{s} < s_1 < s_2 < S$, $\psi \equiv 1$ on $[0, s_1]$, $\psi \equiv 0$ on $[s_2, S]$, ψ decreasing on $[s_1, s_2]$. [4] Observe that $\psi \in D$. Then we may define the linear continuous functional[5] $L\colon\ D' \to \mathbb{R}$, $\quad c \mapsto Lc := \langle c, \psi\rangle$, express the boundary condition as $x(t, 0) = Lc(t) = \langle c(t), \psi\rangle$ and, by standard techniques, include Lc in the definition of the control operator B

$$B\colon D' \to D', \quad Bc := -c + (Lc)\,\delta_0 = -c + \langle c(t), \psi\rangle \delta_0 \tag{4}$$

[4] For instance, $\psi \equiv 1$ on $\left[0, \bar{s} + \frac{S - \bar{s}}{4}\right]$, and ψ decreasing on $\left[\bar{s} + \frac{S-\bar{s}}{4}, \bar{s} + \frac{3}{4}(S - \bar{s})\right]$.

[5] Note that whenever $c(t, \cdot)$ is a function in L^2, then $\langle c(t, \cdot), \psi\rangle = \int_0^S c(t, s)ds$ is the initial boundary condition.

where δ_0 is the Dirac's delta at 0. Then (1) can be written as

$$\begin{cases} x'(t) = Ax(t) + Bc(t), & t > 0 \\ x(0) = x_0 \end{cases} \tag{5}$$

and rewritten in the mild form by means of the variation of constants formula (see e.g. Faggian (2005))

$$x(t) = T(t)x_0 + \int_0^t T(t-\tau)Bc(\tau)\,\mathrm{d}\tau. \tag{6}$$

It is easy also to verify that the adjoint operator of B, namely B^*, is given by

$$B^*: D \to D, \ \textit{with}\ B^*v := -v + \langle \delta_0, v\rangle\,\psi.$$

Remark 3.1 Observe also that both the cut-off function ψ and the function f defined in the example in the previous section belong to the set

$$D^2 \equiv D(A^{*2}) = \{g \in D \ : \ g' \in D\} = \{g \in H^2(0,S) \ : \ g(S) = g'(S) = 0\}, \tag{7}$$

that is, the domain of the generator of the adjoint semigroup $T^*(t)$ restricted to D.
Regarding the objective functional U_T, we assume

$$f \in C^0([0,S], \mathbb{R}^+), \ \exists \lambda > 0 \ : \ \textit{supp}(f) \subset [\lambda, \bar{s}] \tag{8}$$

$$u \in C^1(\mathbb{R}^+, \mathbb{R}^+) \ \textit{and concave} \tag{9}$$

Note that the quantity of wood harvested at time t (2) may be written by means of the duality pairing as

$$w(c(t)) = \langle c(t), f\rangle$$

with $w(c) = \int_0^S c(s)f(s)ds$ whenever c is a function in $L^2(0,S)$. Then, whenever finite, U_T may be written as

$$U_T(c) = \int_0^T e^{-\rho t} u\left(\langle c(t), f\rangle\right)\mathrm{d}t \tag{10}$$

with $0 \le T \le +\infty$.

3.2 Admissible controls and initial data

We denote a trajectory starting at x_0 and driven by a control c as $x(\cdot; x_0, c)$ or $x_{x_0,c}(\cdot)$. Although the abstract problem is set in D', we assume for technical reasons that x_0 is a function in $L^2(0,S)$ (and not a distribution), aware of the fact that initial densities where the mass is concentrated at certain ages do not match the requirements. Moreover, we assume that x_0 is compactly supported in $[0, \bar{s}]$, where $\bar{s}$ is the age above which trees will be considered not productive (see (8)). Such assumption allows for a simplification of the mathematics of the problem, as one expects that an optimal trajectory starting from such an initial data x_0 has to preserve the property of being null at ages greater than $\bar{s}$, that is, to be compactly supported in $[0, \bar{s}]$ as a function of s. Accordingly, the set admissible controls will be chosen so that $c(t)$ is a positive distributions in D', which is null for $s \ge \bar{s}$ and yields trajectories which are null for $s \ge \bar{s}$. All of these requirements need to be formally stated, as we do next.

Initial data. The initial densities x_0 are chosen in the set

$$\Pi := \left\{ x \in L^2(0,S) \ : \ x \geq 0, \ \ \mathrm{supp}(x) \subseteq [0,\bar{s}], \ \ \int_0^{\bar{s}} x(s)\,\mathrm{d}s = 1 \right\}. \tag{11}$$

where *supp*(g) denotes the support of any function/distribution g of variable s.

Admissible control strategies. The set $\mathcal{U}_{x_0}$ of admissible control strategies at some initial $x_0 \in \Pi$ is chosen as

$$\mathcal{U}_{x_0} := \left\{ c \in L^2_{loc}([0,+\infty); D') \ : \ \begin{array}{l} supp(c(t)), supp(x(t,\cdot)) \subseteq [0,\bar{s}] \ \forall t \geq 0 \\ c(t) \text{ and } x(t;x_0,c) \text{ lie in } \mathcal{R}, \ \forall t \geq 0 \end{array} \right\} \tag{12}$$

Remark 3.2 Note that the condition "$c(t)$ and $x(t;x_0,c)$ lie in $\mathcal{R}$" reads "$c(t)$ and $x(t;x_0,c)$ are positive measures", which implicitly says that the definition of admissible controls does not require the associated trajectory to be a function rather than a positive measure. Nonetheless such property holds true whenever one chooses a control in $\mathcal{U}_{x_0}$ and the initial datum in $L^2(0,S)$, as established in the following proposition. Moreover (and this is a mathematical technicality) although we consider controls which are $\mathcal{R}$-valued, we require their integrability with respect to the norm in D' by assuming $c \in L^2_{loc}([0,+\infty); D')$. Indeed, not only is in fact any positive distribution a positive Radon measure (see *e.g.* Proposition 2.3 page 270 of Hirsch and Lacombe (1999)), but also the D'–norm is dominated by the $\mathcal{R}$–norm, so that integrability with respect to D'–norm is a less restrictive requirement. This allows for a wider class of admissible controls.

Proposition 3.3 *Consider an initial datum x_0 in $L^2(0,S)$ and a control $c \in \mathcal{U}_{x_0}$. Then there exists a unique solution $x(\cdot)$ of (5) given by (6) and it belongs to $C(0,+\infty; D')$. Moreover, for any $t \in [0,+\infty)$, $x(t)$ belongs to $L^2(0,S)$.*[6]

Moreover we may formally prove that the surface of the land covered with trees stays constant in time, as stated in the next proposition.

Proposition 3.4 *Let us consider an initial datum $x_0 \in \Pi$ and $c \in \mathcal{U}_{x_0}$, with supp$(c(t)) \subseteq [0,\bar{s}]$, $c(t) \in \mathcal{R}$. Then the solution $x(\cdot) := x(\cdot;x_0,c)$ of (5) satisfies*

$$\langle x(t), \psi \rangle \equiv \langle x_0, \psi \rangle \ \ for\ all\ t \geq 0,$$

that implies that the solution x of (1)

$$\int_0^{\bar{s}} x(t,s)\,\mathrm{d}s = \int_0^{\bar{s}} x_0(s)\,\mathrm{d}s = 1 \tag{13}$$

In some cases we will restrict the class of admissible controls to smaller sets:

$$\mathcal{U}^{\lambda}_{x_0} := \left\{ c \in L^{\infty}([0,+\infty); \mathcal{R}) \ : \ \begin{array}{l} supp(c(t)) \subseteq [\lambda,\bar{s}], \ supp(x(t,\cdot)) \subseteq [0,\bar{s}], \ \forall t \geq 0 \\ c(t), x(t;x_0,c) \in \mathcal{R}, \ \forall t \geq 0 \end{array} \right\} \tag{14}$$

Note that here the controls are bounded in the $\mathcal{R}$–norm. Note also that assuming *supp*$(c^*(t)) \subseteq [\lambda,\bar{s}]$ seems natural, as by (8) any optimal control $c^*(t)$ is expected to satisfy

[6] Note that the solution coincides with the simplified formula (1) given by (3) only if in addition $c \in L^2_{loc}(0,+\infty; L^2(0,\bar{s}))$. Unfortunately, meaningful controls never fall into that class.

$supp(c^*(t)) \subseteq [\lambda, \bar{s}]$ for almost all $t \geq 0$ even if not required in the definition of the admissible class. A further restriction will be

$$\mathcal{U}_{x_0}^{\lambda,K} := \left\{ c \in L^\infty([0,+\infty);\mathcal{R}) \ : \ \begin{array}{l} supp(c(t)) \subseteq [\lambda, \bar{s}], \ supp(x(t,\cdot)) \subseteq [0,\bar{s}] \ \forall t \geq 0 \\ c(t), x(t;x_0,c) \in \mathcal{R}, \quad |c(t)|_{\mathcal{R}} \leq K, \quad \forall t \geq 0 \end{array} \right\} \tag{15}$$

where not only are the controls assumed bounded in $\mathcal{R}$–norm, but also they are by means of the same constant K.

3.3 Optimal, Maximal and Stationary programs

Here we formalize the definition of optimal and maximal control.

Definition 3.5 *Given $x_0 \in L^2(0,S)$ with $supp(x_0) \subseteq [0,\bar{s}]$ we say that $c^* \in \mathcal{U}_{x_0}$ is* optimal *at x_0 if, given any other control $c \in \mathcal{U}_{x_0}$, one has*

$$\liminf_{T\to\infty} \big(U_T(c^*) - U_T(c)\big) \geq 0.$$

and that $c^* \in \mathcal{U}_{x_0}$ is maximal *at x_0 if, given any other control $c \in \mathcal{U}_{x_0}$, one has*

$$\limsup_{T\to\infty} \big(U_T(c^*) - U_T(c)\big) \geq 0.$$

The trajectory x^* associated to an optimal (respectively, maximal) control c^* will be called an optimal *(*maximal*)* trajectory*, while the couple (x^*, c^*) will be referred to as* optimal (maximal) couple *or* program.

A last definition completes the abstract framework, that of stationary program. We require that a stationary program (x,c) to be time independent and, when $c(t) \equiv c \in L^2(0,S)$, to satisfy an equation like

$$x(s) = \left\{ \begin{array}{ll} x\,(s-t) - \int_{s-t}^{s} c(r)\mathrm{d}r & s \geq t \\ \int_{s}^{\bar{s}} c(r)\mathrm{d}r & 0 \leq s < t. \end{array} \right. \tag{16}$$

that is, to solve (1) with a null derivative of x with respect to time t. When c is instead a measure, (16) may be interpreted in the following abstract way.

Definition 3.6 *We say that $(\tilde{x}, \tilde{c}) \in \Pi \times \mathcal{R}$ is a* stationary couple *if, for all $t \geq 0$,*

$$\tilde{x} = T(t)\tilde{x} + \int_0^t T(t-s)B\tilde{c}\,\mathrm{d}s.$$

A stationary couple $(\tilde{x}, \tilde{c})$ is optimal *if $\tilde{c}$ is optimal at $\tilde{x}$.*

Note that both $\tilde{x}$ and $\tilde{c}$ need be time independent, although the dependence on s is implied.

In the following Lemma we identify all stationary couples of the problem.

Lemma 3.7 *Some admissible couple $(\tilde{x}, \tilde{c}) \in \Pi \times \mathcal{R}$ is stationary if and only if $-\tilde{x}$ is weakly increasing and $\tilde{c}$ is its Stieltjes derivative. In this case $\tilde{c}$ is also the derivative of $-\tilde{x}$ in the following distributional sense:*

$$\langle \tilde{x}, \phi \rangle_{L^2} = \langle \tilde{c}, \phi' \rangle_{D' \times D}$$

for any regular function $\phi \colon [0,S] \to \mathbb{R}$ compactly supported in $(0,S)$.

4 The Faustmann problem and the Golden Rules

We recall that, in the discrete time formulation in Mitra and Wan (1985), Mitra and Wan obtained that in a large set of circumstances the optimal policy would be that of *cutting any tree that reaches the critical age that maximizes the present value of bare forest land subject to an infinite sequence of planting cycles.* In our case, that corresponds to considering ages maximizing the function

$$\sum_{n=1}^{\infty} e^{-\rho n s} f(s) = \frac{f(s)}{e^{\rho s}-1}$$

when $\rho > 0$, which can be described as the value of an infinite sequence of planting cycles with harvesting at age s. Note that for f satisfying (8) the following sets coincide and are nonempty

$$\operatorname{argmax}\left\{\frac{f(s)}{e^{\rho s}-1} : s \in [0,\bar{s}]\right\} = \operatorname{argmax}\left\{\frac{1-e^{-\rho}}{e^{\rho s}-1} f(s) : s \in [0,\bar{s}]\right\},$$

moreover if we call $g_\rho(s) := \frac{f(s)}{e^{\rho s}-1}$ we have that $\lim_{\rho \to 0^+} g_\rho(s)(1-e^{-\rho}) = f(s)/s$. Therefore the analysis may be extended to the case $\rho = 0$ by maximizing, rather than $f(s)/(e^{\rho s}-1)$, the following function

$$G_\rho(s) = \begin{cases} \frac{1-e^{-\rho}}{e^{\rho s}-1} f(s) & \rho > 0 \\ f(s)/s & \rho = 0. \end{cases}$$

The Faustmann problem is that of choosing a maximizer M_ρ in

$$\mathcal{A}_\rho \equiv \operatorname{argmax}\{G_\rho(s) : s \in [0,\bar{s}]\}, \quad \forall \rho \geq 0.$$

Maximizers enjoy some interesting properties. In order to study them, once chosen a particular selection M_ρ in $\mathcal{A}_\rho$, it is useful to define the following *support function*

$$h_\rho(s) = \begin{cases} g_\rho(M_\rho)(e^{\rho s}-1) & \rho > 0 \\ f(M_0){M_0}^{-1} s & \rho = 0 \end{cases} \tag{17}$$

Remark 4.1 Indeed, since $h_\rho(s) \geq f(s)$, for all $s \in (0, S]$ and $h_\rho(M_\rho) = f(M_\rho)$, one has

$$\mathcal{A}_\rho = \{s \ \in (0,S] : \ h_\rho(s) = f(s)\}. \tag{18}$$

With reference to Figure 1, $\mathcal{A}_\rho$ is the set where the graph of f touches (from below) that of h_ρ.

Proposition 4.2 *Assume f satisfies (8). Then $\mathcal{A}_\rho \subset (0,S]$, and $\mathcal{A}_\rho \neq \emptyset$, for all $\rho \geq 0$. Moreover, if $0 < \rho_B < \rho_A$, then:*

(*i*) *There exists $\widetilde{s} \in (0,S]$ such that $\mathcal{A}_{\rho_A} \subseteq (0,\widetilde{s}]$ and $\mathcal{A}_{\rho_B} \subseteq [\widetilde{s},S]$. Moreover, $\mathcal{A}_{\rho_A}$ and $\mathcal{A}_{\rho_B}$ may be non-disjoint only if f is not differentiable at $\widetilde{s}$.*

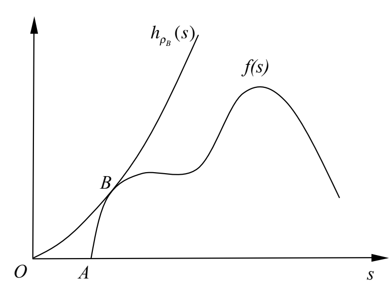

Figure 1: The support function when $\mathcal{A}_\rho$ is singleton.

(*ii*) *As a consequence, for any chosen* $M_\rho \in \mathcal{A}_\rho$*, the selections* $\rho \mapsto M_\rho$ *and* $\rho \mapsto \frac{f(M_\rho)}{M_\rho}$ *are nonincreasing. Moreover* $\mathcal{A}$ *is a not singleton for at most a numerable number of values* $\rho \geq 0$*.*

(*iii*) *For every selection* M_ρ *in* $\mathcal{A}_\rho$*, there exists* $\lim_{\rho \downarrow 0^+} M_\rho = m_0$*. Moreover there exists a selection* $\bar{M}_\rho$ *such that* $\lim_{\rho \downarrow 0^+} \bar{M}_\rho = \min \mathcal{A}_0$*.*

Remark 4.3 With reference to Figure 1, if $\mathcal{A}_{\rho_B}$ has a minimum M_{ρ_B}, then B has coordinates $(M_{\rho_B}, f(M_{\rho_B}))$ and, for all $\rho > \rho_B$, $(M_\rho, f(M_\rho))$ lies on the portion of the graph of f delimited by A and B. Moreover, this fact paired with that of $f(M_\rho)/M_\rho$ decreasing in ρ, suggests that multiplicity of maxima is a structurally unstable property as underlined in (*ii*) of Proposition 4.2.

4.1 The Golden Rule

A *modified golden rule* (x_ρ, c_ρ) (or simply *golden rule* when $\rho = 0$) is a candidate optimal/maximal stationary program of the problem so defined

$$x_\rho(s) := \frac{1}{M_\rho} \chi_{[0, M_\rho]}(s), \tag{19}$$

meaning that all ages in the range $[0, M_\rho]$ are uniformly distributed and equal to $1/M_\rho$, while those in the range $[M_\rho, S]$ are (almost everywhere) null. Note that x_ρ is a function in Π (see (11)). We also consider the stationary control strategy given by

$$c_\rho(t, s) \equiv \frac{1}{M_\rho} \delta_{M_\rho} \tag{20}$$

where δ_{M_ρ} is the Dirac Delta at point M_ρ, that is, the action undertaken by c_ρ is cutting exactly the trees reaching age M_ρ. Note that c_ρ is not a function of s but a distribution.

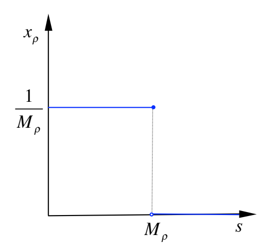

Figure 2: The golden rule stationary path.

The choice of the control c_ρ in the problem affects the quantity of wood, by definition of Dirac's Delta, as follows

$$w(c_\rho) = \langle c_\rho, f \rangle = \frac{1}{M_\rho} \langle \delta_{M_\rho}, f \rangle = \frac{1}{M_\rho} f(M_\rho). \tag{21}$$

Note that the set $\mathcal{A}_\rho$ may be not singleton. We will take such fact into account and derive different results accordingly.
Define now

$$\beta_\rho := \langle c_\rho, f \rangle = \frac{f(M_\rho)}{M_\rho}, \quad \eta_\rho := \frac{f(M_\rho)}{e^{\rho M_\rho} - 1}. \tag{22}$$

and $p_\rho \colon [0, S] \to \mathbb{R}^+$ as

$$p_\rho(s) := \begin{cases} \eta_\rho \left(e^{\rho s} - 1\right) \psi(s) & \rho > 0 \\ \beta_0 s \psi(s) & \rho = 0. \end{cases} \tag{23}$$

Note that $p_0(s) = \lim_{\rho \to 0^+} p_\rho(s)$ and that, for any $\rho \geq 0$, p_ρ is twice differentiable with $p_\rho(S) = {p_\rho}'(S) = 0$, which implies p_ρ is in D^2.

The dual variables in (23) have a straightforward interpretation as stationary competitive prices associated with a golden rule path (see Cass and Shell (1976)). Indeed, assume we interpret $p_\rho(s)$ as the (infinite dimensional) vector of the prices of capital goods (i.e, the prices of the different vintages s of trees) and set $R = \rho\eta_\rho$ the rent rate of the land on which the trees are planted (when $\rho = 0$, define $R = \lim_{\rho \to 0^+} \rho\eta_\rho = \beta_0$). Then by definition (23) one has

$$f(s) \leq p_\rho(s), \ s \in [0, \bar{s}], \quad f(M_\rho) = p_\rho(M_\rho),$$

where the first inequality means that no cutting process yields a positive profit, while the second says that the only cutting processes that do not generate losses are those that operates at the Faustmann ages. Thus, the golden rule controls maximizes the short run profits. In addition, since $p'_\rho(s) = \rho p_\rho(s) + R$ for $M_\rho \geq s \geq 0$ and $p'_\rho(s) \leq \rho p_\rho(s) + R$ for $s \geq M_\rho$, then the asset-market-clearing conditions that hold under competitive arbitrage are satisfied. Clearly, the arbitrage condition in a golden rule takes the form of a "modified Hotelling rule" because a piece of land needs to be rented in order to hold a tree of a given age *in situ*.

Remark 4.4 The bulk of the literature on optimal forest management in discrete time evolved from the approach to optimal growth that is contained in Brock (1970), where duality analysis is restricted to the stationary states and the full set of the Hamiltonian conditions is not studied. In this paper, we show that even in continuous time the analysis of the model can be carried out without the introduction of the Hamiltonian formalism. Accordingly, no attempt is done in this paper to establish some form of the Maximum Principle.

It is not difficult to guess that any golden rule is a stationary couple, as the amount of trees cut at age M_ρ is instantaneously replanted at age 0, preserving the distribution among different ages unaltered. Nonetheless the following result identifies completely the shape of stationary couples of the problem.

Theorem 4.5 *Assume $\rho \geq 0$, and f and u satisfying (8) (9) respectively. Consider the trajectory of system (5) starting at x_ρ. Then (x_ρ, c_ρ) is a stationary couple in the sense of Definition 3.6. Moreover, if $\mathcal{A}_\rho = \{M_\rho\}$, then the unique optimal stationary couple is (x_ρ, c_ρ).*

4.1.1 Golden Rules for positive discounts

In the following sections we will classify the behavior of candidate optimal or maximal programs, assuming either the discount ρ is positive or null, the utility function u is linear or strictly concave, $\mathcal{A}_\rho$ is singleton or multivalued. Nonetheless, when $\rho > 0$ is strictly positive, optimality of the golden rule is a general property (holding for a general concave utility u and a possibly multivalued $\mathcal{A}_\rho$), as stated in the next theorem.

Theorem 4.6 *Assume $\rho > 0$, $M_\rho \in \mathcal{A}_\rho$, and f and u satisfying (8) (9) respectively. Consider the trajectory of system (5) starting at x_ρ. Then c_ρ is optimal at x_ρ in the sense of Definition 3.5.*

Remark 4.7 Note that the assertion holds for any choice of M_ρ in $\mathcal{A}_\rho$. It is easy to prove that the same property holds for any convex linear combination of golden rules, that is, if $\mathcal{A}_\rho = \{M_\rho^1, ..., M_\rho^n\}$, and (x_ρ^i, c_ρ^i) is the golden rule associated to M_ρ^i, then

$$\tilde{x} = \sum_{i=1}^{n} \lambda_i x_\rho^i, \quad \tilde{c} = \sum_{i=1}^{n} \lambda_i c_\rho^i,$$

where $\lambda_i \geq 0, \sum_{i=1}^{n} \lambda_i = 1$, is also an optimal stationary program.

Remark 4.8 Note that by (21) it is $U_T(c_\rho) = \rho^{-1}(1 - e^{-\rho T})u(\beta_\rho)$, when $\rho > 0$ the golden rule is optimal when starting at x_ρ, with maximal overall utility given by

$$\max_{c \in \mathcal{U}_{x_\rho}} U(c) = U(c_\rho) = \lim_{T \to +\infty} U_T(c_\rho) = \frac{u(\beta_\rho)}{\rho},$$

The proof of Theorem 4.6 and of other theorems in the following sections relays on the construction of the *value-loss function*

$$\theta_\rho(c(t), x(t)) = u(\beta_\rho) - u\left(\langle c(t), f\rangle\right) + u'\left(\beta_\rho\right)\left[\rho\left\langle x(t) - x_\rho, p_\rho\right\rangle - \langle x(t), A^* p_\rho\rangle + \langle c(t), p_\rho\rangle\right], \quad (24)$$

which is formally introduced in the appendix, in the statement of Corollary A.6. This function. which gives the value-loss of any admissible couple at the steady state competitive prices, is the analogous of the value-loss function commonly used for finite dimensional optimal growth problem (see McKenzie (1986) for the discrete time case and Magill (1977) for continuous time). The only aspect that is specific to our infinite dimensional setting is that the unit rental costs function contains an element accounting for the ageing process of capital goods. Note that the input-output prices in the value-loss function are expressed in terms of marginal utilities, while capital goods prices in (23) are given in terms of timber. However, since there is a single final good, expressing the prices in the new numeraire amounts to rescaling them by means of the factor $u'(\langle c_\rho, f\rangle)$. An important consequence of this fact is that golden rules are independent from the instantaneous utility function (see (Mitra and Wan (1985)). On the contrary, the form of the instantaneous utility function matters in the analysis of the stationary states of the undiscounted model.

4.2 The Faustmann solution

Besides the golden rule, other controls are candidates to be optimal or maximal when starting at a general initial datum x_0. Indeed, the golden rule may fail even to be admissible at x_0.

Nonetheless, given an initial datum $x_0 \in \Pi$, if M_ρ represents a preferable cutting age providing a maximal harvesting, one may attempt to use the feedback strategy $(\hat{x}, \hat{c})$, where

$$\hat{c}(t) = \hat{x}(t, M_\rho)\delta_{M_\rho}, \quad \forall t \geq 0, \tag{25}$$

that is, $\hat{c}$ cuts existing trees reaching age M_ρ. Such trees vary in time depending on the initial distribution x_0. We remark that a trajectory of the system starting from an initial datum in Π is in L^2, as a function of s, so that $\hat{x}(t, M_\rho)$ is not well defined, as well as the control $\widehat{c}$. Nonetheless, in the following proposition we are able to give meaning to both.

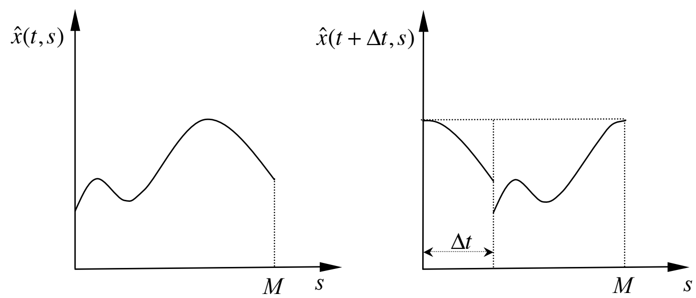

Figure 3: The Faustmann Solution. The effect of cutting at age M and replanting at age 0 induces cycling of the trajectory in a time length of M. After a time length $\Delta t < M$, the graph of the trajectory is translated forward of Δt and the portion exceeding age M reappears for s in $[0, \Delta t]$.

Lemma 4.9 *Assume x_0 satisfies supp$(x_0) \subset [0, M_\rho]$. Set*

$$\widehat{x}(t,s) = \widehat{x}(t)(s) = x_0(s - \sigma(t))\chi_{[\sigma(t),M_\rho]}(s) + x_0(s + M_\rho - \sigma(t))\chi_{[0,\sigma(t)]}(s) \tag{26}$$

where $\sigma(t) = \left\{\frac{t}{M_\rho}\right\} M_\rho = t - \left[\frac{t}{M_\rho}\right] M_\rho$, $[a]$ and $\{a\}$ denote respectively the integer and the fractional part of the real number a. Then $\hat{x}$ is M_ρ-periodic, that is $\hat{x}(t + M_\rho) = \hat{x}(t)$, for all $t \geq 0$, the control $\hat{c}(t) = \hat{x}(t, M_\rho)\delta_{M_\rho}$ is admissible at x_0 and $\hat{x}$ solves the closed loop equation

$$x(t) = T(t)x_0 + \int_0^t T(t-\tau)Bx(\tau, M_\rho)\, \delta_{M_\rho} \mathrm{d}\tau. \tag{27}$$

We will refer to $(\hat{x}, \hat{c})$ as to the *Faustmann solution* or *Faustmann Policy.*

Remark 4.10 Note that the golden rule is the Faustmann solution associated to the initial datum x_ρ.

Remark 4.11 Note that we require that the support of the initial datum x_0 lies in $[0, M_\rho]$ in order to prove the Lemma. Such assumption also implies $\int_0^S \widehat{x}(t,s)ds = 1$ indeed, since the solution is M_ρ periodic

$$\begin{aligned}\int_0^S \widehat{x}(t,s)ds &= \int_0^{\sigma(t)} \widehat{x}(t,s)ds + \int_{\sigma(t)}^{M_\rho} \widehat{x}(t,s)ds \\ &= \int_0^t x_0(s-t)ds + \int_t^{M_\rho} x_0(s + M_\rho - t)ds = \int_0^{M_\rho} x_0(r)dr.\end{aligned}$$

Lemma 4.12 *Assume supp$(x_0) \subset [0, M_\rho]$, $\hat{c}$ the Faustmann policy, $T \geq 0$, and set $n \equiv [T/M_\rho]$, $\sigma(T) = \{T/M_\rho\} M_\rho$ and*

$$U_1^\rho := \int_0^{M_\rho} e^{\rho\tau} u\left(f(M_\rho)x_0(\tau)\right) d\tau, \quad U_2^\rho(T) := \int_{M_\rho - \sigma(T)}^{M_\rho} e^{\rho\tau}\, u\left(f(M_\rho)x_0(\tau)\right) d\tau$$

Then

$$U_T(\widehat{c}) = \begin{cases} \frac{1-e^{-\rho n M_\rho}}{e^{\rho M_\rho}-1}\chi_{[M_\rho,\infty)}(T)\, U_1^\rho + e^{-\rho(n+1)M_\rho}U_2^\rho(T), & \rho > 0 \\ nU_1^0 + U_2^0(T), & \rho = 0 \end{cases} \tag{28}$$

Remark 4.13 Note that when $\rho > 0$ the overall utility is finite

$$U(\hat{c}) = \lim_{T\to\infty} U_T(\hat{c}) = U_1^\rho(e^{\rho M_\rho} - 1)^{-1},$$

contrary to the case $\rho = 0$ where it is not. The formula is consistent with those contained in Remark 4.8 when $x_0 = x_\rho$.

4.3 Null discounts and Good Controls

The case when $\rho = 0$ appears immediately as more complicated than the case of positive discount. For instance, if we compute the utility over a finite horizon T associated to the golden rule $(\bar{x}, \bar{c})$, we obtain

$$U_T(\bar{c}) = T\ u(\beta_0)$$

so that for null discount the utility over an infinite horizon fails to be finite, when evaluated at $\bar{c}$.

With null discount the notion of *good controls*, which we give next, will prove useful.

Definition 4.14 *Assume $\rho = 0$, and denote by $(\bar{x}, \bar{c})$ the associated golden rule, that is $\bar{x} = \frac{1}{M}\chi_{[0,M]}$ and $\bar{c} = \frac{1}{M}\delta_M$. A control $c \in \mathcal{U}_{x_0}$ is* good *if there exists $\theta \in \mathbb{R}$ s.t.*

$$\inf_{T\geq 0} \big(U_T(c) - U_T(\bar{c})\big) \geq -\theta.$$

We recall that the notion of good controls was introduced by Gale (1967) for the undiscounted n sector optimal growth model in discrete time. Note that a control is defined "good" in comparison to the golden rule $\bar{c}$, and that such comparison is performed although the golden rule may fail to be admissible at some arbitrary initial datum x_0. Note also that an equivalent way of giving the definition is to say that a control c is good if

$$\exists\theta \in \mathbb{R}\ :\ \forall T \geq 0,\ U_T(c) \geq U_T(\bar{c}) - \theta.$$

meaning that the utility (over an arbitrary finite horizon T) achieved by means of a good control differs from that obtained at $\bar{c}$ by at most a (possibly negative) finite quantity θ.

The following proposition compares good and optimal controls.

Proposition 4.15 *If $c^* \in \mathcal{U}_{x_0}$ is maximal (and, in particular, optimal) control then it is good.*

The previous result allows to seek for optimal or maximal programs in the class of good controls, as no control which is not good may be optimal or maximal.[7]

[7] We will exploit this result in proving Theorem 5.10, establishing optimality of the golden rule in the case of a single maximum for $f(s)/s$ and for $\rho = 0$ and u strictly concave.

5 Classification of optimal programs

5.1 Linear utility, positive discount

In Theorem 4.6 we already established that, when $\rho > 0$, the golden rules are optimal in all sets of assumptions. In particular this holds true for u linear, say $u(r) = r$. In the following theorem we establish that, in the particular case of $\rho > 0$ and u linear, the Faustmann solution is an optimal program, that is, the optimal policy is cutting trees reaching age M_ρ, regardless the initial distribution x_0, as long as x_0 does not contain trees older than the optimal age M_ρ. This is consistent with Theorem 4.6, as the Faustmann solution coincides with the golden rule when the initial datum is x_ρ.

Theorem 5.1 *Assume $\rho > 0$, f satisfying (8), and $u(r) = r$, $r \geq 0$. Consider an initial datum x_0 in Π with supp$(x_0) \subseteq [0, M_\rho]$. Then the Faustmann Solution $(\hat{x}, \hat{c})$, given by (26) (25) is optimal at x_0.*

Remark 5.2 The proof may be easily adapted to the case of affine utility $u(r) = ar + b$.

Remark 5.3 Note that, for a wide class of initial data, all those supported in $[0, M_\rho]$, the optimal trajectory is cyclic. As a counterpart, the modified golden rule is a stationary solution – an equilibrium – but not an asymptotic equilibrium. Optimal trajectories do not tend to any stationary solution, except when starting already at it.

Remark 5.4 In proving Theorem 5.1 we establish that the optimal value function is linear. The linearity of the function (i. e., all differences in value from the steady state reduces to the difference in value of the initial forest from the stationary forest) explains the lack of convergence of optimal trajectories to the modified golden rule. It is well known indeed that the clustering of solutions in optimal growth models is driven by second order differences due to strict concavity of the value function.

5.2 Linear utility, null discount

In this subsection we are going to show that, with null discount and linear u, the Faustmann solution is maximal but not optimal. Anyway, this is the best one may expect, as one shows that optimal programs do not exist.

Moreover through this and the following subsections we always require the assumption

$$\mathcal{A}_0 \quad \text{is singleton, } \mathcal{A}_0 \equiv \{M\} \tag{29}$$

and only in this case we discuss optimality and/or maximality of steady states and of the Faustmann Solution. The case of multivalued $\mathcal{A}_0$ remains unsolved, nonetheless multiplicity of maxima should vanish with small perturbations of the data as suggested in Remark 4.3.

Theorem 5.5 *Assume that (29) is satisfied, $\rho = 0$ and $u(r) = r$ for all $r \geq 0$. Consider an initial datum $x_0 \in \Pi$ with supp$(x_0) \subseteq [0, M]$. Then the Faustmann Solution $(\hat{x}, \hat{c})$, given by (26) (25) is maximal, although it is not optimal. Indeed no optimal control exists for the problem in this set of data.*

The fact applies to the particular case of the golden rule.

Corollary 5.6 *In the assumptions of Theorem 5.5, the golden rule $(\bar{x}, \bar{c})$ is a maximal, but not optimal, program at $\bar{x}$. Moreover no admissible control at $\bar{x}$ may be optimal.*

As a direct proof of the assertion that $\bar{c}$ is not optimal, nor an optimal control exists, one may build the following example and remark that the proof of Theorem 5.5 is based on a similar construction. The control c_1 defined by means of (30), admissible at $\bar{x}$, is not overtaken by $\bar{c}$: the control c_1 behaves *on average* like $\bar{c}$ but delayed of some initial time interval. We show that the difference in utilities yielded by $\bar{c}$ and c_1 coincide repeatedly with their difference in the initial time interval, precisely because $\rho = 0$ and u is linear.

Example 5.7 Let N be a natural number greater than 1. Define $s_j := jM/N$, for $j = 1,..,N$ and consider a control c_1 and associated trajectory x_1 so defined: when $t \leq M/N$, c_1 cuts the quantity $x_1(t, s_j)$ of available trees of age s_j, subsequently when $t \geq M/N$, c_1 cuts the quantity $x_1(t, M)$ of trees reaching age M, that is

$$c_1(t) = \begin{cases} \sum_{j=1}^{N} x_1(t, s_j)\delta_{s_j}, & 0 \leq t < \frac{M}{N} \\ x_1(t, M)\delta_M, & t \geq \frac{M}{N} \end{cases} \qquad (30)$$

It is easy to check that c_1 is admissible at any x_0 with $supp(x_0) \subset [0, M]$, in particular for $x_0 = \bar{x}$. In the latter case, the associated trajectory $x_1(t, s; c_1, \bar{x}) \equiv x_1(t, s)$ is (a.e.) given by the explicit formula

$$x_1(t, s) = \frac{N}{M}\chi_{[0,t]}(s) + \frac{1}{M}\sum_{j=1}^{N}\chi_{[s_{j-1}+t, s_j]}(s)$$

when $t \in \left[0, \frac{M}{N}\right]$, $s \geq 0$, while in the following interval of length M it is equal to

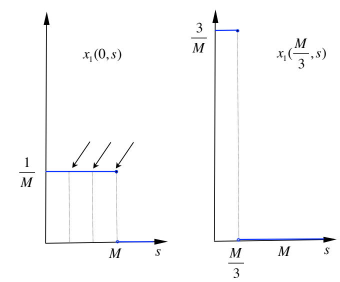

Figure 4: The control c_1 with $N = 3$. From time 0 to time $M/3$ the trees of age $M/3$, $(2M)/3$, and M are cut, yielding a resulting distribution, at time $M/3$, $x_1(M/3, s) = \frac{3}{M}\chi_{[0,M/3]}$.

$$x_1(t, s) = \begin{cases} \frac{N}{M}\chi_{[t-\frac{M}{N},t]}(s) & t \in \left[\frac{M}{N}, M\right] \\ \frac{N}{M}\left[\chi_{[0,t-M]}(s) + \chi_{[t-\frac{M}{N},M]}(s)\right] & t \in \left[M, M + \frac{M}{N}\right] \end{cases} \qquad (31)$$

It is easy understood that, from $t = M/N$ on, the trajectory is periodic with period M, and attains the values described in (31) in all intervals of type $[T_i, T_{i+1}]$ where $T_i = M/N + iM$, with $i \in \mathbb{N}$.

Note that for during every time interval $[T_i, T_{i+1}]$, the control c_1 cuts an amount N/M for a time length M/N, while $\bar{c}$ cuts the amount $1/M$ for a time length M. Then, except for the quantity obtained at the initial time interval, the utilities yielded by $\bar{c}$ and c_1 on a period length interval are both equal to $f(M)$, as no discount is applied and u is linear. As a result, the difference between such utilities is periodically equal to the difference yielded on $[0, M/N]$, that is

$$U_{\frac{M}{N}}(c_1) - U_{\frac{M}{N}}(\bar{c}) = \frac{1}{N}\sum_{j=1}^{N-1} f(s_j)$$

and which may be assumed strictly positive, provided f is not null everywhere, as for a suitable choice of N one may the infer that $f(s_j) > 0$ for at least one j. Such fact is equivalent by definition to stating that the control $\bar{c}$ cannot be (definitively) overtaking c_1, and by means of the same idea one is also able to contradict the existence of an optimal control. For details we refer the reader to the proof of the general case, Theorem 5.5 in the appendix.

Remark 5.8 As it is shown later in Theorem 5.11 (ii), for a utility u which is concave but not necessarily strictly concave, one may prove existence of a maximal control when the admissible set is $\mathcal{U}_{x_0}^{\lambda,K}$ defined in (15). In particular, the result applies when u is linear (and $\rho = 0$).

5.3 Strictly concave utility, null discount

As in the previous, is this subsection we assume that (29) is satisfied. Moreover all statements are proved not for $\mathcal{U}_{x_0}$ as admissible class of strategies, but on the subsets $\mathcal{U}_{x_0}^{\lambda}$ and $\mathcal{U}_{x_0}^{K,\lambda}$ defined in (14) (15).

Theorem 5.9 *Assume $\rho = 0$, and that (9)(8)(29) are satisfied. Assume moreover that u is strictly concave. Then, along the trajectory $x(\cdot)$ starting from some $x_0 \in \Pi$ and driven by a good control $c(\cdot) \in \mathcal{U}_{x_0}^{K,\lambda}$ one has*

$$x(t) \xrightarrow[L^2(0,S)]{t\to\infty} \bar{x}$$

that is, the trajectory $x(t)$ converges to the golden rule $\bar{x}$ in $L^2(0,S)$ norm.

A straightforward consequence of Theorem 5.9 and of is that its assertion holds also for optimal trajectories, i.e. trajectories driven by an optimal control, which is good by means of Proposition 4.15.

Theorem 5.10 *Assume $\rho = 0$, and that* (8) (9) (29) *hold. Assume moreover that u is strictly concave. Then the golden rule $(\bar{x}, \bar{c})$ is an optimal stationary couple.*

As a consequence of Theorem 5.9 is that Faustmann solutions (except for the Golden Rule itself), which were optimal or maximal for linear u, are not optimal for strictly concave u. Indeed, if they were, the associated trajectory would be converging to the Golden Rule, and not cycling.

Theorem 5.11 *Let $x_0 \in \Pi$, $\rho = 0$, and assume (29) is satisfied. Let $\mathcal{U}_{x_0}^{K,\lambda}$ be the space of admissible control defined in (15). Then:*

- (i) *if u is strictly concave, then there exists an optimal control in $\mathcal{U}_{x_0}^{K,\lambda}$;*
- (ii) *if u is concave (but not necessarily strictly concave), then there exists a maximal control in $\mathcal{U}_{x_0}^{K,\lambda}$.*

Note that Theorem 5.9 and Proposition 4.15 imply the following Corollary.

Corollary 5.12 *In the assumptions of Theorem 5.9, the Faustmann solution can be neither an optimal nor a maximal program, except for the particular case of the Golden Rule.*

5.4 Strictly concave utility, positive discount

It is a simple corollary of Theorem 5.9 that the Faustmann Policy is not optimal for the case of a strictly concave utility function and null discount. However, this result does not preclude the possibility that the Faustmann policy turns out optimal for the discounted model. Indeed, for the discrete time model with a strictly concave utility and discounted future utilities, Mitra and Wan in Mitra and Wan (1985) provided a couple of examples in which the Faustmann Policy was in fact optimal, and Wan in Wan Jr (1994) and Salo and Tahvonen in Salo and Tahvonen (2002) and Salo and Tahvonen (2003) have taken the issue further (see also Mitra et al. (1991) for similar results in a different vintage capital model) by showing that optimal Faustmann cycles persist in a neighborhood of the (modified) golden rule even if the discount factor approaches unity. Proposition 1 in Salo and Tahvonen Salo and Tahvonen (2003), in particular, states that for any discount factor less than one the Faustmann Policy is optimal for all initial forests that are sufficiently close to the uniform steady state forest.

On the contrary, for the strictly concave continuous time discounted model the issue of the optimality of cyclical Faustmann solutions is still open: neither a convergence result for the continuous time discounted model is available, nor a case in which the Faustmann Policy is definitely optimal has been found. However, we can establish a partial result by proving that Proposition 1 in Salo and Tahvonen (2003) does not carry over to our continuous time formulation and, hence, that the continuous time model behaves differently from the discrete time model. To illustrate the point consider the following simple example in which a Most Rapid Approach Path to the steady state dominated the path generated by the Faustmann Policy for a set of initial distributions that contains elements arbitrary close to the (modified) golden rule. By suitable rescaling, one may assume that $M_\rho = 1$ is the unique Faustmann maturity age, and that $f(1) = 1$. Consider the following initial density of forest

$$x_a(s) = \begin{cases} 1 & s \in [0, 1-2a[\\ 1+a & s \in [1-2a, 1-a[\\ 1-a & s \in [1-a, 1] \\ 0 & s \in]1, +\infty), \end{cases}$$

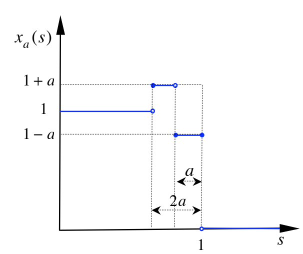

Figure 5: The initial state $x_a(s)$.

where $0 \le a \le \frac{1}{2}$ (note that for $a = 0$ we obtain the golden rule forest). We intend to show that for any a at a neighborhood of 0 the Faustmann Policy is not optimal starting at $x_a(s)$. As a consequence, the analogous of Proposition 1 in Salo and Tahvonen (2003) does not hold in continuous time.

We first compute the utility associated to the Faustmann Policy $\hat{c}_a(t) = \hat{x}_a(t,1)\delta$. Note that each Faustmann cycle comprises three phases of constant timber consumption: an initial phase of length a during which $1-a$ units of timber are consumed followed by a phase of length a during which timber consumption rises to $1+a$ and a final phase of length $1-2a$ during which consumption is constant at the golden rule level. Then on the first cycle, that is for $t \in [0,1]$, one has

$$\langle \hat{c}_a(t), f \rangle = f(1)x_a(t,1) = x_a(t,1) = (1-a)\,\chi_{[0,a]}(t) + (1+a)\,\chi_{[a,2a]}(t) + \chi_{[2a,1]}(t)$$

so that the utility at horizon $T = 1$ is given by

$$
\begin{aligned}
U_1(\hat{c}_a) &= \int_0^a u(1-a)e^{-\rho t}\mathrm{d}s + e^{-\rho a}\int_0^a u(1+a)e^{-\rho t}\mathrm{d}s + e^{-2\rho a}\int_0^{1-2a} u(1)e^{-\rho t} \\
&= \frac{u(1-a)+u(1+a)e^{-\rho a}}{\rho}\left(1-e^{-\rho a}\right) + \frac{u(1)}{\rho}(e^{-2\rho a} - e^{-\rho})
\end{aligned}
$$

Next we note that, starting from any forest $x_a(s)$, a feasible Most Rapid Approach Path to the steady state, namely

$$
\begin{aligned}
\langle c_{mra}(t), f\rangle &= \left[(1-a)f(1) + af(1-a)\right]\chi_{[0,a]}(t) + f(1)\chi_{[a,\infty)}(t) \\
&= \left[(1-a) + af(1-a)\right]\chi_{[0,a]}(t) + \chi_{[a,\infty)}(t)
\end{aligned}
$$

reaches the golden rule after a units of time by continuously clearing the $1-a$ units of land on which mature trees are planted and the a units of land in excess on which trees of age $1-a$ are grown.

The utility associated to the MRA policy at horizon $T = 1$ is then

$$
\begin{aligned}
U_1(c_{mra}) = \int_0^a u\left((1-a)+af(1-a)\right)e^{-\rho t}\mathrm{d}s &+ e^{-\rho a}\int_0^a u(1)e^{-\rho t}\mathrm{d}s \\
&+ e^{-2\rho a}\int_0^{1-2a} u(1)e^{-\rho t}\mathrm{d}s \\
&= \frac{u\left((1-a)+af(1-a)\right)}{\rho}\left(1-e^{-\rho a}\right) + \frac{u(1)}{\rho}(e^{-\rho a} - e^{-\rho})
\end{aligned}
$$

Instead, when $T = n \geq 2$, one has

$$
\begin{aligned}
U_n(\hat{c}_a) &= \sum_{i=0}^{n} e^{-i\rho}U_1(\hat{c}_a) = U_1(\hat{c}_a) + U_1(\hat{c}_a)\left(\frac{1-e^{-n\rho}}{1-e^{-\rho}} - 1\right) \\
U_n(c_{mra}) &= U_1(c_{mra}) + \sum_{i=1}^{n} e^{-i\rho}\frac{u(1)}{\rho}\left(1-e^{-\rho}\right) \\
&= U_1(c_{mra}) + \frac{u(1)}{\rho}\left(1-e^{-\rho}\right)\left(\frac{1-e^{-n\rho}}{1-e^{-\rho}} - 1\right)
\end{aligned}
$$

As a consequence

$$
\frac{u(1)}{\rho}(1-e^{-\rho}) - U_1(c_{mra}) = \frac{\left(u(1) - u(1-a+af(1-a)\right)}{\rho}\left(1-e^{-\rho a}\right)
$$

$$
\begin{aligned}
U_1(\hat{c}_a) - \frac{u(1)}{\rho}(1-e^{-\rho}) &= \frac{u(1-a)+u(1+a)e^{-\rho a}}{\rho}\left(1-e^{-\rho a}\right) + \frac{u(1)}{\rho}(e^{-2\rho a} - 1) \\
&= \frac{u(1-a) - u(1) + (u(1+a)-u(1))e^{-\rho a}}{\rho}\left(1-e^{-\rho a}\right)
\end{aligned}
$$

and

$$\lim_{n\to\infty}\left(U_n(\hat{c}_a)-U_n(c_{mra})\right)=U_1(\hat{c}_a)-U_1(c_{mra})+\ \left(U_1(\hat{c}_a)-\frac{u(1)}{\rho}\left(1-e^{-\rho}\right)\right)\frac{e^{-\rho}}{1-e^{-\rho}}$$

$$=\frac{u(1)}{\rho}(1-e^{-\rho})-U_1(c_{mra})+\ \left(U_1(\hat{c}_a)-\frac{u(1)}{\rho}\left(1-e^{-\rho}\right)\right)\frac{1}{1-e^{-\rho}}$$

$$=\frac{1-e^{-\rho a}}{\rho}\left[\frac{u(1-a)-u(1)}{1-e^{-\rho}}+\frac{u(1+a)-u(1)}{1-e^{-\rho}}e^{-\rho a}-u(1-a+af(1-a))+u(1)\right].$$

We now note that $\frac{1-e^{-\rho a}}{\rho}$ is strictly positive for $a>0$, so the sign of the above expression is the sign of the sum in the square brackets. Such sum is null at $a=0$, its sign for $a>0$ in a neighborhood of 0 is given by the sign of its lowest order non-zero derivative evaluated at the steady state. Simple calculations show that the first derivative is null, while the second is given by

$$2(u''(1)-\rho u'(1))+2f'(1)u'(1)(1-e^{-\rho})=2u''(1),$$

which is strictly negative, as $f'(1)(1-e^{-\rho})=\rho$ because $M_\rho=1$ is a solution of the Faustmann problem. One can therefore conclude that the Faustmann Policy is not optimal starting at $x_a(s)$ in a neighborhood of the steady state.

6 Conclusions

In this paper we developed and analyzed a continuous time version of the Mitra-Wan (Mitra and Wan (1985)) model of optimal forest management. Following the methodological precept (Foley (1975)) that discrete and continuous-time modeling should lead to the same predictions, our main purpose was to isolate the set of phenomena that in the optimal management of a forest are independent of the way time is modeled. Table 1 gives an overview of the results we have established for the continuous-time model and that can be compared with the results that have been obtained in the discrete time model. It turned out that many of the discrete time results carry over to the continuous time version of the model with an important exception: the cyclical optimal solutions that are characteristics of the discounted strictly-concave discrete model disappear in continuous time. We have also established for the continuous time model a set of results that the literature in discrete time has not yet proved: that the Faustmann solution is maximal but not optimal for the undiscounted model with a linear utility function and that in steady states both the Faustmann age and timber production decrease monotonically with the increase of the discount rate.

Unlike in discrete time, modeling timber production in continuous time required a quantum leap from the received vintage capital theory. Indeed, in the typical vintage capital model in continuous time, only irreversible investment in new machines is possible, so that distributed controls can be avoided altogether. Moreover, in the few instances in which investment in older machines is considered (e.g. Feichtinger et al. (2006)) it turned out that optimal investment is spread over a continuum of ages, so that the controls can be functions. In continuous time, however, timber production cannot be modeled this way, because the Faustmann condition implies that generically it is optimal to fell down only trees of a single age. Therefore, to handle the case of forest management in continuous time we had to develop an entirely new class of

<table>
<tr><th></th><th>$\rho = \mathbf{0}$</th><th>$\rho > \mathbf{0}$</th></tr>
<tr><td>u linear</td><td>If $\mathcal{A}_0$ is singleton the GR is maximal, but not optimal
If $\mathcal{A}_0$ is singleton the FS is maximal at any admissible x_0 satisfying $supp(x_0) \subset [0, M_\rho]$
There do not exist optimal controls</td><td>Any MGR is optimal
FS is optimal at any admissible x_0 satisfying $supp(x_0) \subset [0, M_\rho]$</td></tr>
<tr><td>u strictly concave</td><td>Assume $\mathcal{A}_0$ is singleton:
GR is the unique optimal stationary couple
There exists an optimal control
Any optimal trajectory converges (in L^2 norm) to the GR</td><td>Any MGR is optimal
It is not true (as in discrete time) that FS is optimal for all initial forests close to the MGR. There is a counterexample</td></tr>
<tr><td>u concave</td><td>Assume $\mathcal{A}_0$ is singleton:
GR is the unique maximal stationary couple
There exists a maximal (admissible) control</td><td>Any MGR is optimal</td></tr>
</table>

Table 1: Results at one glance: FS stands for Faustmann Solution, GR for Golden Rule, MGR for modified golden rule.

vintage models in which measure-valued controls are allowed. Since this is the first attempt to formulate the Mitra-Wan model in continuous time, we have been concentrating on the basic features of the model, without attempting to use minimal assumptions and without taking into account recent refinements of the theory (Khan and Piazza (2012)). Discussion of these issues is left for future work. We also used an oversimplified framework that, however, can be generalized in several directions. For example, it would be straightforward considering positive cutting costs, adding environmental well-being variables into the objective, and allowing for alternative use of the forest land.

Acknowledgments

We we like to thank Raouf Boucekkine, Gustav Feichtinger, Wilfred Nyangena, Thomas Mariotti, Olli Tahvonen, Yannis Vailakis, and Vladimir Veliov for their useful comments and suggestions.

References

R. B. Ash. *Probability and measure theory.* Harcourt/Academic Press, Burlington, MA, second edition, 2000. With contributions by Catherine Doléans-Dade.

E. Barucci and F. Gozzi. Optimal investment in a vintage capital model. *Research in Economics*, 52:159–188, 1998.

E. Barucci and F. Gozzi. Optimal advertising with a continuum of goods. *Annals of Operations Research*, 88:15–29, 1999.

E. Barucci and F. Gozzi. Technology adoption and accumulation in a vintage-capital model. *Journal of economics*, 74(1):1–38, 2001.

A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter. *Representation and control of infinite dimensional systems.* Systems and Control: Foundations and Applications. Birkhäuser Boston Inc., Boston, MA, second edition, 2007.

R. Boucekkine, M. Germain, and O. Licandro. Replacement echoes in the vintage capital growth model. *Journal of Economic Theory*, 74(2):333–348, 1997.

R. Boucekkine, M. Germain, O. Licandro, and A. Magnus. Creative destruction, investment volatility, and the average age of capital. *Journal of Economic Growth*, 3(4):361–384, 1998.

R. Boucekkine, D. De La Croix, and O. Licandro. Vintage capital growth theory: Three breakthroughs. In Olivier de La Grandville, editor, *Economic Growth and Development*, volume 11 of *Frontiers of Economics and Globalization*, chapter 5, pages 87–116. Emerald Group Publishing Limited, 2011.

W. A. Brock. On existence of weakly maximal programmes in a multi-sector economy. *Review of Economic Studies*, 37(2):275–280., 1970.

W. A. Brock and A. Haurie. On existence of overtaking optimal trajectories over an infinite time horizon. *Mathematics of Operations Research*, 1(4):337–346, 1976.

D. A. Carlson, A. Haurie, and A. Jabrane. Existence of overtaking solutions to infinite dimensional control problems on unbounded time intervals. *SIAM journal on control and optimization*, 25(6):1517–1541, 1987.

D. Cass. On the wicksellian point-input, point-output model of capital accumulation: A modern view (or, neoclassicism slightly vindicated). *Journal of Political Economy*, 81(1):71–97, 1973.

D. Cass and K. Shell. Introduction to hamiltonian dynamics in economics. *Journal of Economic Theory*, 12(1):1–10, 1976.

W.L. Chan and G. B. Zhu. Overtaking optimal control problem of age-dependent populations with infinite horizon. *Journal of Mathematical Analysis and Applications*, 150(1):41–53, 1990.

Kai Lai Chung. *A course in probability theory. Third edition.* Access Online via Elsevier, 2001.

S. Dasgupta and T. Mitra. On optimal forest management: A bifurcation analysis. In K. G. Dastidar, H. Mukhopadhyay, and U. B. Sinha, editors, *Dimensions of economic theory and policy: essays for Anjan Mukherji*. 2011.

K. J. Engel and R. Nagel. *One-parameter semigroups for linear evolution equations*, volume 194 of *Graduate Texts in Mathematics*. Springer, 1999.

S. Faggian. Regular solutions of first-order Hamilton-Jacobi equations for boundary control problems and applications to economics. *Applied Mathematics and Optimization*, 51(2):123–162, 2005. ISSN 0095-4616.

S. Faggian. Hamilton-Jacobi equations arising from boundary control problems with state constraints. *SIAM Journal on Control and Optimization*, 47(4):2157–2178, 2008.

S. Faggian and F. Gozzi. On the dynamic programming approach for optimal control problems of PDE's with age structure. *Mathematical Population Studies*, 11(3-4):233–270, 2004.

S. Faggian and F. Gozzi. Optimal investment models with vintage capital: Dynamic programming approach. *Journal of Mathematical Economics*, 46(4):416–437, 2010.

S. Faggian and L. Grosset. Optimal advertising strategies with age-structured goodwill: segment vs. single medium scenario. *Annals of Operations Research*, 2013. to appear.

G. Feichtinger, G. Tragler, and V.M. Veliov. Optimality conditions for age-structured control systems. *Journal of Mathematical Analysis and Applications*, 288(1):47–68, 2003.

G. Feichtinger, R. F. Hartl, P. M. Kort, and V. M. Veliov. Anticipation effects of technological progress on capital accumulation: a vintage capital approach. *J. Econom. Theory*, 126(1): 143–164, 2006.

D.K. Foley. On two specifications of asset equilibrium in macroeconomic models. *Journal of Political Economy*, 83(2):303–324, 1975.

D. Gale. On optimal development in a multi-sector economy. *The Review of Economic Studies*, 34(1):1–18, 1967.

H. Halkin. Necessary conditions for optimal control problems with infinite horizons. *Econometrica*, 42(2):267–272, 1974.

T. Heaps. The forestry maximum principle. *Journal of Economic Dynamics and Control*, 7(2): 131–151, 1984.

T. Heaps. An extension of the forestry maximum principle. working paper, 2006.

B.J. Heijdra and W.E. Romp. Retirement, pensions, and ageing. *Journal of Public Economics*, 93(3):586–604, 2009.

E. Hewitt. Integration by parts for stieltjes integrals. *The American Mathematical Monthly*, 67 (5):419–423, 1960.

F. Hirsch and G. Lacombe. *Elements of functional analysis*, volume 192 of *Graduate Text in Mathematics*. Springer, 1999.

N. Hritonenko and Y. Yatsenko. From linear to nonlinear utility in vintage capital models. *Mathematical Population Studies*, 15(4):230–248, 2008.

M. A. Khan and A. Piazza. On uniform convergence of undiscounted optimal programs in the mitra–wan forestry model: the strictly concave case. *International Journal of Economic Theory*, 6(1):57–76, 2010.

M. A. Khan and A. Piazza. An overview of turnpike theory: towards the discounted deterministic case. *Advances in Mathematical Economics*, 4(1):39–67, 2011.

M. A. Khan and A. Piazza. On the mitra–wan forestry model: a unified analysis. *Journal of Economic Theory*, 147(1):230–260, 2012.

M. J .P. Magill. Some new results on the local stability of the process of capital accumulation. *Journal of Economic Theory*, 15(1):174–210, 1977.

A. Mas-Colell, M.D. Whinston, and J.R. Green. *Microeconomic theory*. Oxford university press New York, 1995.

L.W. McKenzie. Optimal economic growth, turnpike theorems and comparative dynamics. In *Handbook of mathematical economics 3*, pages 1281–1355. Elsevier, 1986.

L.W. McKenzie. Accumulation programs of maximum utility and the von neumann facet. In T. Mitra and K. Nishimura, editors, *Equilibrium, Trade, and Growth: Selected Papers of Lionel W. McKenzie*, pages 345–376. MIT Press, 2009.

T. Mitra and H. Y. Wan. Some theoretical results on the economics of forestry. *The Review of Economic Studies*, 52(2):263–282, 1985.

T. Mitra and H. Y. Wan. On the Faustmann solution to the forest management problem. *Journal of Economic Theory*, 40(2):229–249, 1986.

T. Mitra, D. Ray, and R. Roy. The economics of orchards: an exercise in point-input, flow-output capital theory. *Journal of Economic Theory*, 53(1):12–50, 1991.

A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

B. Peleg. A weakly maximal golden-rule program for a multi-sector economy. *International Economic Review*, 14(3):574–579, 1973.

W. Rudin. *Real and complex analysis. Third edition.* Tata McGraw-Hill Education, 1987.

S. W. Salant. The equilibrium price path of timber in the absence of replanting: Does hotelling rule the forests too? *Resource and Energy Economics*, 35(4):572–581, 2013.

S. Salo and O. Tahvonen. On equilibrium cycles and normal forests in optimal harvesting of tree vintages. *Journal of Environmental Economics and Management*, 44(1):1–22, 2002.

S. Salo and O. Tahvonen. On the economics of forest vintages. *Journal of Economic Dynamics and Control*, 27(8):1411–1435, 2003.

P. A. Samuelson. Economics of forestry in an evolving society. *Journal of Forest Economics*, 1 (1):115–150, 1995. (reprinted from Economic Inquiry 14, 466-492, 1976).

R. M. Solow, J. Tobin, C. C. von Weizsäcker, and M. Yaari. Neoclassical growth with fixed factor proportions. *The Review of Economic Studies*, 33(2):79–115, 1966.

O. Tahvonen. Optimal harvesting of forest age classes: a survey of some recent results. *Mathematical Population Studies*, 11(3-4):205–232, 2004.

O. Tahvonen. Economics of harvesting age-structured fish populations. *Journal of Environmental Economics and Management*, 58(3):281–299, 2009.

O. Tahvonen and S. Salo. Optimal forest rotation with in situ preferences. *Journal of Environmental Economics and Management*, 37(1):106–128, 1999.

O. Tahvonen, S. Salo, and J. Kuuluvainen. Optimal forest rotation and land values under a borrowing constraint. *Journal of Economic Dynamics and Control*, 25(10):1595–1627, 2001.

C. C. Von Weizsäcker. Existence of optimal programs of accumulation for an infinite time horizon. *The Review of Economic Studies*, 32(2):85–104, 1965.

H. Wan Jr. Revisiting the mitra-wan tree farm. *International Economic Review*, 35(1):193–198, 1994.

A. L. Zaslawski. *Turnpike properties in the calculus of variations and optimal control*. Springer, Berlin, 2006.

W. P. Ziemer. Modern real analysis. *Online lecture notes*, 2004. www.indiana.edu/ mathwz/PRbook.pdf.

A Appendix: proofs

In this appendix we complete our exposition with the proof of the aforementioned results.

A.1 Proofs for Section 3

The subsection contains the proofs of the results which were stated in Section 3 and of some other useful ones. Those results are mostly well known in semigroup theory applied to control in infinite dimensions. In this respect, our main reference is Bensoussan et al. (2007).

Proposition A.1 *The operator $\mathcal{C}\colon L^2([0,T];D') \to C([0,T];D')$ given by $\mathcal{C}(c)(t) := \int_0^t e^{(t-s)A}Bc(s)\,\mathrm{d}s$ is continuous. As a consequence $\mathcal{S}\colon D' \times L^2([0,T];D') \to C([0,T];D')$ defined by by $\mathcal{S}(x_0,c)(t) := T(t)x_0 + \mathcal{C}(c)(t)$ is continuous. Moreover, for any $x_0 \in D'$, and for any $c \in \mathcal{U}_{x_0}$, the function $[0,T] \to D'$, $t \mapsto T(t)x_0 + \mathcal{C}(c)(t)$ is also continuous.*

Proof. See e.g. Bensoussan et al. (2007) Section II.1.3. □

Proof of Proposition 3.3. The proof of the first fact is a consequence of Proposition A.1. Next we need to show that, for any $t \in [0, +\infty)$, $x(t)$ lies in $L^2(0, S)$. Expanding in (6) the definition of B we have

$$x(t) = T(t)x_0 + \int_0^t \langle c(\tau), \psi \rangle \, T(t-\tau)\delta_0 \, \mathrm{d}\tau - \int_0^t T(t-\tau)c(\tau) \, \mathrm{d}\tau. \tag{32}$$

(*i*) Since $x_0 \in L^2(0, S)$ then $T(t)x_0 \in L^2(0, S)$, as $T(t)$ coincides on $L^2(0, S)$ with the translation semigroup. By definition, $T(\tau)x_0$ is then a positive function;

(*ii*) The term $\int_0^t \langle c(\tau), \psi \rangle \, T(t-\tau)\delta_0 \, \mathrm{d}\tau$ belongs to $L^2(0, S)$ as a consequence of Proposition 3.1 in Bensoussan et al. (2007), p. 212, once we have proven that Hypothesis 3.1 page 212 in Bensoussan et al. (2007) is satisfied. To this extent, we denote with B_1 the operator $B_1 \colon D' \to D'$ given by $B_1\phi := \langle \phi, \psi \rangle \, \delta_0$ so that $B_1^* \colon D \to D$ is given by $B_1^* h := \langle \delta_0, h \rangle \psi$. As a consequence

$$|B_1^* T^*(\tau)h|_D \ \leq |\langle \delta_0, T^*(\tau)h \rangle|_{\mathbb{R}} \, |\psi|_D \leq |h(\tau)|_{\mathbb{R}} \, |\psi|_D$$

so that

$$\int_0^T |B_1^* T^*(\tau)h|_D^2 \, \mathrm{d}\tau \leq |\psi|_D^2 \int_0^S |h(\tau)|_{\mathbb{R}}^2 \, \mathrm{d}\tau = |\psi|_D^2 |h|_{L^2(0,S)}^2.$$

Moreover, since all the integrands are positive, B is a positive function.

(*iii*) Last we observe that $-\int_0^t T(t-\tau)c(s) \, \mathrm{d}\tau$ is a negative distribution in D' and by Proposition 2.3 page 270 of Hirsch and Lacombe (1999) is a negative measure on $[0, S]$. Then, by means of Lebesgue decomposition theorem, it may be decomposed into a (negative) part, absolutely continuous w.r.t. the Lebesgue measure on $[0, S]$, and a (negative) singular part.

As a result of the previous analysis, the measure defined by the right side of (32) has singular part coinciding with that described in (*iii*). At the same time, the singular part of the measure defined by the left side of (32) need be positive, then the singular part on both sides is null. Then for any $t \geq 0$, the term $-\int_0^t T(\tau)c(\tau) \, \mathrm{d}\tau$ represents (a measure which is an absolutely continuous w.r.t. the Lebesgue measure on $[0, s]$) which also may be written by means of (32) as the sum of two functions of $L^2(0, S)$, hence a function in $L^2(0, S)$. The explicit formula solution when $c(\cdot) \in L^2_{loc}(0, +\infty; L^2(0, S))$ can be computed by standard calculations. □

Proof of Proposition 3.4. Consider first the case in which $c \in L^2_{loc}(0, +\infty); L^2(0, S))$. Then by means of (3) we have that

$$\langle x(t), \psi \rangle = \langle x_0, \psi \rangle - \int_0^{\bar{s}} \int_0^{t \wedge s} c(t-\tau, s-\tau) \, \mathrm{d}\tau \, \mathrm{d}s + \int_0^{t \wedge s} \int_0^{\bar{s}} c(t-s, \tau) \, \mathrm{d}\tau \, \mathrm{d}s$$

But, changing the variables in the first of the integral above we have

$$\int_0^{\bar{s}} \int_0^{t \wedge s} c(t-\tau, s-\tau) \, \mathrm{d}\tau \, \mathrm{d}s = \int_0^t \int_0^{\bar{s}} c(s, \tau) \, \mathrm{d}\tau \, \mathrm{d}s = \int_0^t \int_0^{\bar{s}} c(t-s, \tau) \, \mathrm{d}\tau \, \mathrm{d}s$$

So, from the previous two equations:

$$\langle x(t), \psi \rangle = \langle x_0, \psi \rangle .$$

The claim for a general $c \in \mathcal{U}_{x_0}$ follows by density and by continuity of the operator $\mathcal{S}$ defined in Proposition A.1. □

It is a well known fact (see Bensoussan et al. (2007) Section II.3.1, pp. 201-204)[8] that if $x(t)$ is a solution of (6) then it is also a *weak solution* of the same equation, where by weak we mean that the the left and right hand sides of (5) are equal when evaluated at any test function $p \in D^2$, with D^2 defined in (7), which is stated in the following proposition.

Proposition A.2 *Let x be the solution to (6) when $x_0 \in \Pi$, $c \in \mathcal{U}_{x_0}$. Let also T be any finite horizon and p be any function in D^2. Then*

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} \langle x(t), p \rangle = \langle x(t), A^* p \rangle - \langle c(t), p \rangle + \langle \delta_0, p \rangle \langle \psi, c(t) \rangle , & \forall t \in (0, T] \\ \langle x(0), p \rangle = \langle x_0, p \rangle = 1. \end{cases}$$

Proof of Lemma 3.7 A stationary couple $(\tilde{x}, \tilde{c})$ needs, by definition, to satisfy, for any $t \geq 0$, the following identity:

$$\tilde{x} = T(t)\tilde{x} + \int_0^t T(t-s)B\tilde{c}\,\mathrm{d}s.$$

Taking a generic real regular ϕ compactly supported in $(0, S)$ we have

$$\langle \tilde{x}, \phi \rangle_{L^2} = \left\langle T(t)\tilde{x} + \int_0^t T(t-s)B\tilde{c}\,\mathrm{d}s, \phi \right\rangle_{L^2} .$$

Observe that in particular $\phi(S) = 0$ so we have that $\phi \in D$. We consider the derivative with respect to t of both the sides of the previous expression and we compute it in $t = 0$. We obtain

$$0 = \langle A^*\phi, \tilde{x} \rangle_{L^2} + \langle \phi, B\tilde{c} \rangle_{D \times D'} = \langle A^*\phi, \tilde{x} \rangle - \langle \phi, \tilde{c} \rangle$$

(where last equality holds since $\phi(0) = 0$). Since $A^*\phi$ is exactly the derivative of ϕ we have proved that, in each stationary couple $(\tilde{x}, \tilde{c})$, $\tilde{c}$ is the distributional derivative of $-\tilde{x}$ in the specified sense. From this fact and since $\tilde{c}$ is a positive measure, we have that $-\tilde{x}$ needs to be (a part for a set of zero (Lebesgue) measure) an increasing function and $\tilde{c}$ is exactly (Theorem 346.1 in Ziemer (2004)) its Stieltjes derivative.

On the other side, assume that $x \in \Pi$, $-\tilde{x}$ is an increasing function on $[0, S]$ and $\tilde{c}$ its Stieltjes derivative. Given a $\phi \in D$ we compute, for any $t \geq 0$,

$$\frac{d}{dt} \left\langle T(t)\tilde{x} + \int_0^t T(t-s)B\tilde{c}\,\mathrm{d}s, \phi \right\rangle_{L^2}$$

If we prove that it is always equal to zero we have the claim. The previous derivative equals

$$\left\langle \tilde{x}, A^* e^{tA^*}\phi \right\rangle_{L^2} + \left\langle B\tilde{c}, e^{tA^*}\phi \right\rangle_{D' \times D}$$
$$= \int_0^S \partial(e^{tA^*}\phi)(r)x(r)\mathrm{d}r - (e^{tA^*}\phi)(0) \int_0^S \partial x(r) + \int_0^S (e^{tA^*}\phi)(r)\partial x(r)$$

[8]More precisely one has to repeat, as indicated at p. 204, the construction at p. 203 with $k \in D^2$ of the weak solution.

where we denoted by $\partial(e^{tA^*}\phi)$ the derivative of $e^{tA^*}\phi$ and we used, in the integral, that $\tilde{c} = -\partial\tilde{x}$ (in the Stieltjes sense, the notation most diffused in the literature to express this fact is $\tilde{c} = \mu_{\tilde{x}}$). Using the integration by part formula for Stieltjes integrals (see e.g. Hewitt (1960)), the previous expression is equal to

$$\int_0^S \partial(e^{tA^*}\phi)(r)x(r)\mathrm{d}r - \Big((e^{tA^*}\phi)(0)(x(S)-x(0))\Big) \\ + \left((e^{tA^*}\phi)(S)x(S) - (e^{tA^*}\phi)(0)x(0) - \int_0^S \partial(e^{tA^*}\phi)(r)x(r)\mathrm{d}r\right)$$

that vanishes because, for any $t \geq 0$, $x(S) = (e^{tA^*}\phi)(S) = 0$. This gives the claim. □

A.2 Proofs for Section 4

Proof of Proposition 4.2. Note that $\mathcal{A}_\rho \neq \emptyset$ is a consequence of the fact that $g_\rho(\cdot)$ is continuous and possibly non-zero only on the compact set $[\lambda, \bar{s}]$, while (8) implies $0 \notin \mathcal{A}$ for any $\rho \geq 0$. Now let $M_{\rho_A} \in \mathcal{A}_{\rho_A}$ and $M_{\rho_B} \in \mathcal{A}_{\rho_B}$ be arbitrarily chosen.

Now, note that the support function h_ρ defined in (17) is an exponential function, increasing and convex, so that $\rho_B < \rho_A$ implies h_{ρ_A} is definitively greater than h_{ρ_B} for increasing values of ρ. Hence, $h_{\rho_A}(0) = h_{\rho_B}(0) = 0$ and: either 1) $h_{\rho_A}(s) > h_{\rho_B}(s)$ for all $s \in (0, S]$, or 2) there exists $\widetilde{s} \in (0, +\infty)$ such that $h_{\rho_A}(\widetilde{s}) = h_{\rho_B}(\widetilde{s})$, $h_{\rho_A}(s) < h_{\rho_B}(s)$ for all $s \in (0, \widetilde{s})$, and $h_{\rho_A}(s) > h_{\rho_B}(s)$ for all $s \in (\widetilde{s}, +\infty)$. Nonetheless, the former never takes place, as g_{ρ_B} is maximal at M_{ρ_B} implies

$$h_{\rho_A}(M_{\rho_A}) - h_{\rho_B}(M_{\rho_A}) = (e^{\rho_B M_{\rho_B}} - 1)(g_{\rho_B}(M_{\rho_A}) - g_{\rho_B}(M_{\rho_B})) \leq 0$$

so that

$$h_{\rho_A}(M_{\rho_A}) \leq h_{\rho_B}(M_{\rho_A}), \text{ with } M_{\rho_A} > 0.$$

As a consequence,

$$M_{\rho_A} \in (0, \widetilde{s}]$$

Similarly, from the maximality of g_{ρ_A} at M_{ρ_A}, ore derives $h_{\rho_A}(M_{\rho_B}) \geq h_{\rho_B}(M_{\rho_B})$,which implies $\widetilde{s} \leq S$ and

$$M_{\rho_B} \in [\widetilde{s}, S]$$

and, since the selections M_{ρ_A} and M_{ρ_B} were arbitrarily chosen in $\mathcal{A}_{\rho_A}$ and $\mathcal{A}_{\rho_B}$ respectively, then the first asserton in (i) is proven. If in addition f is differentiable at $\widetilde{s}$, assume by contradiction that $\mathcal{A}_{\rho_A} \cap \mathcal{A}_{\rho_B} = \{\widetilde{s}\}$. Then by Remark 4.3 $f(\widetilde{s}) = h_{\rho_A}(\widetilde{s}) = h_{\rho_B}(\widetilde{s})$, and $h'_{\rho_A}(\widetilde{s}) > h'_{\rho_B}(\widetilde{s})$ which contradicts the fact that the graph of f lies underneath the graph of both support functions.

Next we prove (ii). The fact that the selection $\rho \mapsto M_\rho$ is nonincreasing is a direct consequence of (i). Now note that from $M_{\rho_A} \leq M_{\rho_B}$ and the convexity of $h_{\rho_B}(s)$ (see also Figure 1) follows

$$\frac{h_{\rho_B}(M_{\rho_B})}{M_{\rho_B}} \geq \frac{h_{\rho_B}(M_{\rho_A})}{M_{\rho_A}} \tag{33}$$

while Remark 4.1 implies

$$f(M_{\rho_B}) = h_{\rho_B}(M_{\rho_B}) \text{ and } \; h_{\rho_B}(M_{\rho_A}) \geq f(M_{\rho_A}) \tag{34}$$

so that (33) and (34) give

$$\frac{f(M_{\rho_B})}{M_{\rho_B}} \geq \frac{f(M_{\rho_A})}{M_{\rho_A}}.$$

The last claim in (*ii*) is a consequence of (*i*) and of countability of the discontinuities of a decreasing function (see e.g. Chung (2001) page 4).

The limit m_0 in (*iii*) exists and is contained in $[0, S]$, as any selection M_ρ in $\mathcal{A}_\rho$ is nonincreasing and there contained. Note that by continuity of f any $\mathcal{A}_\rho$ necessarily has a positive minimum, and that $\mathcal{A}_\rho \cap \mathcal{A}_0$ contains at most one element, which implies $m_0 \leq \min \mathcal{A}_0$. Suppose by contradiction that $m_0 < \min \mathcal{A}_0$. Then $h_0(s) - f(s)$ is always strictly positive on $[\lambda, m_0]$. Moreover, if we define

$$k_\rho(s) = \frac{f(\min \mathcal{A}_0)}{e^{\rho \min \mathcal{A}_0} - 1}(e^{\rho s} - 1). \tag{35}$$

we can observe that

1. $h_\rho(s) \geq k_\rho(s)$,
2. $k_\rho(s) - h_0(s)$ converges uniformly to 0 on $[\lambda, m_0]$ when $\rho \to 0$

So there exists $\hat{\rho}$ small enough such that,

$$h_{\hat{\rho}}(s) \geq k_{\hat{\rho}}(s) > f(s) \tag{36}$$

for any $s \in [\lambda, m_0]$. This implies $\mathcal{A}_{\hat{\rho}} \subset (m_0, \min \mathcal{A}_0]$ that contradicts the hypothesis, and (*iii*) is proved.

□

Prior to demonstrate some of the theorems in Section 4 we need a series of preliminary results, which we state and prove hereby.

Lemma A.3 *Given $a, b \in [0, S]$, $a \leq b$, one has:*

$$T(t)\chi_{[a,b]} = \chi_{[a+t,(b+t)\wedge S]}, \quad \textit{and} \quad \int_0^t T(\tau)\delta_b \, \mathrm{d}\tau = \chi_{[b,(b+t)\wedge S]} \tag{37}$$

Proof. The first assertion follows from

$$T(t)\chi_{[a,b]}(s) = \chi_{[a,b]}(s-t)\chi_{[t,S]}(s) = \chi_{[a+t,b+t]}(s)\chi_{[t,S]}(s) = \chi_{[a+t,(b+t)\wedge S]}(s).$$

For the second, note that, if ϕ is any test function in D, one has $\langle \delta_b, T^*(\tau)\phi \rangle = \phi(b+\tau)$ if $b + \tau \leq S$, and 0 otherwise, so that, by changing the variable in the integral with $\sigma = \tau + b$, one obtains

$$\left\langle \int_0^t T(\tau)\delta_b \, \mathrm{d}\tau, \phi \right\rangle = \int_0^t \langle \delta_b, T^*(\tau)\phi \rangle \; \mathrm{d}\tau = \int_b^{(b+t)\wedge S} \phi(\sigma) \, \mathrm{d}\sigma = \langle \chi_{[b,(b+t)\wedge S]}, \phi \rangle$$

which implies the claim. □

Lemma A.4 *Given $g \in L^2_{loc}([0,+\infty);\mathbb{R})$, $a \in [0,S)$, one has, for any $t \in [0, S-a]$, $s \in [0,S]$:*

$$\left(\int_0^t g(\tau)T(t-\tau)\delta_a \,\mathrm{d}\tau\right)(s) = \chi_{[a,t+a]}(s)g(t+a-s). \tag{38}$$

Proof. The proof is a consequence of the previous Lemma and arguing by density after applying (38) to approximating step-functions. □

Lemma A.5 *Assume $\rho \geq 0$, and x_ρ, c_ρ p_ρ are defined by means of* (19) (20) (23) *respectively. Assume* (9) (8). *Consider any trajectory x of system* (5) *starting at $x_0 \in \Pi$, and driven by a control $c \in \mathcal{U}_{x_0}$. Then, for all $t \geq 0$*

$$\langle c(t) - c_\rho, f - p\rangle \leq \rho \langle x(t) - x_\rho, p_\rho\rangle - \langle x(t) - x_\rho, A^* p_\rho\rangle . \tag{39}$$

Proof. From Proposition 4.2 we derive that there exists M_ρ in the nonempty set $\mathcal{A}_\rho$. Moreover, by definition of p_ρ one has that $f(s) \leq p_\rho(s)$ for all $s \in [0,\bar{s}]$, and $f(M_\rho) = p_\rho(M_\rho)$, so so that for all positive measure $c \in D'$, one derives $\langle f, c\rangle \leq \langle p_\rho, c\rangle$ with the equality holding at $c = \gamma\delta_{M_\rho}$, with γ any nonnegative constant. In particular, for $\gamma = 1$, one has

$$\langle c(t), f - p_\rho\rangle \leq 0 = \langle c_\rho, f - p_\rho, \rangle, \quad \forall t \geq 0 \tag{40}$$

Now we recall that by means of Proposition 3.3, $x(t)$ lies in $L^2(0,S)$, so that the duality pairing with $x(t)$ coincides with scalar product with $x(t)$ in $L^2(0,S)$, moreover $A^* p_\rho = {p_\rho}'$ so that

$$-\rho \langle x(t), p_\rho\rangle + \langle x(t), A^* p_\rho\rangle = -\rho \int_0^{\bar{s}} p_\rho(s) x(t,s)\,\mathrm{d}s + \int_0^{\bar{s}} {p_\rho}'(s) x(t,s)\,\mathrm{d}s =: \Delta(\rho). \tag{41}$$

When $\rho > 0$, we have , *supp* $x(t) \subset [0,\bar{s}]$, ${p_\rho}'(s) = \rho\eta_\rho e^{\rho s}$ on $[0,\bar{s}]$, and Proposition 3.4 holds, so that

$$\Delta(\rho) = \rho\eta_\rho \int_0^{\bar{s}} x(t,s)\,\mathrm{d}s = \rho\eta_\rho \int_0^{\bar{s}} x_0(s)\,\mathrm{d}s = \rho\eta_\rho \tag{42}$$

that is, the quantity on the left hand side is constant for all trajectories (i.e., for all admissible controls c, and all initial data x_0 covering a unitary area). In particular the property holds true when x is set equal to x_ρ. Then the claim follows by means of (40), (41) and (42). We proceed similarly for the case $\rho = 0$, as this time $p_0'(s) = \beta_0$ on $[0,\bar{s}]$, so that

$$\Delta(0) = \int_0^{\bar{s}} p_0'(s) x(t,s)\,\mathrm{d}s = \beta_0 \int_0^{\bar{s}} x(t,s)\,\mathrm{d}s = \beta_0,$$

which leads to the same conclusion. □

Corollary A.6 *In the assumption of Lemma A.5, and for u satisfying (9), set $\beta_\rho := f(M_\rho)/M_\rho$, and $\alpha_\rho := u'(\beta_\rho)$. Then for all $t \geq 0$, the value-loss function defined in (24) satisfies*

$$\theta_\rho(c(t), x(t)) \equiv u(\beta_\rho) - u(\langle c(t), f\rangle) + \alpha_\rho\big[\rho\langle x(t) - x_\rho, p_\rho\rangle - \langle x(t), A^* p_\rho\rangle + \langle c(t), p_\rho\rangle\big] \geq 0$$

Proof. For all c in D', define $h(c) := u(\langle c, f\rangle)$, so that $h(c_\rho) = u(\langle c_\rho, f\rangle) = u(\beta_\rho)$, moreover h is differentiable with $h'(c) = u'(\langle c, f\rangle)f \in D$, $h'(c_\rho) = u'(\beta_\rho)f = \alpha_\rho f$. Since h is concave, we have $u(\langle c, f\rangle) - u(\beta_\rho) \leq \alpha_\rho \langle c - c_\rho, f\rangle$, for all $c \in D'$. Then Lemma A.5 implies

$$u(\langle c(t), f\rangle) \leq u(\beta_\rho) + \alpha_\rho \big[\langle c(t) - c_\rho, p_\rho\rangle + \rho \langle x(t) - x_\rho, p_\rho\rangle - \langle x(t) - x_\rho, A^* p_\rho\rangle \big],$$

and, to complete the proof, we need to show that $-\langle c_\rho, p_\rho\rangle + \langle x_\rho, A^* p_\rho\rangle = 0$. This holds true when $\rho > 0$ as

$$-\langle c_\rho, p_\rho\rangle + \langle x_\rho, A^* p_\rho\rangle = -\frac{\eta_\rho}{M_\rho}\left(e^{\rho M_\rho} - 1\right) + \frac{\rho\eta_\rho}{M_\rho}\int_0^{M_\rho} e^{\rho s}\,\mathrm{d}s = 0.$$

When instead $\rho = 0$, the claim follows from

$$\langle x_\rho, A^* p_0\rangle = \int_0^{\bar{s}} x(t,s)\beta_0 ds = \beta_0 = \langle c_\rho, p_0\rangle. \tag{43}$$

□

Remark A.7 Note that the proofs of the previous results remain true when c_ρ is replaced by a positive multiple $\gamma\delta_{M_\rho}$ of the Dirac's delta at M_ρ. If moreover $\rho = 0$, (39) holds true for a general initial datum x_0 in place of x_ρ. Indeed $\langle x, A^* p_0\rangle = \beta_0$ for all $x \in L^2(0, S)$ having support in $[0, \bar{s}]$ and unitary extension $\int_0^{\bar{s}} x(s)ds = 1$. We summarize these facts in the following generalized version of Lemma A.5 for the case $\rho = 0$.

Corollary A.8 *Let be $\rho = 0$. Consider any trajectory $x(t)$ starting at $x_0 \in \Pi$ and driven by a control $c \in \mathcal{U}_{x_0}$. Then, for all $t \geq 0$, $\gamma \geq 0$,*

$$\langle c(t) - \gamma\delta_{M_0}, f - p_0\rangle \leq -\langle x(t) - x_0, A^* p_0\rangle .$$

Corollary A.9 *In the assumption of Corollary A.6*

$$U_T(c_\rho) - U_T(c) \geq \alpha_\rho \left(\langle x_\rho - x_0, p_\rho\rangle - e^{-\rho T}\langle x_\rho - x(T), p_\rho\rangle\right),$$

Proof. From Corollary A.6 we derive

$$u(\beta_\rho) - u(\langle c(t), f\rangle) \geq \alpha_\rho\big[\rho \langle x_\rho - x(t), p_\rho\rangle - \langle c(t), p_\rho\rangle + \langle x(t), A^* p_\rho\rangle \big] = e^{\rho t}\frac{\mathrm{d}}{\mathrm{d}t}\left\langle x(t) - x_\rho, e^{-\rho t} p_\rho\right\rangle.$$

which promptly implies the thesis. □

Lemma A.10 *Assume x_ρ, p_ρ are defined by means of (19)(23) respectively, and let $\rho > 0$. Consider any trajectory x of system (5) starting at x_ρ, and driven by a control $c \in \mathcal{U}_{x_\rho}$. Then*

$$\lim_{T\to+\infty}\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\left[\left\langle x_\rho - x(t), e^{-\rho t} p\right\rangle\right]\,\mathrm{d}t = 0.$$

Proof. Note that

$$\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\left[\left\langle x_\rho - x(t), e^{-\rho t} p\right\rangle\right]\,\mathrm{d}t = e^{-\rho T}\left\langle x_\rho - x(T), p_\rho\right\rangle$$

so that recalling that $x(T)$ and x_ρ are supported in $[0,\bar{s}]$ and that (13) holds, one gets

$$\left|e^{-\rho T}\left\langle x_\rho - x(T), p_\rho\right\rangle\right| \le e^{-\rho T}\int_0^S |x_\rho(s) - x(T,s)||p_\rho(s)|ds \le 2e^{-\rho T}\eta_\rho\left(e^{\rho S}-1\right) \xrightarrow{T\to+\infty} 0$$

which implies the claim. □

Proof of Theorem 4.5. We need to show that (c_ρ, x_ρ) satisfies Definition 3.6. Note that by means of (4) we have

$$Bc_\rho = -\frac{1}{M_\rho}\delta_{M_\rho} + \langle \frac{1}{M_\rho}\delta_{M_\rho}, \psi\rangle \delta_0 = \frac{1}{M_\rho}\left(\delta_0 - \delta_{M_\rho}\right)$$

so that, by making use of (37) one obtains

$$\begin{aligned} T(t)x_\rho + \int_0^t T(t-\tau)Bc_\rho \,\mathrm{d}\tau &= \frac{1}{M_\rho}\chi_{[t,(t+M_\rho)\wedge S]} + \frac{1}{M_\rho}\int_0^t T(\sigma)(\delta_0 - \delta_{M_\rho})\,\mathrm{d}\sigma \\ &= \frac{1}{M_\rho}\chi_{[t,(t+M_\rho)\wedge S]} + \frac{1}{M_\rho}\left(\chi_{[0,t\wedge S]} - \chi_{[M_\rho,(M_\rho+t)\wedge S]}\right) = \frac{1}{M_\rho}\chi_{[0,M_\rho]} = x_\rho \end{aligned}$$

which implies the thesis.

Now we need to prove that when A_ρ is singleton then the unique optimal stationary couple is given by the golden rule. The claim in the case $\rho = 0$ will follow as a consequence of Theorem 5.9.

Assume then $\rho > 0$, and $(\tilde{x},\tilde{c})$ is an optimal stationary couple.

We show first that $supp(\tilde{c}) = \{M_\rho\}$. In particular we show that $supp(\tilde{c}) \cap [0, M_\rho) = \emptyset$ as, by a similar argument, it may be proven that $supp(\tilde{c}) \cap (0,\bar{s}] = \emptyset$.

Assume by contradiction that $supp(\tilde{c}) \cap [0, M_\rho) \neq \emptyset$ and define, for $\epsilon > 0$, the control[9]

$$c_\epsilon(t)(s) := \begin{cases} (1-\epsilon)\chi_{[0,M_\rho)}(s)\tilde{c}(s) + \chi_{(M_\rho,\bar{s}]}(s)\tilde{c}(s) + \epsilon\delta_{M_\rho}\int_0^t \tilde{c}(M_\rho - s)\,\mathrm{d}s & t\in[0,M_\rho) \\ (1-\epsilon)\chi_{[0,M_\rho)}(s)\tilde{c}(s) + \chi_{(M_\rho,\bar{s}]}(s)\tilde{c}(s) + \epsilon\delta_{M_\rho}\int_0^{M_\rho}\tilde{c}(s)\,\mathrm{d}s & t\in[M_\rho,+\infty) \end{cases} \tag{44}$$

One can easily see that that c_ϵ is admissible at $\tilde{x}$. If we show that $\frac{\mathrm{d}U(c_\epsilon)}{\mathrm{d}\epsilon}|_{\epsilon=0} > 0$, such fact

[9] Observe that $\tilde{x}$ is a decreasing function (and then of bounded variation) so the derivative $\tilde{c} = -\partial x$ is a Radon measure. For this reason the integral $\int_0^{M_\rho}\tilde{c}(s)\,\mathrm{d}s$ appearing in (44) needs to be understood as a Lebesgue-Stieltjes integral (see e.g. Ash (2000) Section 1.5 page 35). Such notation is also common in Physics.

implies that for ϵ small enough, $U(c_\epsilon) > U(\tilde{c})$ and, a s a consequence, $\tilde{c}$ is not optimal. Then

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\epsilon} U(c_\epsilon)\Big|_{\epsilon=0} &= \frac{\mathrm{d}}{\mathrm{d}\epsilon}\left(\int_0^{+\infty} e^{-\rho t} u\left(\langle c_\epsilon(t), f\rangle\right) \mathrm{d}t\right)\Big|_{\epsilon=0} \\
&= u'\left(\langle \tilde{c}, f\rangle\right)\Bigg(\int_0^{+\infty} e^{-\rho t}\Big\langle -\tilde{c}\chi_{[0,M_\rho]} + \delta_{M_\rho}\Big[\chi_{[0,M_\rho)}(t)\int_0^t \tilde{c}(M_\rho - s)\,\mathrm{d}s \\
&\qquad + \chi_{[M_\rho,+\infty)}(t)\int_0^{M_\rho}\tilde{c}(s)\,\mathrm{d}s\Big], f\Big\rangle \,\mathrm{d}t\Bigg) \\
&= u'\left(\langle \tilde{c}, f\rangle\right)\Bigg[\int_0^{+\infty} e^{-\rho t}\left\langle -\tilde{c}\chi_{[0,M_\rho]}, f\right\rangle \mathrm{d}t + \\
&\quad + \int_0^{M_\rho} e^{-\rho t} f(M_\rho)\left(\int_0^t \tilde{c}(M_\rho - s)\,\mathrm{d}s\right)\mathrm{d}t + \left(\int_{M_\rho}^{+\infty} e^{-\rho t} f(M_\rho)\,\mathrm{d}t\right)\left(\int_0^{M_\rho}\tilde{c}(s)\,\mathrm{d}s\right)\Bigg]
\end{aligned}$$

Integrating with respect to t the first and the third addenda and integrating by parts the second, we obtain

$$\begin{aligned}
&= u'\left(\langle \tilde{c}, f\rangle\right)\Bigg[-\frac{1}{\rho}\left\langle f, \tilde{c}\chi_{[0,M_\rho]}\right\rangle + \Bigg(-\frac{f(M_\rho)}{\rho}e^{-\rho M_\rho}\int_0^{M_\rho}\tilde{c}(s)\,\mathrm{d}s \\
&\qquad + \frac{f(M_\rho)}{\rho}\int_0^{M_\rho} e^{-\rho t}\tilde{c}(M_\rho - t)\,\mathrm{d}t\Bigg) + \frac{f(M_\rho)}{\rho}e^{-\rho M_\rho}\int_0^{M_\rho}\tilde{c}(s)\,\mathrm{d}s\Bigg] \\
&= u'\left(\langle \tilde{c}, f\rangle\right)\Bigg[-\frac{1}{\rho}\left\langle f, \tilde{c}\chi_{[0,M_\rho]}\right\rangle + \frac{f(M_\rho)}{\rho}\int_0^{M_\rho} e^{-\rho t} c(M_\rho - t)\,\mathrm{d}t\Bigg] \\
&= u'\left(\langle \tilde{c}, f\rangle\right)\Bigg[-\frac{1}{\rho}\left\langle f, \tilde{c}\chi_{[0,M_\rho]}\right\rangle + \frac{f(M_\rho)}{\rho}e^{-\rho M_\rho}\int_0^{M_\rho} e^{\rho s}\tilde{c}(s)\,\mathrm{d}s\Bigg]
\end{aligned}$$

that, recalling (22) and (23), can be rewritten as

$$\begin{aligned}
&= u'\left(\langle \tilde{c}, f\rangle\right)\Bigg[-\frac{1}{\rho}\left\langle f, \tilde{c}\chi_{[0,M_\rho]}\right\rangle + \frac{f(M_\rho)e^{-\rho M_\rho}}{\rho}\int_0^{M_\rho}\left(\frac{p_\rho(s)}{\eta_\rho(s)} + 1\right) c(s)\,\mathrm{d}s\Bigg] \\
&= \frac{u'\left(\langle \tilde{c}, f\rangle\right)}{\rho}\left[-\left\langle f, \tilde{c}\chi_{[0,M_\rho]}\right\rangle + \left(1 - e^{-\rho M_\rho}\right)\left\langle p_\rho, \tilde{c}\chi_{[0,M_\rho]}\right\rangle + f(M_\rho)e^{-\rho M_\rho}\left\langle \psi, \tilde{c}\chi_{[0,M_\rho]}\right\rangle\right] \\
&= \frac{u'\left(\langle \tilde{c}, f\rangle\right)}{\rho}\left[\left\langle p_\rho - f, \tilde{c}\chi_{[0,M_\rho]}\right\rangle + e^{-\rho M_\rho}\left\langle f(M_\rho)\psi - p_\rho, \tilde{c}\chi_{[0,M_\rho]}\right\rangle\right].
\end{aligned}$$

Since $u' > 0$ and $p_\rho \geq f$, the last expression is greater or equal than

$$\geq \frac{u'\left(\langle \tilde{c}, f\rangle\right)}{\rho}\left[e^{-\rho M_\rho}\left\langle f(M_\rho)\psi - p_\rho, \tilde{c}\right\rangle\right] > 0$$

where the strict positivity of last inequality follows from the following facts: (i) $p_\rho(s) \geq f(M_\rho)$ for all $s \in [0, M_\rho]$, (ii) $p_\rho(s) < f(M_\rho)$ for all $s \in [0, M_\rho)$, (iii) $supp(\tilde{c}) \cap [0, M_\rho) \neq \emptyset$. This shows that $\frac{\mathrm{d}}{\mathrm{d}\epsilon}\big|_{\epsilon=0} U(c_\epsilon) > 0$ and than that $\tilde{c}$ is not optimal at $\tilde{x}$.

Using a similar argument one may prove that, if $supp(\tilde{c}) \cap (0, \bar{s}] \neq \emptyset$, then $(\tilde{x}, \tilde{c})$ cannot be optimal as well, so that necessarily $supp(\tilde{c}) = \{M_\rho\}$. Since the only probability measures whose support is $\{M_\rho\}$ are of type $\tilde{c} = \alpha\delta_{M_\rho}$, for some real number $\alpha \geq 0$. In order for $\tilde{x}$ to be in Π, such α needs to be chosen equal to $1/M_\rho$, that is

$\tilde{c} = c_\rho$ defined in (20) and then that $\tilde{x} = x_\rho$ defined in (19), as we were meant to prove. □

Proof of Theorem 4.6. We make use of definition 3.5 to show that c_ρ is optimal. Let U_T be defined by means of (10) and let c be any control admissible at the initial datum x_ρ, and let $x(t) = x_{x_\rho,c}(t)$ be the associated trajectory. Define

$$\Gamma(T) := U_T(c_\rho) - U_T(c) + \alpha_\rho \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} \left[\left\langle x_\rho - x(t), e^{-\rho t} p_\rho \right\rangle \right] \mathrm{d}t =$$

$$= \int_0^T e^{-\rho t} \big[u(\beta_\rho) - u\left(\langle c(t), f\rangle\right) + \alpha_\rho \big(\rho \left\langle x_\rho - x(t), p_\rho \right\rangle - \langle x(t), A^* p_\rho\rangle + \langle c(t), p_\rho\rangle \big) \big] \,\mathrm{d}t$$

$$- \int_0^T e^{-\rho t} \langle c(t), \psi\rangle \langle \delta_0, p_\rho\rangle \,\mathrm{d}t. \quad (45)$$

where the last equality is obtained by applying Proposition A.2. By means of Corollary A.6 the first addendum in the right hand side is positive, while the second is null, due to the fact that $\langle p, \delta_0\rangle = 0$. Hence $\Gamma(T) \geq 0$, which by means of Lemma A.10 implies

$$\liminf_{T\to+\infty} \left(U_T(c_\rho) - U_T(c)\right) \geq 0,$$

as required. □

Proof of Remark 4.7. By linearity, $\tilde{x}$ satisfies Definition 3.6 and hence is a stationary program. What is left to show is that $(\tilde{c}, \tilde{x})$ is optimal. Let c be any control admissible at $\tilde{x}$ and let x be the associated trajectory. Then using concavity, optimality of c^i_ρ and (45), one gets

$$\liminf_{T\to+\infty} \left(U_T(c_\rho) - U_T(c)\right) \geq \sum_{i=1}^n \lambda_i \liminf_{T\to+\infty} \left(U(c^i_\rho) - U_T(c)\right) \geq 0.$$

□

Proof of Lemma 4.9.

It is straightforward from definition that $\hat{x}$ is M_ρ-periodic. Then we need to show that $\hat{x}$ solves (27), more precisely that

$$\langle \hat{x}(t) - T(t)x_0, \varphi\rangle = \left\langle \int_0^t T(t-\tau) B\hat{x}(\tau, M_\rho)\, \delta_{M_\rho} \mathrm{d}\tau, \varphi \right\rangle, \quad \forall \varphi \in D. \quad (46)$$

Since $\hat{x}$ is M_ρ-periodic, we may assume $t \in (0, M_\rho)$. We note that $\langle T(t)x_0, \varphi\rangle_{D',D} = \int_t^S x_0(s - t)\varphi(s)\,\mathrm{d}s$ while

$$\langle \hat{x}(t), \varphi\rangle_{D',D} = \int_0^S \hat{x}(t,s)\ \varphi(s)\mathrm{d}s = \int_t^M x_0(s-t)\varphi(s)\,\mathrm{d}s + \int_0^t x_0(s + M - t)\varphi(s)\,\mathrm{d}s$$

so that the left hand side in (46) may be rewritten as follows

$$\langle \hat{x}(t) - T(t)x_0, \varphi \rangle_{D',D} = -\int_M^S x_0(s-t)\varphi(s)\,\mathrm{d}s + \int_0^t x_0(M-\tau)\varphi(t-\tau)\,\mathrm{d}\tau \tag{47}$$

On the other hand $\widehat{x}(t, M) = x_0(M-t)$, $\hat{c}(t) = x_0(M-t)\delta_{M_\rho}$, and $\langle \delta_{M_\rho}, \psi \rangle = \psi(M) = 1$ so that

$$B\hat{x}(\tau, M_\rho)\,\delta_{M_\rho} = \hat{x}(\tau, M_\rho)\,\delta_{M_\rho} + \langle \hat{x}(\tau, M_\rho)\,\delta_{M_\rho}, \psi \rangle\,\delta_0 = x_0(M-t)(\delta_0 - \delta_{M_\rho})$$

and the right hand side in (46) is

$$\begin{aligned}\int_0^t \langle\, B\hat{x}(\tau, M_\rho)\,\delta_{M_\rho}, T^*(t-\tau)\varphi \rangle\,\mathrm{d}\tau &= \int_0^t\; x_0(M-t)\,\langle (\delta_0 - \delta_{M_\rho}), T^*(t-\tau)\varphi \rangle\,\mathrm{d}\tau \\ &= \int_0^t\; x_0(M-t)\,([T^*(t-\tau)\varphi]\,(0) - [T^*(t-\tau)\varphi]\,(M_\rho))\,\mathrm{d}\tau \\ &= \int_0^t x_0(M-\tau)\varphi(t-\tau)\,\mathrm{d}\tau - \int_{M+t-S}^t x_0(M-\tau)\varphi(t-\tau+M)\,\mathrm{d}\tau\end{aligned}$$

which is equal, by means of a change of variables, to the right hand side in (47). □**Proof of Lemma 4.12**. Recalling that $\widehat{x}$ is periodic of period M_ρ, and that (Lemma 4.9) $\widehat{x}(t, M_\rho) = x_0(M_\rho - \sigma(t))$, one has $\langle \widehat{c}(t), f \rangle = \widehat{x}(t, M_\rho)f(M_\rho) = f(M_\rho)x_0(M_\rho - \sigma(t))$, so that, once set $n = [T/M_\rho]$, by suitable changes of variables we have (note that if $T < M_\rho$ then the first sum is null)

$$\begin{aligned}U_T(\widehat{c}) &= \sum_{i=0}^{n-1} e^{-\rho i M_\rho} \int_0^{M_\rho} e^{-\rho t} u\,(f(M_\rho)x_0(M_\rho - t))\,dt + e^{-\rho n M_\rho} \int_0^{T-nM_\rho} e^{-\rho t}\,u\,(f(M_\rho)x_0(M_\rho - t))\,dt \\ &= \frac{1 - e^{-n\rho M_\rho}}{e^{\rho M_\rho} - 1} U_1^\rho + e^{-\rho(n+1)M_\rho} U_2^\rho(T)\end{aligned}$$

When $\rho = 0$, similarly

$$U_T(\widehat{c}) = \sum_{i=0}^{n-1} \int_0^{M_\rho} u\,(\,f(M_\rho)x_0(M_\rho - t))\,dt + \int_0^{T-nM_\rho} u\,(f(M_\rho)x_0(M_\rho - t))\,dt = nU_1^0 + e^{-\rho M_\rho} U_2^0(T)$$

□

A.2.1 Optimality of good controls

In order to prove Proposition 4.15, we need some preliminary results, contained in the following lemmata.

Lemma A.11 *Assume (8) and (9) are satisfied. Let $x_0 \in \Pi$ and $T > 0$. Then there exists a positive constant B_T depending on T (and not depending on c and x_0), such that*

$$U_T(c) \le B_T, \quad \forall c \in \mathcal{U}_{x_0},\; \forall x_0 \in \Pi.$$

Proof. We prove the fact in the case $\rho = 0$, as in the case $\rho > 0$ the result holds a fortiori. Set $\varepsilon := \min\{\lambda, \sigma_1\}$. We have $\psi \geq \chi_{[0,\lambda]}$, so that using Proposition 3.4 and expanding $x(\varepsilon)$ by means of (6) we obtain

$$\left\langle T(\varepsilon)x_0, \chi_{[0,\lambda]} \right\rangle + \int_0^\varepsilon \left\langle c(\tau), B^* T^*(\varepsilon - \tau)\chi_{[0,\lambda]} \right\rangle \mathrm{d}\tau = \langle x(\varepsilon), \chi_{[0,\lambda]} \rangle \leq \langle x(\varepsilon), \psi \rangle = 1. \tag{48}$$

Now by definition of T^* we have

$$[T^*(\varepsilon - \tau)\chi_{[0,\lambda]}](s) = \chi_{[0,\lambda]}(s + \varepsilon - \tau)\chi_{[0,S-(\varepsilon-\tau)]}(s) = \chi_{[0,\lambda-(\varepsilon-\tau)]}(s),$$

and by definition of B^*

$$B^* T^*(\varepsilon - \tau)\chi_{[0,\lambda]} = -\chi_{[0,\lambda-(\varepsilon-\tau)]} + \langle \delta_0, \chi_{[0,\lambda-(\varepsilon-\tau)]} \rangle \psi = -\chi_{[0,\lambda-(\varepsilon-\tau)]} + \psi$$

which imply, together with the fact that the support of c lies in $[0, \bar{s}]$, the following inequality

$$\langle c(\tau), B^* T^*(\varepsilon - \tau)\chi_{[0,\lambda]} \rangle = \langle c(\tau), \chi_{[\lambda-(\varepsilon-\tau),\bar{s}]} \rangle \geq \langle c(\tau), \chi_{[\lambda,\bar{s}]} \rangle, \quad \forall \tau \leq \varepsilon$$

Then observing that $\langle T(\varepsilon)x_0, \chi_{[0,\lambda]} \rangle \geq 0$, from the latter and (48) we derive

$$\int_0^\varepsilon \left\langle c(\tau), \chi_{[\lambda,\bar{s}]} \right\rangle \mathrm{d}\tau \leq 1$$

Now we iterate the argument: we use the fact that $x(t) = T(t - r)x(r) + \int_r^t T(t - \tau)Bc(\tau)\,\mathrm{d}\tau$ with $t = r + \varepsilon$, $r = n\varepsilon$ and $n \in \{0, 1, .., [T/\varepsilon]\}$, and obtain similarly that

$$\int_{n\varepsilon}^{(n+1)\varepsilon} \left\langle c(\tau), \chi_{[\lambda,\bar{s}]} \right\rangle \mathrm{d}\tau \leq \int_{n\varepsilon}^{(n+1)\varepsilon} \left\langle c(\tau), \chi_{[\lambda-((n+1)\varepsilon-\tau),\bar{s}]} \right\rangle \mathrm{d}\tau \leq 1$$

so that

$$\int_0^T \left\langle \chi_{[\lambda,\bar{s}]}, c(s) \right\rangle \mathrm{d}s \leq \sum_{n=0}^{[T/\varepsilon]} \int_{n\varepsilon}^{(n+1)\varepsilon} \left\langle c(\tau), \chi_{[\lambda,\bar{s}]} \right\rangle \mathrm{d}\tau \leq \frac{T}{\varepsilon} + 1. \tag{49}$$

Since u is concave, there exist real constants a and b such that $u(q) \leq a + bq$ for all $q \in \mathbb{R}^+$, and such that $b \geq \max_{s\in[\lambda,\bar{s}]} f(s)$ so that $b\chi_{[\lambda,\bar{s}]} \geq f$. Then one has

$$u(\langle c(t), f \rangle) \leq a + b \langle c(t), f \rangle \leq a + b \left\langle c(t), \chi_{[\lambda,\bar{s}]} \right\rangle$$

and hence by (49)

$$U_T(c) = \int_0^T u(\langle c(t), f \rangle)\,\mathrm{d}t \leq aT + b\left(1 + \frac{T}{\varepsilon}\right) =: B_T,$$

with B_T is independent of the chosen control and on the initial datum.

□

We here recall a notion that will be useful in the sequel, that of *local modulus.*[10]. A local modulus is a continuous function $\omega\colon \mathbb{R}^+\times\mathbb{R}^+ \to \mathbb{R}^+$ such that for all $b>0$, $\lim_{a\to 0^+}\omega(a;b)=0$. Throughout this section we will denote by $\omega(\cdot;\cdot)$ any function having these properties, or by $\omega(\cdot)$, if there is no explicit dependence from a parameter b.

Note that for $\rho = 0$ the value-loss function defined in (24) and in Corollary A.6 satisfies

$$\theta(c) \equiv \theta_0(c,x) = -[u(\langle c,f\rangle) - u(\langle \bar{c},f\rangle) - u'(\langle \bar{c},f\rangle)\langle c-\bar{c},f\rangle] \\ = u(\beta_0) - u(\langle c,f\rangle) + \alpha_0 \langle c-\bar{c},p_0\rangle \quad (50)$$

which is positive or null due to the fact that u is a concave function. In particular $\theta(\bar{c}) = 0$.

Remark A.12 Given a control $c \in \mathcal{U}_{x_0}$ and the associated trajectory x, and $x \in \Pi$ (and then, in particular, along an admissible trajectory $x(t)$ driven by an admissible control c), (43), implies $\alpha_0 \langle x, A^* p_0\rangle = \alpha_0\beta_0$ so that

$$\theta(c) = u(\beta_0) - u(\langle c,f\rangle) + \alpha_0\langle c,p_0\rangle - \alpha_0\langle x, A^*p_0\rangle$$

It is also straightforward that θ s a continuous function, indeed

$$|\theta(c)-\theta(c_1)| \le |u(\langle c,f\rangle) - u(\langle c_1,f\rangle)| + \alpha_0|c-c_1|_{D'}|p_o|_D \le \omega_\theta(|c-c_1|_{D'})$$

for some modulus ω_θ, for any $c, c_1 \in D'$ (note indeed that u is a uniformly continuous function).

Lemma A.13 *Given a good control $c(\cdot) \in \mathcal{U}_{x_0}$ then the following limit exists and is finite*

$$L_c := \lim_{T\to\infty} \int_0^T \theta(c(t))\,\mathrm{d}t \in [0,+\infty).$$

Remark A.14 As a consequence of the previous lemma, for any fixed positive constant A, we have

$$\int_{t-A}^{t} \theta(c(\tau))\,\mathrm{d}\tau \le \omega(1/t)$$

for a suitable modulus ω, that is, the integral is infinitesimal as t tends to $+\infty$.

Proof. Since $\theta(c(t)) \ge 0$, the sequence $T \mapsto \Delta_T = \int_0^T \theta(c(t))\,\mathrm{d}t$ is positive and increasing and therefore the limit exists, and is positive. Now observe that Proposition A.2 implies

$$\Delta_T = U_T(\bar{c}) - U_T(c) + \alpha_0 \int_0^T \langle c(t),p_0\rangle - \langle x(t), A^*p_0\rangle\ \mathrm{d}t \\ = U_T(\bar{c}) - U_T(c) - \alpha_0\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}[\langle x(t),p_0\rangle]\ \mathrm{d}t \\ = U_T(\bar{c}) - U_T(c) - \alpha_0\langle x(T)-x_0,p_0\rangle. \quad (51)$$

Recalling that c is good, by definition there exists $\theta \in \mathbb{R}$ such that for all T is $U_T(\bar{c}) - U_T(c) \le \theta$. Moreover, since $x(T) - x_0 \in L^2(0,S)$, in (51) duality pairing coincides with the scalar product in

[10] Actually in the definition of local modulus often other properties are required.

$L^2(0,S)$, and observing that p_0 lies in $L^\infty(0,S)$ (i.e. it is a bounded function), with $\|p_0\|_{L^\infty} \leq \beta_0 S$, by means of Hölder inequality one gets

$$|\langle x(T) - x_0, p_0\rangle| = |(x(T) - x_0|p_0)_{L^2(0,S)}| \leq \|p_0\|_{L^\infty}(\|x(T)\|_{L^1} + \|x_0\|_{L^1}) = 2\|p_0\|_{L^\infty},$$

so that

$$\Delta_T \leq \theta + 2\alpha_0\beta_0 S$$

and the proof is complete. □

Lemma A.15 *For any given x_0 and x_1 in Π, there exists a control $\hat{c} \in \mathcal{U}_{x_0}$, denoted with $\hat{c}(\cdot) = \hat{c}(\cdot; x_0, x_1, \bar{s})$ that drives the system from x_0 to x_1 in a time length less or equal to $\bar{s}$.*

Proof. We define

$$d^+(s) := (x_0(s) - x_1(s)) \vee 0, \qquad d^-(s) := (x_1(s) - x_0(s)) \vee 0, \qquad \text{for } s \in [0,S],$$

so that $d^+(s)$ (respectively, $d^-(s)$) is strictly positive at those points where x_0 is strictly bigger (respectively, smaller) than x_1. In some sense d^+ and d^- represent the values that need to be compensated by the choice of a suitable control. Since $\int_0^S x_0(s)\,\mathrm{d}s = \int_0^S x_1(s)\,\mathrm{d}s = 1$ we have that

$$J := \int_0^S d^+(s)\,\mathrm{d}s = \int_0^S d^-(s)\,\mathrm{d}s = \int_0^{\bar{s}} d^+(s)\,\mathrm{d}s = \int_0^{\bar{s}} d^-(s)\,\mathrm{d}s.$$

where the last equalities derive from the fact that both x_0 and x_1 are supported in $[0,\bar{s}]$.

If $J = 0$, then $x_0 = x_1$ and there is nothing to prove. Now assume $J > 0$. We define, for $t \in [0,\bar{s}]$

$$e^-(t) := \int_{\bar{s}-t}^{\bar{s}} d^-(\tau)\,\mathrm{d}\tau. \tag{52}$$

where e^- measures the mass of trees to be compensated, with age in the interval $[\bar{s}-t,\bar{s}]$. Note that $e^-(0) = 0$, $e^-(\bar{s}) = J$, and e^- is an increasing function. We define also, for $t \in [0,\bar{s}]$

$$D^+(t,s) := \begin{cases} 0, & s > \bar{s} \\ d^+(s-t), & s \in [t,\bar{s}] \\ d^+(s-t+\bar{s}), & s \in [0,t). \end{cases} \tag{53}$$

This function represents the translation with replanting of the exceeding part of the initial forest x_0. We want to prove that the following control satisfies the claim of the lemma in a time length of $\bar{s}$:

$$\hat{c}(t,\cdot) = \left(x_0(\bar{s}-t) - \frac{d^+(\bar{s}-t)e^-(t)}{J}\right)\delta_{\bar{s}} + d^-(\bar{s}-t)\frac{D^+(t,\cdot)}{J}, \qquad t \in [0,\bar{s}].$$

Note that $x_0, x_1 \in L^2(0,S)$ imply $d^+, d^- \in L^2(0,S)$, $D^+(t) \in L^2(0,S)$ for any $t \in [0,\bar{s}]$, and $t \mapsto D(T)$ belongs to $C(0,\bar{s};L^2(0,S))$. Moreover $e^- \in C(0,\bar{s})$, implies $\hat{c} \in L^2(0,\bar{s},D')$ (in fact, $c \in L^2(0,\bar{s},\mathcal{R})$). Besides this, the control $\hat{c}$, roughly speaking, acts as follows: (i) the term $x_0(\bar{s}-t)\delta_{\bar{s}}$ cyclicly cuts and replants at age 0 the trees reaching age $\bar{s}$ (it is indeed the Faustmann policy for $M = \bar{s}$); (ii) Note that if at time t the trees aged $\bar{s}$ are not enough (i.e. if $d^-(\bar{s}-t) > 0$), one needs to cut more trees in order to reach the level $x_1(\bar{s}-t)$. The term

$d^-(\bar{s}-t)\frac{D^+(s)}{J}$ does the job, taking from those ages where trees are exceeding the final target x_1 and represented by $D^+(t)$ (note that the trees (re)planted at time t are those having age $\bar{s}-t$ at time $\bar{s}$, the final time in which we want the forest to have the configuration x_1); (iv) The term $-\frac{d^+(\bar{s}-t)e^-(t)}{J}\delta_{\bar{s}}$ takes into account the fact that part of exceeding trees described in (ii) (for which d^+ is strictly positive) were already cut in $[0,t)$, so that one needs only to cut (by means of $\delta_{\bar{s}}$) the remaining ones.

Let us now give a formal proof of what stated above. We need to prove that the trajectory $\hat{x}(t)$ associated to the control $\hat{c}(t)$ satisfies $\hat{x}(\bar{s},s) = x_1(s)$ for all s. We have

$$\hat{x}(t) = I_1(t) + I_2(t) + I_3(t) := \left[T(t)x_0 + \int_0^t T(t-\tau)Bx_0(t-\tau)\delta_{\bar{s}}\,\mathrm{d}\tau\right]$$
$$-\frac{1}{J}\left[\int_0^t d^+(\bar{s}-\tau)e^-(\tau)T(t-\tau)B\delta_{\bar{s}}\,\mathrm{d}\tau\right] + \frac{1}{J}\left[\int_0^t d^-(\bar{s}-\tau)T(t-\tau)BD^+(\tau)\,\mathrm{d}\tau\right].$$

Recalling the Faustmann solution (see 26), we have $I_1(\bar{s}) = x_0$. Regarding $I_3(t)$, note that

$$T(t-\tau)BD^+(\tau) = T(t-\tau)\left\langle D^+(\tau),\psi\right\rangle\delta_0 - T(t-\tau)D^+(\tau) = J\delta_0 - T(t-\tau)D^+(\tau)$$

so that $I_3(t) = I_{31}(t) + I_{32}(t)$ with

$$I_{31}(t,s) = \int_0^t d^-(\bar{s}-\tau)T(t-\tau)\delta_0\,\mathrm{d}\tau = d^-(\bar{s}-t+s)\chi_{[0,t]}(s)$$

where the last equality is derived by means of (38), while

$$I_{32}(t) = -\int_0^t d^-(\bar{s}-\tau)T(t-\tau)\frac{D^+(\tau)}{J}\,\mathrm{d}\tau.$$

Now note that $T(t-\tau)D^+(\tau)(s) = D^+(\tau)(s-t+\tau)$ if $s-t+\tau \geq 0$ and 0 if $s-t+\tau < 0$, so that the last expression, evaluated at s, gives

$$I_{32}(t,s) = -\frac{1}{J}\int_{(t-s)\vee 0}^t d^-(\bar{s}-\tau)D^+(\tau)(s-t+\tau)\,\mathrm{d}\tau = -\frac{1}{J}\int_{(t-s)\vee 0}^{t\wedge(\bar{s}+t-s)} d^-(\bar{s}-\tau)D^+(\tau)(s-t+\tau)\,\mathrm{d}\tau.$$

By means of the definition of e^- and $D^+(\tau)$ given in (52) (53) respectively, the latter may be made explicit as follow

$$I_{32}(t)(s) = \begin{cases} -\frac{1}{J}\int_{t-s}^t d^-(\bar{s}-\tau)d^+(s-t+\bar{s})\,\mathrm{d}\tau = -\frac{1}{J}d^+(s-t+\bar{s})\left(e^-(t)-e^-(t-s)\right) & s\in[0,t) \\ -\frac{1}{J}\int_0^t d^-(\bar{s}-\tau)d^+(s-t)\,\mathrm{d}\tau = -\frac{1}{J}d^+(s-t)e^-(t) & s\in(t,\bar{s}] \\ -\frac{1}{J}\int_0^{(\bar{s}+t-s)} d^-(\bar{s}-\tau)d^+(s-t)\,\mathrm{d}\tau = -\frac{1}{J}d^+(s-t)e^-(\bar{s}+t-s) & s>\bar{s} \end{cases}$$

Regarding $I_2(t)$, since $-B\delta_{\bar{s}} = \delta_{\bar{s}} - \delta_0$, we may apply again (38) and derive

$$I_2(t,s) = \frac{1}{J}\left[d^+(s-t)e^-(t+\bar{s}-s)\chi_{[\bar{s},\bar{s}+t)}(s) - d^+(\bar{s}+s-t)e^-(t-s)\chi_{[0,t)}(s)\right]$$
$$= \begin{cases} \frac{1}{J}d^+(\bar{s}+s-t)e^-(t-s) & \text{if } s\in[0,t) \\ 0 & \text{if } s\in[t,\bar{s}] \\ \frac{1}{J}d^+(s-t)e^-(t+\bar{s}-s) & \text{if } s>\bar{s} \end{cases}$$

As a whole

$$\hat{x}(t,s) = \begin{cases} x_0(s) + d^-(\bar{s}-t+s) - d^+(\bar{s}-t+s)\frac{e^-(t)}{J} & \text{if } \ s \in [0,t) \\ x_0(s) - d^+(s-t)\frac{e^-(t)}{J} & \text{if } \ s \in [t,\bar{s}] \\ 0 & \text{if } \ s > \bar{s}. \end{cases}$$

so that at any time t the support of $x(t)$ is in $[0,\bar{s}]$, moreover $e^-(\bar{s}) = J$ implies at $t = \bar{s}$ that

$$\hat{x}(\bar{s},s) = \begin{cases} x_0(s) + d^-(s) - d^+(s) & \text{if } s \in [0,\bar{s}] \\ 0 & \text{if } s > \bar{s} \end{cases}$$

which turn to equal $x_1(s)$ by means of the definitions of d^+ and d^-. □

Now we are ready to prove that maximal (and the in particular any optimal) controls are good controls.

Proof of Proposition 4.15. Assume by contradiction that the maximal control c^* is not good, and denote by x^* the associated trajectory.

Then, given any $\theta \in \mathbb{R}$, there exists $T_\theta \geq 0$ with

$$U_{T_\theta}(c^*) - U_{T_\theta}(\bar{c}) < -\theta$$

Next we show that T_θ may be chosen arbitrarily large, for instance $T_\theta > 2\bar{s}$, if θ is chosen sufficiently large. Indeed by means of Lemma A.11 one has $\sup_{t\in[0,2\bar{s}]} |U_t(c^*) - U_t(\bar{c})| \leq B_{2\bar{s}}$ so that for $\theta > B_{2\bar{s}}$, we have $U_T(c^*) - U_T(\bar{c}) < -\theta$ only for values of T which are greater than $2\bar{s}$. But one such T does exist by assumption.

Hence we select $\theta > 2B_{\bar{s}} > B_{2\bar{s}}$ and $T_\theta > 2\bar{s}$ and define, with the notation of the previous Lemma, the following controls: $c_1(t) = c(t;x_0,\bar{x})$ stirring the system from x_0 to $\bar{x}$ in time $\bar{s}$, and $c_2(t) = c(t;\bar{x},x^*(T_\theta))$, stirring the system from $\bar{x}$ to $x^*(T)$ in time $\bar{s}$ and moreover

$$\tilde{c}(t) = \begin{cases} c_1(t) & \text{if } t \in [0,\bar{s}) \\ \bar{c} & \text{if } t \in [\bar{s}, T_\theta - \bar{s}) \\ c_2(t) & \text{if } t \in [T_\theta - \bar{s}, T_\theta) \\ c^*(t) & \text{if } t \geq T_\theta \end{cases}$$

We show that c^* is definitely overtaken by $\tilde{c}$, so that c^* cannot be maximal, yielding a contradiction. To do so it is enough to observe that, for any $T \geq T_\theta$, one has

$$\begin{aligned} U_T(\tilde{c}) - U_T(c^*) &= U_{T_\theta}(\tilde{c}) - U_{T_\theta}(c^*) = \\ &= U_{\bar{s}}(\tilde{c}) - U_{\bar{s}}(\bar{c}) + U_{T_\theta}(\bar{c}) - U_{T_\theta}(c^*) + \int_{T-\bar{s}}^{T} [u(\langle \tilde{c}(t), f\rangle) - u(\langle \bar{c}, f\rangle)]\,\mathrm{d}t \\ &\geq U_{\bar{s}}(\tilde{c}) - U_{\bar{s}}(\bar{c}) + \theta + U_{\bar{s}}(\tilde{c}(\cdot + T - \bar{s})) - U_{\bar{s}}(\langle \bar{c}, f\rangle) \geq \theta - 2B_{\bar{s}} > 0. \end{aligned}$$

□

A.3 Proofs for Section 5

A.3.1 Linear utility, positive discount

Proof of Theorem 5.1. Consider (28) when $u(r) = r$. Note that $\widehat{c}(t)$ coincides with c_ρ when $x_0 = x_\rho$, so that (28) applies also with (x_ρ, c_ρ) in place of $(x^*, \widehat{c})$. If $n = [T/M_\rho]$, $\sigma(t) = \{T/M_\rho\}M_\rho$ then

$$U_T(\widehat{c}) - U_T(c_\rho) = \eta_\rho(1 - e^{-\rho n M_\rho}) \int_0^{M_\rho} e^{\rho\sigma} (x_0(\sigma) - 1/M_\rho) + e^{-\rho(n+1)M_\rho} \int_{M_\rho - \sigma(T)}^{M_\rho} e^{\rho\sigma} (x_0(\sigma) - 1/M_\rho) .$$

Hence when $T \to +\infty$, and once set $\phi(t) = e^{\rho t}$, we derive

$$U(\widehat{c}) - U(c_\rho) = \lim_{T \to +\infty} \big(U_T(\widehat{c}) - U_T(c_\rho)\big) = \eta_\rho \langle x_0 - x_\rho, \phi\rangle = \langle x_0 - x_\rho, p_\rho\rangle. \tag{54}$$

Now let c be arbitrarily chosen in $\mathcal{U}_{x_0}$, with $x(t) = x(t; c, x_0)$ the associated trajectory. Let $T > 0$, and note that Corollary A.9 implies (for $u(r) = r$, is $\alpha_\rho = 1$)

$$U_T(c_\rho) - U_T(c) \geq -e^{-\rho T}\langle x_\rho - x(T), p_\rho\rangle + \langle x_\rho - x_0, p_\rho\rangle , \tag{55}$$

Coupling the previous relation with (54) we derive

$$U_T(\hat{c}) - U_T(c) \geq \sigma(T)$$

for a suitable function σ, $\sigma(T) \to 0$ as $T \to +\infty$, which implies the thesis. □

A.3.2 Linear Utility, Null discount

Average of a trajectory. Assume x is the trajectory associated to some initial datum x_0 and driven by an admissible control c. We denote by means of $x^A(t)$ the *average* of the trajectory over a time interval $[0, t]$ that is

$$x^A(t) := \frac{1}{t} \int_0^t x(s; x_0, c)\, \mathrm{d}s. \tag{56}$$

Lemma A.16 *Assume $\rho = 0$, $c \in \mathcal{U}_{x_0} \cap L^\infty([0, +\infty); D')$ a good control, and $x_0 \in \Pi$. Then*

$$\big\langle x^A(t), h\big\rangle \to \langle \bar{x}, h\rangle\, , \text{ as } t \to +\infty. \tag{57}$$

Proof. The proof is obtained by applying Theorem 9.1.3 in Zaslawski (2006) to the modified objective functional $\tilde{U}_T(c) = \int_0^T u\left(\langle c(t), f\rangle\right) - u(\langle c_\rho, f\rangle)\, \mathrm{d}t$, so that controls which are good for $\tilde{U}_T(c)$ satisfy Hypothesis (b) of Theorem 9.1.3. Note that the weak convergence derived in Zaslawski (2006) is in our case the weak convergence in the norm of D'. Note also that, since the control is bounded in D' and so is the trajectory, we can modify the functional $\tilde{U}_T(c)$ outside a ball in D' where both are contained in order to verify the coercivity assumption (1.9) page page 260 (although not recalled in the statement of the theorem, it is indeed needed). Eventually, using the arguments of Corollary A.6 we can see that Assumption 1 page 259 of Zaslawski (2006) is satisfied and the maximum (there cited as a minimum) is attained by the golden rule couple $(\bar{x}, \bar{c})$ (which is true *a fortiori* for the modified functional). □

Proof of Theorem 5.5. We divide the long proof into several steps.
Claim 1: $\hat{c}$ is a maximal control. We consider the trajectory $\hat{x} = x(\cdot; x_0, \hat{c})$, starting at x_0 and driven by the control $\hat{c}$. The control $\hat{c}$ is good. Indeed, we let $T > 0$ be arbitrarily fixed, and we apply Lemma 4.12 with $\rho = 0$ and $u(r) = r$ both to $\hat{c}$ and $\bar{c}$ deriving

$$U_T(\widehat{c}) - U_T(\overline{c}) = f(M) \int_{M-\sigma(T)}^{M} \left(x_0(\tau) - \frac{1}{M} \right) d\tau = f(M) \left[\int_{M-\sigma(T)}^{M} x_0(\tau) d\tau - \left\{ \frac{T}{M} \right\} \right] \geq -f(M)$$

which implies $\hat{c}$ is good. If by contradiction $\hat{c}$ is not maximal, then there exists a control $\tilde{c}$ in $\mathcal{U}_{x_0}$ and some $\hat{T}, a > 0$ such that for all $T \geq \hat{T}$

$$U_T(\hat{c}) - U_T(\tilde{c}) < -a. \tag{58}$$

Now assume $R \geq 3\hat{T}$. We integrate on $[0, R]$ and divide by R the previous inequality and obtain

$$\frac{1}{R} \int_0^R (U_T(\hat{c}) - U_T(\tilde{c}))\ \mathrm{d}T = \frac{1}{R} \int_0^{\hat{T}} (U_T(\hat{c}) - U_T(\tilde{c}))\ \mathrm{d}T + \frac{1}{R} \int_{\hat{T}}^R (U_T(\hat{c}) - U_T(\tilde{c}))\ \mathrm{d}T \tag{59}$$

where the first integral converges to 0 for $R \to \infty$, while the second is smaller than $-\frac{2}{3}a$ as a consequence of (58). Then for a sufficiently large R one has

$$\frac{1}{R} \int_0^R (U_T(\hat{c}) - U_T(\tilde{c}))\ \mathrm{d}T < -\frac{2}{3}a \tag{60}$$

On the other hand, if $\tilde{x}$ is the trajectory starting at x_0 and driven by control $\tilde{c}$, and $\tilde{x}^A(t)$ the respective average, from Corollary A.8, one has

$$U_T(\hat{c}) - U_T(\tilde{c}) \geq \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} \langle \hat{x}(t) - \tilde{x}(t), p \rangle\ \mathrm{d}t = \langle \hat{x}(T) - \tilde{x}(T), p \rangle \tag{61}$$

so that, integrating on $[0, R]$ and dividing by R, one gets

$$\frac{1}{R} \int_0^R (U_T(\hat{c}) - U_T(\tilde{c}))\ \mathrm{d}T \geq \left\langle \hat{x}^A(R) - \tilde{x}^A(R), p_0 \right\rangle \geq -\frac{1}{3}a$$

for a sufficiently large R, as $\left\langle \hat{x}^A(R) - \tilde{x}^A(R), p_0 \right\rangle \to 0$, as $R \to +\infty$, in view of (57): a contradiction.

Claim 2: the control $\hat{c}$ is not optimal. Assume c_1 is the control defined in (30), and $x_1(t, s) \equiv x_1(t, s; c_1, x_0)$, the associated trajectory. With reference to the notation there introduced, we assume also (and the reasons will be clear in a short while) that N is big enough so that $f(s_{N-1}) > 0$, which is true as f is continuous and $f(M) > 0$. Moreover we assume that x_0 satisfies the condition

$$\int_{s_{N-2}}^{s_{N-1}} x_0(r) dr > 0. \tag{62}$$

In order to prove $\hat{c}$ not optimal, it is sufficient to show that there exists $a > 0$ such that

$$U_{T_n}(\widehat{c}) - U_{T_n}(c_1) = -a, \quad \forall T_n = \frac{M}{N} + nM, \text{with } n \in \mathbb{N}.$$

At some initial time interval, that is, for t in $[0, M/N]$, we have

$$x_1(t,s) = \sum_{j=0}^{N} x_0(s_j + s - t)\chi_{[0,t]}(s) + x_0(s-t)\sum_{j=1}^{N}\chi_{[s_j,s_j+t]}(s)$$

Afterwards, the solution becomes periodic of period M, and repeatedly equal to

$$x_1(t,s) = \begin{cases} \chi_{\left[t-\frac{M}{N},t\right]}(s)\sum_{j=1}^{N} x_0(s_{j-1}+s+\frac{M}{N}-t) & t\in\left[\frac{M}{N},M\right] \\ \chi_{[0,t-M]}(s)\sum_{j=1}^{N} x_0(s_{j-1}+s+\frac{N+1}{N}M-t)+ & \\ \quad +\chi_{\left[t-\frac{M}{N},M\right]}(s)\sum_{j=1}^{N} x_0(s_{j-1}+s+\frac{M}{N}-t) & t\in\left[M,M+\frac{M}{N}\right] \end{cases}$$

(the general formula is obtained by replacing t with $\xi(t) = t - \left[\frac{M}{N}\right]M - \frac{M}{N}$ in the right hand side). Recalling that for any t in $[0, M/N]$, we have $x_1(t,s_j) = x_0(s_j - t)$, and that $\langle \delta_{s_j}, f\rangle = f(s_j)$

$$U_{\frac{M}{N}}(c_1) = \sum_{j=1}^{N} f(s_j)\int_0^{\frac{M}{N}} x_0(s_j-t)dt = \sum_{j=1}^{N} f(s_j)\int_{s_{j-1}}^{s_j} x_0(r)dr$$

By means of the periodicity of x_1 for $t \geq M/N$, we then derive

$$U_{T_n}(c_1) = U_{\frac{M}{N}}(c_1) + n\,f(M)\left[\int_{\frac{M}{N}}^{M} x_1(t,M)dt + \int_{M}^{M+\frac{M}{N}} x_1(t,M)dt\right]$$

Note that for $\frac{M}{N} < t < M$, we have $x_1(t,M) = 0$, while for $M < t < M + \frac{M}{N}$ we have $x_1(t,M) = \sum_{j=1}^{N} x_0(s_{j-1} + M + \frac{M}{N} - t)$ so that

$$\int_{M}^{M+\frac{M}{N}} x_1(t,M)dt = \sum_{j=1}^{N}\int_{M}^{M+\frac{M}{N}} x_0(s_{j-1}+M+\frac{M}{N}-t)dt = \sum_{j=1}^{N}\int_{s_{j-1}}^{s_j} x_0(r)dr = 1$$

which implies

$$U_{T_n}(c_1) = \sum_{j=1}^{N} f(s_j)\int_{s_{j-1}}^{s_j} x_0(r)dr + n\,f(M)$$

The difference between such utility and that yielded by means of the Faustmann policy $\widehat{c}$ is then

$$U_{T_n}(\widehat{c}) - U_{T_n}(c_1) = f(M)\int_{M-\frac{M}{n}}^{M} x_0(r)dr - \sum_{j=1}^{N} f(s_j)\int_{s_{j-1}}^{s_j} x_0(r)dr = -\sum_{j=1}^{N-1} f(s_j)\int_{s_{j-1}}^{s_j} x_0(r)dr =: -a \tag{63}$$

Note that $a > 0$, as the last term of the sum above is strictly positive in view of (62). As a consequence, $\widehat{c}$ is not optimal. The proof for the case when (62) is not satisfied is easily obtained by applying a control c_2 in place of c_1 shaped as follows. More precisely, if

$$m := \max\left\{ j \;:\; 1 \leq j \leq N-1,\; \int_{s_{j-1}}^{s_j} x_0(r)dr > 0 \right\}$$

(such a maximal index exists as the forest has positive density and extension 1), we define $\tau := \frac{M}{N}(N-1-m)$, c_2 as

$$c_2(t) = \bar{c}\,\chi_{[0,\tau]}(t) + c_1(t-\tau)\chi_{[\tau,+\infty)}(t) \tag{64}$$

that is, the control which coincides with $\bar{c}$ until the associated trajectory x_2 yields a positive integral (with respect to s) on $[s_{N-2}, s_{N-1}]$, and behaves like c_1 afterwards.

Claim 3: an optimal control does not exist. We assume by contradiction that $\tilde{c}(t) \in U_{x_0}$ is an optimal control. Then in particular, given any $\varepsilon > 0$, there exists T_ε such that

$$U_T(\tilde{c}) - U_T(\widehat{c}) \geq -\varepsilon \text{ and } U_T(\tilde{c}) - U_T(c_1) \geq -\varepsilon \qquad \forall T \geq T_\varepsilon. \tag{65}$$

On the other hand (63) implies , for a sufficiently small $\nu \in [0, M]$ not depending on n, that

$$U_T(c_1) - U_T(\widehat{c}) \geq \frac{a}{2}, \quad \forall T \in [T_n, T_n + \nu].$$

from which, if $n_\varepsilon \in N$ is such that $T_n > T_\varepsilon$ for all $n \geq n_\varepsilon$, we derive also

$$U_T(\tilde{c}) - U_T(\widehat{c}) \geq \frac{a}{2} - \varepsilon, \quad \forall T \in [T_n, T_n + \nu], \;\; \forall n \geq n_\varepsilon. \tag{66}$$

We show first that

$$\lim \inf_{n \to +\infty} \frac{1}{T_n + \nu} \int_0^{T_n+\nu} (U_T(\tilde{c}) - U_T(\hat{c}))\, dT \geq \frac{\nu a}{4}. \tag{67}$$

Indeed $\int_0^{T_n+\nu} (U_T(\tilde{c}) - U_T(\hat{c}))\, dT = A + B_n + C_n$, where

$$A \equiv \int_0^{T_{n_\varepsilon}} (U_T(\tilde{c}) - U_T(\hat{c}))\, dT$$

and in view of (65)

$$B_n \equiv \sum_{i=n_\varepsilon}^{n-1} \int_{T_i+\nu}^{T_{i+1}} (U_T(\tilde{c}) - U_T(\hat{c}))\, dT \geq -\varepsilon\, \nu\, (n - n_\varepsilon)$$

and in view of (66)

$$C_n \equiv \sum_{i=n_\varepsilon}^{n} \int_{T_i}^{T_i+\nu} (U_T(\tilde{c}) - U_T(\hat{c}))\, dT \geq \left(\frac{a}{2} - \varepsilon\right) \nu\, (n - n_\varepsilon + 1)$$

so that (recall that $T_n = nM + M/N$), if $o(1/n)$ is infinitesimal as $n \to \infty$, we have

$$\frac{1}{T_n + \nu} \int_0^{T_n+\nu} (U_T(\tilde{c}) - U_T(\hat{c}))\, dT \geq \frac{\nu}{M} \left(\frac{a}{2} - 2\varepsilon\right) + o(1/n)$$

so that, choosing $\varepsilon \leq a(2 - M/2)$, and passing to limits we obtain (67).

On the other hand, if we choose $\gamma = \widehat{x}(t, M)$ in Lemma A.8, we make use of Proposition A.2 (and recall that $p_0(s) = \beta_0 s\psi(s)$, $\langle \delta_0, p_0 \rangle = 0$, $A^* p_0(s) = \beta_0 \chi_{[0,S]}(s)$) we obtain

$$\begin{aligned} \langle \tilde{c}(t) - \widehat{c}(t), f \rangle &\leq \langle \tilde{c}(t) - \widehat{c}(t), p_0 \rangle - \langle \tilde{x}(t) - x_0, A^* p_0 \rangle & (68) \\ &= -\frac{d}{dt} \langle \tilde{x}(t) - \widehat{x}(t), p_0 \rangle - \langle \widehat{x}(t) - x_0, A^* p_0 \rangle = -\frac{d}{dt} \langle \tilde{x}(t) - \widehat{x}(t), p_0 \rangle & (69) \end{aligned}$$

where $\langle \widehat{x}(t) - x_0, A^* p_0 \rangle = 0$ follows from A.7. Then, for all T, one integrates the last inequality on $[0, T]$ and derives

$$U_T(\widetilde{c}) - U_T(\widehat{c}) \leq \int_0^T \frac{d}{dt} \langle \widehat{x}(t) - \tilde{x}(t), p_0 \rangle \, dt = \langle \tilde{x}(T) - \widehat{x}(T), p_0 \rangle \tag{70}$$

Now, we denote with $x^A(t)$ the average of the trajectory x, as defined in (56). Taking the integral on $[0, S]$ in both sides of (70) and dividing by S one has

$$\frac{1}{S} \int_0^S (U_T(\widetilde{c}) - U_T(\widehat{c})) \, dT \leq \left\langle \tilde{x}^A(S) - \widehat{x}^A(S), p_0 \right\rangle \tag{71}$$

Since $\widehat{c}$ and $\tilde{c}$ are good, by Lemma A.16 one has

$$\left\langle \tilde{x}^A(S) - \widehat{x}^A(S), p_0 \right\rangle \xrightarrow{S \to \infty} 0, \quad S \to \infty$$

which together with (71) contradicts (67). □

A.3.3 Strictly concave utility, null discount

Proof of Theorem 5.9. We set $H := L^2(0, S)$, noting that $D \hookrightarrow L^2(0, S)$ (with continuous inclusion) and use this property when necessary without further notice. Let $\varepsilon > 0$ be fixed. We have to prove that there exists $t(\varepsilon) > 0$, such that

$$i(t) := |x(t) - \bar{x}|_H \leq \varepsilon, \quad \text{for all } t \geq t(\varepsilon). \tag{72}$$

We will make use of the following decomposition of c.

We recall that by assumption M is the unique positive maximum point of $f(s)/s$ (which implies $p_0(s) - f(s) \geq 0$ for all $s \in [0, S]$, and with the equality holding only at $s = 0$ and $s = M$). Then, from the continuity of f and for a sufficiently small $\xi > 0$, there exists $\zeta(\xi) > 0$, with $\lambda < M - \zeta(\xi)$, such that $|s - M| \geq \zeta(\xi)$ implies $p(s) - f(s) \geq \xi$. Note that $\zeta(\xi) \xrightarrow{\xi \to 0} 0$. Furthermore, since c is a positive measure with $supp(c) \subseteq [\lambda, S]$, then

$$c = c_n + c_f, \text{ with } c_n = c\nu_\xi \textit{ and } c_f = c(1 - \nu_\xi),$$

where ν_ξ is a $[0, 1]$-valued smooth cut-off function with $\nu_\xi(s) \equiv 1$ for $|s - M| \leq \zeta(\xi)/2$ and $\nu_\xi(s) \equiv 0$ when $|s - M| \geq \zeta(\xi)$.

Now, since $t \to +\infty$, we may assume $t > S$. As a consequence, in (6) we have $T(t)x_0[s] = 0$ for all $s \in [0, S]$, and $T(t-\tau)Bc(\tau)[s] = Bc(\tau)[s - t + \tau]\chi_{[t-\tau,S]}[s] = 0$ for all $\tau \le t - S$, so that in (72)

$$x(t) - \bar{x} = \int_{t-S}^{t} T(t-\tau)Bc(\tau)\,\mathrm{d}\tau - \bar{x} = \int_{t-S}^{t} T(t-\tau)B\left(c(\tau) - \bar{c}\right)\,\mathrm{d}\tau$$
$$= I_1(t,\xi) + I_2(t,\xi) + I_3(t,\xi), \quad (73)$$

where

$$I_1(t,\xi) := \int_{t-S}^{t} T(t-\tau)Bc_f(\tau)\,\mathrm{d}\tau, \quad I_2(t,\xi) := \int_{t-S}^{t} T(t-\tau)B(c_n(\tau) - |c_n(\tau)|_{\mathcal{R}}\delta_M)\,\mathrm{d}\tau$$
$$\text{and } I_3(t,\xi) := \int_{t-S}^{t} T(t-\tau)B(|c_n(\tau)|_{\mathcal{R}}\delta_M - \bar{c})\,\mathrm{d}\tau.$$

In next steps we will estimate the H-norm of $I_1(t,\xi)$, $I_2(t,\xi)$ and $I_3(t,\xi)$.

Step 1: A preliminary estimate. We start by estimating the quantity $|c_f|_{D'}$. Given $x \in S$, and $c \in D'_+$, from Remark A.12 and from the concavity of u follows

$$\theta(c) = \theta(c) - \theta(\bar{c}) = u(\langle \bar{c}, f\rangle) - u(\langle c, f\rangle) + \alpha_0\langle p, c - \bar{c}\rangle$$
$$\ge -\alpha_0\langle c - \bar{c}, f\rangle + \alpha_0\langle c - \bar{c}, p_0\rangle = \alpha_0\langle c - \bar{c}, p_0 - f\rangle$$
$$= \alpha_0\int_0^S (p(s) - f(s))dc(s) \ge \alpha_0\xi|c_f|_{\mathcal{R}},$$

as c is positive. Observe that $D \hookrightarrow C([0,S])$ with continuous inclusion so that $\mathcal{R} \hookrightarrow D'$, and in particular $|c_f|_{D'} \le C|c_f|_{\mathcal{R}}$ for some fixed positive constant $C > 0$. Then from the previous inequalities follows that

$$|c_f|_{D'} \le C|c_f|_{\mathcal{R}} \le \frac{C}{\alpha_0\xi}\theta(c), \ \forall x \in \Pi, \ \forall c \in D'_+. \quad (74)$$

As observed in Remark A.12, the estimate does not depend on x.
Step 2: Estimate on $I_1(t,\xi)$. Note that $\|T(t)\|_{L(D')} \le 1$, so that (74) implies

$$|I_1(t,\xi)|_H \le \|B\|_{L(D')}\int_{t-S}^{t}|c_f(\tau)|_{D'}\,\mathrm{d}\tau \le \frac{C\|B\|_{L(D')}}{\alpha_0\xi}\int_{t-S}^{t}\theta(c(\tau))\,\mathrm{d}\tau$$

which implies that for some modulus ω_1 one has

$$|I_1(t,\xi)|_H \le \omega_1(1/t;\xi).$$

Step 3: Estimate on I_2. Given $\phi \in D$ one has

$$|\langle c_n - |c_n|_{\mathcal{R}}\delta_M, \phi\rangle| \le |c_n|_{\mathcal{R}}\max_{|s-M|\le\zeta(\xi)}|\phi(s) - \phi(M)| \le |c_n|_{\mathcal{R}}\int_{M-\zeta(\xi)}^{M+\zeta(\xi)}|\phi'(s)|\,\mathrm{d}s$$
$$\le \sqrt{2\zeta(\xi)}\,|c_n|_{\mathcal{R}}\left(\int_{M-\zeta(\xi)}^{M+\zeta(\xi)}|\phi'(s)|^2\,\mathrm{d}s\right)^{1/2} \le \sqrt{2\zeta(\xi)}\,|c_n|_{\mathcal{R}}|\phi|_D$$

which by definition gives

$$|(c_n - |c_n|_{\mathcal{R}}\delta_M)|_{D'} \leq \sqrt{2\zeta(\xi)}\, |c_n|_{\mathcal{R}} \tag{75}$$

so that

$$|I_2(t,\xi)|_H \leq \sqrt{2\zeta(\xi)}\, \|B\|_{L(D')}\, SK \leq \omega_2(\xi)$$

for some modulus ω_2 (with $\omega_2(\xi) \xrightarrow{\xi\to 0} 0$, uniformly with respect to t).

Step 4: Estimate on $I_3(t,\xi)$. In order to estimate $I_3(t,\xi)$ we need define, besides ξ, some other parameters. Since u is strictly convex and differentiable, recalling that $\alpha_0 = u'(\langle \bar{c}, p_0\rangle)$, $\beta_0 = \langle \bar{c}, p_0\rangle$ and defining $\beta_\eta = (1+\eta)\langle \bar{c}, p_0\rangle = (1+\eta)\beta_0$, one may consider $\gamma > 0$ and $0 < \eta < 1$ such that

$$\gamma = u(\beta_0) - u(\beta_\eta) + \alpha_0\eta\beta_0 > 0.$$

and moreover

$$\Delta = -[u'(\beta_\eta) - \alpha_0] < 0$$

since u' is strictly decreasing. Note that γ as a function of η is strictly increasing and attains the value zero at $\eta = 0$, so that its inverse $\eta(\gamma)$ is well defined and enjoys the same property, in particular $\eta(\gamma) \xrightarrow{\gamma\to 0} 0$. As a consequence, Δ may itself be regarded as a function of γ, with $\Delta(\gamma) \xrightarrow{\gamma\to 0} 0$.

Now we rewrite $I_3(t,\xi)$ as the sum of four terms and estimate them separately:

$$\begin{aligned} I_3(t,\xi) = \int_{t-S}^{t} T(t-\tau)B(|c_n(\tau)|_{\mathcal{R}}\delta_M - \bar{c})\,\mathrm{d}\tau &\equiv I_{31}(t,\xi) + I_{32}(t,\xi) + I_{33}(t,\xi) + I_{34}(t,\xi) \\ = \int_{t-S}^{t} \left(|c_n(\tau)|_{\mathcal{R}} - \frac{1}{M} - \eta - \frac{\theta(|c_n(\tau)|_{\mathcal{R}}\delta_M)}{\Delta}\right) & T(t-\tau)B\delta_M\,\mathrm{d}\tau \\ + \int_{t-S}^{t} \left(\frac{\theta(|c_n(\tau)|_{\mathcal{R}}\delta_M) - \theta(c(\tau))}{\Delta}\right) & T(t-\tau)B\delta_M\,\mathrm{d}\tau \\ + \int_{t-S}^{t} \left(\frac{\theta(c(\tau))}{\Delta}\right) T(t-\tau)B\delta_M\,\mathrm{d}\tau & + \eta\int_{t-S}^{t} T(t-\tau)B\delta_M\,\mathrm{d}\tau. \end{aligned} \tag{76}$$

To estimate $I_{34}(t,\xi)$ it suffices to observe that for every we fixed γ,

$$|I_{34}(t,\xi)|_H \leq \|B\||\delta_M|_{D'} S\eta =: \omega_{34}(\gamma;\xi) \tag{77}$$

where ω_{34} is a local modulus.

Next, to estimate $I_{33}(t,\xi)$ Remark A.14 implies

$$|I_{33}(t,\xi)|_H \leq \|B\||\delta_M|_{D'} \int_{t-S}^{t} \left|\frac{\theta(c(\tau))}{\Delta}\right| \mathrm{d}\tau \leq \omega_{33}(1/t;\gamma,\xi). \tag{78}$$

for some local modulus ω_{33}.

Then, to estimate $I_{32}(t,\xi)$, we note that

$$|\theta(c(\tau)) - \theta(c_n(\tau))| \leq \omega_\theta(|c_f|_{D'}) \leq \omega_\theta\left((C/\alpha_0\xi)\theta(c(\tau))\right)$$

so that, as a consequence of Remark A.14 for a sufficiently large t, the quantity $\theta(c(\tau))$ is infinitesimal, at least outside a subset of $[t-S,t]$ of arbitrarily small Lebesgue measure, so that one has

$$\int_{t-S}^{t} \frac{|\theta(c(\tau)) - \theta(c_n(\tau))|}{\Delta}\,\mathrm{d}\tau \leq \hat{\omega}(1/t;\gamma,\xi)$$

for some local modulus $\hat{\omega}$. Moreover, in view of (75) one has

$$|\theta(|c_n(\tau)|_{\mathcal{R}}\delta_M) - \theta(c_n(\tau))| \leq \omega_\theta(||c_n(\tau)|_{\mathcal{R}}\delta_M - c_n(\tau)||_{D'}) \leq \omega_\theta(\sqrt{2\zeta(\xi)}K)$$

so that

$$\int_{t-S}^{t} \frac{|\theta(|c_n(\tau)|_{\mathcal{R}}\delta_M) - \theta(c_n(\tau))|}{\Delta}\,\mathrm{d}\tau \leq \check{\omega}(\gamma;\xi),$$

for some modulus $\check{\omega}$. Hence, once set $\omega_{32} = \|B\||\delta_M|_{D'}\,(\hat{\omega}+\check{\omega})$, one derives

$$I_{32}(t,\xi) \leq \omega_{32}(1/t;\gamma,\xi). \tag{79}$$

We are left with the estimate on I_{31}. By definition of $\theta(c)$ and concavity of u

$$\begin{aligned} &-\theta(c) + \alpha_0\,\langle c-\bar{c},p_0\rangle + \gamma - \alpha_0\eta\,\langle \bar{c},p_0\rangle = \\ &\qquad = u\,(\beta_0) - u\,(\beta_\eta) \leq u'\,(\beta_\eta)\,\langle c-(1+\eta)\bar{c},f\rangle \end{aligned}$$

where we used first the definition of $\theta(c)$ and the strict concavity of u. Recalling that $\langle \bar{c},p\rangle = \langle \bar{c},f\rangle = \beta_0$, we obtain

$$\theta(c) \geq \gamma + \alpha_0\,\langle c,p_0\rangle - \alpha_0\beta_0 - \alpha_0\eta\beta_0 - u'(\beta_\eta)\,\langle f,c\rangle + u'(\beta_\eta)(1+\eta)\beta_0$$

so that, once set the expression above becomes

$$\theta(c) \geq \gamma - \eta\alpha_0\Delta + \langle c-\bar{c},\alpha_0 p - u'(\beta_\eta)f\rangle\,.$$

For $c \equiv |c_n(t)|_{\mathcal{R}}\delta_M$ the previous inequality reads

$$\varphi(\tau) := |c_n(t)|_{\mathcal{R}} - \frac{1}{M} - \eta - \frac{\theta(|c_n(t)|_{\mathcal{R}}\delta_M)}{\Delta} \leq 0.$$

Now note that, as a consequence of (73), step 2 and step 3, (76) (77) (78) (79), Hölder inequality and the fact that $\langle x(t)-\bar{x},\psi\rangle = 0$, one has

$$\begin{aligned} |\langle x(t)-\bar{x}-I_{31}(t,\xi),\psi\rangle| &= \left|\int_{t-S}^{t} \varphi(\tau)\,\langle T(t-\tau)B\delta_M,\psi\rangle\,\mathrm{d}s\right| + \\ &\qquad + |\langle I_1+I_2+I_{32}+I_{33}+I_{34},\psi\rangle| \leq \omega_{31}(1/t;\gamma,\xi), \end{aligned}$$

with $\omega_{31} = \sqrt{S}(\omega_1+\omega_2+\omega_{32}+\omega_{33}+\omega_{34})$. Now since

$$\langle T(t-\tau)B\delta_M,\psi\rangle = \langle \delta_0-\delta_M, T^*(t-\tau)\psi\rangle = \psi(t-\tau) - \psi(t-\tau+M)$$

the previous estimate may be rewritten as

$$\left|\int_{t-S}^{t} \varphi(\tau)(\psi(t-\tau)-\psi(t-\tau+M))\,\mathrm{d}\tau\right| \leq \omega_5(1/t;\gamma,\xi)$$

By definition of ψ we have $0 \leq \psi(t-\tau) - \psi(t-\tau+M) \leq 1$, moreover $\varphi(\tau) \leq 0$, so that the integrand of the last equation is always negative. Moreover, from the definition of $\psi(s)$ it is easily shown that on some interval $[t_1, t_2] \subseteq [0, S]$ one has $\psi(t-\tau) - \psi(t-\tau+M) \geq c$ for a suitable $c > 0$. As a consequence

$$c \int_{t_1}^{t_2} |\varphi(t-\sigma)| \, \mathrm{d}\sigma \leq \int_{t_1}^{t_2} |\varphi(t-\sigma)| \, |(\psi(\sigma) - \psi(\sigma+M))| \, \mathrm{d}\sigma \leq \omega_5(1/t; \gamma, \xi).$$

Now, since the previous equation holds for all t iterating the argument $[S/(t_2 - t_1)] + 1$ times, we obtain

$$\int_0^S |\varphi(t-\sigma)| \, \mathrm{d}\sigma \leq \frac{1}{c} \frac{S}{(t_2 - t_1)} \omega_5(1/t; \gamma, \xi) = \omega_4(1/t; \gamma, \xi).$$

We are now ready to draw the conclusion. It is sufficient to choose, in order, ξ, γ sufficiently small and $t(\varepsilon)$ sufficiently large to derive from all of the previous steps that $t \geq (\varepsilon)$ implies $i(t) \leq \varepsilon$, as we intended to show. $\square$

Proof of Theorem 5.10. In view of Proposition 4.15 it is sufficient that we compare $\bar{c}$ to good controls. Indeed for any good control c, Theorem 5.9 implies

$$\lim_{T \to +\infty} \int_0^T \frac{\mathrm{d}}{\mathrm{d}t} \left[\langle \bar{x} - x_{\bar{x},c}(t), \alpha_0 p_0 \rangle \right] \mathrm{d}t = \lim_{T \to +\infty} \langle \bar{x} - x_{\bar{x},c}(T), \alpha_0 p_0 \rangle = 0.$$

so that

$$\begin{aligned} \liminf_{T \to \infty} \big(U_T(\bar{c}) - U_T(c) \big) &= \liminf_{T \to \infty} \int_0^T \Big(U_T(\bar{c}) - U_T(c) + \frac{\mathrm{d}}{\mathrm{d}t} \left[\langle \bar{x} - x_{\bar{x},c}(t), \alpha_0 p_0 \rangle \right] \Big) \mathrm{d}t \\ &= \liminf_{T \to \infty} \int_0^T \left[u\left(\langle f, \bar{c} \rangle \right) - u\left(\langle f, c(t) \rangle \right) - \alpha_0 \langle x(t), A^* p_0 \rangle + \langle c(t), p_0 \rangle \right] \geq 0 \end{aligned}$$

where the last inequality follows from Corollary A.6. $\square$

Proof of Theorem 5.11 To prove (i) we first build the candidate optimal control $\tilde{c}$ as limit of a suitable sequence. We consider the quantity

$$S \equiv \sup_{c \in \mathcal{U}_{x_0}^{K,\lambda}} \left(\limsup_{T \to +\infty} \left[U_T(c) - U_T(\bar{c}) \right] \right), \tag{80}$$

(S possibly equal to $+\infty$). Let $\{c_n\}$ be a maximizing sequence in $\mathcal{U}_{x_0}^{K,\lambda}$, and let θ be the function defined in (50). Then for $T > 0$, we have

$$\begin{aligned} U_T(c_n) - U_T(\bar{c}) = -\int_0^T \left(\theta(c_n(t)) + \frac{\mathrm{d}}{\mathrm{d}t} \alpha_0 \langle x_n(t), p_0 \rangle \right) \mathrm{d}t \\ = -\int_0^T \theta(c_n(t)) \, \mathrm{d}t - \alpha_0 \langle x_n(T) - x_0, p_0 \rangle . \end{aligned} \tag{81}$$

Since $\|p_0\|_\infty < +\infty$ and $\|x_n(t)\|_{L^1} = \|x_0\|_{L^1} = 1$ then $|\alpha_0 \langle x_n(t) - x_0, p_0 \rangle| \leq 2\alpha_0 \|p_0\|_\infty$, so that, being $\theta(c_n(t))$ positive for all t, it may happen either (i) $\lim_{T \to +\infty} (U_T(c_n) - U_T(\bar{c})) = -\infty$, which we may exclude as $\{c_n\}$ is a maximizing sequence, or (ii) $\liminf_{T\to+\infty}(U_T(c_n) - U_T(\bar{c})) \geq -\infty$, the latter implying c_n is a good control. Note also that from (81) and the positivity of θ follows also

$$U_T(c_n) - U_T(\bar{c}) \geq -2\alpha_0\|p_0\|_\infty$$

implying that $S > +\infty$. Hence with no loss of generality, we may assume that c_n are good controls. Note also that Lemma A.13 and Theorem 5.9 imply that for any good control c the following limit exists and is finite:

$$\lim_{T\to+\infty}(U_T(c) - U_T(\bar{c})) = -\lim_{T\to+\infty}\int_0^T \left(\theta(c(t)) + \frac{\mathrm{d}}{\mathrm{d}t}\alpha_0\langle x(t), p_0\rangle\right)\mathrm{d}t = -L_c - \alpha_0\langle \bar{x} - x_0, p_0\rangle,$$

so that (80) implies

$$S = \lim_{n\to\infty}\lim_{T\to+\infty}[U_T(c_n) - U_T(\bar{c})]. \tag{82}$$

Now, set $h > 0$ and $L_h^2([0,+\infty); D')$ the Hilbert space of all functions $\phi\colon [0,+\infty) \to D'$ such that the norm $\int_0^{+\infty} e^{-ht}|\phi(t)|_{D'}^2\,\mathrm{d}t < +\infty$. It is a tedious but standard proof that $\mathcal{U}_{x_0}^{K,\lambda}$ is a sequentially weakly compact subset of $L_h^2([0,+\infty); D')$. Hence from $\{c_n(\cdot)\}$ one may extract a subsequence weakly converging to some $\tilde{c}(\cdot) \in L_h^2([0,+\infty); D')$ and $\tilde{c}(\cdot) \in \mathcal{U}_{x_0}^{K,\lambda}$.

Next we show that $\tilde{c}$ is optimal. Note that

$$\begin{aligned}\liminf_{T\to+\infty}(U_T(\tilde{c}) - U_T(c)) &\geq \liminf_{T\to+\infty}[U_T(\tilde{c}) - U_T(\bar{c})] - \limsup_{T\to+\infty}[U_T(c) - U_T(\bar{c})] \\ &\geq \liminf_{T\to+\infty}[U_T(\tilde{c}) - U_T(\bar{c})] - S\end{aligned}$$

so that it is enough to prove that $\liminf_{T\to+\infty}[U_T(\tilde{c}) - U_T(\bar{c})] \geq S$ to derive the positivity of the right hand side, and the conclusion. We start by proving that

$$\limsup_{T\to+\infty}[U_T(\tilde{c}) - U_T(\bar{c})] = S.$$

By definition of S, the left hand side is smaller than the right hand side, while the reverse inequality is obtained by observing that $c \mapsto \limsup_{T\to+\infty}[U_T(c) - U_T(\bar{c})]$ is a concave functional on $\mathcal{U}_{x_0}^{K,\lambda}$ (note that $\mathcal{U}_{x_0}^{K,\lambda}$ is a convex subset of $L_h^2([0,+\infty); D')$) so that passing to limits one obtains

$$\lim_n \lim_{T\to+\infty}[U_T(c_n) - U_T(\bar{c})] \leq S.$$

Note that, since $\{c_n(\cdot)\}$ is a maximizing sequence for the limsup, it is also maximizing for the liminf, more precisely

$$\sup_{c\in\mathcal{U}_{x_0}^{K,\lambda}}\left[\liminf_{T\to+\infty}(U_T(c) - U_T(\bar{c}))\right] = \sup_{c\in\mathcal{U}_{x_0}^{K,\lambda}}\left[\limsup_{T\to+\infty}(U_T(c) - U_T(\bar{c}))\right] = S.$$

Then arguing as before on the concave functional $c \mapsto \liminf_{T\to+\infty}[U_T(c) - U_T(\bar{c})]$, one derives

$$\liminf_{T\to+\infty}[U_T(\tilde{c}) - U_T(\bar{c})] = S$$

and the conclusion. □

(ii) Assume that c is a good admissible control, denote by $x(\cdot)$ the associated trajectory, set $R > 0$. Then

$$\begin{aligned}\frac{1}{R}\int_0^R (U_T(c) - U_T(\bar{c}))\,\mathrm{d}T &= \frac{1}{R}\int_0^R \int_0^T \theta(c(t)) + \frac{\mathrm{d}}{\mathrm{d}t}\alpha_0\,\langle x(t), p_0\rangle\,\mathrm{d}t\,\mathrm{d}T \\ &= \frac{1}{R}\int_0^R \int_0^T \theta(c(t))\,\mathrm{d}t\,\mathrm{d}T + \frac{\alpha_0}{R}\int_0^R [\langle x(T) - \bar{x}, p_0\rangle - \langle x_0 - \bar{x}, p_0\rangle]\,\mathrm{d}T.\end{aligned}$$

One one side, Lemma A.13 implies the first integral of the last expression converges to L_c when R goes to infinity, on the other side

$$\frac{1}{R}\int_0^R \langle x(T) - \bar{x}, p_0\rangle\,\mathrm{d}T = \left\langle \frac{1}{R}\int_0^R x(T)\,\mathrm{d}T - \bar{x}, p_0\right\rangle \to 0, \quad R \to \infty$$

as a consequence of Theorem 9.1.3 page 260 in Zaslawski (2006) (crf. proof of Proposition 5.5 for more details). Hence

$$\lim_{R\to+\infty} \frac{1}{R}\int_0^R (U_T(c) - U_T(\bar{c}))\,\mathrm{d}T = L_c + \alpha_0\,\langle \bar{x} - x_0, p_0\rangle\,,$$

so that the limit exists and is finite. We can now argue as in the proof of part (i) and prove that there exists a control $\tilde{c}(\cdot)$ that maximizes $\mathrm{limsup}_{R\to+\infty}\,\frac{1}{R}\int_0^R (U_T(c) - U_T(\bar{c}))\,\mathrm{d}T$ in $\mathcal{U}_{x_0}^{K,\lambda}$. So, given another control $c(\cdot)$ in $\mathcal{U}_{x_0}^{K,\lambda}$, we have

$$\begin{aligned}&\limsup_{R\to+\infty} \frac{1}{R}\int_0^R (U_T(\tilde{c}) - U_T(c))\,\mathrm{d}T \\ &\quad\geq \limsup_{R\to+\infty} \frac{1}{R}\int_0^R (U_T(\tilde{c}) - U_T(\bar{c}))\,\mathrm{d}T - \limsup_{R\to+\infty} \frac{1}{R}\int_0^R (U_T(c) - U_T(\bar{c}))\,\mathrm{d}T \geq 0\end{aligned}$$

and this implies that

$$\limsup_{T\to\infty} \big(U_T(\tilde{c}) - U_T(c)\big) \geq 0.$$

Since this is true for all c in $\mathcal{U}_{x_0}^{K,\lambda}$, $\tilde{c}$ is maximal in $\mathcal{U}_{x_0}^{K,\lambda}$. □

Remark A.17 The control that we have proved to be maximal is exactly the one that minimizes

$$\lim_{T\to\infty} \int_0^T \theta(c(t))\,\mathrm{d}t.$$