Adv. Math. Econ. 14, 39–67 (2011)

Advances in
MATHEMATICAL
ECONOMICS

An overview of turnpike theory: towards the discounted deterministic case

M. Ali Khan[1] **and Adriana Piazza**[2,*]

[1] Department of Economics, The Johns Hopkins University, Baltimore, MD 21218, USA
(e-mail: akhan@jhu.edu)

[2] Departamento de Matemática, Universidad Técnica Federico Santa María, Avda. España 1680, Casilla 110-V, Valparaíso, Chile
(e-mail: adriana.piazza@usm.cl)

Received: July 22, 2010
Revised: July 30, 2010

JEL classification: C62, D90, Q23

Mathematics Subject Classification (2010): 91B62, 49J45, 49O2

Abstract. In the last 5 years, there has been extensive work on the existence and characterization of solutions of undiscounted optimal programs in simple discrete-time models of the 'choice of technique' in development planning and of lumber extraction in the economics of forestry. In this expository essay, we present a unified treatment of the characterization results in two of these models. Furthermore, with an eye towards extensions to the discounted setting, we present the general theory, both with or without "smoothness hypotheses" on the felicity function, and in continuous-time and discrete-time taking special care to distinguish asymptotic convergence of optimal programs from their classical turnpike properties. We show

* A preliminary version of this work was presented at the *Workshop on Mathematical Economics* held in Keio University, Tokyo during November 13–15, 2009. The authors are grateful to Professor Toru Maruyama for his invitation, and to him and Professors Alexander Ioffe, Alejandro Jofre, Boris Mordukhovich and Haru Takahashi for discussion and encouragement. Adriana Piazza gratefully acknowledges the financial support of Programa Basal PFB 03, *Centro de Modelamiento Matemático (CMM), Universidad de Chile* and that of FONDECYT under project 11090254. Ali Khan would like to acknowledge, in addition, the hospitality of the Departamento de Matemática, *Universidad Técnica Federico Santa María*, of the *CMM* and *Facultad de Economía y Negocios* at the *Universidad de Chile* during his visit in July 2010, when the final version of this work was completed.

S. Kusuoka, T. Maruyama (eds.), *Advances in Mathematical Economics Volume 14,*
DOI: 10.1007/978-4-431-53883-7_3,

how the general results do not translate directly to the particular toy models examined here, and thereby suggest open problems for a more universal theory.

Key words: Intertemporal resource allocation, discount factor, optimal programs, asymptotic convergence, classical turnpike theory, neighborhood turnpike theorem, development planning, forest management

> We wish to describe briefly what we shall mean by qualitative properties. The number ρ is called the *discount rate,* but we do not insist that ρ be less than one. As will be seen, the magnitude of ρ makes no difference to the analysis to follow. One need not even use the Kuhn–Tucker Theorem and can do the whole thing with elementary calculus.[1]
>
> Gale (1970)

1. Introduction

Turnpike theory, even without adjectival qualifications, is a somewhat diffused and misunderstood subject that deters entry to outsiders, hampers interdisciplinary dialogue between economists and mathematicians, and with its attendant terminological proliferation, renders the evaluation of marginal contribution of ongoing research exceedingly difficult to determine. It is not at all clear where the non-initiated reader, wanting to learn about the state of the subject and keep up with current work, ought to go. McKenzie's entry on *turnpike theory* in the 1987 *Palgrave* [13] is dropped in the 2008 version [12], and a search of the word *turnpike* in the latter turns up ten entries, none of them explicitly devoted to the theory and four on individual authors: Gale, Nikaido, Samuelson and Solow, with McKenzie and Radner conspicuously missing. Substantive entries which discuss the term are on capital theory [Becker], golden-rule [Phelps], (multisectoral) growth models [Brock–Dechert], neoclassical growth theory [Manuelli], and Pontryagin's principle of optimality [Zelkina], but except for the entry of Brock–Dechert, in each of these it turns up in a passing and cursory way. And, in another search engine, extensive references are furnished in Google Scholar to the *Townsend turnpike* in the field of monetary economics, an appropriation of the Samuelsonian term that takes it away to a total different subject matter. It is clear that the subject could do with an overview and the stables with some cleaning.

A point to begin is perhaps Bewley's 2007 text on *General Equilibrium, Overlapping Generations Models and Optimal Growth Theory* in which there are two sections devoted to the turnpike theorem in the final chapter, and the subject is introduced as follows:

[1] For the precise references to these sentences in [15], see Footnote 34 below.

> The term *turnpike theorem* arose early in work on optimal growth theory. We may visualize the turnpike in a model in which the initial and terminal capital stocks are held fixed. Suppose that the initial and terminal periods are -1 and T, respectively, so that $\underline{K}_{-1}$ and $\underline{K}_T$ are the initial and terminal capital stocks. Radner (1961) showed that if T was large, then the optimal path of capital approaches a unique optimal stationary level, stays close to it for a large fraction of the $T + 2$ periods, and veers away toward the terminal stock only in the final periods. The image resembles a map of an interstate highway or turnpike with entrance and exit ramps.[2]

It is important to note that this part of the text is focussed on optimal finite-horizon programs, a focus that is relinquished in the subsequent sentences.

> In other work, researchers assume that the horizon is infinite, so that there is only an entrance ramp, and the theorem says that optimal paths asymptotically approach the unique stationary optimal path. The term *turnpike theorem* refers to any such assertion. Turnpike theorems give structure to optimal paths and focus attention on the stationary optimal one. Since convergence to the unique optimal stationary path is quite rapid in most fully specified examples, we are encouraged to think of modern economics as always near the stationary optimal state.

There are several lessons to be taken from this text. First, as already noted, the important distinction between finite-horizon and infinite-horizon resource allocation programs, and the corresponding inclusion of the latter with only an entrance ramp well within the rubric of turnpike theory. This inclusion is originally due to McKenzie,[3] and surely represents a substantial terminological over-reaching of the theory. Second, there is a focus on the uniqueness of the optimal stationary program, an assumption that is not necessary to turnpike theory, with or without an exit ramp, which at least according to McKenzie, can be viewed, rather broadly as the *bunching* or *clustering* of optimal trajectories around a certain benchmark set as the time horizon increases.[4] Third, the entire discussion is conducted in terms of

[2] See [4, pp. 542–543], also his Fig. 10.14 with time on the horizontal axis and capital stocks on the vertical.

[3] See [34] where McKenzie distinguishes three kinds of turnpikes: the *early, late* and *middle turnpikes.* In terms of this categorization, Bewley considers only the last two kinds of turnpikes. It is also worth pointing out that in [39], McKenzie uses a different classification scheme based on the *Samuelson* and *Ramsey turnpike.* It may be worth stating that in this essay, we are using the word as an adjective, as in "turnpike theory" or in "turnpike theorem", rather than as a noun.

[4] In this connection, see [25] for detailed references to the McKenzie's *oeuvre.*

discrete-time resource allocation problems, and the asymptotic implementation of the continuous-time results as being derived from their discrete-time counterparts is not raised. This must surely remain an important desideratum for the theory, and especially when the subject moves to its *raison d'etre* in computation. Finally, a complete silence as regards the discount factor being a crucial determinant of the shape of the theory; the independence of the dynamics on it.

As regards the discount factor, Bewley presents a result in the discounted setting on the global asymptotic stability of the solution to a one-sector infinite-horizon growth model due to Ramsey (1928), Cass (1965), Koopmans (1965) and Samuelson (1965).[5] He writes:

> A disappointing feature of growth theory is that this turnpike theorem does not generalize fully to models with many commodities or with uncertainty. The turnpike theorem with a positive discount rate, r, seems to require special assumptions in these interesting settings, and the generalizations are true only when r is sufficiently small. The turnpike theorems with $r = 0$ do generalize to models with many commodities and uncertainty. For this reason they probably should be thought of as the true turnpike theorems.[6]

And again.

> Although the theory of optimal allocations and programs becomes more difficult when utility is not discounted, the theory does become more robust. For instance, the turnpike theorem does not apply fully when there is more than one commodity and future utility is discounted, but it does apply when there are multiple commodities and there is no discounting.[7]

In his 1987 *Palgrave* entry, McKenzie had already noted that the "spirit of the original turnpike theorem is not well preserved in the aggregative model since the emphasis in the original theorem lies on the relative composition of the capital stock."

This emphasis on the spirit of the original theorem is precisely the point of departure for this essay. The original Samuelsonian motivation of turnpike theory was to cope with the indeterminacy of the terminal capital stock in finite-horizon resource allocation problems by arguing for its irrelevance

---

[5] Also referred to as the "aggregate model", or in the macroeconomic literature, as the Ramsey–Cass–Koopmans (RCK) model.

[6] See [4, p. 544] where Bewley refers to his Theorems 10.79 and 10.80 in this connection.

[7] See [4, p. 551] where the fact that the undiscounted "theory with multiple commodities" is not presented is footnoted. Also see a repetition of this statement on page 577.

in the setting of a "large enough" finite-horizon. The two turnpike theorems that Bewley presents for the undiscounted setting appropriate the term solely for the global asymptotic stability of solutions to infinite-horizon programming problems, and concern the "overlapping generations model" and the one-sector aggregate growth model.[8] To be sure, the optimal stationary program, being a solution to a static optimization problem, is a possible benchmark for either the finite- or the infinite-horizon problem, and the emphasis on an infinite-horizon problem is a complementary approach to the difficulty mentioned above of deciding on the 'optimal value of the terminal capital stock'.[9] Both approaches draw the relevance of the von Neumann maximal growth path, or the stationary state in the writings of the classical economists before him, to optimal solutions to resource allocation problems.[10] But there is an important analytical difference. In the first case, this relevance is established by turnpike theorems for solutions to finite-horizon problems, and in the second, by asymptotic stability theorems for solutions to infinite-horizon problems. The first involves explicit quantification of the unknowns, and is logically antecedent to the qualitative treatment involved in the second approach.[11] As we shall see in the sequel, this appropriation is very much in tune with the current research literature. This may be dismissed as a semantic issue hardly of any substantive import.[12] However, one of the reasons for striving for terminological clarification is also for the sake of communication: simply to be clear to oneself and to one's readers the precise subject matter that is being discussed. Furthermore, a rich terminology for asymptotic stability is already available in the classical theory of difference and differential equations.[13] Anyhow, what is especially interesting about Bewley's text is not its discussion of turnpike theory in the large but in its integration of static and dynamic general equilibrium theory, a 2007 follow-through perhaps of the

---

8 These are the theorems referred to in Footnote 6. Subsequent to their presentation, he re-emphasizes the meaning of the phrase "true turnpike theorems"; see [4, p. 577, paragraph 4].

9 See, for example, Gale's [14] justification for considering an infinite-horizon program, one that surely was in the mind of Ramsey, and goes back to [45].

10 In the context of Footnote 3, the word "turnpike" is being used as a noun.

11 For this qualitative–quantitative distinction, we follow Pietsch's monumental history of Banach spaces; see [44].

12 We shall have occasion to return to this blanket claim in the conclusion to this essay, and qualify it; see Footnote 36 below and the text it footnotes.

13 See the bibliography of [6] and [3] for references to, and the reliance on, this classical literature even in the context of the particular subject-matter we discuss here.

1958 Dosso text [11]. In a brief and cursory footnote [4, Footnote 5; p. 543], he defers to McKenzie (1976) for a history of the term and to the survey of the attendant literature.

And so, as a second to attempt to get at an overview of the theory, one can perhaps proceed to Fischer's *Palgrave* 1987 entry[14] on Paul Samuelson in [13]. In his fourth section on the categorization of Samuelson's work in capital theory, Fischer writes:

> The theory of capital and growth sections of the first four volumes of *CSP* account for 38 papers, the largest single category.[15] Although capital theory is the branch of economics most vulnerable to Samuelson's comparative technical advantage, and although both his earliest 1937 papers are placed in that category in *CSP* (I:17, I:20), the output in this area is concentrated in *CSP* (III) covering the years from 1965 to 1970. Solow (1983) [53] provides a fine review of this part of Samuelson's research, some of which he co-authored. As Solow (1983) notes, much of the capital theory in *CSP* is related to developments in Dorfman, Samuelson and Solow (1958), which itself grew out of a 1949 Samuelson three-part memorandum for the Rand Corporation.

And as is well-known, especially to readers of Lionel McKenzie, the formal study of the turnpike theorem – also referred to as the "turnpike conjecture" at the time – was launched with the publication of Dosso.[16] But to get back to Fischer:

> Notable among the contributions is a variety of turnpike theorems. A turnpike theorem is conjectured in the 1949 Samuelson memorandum, and fully worked out in the 1958 volume. The theorem states that for any accumulation programme, starting from an initial vector of capital goods, and with specified terminal conditions, as the horizon lengthens the optimal programme spends an increasing proportion of its time near the von Neumann ray; more generally in

[14] This is reprinted without change in [12].

[15] Fischer uses the abbreviation *CSP* to refer to the five volumes of Samuelson's *Collected Papers* with the Roman numeral referring to the particular volume, and the Arabic numeral referring to the chapter in it.

[16] See McKenzie [34] for a detailed discussion to Dosso; also Mckenzie's 1963 papers [31, 32] and his 1987 *Palgrave* entry. The 1963 papers are masterly introductions to the pre-Ramsey literature, and devoted solely to what is being termed the "classical turnpike theorem" in this essay. Their introductions are useful complements to Fischer's "fully worked out" phrase in the quotation below. Also see McKenzie's description of the "Hicks' pilgrimage" in [31, Footnote 1], and a terminological clarification of the term *balanced growth* in the subsequent footnote.

> problems with intermediate consumption, the economy spends time near the modified golden rule. Several of the papers in the capital and growth section of *CSP*(III) contain turnpike theorems. A periodic turnpike result is reported in 1976, *CSP*(IV:224).

By seeing all the results from 1949 to the 1970s under a unified perspective, Fischer's entry then glosses over two points: first, an analytical turn, and a shift of emphasis, in the years 1965 to 1967 from a performance functional depending solely on the terminal capital stock of a finite-horizon allocation problem, an emphasis deriving from von Neumann's paper [55],[17] to one depending on the aggregate felicity obtained over the *entire* planning period, be it finite or infinite, deriving from Ramsey's paper [45]; and second, the dissonance between the undiscounted and discounted settings, a dissonance hinging precisely on the procedure of aggregation that is involved in the first shift. What Fischer does keep clear, however, is the linguistic subscription to what is being termed classical turnpike theory in this essay: he does not discuss asymptotic stability or include it within his outline of the subject.

Our attempt to get an up-to-date overview of turnpike theory began with Bewley's 2007 text, and moved on to Fischer's 1987 *Palgrave* entry. We now make a third attempt by going to the 2008 *Palgrave* entry of Brock–Dechert, experts who have also put their own particular stamp on the subject.

> A key property of the one-sector model that promotes its use in real business cycle applications as well as intertemporal general equilibrium asset pricing applications is the stochastic analog of the turnpike theorem. This theorem states that optimal capital stock and optimal consumption converge in a stochastic sense to a unique stochastic limit under standard assumptions of concavity of the payoff function (for example, the planner's preferences) and of the production function and modest assumptions on the structure of the stochastic shocks. It is much more difficult to obtain such results for general multisector stochastic models, and even for deterministic versions of those models.

Thus, classical turnpike theory is being ruled out from the outset in favor of asymptotic stability.[18] And in this context, Brock–Dechert consider the undiscounted case first, and write:

> The basic idea behind these results, called 'turnpike' results, is to first observe that, if the discount rate on the future is zero, the dynamic optimization problem will attempt to maximize a long-run

[17] See the references in the papers cited in Footnote 16 above.

[18] Thus we now have three different, albeit overlapping definitions of the term "turnpike theory."

> 'static' objective in order to avoid infinite 'value loss' if it failed to do so. Making this intuition mathematically precise requires introduction of a partial ordering called the 'overtaking ordering' and making assumptions on the objective function and the dynamics so that avoidance of infinite value loss results in convergence of the optimal quantities to a unique long-run limit.

In keeping with the thrust of this essay, the point of course is how to parlay the undiscounted results, and the associated mathematical techniques, into results for the discounted case. As we shall see in the sequel, this is not possible in the large, and one has to be satisfied with the local case. To quote Brock–Dechert again,

> Once one has results well in hand for the case of zero discounting on the future, intuition suggests that there should be a notion of 'continuity' that would enable one to prove that, if the discount rate is close enough to zero, convergence would still hold. Unfortunately, turning such intuition into precise mathematics turns out to be rather difficult (see McKenzie (1986, 2002), for deterministic literature and Arkin–Evstigneev (1987), the papers in Dechert (2001) [10], and Marimon (1989) [30] for the stochastic case).

And so, even if one confines oneself to the deterministic literature, as we do in this essay,[19] one would think that a natural point to begin would be McKenzie's 1987 and 2002 texts [38, 40], and then to move on to relevant papers in the current scholarly literature.[20] It is worthy of note in this context that there is no chapter on "turnpike theory in the 2006 Handbook [9]. However, a paper that is especially relevant is Mitra's 2005 "complete characterization of the turnpike property of optimal paths in the (reduced form) aggregative model of intertemporal allocation," [41], but in the very first sentence of his definitive paper, Mitra defines the "turnpike property" as equivalent to "global asymptotic stability of a non-trivial stationary optimal stock".[21] In sum, this literature is not always clear from the viewpoint of the

---

[19] This stochastic literature is huge, and deserving of its own overview. Our basic point is that the deterministic results ought to be antecedent to it as a matter of logical priority. At any rate, the 2008 *Palgrave* entry of Brock–Dechert is again worth quoting: "infinite horizon stochastic multisector models are also basic in constructing econometrically tractable models to use in analysing data. Here, especially, is where stochastic versions of the turnpike theorem (explained below) are used. For example, it is used to justify use of laws of large numbers and central limit theorems in econometric time-series applications."

[20] See [6, Chap. 6] for a comparative results in continuous time.

[21] Given the clarity of the phrase, one is led to inquire into the need for the turnpike terminology in the first place. It is almost as if it has been endowed with incantatory

distinction between classical turnpike theory and the local or global asymptotic stability of optimal programs. In any case, as the 2008 *Palgrave* entries attest, none of the theorems are simple enough to be presented without a detailed discussion of the hypotheses and the conclusions they engender, a task that we pursue here.

The outline of this essay is then as follows. In Sect. 2, we consider relevant results from the 1991 text by Carlson et al. [6] that considers both classical turnpike theory as well as the qualitative behavior of solutions to optimal control problems over an unbounded time interval. Given the primary integrative motivation behind this essay, the fact that this text confines itself to continuous time, and has both the calculus of variations and Pontryagin's maximum principle as background, is an added advantage. We supplement the discussion by a recent survey paper of Rockafellar's [46]. In Sect. 3 we present the model in discrete time. We follow the economic literature but also [3, 59]. In Sect. 4, we turn to the results on asymptotic stability based on the "smoothness hypotheses" on felicity functions as presented in the initial 1976 paper of Scheinkman's [52], and the subsequent extensions by Araujo–Scheinkman and McKenzie in the discrete-time framework; see [1, Definition 4.2]. These results constitute a remarkably successful application of the implicit function theorem for ℓ_∞ through an exploitation of "smoothness hypotheses" on the felicity function.[22] Section 5 moves away from the smoothness hypothesis, and considers McKenzie's refinements of the interiority assumptions. We also present his version of the "neighborhood turnpike theorem." Sections 6 and 7 then turn to recent results on the RSS and MW models, and discuss them in the light of previous work. Section 8 concludes the essay.

2. Relevant results in continuous time

The 1991 text by Carlson et al. [6] is scrupulous in maintaining the distinction between classical turnpike theory and the global asymptotic stability of solutions to infinite-horizon allocation problems,[23] and even though this essay is limited to allocation problems in discrete time, it is worthwhile to

---

power, for approval or for dismissal of the literature! We make a serious attempt in this essay to avoid the turnpike terminology when we can satisfactorily do so.

[22] See the comprehensive treatment of this theorem in Krantz–Parks [28].

[23] After a consideration of examples in Chap. 3, titled "asymptotic stability and the turnpike property in some simple control problems," the authors turn to the development of the general theory (in continuous-time) in in the subsequent four chapters. Their Sect. 6.7, and Chap. 7 is particularly relevant to the delineation of the discounted case that is being attempted in this essay.

consider their results. At the end of this section we shall supplement our discussion with a consideration of the recent survey of Rockafellar (2009), and his reliance on his own work, as well as that of Samuelson [50] and Levhari–Leviatan [29].

Consider the following problem subsequently referred to as Problem $\mathcal{P}_c$.

$$\max J_T(x_o, w(\cdot)) = \int_0^T f_0(x(t), w(t))dt \text{ subject to } \dot{x}(t) = f(x(t), w(t)) \text{ and } x(0) = x_o,$$

where $f(\cdot, \cdot) : \mathbb{R}^n \times R^m \longrightarrow \mathbb{R}^n$ is a continuous mapping with respect to both arguments $x(t)$ and $w(t)$, $W(x) \subset \mathbb{R}^m$ is compact, and that the mapping $x \to W(x)$ is upper semicontinuous.[24] This problem was originally considered in 1956 by Samuelson–Solow [51] as a multi-sectoral generalization of the problem of Ramsey [45].

Definition 2.1. *A pair of functions* $(x(\cdot), w(\cdot)) : [0, \infty) \longrightarrow \mathbb{R}^m$ *is said to be admissible if* $x(\cdot)$ *is absolutely continuous,* $w(\cdot)$ *is measurable,* $\dot{x}(t) = f(x(\cdot), w(\cdot))$, $w(t) \in W(x(t))$ *almost everywhere in* $[0, \infty)$ *and* $x(0) = x_o$. *Given an admissible pair of functions* $(x(\cdot), w(\cdot))$, *we shall refer to* $x(\cdot)$ *as a trajectory from* x_o *and generated by the admissible control* $w(\cdot)$.

We now turn to optimality criteria for infinite-horizon problems.

Definition 2.2. *A trajectory* $x^*(\cdot)$ *from* x_o *is said to be maximal if for any other trajectory from* x_o,

$$\liminf_{T\to\infty} \int_{t=0}^T J_T(x_o, w(\cdot)) - J_T(x_o, w^*(\cdot)) \leq 0.$$

A program is optimal if the lim sup *operator above is substituted for* lim inf .

The reader should note that *maximal* trajectories are referred to in the mathematical literature as *weakly overtaking optimal*, and *optimal* trajectories as *overtaking optimal* trajectories.[25]

For the Problem $\mathcal{P}_c$, let $(\hat{x}, \hat{w})$ be the solution to the following associated static problem:

$$\max f_o(x, w) \text{ such that } f(x, w) = 0 \text{ and } w \in W(x). \quad (1)$$

[24] This is the basic autonomous system considered in [6, Sect. 4.2] which should be referred to for the definitions and for simple examples. Note, however, that by following this text, we are not imposing non-negativity constraints of the state variables x.

[25] See [6, Definition 1.2] and the discussion following it. Rockafellar [46] also considers *optimal* trajectories where the inequality is strict and refers to them as satisfying a "strong" overtaking property.

Assumption 2.1. *Let (i) $\hat{x}$ be a unique solution to the stationary problem (1), (ii) there exist $\hat{p} \in \mathbb{R}^n$ such that $f_o(\hat{x}, \hat{w}) = \max\{f_o(x, w) + \hat{p} f(x, w) : x \in \mathbb{R}^n, w \in W(x)\}$, and (iii) for any $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$||x - \hat{X}(\hat{x}, \hat{p})|| > \varepsilon \Longrightarrow \text{for all } w \in W(x),\ f_o(\hat{x}, \hat{w}) > (f_o(x, w) + \hat{p} f(x, w)) + \delta,$$

where $\hat{X}(\hat{x}, \hat{p}) \equiv \{x :$ there exists $w \in W(x)$ with $f_o(x, w) + \hat{p} f(x, w) = f_o(\hat{x}, \hat{w})\}$.

We can now present a basic result on the asymptotic stability of the solutions to the Problem $\mathcal{P}_c$.

Theorem 2.1. *Under Assumption 2.1 for the Problem $\mathcal{P}_c$, for any initial state x_o such that $\hat{x}$ is reachable in a finite time $T(x_o)$, any bounded maximal trajectory starting from x_o must satisfy $\lim_{T\to\infty} x^*(t) \in \hat{X}(\hat{x}, \hat{p})$.*

This theorem is presented in [6, Sect. 4.3] under the title "convergence toward the von Neumann set for weakly overtaking trajectories", and it is emphasized that the "classical maximum principle does not necessarily hold for the class" of functions for which the theorem is stated. In particular no differentiability assumptions on the function f_0 and f are made. And in the next section, the authors consider the "optimal control problem with fixed terminal time and prove a general turnpike theorem under the same assumptions which guarantee asymptotic stability of weak overtaking trajectories."

Theorem 2.2. *Under Assumption 2.1 for the Problem $\mathcal{P}_c$, for any initial state x_o such that $\hat{x}$ is reachable in a finite time $T_1(x_o)$, and any terminal state x_T reachable from $\hat{x}$ in a finite time $T_2(x_T)$, for any optimal trajectory $x^*(\cdot)$ and any $T > T_1(x_o) + T_2(x_T)$, the following property holds:*

$$\forall \varepsilon > 0 \text{ there exists } \tau(\varepsilon) > 0 \text{ such that } \mu[\{t \in [0, T] : ||x^*(t) - \hat{X}|| > \varepsilon\}] < \tau(\varepsilon),$$

where $\tau(\varepsilon)$ does not depend on T, and μ denotes Lebesgue measure.

In the remainder of Chap. 4 of [6], the authors consider the infinite-horizon version of the control problem $\mathcal{P}_c$, and using the results of Halkin [16] and Yano [56], present theorems on the existence and asymptotic stability of optimal and maximal trajectories. In particular, the results they present rely on what they refer to as the S property of the analogue of the von Neumann set. In Chap. 6 of their book [6], the authors consider both classical turnpike theory and local and global asymptotic stability of the solutions of the discounted version of Problem $\mathcal{P}_c$. We refer the reader to these results and satisfy ourselves with the following 2009 judgement of Rockafellar in [46].

> The infinite-horizon problem we wish to investigate in general, as an extension of the finite-horizon problem
>
> $$\max \int_0^T L_0(x(t), \dot{x}(t))dt \text{ subject to } x(0) = x_o, x(T) = x_T,$$
>
> has the form
>
> $$\max \int_0^\infty L_0(x(t), \dot{x}(t))e^{-\rho t}dt \text{ subject to } x(0) = x_o.$$
>
> Do the results for the [first] problem have some counterpart for the [second]? A particular trouble-spot is what to make of the absence of a terminal constraint. Should one just allow $x(T)$ to do anything as $T \longrightarrow \infty$, or should a sort of infinite-horizon terminal constraint be imposed? Mathematical economists have had to contend with these issues unaided by much technical literature, and various difficulties have not been resolved to satisfaction. We need to understand better the infinite-horizon integral we are contemplating in the [second problem].

3. The basic model in discrete time

We now turn to the discrete-time setting, and consider the corresponding version of the Problem $\mathcal{P}_c$. We shall refer to it as the Problem $\mathcal{P}_d$, or $\mathcal{P}_d(x_0, \rho)$ for emphasis.

$$V_T(x_o, \rho) = \max \sum_0^\infty \rho^t u(x_t, x_{t+1}) \text{ subject to a given } x_0 = x_o$$

and where $u : \Omega \subset \mathbb{R}_+^{2n} \longrightarrow \mathbb{R}$ and $0 < \rho < 1$. We shall also consider the case where $\rho = 1$ and refer to it as the undiscounted version of the Problem $\mathcal{P}_d$.

Araujo–Scheinkman [1] ascribe the original formulation of this problem to Samuelson–Solow [51], in the continuous-time version, and to Gale [14] and McKenzie [33] in the discrete-time version. They write:

> The reader more familiar with the "one-sector" growth theory would think of u as a "derived" utility function since there the utility function is defined in terms of consumption. There is no a priori reason to think of this as being a "derived" utility function since in many of the applications (e.g. the human capital theory) the utility function

does depend on the "state variables." Of course, the most general case is the one where the "original" utility depends on *both* "flow" and "stock" variables.[26]

The following basic concepts apply to the Problem $\mathcal{P}_d$ in either the undiscounted or discounted version.

Definition 3.1. *A program from x_o is a sequence $\{x(t)\}$ such that $x(0) = x_o$, and for all $t \in \mathbb{N}$, $(x(t), x(t+1)) \in \Omega$. A program $\{x(t)\}$ is simply a program from $x(0)$. A program $\{x(t)\}$ is called stationary if for all $t \in \mathbb{N}$, $(x(t)) = (x(t+1))$.*

We now turn to optimality criteria for a program. In the discounted case, with assumptions on Ω that ensure boundedness of programs, there is no issue of convergence of the performance functional.

Definition 3.2. *For all $0 < \rho < 1$, a program $\{x^*(t)\}$ from x_o is said to be optimal if*

$$\sum_{t=0}^{\infty} \rho^t [u(x(t), x(t+1)) - u(x^*(t), x^*(t+1))] \leq 0$$

for every program $\{x(t)\}$ from x_o. A stationary optimal program is a program that is stationary and optimal.

As first made explicit by Ramsey [45], the issue of the convergence of the performance functional becomes pressing only for the undiscounted case, which is to say, for the case $\rho = 1$. We can present the following analogue of Definition 2.2 for the discrete-time case.

Definition 3.3. *A program $\{x^*(t)\}$ from x_o is said to be maximal if for any other program from $\{x(t)\}$ from x_o,*

$$\liminf_{T\to\infty} \textstyle\sum_{t=0}^{T} u(x(t), x(t+1)) - u(x^*(t), x^*(t+1)) \leq 0.$$

A program is optimal if the lim sup *operator above is substituted for* lim inf .

We conclude this section by considering the discrete-time analogue of static problem considered above as (1).

$$\max\ u(x, x') \text{ such that } x' \geq x \text{ and } (x, x') \in \Omega. \tag{2}$$

[26] See [1, Footnote 2]. This is a rather prescient observation in the light of current interest in environmental economics. Perhaps such a more general model could be refereed to as the *MRSS* model: the *multi-sectoral Ramsey–Samuelson–Solow* model.

The interesting questions concern the discounted case where one solves the following problem:

$$\text{Find } (\hat{x}^{\rho}, \hat{x}^{\rho}) \in \Omega \text{ which solves } \max\ u(x, x') \text{ such that } x \leq (1-\rho)\hat{x}_{\rho} + \rho x' \text{ and } (x, x') \in \Omega. \tag{3}$$

Note that the above problem is a fixed point problem which reduces to the maximization problem represented in (2) only in the case $\rho = 1$. In the economic literature, the solution to (2) is referred to as the *golden-rule* stock, and to that of (3) as the *modified golden-rule* stock. For details as to these concepts, the reader can see McKenzie [38] and his references.[27] Also see the *weak* and *strong* turnpike distinction in [3].

4. Smooth felicities and asymptotic stability

In 1976, Scheinkman [52] assumes "smoothness hypotheses" on the felicity function u, and investigates whether the global asymptotic stability property of the optimal solutions to an undiscounted infinite-horizon problem can be translated to the optimal solutions of a discounted problem for "small" discount factors. This is to ask for a "neighborhood global stability theorem." His results were extended by Araujo–Scheinkman [1, 2] and McKenzie [35, 37, 38] to general perturbation results for arbitrary discount factors. These results constitute a remarkably successful application of the implicit function theorem for ℓ_∞, one that overcomes an important difficulty. The crucial issue can best be put in the words of Araujo–Scheinkman [1, p. 602].

> Since the implicit function theorem for such a space [the space of bounded sequences] reads exactly as in Euclidean spaces, the task becomes once again to study the inverse of a linear map. Given a linear map from a finite dimensional space into itself, the map is invertible if and only if it is one to one. This is no longer true in infinite dimensional spaces. Hence, concavity type restrictions . . . are not enough.

4.1. Scheinkman's "neighborhood" stability theorem

In an important paper in 1972, Samuelson [50], and following him, Levhari–Leviatan [29], had observed the saddle point property of optimal control motions. Rockafellar [46] discusses how the saddle-point property of the

[27] In the continuous-time case, the problem represented in (3) is referred to as an *implicit programming problem* in [6, Sect. 6.7].

associated Hamiltonian of the optimal control problem at the rest point translates into an "associated saddle point behavior of the Hamiltonian system in the dynamical sense" in the one-dimensional case, and asks:

> Could this somehow carry over to n-dimensions and in perturbation to positive discounting, at least if ρ isn't too high?

It is precisely an answer to this theorem that is provided in [52] as an application of the Hirsch–Pugh theorem.

We begin with the two basic assumptions of the texts [38, 40] that have now become standard for the general theory.

Assumption 4.1. *There exist* $M > 0$ *and* $N < 1$ *such that for all* $(x, x') \in \Omega$, $||x|| < M$ *implies* $||x'|| < N||x||$.

Assumption 4.2. *If* $(x, x') \in \Omega$, *then* $(y, y') \in \Omega$ *for all* $y \geq x$, $0 \leq y' \leq x'$, *and* $u(y, y') \geq u(x, x')$ *holds.*

Assumption 4.3. *There exists* $(\bar{x}, \bar{x}') \in \Omega$ *such that* $\bar{x}' >> \bar{x}$. *In this case,* $\bar{x}$ *is said to be expansible.*

We now present an assumption due to Brock [5] and fundamental to the theory; we already saw it in Assumption 2.1(i) above.

Assumption 4.4. *The golden-rule stock* $\hat{x}$ *is unique and expansible*

Assumption 4.5. *(i)* $u(\cdot, \cdot)$ *is* C^3 *in Int* Ω *and concave. (ii) The matrix*

$$\begin{bmatrix} u_{xx}(\hat{x}_\rho, \hat{x}_\rho) & u_{xx'}(\hat{x}_\rho, \hat{x}_\rho) \\ u_{x'x}(\hat{x}_\rho, \hat{x}_\rho) & u_{x'x'}(\hat{x}_\rho, \hat{x}_\rho) \end{bmatrix} \equiv \begin{bmatrix} A_\rho & B_\rho \\ B'_\rho & C_\rho \end{bmatrix}$$

is negative definite at $(\hat{x}, \hat{x})$. *(iii) All programs are interior programs so that the Euler difference equation system*

$$u_{x'}(x(t-1), x(t)) + \rho u_x(x(t), x(t+1)) = 0 \text{ for all } t \geq 1 \qquad (4)$$

is satisfied for an optimal program. (iv) The characteristic equation of the Euler difference equation system (4) *does not have a zero root. (v) There exists* $\rho_\ell < 1$ *such that for all* $\rho \geq \rho_\ell$, *there exists an expansible stationary program* $\underline{x}$ *such that* $\hat{x}^\rho >> \underline{x}$.

Theorem 4.1 (**Scheinkman**). *Under Assumptions 4.1, 4.2, 4.3, 4.4 and 4.5, for any expansible* $\underline{x}$, *there exists* $\bar{\rho} < 1$ *such that* $1 \geq \rho \geq \bar{\rho}$ *and* $x_o \geq \underline{x}$ *implies that there exists an optimal program* $x(t, x_o, \rho)$ *such that* $\lim_{t\to\infty} x(t, x_o, \rho) = \hat{x}^\rho$ *where* $\hat{x}^\rho$ *is the unique modified golden-rule stock associated with* ρ.

The reader is referred to [37, pp. 347–350] for a description of Scheinkman's theorem, and to [52] for its proof. The "visit lemma" in the proof does not require any smoothness hypotheses, and plays an important role in [22, 25], as discussed below.

4.2. Refinements of the "neighborhood" theorem

We now turn to the results of Araujo–Scheinkman [1]. We begin by providing an alternative to Assumption 4.5 which will be used as the Standing Hypothesis for all of the remaining results in this section.

Assumption 4.6. *The felicity function u is strictly concave in Ω and C^2, with $D^2 u$ negative definite in Int (Ω).*

Definition 4.1. *A program $\{x(t)\}$ is said to be a regular program if there exists $\varepsilon > 0$ such that*

$$\inf\{d_H[(x_t, x_{t+1}), \partial\Omega] : t = 0, 1, \cdots\} > 0,$$

where d_H denotes the Hausdorff distance, and $\partial\Omega$ the boundary of Ω.

We shall denote by the first projection of Ω by $\Pi_1(\Omega)$ and by $\mathcal{S}(\mathcal{S}^r)$ the set of all pairs of stocks and discount factors (x, ρ), $x \in \Pi_1(\Omega)$, $\rho \in (0, 1)$ for which there exists a solution (regular solution) to $\mathcal{P}_d(x, \rho)$. The optimal program function $x : \mathcal{S} \longrightarrow \ell_\infty^n$, $x(x_o, \rho)$ is a solution to $\mathcal{P}_d(x_o, \rho)$. Let

$$\tau = \{(x, \{x(t)\}) : \{x(t)\} \text{ is a regular program from } x\}.$$

We shall need the mapping

$$\Phi : \tau \times (0, 1) \longrightarrow \ell_\infty^n \text{ by } [\Phi(x_o, \{x(t)\}, \rho)]_t = u_{x'}(x(t-1), x(t)) \\ +\rho u_x(x(t), x(t+1)).$$

Note that for a given ρ, Φ maps a regular program into the left hand side of the Euler difference equation system (4). A crucial assumption of the analysis is the following[28]

Assumption 4.7. *$T = D_x\Phi(x_o, \{x(t)\}, \rho)$ has dominant diagonal blocks.*

Theorem 4.2. *Under Assumption 4.6, for any $(x_o, \rho) \in \mathcal{S}^r$ satisfying Assumption 4.7, there exists $\varepsilon > 0$ such that $B_\varepsilon(x_o, \rho) \subset \mathcal{S}^r$ and $x : B_\varepsilon(x_o, \rho) \longrightarrow \ell_\infty^n$ is C^1.*

We shall now need the following concepts.

Definition 4.2. *For any $\rho \in (0, 1)$, a modified golden-rule stock $\hat{x}^\rho$ is said to be locally asymptotically stable if there exists $\varepsilon > 0$ such that*

$$x_o \in \Pi_1(\Omega),\ |x_o - \hat{x}^\rho| < \varepsilon \Longrightarrow \lim_{t\to\infty} x(t, x_o, \rho) = \hat{x}^\rho.$$

[28] The reader is referred to [1, Definition 2.4] for the precise definition. We also note that the symbol x is being used in two senses: the value of a stock at a particular time, as well as a function from the initial stock and the discount factor to ℓ_∞^n.

Definition 4.3. *For any $X \subset \Pi_1(Int\ \Omega)$ and a $\rho \in (0, 1)$, a modified golden-rule stock $\hat{x}^\rho$ is said to be X-globally asymptotically stable at ρ if*

$$x_o \in X,\ (x_o, \rho) \in \mathcal{S}^r \implies \lim_{t\to\infty} x(t, x_o, \rho) = \hat{x}^\rho.$$

If X equals $\Pi_1(Int\ \Omega)$ in the above definition, $\hat{x}^\rho$ is said to be globally asymptotically stable at ρ.

The next result attests to the power of the continuity hypothesis in the strengthening the local asymptotic stability of a modified golden-rule program to a larger set of initial stocks.

Theorem 4.3. *For any $\rho \in (0, 1)$ and any connected $X \subset \Pi_1(\Omega)$ such that (i) $(x_o, \rho) \in \mathcal{S}^r$ for all $x_o \in X$, (ii) $x(\cdot, \rho) : X \longrightarrow \ell^n_\infty$ is continuous, and (iii) $\hat{x}^\rho$ is locally asymptotically stable, $\hat{x}^\rho$ is X-globally asymptotically stable.*

The question then devolves to a result on global asymptotic stability without the continuity assumption, which is to ask for a result with assumptions on the primitives. We present this next.

Theorem 4.4. *Under Assumption 4.6, for any $\rho \in (0, 1)$ such that (i) a modified golden-rule stock $(x^\rho, x^\rho) \in Int\ (\Omega)$, (ii) $(x_o, \rho) \in \mathcal{S}^r$, (iii) Assumption 4.7 holds for all $x_o \in Int\ (\Omega)$, x^ρ is globally asymptotically stable.*

In order to develop the next assumption, we shall abbreviate our notation to denote by u_1 the vector of partial derivatives of u with respect to the first n coordinates of $\mathbb{R}^{2n}$, and by u_2 the vector of partial derivatives of u with respect to the last n coordinates of $\mathbb{R}^{2n}$. Correspondingly, we shall denote by $u_{i,j},\ i = 1, 2;\ j = 1, 2$, the matrices of partial derivatives of $u_i,\ i = 1, 2$. Thus, in terms of the earlier notation, u_{21} denotes $u_{x'x}$. Finally, we shall denote by u^ρ_{ij} the matrix $u_{x'x}(\hat{x}^\rho, \hat{x}^r ho)$.

Assumption 4.8. *The modified golden-rule stock $(\hat{x}^\rho, \hat{x}^\rho) \in Int\ (\Omega)$, $det\ u_{x'x}(\hat{x}^\rho, \hat{x}^\rho) \neq 0$, and there exist n roots $\lambda_1, \cdots, \lambda_n$, all of modulus less than one, of the characteristic equation*

$$det\ (u^\rho_{21} + \lambda[u^\rho_{22} + \rho u^\rho_{11}] + \lambda^2 \rho u^\rho_{12}) = 0.$$

Theorem 4.5 (Araujo–Scheinkman). *Under Assumption 4.6, for any $\rho \in (0, 1)$ satisfying Assumption 4.8, and $x_o \in Int\ (\Omega)$ such that $(x_o, \rho) \in \mathcal{S}^r$ and $\lim_{t\to\infty} x(t, x_o, p) = \hat{x}^\rho$, $T(x_o, \rho)$ is an isomorphism. Consequently, there exists $\varepsilon > 0$ such that $B_\varepsilon(x_o, \rho) \subset \mathcal{S}^r$ and the restriction of the optimal program function x to $B_\varepsilon(x_o, \rho)$ is C^1.*

By strengthening the asymptotic stability of the modified golden-rule stock x^{ρ} to global asymptotic stability in Theorem 4.5, Araujo–Scheinkman state two important corollaries of the result for the given ρ: (i) that the C^1 property of the optimal program holds for all $x_o \in \text{Int}(\Omega)$, (ii) that on the optimal program, the optimal policy function from the stock in one period to that in the next is also C^1. But the *piéce de resistance* is the following theorem in the neighborhood of the discount factor.

Theorem 4.6. *Under Assumption 4.6, for any $\bar{\rho} \in (0, 1)$ satisfying Assumption 4.8, $x_o \in Int(\Omega)$, $(x_o, \bar{\rho}) \in \mathcal{S}^r$, the modified golden-rule stock $\hat{x}^{\bar{\rho}}$ is globally asymptotically stable, and any compact set $X \subset \Pi_1(Int\ \Omega)$, there exists $\varepsilon > 0$ such that for any $\rho \in B_\varepsilon(\bar{\rho})$, x^{ρ} is X-asymptotically stable at ρ.*

Araujo–Scheinkman remark that Theorem 4.6 can be strengthened to a result on global asymptotic stability if the Visiting Lemma holds. This is to say that for any $x_o \in \text{Int}(\Omega)$ and any $\rho \in B_\varepsilon(\bar{\rho})$, and for a suitable set of stocks K, there exists $t(x_o, \rho)$ such that $x_{t(x_o,\rho)} \in K$.

5. The theory in non-smooth environments

In the chapter on *Competitive Equilibrium over Time,* McKenzie [40] presents results for a "generalized Ramsey model", and his work is distinguished by its reliance on a bounded assumption on the technology Ω, and a joint assumption on the pair (u, Ω) guaranteeing *free-disposal* and *monotonicity.*

Under Assumptions 4.1 and 4.2, he presents theorems for both the discounted and undiscounted cases, seeing the latter as logically antecedent to the former.

5.1. Theorems under refined interiority assumptions

We can now present a classical turnpike theorem, $\text{Card}(A)$ denoting the cardinality of a finite set A.

Theorem 5.1. *Let Assumptions 4.1, 4.2 and 4.3 hold and that $\hat{x}$ is expansible. Then, there exists $\varepsilon > 0$ such that for all $T > L$, and all optimal programs $\{x_t\}_{t=0}^{t=T}$ such that x_0 and x_T are expansible,*

$$Card\Big\{i \in [0, \cdots, T-1] : \|x(t) - X(\hat{x}, \hat{p})\| > \varepsilon\Big\} \leq L.$$

This is presented in [40, Theorem 3, Chap. 7] and proved there.

Mckenzie introduces his other theorem for the undiscounted Ramsey setting as “an asymptotic theorem for infinite optimal paths.” For this, he needs a particular set of initial stocks K from which there start programs $\{x(t)\}$ such that

$$\liminf_{T\to\infty} \sum_{t=0}^{T} u(x(t), x(t+1)) - u(x^*(t), x^*(t+1)) \geq -\infty.$$

These are the *good* programs of Gale [14] and Brock [5]; see Assumptions 2.1 and 4.4.

Theorem 5.2. *Let Assumptions 4.1, 4.2 and 4.3 hold and that $\hat{x}$ is unique. Let $\{x(t)\}$ be an optimal program that starts from a stock in the interior of K. Then, for any $\varepsilon > 0$, there exists T_o such that $||x(t) - \hat{x}|| < \varepsilon$ for all $t \geq T_o$.*

5.2. McKenzie’s “neighborhood turnpike” theorem

We shall now make assumptions on the discount factor. We shall need the following additional notation. Let $D = \{x \in \mathbb{R}^n_+ : (x, x) \in \Omega\}$, and

$$D_\rho = \{x \in D : u(x, x) \geq u(\bar{x}, \bar{x}') \text{ for all } (\bar{x}, \bar{x}') \in \Omega \text{ such that } \rho\bar{x}' >> x\}.$$

Before introducing the next assumption, we shall need McKenzie’s definition of *uniform strict concavity* of a function taken from [40, Sect. 7.6].

Definition 5.1. *The utility function u is said to be uniformly strictly concave over D_ρ if for any $bx \in D_\rho$, and any $(w, z) \in \Omega$, the following holds: For any $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$|(x, x) - (w, z)| > \varepsilon \Longrightarrow u\left(\frac{1}{2}(x, x) + \frac{1}{2}(w, z)\right) - \frac{1}{2}(u(x, x) + u(w, z)) > \delta.$$

We can now present

Assumption 5.1. *There exists $\rho_\ell < 1$ and $(\bar{x}, \bar{x}') \in \Omega$ such that $\rho_\ell \bar{x}' >> \bar{x}$. We assume $\rho_\ell \leq \rho \leq 1$, u is uniformly strictly concave over D_ρ, and that D_ρ is in the relative interior of D.*

Assumption 5.2. *$\hat{x}$ is expansible which is to say that there exists $(\hat{x}, x') \in \Omega$ such that $x' >> \hat{x}$. Furthermore D_ρ is in the relative interior of D.*

Definition 5.2. *We shall say that an initial stock $x_o \in \mathbb{R}^n_+$ is sufficient if there exists a finite program $\{x_t\}_{t=0}^{t=T}$ such that x_T is expansible.*

We can now present McKenzie’s so-called “neighborhood turnpike theorem” for any expansible $\hat{x}$.

Theorem 5.3. *Under Assumptions 4.1 and 4.2, for any sufficient $x_o \in \mathbb{R}^n$, and any ρ satisfying Assumptions 5.1 and 5.2, and any $\varepsilon > 0$, there exists $1 > \bar{\rho} > 0$ such that $\hat{x}^{\rho}$ is asymptotically stable at (ρ, x_o) for all $\bar{\rho} \leq \rho \leq 1$.*

For the proof, the reader is referred to that of Theorem 6 in [40, Sect. 7.6]. Note that in [36, Theorem 4] McKenzie presents another theorem which he introduces as the "neighborhood turnpike" and refers to it as the "basic result of the subject." However, he does not report this theorem in his text [40]. It is clear that further comparative work needs to be done.

6. Particular undiscounted cases: the RSS model

There is by now an extensive literature on the so-called RSS model, and we refer the reader to the relevant references: for its underpinnings and interpretation as a model in development planning and the antecedent literature due to Robinson, Solow and Srinivasan, see [17]; for details of the important two-sector RSS case, see [18, 20] and their references; for the continuous time analysis, see [19, 54]. Our main concern here is to bring out the fact that the RSS model is a particular case of the model spelt out in Sect. 3.[29]

The model requires the total labor force to be stationary, perfectly divisible and positive, and thereby normalized to unity. There are n types of machines, produced using only labor, with $a_i > 0$ units of labor are needed to produce one unit of a perfectly divisible machine of type i, for $i = 1, \ldots, n$. Let a denote the n-vector $(a_1, \cdots, a_n)$. The *state* of the system will be represented by a point $x \in \mathbb{R}^n_+$, where x_i represents the current stock of machines of type $i = 1, \ldots, n$. If x represents the stock of machines today, x' the stock tomorrow and $d \in (0, 1)$ the common rate of depreciation, then $z = x' - (1 - d)x$ stands for the number of machines produced during the period. The constraints $z \geq 0$ and $az \leq 1$ represent respectively the irreversibility of investment and the maximum labor available, where az denotes inner product. With these definitions, the *transition possibility set* is given by

$$\Omega = \{(x, x') \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ : x' - (1-d)x \geq 0 \text{ and } a(x' - (1-d)x) \leq 1\}. \quad (5)$$

Given the pair $(x, x') \in \Omega$ and denoting by e the vector $(1, 1, \cdots, 1) \in \mathbb{R}^n$, the stock of machines that may be devoted to the consumption goods sector is given by the correspondence

$$\Lambda(x, x') = \{y \in \mathbb{R}^n : 0 \leq y \leq x \text{ and } ey \leq 1 - a(x' - (1-d)x)\}. \quad (6)$$

[29] We leave it to the reader as an exercise to show that the continuous-time version of the model, as presented in [19, 54] and their references, is a particular case of the model spelt out in Sect. 2.

One unit of labor together with one unit of machine of type $i = 1, \ldots, n$, can produce $b_i > 0$ units of a single consumption good, thus, defining the output-coefficients vector $b = (b_1, \ldots, b_n) \in \mathbb{R}^n_+$, the total good production is given by by, with $y \in \Lambda(x, x')$.

Welfare is derived only from the consumption good and is represented by a felicity function $w : \mathbb{R}_+ \to \mathbb{R}_+$ that is assumed to be continuous, strictly increasing, concave and differentiable. In terms of the triple (Ω, u, x_o) pertaining to the general theory, and mentioned in the first sentence of the introduction, we can now define

$$u(x, x') = \arg\max\{w(by) : y \in \Lambda(x, x')\} \text{ for all } (x, x') \in \Omega.$$

In the results reported in this section, we shall make the following assumption:

Assumption 6.1. *There exists $\sigma \in \{1, \ldots, n\}$ such that for all $i \in \{1, \ldots, n\} \setminus \{\sigma\}$, $c_\sigma > c_i$, where $c_i = b_i/(1 + da_i)$, $i = 1, \ldots, n$.*

This is a fundamental assumption that guarantees a unique golden-rule stock, and thereby a unique von-Neumann ray. This assumption goes back to Brock [5], and for a detailed discussion in the context of the RSS model, the reader is referred to [17]; also see [57].

We shall also work with the basic parameter (and attendant notation) that was identified in the discrete-time context [17], and was totally missed in the continuous-time setting originally studied by Stiglitz [54].

$$\xi_i = (1/a_i) - (1 - d) \text{ for all } i = 1, \cdots, n. \tag{7}$$

Even though a special case of the general theory, the model does not fulfill Assumptions 4.5 and 4.6 that the theory requires.

6.1. A theorem on asymptotic stability

We now present a basic result on the asymptotic stability

Theorem 6.1. *Assume that (i) w is strictly concave, or that (ii) $\xi_\sigma \neq 1$. Let $M_0, \epsilon > 0$. Then there exists a natural number T_0 such that for each optimal program $\{x(t), y(t)\}_{t=0}^{\infty}$ satisfying $x(0) \leq M_0 e$ and each integer $t \geq T_0$*

$$||x(t) - \widehat{x}||,\ ||y(t) - \widehat{x}|| \leq \epsilon.$$

Note that the time period T_o is does not depend on the initial stock, lying in the given range of initial stocks, from which the optimal program starts. For a proof, the reader is referred to [24].

6.2. Classical turnpike theorems

The basic results are as follows.

Theorem 6.2. *Let M, ϵ be positive numbers and $\Gamma \in (0,1)$. Then there exists a natural number L such that for each integer $T > L$, each $z_0, z_1 \in R^n_+$ satisfying $z_0 \le Me$ and $az_1 \le \Gamma d^{-1}$ and each program $(\{x(t)\}_{t=0}^{T}, \{y(t)\}_{t=0}^{T-1})$ which satisfies*

$$x(0) = z_0,\ x(T) \ge z_1, \quad \sum_{t=0}^{T-1} w(by(t)) \ge U(z_0, z_1, 0, T) - M,$$

the following inequality holds:

$$Card\{i \in \{0, \dots, T-1\} :\ \max\{||x(t) - \widehat{x}||,\ ||y(t) - \widehat{x}||\} > \epsilon\} \le L.$$

Theorem 6.3. *Let M, ϵ be positive numbers and $\Gamma \in (0,1)$. Then there exist a natural number L and a positive number γ such that for each integer $T > 2L$, each $z_0, z_1 \in R^n_+$ satisfying $z_0 \le Me$ and $az_1 \le \Gamma d^{-1}$ and each program $(\{x(t)\}_{t=0}^{T}, \{y(t)\}_{t=0}^{T-1})$ which satisfies*

$$x(0) = z_0,\ x(T) \ge z_1, \quad \sum_{t=0}^{T-1} w(by(t)) \ge U(z_0, z_1, 0, T) - \gamma,$$

there are integers τ_1, τ_2 such that $\tau_1 \in [0, L],\ \tau_2 \in [T-L, T]$,

$||x(t) - \widehat{x}||,\ ||y(t) - \widehat{x}|| \le \epsilon$ for all $t = \tau_1, \dots, \tau_2 - 1$ and $||x(\tau_2) - \widehat{x}|| \le \epsilon$.

Moreover if $||x(0) - \widehat{x}|| \le \gamma$ then $\tau_1 = 0$.

For proofs and notation, the reader is referred to [25]. This reference also derives the asymptotic stability result as a corollary of these theorems. Many turnpike results of this kind motivated by non-economic applications are collected in [38].

7. Particular undiscounted cases: the MW model

There is by now a vigorously-developing literature on the so-called MW model, and we refer the reader to the relevant references: the pioneering papers are of Mitra–Wan [42, 43] subsequent to those of Faustman and Samuelson.[30] As in Sect. 6, our main concern here is to bring out the fact that the MW model is a particular case of the model spelt out in Sect. 3.[31]

[30] For the recent papers, see [47, 48] and [21, 22] and their references.

[31] The continuous-time version of the model is difficult, and no longer an easily formulated particular version of the model spelt out in Sect. 2; see [42, 43] for the early formulations in the work of Kemp and Wan.

We shall work in the $(n-1)$-dimensional simplex $\Delta = \{x \in \mathbb{R}_+ : \sum_{i=1}^n x_i = 1\}$, let the supremum norm of x be $\|x\|_\infty$. Let the total surface be unity and n the age after which a tree dies or losses its economic value. We consider that the timber content per unit of area is related only to the age of the trees, through the biomass coefficient vector $b = (b_1, \cdots, b_n)$. We make an assumption analogous to Assumption 6.1 that:

Assumption 7.1. *There exists* $\sigma = \{1, \cdots, n\}$ *such that* $b_\sigma/\sigma > b_i/i$ *for all* $i \in \{1, \cdots, n\} \setminus \{\sigma\}$.

For each period $t \in \mathbb{N}$ we denote $x_i(t) \geq 0, i = 1, \cdots, n$ the surface occupied by trees of age i at time t. At every stage we must decide how much land to harvest of every age-class, $c(t) = (c_1(t), \cdots, c_n(t))$ where $c_i(t) \in [0, x_i(t)]$. As we know that after n a tree has no value, we assume that $c_n(t) = x_n(t)$ for all t. By the end of period $t+1$, the state will be exactly

$$x(t+1) = \left(\sum_{i=1}^n c_i(t), x_1(t) - c_1(t), \cdots, x_{n-1}(t) - c_{n-1}(t)\right).$$

Define the transition possibility set Ω as the collection of pairs $(x, x') \in \Delta \times \Delta$ such that it is possible to go from the state x in the current period (today) to the state of the forest x' in the next period (tomorrow) fulfilling relations. Formally,

$$\Omega = \{(x, x') \in \Delta \times \Delta : x_i \geq x'_{i+1} \text{ for all } i = 1, \cdots, n1\}$$

The vector of harvests needed to perform this transition is given by the function $\lambda : \Omega \to \mathbb{R}^n_+$, $\lambda(x, x') = (x_1 - x'_2, x_2 - x'_3, \cdots, x_{n-1} - x'_n, x_n)$. In addition, it is easy to see that $(x, x') \in \Omega \Leftrightarrow x, x' \in \Delta$ and $\lambda(x, x') \geq 0$.

The preferences of the planner are represented by a felicity function, $w : [0, \infty) \to \mathbb{R}$ which is assumed to be continuous, strictly increasing and concave. Define for any $(x, x') \in \Omega$ the function $u(x, x')$ as

$$u(x, x') = w(bc) \text{ where } c = \lambda(x, x')$$

This model is in fact an equivalent formulation of the one proposed in [42, 43],[32] and one that does not fulfill assumptions 4.2 and 4.4 of the general theory. The set Ω has no interior point in $\mathbb{R}^{2n}$, and the natural preorder does not apply to it. A reformulation fulfills Assumption 4.3 but not 4.4.

In the discounted case, the modified golden-rule stock is not asymptotically stable, see [47, 48]. However the validity of the turnpike theorems is still an open question. Here again Assumptions 4.6 and 4.8 are not fulfilled: non-interiority and non-concavity remains a thorny issue.

[32] For differences in the model as presented in [42, 43] and here, see [22].

7.1. A theorem on asymptotic stability

We first present an asymptotic stability result taken from [23], a reference that contains further discussion and proof.

Theorem 7.1. *Let $\varepsilon > 0$. There exists a natural number T_0 such that for each optimal program $\{x(t)\}$ the following inequality holds:*

$$\|x(t) - \hat{x}\|_\infty < \varepsilon \text{ for all } t \geq T_0.$$

For a proof of this theorem, the reader is referred to [23].

7.2. Classical turnpike results

Next, we present classical turnpike results taken from [22], again a reference that contains further discussion and proof.

Theorem 7.2. *Given $M > 0$ and $\varepsilon > 0$ there exists $L \in \mathbb{N}$ such that for all $T > L$ and each program $\{x(t)\}_{t=0}^{T}$ satisfying*

$$\textstyle\sum_{t=0}^{T-1} w(bc(t)) \geq U\big(x(0), x(T), 0, T\big) - M,$$

$$Card\big\{i \in [0, \cdots, T-1] : \|x(t) - \hat{x}\| > \varepsilon\big\} \leq L.$$

Theorem 7.3. *Let $\varepsilon > 0$. Then there exist $L \in \mathbb{N}$ and $M > 0$ such that for all $T > 2L + n + \sigma$ and each program $\{x(t)\}_{t=0}^{T}$ satisfying*

$$\textstyle\sum_{t=0}^{T-1} w(bc(t)) \geq U\big(x(0), x(T), 0, T\big) - M,$$

there are τ_1, τ_2 such that $\tau_1 \in [0, L]$, $\tau_2 \in [T - L, T]$ and $\|x(t) - \hat{x}\| \leq \varepsilon$ for all $t = \tau_1, \cdots, \tau_2$. Moreover, if $\|x(0) - \hat{x}\| \leq \varepsilon/n^2$ then $\tau_1 = 0$.

8. Conclusion

The alert reader has surely noticed that in the work presented above, we have been silent on the aggregative, one-sector setting, referred to as the RCK model in the introduction. We briefly remedy this deficiency here. As is well-known, the pioneering results on the discounted case in this setting are due to Cass [7, 8], Koopmans [26, 27] and Samuelson [49]. Indeed, Samuelson sketches the theorem in the last footnote of his 1965 paper as follows:

> Suppose society has a systematic subjective rate of time preference ρ, so that the integrand of $u(c)$ is replaced by $u(c)e^{-\rho t}$. Then a new "turnpike" is given by the root of $f'(x) = \rho$ rather than equal zero. Call this x^ρ. Then for T sufficiently large, for all (x_o, x_T), the optimal path $x(t)$ will remain indefinitely near to x^ρ an indefinitely large

> fraction of the time. The motions are not quite balanced catenaries, instead approximating to $k(t) = a_1 e^{\lambda_1 t} + a_2 e^{\lambda_2 t}, \lambda_1 > \lambda_2$. This constitutes an even more general Turnpike Theorem than the one elaborated here, and it reminds us that correct saving must depend on what is assumed about social time preference.[33]

At this point in the exposition we can connect to the epigraph of this essay taken from a neglected paper of Gale's [15].[34] Unlike Ramsey, Samuelson, Solow and Cass, Gale conducts the analysis in discrete-time, and shows that the results on the characterization of solutions of the RCK model are too good to be true: the classical turnpike theorem is valid without any restriction on the discount factor. In a recent paper [41], Mitra provides a synthesis of sufficient conditions for asymptotic stability of optimal programs in the RCK model, but has nothing to say on what we are referring to as classical turnpike theory in this essay.

In conclusion, and as emphasized in the introduction, the work presented above has been governed by two objectives: first, to encourage productive engagement across disciplinary boundaries; and second, to identify open problems for both disciplines. The point is to provide an overview that would benefit and bring together the two communities of economists and mathematicians: the former working in growth theory and economic dynamics, and the latter in dynamical systems and optimal control theory.[35] A terminological distinction between classical turnpike theory and the qualitative behavior of solutions to optimal control problems over an unbounded time interval, in either discrete or continuous time, certainly advances this objective.[36] Ironically, recent work of economists has concerned itself with the qualitative investigation of asymptotic stability of optimal control motions in discrete time, and it has been left to mathematicians to investigate classical turnpike theory in continuous time, with of course notable exceptions in both cases.[37] As regards the second objective, the aspiration remains for a generalized theory pertaining to the model due to Samuelson–Solow (1956), Gale (1967)

33 If footnotes could be labeled "pioneering," surely this footnote (notationally modified) would be high on such a list. It is a fascinating exercise in the history of economic analysis to trace the work surveyed in Sect. 4 as an elaboration of this footnote.

34 In the context of this epigraph, the reader should note that the Kuhn–Tucker theorem is now referred to as the Karush–Kuhn–Tucker theorem, and what Gale refers to as the *discount rate* is the *discount factor.* The first sentence of the epigraph is taken from page 308, the second from pp. 314–315 and the third from p. 310.

35 We see Rockafellar's recent survey [46] as being kindred in this motivation.

36 And so we end as we began: with the importance of having a clear and well-established terminology. See Footnote 12 above and the text it footnotes.

37 See, for example, Carlson et al. [6], Arkin–Evstegneev [3] and Zaslavski [58], and their references, to the work of both economists and mathematicians.

and McKenzie (1968), one that moves seamlessly between the discounted and undiscounted cases, asymptotically implements the continuous-time results in terms of their discrete-time counterparts, and covers both the RSS and MW models, the two concrete instances in development planning and the economics of forestry considered above. As we have seen, non-interiority assumptions are endemic to these applications, and smoothness, strict concavity and corresponding curvature assumptions do not translate to corresponding assumptions on the reduced-form performance functionals in which they are phrased.

References

1. Araujo, A., Scheinkman, J.A.: Comparative dynamics and the turnpike property. Econometrica **45**, 601–620 (1977)
2. Araujo, A., Scheinkman, J.A.: Notes on comparative dynamics. In: Green, J.R., Scheinkman, J.A. (eds.) General Equilibrium, Growth and Trade, pp. 217–226. Academic, New York (1979)
3. Arkin, V., Evstigneev, I.: Stochastic Models of Control and Economic Dynamics. Academic, New York (1987)
4. Bewley, T.F.: General Equilibrium, Overlapping Generations Models and Optimal Growth Theory. Harvard University Press, Cambridge (2007)
5. Brock, W.A.: On existence of weakly maximal programmes in a multi-sector economy. Rev. Econ. Stud. **37**, 275–280 (1970)
6. Carlson, D.A., Haurie, A.B., Leizarowitz, A.: Infinite Horizon Optimal Control: Deterministic and Stochastic Systems. Springer, Berlin (1991)
7. Cass, D.: Optimum growth in an aggregative model of capital accumulation. Rev. Econ. Stud. **32**, 233–240 (1965)
8. Cass, D.: Optimum growth in an aggregative model of capital accumulation: a turnpike theorem. Econometrica **34**, 833–850 (1966)
9. Dana, R.A., Le Van, C., Mitra, T., Nishimura, K. (eds.): Handbook of Optimal Growth, vol. 1. Springer, Berlin (2006)
10. Dechert, W. (ed.): Growth Theory, Nonlinear Dynamics, and Economic Modelling: Scientific Essays of William Allen Brock. Edward Elgar, Cheltenham (2001)
11. Dorfman, R., Solow, R.M., Samuelson, P.A.: Linear Programming and Economic Analysis. McGraw-Hill, New York (1958)
12. Durlauf, S., Blume, L. (eds.): The New Palgrave, vols. 1–7. Macmillan, New York (2008)
13. Eatwell, J., Milgate, M., Newman, P.K. (eds.): The New Palgrave, vols. 1–4. Macmillan, New York (1987)

14. Gale, D.: On optimal development in a multi-sector economy. Rev. Econ. Stud. **34**, 1–18 (1967)
15. Gale, D.: Nonlinear duality and qualitative properties of optimal growth. In: Abadie, J. (ed.) Integer and Nonlinear Programming, pp. 309–319. North Holland, Amsterdam (1970)
16. Halkin, H.: Necessary conditions for optimal control problems with infinite horizon. Econometrica **42**, 267–273 (1974)
17. Khan, M.A., Mitra, T.: On choice of technique in the Robinson–Solow–Srinivasan model. Int. J. Econ. Theory **1**, 83–109 (2005)
18. Khan, M.A., Mitra, T.: Discounted optimal growth in the two-sector RSS model: a geometric investigation. Adv. Math. Econ. **8**, 349–381 (2006)
19. Khan, M.A., Mitra, T.: Optimal growth in the two-sector RSS model: a continuous time analysis. In: Proceedings of the Seventh Portugese Conference on Automatic Control, Lisboa, Portugal. Electronic publication (2006)
20. Khan, M.A., Mitra, T.: Optimal growth in a two-sector RSS model without discounting: a geometric investigation. Jpn. Econ. Rev. **58**, 191–225 (2007)
21. Khan, M.A., Piazza, A.: On the Mitra-Wan Forestry Model: A Unified Analysis. *Publicación Técnica CMM* No. 226. Paper presented at *Workshop in Economic Dynamics,* National University of Singapore, July 30–August 1, 2009 (2009)
22. Khan, M.A., Piazza, A.: Classical turnpike theory and the economics of forestry. J. Behav. Econ. Organ. (2010, forthcoming)
23. Khan, M.A., Piazza, A.: On uniform convergence of undiscounted optimal programs in the Mitra-Wan forestry model: the strictly concave case. Int. J. Econ. Theory **6**, 57–76 (2010)
24. Khan, M.A., Zaslavski, A.J.: On a uniform turnpike of the third kind in the Robinson–Solow–Srinivasan model. J. Econ. **92**, 137–166 (2006)
25. Khan, M.A., Zaslavski, A.J.: On two classical turnpike results for the Robinson–Solow–Srinivasan (RSS) model. Adv. Math. Econ. **13**, 47–97 (2010)
26. Koopmans, T.C.: On the concept of optimal economic growth. In: Pontificae Academiae Scientarum Scripia Veria (ed.) Economic Approach to Development Planning, pp. 225–287. North Holland, Amsterdam (1965)
27. Koopmans, T.C.: Intertemporal distribution and "optimal" aggregate economic growth. In: Fellner, W., et al. (eds.) Ten Economic Studies in the Tradition of Irving Fisher, pp. 95–126. Wiley, New York (1967)
28. Krantz, S.G., Parks, H.R.: The Implicit Function Theorem: History, Theory and Applications. Birkhäuser, Basel (1998)
29. Levhari, D., Leviatan, N.: Stability in the saddlepoint sense. J. Econ. Theory **4**, 88–93 (1972)

30. Marimon, R.: Stochastic turnpike property and stationary equilibrium. J. Econ. Theory **47**, 282–306 (1989)
31. McKenzie, L.W.: The Dorfman–Samuelson–Solow turnpike theorem. Int. Econ. Rev. **4**, 29–43 (1963)
32. McKenzie, L.W.: The turnpike theorem of Morishima. Rev. Econ. Stud. **30**, 169–176 (1963)
33. McKenzie, L.W.: Accumulation programs of maximum utility and the von Neumann facet. In: Wolfe, J.N. (ed.) Value, Capital and Growth, pp. 353–383. Edinburgh University Press, Edinburgh (1968)
34. McKenzie, L.W.: Turnpike theory. Econometrica **43**, 841–865 (1976)
35. McKenzie, L.W.: A new route to the turnpike. In: Henn, R., Moeschlin, L.O. (eds.) Mathemtical Economics and Game Theory. Springer, Berlin (1977)
36. McKenzie, L.W.: Turnpike theory, discounted utility, and the von Neumann facet. J. Econ. Theory **27**, 194–209 (1982)
37. McKenzie, L.W.: A primal route to the turnpike and Lyapunov stability. J. Econ. Theory **30**, 330–352 (1983)
38. McKenzie, L.W.: Optimal economic growth, turnpike theorems and comparative dynamics. In: Arrow, K.J., Intrilligator, M. (eds.) Handbook of Mathematical Economics, vol. 3, pp. 1281–1355. North-Holland, New York (1986)
39. McKenzie, L.W.: Turnpikes. Am. Econ. Rev. (Papers and Proceedings) **88**, 1–14 (1998)
40. McKenzie, L.W.: Classical General Equilibrium Theory. The MIT Press, Cambridge (2002)
41. Mitra, T.: Characterization of the turnpike property of optimal paths in the aggregative model of intertemporal allocation. Int. J. Econ. Theory **1**, 247–275 (2005)
42. Mitra, T., Wan, H.W., Jr.: Some theoretical results on the economics of forestry. Rev. Econ. Stud. **LII**, 263–282 (1985)
43. Mitra, T., Wan, H.W., Jr.: On the Faustmann solution to the forest management problem. J. Econ. Theory **40**, 229–249 (1986)
44. Pietsch, A.: History of Banach Spaces and Linear Operators. Birkhäuser, Berlin (2007)
45. Ramsey, F.: A mathematical theory of savings. Econ. J. **38**, 543–559 (1928)
46. Rockafellar, R.T.: Hamiltonian trajectories and saddle points in mathemtical economics. Contr. Cybern. **4**, 1575–1588 (2009)
47. Salo, S., Tahvonen, O.: On equilibrium cycles and normal forests in optimal harvesting of tree vintages. J. Environ. Econ. Manage. **44**, 1–22 (2002)

48. Salo, S., Tahvonen, O.: On the economics of forest vintages. J. Econ. Dyn. Control **27**, 1411–1435 (2003)
49. Samuelson, P.A.: A catenary turnpike theorem involving consumption and the golden rule. Am. Econ. Rev. **55**, 486–496 (1965). An erratum in 55, 864–866
50. Samuelson, P.A.: The general saddlepoint property of optimal control motions. J. Econ. Theory **5**, 102–120 (1972)
51. Samuelson, P.A., Solow, R.M.: A complete capital model involving heterogeneous capital goods. Q. J. Econ. **27**, 537–562 (1956)
52. Scheinkman, J.A.: On optimal steady states of *n*-sector growth models when utility is discounted. J. Econ. Theory **12**, 11–30 (1976)
53. Solow, R.M.: Modern capital theory. In: Brown, E.C., Solow, R.M. (eds.) Paul Samuelson and Modern Economic Theory. McGraw-Hill, New York (1983)
54. Stiglitz, J.E.: A note on technical choice under full employment in a socialist economy. Econ. J. **78**, 603–609 (1968)
55. von Neumann, J.: Über ein ökonomisches gleichungs-system und eine verallgemeinerung des Brouwerschen fixpunktsatzes. In: Menger, K. (ed.) Ergebnisse eines Mathematischen Kolloquiums, 8. Translated as: A model of general economic equilibrium. Rev. Econ. Stud. **13**, 1945–1946, 1–9 (1935–1936)
56. Yano, M.: Note on the existence of an optimal capital accumulation path. J. Econ. Theory **27**, 421–429 (1981)
57. Zaslavski, A.J.: Optimal programs in the RSS model. Int. J. Econ. Theory **1**, 151–165 (2005)
58. Zaslavski, A.J.: Turnpike Properties in the Calculus of Variations and Optimal Control. Springer, New York (2005)
59. Zaslavski, A.J.: Turnpike results for discrete-time optimal control systems arising in economic dynamics. Nonlinear Anal. **67**, 2024–2049 (2007)

Note added to proof: The Editor of *The New Palgrave* (2008), Professor Steven Durlauf, has very kindly informed the authors that the omission of the entry of the (late) Lionel McKenzie on *turnpike theory* was an oversight, and that it will be reinserted in the on-line version.