Optimal sampled-data control, and generalizations on time scales

Loïc Bourdin, Emmanuel Trélat

▶ To cite this version:

Loïc Bourdin, Emmanuel Trélat. Optimal sampled-data control, and generalizations on time scales. Mathematical Control and Related Fields, AIMS, 2016, 6 (1), pp.53–94. <hal-01110750v2>

HAL Id: hal-01110750

https://hal.archives-ouvertes.fr/hal-01110750v2

Submitted on 6 Dec 2015

Optimal sampled-data control, and generalizations on time scales

Loïc Bourdin*   Emmanuel Trélat†

Abstract

In this paper, we derive a version of the Pontryagin maximum principle for general finite-dimensional nonlinear optimal sampled-data control problems. Our framework is actually much more general, and we treat optimal control problems for which the state variable evolves on a given time scale (arbitrary non-empty closed subset of $\mathbb{R}$), and the control variable evolves on a smaller time scale. Sampled-data systems are then a particular case. Our proof is based on the construction of appropriate needle-like variations and on the Ekeland variational principle.

Keywords: optimal control; sampled-data; Pontryagin maximum principle; time scale.

AMS Classification: 49J15; 93C57; 34N99; 39A12.

1 Introduction

Optimal control theory is concerned with the analysis of controlled dynamical systems, where one aims at steering such a system from a given configuration to some desired target by minimizing some criterion. The Pontryagin maximum principle (in short, PMP), established at the end of the 50's for general finite-dimensional nonlinear continuous-time dynamics (see [46], and see [30] for the history of this discovery), is certainly the milestone of the classical optimal control theory. It provides a first-order necessary condition for optimality, by asserting that any optimal trajectory must be the projection of an extremal. The PMP then reduces the search of optimal trajectories to a boundary value problem posed on extremals. Optimal control theory, and in particular the PMP, has an immense field of applications in various domains, and it is not our aim here to list them.

We speak of a *purely continuous-time optimal control problem*, when both the state q and the control u evolve continuously in time, and the control system under consideration has the form

$$\dot{q}(t) = f(t, q(t), u(t)), \qquad \text{for a.e. } t \in \mathbb{R}^+,$$

where $q(t) \in \mathbb{R}^n$ and $u(t) \in \Omega \subset \mathbb{R}^m$. Such models assume that the control is permanent, that is, the value of $u(t)$ can be chosen at each time $t \in \mathbb{R}^+$. We refer the reader to textbooks on continuous optimal control theory such as [4, 13, 14, 18, 20, 21, 33, 42, 43, 46, 47, 49, 50] for many examples of theoretical or practical applications.

---

*Université de Limoges, Institut de recherche XLIM, Département de Mathématiques et d'Informatique. UMR CNRS 7252. Limoges, France (`loic.bourdin@unilim.fr`).

†Sorbonne Universités, UPMC Univ Paris 06, CNRS UMR 7598, Laboratoire Jacques-Louis Lions, Institut Universitaire de France, F-75005, Paris, France (`emmanuel.trelat@upmc.fr`).

We speak of a *purely discrete-time optimal control problem*, when both the state q and the control u evolve in a discrete way in time, and the control system under consideration has the form

$$q_{k+1} - q_k = f(k, q_k, u_k), \qquad k \in \mathbb{N},$$

where $q_k \in \mathbb{R}^n$ and $u_k \in \Omega \subset \mathbb{R}^m$. As in the continuous case, such models assume that the control is permanent, that is, the value of u_k can be chosen at each time $k \in \mathbb{N}$. A version of the PMP for such discrete-time control systems has been established in [32, 39, 41] under appropriate convexity assumptions. The considerable development of the discrete-time control theory was in particular motivated by the need of considering digital systems or discrete approximations in numerical simulations of differential control systems (see the textbooks [12, 24, 45, 49]). It can be noted that some early works devoted to the discrete-time PMP (like [27]) are mathematically incorrect. Some counterexamples were provided in [12] (see also [45]), showing that, as is now well known, the exact analogous of the continuous-time PMP does not hold at the discrete level. More precisely, the maximization condition of the continuous-time PMP cannot be expected to hold in general in the discrete-time case. Nevertheless, a weaker condition can be derived, in terms of nonpositive gradient condition (see [12, Theorem 42.1]).

We speak of an *optimal sampled-data control problem*, when the state q evolves continuously in time, whereas the control u evolves in a discrete way in time. This hybrid situation is often considered in practice for problems in which the evolution of the state is very quick (and thus can be considered continuous) with respect to that of the control. We often speak, in that case, of *digital control*. This refers to a situation where, due for instance to hardware limitations or to technical difficulties, the value $u(t)$ of the control can be chosen only at times $t = kT$, where $T > 0$ is fixed and $k \in \mathbb{N}$. This means that, once the value $u(kT)$ is fixed, $u(t)$ remains constant over the time interval $[kT, (k+1)T)$. Hence the trajectory q evolves according to

$$\dot{q}(t) = f(t, q(t), u(kT)), \qquad \text{for a.e. } t \in [kT, (k+1)T), \quad k \in \mathbb{N}.$$

In other words, this *sample-and-hold* procedure consists of "freezing" the value of u at each *controlling time* $t = kT$ on the corresponding *sampling time interval* $[kT, (k+1)T)$, where T is called the *sampling period*. In this situation, the control of the system is clearly nonpermanent.

To the best of our knowledge, the classical optimal control theory does not treat general nonlinear optimal sampled-data control problems, but concerns either purely continuous-time, or purely discrete-time optimal control problems. It is one of our objectives to derive, in this paper, a PMP which can be applied to general nonlinear optimal sampled-data control problems.

Actually, we will be able to establish a PMP in the much more general framework of *time scales*, which unifies and extends continuous-time and discrete-time issues. But, before coming to that point, we feel that it is of interest to enunciate a PMP in the particular case of sampled-data systems and where the set Ω of pointwise constraints on the controls is convex.

PMP for optimal sampled-data control problems and Ω convex. Let n, m and j be nonzero integers. Let $T > 0$ be an arbitrary sampling period. In what follows, for any real number t, we denote by $E(t)$ the integer part of t, defined as the unique integer such that $E(t) \leq t < E(t)+1$. Note that $k = E(t/T)$ whenever $kT \leq t < (k+1)T$. We consider the general nonlinear optimal sampled-data control problem

$$(\textbf{OSDCP})_{\mathbb{N}T}^{\mathbb{R}^+} \quad \begin{cases} \min \displaystyle\int_0^{t_f} f^0(\tau, q(\tau), u(k(\tau)T)) \, d\tau, & \text{with } k(\tau) = E(\tau/T), \\ \dot{q}(t) = f(t, q(t), u(k(t)T)), & \text{with } k(t) = E(t/T), \\ u(kT) \in \Omega, \\ g(q(0), q(t_f)) \in \mathrm{S}. \end{cases}$$

Here, $f : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $f^0 : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ are continuous, and of class C^1 in (q, u), $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^j$ is of class C^1, and Ω (resp., S) is a non-empty closed convex subset of $\mathbb{R}^m$ (resp., of $\mathbb{R}^j$). The final time $t_f \geq 0$ can be fixed or not.

Note that, under appropriate (usual) compactness and convexity assumptions, the optimal control problem $(\mathbf{OSDCP})_{\mathbb{N}T}^{\mathbb{R}^+}$ has at least one solution (see Theorem 2 in Section 2.2).

Recall that g is said to be submersive at a point $(q_1, q_2) \in \mathbb{R}^n \times \mathbb{R}^n$ if the differential of g at this point is surjective. We define the Hamiltonian $H : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$, as usual, by $H(t, q, p, p^0, u) = \langle p, f(t, q, u)\rangle_{\mathbb{R}^n} + p^0 f^0(t, q, u)$.

Theorem 1 (Pontryagin maximum principle for $(\mathbf{OSDCP})_{\mathbb{N}T}^{\mathbb{R}^+}$). *If a trajectory q^*, defined on $[0, t_f^*]$ and associated with a sampled-data control u^*, is an optimal solution of $(\mathbf{OSDCP})_{\mathbb{N}T}^{\mathbb{R}^+}$, then there exists a nontrivial couple (p, p^0), where $p : [0, t_f^*] \to \mathbb{R}^n$ is an absolutely continuous mapping (called adjoint vector) and $p^0 \leq 0$, such that the following conditions hold:*

- *Extremal equations:*

$$\dot{q}^*(t) = \frac{\partial H}{\partial p}(t, q^*(t), p(t), p^0, u^*(k(t)T)), \qquad \dot{p}(t) = -\frac{\partial H}{\partial q}(t, q^*(t), p(t), p^0, u^*(k(t)T)),$$

  for almost every $t \in [0, t_f^*)$, with $k(t) = E(t/T)$.

- *Maximization condition:*
  For every controlling time $kT \in [0, t_f^*)$ such that $(k+1)T \leq t_f^*$, we have

$$\left\langle \int_{kT}^{(k+1)T} \frac{\partial H}{\partial u}(\tau, q^*(\tau), p(\tau), p^0, u^*(kT))\, d\tau \ , \ y - u^*(kT) \right\rangle_{\mathbb{R}^m} \leq 0, \tag{1}$$

  for every $y \in \Omega$. In the case where $kT \in [0, t_f^*)$ with $(k+1)T > t_f^*$, the above maximization condition is still valid provided $(k+1)T$ is replaced with t_f^*.

- *Transversality conditions on the adjoint vector:*
  If g is submersive at $(q^*(0), q^*(t_f^*))$, then the nontrivial couple (p, p^0) can be selected to satisfy

$$p(0) = -\left(\frac{\partial g}{\partial q_1}(q^*(0), q^*(t_f^*))\right)^{\top} \psi, \qquad p(t_f^*) = \left(\frac{\partial g}{\partial q_2}(q^*(0), q^*(t_f^*))\right)^{\top} \psi,$$

  where $-\psi$ belongs to the orthogonal of S at the point $g(q^*(0), q^*(t_f^*)) \in$ S.

- *Transversality condition on the final time:*
  If the final time is left free in the optimal control problem $(\mathbf{OSDCP})_{\mathbb{N}T}^{\mathbb{R}^+}$, if $t_f^* > 0$ and if f and f^0 are of class C^1 with respect to t in a neighborhood of t_f^*, then the nontrivial couple (p, p^0) can be moreover selected to satisfy

$$H(t_f^*, q^*(t_f^*), p(t_f^*), p^0, u^*(k_f^*T)) = 0,$$

  where $k_f^* = E(t_f^*/T)$ whenever $t_f^* \notin \mathbb{N}T$, and $k_f^* = E(t_f^*/T) - 1$ whenever $t_f^* \in \mathbb{N}T$.

Note that the only difference with the usual statement of the PMP for purely continuous-time optimal control problems is in the maximization condition. Here, for sampled-data control systems, the usual pointwise maximization condition of the Hamiltonian is replaced with the inequality (1). This is not a surprise, because already in the purely discrete case, as mentioned

earlier, the pointwise maximization condition fails to be true in general, and must be replaced with a weaker condition.

The condition (1), which is satisfied for every $y \in \Omega$, gives a necessary condition allowing to compute $u^*(kT)$ in general, and this, for all controlling times $kT \in [0, t_f^*)$. We will provide in Section 3.1 a simple optimal consumption problem with sampled-data control, and show how these computations can be done in a simple way.

Note that the optimal sampled-data control problem $(\textbf{OSDCP})_{\mathbb{N}T}^{\mathbb{R}^+}$ can of course be seen as a finite-dimensional optimization problem where the unknowns are $u^*(kT)$, with $k \in \mathbb{N}$ such that $kT \in [0, t_f^*)$. The same remark holds, by the way, for purely discrete-time optimal control problems. One could then apply classical Lagrange multiplier (or KKT) rules to such optimization problems with constraints (numerically, this leads to direct methods). The Pontryagin maximum principle is a far-reaching version of the Lagrange multiplier rule, yielding more precise information and reducing the initial optimal control problem to a shooting problem (see, *e.g.*, [51] for such a discussion).

Extension to the time scale framework. In this paper, we actually establish a version of Theorem 1 in a much more general framework, allowing for example to study sampled-data control systems where the control can be permanent on a first time interval, then sampled on a finite set, then permanent again, etc. More precisely, Theorem 1 can be extended to a general framework in which the set of controlling times is not $\mathbb{N}T$ but some arbitrary non-empty closed subset of $\mathbb{R}$ (*i.e.*, a *time scale*), and also the state may evolve on another time scale. We will state a PMP for such general systems in Section 2.3 (see Theorem 3). Since such systems, where we have two time scales (one for the state and one for the control), can be viewed as a generalization of sampled-data control systems, we will refer to them as *sampled-data control systems on time scales*.

Let us first recall and motivate the notion of time scale. The time scale theory was introduced in [34] in order to unify discrete and continuous analysis. By definition, a time scale $\mathbb{T}$ is an arbitrary non-empty closed subset of $\mathbb{R}$, and a dynamical system is said to be posed on the time scale $\mathbb{T}$ whenever the time variable evolves along this set $\mathbb{T}$. The time scale theory aims at closing the gap between continuous and discrete cases, and allows one to treat general processes involving both continuous-time and discrete-time variables. The purely continuous-time case corresponds to $\mathbb{T} = \mathbb{R}^+$ and the purely discrete-time case corresponds to $\mathbb{T} = \mathbb{N}$. But a time scale can be much more general (see, *e.g.*, [29, 44] for a study of a seasonally breeding population whose generations do not overlap, and see [6] for applications to economics), and can even be a Cantor set. Many notions of standard calculus have been extended to the time scale framework, and we refer the reader to [1, 2, 10, 11] for details on that theory.

The theory of the calculus of variations on time scales, initiated in [8], has been well studied in the existing literature (see, *e.g.*, [7, 9, 17, 28, 35, 38]). In [36, 37], the authors establish a weak version of the PMP (with a nonpositive gradient condition) for control systems defined on general time scales. In [16], we derived a strong version of the PMP, in a very general time scale setting, encompassing both the purely continuous-time PMP (with a maximization condition) and the purely discrete-time PMP (with a nonpositive gradient condition).

All these works are concerned with control systems defined on general time scales with permanent control. The main objective of the present paper is to handle control systems defined on general time scales with nonpermanent control, that we refer to as *sampled-data control systems on time scales*, and for which we assume that the state and the control are allowed to evolve on different time scales (the time scale of the control being a subset of the time scale of the state). This framework is the natural extension of the classical sampled-data setting, and allows to treat simultaneously many sampling-data control situations.

Our main result is a *PMP for general finite-dimensional nonlinear optimal sampled-data control*

problems on time scales. Note that our result will be derived without any convexity assumption on the set Ω of pointwise constraints on the controls. Our proof is based on the construction of appropriate needle-like variations and on the Ekeland variational principle. In the case of a permanent control, our statement encompasses the time scale version of the PMP obtained in [16], and *a fortiori* it also encompasses the classical continuous and discrete versions of the PMP.

Organization of the paper. In Section 2, after having recalled several basic facts in time scale calculus, we define a general nonlinear optimal sampled-data control problem defined on time scales, and we state a Pontryagin maximum principle (Theorem 3) for such problems. Section 3 is devoted to some applications of Theorem 3 and further comments. Section 4 is devoted to the proof of Theorem 3.

2 Main result

Let $\mathbb{T}$ be a time scale, that is, an arbitrary non-empty closed subset of $\mathbb{R}$. Without loss of generality, we assume that $\mathbb{T}$ is bounded below, denoting by $a = \min \mathbb{T}$, and unbounded above.[1] Throughout the paper, $\mathbb{T}$ will be the time scale on which the state of the control system evolves.

We start the section by recalling some useful notations and basic results of time scale calculus, in particular the notion of Lebesgue Δ-measure and of absolutely continuous function within the time scale setting. The reader already acquainted with time scale calculus may jump directly to Section 2.2.

2.1 Preliminaries on time scale calculus

The forward jump operator $\sigma : \mathbb{T} \to \mathbb{T}$ is defined by $\sigma(t) = \inf\{s \in \mathbb{T} \mid s > t\}$ for every $t \in \mathbb{T}$. A point $t \in \mathbb{T}$ is said to be *right-scattered* whenever $\sigma(t) > t$. A point $t \in \mathbb{T}$ is said to be *right-dense* whenever $\sigma(t) = t$. We denote by RS the set of all right-scattered points of $\mathbb{T}$, and by RD the set of all right-dense points of $\mathbb{T}$. Note that RS is at most countable (see [23, Lemma 3.1]) and that RD is the complement of RS in $\mathbb{T}$. The graininess function $\mu : \mathbb{T} \to \mathbb{R}^+$ is defined by $\mu(t) = \sigma(t) - t$ for every $t \in \mathbb{T}$.

For every subset A of $\mathbb{R}$, we denote by $A_{\mathbb{T}} = A \cap \mathbb{T}$. An interval of $\mathbb{T}$ is defined by $I_{\mathbb{T}}$ where I is an interval of $\mathbb{R}$. For every $b \in \mathbb{T} \backslash \{a\}$ and every $s \in [a,b)_{\mathbb{T}} \cap \mathrm{RD}$, we set

$$\mathcal{V}^{s,b} = \{\beta \geq 0 \mid s + \beta \in [s,b]_{\mathbb{T}}\}. \tag{2}$$

Note that 0 is not isolated in $\mathcal{V}^{s,b}$.

Δ-differentiability. Let $n \in \mathbb{N}^*$. The notations $\|\cdot\|_{\mathbb{R}^n}$ and $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$ respectively stand for the usual Euclidean norm and scalar product of $\mathbb{R}^n$. A function $q : \mathbb{T} \to \mathbb{R}^n$ is said to be Δ-differentiable at $t \in \mathbb{T}$ if the limit

$$q^{\Delta}(t) = \lim_{\substack{s \to t \\ s \in \mathbb{T}}} \frac{q^{\sigma}(t) - q(s)}{\sigma(t) - s}$$

exists in $\mathbb{R}^n$, where $q^{\sigma} = q \circ \sigma$. Recall that, if $t \in \mathrm{RD}$, then q is Δ-differentiable at t if and only if the limit of $\frac{q(t) - q(s)}{t - s}$ as $s \to t$, $s \in \mathbb{T}$, exists; in that case it is equal to $q^{\Delta}(t)$. If $t \in \mathrm{RS}$ and if q is continuous at t, then q is Δ-differentiable at t, and $q^{\Delta}(t) = \frac{q^{\sigma}(t) - q(t)}{\mu(t)}$ (see [10]).

---

[1]In this paper we only work on a bounded subinterval of type $[a,b] \cap \mathbb{T}$ with a, $b \in \mathbb{T}$. It is not restrictive to assume that $a = \min \mathbb{T}$ and that $\mathbb{T}$ is unbounded above. On the other hand, these two assumptions widely simplify the notations introduced in Section 2.1 (otherwise we would have, in all further statements, to distinguish between points of $\mathbb{T} \backslash \{\max \mathbb{T}\}$ and $\max \mathbb{T}$).

If q, $q' : \mathbb{T} \to \mathbb{R}^n$ are both Δ-differentiable at $t \in \mathbb{T}$, then the scalar product $\langle q, q' \rangle_{\mathbb{R}^n}$ is Δ-differentiable at t and

$$\langle q, q' \rangle_{\mathbb{R}^n}^{\Delta}(t) = \langle q^{\Delta}(t), q'^{\sigma}(t) \rangle_{\mathbb{R}^n} + \langle q(t), q'^{\Delta}(t) \rangle_{\mathbb{R}^n} = \langle q^{\Delta}(t), q'(t) \rangle_{\mathbb{R}^n} + \langle q^{\sigma}(t), q'^{\Delta}(t) \rangle_{\mathbb{R}^n}. \quad (3)$$

These equalities are usually called Leibniz formulas (see [10, Theorem 1.20]).

Lebesgue Δ-measure and Lebesgue Δ-integrability. Let μ_{Δ} be the Lebesgue Δ-measure on $\mathbb{T}$ defined in terms of Carathéodory extension in [11, Chapter 5]. We also refer the reader to [3, 5, 23, 31] for more details on the μ_{Δ}-measure theory. For all $(c, d) \in \mathbb{T}^2$ such that $c \leq d$, one has $\mu_{\Delta}([c, d)_{\mathbb{T}}) = d - c$. Recall that $A \subset \mathbb{T}$ is a μ_{Δ}-measurable set of $\mathbb{T}$ if and only if A is an usual μ_L-measurable set of $\mathbb{R}$, where μ_L denotes the usual Lebesgue measure (see [23, Proposition 3.1]). Moreover, if $A \subset \mathbb{T}$, then

$$\mu_{\Delta}(A) = \mu_L(A) + \sum_{r \in A \cap \mathrm{RS}} \mu(r).$$

Let $A \subset \mathbb{T}$. A property is said to hold Δ-almost everywhere (in short, Δ-a.e.) on A if it holds for every $t \in A \backslash A'$, where $A' \subset A$ is some μ_{Δ}-measurable subset of $\mathbb{T}$ satisfying $\mu_{\Delta}(A') = 0$. In particular, since $\mu_{\Delta}(\{r\}) = \mu(r) > 0$ for every $r \in \mathrm{RS}$, we conclude that, if a property holds Δ-a.e. on A, then it holds for every $r \in A \cap \mathrm{RS}$.

Similarly, if $A \subset \mathbb{T}$ is such that $\mu_{\Delta}(A) = 0$, then $A \subset \mathrm{RD}$.

Let $n \in \mathbb{N}^*$ and let $A \subset \mathbb{T}$ be a μ_{Δ}-measurable subset of $\mathbb{T}$. Consider a function q defined Δ-a.e. on A with values in $\mathbb{R}^n$. Let $\tilde{A} = A \cup (r, \sigma(r))_{r \in A \cap \mathrm{RS}}$, and let $\tilde{q}$ be the extension of q defined μ_L-a.e. on $\tilde{A}$ by $\tilde{q}(t) = q(t)$ whenever $t \in A$, and by $\tilde{q}(t) = q(r)$ whenever $t \in (r, \sigma(r))$, for every $r \in A \cap \mathrm{RS}$. Recall that q is μ_{Δ}-measurable on A if and only if $\tilde{q}$ is μ_L-measurable on $\tilde{A}$ (see [23, Proposition 4.1]).

Let $n \in \mathbb{N}^*$ and let $A \subset \mathbb{T}$ be a μ_{Δ}-measurable subset of $\mathbb{T}$. The functional space $\mathrm{L}_{\mathbb{T}}^{\infty}(A, \mathbb{R}^n)$ is the set of all functions q defined Δ-a.e. on A, with values in $\mathbb{R}^n$, that are μ_{Δ}-measurable on A and bounded Δ-almost everywhere. Endowed with the norm $\|q\|_{\mathrm{L}_{\mathbb{T}}^{\infty}(A, \mathbb{R}^n)} = \sup \mathrm{ess}_{\tau \in A} \|q(\tau)\|_{\mathbb{R}^n}$, it is a Banach space (see [3, Theorem 2.5]). The functional space $\mathrm{L}_{\mathbb{T}}^1(A, \mathbb{R}^n)$ is the set of all functions q defined Δ-a.e. on A, with values in $\mathbb{R}^n$, that are μ_{Δ}-measurable on A and such that $\int_A \|q(\tau)\|_{\mathbb{R}^n} \Delta\tau < +\infty$. Endowed with the norm $\|q\|_{\mathrm{L}_{\mathbb{T}}^1(A, \mathbb{R}^n)} = \int_A \|q(\tau)\|_{\mathbb{R}^n} \Delta\tau$, it is a Banach space (see [3, Theorem 2.5]). We recall here that if $q \in \mathrm{L}_{\mathbb{T}}^1(A, \mathbb{R}^n)$ then

$$\int_A q(\tau) \Delta\tau = \int_{\tilde{A}} \tilde{q}(\tau) \, d\tau = \int_A q(\tau) \, d\tau + \sum_{r \in A \cap \mathrm{RS}} \mu(r) q(r),$$

see [23, Theorems 5.1 and 5.2]. Note that if A is bounded then $\mathrm{L}_{\mathbb{T}}^{\infty}(A, \mathbb{R}^n) \subset \mathrm{L}_{\mathbb{T}}^1(A, \mathbb{R}^n)$.

Absolutely continuous functions. Let $n \in \mathbb{N}^*$ and let $(c, d) \in \mathbb{T}^2$ such that $c < d$. Let $\mathrm{C}([c, d]_{\mathbb{T}}, \mathbb{R}^n)$ denote the space of continuous functions defined on $[c, d]_{\mathbb{T}}$ with values in $\mathbb{R}^n$. Endowed with its usual uniform norm $\|\cdot\|_{\infty}$, it is a Banach space. Let $\mathrm{AC}([c, d]_{\mathbb{T}}, \mathbb{R}^n)$ denote the subspace of absolutely continuous functions.

Let $t_0 \in [c, d]_{\mathbb{T}}$ and $q : [c, d]_{\mathbb{T}} \to \mathbb{R}^n$. It is easily derived from [22, Theorem 4.1] that $q \in \mathrm{AC}([c, d]_{\mathbb{T}}, \mathbb{R}^n)$ if and only if q is Δ-differentiable Δ-a.e. on $[c, d)_{\mathbb{T}}$ and satisfies $q^{\Delta} \in \mathrm{L}_{\mathbb{T}}^1([c, d)_{\mathbb{T}}, \mathbb{R}^n)$, and for every $t \in [c, d]_{\mathbb{T}}$ one has $q(t) = q(t_0) + \int_{[t_0, t)_{\mathbb{T}}} q^{\Delta}(\tau) \Delta\tau$ whenever $t \geq t_0$, and $q(t) = q(t_0) - \int_{[t, t_0)_{\mathbb{T}}} q^{\Delta}(\tau) \Delta\tau$ whenever $t \leq t_0$.

Assume that $q \in \mathrm{L}_{\mathbb{T}}^1([c, d)_{\mathbb{T}}, \mathbb{R}^n)$, and let Q be the function defined on $[c, d]_{\mathbb{T}}$ by $Q(t) = \int_{[t_0, t)_{\mathbb{T}}} q(\tau) \Delta\tau$ whenever $t \geq t_0$, and by $Q(t) = -\int_{[t, t_0)_{\mathbb{T}}} q(\tau) \Delta\tau$ whenever $t \leq t_0$. Then $Q \in \mathrm{AC}([c, d]_{\mathbb{T}})$ and $Q^{\Delta} = q$ Δ-a.e. on $[c, d)_{\mathbb{T}}$.

Note that, if $q \in \mathrm{AC}([c,d]_\mathbb{T}, \mathbb{R}^n)$ is such that $q^\Delta = 0$ Δ-a.e. on $[c,d)_\mathbb{T}$, then q is constant on $[c,d]_\mathbb{T}$, and that, if $q, q' \in \mathrm{AC}([c,d]_\mathbb{T}, \mathbb{R}^n)$, then $\langle q, q' \rangle_{\mathbb{R}^n} \in \mathrm{AC}([c,d]_\mathbb{T}, \mathbb{R})$ and the Leibniz formula (3) is available Δ-a.e. on $[c,d)_\mathbb{T}$.

For every $q \in \mathrm{L}_\mathbb{T}^1([c,d)_\mathbb{T}, \mathbb{R}^n)$, we denote by $\mathscr{L}_{[c,d)_\mathbb{T}}(q)$ the set of points $t \in [c,d)_\mathbb{T}$ that are Δ-Lebesgue points of q. It holds $\mu_\Delta(\mathscr{L}_{[c,d)_\mathbb{T}}(q)) = \mu_\Delta([c,d)_\mathbb{T}) = d - c$, and

$$\lim_{\substack{\beta \to 0 \\ \beta \in \mathcal{V}^{s,d}}} \frac{1}{\beta} \int_{[s,s+\beta)_\mathbb{T}} q(\tau) \, \Delta\tau = q(s),$$

for every $s \in \mathscr{L}_{[c,d)_\mathbb{T}}(q) \cap \mathrm{RD}$, where $\mathcal{V}^{s,d}$ is defined by (2).

2.2 Optimal sampled-data control problems on time scales

Let $\mathbb{T}_1$ be another time scale. Throughout the paper, $\mathbb{T}_1$ will be the time scale on which the control evolves. We assume that $\mathbb{T}_1 \subset \mathbb{T}$.[2]

Similarly to $\mathbb{T}$, we assume that $\min \mathbb{T}_1 = a$ and that $\mathbb{T}_1$ is unbounded above. As in the previous paragraph, we introduce the notations σ_1, RS_1, RD_1, $\mathcal{V}_1^{s,b}$, Δ_1, etc., associated with the time scale $\mathbb{T}_1$. Since $\mathbb{T}_1 \subset \mathbb{T}$, note that $\mathrm{RS} \subset \mathrm{RS}_1$ and $\mathrm{RD}_1 \subset \mathrm{RD}$. We define the map

$$\Phi : \quad \mathbb{T} \quad \longrightarrow \quad \mathbb{T}_1$$
$$t \quad \longmapsto \quad \Phi(t) = \sup\{s \in \mathbb{T}_1 \,|\, s \le t\}.$$

For every $t \in \mathbb{T}_1$, we have $\Phi(t) = t$. For every $t \in \mathbb{T} \backslash \mathbb{T}_1$, we have $\Phi(t) \in \mathrm{RS}_1$ and $\Phi(t) < t < \sigma_1(\Phi(t))$. Note that, if $t \in \mathbb{T}$ is such that $\Phi(t) \in \mathrm{RD}_1$, then $t \in \mathbb{T}_1$.

In what follows, given a function $u : \mathbb{T}_1 \to \mathbb{R}$, we denote by u^Φ the composition $u \circ \Phi : \mathbb{T} \to \mathbb{R}$. Of course, when dealing with functions having multiple components, this composition is applied to each component. Let us mention, at this step, that if $u \in \mathrm{L}_{\mathbb{T}_1}^\infty(\mathbb{T}_1, \mathbb{R})$ then $u^\Phi \in \mathrm{L}_\mathbb{T}^\infty(\mathbb{T}, \mathbb{R})$ (see Proposition 1 and more properties in Section 4.1.1).

Let n, m and j be nonzero integers. We consider the general nonlinear optimal sampled-data control problem on time scales

$$(\mathbf{OSDCP})_{\mathbb{T}_1}^\mathbb{T} \quad \begin{cases} \min \displaystyle\int_{[a,b)_\mathbb{T}} f^0(\tau, q(\tau), u^\Phi(\tau)) \, \Delta\tau, \\ q^\Delta(t) = f(t, q(t), u^\Phi(t)), \\ u \in \mathrm{L}_{\mathbb{T}_1}^\infty(\mathbb{T}_1, \Omega), \\ g(q(a), q(b)) \in \mathrm{S}. \end{cases} \tag{4}$$

Here, the trajectory of the system is $q : \mathbb{T} \to \mathbb{R}^n$, the mappings $f : \mathbb{T} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $f^0 : \mathbb{T} \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ are continuous, of class C^1 in (q, u), the mapping $g : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^j$ is of class C^1, Ω is a non-empty closed subset of $\mathbb{R}^m$, and S is a non-empty closed convex subset of $\mathbb{R}^j$. The final time $b \in \mathbb{T}$ can be fixed or not.

Remark 1. We recall that, given $u \in \mathrm{L}_{\mathbb{T}_1}^\infty(\mathbb{T}_1, \mathbb{R}^m)$, we say that q is a solution of (4) on $I_\mathbb{T}$ if:

1. $I_\mathbb{T}$ is an interval of $\mathbb{T}$ satisfying $a \in I_\mathbb{T}$ and $I_\mathbb{T} \backslash \{a\} \ne \emptyset$;

---

[2]Indeed, it is not natural to consider controlling times $t \in \mathbb{T}_1$ at which the dynamics does not evolve, that is, at which $t \notin \mathbb{T}$. The value of the control at such times $t \in \mathbb{T}_1 \backslash \mathbb{T}$ would not influence the dynamics, or, maybe, only on $[t^*, +\infty[_\mathbb{T}$ where $t^* = \inf\{s \in \mathbb{T} \,|\, s \ge t\}$. In this last case, note that $t^* \in \mathbb{T}$ and we can replace $\mathbb{T}_1$ by $(\mathbb{T}_1 \cup \{t^*\}) \backslash \{t\}$ without loss of generality.

2. For every $c \in I_\mathbb{T} \backslash \{a\}$, $q \in \mathrm{AC}([a,c]_\mathbb{T}, \mathbb{R}^n)$ and (4) holds for Δ-a.e. $t \in [a,c)_\mathbb{T}$.

Existence and uniqueness of solutions (Cauchy-Lipschitz theorem on time scales) have been established in [15], and useful results are recalled in Section 4.1.2.

Remark 2. The time scale $\mathbb{T}_1$ stands for the set of controlling times of the control system (4). If $\mathbb{T} = \mathbb{T}_1$, then the control is permanent. The case $\mathbb{T} = \mathbb{T}_1 = \mathbb{R}^+$ corresponds to the classical continuous case, whereas $\mathbb{T} = \mathbb{T}_1 = \mathbb{N}$ coincides with the classical discrete case. If $\mathbb{T}_1 \subsetneq \mathbb{T}$, the control is nonpermanent and sampled. In that case, the sampling times are given by $t \in \mathrm{RS}_1$ such that $\sigma(t) < \sigma_1(t)$ and the corresponding sampling time intervals are given by $[t, \sigma_1(t))_\mathbb{T}$. The classical optimal sampled-data control problem $(\mathbf{OSDCP})_{\mathbb{N}T}^{\mathbb{R}^+}$ investigated in Theorem 1 corresponds to $\mathbb{T} = \mathbb{R}^+$ and $\mathbb{T}_1 = \mathbb{N}T$, with $T > 0$.

Remark 3. Let us consider two optimal control problems $(\mathbf{OSDCP})_{\mathbb{T}_1}^\mathbb{T}$ and $(\mathbf{OSDCP})_{\mathbb{T}_2}^\mathbb{T}$, posed on the same general time scale $\mathbb{T}$ for the trajectories, but with two different sets of controlling times $\mathbb{T}_1$ and $\mathbb{T}_2$, and let us assume that $\mathbb{T}_2 \subset \mathbb{T}_1$. We denote by Φ_1 and Φ_2 the corresponding mappings from $\mathbb{T}$ to $\mathbb{T}_1$ and $\mathbb{T}_2$ respectively. If $u_1 \in \mathrm{L}_{\mathbb{T}_1}^\infty(\mathbb{T}_1, \Omega)$ is an optimal control for $(\mathbf{OSDCP})_{\mathbb{T}_1}^\mathbb{T}$ and if there exists $u_2 \in \mathrm{L}_{\mathbb{T}_2}^\infty(\mathbb{T}_2, \Omega)$ such that $u_2^{\Phi_2}(t) = u_1^{\Phi_1}(t)$ for Δ-a.e. $t \in \mathbb{T}$, it is clear that u_2 is an optimal control for $(\mathbf{OSDCP})_{\mathbb{T}_2}^\mathbb{T}$. We refer to Section 3.1 for examples.

Remark 4. The framework of $(\mathbf{OSDCP})_{\mathbb{T}_1}^\mathbb{T}$ encompasses optimal parameter problems. Indeed, let us consider the parametrized dynamical system

$$q^\Delta(t) = f(t, q(t), \lambda), \quad \Delta\text{-a.e. } t \in \mathbb{T}, \tag{5}$$

with $\lambda \in \Omega$. Then, considering $\mathbb{T}_1 = \{a\} \cup [b, +\infty[_\mathbb{T}$, (4) coincides with (5) where $u(a)$ plays the role of λ. In this situation, Theorem 3 (stated in Section 2.3) provides necessary conditions for optimal parameters λ. We refer to Section 3.1 for examples.

Remark 5. A possible extension is to study dynamical systems on time scales with several sampled-data controls but with different sets of controlling times:

$$q^\Delta(t) = f(t, q(t), u_1^{\Phi_1}(t), u_2^{\Phi_2}(t)), \quad \Delta\text{-a.e. } t \in \mathbb{T},$$

where $\mathbb{T}_1$ and $\mathbb{T}_2$ are general time scales contained in $\mathbb{T}$, and Φ_1 and Φ_2 are the corresponding mappings from $\mathbb{T}$ to $\mathbb{T}_1$ and $\mathbb{T}_2$. Our main result (Theorem 3) can be easily extended to this framework. Actually, this *multiscale version* will be useful in order to derive the transversality condition on the final time (see Remark 30).

Remark 6. Another possible extension is to study dynamical systems on time scales with sampled-data control where the state q and the constraint function f are also sampled:

$$q^\Delta(t) = f(\Phi_1(t), q^{\Phi_2}(t), u^{\Phi_3}(t)), \quad \Delta\text{-a.e. } t \in \mathbb{T},$$

where $\mathbb{T}_1$, $\mathbb{T}_2$ and $\mathbb{T}_3$ are general time scales contained in $\mathbb{T}$, and Φ_1, Φ_2 and Φ_3 are the corresponding mappings from $\mathbb{T}$ to $\mathbb{T}_1$, $\mathbb{T}_2$ and $\mathbb{T}_3$ respectively. In particular, the setting of [16] corresponds to the above framework with $\mathbb{T} = \mathbb{R}^+$ and $\mathbb{T}_1 = \mathbb{T}_2 = \mathbb{T}_3$ a general time scale.

Although this is not the main objective of our paper, we provide hereafter a result stating the existence of optimal solutions for $(\mathbf{OSDCP})_{\mathbb{T}_1}^\mathbb{T}$, under some appropriate compactness and convexity assumptions. Actually, if the existence of solutions is stated, the necessary conditions provided in Theorem 3, allowing to compute explicitly optimal sampled-data controls, may prove the uniqueness of the optimal solution. We refer to Section 3.1 for examples.

Let $\mathcal{M}$ stand for the set of trajectories q, associated with $b \in \mathbb{T}$ and with a sampled-data control $u \in \mathrm{L}_{\mathbb{T}_1}^\infty(\mathbb{T}_1, \Omega)$, satisfying (4) Δ-a.e. on $[a,b]_\mathbb{T}$ and $g(q(a), q(b)) \in \mathrm{S}$. We define the set of extended velocities $\mathcal{W}(t, q) = \{(f(t,q,u), f^0(t,q,u))^\top \mid u \in \Omega\}$ for every $(t,q) \in \mathbb{T} \times \mathbb{R}^n$.

Theorem 2. *If Ω is compact, $\mathcal{M}$ is non-empty, $\|q\|_\infty + b \leq M$ for every $q \in \mathcal{M}$ and for some $M \geq 0$, and if $\mathcal{W}(t,q)$ is convex for every $(t,q) \in \mathbb{T} \times \mathbb{R}^n$, then $(\mathbf{OSDCP})_{\mathbb{T}_1}^{\mathbb{T}}$ has at least one optimal solution.*

The proof of Theorem 2 is done in Section 4.4. Note that, in this theorem, it suffices to assume that g is continuous. Besides, the assumption on the boundedness of trajectories can be weakened, by assuming, for instance, that the extended dynamics have a sublinear growth at infinity (see, *e.g.*, [25]; many other easy and standard extensions are possible).

2.3 Pontryagin maximum principle for $(\mathbf{OSDCP})_{\mathbb{T}_1}^{\mathbb{T}}$

2.3.1 Preliminaries on convexity and stable Ω-dense directions

The orthogonal of the closed convex set S at a point $x \in$ S is defined by

$$\mathcal{O}_{\mathrm{S}}[x] = \{x' \in \mathbb{R}^j \mid \forall x'' \in \mathrm{S}, \ \langle x', x'' - x\rangle_{\mathbb{R}^j} \leq 0\}.$$

It is a closed convex cone containing 0.

We denote by d_{S} the distance function to S defined by $d_{\mathrm{S}}(x) = \inf_{x' \in \mathrm{S}} \|x - x'\|_{\mathbb{R}^j}$, for every $x \in \mathbb{R}^j$. Recall that, for every $x \in \mathbb{R}^j$, there exists a unique element $\mathrm{P}_{\mathrm{S}}(x) \in$ S (projection of x onto S) such that $d_{\mathrm{S}}(x) = \|x - \mathrm{P}_{\mathrm{S}}(x)\|_{\mathbb{R}^j}$. It is characterized by the property $\langle x - \mathrm{P}_{\mathrm{S}}(x), x' - \mathrm{P}_{\mathrm{S}}(x)\rangle_{\mathbb{R}^j} \leq 0$ for every $x' \in$ S. In particular, $x - \mathrm{P}_{\mathrm{S}}(x) \in \mathcal{O}_{\mathrm{S}}[\mathrm{P}_{\mathrm{S}}(x)]$. The function P_{S} is 1-Lipschitz continuous. We recall the following obvious lemmas.

Lemma 1. *Let $(x_k)_{k\in\mathbb{N}}$ be a sequence of points of $\mathbb{R}^j$ and $(\zeta_k)_{k\in\mathbb{N}}$ be a sequence of nonnegative real numbers such that $x_k \to x \in$ S and $\zeta_k(x_k - \mathrm{P}_{\mathrm{S}}(x_k)) \to x' \in \mathbb{R}^j$ as $k \to +\infty$. Then $x' \in \mathcal{O}_{\mathrm{S}}[x]$.*

Lemma 2. *The function $d_{\mathrm{S}}^2 : x \mapsto d_{\mathrm{S}}(x)^2$ is differentiable on $\mathbb{R}^j$, with $\mathrm{d}d_{\mathrm{S}}^2(x)(x') = 2\langle x - \mathrm{P}_{\mathrm{S}}(x), x'\rangle_{\mathbb{R}^j}$.*

Hereafter we recall the notion of stable Ω-dense directions and we state an obvious lemma. We refer to [16, Section 2.2] for more details.

Definition 1. *Let $v \in \Omega$. A direction $y \in \Omega$ is said to be a stable Ω-dense direction from v if there exists $\varepsilon > 0$ such that 0 is not isolated in $\{\alpha \in [0,1], \ v' + \alpha(y-v') \in \Omega\}$ for every $v' \in \overline{\mathrm{B}}_{\mathbb{R}^m}(v,\varepsilon) \cap \Omega$. The set of all stable Ω-dense directions from v is denoted by $\mathrm{D}_{\mathrm{stab}}^\Omega(v)$.*

Lemma 3. *If Ω is convex, then $\mathrm{D}_{\mathrm{stab}}^\Omega(v) = \Omega$ for every $v \in \Omega$.*

2.3.2 Main result

Recall that g is said to be submersive at a point $(q_1, q_2) \in \mathbb{R}^n \times \mathbb{R}^n$ if the differential of g at this point is surjective. We define the Hamiltonian $H : \mathbb{T} \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$ of $(\mathbf{OSDCP})_{\mathbb{T}_1}^{\mathbb{T}}$ by $H(t,q,p,p^0,u) = \langle p, f(t,q,u)\rangle_{\mathbb{R}^n} + p^0 f^0(t,q,u)$.

Theorem 3 (Pontryagin maximum principle for $(\mathbf{OSDCP})_{\mathbb{T}_1}^{\mathbb{T}}$). *If a trajectory q^*, defined on $[a,b^*]_{\mathbb{T}}$ and associated with a sampled-data control $u^* \in \mathrm{L}_{\mathbb{T}_1}^\infty(\mathbb{T}_1, \Omega)$, is an optimal solution of $(\mathbf{OSDCP})_{\mathbb{T}_1}^{\mathbb{T}}$, then there exists a nontrivial couple (p, p^0), where $p \in \mathrm{AC}([a,b^*]_{\mathbb{T}}, \mathbb{R}^n)$ (called adjoint vector) and $p^0 \leq 0$, such that the following conditions hold:*

- *Extremal equations:*

$$q^{*\Delta}(t) = \frac{\partial H}{\partial p}(t, q^*(t), p^\sigma(t), p^0, u^{*\Phi}(t)), \qquad p^\Delta(t) = -\frac{\partial H}{\partial q}(t, q^*(t), p^\sigma(t), p^0, u^{*\Phi}(t)), \quad (6)$$

for Δ-a.e. $t \in [a,b^*)_{\mathbb{T}}$.

- *Maximization condition:*

  - *For Δ_1-a.e. $s \in [a, b^*)_{\mathbb{T}_1} \cap \mathrm{RD}_1$, we have*

  $$u^*(s) \in \arg\max_{z \in \Omega} H(s, q^*(s), p(s), p^0, z).$$

  - *For every $r \in [a, b^*)_{\mathbb{T}_1} \cap \mathrm{RS}_1$ such that $\sigma_1(r) \leq b^*$, we have*

  $$\left\langle \int_{[r, \sigma_1(r))_{\mathbb{T}}} \frac{\partial H}{\partial u}(\tau, q^*(\tau), p^\sigma(\tau), p^0, u^*(r)) \, \Delta\tau \, , \, y - u^*(r) \right\rangle_{\mathbb{R}^m} \leq 0, \tag{7}$$

  for every $y \in \mathrm{D}^\Omega_{\mathrm{stab}}(u^*(r))$. In the case where $r \in [a, b^*)_{\mathbb{T}_1} \cap \mathrm{RS}_1$ with $\sigma_1(r) > b^*$, the above maximization condition is still valid provided $\sigma_1(r)$ is replaced with b^*.

- *Transversality conditions on the adjoint vector:*
  If g is submersive at $(q^*(a), q^*(b^*))$, then the nontrivial couple (p, p^0) can be selected to satisfy

  $$p(a) = -\left(\frac{\partial g}{\partial q_1}(q^*(a), q^*(b^*))\right)^\top \psi, \qquad p(b^*) = \left(\frac{\partial g}{\partial q_2}(q^*(a), q^*(b^*))\right)^\top \psi, \tag{8}$$

  where $-\psi \in \mathcal{O}_{\mathrm{S}}[g(q^*(a), q^*(b^*))]$.

- *Transversality condition on the final time:*
  If the final time is left free in the optimal control problem $(\mathbf{OSDCP})^{\mathbb{T}}_{\mathbb{T}_1}$, if b^* belongs to the interior of $\mathbb{T}$ (for the topology of $\mathbb{R}$), and if f and f^0 are of class C^1 with respect to t in a neighborhood of b^*, then the nontrivial couple (p, p^0) can be moreover selected such that the Hamiltonian function $t \mapsto H(t, q^*(t), p(t), p^0, u^{*\Phi}(t))$ coincides almost everywhere, in some neighborhood of b^*, with a continuous function vanishing at $t = b^*$.

  In particular, if $u^{*\Phi}(t)$ has a left-limit at $t = b^*$ (denoted by $u^{*\Phi}(b^*_-)$), then the transversality condition can be written as

  $$H(b^*, q^*(b^*), p(b^*), p^0, u^{*\Phi}(b^*_-)) = 0.$$

Theorem 3 is proved in Section 4. Several remarks are in order.

Remark 7. As is well known, the nontrivial couple (p, p^0) of Theorem 3, which is a Lagrange multiplier, is defined up to a multiplicative scalar. Defining as usual an *extremal* as a quadruple (q, p, p^0, u) solution of the extremal equations (6), an extremal is said to be *normal* whenever $p^0 \neq 0$ and *abnormal* whenever $p^0 = 0$. In the normal case $p^0 \neq 0$, it is usual to normalize the Lagrange multiplier so that $p^0 = -1$.

Remark 8. Theorem 3 encompasses the time scale version of the PMP derived in [16] when the control is permanent, that is, when $\mathbb{T}_1 = \mathbb{T}$. Indeed, in that case, for every $r \in [a, b^*)_{\mathbb{T}} \cap \mathrm{RS}$, $r \in [a, b^*)_{\mathbb{T}_1} \cap \mathrm{RS}_1$ and $\sigma_1(r) = \sigma(r) \leq b^*$. Then the condition (7) can be written as the nonpositive gradient condition

$$\left\langle \frac{\partial H}{\partial u}(r, q^*(r), p(\sigma(r)), p^0, u^*(r)) \, , \, y - u^*(r) \right\rangle_{\mathbb{R}^m} \leq 0,$$

for every $y \in \mathrm{D}^\Omega_{\mathrm{stab}}(u^*(r))$. Moreover, in the case of a free final time, under the assumptions made in the fourth item of Theorem 3, b^* also belongs to the interior of $\mathbb{T}_1 = \mathbb{T}$, and then in that case we recover the classical condition

$$\max_{z \in \Omega} H(b^*, q^*(b^*), p(b^*), p^0, z) = 0.$$

A fortiori, Theorem 3 encompasses both the classical continuous-time and discrete-time versions of the PMP, that is respectively, when $\mathbb{T} = \mathbb{T}_1 = \mathbb{R}^+$ and $\mathbb{T} = \mathbb{T}_1 = \mathbb{N}$.

Remark 9. If the final time is free, under the assumptions made in the fourth item of Theorem 3, and if moreover $b^* \notin \mathbb{T}_1$ (or if b^* is a left-scattered point of $\mathbb{T}_1$), $u^{*\Phi}$ is constant on $[\rho_1(b^*), b^*)_\mathbb{T}$ where $\rho_1(b^*) = \max\{s \in \mathbb{T}_1 \mid s < b^*\} < b^*$ (and thus, in particular, $u^{*\Phi}(t)$ has a left-limit at $t = b^*$), and therefore $H(b^*, q^*(b^*), p(b^*), p^0, u^*(\rho_1(b^*))) = 0$. This is similar to the situation of Theorem 1.

Remark 10. Let us describe some typical situations of terminal conditions $g(q(a), q(b)) \in \mathrm{S}$ in $(\mathbf{OSDCP})_{\mathbb{T}_1}^\mathbb{T}$, and of the corresponding transversality conditions on the adjoint vector.

- If the initial and final points are fixed in $(\mathbf{OSDCP})_{\mathbb{T}_1}^\mathbb{T}$, that is, if we impose $q(a) = q_a$ and $q(b) = q_b$, then $j = 2n$, $g(q_1, q_2) = (q_1, q_2)$ and $\mathrm{S} = \{q_a\} \times \{q_b\}$. In that case, the transversality conditions on the adjoint vector give no additional information.

- If the initial point is fixed, that is, if we impose $q(a) = q_a$, and if the final point is left free in $(\mathbf{OSDCP})_{\mathbb{T}_1}^\mathbb{T}$, then $j = n$, $g(q_1, q_2) = q_1$ and $\mathrm{S} = \{q_a\}$. In that case, the transversality conditions on the adjoint vector imply that $p(b^*) = 0$. Moreover, we have $p^0 \neq 0$[3] and we can normalize the Lagrange multiplier so that $p^0 = -1$ (see Remark 7).

- If the initial point is fixed, that is, if we impose $q(a) = q_a$, and if the final point is subject to the constraint $G(q(b)) \in (\mathbb{R}^+)^k$ in $(\mathbf{OSDCP})_{\mathbb{T}_1}^\mathbb{T}$, where $G = (G^1, \ldots, G^k) : \mathbb{R}^n \to \mathbb{R}^k$ is of class C^1 and is submersive at any point $q_2 \in G^{-1}((\mathbb{R}^+)^k)$, then $j = n + k$, $g(q_1, q_2) = (q_1, G(q_2))$ and $\mathrm{S} = \{q_a\} \times (\mathbb{R}^+)^k$. The transversality conditions on the adjoint vector imply that

$$p(b^*) = \sum_{i=1}^k \lambda_i \partial_{q_2} G^i(q^*(b^*)),$$

  with $\lambda_i \geq 0$, $i = 1, \ldots, k$.

- If the periodic condition $q(a) = q(b)$ is imposed in $(\mathbf{OSDCP})_{\mathbb{T}_1}^\mathbb{T}$, then $j = n$, $g(q_1, q_2) = q_1 - q_2$ and $\mathrm{S} = \{0\}$. In that case, the transversality conditions on the adjoint vector yield that $p(a) = p(b^*)$.

We stress that, in all examples above, the function g is indeed a submersion.

Remark 11. In the case where g is not submersive at $(q^*(a), q^*(b^*))$, to obtain transversality conditions on the adjoint vector, Theorem 3 can be reformulated as follows:

> *If a trajectory q^*, defined on $[a, b^*]_\mathbb{T}$ and associated with a sampled-data control $u^* \in \mathrm{L}_{\mathbb{T}_1}^\infty(\mathbb{T}_1, \Omega)$, is an optimal solution of $(\mathbf{OSDCP})_{\mathbb{T}_1}^\mathbb{T}$, then there exists a nontrivial couple $(\psi, p^0) \in \mathbb{R}^j \times \mathbb{R}$, with $-\psi \in \mathcal{O}_\mathrm{S}[g(q^*(a), q^*(b^*))]$ and $p^0 \leq 0$, and there exists $p \in \mathrm{AC}([a, b^*]_\mathbb{T}, \mathbb{R}^n)$ such that the extremal equations, the maximization conditions and the transversality conditions are satisfied.*

However, with this formulation, the couple (p, p^0) may be trivial[4] and, as a consequence, the result may not provide any information. We refer to Sections 4.3.3 and 4.3.4 for more details.

---

[3]Indeed, if $p^0 = 0$, then the adjoint vector p is trivial from the extremal equation and from the final condition $p(b^*) = 0$. This leads to a contradiction since the couple (p, p^0) has to be nontrivial.

[4]Indeed, if $p^0 = 0$ and ψ belongs to the kernel of $(\partial_{q_2} g(q^*(a), q^*(b^*)))^\top$, then the couple (p, p^0) is trivial. This situation leads to a contradiction if g is submersive at $(q^*(a), q^*(b^*))$.

Remark 12. In this paper, the closedness of Ω is used in a crucial way in our proof of the PMP. Indeed, the closure of Ω allows us to define the Ekeland functional on a complete metric space (see Section 4.3.2). However, if the initial point is fixed, that is, if we impose $q(a) = q_a$, and if the final point is left free in $(\mathbf{OSDCP})_{\mathbb{T}_1}^{\mathbb{T}}$, then Theorem 3 can be proved with a simple calculus of variations, without using the Ekeland variational principle. In this particular case, the closedness assumption can be removed.

Remark 13. If the cost functional to be minimized in $(\mathbf{OSDCP})_{\mathbb{T}_1}^{\mathbb{T}}$ is $\int_{[a,b)_{\mathbb{T}}} f^0(\tau, q(\tau), u^{\Phi}(\tau)) \, \Delta\tau +$ $\ell(b, q(a), q(b))$, where $\ell : \mathbb{T} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is continuous, of class C^1 in (q_1, q_2), then the transversality conditions on the adjoint vector become

$$p(a) = - \left(\partial_{q_1} g(q^*(a), q^*(b^*)) \right)^{\top} \psi - p^0 \partial_{q_1} \ell(b^*, q^*(a), q^*(b^*)),$$
$$\text{and } p(b^*) = \left(\partial_{q_2} g(q^*(a), q^*(b^*)) \right)^{\top} \psi + p^0 \partial_{q_2} \ell(b^*, q^*(a), q^*(b^*)).$$

Moreover, in the fourth item of Theorem 3, if ℓ is of class C^1 in a neighborhood of b^*, the transversality condition on the final time must be replaced by:

The nontrivial couple (p, p^0) can be selected such that the Hamiltonian function $t \mapsto$ $H(t, q^*(t), p(t), p^0, u^{*\Phi}(t))$ coincides almost everywhere, in some neighborhood of b^*, with a continuous function that is equal to $-p^0 \partial_b \ell(b^*, q^*(a), q^*(b^*))$ at $t = b^*$.

To prove this claim, it suffices to modify accordingly the Ekeland functional in the proof of Theorem 3 (see Section 4.3.2).

3 Applications and further comments

In this section, we first give, in Section 3.1, a very simple example of an optimal control problem on time scales with sampled-data control, that we treat in details and on which all computations are explicit. The interest is that this example provides as well a simple situation where it is evident that some of the properties that are valid in the classical continous-time PMP do not hold anymore in the time-scale context. We gather these remarks in Section 3.2.

3.1 A model for optimal consumption with sampled-data control

Throughout this subsection, $\mathbb{T}$ and $\mathbb{T}_1$ are two time scales, unbounded above, satisfying $\mathbb{T}_1 \subset \mathbb{T}$, $\min \mathbb{T} = \min \mathbb{T}_1 = 0$ and $12 \in \mathbb{T}$. In the sequel, we study the following one-dimensional dynamical system with sampled-data control on time scales:

$$q^{\Delta}(t) = u^{\Phi}(t) q(t), \qquad \Delta\text{-a.e. } t \in [0, 12)_{\mathbb{T}}, \tag{9}$$

with the initial condition $q(0) = 1$, and subject to the constraint $u(t) \in [0, 1]$ for Δ_1-a.e. $t \in [0, 12)_{\mathbb{T}_1}$. Since the final time $b = 12$ is fixed, we can assume that $12 \in \mathbb{T}_1$ without loss of generality.

The above model is a classical model for the evolution of a controlled output of a factory during the time interval $[0, 12]_{\mathbb{T}}$ (corresponding to the twelve months of a year). Precisely, $q(t) \in \mathbb{R}$ stands for the output at time $t \in \mathbb{T}$ and $u(t) \in [0, 1]$ stands for the fraction of the output reinvested at each controlling time $t \in \mathbb{T}_1$. We assume that this fraction is sampled at each sampling time $t \in \mathbb{T}_1$ such that $t \in \mathrm{RS}_1$ and $\sigma(t) < \sigma_1(t)$, on the corresponding sampling interval $[t, \sigma_1(t))_{\mathbb{T}}$. In the sequel, our goal is to maximize the total consumption

$$C(u) = \int_{[0,12)_{\mathbb{T}}} (1 - u^{\Phi}(\tau)) q(\tau) \, \Delta\tau. \tag{10}$$

In other words, our aim is to maximize the quantity of the output that we do not reinvest.

Remark 14. In the continuous case and with a permanent control (that is, with $\mathbb{T} = \mathbb{T}_1 = \mathbb{R}^+$), the above optimal control problem is a very well-known application of the classical Pontryagin maximum principle. We refer for example to [48, Exercice 2.3.3. p.82] or [40, p.92]. In this section, our aim is to solve this optimal control problem in cases where the control is nonpermanent and sampled. We first treat some examples in the continuous-time setting $\mathbb{T} = \mathbb{R}^+$ and then in the discrete-time setting $\mathbb{T} = \mathbb{N}$.

The above optimal sampled-data control problem corresponds to $(\mathbf{OSDCP})_{\mathbb{T}_1}^{\mathbb{T}}$ with $n = m = j = 1$, $a = 0$, $b = 12$ (fixed final time), $\Omega = [0, 1]$ (convex), $g(q_1, q_2) = q_1$ and $\mathrm{S} = \{1\}$ (fixed initial value $q(0) = 1$ and free final value $q(12)$), $f(t, q, u) = uq$ and $f^0(t, q, u) = (u - 1)q$ (our aim is to minimize $-C(u)$).

Since f and f^0 are affine in u and since Ω is compact, Theorem 2 asserts that $(\mathbf{OSDCP})_{\mathbb{T}_1}^{\mathbb{T}}$ admits an optimal solution q^*, defined on $[0, 12]_{\mathbb{T}}$ and associated with a sampled-data control $u^* \in \mathrm{L}_{\mathbb{T}_1}^{\infty}(\mathbb{T}_1, [0, 1])$. We now apply Theorem 3 in order to compute explicitly the values of u^* at each controlling time $t \in [0, 12)_{\mathbb{T}_1}$. The nontrivial couple (p, p^0) satisfies $p^0 = -1$ and $p(12) = 0$ (see Remark 10). The adjoint vector $p \in \mathrm{AC}([0, 12]_{\mathbb{T}}, \mathbb{R})$ is a solution of

$$p^{\Delta}(t) = -u^{*\Phi}(t)p^{\sigma}(t) + u^{*\Phi}(t) - 1, \qquad \Delta\text{-a.e. } t \in [0, 12)_{\mathbb{T}}. \tag{11}$$

Moreover, one has the following maximization conditions:

1. for Δ_1-a.e. $s \in [0, 12)_{\mathbb{T}_1} \cap \mathrm{RD}_1$,

$$u^*(s) \in \arg\max_{z \in [0,1]} \left(z(p(s) - 1) + 1 \right) q^*(s). \tag{12}$$

2. for every $r \in [0, 12)_{\mathbb{T}_1} \cap \mathrm{RS}_1$,

$$\int_{[r, \sigma_1(r))_{\mathbb{T}}} q^*(\tau)(p^{\sigma}(\tau) - 1)(y - u^*(r)) \, \Delta\tau \leq 0, \tag{13}$$

for every $y \in \mathrm{D}_{\mathrm{stab}}^{\Omega}(u^*(r)) = \Omega = [0, 1]$ (since Ω is convex, see Lemma 3).

Since q^* is a solution of (9) and satisfies $q^*(0) = 1$, one can easily see that q^* is monotonically increasing on $[0, 12]_{\mathbb{T}}$ and then q^* has positive values. From (11) and since $p(12) = 0$, one can easily obtain that p is monotonically decreasing on $[0, 12]_{\mathbb{T}}$ and then p has nonnegative values.

3.1.1 Continuous-time setting $\mathbb{T} = \mathbb{R}^+$

Case $\mathbb{T}_1 = \mathbb{R}^+$ (permanent control). Since q^* has positive values, one can conclude from (11) and (12) that:

$$u^*(t) = \begin{cases} 1 & \text{for a.e. } t \text{ such that } p(t) - 1 > 0, \\ 0 & \text{for a.e. } t \text{ such that } p(t) - 1 < 0, \end{cases} \quad \dot{p}(t) = \begin{cases} -p(t) & \text{for a.e. } t \text{ such that } u^*(t) = 1, \\ -1 & \text{for a.e. } t \text{ such that } u^*(t) = 0. \end{cases}$$

Since $p(12) = 0$, one can easily prove that the optimal (permanent) control u is unique and given by $u^*(t) = 1$ for a.e. $t \in [0, 11)$ and $u^*(t) = 0$ for a.e. $t \in [11, 12)$. The associated optimal consumption is $C(u^*) = e^{11} \simeq 59874.142$. We refer to [48, Exercice 2.3.3. p.82] or [40, p.92] for more details.

Case $\mathbb{T}_1$ discrete (sampled-data control). Solving the differential equations (9) and (11) leads to $q^*(t) = q^*(\sigma_1(r))e^{u^*(r)(t-\sigma_1(r))}$ and

$$p^\sigma(t) = p(t) = \begin{cases} p(\sigma_1(r))e^{-u^*(r)(t-\sigma_1(r))} + \frac{u^*(r)-1}{u^*(r)}(1 - e^{-u^*(r)(t-\sigma_1(r))}) & \text{if } u^*(r) \neq 0, \\ p(\sigma_1(r)) + \sigma_1(r) - t & \text{if } u^*(r) = 0, \end{cases}$$

for every $r \in [0,12)_{\mathbb{T}_1}$ and every $t \in [r, \sigma_1(r)]$. Then, (13) can be written as $q^*(\sigma_1(r))(y - u^*(r))\Gamma_r(u^*(r)) \leq 0$, where $\Gamma_r : [0,1] \to \mathbb{R}$ is the continuous function given by:

$$\Gamma_r(x) = \begin{cases} \frac{1}{x^2}\left[e^{-\mu_1(r)x} - \left(1 + \mu_1(r)x(x(1 - p(\sigma_1(r))) - 1)\right)\right] & \text{if } x \neq 0, \\ \mu_1(r)\left(p(\sigma_1(r)) + \frac{\mu_1(r)}{2} - 1\right) & \text{if } x = 0. \end{cases}$$

Since (13) holds true for every $y \in [0,1]$ and since q^* has positive values, the following properties are satisfied for every $r \in [0,12)_{\mathbb{T}_1}$:

- if Γ_r has negative values on $(0,1]$, then $u^*(r) = 0$;

- if Γ_r has positive values on $[0,1)$, then $u^*(r) = 1$;

- if $\Gamma_r(0) > 0$ and $\Gamma_r(1) < 0$, then $u^*(r) \in (0,1)$ is a solution of the nonlinear equation $\Gamma_r(x) = 0$.

Note that Γ_r depends only on $\mu_1(r)$ and $p(\sigma_1(r))$. As a consequence, the knowledge of the value $p(12) = 0$ and the above properties allow to compute $u^*(r_0)$ where r_0 is the element of $[0,12)_{\mathbb{T}_1}$ such that $\sigma_1(r_0) = 12$ (and $\mu_1(r_0) = 12 - r_0$). Then, the knowledge of $u^*(r_0)$ allows to compute $p(r_0)$ from (11). Then, the knowledge of $p(r_0)$ and the above properties allow to compute $u^*(r_1)$ where r_1 is the element of $[0,12)_{\mathbb{T}_1}$ such that $\sigma_1(r_1) = r_0$ (and $\mu_1(r_1) = r_0 - r_1$), etc. Actually, this recursive procedure allows to compute $u^*(r)$ for every $r \in [0,12)_{\mathbb{T}_1}$. Numerically, we obtain the following results:

$\mathbb{T}_1 = \mathbb{N}$	$u^*(t) = \begin{cases} 1 & \text{if } t \in \{0,1,\ldots,10\} \\ 0 & \text{if } t = 11 \end{cases}$	$C(u^*) = e^{11} \simeq 59874.142$
$\mathbb{T}_1 = 2\mathbb{N}$	$u^*(t) = \begin{cases} 1 & \text{if } t \in \{0,2,4,6,8\} \\ 0 & \text{if } t = 10 \end{cases}$	$C(u^*) = 2e^{10} \simeq 44052.932$
$\mathbb{T}_1 = 3\mathbb{N}$	$u^*(t) = \begin{cases} 1 & \text{if } t \in \{0,3,6\} \\ 0.4536 & \text{if } t = 9 \end{cases}$	$C(u^*) \simeq 28299.767$
$\mathbb{T}_1 = 4\mathbb{N}$	$u^*(t) = \begin{cases} 1 & \text{if } t \in \{0,4\} \\ 0.6392 & \text{if } t = 8 \end{cases}$	$C(u^*) \simeq 20013.885$
$\mathbb{T}_1 = 9\mathbb{N}$	$u^*(t) = \begin{cases} 1 & \text{if } t = 0 \\ 0.4536 & \text{if } t = 9 \end{cases}$	$C(u^*) \simeq 28299.767$
$\mathbb{T}_1 = 12\mathbb{N}$	$u^*(0) \simeq 0.9083$	$C(u^*) \simeq 5467.24$
$\mathbb{T}_1 = 12\mathbb{N} \cup \{10, 11.5\}$	$u^*(t) = \begin{cases} 1 & \text{if } t = 0 \\ 0.9072 & \text{if } t = 10 \\ 0 & \text{if } t = 11.5 \end{cases}$	$C(u^*) \simeq 49476.604$

Remark 15. In this example, Theorem 2 states the existence of an optimal solution. In all cases above studied, the Pontryagin maximum principle proves that the optimal solution is moreover unique.

Remark 16. The case $\mathbb{T}_1 = \mathbb{N}$ can be easily deduced from the permanent case $\mathbb{T}_1 = \mathbb{R}^+$ (see Remark 3). Similarly, the case $\mathbb{T}_1 = 9\mathbb{N}$ can be deduced from the case $\mathbb{T}_1 = 3\mathbb{N}$.

Remark 17. The case $\mathbb{T}_1 = 12\mathbb{N}$ corresponds to an optimal parameter problem (see Remark 4).

Remark 18. For the needs of Section 3.2, let us give some details on the case $\mathbb{T}_1 = 2\mathbb{N}$. In that case, $q^*(t) = e^t$ on $[0, 10]$ and $q^*(t) = e^{10}$ on $[10, 12]$. Moreover, $p(t) = 2e^{10-t}$ on $[0, 10]$ and $p(t) = 12 - t$ on $[10, 12]$. Hence, for every $t \in [10, 11)$, $\arg\max_{z \in \Omega} H(t, q^*(t), p(t), p^0, z) = \{1\}$. The Hamiltonian $t \mapsto H(t, q^*(t), p(t), p^0, u^{*\Phi}(t))$ is equal to $t \mapsto 2e^{10}$ a.e. on $[0, 10)$ and is equal to $t \mapsto e^{10}$ a.e. on $[10, 12]$. Finally, the maximized Hamiltonian $t \mapsto \max_{z \in \Omega} H(t, q^*(t), p^\sigma(t), p^0, z)$ is equal to $t \mapsto 2e^{10}$ on $[0, 10)$, to $t \mapsto (12 - t)e^{10}$ on $[10, 11)$ and to $t \mapsto e^{10}$ on $[11, 12]$.

Case $\mathbb{T}_1$ hybrid (sampled-data control). In this paragraph, we study the hybrid case $\mathbb{T}_1 = [0, 6] \cup \{10\} \cup [11.5, +\infty)$. Similarly to the permanent case $\mathbb{T}_1 = \mathbb{R}^+$, one can easily conclude from (12) that:

$$u^*(t) = \begin{cases} 1 & \text{for a.e. } t \in [0, 6) \cup [11.5, 12) \text{ such that } p(t) - 1 > 0, \\ 0 & \text{for a.e. } t \in [0, 6) \cup [11.5, 12) \text{ such that } p(t) - 1 < 0. \end{cases}$$

Since $p(12) = 0$, one can easily prove that $u^*(t) = 0$ for a.e. $t \in [11.5, 12)$, and then $p(11.5) = 0.5$. From the knowledge of $p(11.5) = 0.5$ and using similar arguments than in the previous paragraph, one can compute $u^*(10) \simeq 0.9072$. From (11), it gives $p(10) \simeq 2.2462$. From the knowledge of $p(10) \simeq 2.2462$, one can compute $u^*(6) = 1$. From (11), it gives $p(6) \simeq 122.6402$. Since p is monotonically decreasing, we conclude that $p(t) - 1 > 0$ for every $t \in [0, 6)$. We finally conclude that the optimal sampled-data control u^* is unique and is given by $u^*(t) = 1$ for a.e. $t \in [0, 6)$, $u^*(6) = 1$, $u^*(10) \simeq 0.9072$ and $u^*(t) = 0$ for a.e. $t \in [11.5, 12)$. The associated optimal consumption is $C(u^*) \simeq 49476.604$.

Remark 19. The discrete case $\mathbb{T}_1 = 12\mathbb{N} \cup \{10, 11.5\}$ can now be seen as a consequence of the hybrid case $\mathbb{T}_1 = [0, 6] \cup \{10\} \cup [11.5, +\infty)$ (see Remark 3).

3.1.2 Discrete-time setting $\mathbb{T} = \mathbb{N}$

In the discrete-time setting $\mathbb{T} = \mathbb{N}$, similarly to the continuous-time one, we can prove that (13) can be written as $q^*(r)(y - u^*(r))\Lambda_r(u^*(r)) \leq 0$, where $\Lambda_r : [0, 1] \to \mathbb{R}$ is the continuous function given by:

$$\Lambda_r(x) = \begin{cases} \dfrac{1}{x^2}\left[1 - (1 + x)^{\mu_1(r)-1}\left(1 + \mu_1(r)x(x(1 - p(\sigma_1(r))) - 1) + x\right)\right] & \text{if } x \neq 0, \\ \mu_1(r)\left(p(\sigma_1(r)) + \dfrac{\mu_1(r) - 3}{2}\right) & \text{if } x = 0. \end{cases}$$

The same recursive procedure (than in the previous section) allows to compute $u^*(r)$ for every $r \in [0, 12)_{\mathbb{T}_1}$. Numerically, we obtain the following results:

$\mathbb{T}_1 = 2\mathbb{N}$	$u^*(t) = \begin{cases} 1 & \text{if } t \in \{0, 2, 4, 6, 8\} \\ 0 & \text{if } t = 10 \end{cases}$	$C(u^*) = 2^{11} = 2048$
$\mathbb{T}_1 = 3\mathbb{N}$	$u^*(t) = \begin{cases} 1 & \text{if } t \in \{0, 3, 6\} \\ 0 & \text{if } t = 9 \end{cases}$	$C(u^*) = 3 \cdot 2^9 = 1536$
$\mathbb{T}_1 = 4\mathbb{N}$	$u^*(t) = \begin{cases} 1 & \text{if } t \in \{0, 4\} \\ 0.2886 & \text{if } t = 8 \end{cases}$	$C(u^*) \simeq 1108.882$
$\mathbb{T}_1 = 6\mathbb{N}$	$u^*(t) = \begin{cases} 1 & \text{if } t = 0 \\ 0.5725 & \text{if } t = 6 \end{cases}$	$C(u^*) \simeq 674.787$
$\mathbb{T}_1 = 12\mathbb{N}$	$u^*(0) \simeq 0.8145$	$C(u^*) \simeq 159.647$

15

Remark 20. In the permanent discrete case $\mathbb{T} = \mathbb{T}_1 = \mathbb{N}$, the optimal control is not unique. Indeed, every control u^* satisfying $u^*(t) = 1$ for every $t \in \{0, 1, \ldots, 9\}$, $u^*(10) \in [0, 1]$ and $u^*(11) = 0$ is optimal. Let us give some details on the proof and let us note that the PMP does not provide any constraint on the value $u^*(10)$ in that case.

With $r = 11$ and $p(12) = 0$, Λ_r has negative values on $(0, 1]$ (constantly equal to -1), then $u^*(11) = 0$ from (13). From (11), we compute $p(11) = 1$. With $r = 10$ and $p(11) = 1$, Λ_r is constantly equal to 0 and then (13) does not provide any constraint on the value $u^*(10)$. Actually, it does not matter since $p(10)$ can still be computed from (11) ($p(10) = 2$) and the recursive procedure can be pursued. We obtain $u^*(t) = 1$ for every $t \in \{0, 1, \ldots, 9\}$. From (10), we have

$$C(u^*) = (1 - u^*(10))q^*(10) + q^*(11) = (1 - u^*(10)) \cdot 2^{10} + (1 + u^*(10))q^*(10) = 2 \cdot 2^{10} = 2^{11}.$$

Finally, if $u^*(11) = 0$ and $u^*(t) = 1$ for every $t \in \{0, 1, \ldots, 9\}$, the value of $u^*(10) \in [0, 1]$ does not influence $C(u^*)$. This concludes the proof and the remark.

Remark 21. The case $\mathbb{T}_1 = 2\mathbb{N}$ can be seen as a consequence of the permanent case $\mathbb{T}_1 = \mathbb{N}$ (see Remarks 3 and 20).

Remark 22. The case $\mathbb{T}_1 = 12\mathbb{N}$ corresponds to an optimal parameter problem (see Remark 4).

3.2 Non-extension of several classical properties

In this section, we recall some basic properties that occur in classical optimal control theory in the continuous-time setting and with a permanent control, that is, with $\mathbb{T} = \mathbb{T}_1 = \mathbb{R}^+$. Our aim is to discuss their extension (or their failure) to the general time scale setting and to the nonpermament control case. We will provide several counterexamples in the discrete-time setting with a permanent control ($\mathbb{T} = \mathbb{T}_1 = \mathbb{N}$) and in the continuous-time setting with a nonpermanent control ($\mathbb{T}_1 \subsetneq \mathbb{T} = \mathbb{R}^+$). In the following paragraphs, except the last one, the final time $b \in \mathbb{T}$ can be fixed or not.

Pointwise maximization condition of the Hamiltonian. In the case $\mathbb{T} = \mathbb{T}_1 = \mathbb{R}^+$, it is well known that an optimal (permanent) control u^* satisfies the maximization condition $u^*(t) \in \arg\max_{z \in \Omega} H(t, q^*(t), p(t), p^0, z)$ for a.e. $t \in [0, b^*)$. We refer to [16, Example 7] for a counterexample showing the failure of this maximization condition in the case $\mathbb{T} = \mathbb{T}_1 = \mathbb{N}$. We refer to Remark 18 for a counterexample in the case $\mathbb{T}_1 \subsetneq \mathbb{T} = \mathbb{R}^+$.

Continuity of the Hamiltonian. In the case $\mathbb{T} = \mathbb{T}_1 = \mathbb{R}^+$, it is well known that the Hamiltonian function $t \mapsto H(t, q^*(t), p(t), p^0, u^*(t))$ coincides almost everywhere on $[0, b^*]$ with the continuous function $t \mapsto \max_{z \in \Omega} H(t, q^*(t), p(t), p^0, z)$. Remark 18 provides a counterexample showing the failure of this regularity property in the case $\mathbb{T}_1 \subsetneq \mathbb{T} = \mathbb{R}^+$.

Remark 23. Nevertheless, in the case of a free final time, under the assumptions of the fourth item of Theorem 3, the Hamiltonian function $t \mapsto H(t, q^*(t), p(t), p^0, u^{*\Phi}(t))$ coincides almost everywhere, in some neighborhood of b^*, with a continuous function.

The autonomous case. In the case $\mathbb{T} = \mathbb{T}_1 = \mathbb{R}^+$, if the Hamiltonian H is autonomous (that is, does not depend on t), it is well known that the function $t \mapsto H(q^*(t), p(t), p^0, u^*(t))$ is almost everywhere constant on $[0, b^*]$, this constant being equal to the maximized Hamiltonian. We refer to [16, Example 8] for a counterexample showing the failure of this constantness property in the case $\mathbb{T} = \mathbb{T}_1 = \mathbb{N}$, and we refer to Remark 18 for a counterexample in the case $\mathbb{T}_1 \subsetneq \mathbb{T} = \mathbb{R}^+$ (and there, even the maximized autonomous Hamiltonian is not constant).

16

Saturated constraint set Ω for Hamiltonian affine in u. In this paragraph, we assume that Ω is convex. In the case $\mathbb{T} = \mathbb{T}_1 = \mathbb{R}^+$, if the Hamiltonian is *affine in u*, that is, if it can be written as

$$H(t, q, p, p^0, u) = \langle H_1(t, q, p, p^0), u \rangle_{\mathbb{R}^m} + H_2(t, q, p, p^0),$$

one can easily prove that $H_1(t, q^*(t), p(t), p^0) \in \mathcal{O}_\Omega[u^*(t)]$ for almost every $t \in [0, b^*)$. It follows that an optimal (permanent) control u^* must take its values at the boundary of Ω (saturation of the constraints) for almost every $t \in [0, b^*)$ such that $H_1(t, q^*(t), p(t), p^0) \neq 0_{\mathbb{R}^m}$.

Remark 24. This classical property can be extended to the case $\mathbb{T} = \mathbb{T}_1 = \mathbb{N}$. Indeed, in that case, the nonpositive gradient condition is given by

$$\left\langle \frac{\partial H}{\partial u}(t, q^*(t), p(t+1), p^0, u^*(t)), y - u^*(t) \right\rangle_{\mathbb{R}^m} = \langle H_1(t, q^*(t), p(t+1), p^0), y - u^*(t) \rangle_{\mathbb{R}^m} \leq 0,$$

for every $y \in \Omega$, that is, $H_1(t, q^*(t), p(t+1), p^0) \in \mathcal{O}_\Omega[u^*(t)]$.

Remark 25. Remark 20 provides an interesting example in the case $\mathbb{T} = \mathbb{T}_1 = \mathbb{N}$. Indeed, in that case, the control defined by $u^*(t) = 0$ for every $t \in \{0, \ldots, 9\}$, $u^*(10) = 1/2$ and $u^*(11) = 1$ is an optimal (permanent) control. However, it does not saturate the constraint set Ω at $t = 10$. It is not a surprise since, in that case, $H_1(t, q^*(t), p(t+1), p^0) = 0$ at $t = 10$.

Note that, in the case $\mathbb{T}_1 = 4\mathbb{N}$, Section 3.1.2 provides a counterexample showing the failure of this classical property in the case $\mathbb{T}_1 \subsetneq \mathbb{T} = \mathbb{N}$. Similarly, Section 3.1.1 in the case $\mathbb{T}_1 = 3\mathbb{N}$ provides a counterexample in the case $\mathbb{T}_1 \subsetneq \mathbb{T} = \mathbb{R}^+$.

Remark 26. Figure 1 represents the values of the optimal sampled-data control u^* of Section 3.1.1, in the case $\mathbb{T}_1 = 12\mathbb{N} \cup \{\lambda\}$, where λ is a parameter evolving in $(0, 12)$. In that case, $u^*(0)$ (resp., $u^*(\lambda)$) saturates the constraint set $\Omega = [0, 1]$ approximately for $\lambda \in (0, 11.9245)$ (resp., for $\lambda \in (9.9866, 12)$).
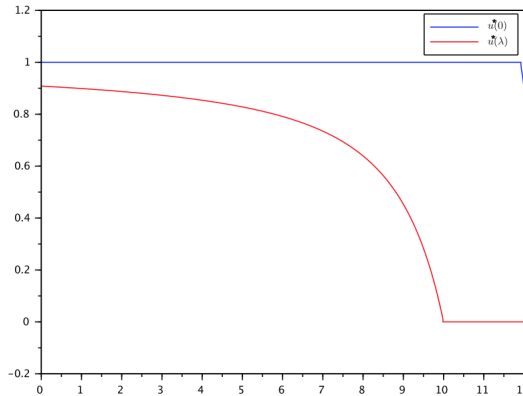

Figure 1: Optimal sampled-data control u^*.

Vanishing of the maximized Hamiltonian at the final time. In this paragraph, we assume that the final time is left free. In the case $\mathbb{T} = \mathbb{T}_1 = \mathbb{R}^+$, under the assumptions of the fourth item of Theorem 3, it is well known that the maximized Hamiltonian vanishes at $t = b^*$ (see Remark 8).

In Theorem 3, we have established that this property is still valid in the time scale setting under some appropriate conditions, the main one being that b^* must belong to the interior of $\mathbb{T}$. In the discrete case $\mathbb{T} = \mathbb{N}$, the interior of $\mathbb{T}$ is empty and then the latter assumption is never satisfied. The Hamiltonian at the final time may then not vanish, and we refer to [16, Example 8] for a counterexample with $\mathbb{T} = \mathbb{T}_1 = \mathbb{N}$.

4 Proofs

The section is structured as follows. Subsections 4.1, 4.2 and 4.3 are devoted to the proof of Theorem 3. In Subsection 4.1, we recall some known Cauchy-Lipschitz results on time scales and we establish some preliminary results on the relations between u and u^Φ. In Subsection 4.2, we introduce appropriate needle-like variations of the control. Finally, in Subsection 4.3, we apply the Ekeland variational principle to an adequate functional in an appropriate complete metric space, and then we prove the PMP. In Subsection 4.4 (that the reader can read independently of the rest of Section 4), we detail the proof of Theorem 2.

4.1 Preliminaries

4.1.1 Relations between u and u^Φ

We start with a lemma whose arguments of proof will be used several times.

Lemma 4. *Let $m \in \mathbb{N}^*$ and let $c < d$ be two elements of $\mathbb{T}$ with $c \in \mathbb{T}_1$. Let u, $v : [c,d)_{\mathbb{T}_1} \to \mathbb{R}^m$ be two functions. Then, $u = v$ Δ_1-a.e. on $[c,d)_{\mathbb{T}_1}$ if and only if $u^\Phi = v^\Phi$ Δ-a.e. on $[c,d)_{\mathbb{T}}$.*

Proof. Without loss of generality, we can assume that v is constant equal to $0_{\mathbb{R}^m}$. Let us define $A = \{t \in [c,d)_{\mathbb{T}_1} \,|\, u(t) \neq 0_{\mathbb{R}^m}\}$ and $B = \{t \in [c,d)_{\mathbb{T}} \,|\, u^\Phi(t) \neq 0_{\mathbb{R}^m}\}$. Since $\Phi(t) = t$ for every $t \in \mathbb{T}_1$, the inclusion $A \subset B$ holds. Firstly, let us assume that $\mu_{\Delta_1}(A) = 0$. Hence, $A \subset \mathrm{RD}_1$ and $\mu_{\Delta_1}(A) = \mu_L(A) = 0$. Since $A \subset \mathrm{RD}_1 \subset \mathrm{RD}$, we deduce that $\mu_\Delta(A) = \mu_L(A) = 0$. On the other hand, for every $t \in B$, $\Phi(t) \in A \subset \mathrm{RD}_1$, then $t \in \mathbb{T}_1$ and $t = \Phi(t) \in A$. We conclude that $A = B$ and $\mu_\Delta(B) = 0$. Secondly, let us assume that $\mu_\Delta(B) = 0$ and $\mu_{\Delta_1}(A) > 0$. Since $A \subset B$, we deduce that $\mu_\Delta(A) = 0$, $A \subset \mathrm{RD}$ and $\mu_\Delta(A) = \mu_L(A) = 0$. Since $\mu_L(A) = 0$ and $\mu_{\Delta_1}(A) > 0$, we conclude that there exists $t_0 \in \mathrm{RS}_1 \cap A \subset \mathrm{RS}_1 \cap \mathrm{RD}$. Consequently, u^Φ is constant (different of $0_{\mathbb{R}^m}$) on $[t_0, \min(\sigma_1(t_0), d))_{\mathbb{T}}$. As a consequence, $\mu_\Delta(B) \geq \min(\sigma_1(t_0), d) - t_0 > 0$. This raises a contradiction. $\qquad \square$

Proposition 1. *Let $m \in \mathbb{N}^*$ and let $c < d$ be two elements of $\mathbb{T}$ with $c \in \mathbb{T}_1$.*

1. *For every $u \in \mathrm{L}^1_{\mathbb{T}_1}([c,d)_{\mathbb{T}_1}, \mathbb{R}^m)$, we have $u^\Phi \in \mathrm{L}^1_{\mathbb{T}}([c,d)_{\mathbb{T}}, \mathbb{R}^m)$ and*

$$\|u^\Phi\|_{\mathrm{L}^1_{\mathbb{T}}([c,d)_{\mathbb{T}}, \mathbb{R}^m)} \leq \|u\|_{\mathrm{L}^1_{\mathbb{T}_1}([c,d)_{\mathbb{T}_1}, \mathbb{R}^m)}. \tag{14}$$

2. *For every $u \in \mathrm{L}^\infty_{\mathbb{T}_1}([c,d)_{\mathbb{T}_1}, \mathbb{R}^m)$, we have $u^\Phi \in \mathrm{L}^\infty_{\mathbb{T}}([c,d)_{\mathbb{T}}, \mathbb{R}^m)$ and*

$$\|u^\Phi\|_{\mathrm{L}^\infty_{\mathbb{T}}([c,d)_{\mathbb{T}}, \mathbb{R}^m)} = \|u\|_{\mathrm{L}^\infty_{\mathbb{T}_1}([c,d)_{\mathbb{T}_1}, \mathbb{R}^m)}.$$

Proof. Let $u \in \mathrm{L}^1_{\mathbb{T}_1}([c,d)_{\mathbb{T}_1}, \mathbb{R}^m)$. We first treat the μ_Δ-measurability of u^Φ. From Lemma 4, we can consider that u is defined everywhere on $[c,d)_{\mathbb{T}_1}$ and is μ_{Δ_1}-measurable on $[c,d)_{\mathbb{T}_1}$. Let us prove that u^Φ is μ_Δ-measurable on $[c,d)_{\mathbb{T}}$. We introduce $d' = \inf\{s \in \mathbb{T}_1 \,|\, s \geq d\} \in \mathbb{T}_1$. Note that

18

$d' \geq d$, with equality if and only if $d \in \mathbb{T}_1$. From [23], u is μ_{Δ_1}-measurable on $[c,d)_{\mathbb{T}_1}$ if and only if the extension $\tilde{u}$ defined on $[c,d')$ by

$$\tilde{u}(t) = \begin{cases} u(t) & \text{if } t \in \mathbb{T}_1, \\ u(r) & \text{if } t \in (r, \sigma_1(r)), \text{ where } r \in \mathrm{RS}_1, \end{cases}$$

is μ_L-measurable on $[c,d')$. By hypothesis, $\tilde{u}$ is μ_L-measurable on $[c,d')$ and consequently, the restriction $\tilde{u}_{|[c,d)}$ is μ_L-measurable on $[c,d)$. Still from [23], u^{Φ} is μ_Δ-measurable on $[c,d)_{\mathbb{T}}$ if and only if the extension $\widetilde{u^{\Phi}}$ defined on $[c,d)$ by

$$\widetilde{u^{\Phi}}(t) = \begin{cases} u^{\Phi}(t) & \text{if } t \in \mathbb{T}, \\ u^{\Phi}(r) & \text{if } t \in (r, \sigma(r)), \text{ where } r \in \mathrm{RS}, \end{cases}$$

is μ_L-measurable on $[c,d)$. Hence, it is sufficient to see that $\widetilde{u^{\Phi}} = \tilde{u}_{|[c,d)}$. This is true since, for any $t \in [c,d)$, we have:

- either $t \in \mathbb{T}_1$, and then $t \in \mathbb{T}$ and thus $\widetilde{u^{\Phi}}(t) = u^{\Phi}(t) = u(\Phi(t)) = u(t) = \tilde{u}(t)$.

- either $t \in \mathbb{T} \backslash \mathbb{T}_1$, and then $t \in (\Phi(t), \sigma_1(\Phi(t)))$ where $\Phi(t) \in \mathrm{RS}_1$ and thus $\widetilde{u^{\Phi}}(t) = u^{\Phi}(t) = u(\Phi(t)) = \tilde{u}(t)$.

- either $t \notin \mathbb{T}$, and then $t \in (r, \sigma(r))$ where $r \in \mathrm{RS}$ and then $t \in (\Phi(r), \sigma_1(\Phi(r)))$ where $\Phi(r) \in \mathrm{RS}_1$ and thus $\widetilde{u^{\Phi}}(t) = u^{\Phi}(r) = u(\Phi(r)) = \tilde{u}(t)$.

This establishes the μ_Δ-measurability of u^{Φ} on $[c,d)_{\mathbb{T}}$.

Considering $\|u\|_{\mathbb{R}^m}$ instead of u, we can consider that $m = 1$ and $u \in \mathrm{L}^1_{\mathbb{T}_1}([c,d)_{\mathbb{T}_1}, \mathbb{R}^+)$. From [23], we have

$$\int_{[c,d)_{\mathbb{T}}} u^{\Phi}(\tau) \, \Delta\tau = \int_{[c,d)} \widetilde{u^{\Phi}}(\tau) \, d\tau = \int_{[c,d)} \tilde{u}_{|[c,d)}(\tau) \, d\tau$$
$$= \int_{[c,d')} \tilde{u}(\tau) \, d\tau - \int_{[d,d')} \tilde{u}(\tau) \, d\tau = \int_{[c,d)_{\mathbb{T}_1}} u(\tau) \, \Delta_1\tau - (d' - d)u(\Phi(d)).$$

Noting that $(d' - d)u(\Phi(d)) \geq 0$ concludes the proof of the first point.

Let us prove the second point. Let $u \in \mathrm{L}^\infty_{\mathbb{T}_1}([c,d)_{\mathbb{T}_1}, \mathbb{R}^m)$. The μ_Δ-measurability of u^{Φ} is already proved since $\mathrm{L}^\infty_{\mathbb{T}_1}([c,d)_{\mathbb{T}_1}, \mathbb{R}^m) \subset \mathrm{L}^1_{\mathbb{T}_1}([c,d)_{\mathbb{T}_1}, \mathbb{R}^m)$. Let $M \geq 0$ be a constant. With the same arguments as in the proof of Lemma 4, we can prove that $\|u(t)\|_{\mathbb{R}^m} \leq M$ for Δ_1-a.e. $t \in [c,d)_{\mathbb{T}_1}$ if and only if $\|u^{\Phi}(t)\|_{\mathbb{R}^m} \leq M$ for Δ-a.e. $t \in [c,d)_{\mathbb{T}}$. As a consequence, we get $u^{\Phi} \in \mathrm{L}^\infty_{\mathbb{T}}([c,d)_{\mathbb{T}}, \mathbb{R}^m)$ and $\|u^{\Phi}\|_{\mathrm{L}^\infty_{\mathbb{T}}([c,d)_{\mathbb{T}}, \mathbb{R}^m)} = \|u\|_{\mathrm{L}^\infty_{\mathbb{T}_1}([c,d)_{\mathbb{T}_1}, \mathbb{R}^m)}$. $\qquad\square$

Remark 27. From the proof above, we see that the inequality (14) is an equality if and only if $d \in \mathbb{T}_1$ or $u(\Phi(d)) = 0$. Then, considering $\mathbb{T} = \mathbb{N}$, $\mathbb{T}_1 = 2\mathbb{N}$, $m = 1$, $c = 0$, $d = 1$ and u the constant function equal to 1 provides a counterexample. Indeed, in that case, we have $u \in \mathrm{L}^1_{\mathbb{T}_1}([c,d)_{\mathbb{T}_1}, \mathbb{R}^m)$ and $u^{\Phi} \in \mathrm{L}^1_{\mathbb{T}}([c,d)_{\mathbb{T}}, \mathbb{R}^m)$ with $\|u^{\Phi}\|_{\mathrm{L}^1_{\mathbb{T}}([c,d)_{\mathbb{T}}, \mathbb{R}^m)} = 1 < 2 = \|u\|_{\mathrm{L}^1_{\mathbb{T}_1}([c,d)_{\mathbb{T}_1}, \mathbb{R}^m)}$.

4.1.2 Recalls on Δ-Cauchy-Lipschitz results

According to [15, Theorem 1], for every control $u \in \mathrm{L}^\infty_{\mathbb{T}_1}(\mathbb{T}_1, \mathbb{R}^m)$ and every initial condition $q_a \in \mathbb{R}^n$, there exists a unique maximal solution of (4) such that $q(a) = q_a$, denoted by $q(\cdot, u, q_a)$, and

defined on a maximal interval, denoted by $I_\mathbb{T}(u, q_a)$. The word *maximal* means that $q(\cdot, u, q_a)$ is an extension of any other solution. Moreover, we recall that (see [15, Lemma 1])

$$\forall t \in I_\mathbb{T}(u, q_a), \quad q(t, u, q_a) = q_a + \int_{[a,t)_\mathbb{T}} f(\tau, q(\tau, u, q_a), u^\Phi(\tau)) \, \Delta\tau.$$

Finally, either $I_\mathbb{T}(u, q_a) = \mathbb{T}$, that is, $q(\cdot, u, q_a)$ is a *global* solution of (4), or $I_\mathbb{T}(u, q_a) = [a, c)_\mathbb{T}$ where c is a left-dense point of $\mathbb{T}$, and in that case, $q(\cdot, u, q_a)$ is unbounded on $I_\mathbb{T}(u, q_a)$ (see [15, Theorem 2]).

Definition 2. For a given $b \in \mathbb{T}$, a couple $(u, q_a) \in \mathrm{L}^\infty_{\mathbb{T}_1}(\mathbb{T}_1, \mathbb{R}^m) \times \mathbb{R}^n$ is said to be *admissible* on $[a, b]_\mathbb{T}$ whenever $b \in I_\mathbb{T}(u, q_a)$.

For a given $b \in \mathbb{T}$, we denote by $\mathcal{UQ}^b$ the set of all admissible couples (u, q_a) on $[a, b]_\mathbb{T}$. It is endowed with the norm $\|(u, q_a)\|_{\mathcal{UQ}^b} = \|u\|_{\mathrm{L}^1_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)} + \|q_a\|_{\mathbb{R}^n}$.

4.2 Needle-like variations of the control, and variation of the initial condition

Throughout this section, we consider $b \in \mathbb{T}$ and $(u, q_a) \in \mathcal{UQ}^b$. We are going, in particular, to define appropriate needle-like variations. As in [16], we have to distinguish between right-dense and right-scattered times, along the time scale $\mathbb{T}_1$. We will also define appropriate variation of the initial condition.

In the sequel, the notation $\|\cdot\|$ stands for the usual induced norm of matrices in $\mathbb{R}^{n,n}$, $\mathbb{R}^{n,m}$ and $\mathbb{R}^{m,m}$.

4.2.1 General variation of (u, q_a)

In the first lemma below, we prove that $\mathcal{UQ}^b$ is open. Actually we prove a stronger result, by showing that $\mathcal{UQ}^b$ contains a neighborhood of any of its point in $\mathrm{L}^1_{\mathbb{T}_1} \times \mathbb{R}^n$ topology, which will be useful in order to define needle-like variations.

Lemma 5. *Let $R > \|u\|_{\mathrm{L}^\infty_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)}$. There exist $\nu_R > 0$ and $\eta_R > 0$ such that the set*

$$\mathrm{E}_R(u, q_a) = \left(\overline{\mathrm{B}}_{\mathrm{L}^\infty_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)}(0, R) \cap \overline{\mathrm{B}}_{\mathrm{L}^1_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)}(u, \nu_R) \right) \times \overline{\mathrm{B}}_{\mathbb{R}^n}(q_a, \eta_R)$$

is contained in $\mathcal{UQ}^b$.

Before proving this lemma, let us recall a time scale version of Gronwall's Lemma (see [10, Chapter 6.1]). The generalized exponential function is defined by $e_L(t, c) = \exp(\int_{[c,t)_\mathbb{T}} \xi_{\mu(\tau)}(L) \, \Delta\tau)$, for every $L \geq 0$, every $c \in \mathbb{T}$ and every $t \in [c, +\infty)_\mathbb{T}$, where $\xi_{\mu(\tau)}(L) = \log(1 + L\mu(\tau))/\mu(\tau)$ whenever $\mu(\tau) > 0$, and $\xi_{\mu(\tau)}(L) = L$ whenever $\mu(\tau) = 0$ (see [10, Chapter 2.2]). Note that, for every $L \geq 0$ and every $c \in \mathbb{T}$, the function $e_L(\cdot, c)$ is positive and increasing on $[c, +\infty)_\mathbb{T}$.

Lemma 6 ([10])**.** *Let $c < d$ be two elements of $\mathbb{T}$, let L_1 and L_2 be two nonnegative real numbers, and let $q \in \mathrm{C}([c, d]_\mathbb{T}, \mathbb{R})$ satisfying $0 \leq q(t) \leq L_1 + L_2 \int_{[c,t)_\mathbb{T}} q(\tau) \, \Delta\tau$, for every $t \in [c, d]_\mathbb{T}$. Then $0 \leq q(t) \leq L_1 e_{L_2}(t, c)$, for every $t \in [c, d]_\mathbb{T}$.*

Proof of Lemma 5. By continuity of $q(\cdot, u, q_a)$ on $[a, b]_\mathbb{T}$, the set

$$K_R = \{(t, x, v) \in [a, b]_\mathbb{T} \times \mathbb{R}^n \times \overline{\mathrm{B}}_{\mathbb{R}^m}(0, R) \mid \|x - q(t, u, q_a)\|_{\mathbb{R}^n} \leq 1\}$$

is a compact subset of $\mathbb{T} \times \mathbb{R}^n \times \mathbb{R}^m$. Therefore $\|\partial f/\partial q\|$ and $\|\partial f/\partial u\|$ are bounded on K_R by some $L_R \geq 0$ and then, from convexity and from the mean value inequality, it holds:

$$\|f(t, x_2, v_2) - f(t, x_1, v_1)\|_{\mathbb{R}^n} \leq L_R(\|x_2 - x_1\|_{\mathbb{R}^n} + \|v_2 - v_1\|_{\mathbb{R}^m}), \qquad (15)$$

for all $(t, x_1, v_1), (t, x_2, v_2) \in K_R$. Let $\nu_R > 0$ and $0 < \eta_R < 1$ such that $(\eta_R + \nu_R L_R)e_{L_R}(b, a) < 1$.

Let $(u', q_a') \in \mathrm{E}_R(u, q_a)$. Our aim is to prove that $b \in I_{\mathbb{T}}(u', q_a')$. By contradiction, assume that the set

$$A = \{t \in I_{\mathbb{T}}(u', q_a') \cap [a, b]_{\mathbb{T}} \mid \|q(t, u', q_a') - q(t, u, q_a)\|_{\mathbb{R}^n} > 1\}$$

is not empty and let $t_0 = \inf A$. Since $\mathbb{T}$ is closed, $t_0 \in I_{\mathbb{T}}(u', q_a') \cap [a, b]_{\mathbb{T}}$ and $[a, t_0]_{\mathbb{T}} \subset I_{\mathbb{T}}(u', q_a') \cap [a, b]_{\mathbb{T}}$. If t_0 is a minimum then $\|q(t_0, u', q_a') - q(t_0, u, q_a)\|_{\mathbb{R}^n} > 1$. If t_0 is not a minimum then $t_0 \in \mathrm{RD}$ and by continuity we have $\|q(t_0, u', q_a') - q(t_0, u, q_a)\|_{\mathbb{R}^n} \geq 1$. Moreover one has $t_0 > a$ since $\|q(a, u', q_a') - q(a, u, q_a)\|_{\mathbb{R}^n} = \|q_a' - q_a\|_{\mathbb{R}^n} \leq \eta_R < 1$. Hence $\|q(\tau, u', q_a') - q(\tau, u, q_a)\|_{\mathbb{R}^n} \leq 1$ for every $\tau \in [a, t_0)_{\mathbb{T}}$. Therefore $(\tau, q(\tau, u', q_a'), u'^\Phi(\tau))$ and $(\tau, q(\tau, u, q_a), u^\Phi(\tau))$ are elements of K_R for Δ-a.e. $\tau \in [a, t_0)_{\mathbb{T}}$. Since one has

$$q(t, u', q_a') - q(t, u, q_a) = q_a' - q_a + \int_{[a,t)_{\mathbb{T}}} \left(f(\tau, q(\tau, u', q_a'), u'^\Phi(\tau)) - f(\tau, q(\tau, u, q_a), u^\Phi(\tau)) \right) \Delta\tau,$$

for every $t \in I_{\mathbb{T}}(u', q_a') \cap [a, b]_{\mathbb{T}}$, it follows from (15) that, for every $t \in [a, t_0]_{\mathbb{T}}$,

$$\|q(t, u', q_a') - q(t, u, q_a)\|_{\mathbb{R}^n} \leq \|q_a' - q_a\|_{\mathbb{R}^n} + L_R \int_{[a,t)_{\mathbb{T}}} \|u'^\Phi(\tau) - u^\Phi(\tau)\|_{\mathbb{R}^m} \Delta\tau$$
$$+ L_R \int_{[a,t)_{\mathbb{T}}} \|q(\tau, u', q_a') - q(\tau, u, q_a)\|_{\mathbb{R}^n} \Delta\tau,$$

which implies from Lemma 6 that, for every $t \in [a, t_0]_{\mathbb{T}}$,

$$\|q(t, u', q_a') - q(t, u, q_a)\|_{\mathbb{R}^n} \leq (\|q_a' - q_a\|_{\mathbb{R}^n} + L_R\|u'^\Phi - u^\Phi\|_{\mathrm{L}^1_{\mathbb{T}}([a,b)_{\mathbb{T}}, \mathbb{R}^m)})e_{L_R}(b, a),$$

which finally implies from Proposition 1 that, for every $t \in [a, t_0]_{\mathbb{T}}$,

$$\|q(t, u', q_a') - q(t, u, q_a)\|_{\mathbb{R}^n} \leq (\|q_a' - q_a\|_{\mathbb{R}^n} + L_R\|u' - u\|_{\mathrm{L}^1_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)})e_{L_R}(b, a)$$
$$\leq (\eta_R + \nu_R L_R)e_{L_R}(b, a) < 1.$$

This raises a contradiction at $t = t_0$. Therefore A is empty and thus $q(\cdot, u', q_a')$ is bounded on $I_{\mathbb{T}}(u', q_a') \cap [a, b]_{\mathbb{T}}$. It follows from [15, Theorem 2] that $b \in I_{\mathbb{T}}(u', q_a')$, that is, $(u', q_a') \in \mathcal{UQ}^b$. $\qquad\square$

Remark 28. Let $R > \|u\|_{\mathrm{L}^\infty_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)}$ and $(u', q_a') \in \mathrm{E}_R(u, q_a)$. With the notations of the above proof, since $[a, b]_{\mathbb{T}} \subset I_{\mathbb{T}}(u', q_a')$ and A is empty, we infer that $\|q(t, u', q_a') - q(t, u, q_a)\|_{\mathbb{R}^n} \leq 1$, for every $t \in [a, b]_{\mathbb{T}}$. Therefore $(\tau, q(\tau, u', q_a'), u'^\Phi(\tau)) \in K_R$ for every $(u', q_a') \in \mathrm{E}_R(u, q_a)$ and for Δ-a.e. $\tau \in [a, b)_{\mathbb{T}}$.

Lemma 7. *Let* $R > \|u\|_{\mathrm{L}^\infty_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)}$. *The mapping*

$$F_R(u, q_a): \quad (\mathrm{E}_R(u, q_a), \|\cdot\|_{\mathcal{UQ}^b}) \quad \longrightarrow \quad (\mathrm{C}([a, b]_{\mathbb{T}}, \mathbb{R}^n), \|\cdot\|_\infty)$$
$$(u', q_a') \quad \longmapsto \quad q(\cdot, u', q_a')$$

is Lipschitzian. In particular, for every $(u', q_a') \in \mathrm{E}_R(u, q_a)$, $q(\cdot, u', q_a')$ *converges uniformly to* $q(\cdot, u, q_a)$ *on* $[a, b]_{\mathbb{T}}$ *when* u' *tends to* u *in* $\mathrm{L}^1_{\mathbb{T}_1}([a, b)_{\mathbb{T}_1}, \mathbb{R}^m)$ *and* q_a' *tends to* q_a *in* $\mathbb{R}^n$.

Proof. Let (u', q_a') and (u'', q_a'') be two elements of $\mathrm{E}_R(u, q_a) \subset \mathcal{UQ}^b$. It follows from Remark 28 that $(\tau, q(\tau, u'', q_a''), u''^\Phi(\tau))$ and $(\tau, q(\tau, u', q_a'), u'^\Phi(\tau))$ are elements of K_R for Δ-a.e. $\tau \in [a, b]_{\mathbb{T}}$. Following the same arguments as in the previous proof, it follows from (15), from Lemma 6 and from Proposition 1 that, for every $t \in [a, b]_{\mathbb{T}}$,

$$\|q(t, u'', q_a'') - q(t, u', q_a')\|_{\mathbb{R}^n} \leq (\|q_a'' - q_a'\|_{\mathbb{R}^n} + L_R\|u'' - u'\|_{\mathrm{L}^1_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)})e_{L_R}(b, a).$$

The lemma follows. $\qquad\square$

4.2.2 Needle-like variation of u at a point $r \in \mathrm{RS}_1$

Let $r \in [a, b)_{\mathbb{T}_1} \cap \mathrm{RS}_1$ and let $y \in \mathbb{R}^m$. We define the needle-like variation $\Pi = (r, y)$ of u at r by

$$u_\Pi(t, \alpha) = \begin{cases} u(r) + \alpha(y - u(r)) & \text{if} \quad t = r, \\ u(t) & \text{if} \quad t \neq r, \end{cases}$$

for Δ_1-a.e. $t \in [a, b)_{\mathbb{T}_1}$ and for every $\alpha \in [0, 1]$. In the sequel, let us denote by $\sigma_1^*(r) = \min(\sigma_1(r), b)$.

Lemma 8. *There exists $0 < \alpha_0 \leq 1$ such that $(u_\Pi(\cdot, \alpha), q_a) \in \mathcal{UQ}^b$ for every $\alpha \in [0, \alpha_0]$.*

Proof. Let $R = \max(\|u\|_{\mathrm{L}^\infty_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)}, \|u(r)\|_{\mathbb{R}^m} + \|y\|_{\mathbb{R}^m}) + 1 > \|u\|_{\mathrm{L}^\infty_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)}$. We use the notations K_R, L_R, ν_R and η_R, defined in Lemma 5 and in its proof. One has $\|u_\Pi(\cdot, \alpha)\|_{\mathrm{L}^\infty_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)} \leq R$ for every $\alpha \in [0, 1]$, and

$$\|u_\Pi(\cdot, \alpha) - u\|_{\mathrm{L}^1_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)} = \mu_1(r)\|u_\Pi(r, \alpha) - u(r)\|_{\mathbb{R}^m} = \alpha\mu_1(r)\|y - u(r)\|_{\mathbb{R}^m}.$$

Hence, there exists $0 < \alpha_0 \leq 1$ such that $\|u_\Pi(\cdot, \alpha) - u\|_{\mathrm{L}^1_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)} \leq \nu_R$ for every $\alpha \in [0, \alpha_0]$, and hence $(u_\Pi(\cdot, \alpha), q_a) \in \mathrm{E}_R(u, q_a)$. The claim follows then from Lemma 5. $\qquad\square$

Lemma 9. *The mapping*

$$F_\Pi(u, q_a): \quad \begin{array}{ccc} ([0, \alpha_0], |\cdot|) & \longrightarrow & (\mathrm{C}([a, b]_{\mathbb{T}}, \mathbb{R}^n), \|\cdot\|_\infty) \\ \alpha & \longmapsto & q(\cdot, u_\Pi(\cdot, \alpha), q_a) \end{array}$$

is Lipschitzian. In particular, for every $\alpha \in [0, \alpha_0]$, $q(\cdot, u_\Pi(\cdot, \alpha), q_a)$ converges uniformly to $q(\cdot, u, q_a)$ on $[a, b]_{\mathbb{T}}$ as α tends to 0.

Proof. We use the notations of the proof of Lemma 8. It follows from Lemma 7 that there exists $C \geq 0$ (the Lipschitz constant of $F_R(u, q_a)$) such that

$$\|q(\cdot, u_\Pi(\cdot, \alpha_2), q_a) - q(\cdot, u_\Pi(\cdot, \alpha_1), q_a)\|_\infty \leq C\|(u_\Pi(\cdot, \alpha_2), q_a) - (u_\Pi(\cdot, \alpha_1), q_a)\|_{\mathcal{UQ}^b}$$
$$= C|\alpha_2 - \alpha_1|\mu_1(r)\|y - u(r)\|_{\mathbb{R}^m},$$

for all α_1 and α_2 in $[0, \alpha_0]$. The lemma follows. $\qquad\square$

We define the so-called *first variation vector* $h_\Pi(\cdot, u, q_a)$ associated with the needle-like variation $\Pi = (r, y)$ as the unique solution on $[r, \sigma_1^*(r)]_{\mathbb{T}}$ of the linear Δ-Cauchy problem

$$h^\Delta(t) = \frac{\partial f}{\partial q}(t, q(t, u, q_a), u(r))\, h(t) + \frac{\partial f}{\partial u}(t, q(t, u, q_a), u(r))\, (y - u(r)), \quad h(r) = 0. \quad (16)$$

The existence and uniqueness of $h_\Pi(\cdot, u, q_a)$ are ensured by [15, Theorem 3].

Proposition 2. *The mapping*

$$F_\Pi(u, q_a): \quad ([0, \alpha_0], |\cdot|) \quad \longrightarrow \quad (\mathrm{C}([r, \sigma_1^*(r)]_\mathbb{T}, \mathbb{R}^n), \|\cdot\|_\infty)$$
$$\alpha \quad \longmapsto \quad q(\cdot, u_\Pi(\cdot, \alpha), q_a)$$

is differentiable at 0, and one has $DF_\Pi(u, q_a)(0) = h_\Pi(\cdot, u, q_a)$.

Proof. We use the notations of the proof of Lemma 8. Recall that $(\tau, q(\tau, u_\Pi(\cdot, \alpha), q_a), u_\Pi^\Phi(\tau, \alpha)) \in K_R$ for Δ-a.e. $\tau \in [a, b)_\mathbb{T}$ and for every $\alpha \in [0, \alpha_0]$ (see Remark 28). For every $\alpha \in (0, \alpha_0]$ and every $t \in [r, \sigma_1^*(r)]_\mathbb{T}$, we define

$$\varepsilon_\Pi(t, \alpha) = \frac{q(t, u_\Pi(\cdot, \alpha), q_a) - q(t, u, q_a)}{\alpha} - h_\Pi(t, u, q_a).$$

It suffices to prove that $\varepsilon_\Pi(\cdot, \alpha)$ converges uniformly to 0 on $[r, \sigma_1^*(r)]_\mathbb{T}$ as α tends to 0. For every $\alpha \in (0, \alpha_0]$, since the function $\varepsilon_\Pi(\cdot, \alpha)$ vanishes at $t = r$ and is absolutely continuous on $[r, \sigma_1^*(r)]_\mathbb{T}$, $\varepsilon_\Pi(t, \alpha) = \int_{[r,t)_\mathbb{T}} \varepsilon_\Pi^\Delta(\tau, \alpha) \, \Delta\tau$ for every $t \in [r, \sigma_1^*(r)]_\mathbb{T}$, where

$$\varepsilon_\Pi^\Delta(\tau, \alpha) = \frac{f(\tau, q(\tau, u_\Pi(\cdot, \alpha), q_a), u(r) + \alpha(y - u(r))) - f(\tau, q(\tau, u, q_a), u(r))}{\alpha}$$
$$- \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u(r)) \, h_\Pi(\tau, u, q_a) - \frac{\partial f}{\partial u}(\tau, q(\tau, u, q_a), u(r)) \, (y - u(r)),$$

for Δ-a.e. $\tau \in [r, \sigma_1^*(r))_\mathbb{T}$. Using the Taylor formula with integral remainder, we get

$$\varepsilon_\Pi^\Delta(\tau, \alpha) = \int_0^1 \frac{\partial f}{\partial q}(\star_{\theta\alpha\tau}) \, d\theta \, \varepsilon_\Pi(\tau, \alpha)$$
$$+ \left(\int_0^1 \frac{\partial f}{\partial q}(\star_{\theta\alpha\tau}) \, d\theta - \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u(r)) \right) h_\Pi(\tau, u, q_a)$$
$$+ \left(\int_0^1 \frac{\partial f}{\partial u}(\star_{\theta\alpha\tau}) \, d\theta - \frac{\partial f}{\partial u}(\tau, q(\tau, u, q_a), u(r)) \right) (y - u(r)),$$

where

$$\star_{\theta\alpha\tau} = (\tau, q(\tau, u, q_a) + \theta(q(\tau, u_\Pi(\cdot, \alpha), q_a) - q(\tau, u, q_a)), u(r) + \theta\alpha(y - u(r))) \in K_R.$$

It follows that $\|\varepsilon_\Pi^\Delta(\tau, \alpha)\|_{\mathbb{R}^n} \leq \chi_\Pi(\tau, \alpha) + L_R \|\varepsilon_\Pi(\tau, \alpha)\|_{\mathbb{R}^n}$, where

$$\chi_\Pi(\tau, \alpha) = \left\| \left(\int_0^1 \frac{\partial f}{\partial q}(\star_{\theta\alpha\tau}) \, d\theta - \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u(r)) \right) h_\Pi(\tau, u, q_a) \right\|_{\mathbb{R}^n}$$
$$+ \left\| \left(\int_0^1 \frac{\partial f}{\partial u}(\star_{\theta\alpha\tau}) \, d\theta - \frac{\partial f}{\partial u}(\tau, q(\tau, u, q_a), u(r)) \right) (y - u(r)) \right\|_{\mathbb{R}^n}.$$

Therefore, one has

$$\|\varepsilon_\Pi(t, \alpha)\|_{\mathbb{R}^n} \leq \int_{[r,\sigma_1^*(r))_\mathbb{T}} \chi_\Pi(\tau, \alpha) \, \Delta\tau + L_R \int_{[r,t)_\mathbb{T}} \|\varepsilon_\Pi(\tau, \alpha)\|_{\mathbb{R}^n} \, \Delta\tau,$$

for every $t \in [r, \sigma_1^*(r)]_\mathbb{T}$. From Lemma 6, $\|\varepsilon_\Pi(t, \alpha)\|_{\mathbb{R}^n} \leq \Upsilon_\Pi(\alpha) e_{L_R}(\sigma_1^*(r), r)$, for every $t \in [r, \sigma_1^*(r)]_\mathbb{T}$, where $\Upsilon_\Pi(\alpha) = \int_{[r,\sigma_1^*(r))_\mathbb{T}} \chi_\Pi(\tau, \alpha) \, \Delta\tau$.

To conclude, it remains to prove that $\Upsilon_\Pi(\alpha)$ converges to 0 as α tends to 0. Since $q(\tau, u_\Pi(\cdot, \alpha), q_a)$ converges uniformly to $q(\tau, u, q_a)$ on $[r, \sigma_1^*(r)]_\mathbb{T}$ as α tends to 0 (see Lemma 9) and since $\partial f/\partial q$ and $\partial f/\partial u$ are uniformly continuous on K_R, the conclusion follows. $\square$

Then, we define the so-called *second variation vector* $w_\Pi(\cdot, u, q_a)$ associated with the needle-like variation $\Pi = (r, y)$ as the unique solution on $[\sigma_1^*(r), b]_{\mathbb{T}}$ of the linear Δ-Cauchy problem

$$w^\Delta(t) = \frac{\partial f}{\partial q}(t, q(t, u, q_a), u^\Phi(t)) \, w(t), \quad w(\sigma_1^*(r)) = h_\Pi(\sigma_1^*(r), u, q_a). \tag{17}$$

The existence and uniqueness of $w_\Pi(\cdot, u, q_a)$ are ensured by [15, Theorem 3].

Proposition 3. *The mapping*

$$F_\Pi(u, q_a) : \quad \begin{array}{ccc} ([0, \alpha_0], |\cdot|) & \longrightarrow & (\mathrm{C}([\sigma_1^*(r), b]_{\mathbb{T}}, \mathbb{R}^n), \|\cdot\|_\infty) \\ \alpha & \longmapsto & q(\cdot, u_\Pi(\cdot, \alpha), q_a) \end{array}$$

is differentiable at 0, and one has $DF_\Pi(u, q_a)(0) = w_\Pi(\cdot, u, q_a)$.

Proof. We use the notations of the proof of Lemma 8. From Proposition 2, the case $\sigma_1^*(r) = b$ is already proved. As a consequence, we only focus here on the case $\sigma_1^*(r) = \sigma_1(r) < b$.

Recall that $(\tau, q(\tau, u_\Pi(\cdot, \alpha), q_a), u_\Pi^\Phi(\tau, \alpha)) \in K_R$ for Δ-a.e. $\tau \in [a, b)_{\mathbb{T}}$ and for every $\alpha \in [0, \alpha_0]$ (see Remark 28). For every $\alpha \in (0, \alpha_0]$ and every $t \in [\sigma_1(r), b]_{\mathbb{T}}$, we define

$$\varepsilon_\Pi(t, \alpha) = \frac{q(t, u_\Pi(\cdot, \alpha), q_a) - q(t, u, q_a)}{\alpha} - w_\Pi(t, u, q_a).$$

It suffices to prove that $\varepsilon_\Pi(\cdot, \alpha)$ converges uniformly to 0 on $[\sigma_1(r), b]_{\mathbb{T}}$ as α tends to 0. For every $\alpha \in (0, \alpha_0]$, since the function $\varepsilon_\Pi(\cdot, \alpha)$ is absolutely continuous on $[\sigma_1(r), b]_{\mathbb{T}}$, we have $\varepsilon_\Pi(t, \alpha) = \varepsilon_\Pi(\sigma_1(r), \alpha) + \int_{[\sigma_1(r), t)_{\mathbb{T}}} \varepsilon_\Pi^\Delta(\tau, \alpha) \, \Delta\tau$ for every $t \in [\sigma_1(r), b]_{\mathbb{T}}$, where

$$\varepsilon_\Pi^\Delta(\tau, \alpha) = \frac{f(\tau, q(\tau, u_\Pi(\cdot, \alpha), q_a), u^\Phi(\tau)) - f(\tau, q(\tau, u, q_a), u^\Phi(\tau))}{\alpha}$$
$$- \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u^\Phi(\tau)) \, w_\Pi(\tau, u, q_a),$$

for Δ-a.e. $\tau \in [\sigma_1(r), b)_{\mathbb{T}}$. Using the Taylor formula with integral remainder, we get

$$\varepsilon_\Pi^\Delta(\tau, \alpha) = \int_0^1 \frac{\partial f}{\partial q}(\star_{\theta\alpha\tau}) \, d\theta \, \varepsilon_\Pi(\tau, \alpha)$$
$$+ \left(\int_0^1 \frac{\partial f}{\partial q}(\star_{\theta\alpha\tau}) \, d\theta - \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u^\Phi(\tau)) \right) w_\Pi(\tau, u, q_a),$$

where:

$$\star_{\theta\alpha\tau} = (\tau, q(\tau, u, q_a) + \theta(q(\tau, u_\Pi(\cdot, \alpha), q_a) - q(\tau, u, q_a)), u^\Phi(\tau)) \in K_R.$$

It follows that $\|\varepsilon_\Pi^\Delta(\tau, \alpha)\|_{\mathbb{R}^n} \leq \chi_\Pi(\tau, \alpha) + L_R \|\varepsilon_\Pi(\tau, \alpha)\|_{\mathbb{R}^n}$, where

$$\chi_\Pi(\tau, \alpha) = \left\| \left(\int_0^1 \frac{\partial f}{\partial q}(\star_{\theta\alpha\tau}) \, d\theta - \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u^\Phi(\tau)) \right) w_\Pi(\tau, u, q_a) \right\|_{\mathbb{R}^n}.$$

Therefore, one has

$$\|\varepsilon_\Pi(t, \alpha)\|_{\mathbb{R}^n} \leq \|\varepsilon_\Pi(\sigma_1(r), \alpha)\|_{\mathbb{R}^n} + \int_{[\sigma_1(r), b)_{\mathbb{T}}} \chi_\Pi(\tau, \alpha) \, \Delta\tau + L_R \int_{[\sigma_1(r), t)_{\mathbb{T}}} \|\varepsilon_\Pi(\tau, \alpha)\|_{\mathbb{R}^n} \, \Delta\tau,$$

for every $t \in [\sigma_1(r), b]_{\mathbb{T}}$. It follows from Lemma 6 that $\|\varepsilon_\Pi(t, \alpha)\|_{\mathbb{R}^n} \leq \Upsilon_\Pi(\alpha) e_{L_R}(b, \sigma_1(r))$, for every $t \in [\sigma_1(r), b]_{\mathbb{T}}$, where $\Upsilon_\Pi(\alpha) = \|\varepsilon_\Pi(\sigma_1(r), \alpha)\|_{\mathbb{R}^n} + \int_{[\sigma_1(r), b)_{\mathbb{T}}} \chi_\Pi(\tau, \alpha) \, \Delta\tau$.

To conclude, it remains to prove that $\Upsilon_\Pi(\alpha)$ converges to 0 as α tends to 0. First, from Proposition 2, it is easy to see that $\|\varepsilon_\Pi(\sigma_1(r), \alpha)\|_{\mathbb{R}^n}$ converges to 0 as α tends to 0. Second, since $q(\tau, u_\Pi(\cdot, \alpha), q_a)$ converges uniformly to $q(\tau, u, q_a)$ on $[\sigma_1(r), b]_{\mathbb{T}}$ as α tends to 0 (see Lemma 9) and since $\partial f/\partial q$ is uniformly continuous on K_R, we infer that $\int_{[\sigma_1(r), b)_{\mathbb{T}}} \chi_\Pi(\tau, \alpha) \, \Delta\tau$ converges to 0 as α tends to 0. The conclusion follows. $\qquad\square$

Lemma 10. *Let $R > \|u\|_{L^\infty_{\mathbb{T}_1}([a,b]_{\mathbb{T}_1}, \mathbb{R}^m)}$ and let $(u_k, q_{a,k})_{k\in\mathbb{N}}$ be a sequence of elements of $\mathrm{E}_R(u, q_a)$. If u_k converges to u Δ_1-a.e. on $[a,b)_{\mathbb{T}_1}$ and $q_{a,k}$ converges to q_a in $\mathbb{R}^n$ as k tends to $+\infty$, then $h_\Pi(\cdot, u_k, q_{a,k})$ converges uniformly to $h_\Pi(\cdot, u, q_a)$ on $[r, \sigma_1^*(r)]_{\mathbb{T}}$ as k tends to $+\infty$.*

Proof. We use the notations K_R, L_R, ν_R and η_R, defined in Lemma 5 and in its proof.

Let us consider the absolutely continuous function defined by $\varepsilon_k(\cdot) = h_\Pi(\cdot, u_k, q_{a,k}) - h_\Pi(\cdot, u, q_a)$ on $[r, \sigma_1^*(r)]_{\mathbb{T}}$. Let us prove that ε_k converges uniformly to 0 on $[r, \sigma_1^*(r)]_{\mathbb{T}}$ as k tends to $+\infty$. Since $\varepsilon_k(r) = 0$, one has

$$
\begin{aligned}
\varepsilon_k(t) ={}& \int_{[r,t)_{\mathbb{T}}} \frac{\partial f}{\partial q}(\tau, q(\tau, u_k, q_{a,k}), u_k(r)) \, \varepsilon_k(\tau) \, \Delta\tau \\
&+ \int_{[r,t)_{\mathbb{T}}} \left(\frac{\partial f}{\partial q}(\tau, q(\tau, u_k, q_{a,k}), u_k(r)) - \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u(r)) \right) h_\Pi(\tau, u, q_a) \, \Delta\tau \\
&+ \int_{[r,t)_{\mathbb{T}}} \left(\frac{\partial f}{\partial u}(\tau, q(\tau, u_k, q_{a,k}), u_k(r)) - \frac{\partial f}{\partial u}(\tau, q(\tau, u, q_a), u(r)) \right) (y - u(r)) \, \Delta\tau \\
&+ \int_{[r,t)_{\mathbb{T}}} \frac{\partial f}{\partial u}(\tau, q(\tau, u_k, q_{a,k}), u_k(r)) \, (u(r) - u_k(r)) \, \Delta\tau
\end{aligned}
$$

for every $t \in [r, \sigma_1^*(r)]_{\mathbb{T}}$ and every $k \in \mathbb{N}$. Since $(u_k, q_{a,k}) \in \mathrm{E}_R(u, q_a)$ for every $k \in \mathbb{N}$, it follows from Remark 28 that $(\tau, q(\tau, u_k, q_{a,k}), u_k^\Phi(\tau)) \in K_R$ for Δ-a.e. $\tau \in [a,b)_{\mathbb{T}}$. Hence it follows from Lemma 6 that $\|\varepsilon_k(t)\|_{\mathbb{R}^n} \leq \Upsilon_k e_{L_R}(b, r)$, for every $t \in [r, \sigma_1^*(r)]_{\mathbb{T}}$, where Υ_k is given by

$$
\begin{aligned}
\Upsilon_k ={}& L_R \mu_1(r) \|u(r) - u_k(r)\|_{\mathbb{R}^m} \\
&+ \int_{[r,\sigma_1^*(r))_{\mathbb{T}}} \left\| \frac{\partial f}{\partial q}(\tau, q(\tau, u_k, q_{a,k}), u_k(r)) - \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u(r)) \right\| \|h_\Pi(\tau, u, q_a)\|_{\mathbb{R}^n} \, \Delta\tau \\
&+ \int_{[r,\sigma_1^*(r))_{\mathbb{T}}} \left\| \frac{\partial f}{\partial u}(\tau, q(\tau, u_k, q_{a,k}), u_k(r)) - \frac{\partial f}{\partial u}(\tau, q(\tau, u, q_a), u(r)) \right\| \|y - u(r)\|_{\mathbb{R}^m} \, \Delta\tau.
\end{aligned}
$$

Since $\mu_{\Delta_1}(\{r\}) = \mu_1(r) > 0$, $u_k(r)$ converges to $u(r)$ as k tends to $+\infty$. Moreover, from the Lebesgue dominated convergence theorem, $(u_k, q_{a,k})$ converges to (u, q_a) in $(\mathrm{E}_R(u, q_a), \|\cdot\|_{\mathcal{U}\mathcal{Q}^b})$ and, from Lemma 7, $q(\cdot, u_k, q_{a,k})$ converges uniformly to $q(\cdot, u, q_a)$ on $[a,b]_{\mathbb{T}}$ as k tends to $+\infty$. Since $\partial f/\partial q$ and $\partial f/\partial u$ are uniformly continuous on K_R, we conclude that Υ_k converges to 0 as k tends to $+\infty$. The lemma follows. $\qquad\square$

Lemma 11. *Let $R > \|u\|_{L^\infty_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)}$ and let $(u_k, q_{a,k})_{k\in\mathbb{N}}$ be a sequence of elements of $\mathrm{E}_R(u, q_a)$. If u_k converges to u Δ_1-a.e. on $[a,b)_{\mathbb{T}_1}$ and $q_{a,k}$ converges to q_a in $\mathbb{R}^n$ as k tends to $+\infty$, then $w_\Pi(\cdot, u_k, q_{a,k})$ converges uniformly to $w_\Pi(\cdot, u, q_a)$ on $[\sigma_1^*(r), b]_{\mathbb{T}}$ as k tends to $+\infty$.*

Remark 29. With the same arguments of the proof of Lemma 4, the hypotheses of Lemma 11 give the convergences of u_k^Φ to u^Φ Δ-a.e. on $[a,b)_{\mathbb{T}}$ and, from the Lebesgue dominated convergence theorem, of $(u_k, q_{a,k})$ to (u, q_a) in $(\mathrm{E}_R(u, q_a), \|\cdot\|_{\mathcal{U}\mathcal{Q}^b})$.

Proof of Lemma 11. We use the notations K_R, L_R, ν_R and η_R, defined in Lemma 5 and in its proof. From Lemma 10, the case $\sigma_1^*(r) = b$ is already proved. As a consequence, we only focus here on the case $\sigma_1^*(r) = \sigma_1(r) < b$.

Let us consider the absolutely continuous function defined by $\varepsilon_k(\cdot) = w_{\text{II}}(\cdot, u_k, q_{a,k}) - w_{\text{II}}(\cdot, u, q_a)$ on $[\sigma_1(r), b]_{\mathbb{T}}$. Let us prove that ε_k converges uniformly to 0 on $[\sigma_1(r), b]_{\mathbb{T}}$ as k tends to $+\infty$. One has

$$\varepsilon_k(t) = \varepsilon_k(\sigma_1(r)) + \int_{[\sigma_1(r),t)_{\mathbb{T}}} \frac{\partial f}{\partial q}(\tau, q(\tau, u_k, q_{a,k}), u_k^\Phi(\tau))\, \varepsilon_k(\tau)\, \Delta\tau$$
$$+ \int_{[\sigma_1(r),t)_{\mathbb{T}}} \left(\frac{\partial f}{\partial q}(\tau, q(\tau, u_k, q_{a,k}), u_k^\Phi(\tau)) - \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u^\Phi(\tau)) \right) w_{\text{II}}(\tau, u, q_a)\, \Delta\tau,$$

for every $t \in [\sigma_1(r), b]_{\mathbb{T}}$ and every $k \in \mathbb{N}$. Since $(u_k, q_{a,k}) \in \mathrm{E}_R(u, q_a)$ for every $k \in \mathbb{N}$, it follows from Remark 28 that $(\tau, q(\tau, u_k, q_{a,k}), u_k^\Phi(\tau)) \in K_R$ for Δ-a.e. $\tau \in [\sigma_1(r), b)_{\mathbb{T}}$. Hence it follows from Lemma 6 that
$$\|\varepsilon_k(t)\|_{\mathbb{R}^n} \leq (\varepsilon_k(\sigma_1(r)) + \Upsilon_k)e_{L_R}(b, \sigma_1(r)),$$
for every $t \in [\sigma_1(r), b]_{\mathbb{T}}$, where Υ_k is given by

$$\Upsilon_k = \int_{[\sigma_1(r),b)_{\mathbb{T}}} \left\| \frac{\partial f}{\partial q}(\tau, q(\tau, u_k, q_{a,k}), u_k^\Phi(\tau)) - \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u^\Phi(\tau)) \right\| \|w_{\text{II}}(\tau, u, q_a)\|_{\mathbb{R}^n}\, \Delta\tau.$$

Since $(u_k, q_{a,k})$ converges to (u, q_a) in $(\mathrm{E}_R(u, q_a), \|\cdot\|_{\mathcal{UQ}^b})$ and from Lemma 7, $q(\cdot, u_k, q_{a,k})$ converges uniformly to $q(\cdot, u, q_a)$ on $[a, b]_{\mathbb{T}}$ as k tends to $+\infty$. Since $\partial f/\partial q$ is continuous and bounded on K_R and since u_k^Φ converges to u^Φ Δ-a.e. on $[a, b)_{\mathbb{T}}$, the Lebesgue dominated convergence theorem concludes that Υ_k converges to 0 as k tends to $+\infty$. Finally, from Lemma 10, one has $\varepsilon_k(\sigma_1(r)) = w_{\text{II}}(\sigma_1(r), u_k, q_{a,k}) - w_{\text{II}}(\sigma_1(r), u, q_a) = h_{\text{II}}(\sigma_1(r), u_k, q_{a,k}) - h_{\text{II}}(\sigma_1(r), u, q_a)$ converges to 0 as k tends to $+\infty$. The lemma follows. $\qquad\square$

4.2.3 Needle-like variation of u at a point $s \in \mathrm{RD}_1$

Let $s \in \mathscr{L}_{[a,b)_{\mathbb{T}}}(f(\cdot, q(\cdot, u, q_a), u^\Phi)) \cap \mathrm{RD}_1$ and $z \in \mathbb{R}^m$. Note that $s \in \mathbb{T}_1$ and then $\Phi(s) = s$. We define the needle-like variation $\text{II} = (s, z)$ of u at s by

$$u_{\text{II}}(t, \beta) = \begin{cases} z & \text{if } t \in [s, s+\beta)_{\mathbb{T}_1}, \\ u(t) & \text{if } t \notin [s, s+\beta)_{\mathbb{T}_1}. \end{cases}$$

for Δ_1-a.e. $t \in [a, b)_{\mathbb{T}_1}$ and for every $\beta \in \mathcal{V}_1^{s,b} \subset \mathcal{V}^{s,b}$.

Lemma 12. *There exists $\beta_0 > 0$ such that $(u_{\text{II}}(\cdot, \beta), q_a) \in \mathcal{UQ}^b$ for every $\beta \in \mathcal{V}_1^{s,b} \cap [0, \beta_0]$.*

Proof. Let $R = \max(\|u\|_{\mathrm{L}_{\mathbb{T}_1}^\infty([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)}, \|z\|_{\mathbb{R}^m}) + 1 > \|u\|_{\mathrm{L}_{\mathbb{T}_1}^\infty([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)}$. We use the notations K_R, L_R, ν_R and η_R, defined in the proof of Lemma 5. For every $\beta \in \mathcal{V}_1^{s,b}$, we have $\|u_{\text{II}}(\cdot, \beta)\|_{\mathrm{L}_{\mathbb{T}_1}^\infty([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)} \leq R$ and
$$\|u_{\text{II}}(\cdot, \beta) - u\|_{\mathrm{L}_{\mathbb{T}_1}^1([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)} = \int_{[s,s+\beta)_{\mathbb{T}_1}} \|z - u(\tau)\|_{\mathbb{R}^m}\, \Delta_1\tau \leq 2R\beta.$$

Hence, there exists $\beta_0 > 0$ such that for every $\beta \in \mathcal{V}_1^{s,b} \cap [0, \beta_0]$, $\|u_{\text{II}}(\cdot, \beta) - u\|_{\mathrm{L}_{\mathbb{T}_1}^1([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)} \leq \nu_R$ and thus $(u_{\text{II}}(\cdot, \beta), q_a) \in \mathrm{E}_R(u, q_a)$. The conclusion then follows from Lemma 5. $\qquad\square$

Lemma 13. *The mapping*

$$F_{\text{II}}(u, q_a): \quad (\mathcal{V}_1^{s,b} \cap [0, \beta_0], |\cdot|) \quad \longrightarrow \quad (\mathrm{C}([a,b]_{\mathbb{T}}, \mathbb{R}^n), \|\cdot\|_\infty)$$
$$\beta \quad \longmapsto \quad q(\cdot, u_{\text{II}}(\cdot, \beta), q_a)$$

is Lipschitzian. In particular, for every $\beta \in \mathcal{V}_1^{s,b} \cap [0, \beta_0]$, $q(\cdot, u_{\mathrm{II}}(\cdot, \beta), q_a)$ converges uniformly to $q(\cdot, u, q_a)$ on $[a, b]_{\mathbb{T}}$ as β tends to 0.

Proof. We use the notations of the proof of Lemma 12. From Lemma 7, there exists $C \geq 0$ (Lipschitz constant of $F_R(u, q_a)$) such that

$$\|q(\cdot, u_{\mathrm{II}}(\cdot, \beta_2), q_a) - q(\cdot, u_{\mathrm{II}}(\cdot, \beta_1), q_a)\|_\infty \leq C\|(u_{\mathrm{II}}(\cdot, \beta_2), q_a) - (u_{\mathrm{II}}(\cdot, \beta_1), q_a)\|_{\mathcal{U}\mathcal{Q}^b}$$
$$\leq 2CR|\beta_2 - \beta_1|,$$

for all β_1 and β_2 in $\mathcal{V}_1^{s,b} \cap [0, \beta_0]$. The lemma follows. $\qquad \square$

According to [15, Theorem 3], we define the *variation vector* $w_{\mathrm{II}}(\cdot, u, q_a)$ associated with the needle-like variation $\mathrm{II} = (s, z)$ as the unique solution on $[s, b]_{\mathbb{T}}$ of the linear Δ-Cauchy problem

$$w^\Delta(t) = \frac{\partial f}{\partial q}(t, q(t, u, q_a), u^\Phi(t)) \, w(t), \quad w(s) = f(s, q(s, u, q_a), z) - f(s, q(s, u, q_a), u(s)). \quad (18)$$

Proposition 4. *For every* $\delta \in \mathcal{V}^{s,b} \backslash \{0\}$*, the mapping*

$$F_{\mathrm{II}}(u, q_a): \quad (\mathcal{V}_1^{s,b} \cap [0, \beta_0], |\cdot|) \quad \longrightarrow \quad (\mathrm{C}([s+\delta, b]_{\mathbb{T}}, \mathbb{R}^n), \|\cdot\|_\infty)$$
$$\beta \quad \longmapsto \quad q(\cdot, u_{\mathrm{II}}(\cdot, \beta), q_a)$$

is differentiable at 0, and one has $DF_{\mathrm{II}}(u, q_a)(0) = w_{\mathrm{II}}(\cdot, u, q_a)$.

Proof. We use the notations of the proof of Lemma 12. Recall that $(\tau, q(\tau, u_{\mathrm{II}}(\cdot, \beta), q_a), u_{\mathrm{II}}^\Phi(\tau, \beta))$ and $(\tau, q(\tau, u_{\mathrm{II}}(\cdot, \beta), q_a), z)$ belong to K_R for every $\beta \in \mathcal{V}_1^{s,b} \cap [0, \beta_0]$ and for Δ-a.e. $\tau \in [a, b)_{\mathbb{T}}$ (see Remark 28). For every $\beta \in \mathcal{V}_1^{s,b} \cap (0, \beta_0]$ and every $t \in [s+\beta, b]_{\mathbb{T}}$, we define

$$\varepsilon_{\mathrm{II}}(t, \beta) = \frac{q(t, u_{\mathrm{II}}(\cdot, \beta), q_a) - q(t, u, q_a)}{\beta} - w_{\mathrm{II}}(t, u, q_a).$$

It suffices to prove that $\varepsilon_{\mathrm{II}}(\cdot, \beta)$ converges uniformly to 0 on $[s+\beta, b]_{\mathbb{T}}$ as β tends to 0. Note that, for every $\delta \in \mathcal{V}^{s,b} \backslash \{0\}$, it suffices to consider $\beta \leq \delta$. For every $\beta \in \mathcal{V}_1^{s,b} \cap (0, \beta_0]$, the function $\varepsilon_{\mathrm{II}}(\cdot, \beta)$ is absolutely continuous on $[s+\beta, b]_{\mathbb{T}}$ and $\varepsilon_{\mathrm{II}}(t, \beta) = \varepsilon_{\mathrm{II}}(s+\beta, \beta) + \int_{[s+\beta, t)_{\mathbb{T}}} \varepsilon_{\mathrm{II}}^\Delta(\tau, \beta) \, \Delta\tau$, for every $t \in [s+\beta, b]_{\mathbb{T}}$, where

$$\varepsilon_{\mathrm{II}}^\Delta(\tau, \beta) = \frac{f(\tau, q(\tau, u_{\mathrm{II}}(\cdot, \beta), q_a), u^\Phi(\tau)) - f(\tau, q(\tau, u, q_a), u^\Phi(\tau))}{\beta}$$
$$- \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u^\Phi(\tau)) \, w_{\mathrm{II}}(\tau, u, q_a),$$

for Δ-a.e. $\tau \in [s+\beta, b)_{\mathbb{T}}$. Using the Taylor formula with integral remainder, we get

$$\varepsilon_{\mathrm{II}}^\Delta(\tau, \beta) = \int_0^1 \frac{\partial f}{\partial q}(\star_{\theta\beta\tau}) \, d\theta \, \varepsilon_{\mathrm{II}}(\tau, \beta)$$
$$+ \left(\int_0^1 \frac{\partial f}{\partial q}(\star_{\theta\beta\tau}) \, d\theta - \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u^\Phi(\tau)) \right) w_{\mathrm{II}}(\tau, u, q_a),$$

where:
$$\star_{\theta\beta\tau} = (\tau, q(\tau, u, q_a) + \theta(q(\tau, u_{\mathrm{II}}(\cdot, \beta), q_a) - q(\tau, u, q_a)), u^\Phi(\tau)) \in K_R.$$

It follows that $\|\varepsilon_{\mathrm{II}}^\Delta(\tau, \beta)\|_{\mathbb{R}^n} \leq \chi_{\mathrm{II}}(\tau, \beta) + L_R\|\varepsilon_{\mathrm{II}}(\tau, \beta)\|_{\mathbb{R}^n}$, where

$$\chi_{\mathrm{II}}(\tau, \beta) = \left\| \left(\int_0^1 \frac{\partial f}{\partial q}(\star_{\theta\beta\tau}) \, d\theta - \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u^\Phi(\tau)) \right) w_{\mathrm{II}}(\tau, u, q_a) \right\|_{\mathbb{R}^n}.$$

27

Therefore, one has

$$\|\varepsilon_{\mathrm{II}}(t,\beta)\|_{\mathbb{R}^n} \le \|\varepsilon_{\mathrm{II}}(s+\beta,\beta)\|_{\mathbb{R}^n} + \int_{[s+\beta,b)_{\mathbb{T}}} \chi_{\mathrm{II}}(\tau,\beta)\,\Delta\tau + L_R \int_{[s+\beta,t)_{\mathbb{T}}} \|\varepsilon_{\mathrm{II}}(\tau,\beta)\|_{\mathbb{R}^n}\,\Delta\tau,$$

for every $t \in [s+\beta,b]_{\mathbb{T}}$, and it follows from Lemma 6 that $\|\varepsilon_{\mathrm{II}}(t,\beta)\|_{\mathbb{R}^n} \le \Upsilon_{\mathrm{II}}(\beta)e_{L_R}(b,s)$, for every $t \in [s+\beta,b]_{\mathbb{T}}$, where $\Upsilon_{\mathrm{II}}(\beta) = \|\varepsilon_{\mathrm{II}}(s+\beta,\beta)\|_{\mathbb{R}^n} + \int_{[s+\beta,b)_{\mathbb{T}}} \chi_{\mathrm{II}}(\tau,\beta)\,\Delta\tau$.

To conclude, it remains to prove that $\Upsilon_{\mathrm{II}}(\beta)$ converges to 0 as β tends to 0. Since $q(\cdot, u_{\mathrm{II}}(\cdot,\beta), q_a)$ converges uniformly to $q(\cdot, u, q_a)$ on $[s+\beta, b]_{\mathbb{T}}$ as β tends to 0 (see Lemma 13) and since $\partial f/\partial q$ is uniformly continuous on K_R, we first infer that $\int_{[s+\beta,b)_{\mathbb{T}}} \chi_{\mathrm{II}}(\tau,\beta)\,\Delta\tau$ converges to 0 as β tends to 0. Secondly, let us prove that $\|\varepsilon_{\mathrm{II}}(s+\beta,\beta)\|_{\mathbb{R}^n}$ converges to 0 as β tends to 0. By continuity, $w_{\mathrm{II}}(s+\beta, u, q_a)$ converges to $w_{\mathrm{II}}(s, u, q_a)$ as β to 0. Moreover, since $q(\cdot, u_{\mathrm{II}}(\cdot,\beta), q_a)$ converges uniformly to $q(\cdot, u, q_a)$ on $[a,b]_{\mathbb{T}}$ as β tends to 0 and since f is uniformly continuous on K_R, it follows that $f(\cdot, q(\cdot, u_{\mathrm{II}}(\cdot,\beta), q_a), z)$ converges uniformly to $f(\cdot, q(\cdot, u, q_a), z)$ on $[a,b]_{\mathbb{T}}$ as β tends to 0. Therefore, it suffices to note that

$$\frac{1}{\beta} \int_{[s,s+\beta)_{\mathbb{T}}} \big(f(\tau, q(\tau, u, q_a), z) - f(\tau, q(\tau, u, q_a), u^{\Phi}(\tau))\big)\,\Delta\tau$$

converges to $w_{\mathrm{II}}(s, u, q_a) = f(s, q(s, u, q_a), z) - f(s, q(s, u, q_a), u(s))$ as β tends to 0 since s is a Δ-Lebesgue point of $f(\cdot, q(\cdot, u, q_a), z)$ by continuity and of $f(\cdot, q(\cdot, u, q_a), u^{\Phi}(\cdot))$ by hypothesis. Then $\|\varepsilon_{\mathrm{II}}(s+\beta,\beta)\|_{\mathbb{R}^n}$ converges to 0 as β tends to 0, and hence $\Upsilon_{\mathrm{II}}(\beta)$ converges to 0 as well. $\qquad\square$

Lemma 14. *Let $R > \|u\|_{\mathrm{L}^{\infty}_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)}$ and let $(u_k, q_{a,k})_{k\in\mathbb{N}}$ be a sequence of elements of $\mathrm{E}_R(u, q_a)$. If u_k converges to u Δ_1-a.e. on $[a,b)_{\mathbb{T}_1}$, $u_k(s)$ converges to $u(s)$ and $q_{a,k}$ converges to q_a as k tends to $+\infty$, then $w_{\mathrm{II}}(\cdot, u_k, q_{a,k})$ converges uniformly to $w_{\mathrm{II}}(\cdot, u, q_a)$ on $[s,b]_{\mathbb{T}}$ as k tends to $+\infty$.*

Proof. The proof is similar to the one of Lemma 11, replacing $\sigma_1(r)$ with s. $\qquad\square$

4.2.4 Variation of the initial condition q_a

Let $q'_a \in \mathbb{R}^n$.

Lemma 15. *There exists $\gamma_0 > 0$ such that $(u, q_a + \gamma q'_a) \in \mathcal{U}\mathcal{Q}^b$ for every $\gamma \in [0, \gamma_0]$.*

Proof. Let $R = \|u\|_{\mathrm{L}^{\infty}_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)} + 1 > \|u\|_{\mathrm{L}^{\infty}_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)}$. We use the notations K_R, L_R, ν_R and η_R, defined in the proof of Lemma 5. There exists $\gamma_0 > 0$ such that $\|q_a + \gamma q'_a - q_a\|_{\mathbb{R}^n} = \gamma \|q'_a\|_{\mathbb{R}^n} \le \eta_R$ for every $\gamma \in [0, \gamma_0]$, and hence $(u, q_a + \gamma q'_a) \in \mathrm{E}_R(u, q_a)$. Then the claim follows from Lemma 5. $\qquad\square$

Lemma 16. *The mapping*

$$F_{q'_a}(u, q_a): \quad \begin{array}{ccc} ([0,\gamma_0], |\cdot|) & \longrightarrow & (\mathrm{C}([a,b]_{\mathbb{T}}, \mathbb{R}^n), \|\cdot\|_{\infty}) \\ \gamma & \longmapsto & q(\cdot, u, q_a + \gamma q'_a) \end{array}$$

is Lipschitzian. In particular, for every $\gamma \in [0, \gamma_0]$, $q(\cdot, u, q_a + \gamma q'_a)$ converges uniformly to $q(\cdot, u, q_a)$ on $[a,b]_{\mathbb{T}}$ as γ tends to 0.

Proof. We use the notations of the proof of Lemma 15. From Lemma 7, there exists $C \ge 0$ (Lipschitz constant of $F_R(u, q_a)$) such that

$$\|q(\cdot, u, q_a + \gamma_2 q'_a) - q(\cdot, u, q_a + \gamma_1 q'_a)\|_{\infty} \le C\|(u, q_a + \gamma_2 q'_a) - (u, q_a + \gamma_1 q'_a)\|_{\mathcal{U}\mathcal{Q}^b}$$
$$= C|\gamma_2 - \gamma_1|\|q'_a\|_{\mathbb{R}^n}.$$

for all γ_1 and γ_2 in $[0, \gamma_0]$. $\qquad\square$

According to [15, Theorem 3], we define the *variation vector* $w_{q'_a}(\cdot, u, q_a)$ associated with the perturbation q'_a as the unique solution on $[a,b]_{\mathbb{T}}$ of the linear Δ-Cauchy problem

$$w^{\Delta}(t) = \frac{\partial f}{\partial q}(t, q(t, u, q_a), u^{\Phi}(t))\, w(t), \quad w(a) = q'_a. \tag{19}$$

Proposition 5. *The mapping*

$$
\begin{aligned}
F_{q'_a}(u, q_a): \quad ([0, \gamma_0], |\cdot|) &\longrightarrow (\mathrm{C}([a,b]_{\mathbb{T}}, \mathbb{R}^n), \|\cdot\|_{\infty}) \\
\gamma &\longmapsto q(\cdot, u, q_a + \gamma q'_a)
\end{aligned}
$$

is differentiable at 0, and one has $DF_{q'_a}(u, q_a)(0) = w_{q'_a}(\cdot, u, q_a)$.

Proof. We use the notations of the proof of Lemma 15. Note that, from Remark 28, $(\tau, q(\tau, u, q_a + \gamma q'_a), u^{\Phi}(\tau)) \in K_R$ for every $\gamma \in [0, \gamma_0]$ and for Δ-a.e. $\tau \in [a, b)_{\mathbb{T}}$. For every $\gamma \in (0, \gamma_0]$ and every $t \in [a, b]_{\mathbb{T}}$, we define

$$\varepsilon_{q'_a}(t, \gamma) = \frac{q(t, u, q_a + \gamma q'_a) - q(t, u, q_a)}{\gamma} - w_{q'_a}(t, u, q_a).$$

It suffices to prove that $\varepsilon_{q'_a}(\cdot, \gamma)$ converges uniformly to 0 on $[a, b]_{\mathbb{T}}$ as γ tends to 0. For every $\gamma \in (0, \gamma_0]$, since the function $\varepsilon_{q'_a}(\cdot, \gamma)$ vanishes at $t = a$ and is absolutely continuous on $[a, b]_{\mathbb{T}}$, $\varepsilon_{q'_a}(t, \gamma) = \int_{[a,t)_{\mathbb{T}}} \varepsilon^{\Delta}_{q'_a}(\tau, \gamma)\, \Delta\tau$, for every $t \in [a, b]_{\mathbb{T}}$, where

$$\varepsilon^{\Delta}_{q'_a}(\tau, \gamma) = \frac{f(\tau, q(\tau, u, q_a + \gamma q'_a), u^{\Phi}(\tau)) - f(\tau, q(\tau, u, q_a), u^{\Phi}(\tau))}{\gamma}$$
$$- \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u^{\Phi}(\tau), \tau)\, w_{q'_a}(\tau, u, q_a),$$

for Δ-a.e. $\tau \in [a, b)_{\mathbb{T}}$. Using the Taylor formula with integral remainder, we get

$$\varepsilon^{\Delta}_{q'_a}(\tau, \gamma) = \int_0^1 \frac{\partial f}{\partial q}(\star_{\theta \gamma \tau})\, d\theta \cdot \varepsilon_{q'_a}(\tau, \gamma)$$
$$+ \left(\int_0^1 \frac{\partial f}{\partial q}(\star_{\theta \gamma \tau})\, d\theta - \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u^{\Phi}(\tau)) \right) w_{q'_a}(\tau, u, q_a),$$

where

$$\star_{\theta \gamma \tau} = (\tau, q(\tau, u, q_a) + \theta(q(\tau, u, q_a + \gamma q'_a) - q(\tau, u, q_a)), u^{\Phi}(\tau)) \in K_R.$$

It follows that $\|\varepsilon^{\Delta}_{q'_a}(\tau, \gamma)\|_{\mathbb{R}^n} \leq \chi_{q'_a}(\tau, \gamma) + L_R \|\varepsilon_{q'_a}(\tau, \gamma)\|_{\mathbb{R}^n}$, where

$$\chi_{q'_a}(\tau, \gamma) = \left\| \left(\int_0^1 \frac{\partial f}{\partial q}(\star_{\theta \gamma \tau})\, d\theta - \frac{\partial f}{\partial q}(\tau, q(\tau, u, q_a), u^{\Phi}(\tau)) \right) w_{q'_a}(\tau, u, q_a) \right\|_{\mathbb{R}^n}.$$

Hence

$$\|\varepsilon_{q'_a}(t, \gamma)\|_{\mathbb{R}^n} \leq \int_{[a,b)_{\mathbb{T}}} \chi_{q'_a}(\tau, \gamma)\, \Delta\tau + L_R \int_{[a,t)_{\mathbb{T}}} \|\varepsilon_{q'_a}(\tau, \gamma)\|_{\mathbb{R}^n}\, \Delta\tau,$$

for every $t \in [a, b]_{\mathbb{T}}$, and it follows from Lemma 6 that $\|\varepsilon_{q'_a}(t, \gamma)\|_{\mathbb{R}^n} \leq \Upsilon_{q'_a}(\gamma) e_{L_R}(b, a)$, for every $t \in [a, b]_{\mathbb{T}}$, where $\Upsilon_{q'_a}(\gamma) = \int_{[a,b)_{\mathbb{T}}} \chi_{q'_a}(\tau, \gamma)\, \Delta\tau$.

To conclude, it remains to prove that $\Upsilon_{q'_a}(\gamma)$ converges to 0 as γ tends to 0. Since $q(\cdot, u, q_a + \gamma q'_a)$ converges uniformly to $q(\cdot, u, q_a)$ on $[a, b]_{\mathbb{T}}$ as γ tends to 0 (see Lemma 16) and since $\partial f / \partial q$ is uniformly continuous on K_R, the conclusion follows. $\qquad \square$

Lemma 17. *Let $R > \|u\|_{\mathrm{L}^{\infty}_{\mathbb{T}}([a,b)_{\mathbb{T}}, \mathbb{R}^m)}$ and let $(u_k, q_{a,k})_{k \in \mathbb{N}}$ be a sequence of elements of $\mathrm{E}_R(u, q_a)$. If u_k converges to u Δ_1-a.e. on $[a, b)_{\mathbb{T}_1}$ and $q_{a,k}$ converges to q_a in $\mathbb{R}^n$ as k tends to $+\infty$, then $w_{q'_a}(\cdot, u_k, q_{a,k})$ converges uniformly to $w_{q'_a}(\cdot, u, q_a)$ on $[a, b]_{\mathbb{T}}$ as k tends to $+\infty$.*

Proof. The proof is similar to the one of Lemma 11, replacing $\sigma_1(r)$ with a. $\qquad \square$

4.3 Proof of Theorem 3

We are now in a position to prove the PMP. In the sequel, we consider q^* an optimal trajectory, associated with an optimal sampled-data control u^* and with $b^* \in \mathbb{T}$, with $b^* = b$ if the final time is fixed. We set $q_a^* = q^*(a)$.

4.3.1 The augmented system

As in [43, 46], we consider the *augmented system* in $\mathbb{R}^{n+1}$

$$\bar{q}^{\Delta}(t) = \bar{f}(t, \bar{q}(t), u^{\Phi}(t)), \tag{20}$$

with $\bar{q} = (q, q^0)^{\top}$ the augmented state with values in $\mathbb{R}^n \times \mathbb{R}$, and $\bar{f} : \mathbb{T} \times \mathbb{R}^{n+1} \times \mathbb{R}^m \to \mathbb{R}^{n+1}$, the augmented dynamics, defined by $\bar{f}(t, \bar{q}, u) = (f(t, q, u), f^0(t, q, u))^{\top}$. Note that $\bar{f}$ does not depend on q^0. We will always impose as an initial condition $q^0(a) = 0$, so that $q^0(b) = \int_{[a,b)_{\mathbb{T}}} f^0(\tau, q(\tau), u^{\Phi}(\tau)) \Delta\tau$. Hence, the additional coordinate q^0 stands for the cost.

Denoting $\bar{q}_a^* = (q_a^*, 0)^{\top}$, we have $(u^*, \bar{q}_a^*) \in \mathcal{U}\bar{\mathcal{Q}}^{b^*}$. We set $\bar{q}^* = \bar{q}(\cdot, u^*, \bar{q}_a^*) = (q^*, q^{0*})^{\top}$. Hence, $\bar{q}^*$ is a solution of (20) on $[a, b^*]_{\mathbb{T}}$ associated with the control u^*, satisfying $q^{0*}(a) = 0$ and $g(q^*(a), q^*(b^*)) \in \mathrm{S}$ and minimizing $q^0(b)$ over all possible trajectories $\bar{q}$ solutions of (20) on $[a, b]_{\mathbb{T}}$, where $b \in \mathbb{T}$ and associated with a control $u \in \mathrm{L}^{\infty}_{\mathbb{T}_1}([a, b)_{\mathbb{T}_1}, \Omega)$, satisfying $q^0(a) = 0$ and $g(q(a), q(b)) \in \mathrm{S}$.

4.3.2 Application of the Ekeland variational principle

For the completeness, we recall a simplified (but sufficient) version of the Ekeland variational principle.

Theorem 4 ([26]). *Let* $(\mathrm{E}, d_{\mathrm{E}})$ *be a complete metric space and let* $J : \mathrm{E} \to \mathbb{R}^+$, $\lambda \mapsto J(\lambda)$ *be a continuous nonnegative mapping. Let* $\varepsilon > 0$ *and* $\lambda^* \in \mathrm{E}$ *such that* $J(\lambda^*) \leq \varepsilon$. *Then, there exists* $\lambda_{\varepsilon} \in \mathrm{E}$ *such that* $d_{\mathrm{E}}(\lambda_{\varepsilon}, \lambda^*) \leq \sqrt{\varepsilon}$ *and,* $-\sqrt{\varepsilon}\, d_{\mathrm{E}}(\lambda, \lambda_{\varepsilon}) \leq J(\lambda) - J(\lambda_{\varepsilon})$, *for every* $\lambda \in \mathrm{E}$.

Let $R > \|u^*\|_{\mathrm{L}^{\infty}_{\mathbb{T}_1}([a,b^*)_{\mathbb{T}_1}, \mathbb{R}^m)}$. Recall that $\mathrm{E}_R(u^*, \bar{q}_a^*) \subset \mathcal{U}\bar{\mathcal{Q}}^{b^*}$ (see Lemma 5). We set

$$\mathrm{E}_R^{\Omega,0}(u^*, \bar{q}_a^*) = \{(u, \bar{q}_a) \in \mathrm{E}_R(u^*, \bar{q}_a^*) \mid u \in \mathrm{L}^{\infty}_{\mathbb{T}_1}([a, b^*)_{\mathbb{T}_1}, \Omega), \bar{q}_a = (q_a, 0)\}.$$

Note that $(u^*, \bar{q}_a^*) \in \mathrm{E}_R^{\Omega,0}(u^*, \bar{q}_a^*)$. Since Ω is closed, it follows from the (partial) converse of the Lebesgue dominated convergence theorem that $(\mathrm{E}_R^{\Omega,0}(u^*, \bar{q}_a^*), \|\cdot\|_{\mathcal{U}\bar{\mathcal{Q}}^{b^*}})$ is a closed subset of $\mathrm{L}^1_{\mathbb{T}_1}([a, b^*)_{\mathbb{T}_1}, \mathbb{R}^m) \times \mathbb{R}^{n+1}$ and then is a complete metric space. For every $\varepsilon > 0$, we define the functional $J_{\varepsilon}^R : (\mathrm{E}_R^{\Omega,0}(u^*, \bar{q}_a^*), \|\cdot\|_{\mathcal{U}\bar{\mathcal{Q}}^{b^*}}) \to \mathbb{R}^+$ by

$$J_{\varepsilon}^R(u, \bar{q}_a) = \left(\left(\left(q^0(b^*, u, \bar{q}_a) - q^{0*}(b^*) + \varepsilon \right)^+ \right)^2 + d_{\mathrm{S}}^2 \left(g\left(q_a, q(b^*, u, \bar{q}_a)\right) \right) \right)^{1/2}.$$

Since g and d_{S}^2 are continuous and so is $F_R(u^*, \bar{q}_a^*)$ (see Lemma 7), it follows that J_{ε}^R is continuous on $(\mathrm{E}_R^{\Omega,0}(u^*, \bar{q}_a^*), \|\cdot\|_{\mathcal{U}\bar{\mathcal{Q}}^{b^*}})$. Moreover, one has $J_{\varepsilon}^R(u^*, \bar{q}_a^*) = \varepsilon$ and, from optimality of $q^{0*}(b^*)$, $J_{\varepsilon}^R(u, \bar{q}_a) > 0$ for every $(u, \bar{q}_a) \in \mathrm{E}_R^{\Omega,0}(u^*, \bar{q}_a^*)$. It follows from the Ekeland variational principle that, for every $\varepsilon > 0$, there exists $(u_{\varepsilon}^R, \bar{q}_{a,\varepsilon}^R) \in \mathrm{E}_R^{\Omega,0}(u^*, \bar{q}_a^*)$ such that $\|(u_{\varepsilon}^R, \bar{q}_{a,\varepsilon}^R) - (u^*, \bar{q}_a^*)\|_{\mathcal{U}\bar{\mathcal{Q}}^{b^*}} \leq \sqrt{\varepsilon}$ and

$$-\sqrt{\varepsilon}\, \|(u, \bar{q}_a) - (u_{\varepsilon}^R, \bar{q}_{a,\varepsilon}^R)\|_{\mathcal{U}\bar{\mathcal{Q}}^{b^*}} \leq J_{\varepsilon}^R(u, \bar{q}_a) - J_{\varepsilon}^R(u_{\varepsilon}^R, \bar{q}_{a,\varepsilon}^R), \tag{21}$$

for every $(u, \bar{q}_a) \in \mathrm{E}_R^{\Omega,0}(u^*, \bar{q}_a^*)$. In particular, u_ε^R converges to u^* in $\mathrm{L}_{\mathbb{T}_1}^1([a, b^*)_{\mathbb{T}_1}, \mathbb{R}^m)$ and $\bar{q}_{a,\varepsilon}^R$ converges to $\bar{q}_a^*$ as ε tends to 0. Besides, setting

$$\psi_\varepsilon^{0R} = \frac{-1}{J_\varepsilon^R(u_\varepsilon^R, \bar{q}_{a,\varepsilon}^R)} \left(q^0(b^*, u_\varepsilon^R, \bar{q}_{a,\varepsilon}^R) - q^{0*}(b^*) + \varepsilon \right)^+ \leq 0 \tag{22}$$

and

$$\psi_\varepsilon^R = \frac{-1}{J_\varepsilon^R(u_\varepsilon^R, \bar{q}_{a,\varepsilon}^R)} \left(g(q_{a,\varepsilon}^R, q(b^*, u_\varepsilon^R, \bar{q}_{a,\varepsilon}^R)) - \mathrm{P_S}\big(g(q_{a,\varepsilon}^R, q(b^*, u_\varepsilon^R, \bar{q}_{a,\varepsilon}^R))\big) \right) \in \mathbb{R}^j, \tag{23}$$

note that $|\psi_\varepsilon^{0R}|^2 + \|\psi_\varepsilon^R\|_{\mathbb{R}^j}^2 = 1$ and $-\psi_\varepsilon^R \in \mathcal{O}_\mathrm{S}[\mathrm{P_S}(g(q_{a,\varepsilon}^R, q(b^*, u_\varepsilon^R, \bar{q}_{a,\varepsilon}^R)))]$.

Using a compactness argument, the continuity of $F_R(u^*, \bar{q}_a^*)$ (see Lemma 7), the C^1-regularity of g and the (partial) converse of the Lebesgue dominated convergence theorem, we infer that there exists a sequence $(\varepsilon_k)_{k \in \mathbb{N}}$ of positive real numbers converging to 0 such that $u_{\varepsilon_k}^R$ converges to u^* Δ_1-a.e. on $[a, b^*)_{\mathbb{T}_1}$, $\bar{q}_{a,\varepsilon_k}^R$ converges to $\bar{q}_a^*$, $g(q_{a,\varepsilon_k}^R, q(b^*, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R))$ converges to $g(q_a^*, q^*(b^*)) \in$ S, $\mathrm{d}g(q_{a,\varepsilon_k}^R, q(b^*, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R))$ converges to $\mathrm{d}g(q_a^*, q^*(b^*))$, $\psi_{\varepsilon_k}^{0R}$ converges to some $\psi^{0R} \leq 0$, and $\psi_{\varepsilon_k}^R$ converges to some $\psi^R \in \mathbb{R}^j$ as k tends to $+\infty$, with $|\psi^{0R}|^2 + \|\psi^R\|_{\mathbb{R}^j}^2 = 1$ and $-\psi^R \in \mathcal{O}_\mathrm{S}[g(q_a^*, q^*(b^*))]$ (see Lemma 1).

In the next lemmas, we use the inequality (21) respectively with needle-like variations of $u_{\varepsilon_k}^R$ at right-scattered points of $\mathbb{T}_1$ and at right-dense points of $\mathbb{T}_1$, and then variations of $\bar{q}_{a,\varepsilon_k}^R$. Hence, we infer some important inequalities by taking the limit in k. Note that these variations were defined in Section 4.2 for any dynamics f, and that we apply them here to the augmented system (20), associated with the augmented dynamics $\bar{f}$.

Lemma 18. *For every $r \in [a, b^*)_{\mathbb{T}_1} \cap \mathrm{RS}_1$ and every $y \in \mathrm{D}_{\mathrm{stab}}^\Omega(u^*(r))$, considering the needle-like variation $\Pi = (r, y)$ as defined in Section 4.2.2, one has*

$$\psi^{0R} w_\Pi^0(b^*, u^*, \bar{q}_a^*) + \left\langle \left(\frac{\partial g}{\partial q_2}(q_a^*, q^*(b^*)) \right)^\top \psi^R, w_\Pi(b^*, u^*, \bar{q}_a^*) \right\rangle_{\mathbb{R}^n} \leq 0, \tag{24}$$

where the variation vector $\bar{w}_\Pi = (w_\Pi, w_\Pi^0)^\top$ is defined by (17) (replacing f with $\bar{f}$).

Proof. Since $u_{\varepsilon_k}^R$ converges to u^* Δ_1-a.e. on $[a, b^*)_{\mathbb{T}_1}$, it follows that $u_{\varepsilon_k}^R(r)$ converges to $u^*(r)$ as k tends to $+\infty$, where $\|u^*(r)\|_{\mathbb{R}^m} < R$. It follows that $\|u_{\varepsilon_k}^R(r)\|_{\mathbb{R}^m} < R$ and 0 is not isolated in $\Xi = \{\alpha \in [0, 1], \ u_{\varepsilon_k}^R(r) + \alpha(y - u_{\varepsilon_k}^R(r)) \in \Omega\}$ for any sufficiently large k, see Definition 1. Fixing such a large k, one has $u_{\varepsilon_k, \Pi}^R(\cdot, \alpha) \in \mathrm{L}_{\mathbb{T}_1}^\infty([a, b^*)_{\mathbb{T}_1}, \Omega)$ and

$$\|u_{\varepsilon_k, \Pi}^R(\cdot, \alpha)\|_{\mathrm{L}_{\mathbb{T}_1}^\infty([a, b^*)_{\mathbb{T}_1}, \mathbb{R}^m)} \leq \max(\|u_{\varepsilon_k}^R\|_{\mathrm{L}_{\mathbb{T}_1}^\infty([a, b^*)_{\mathbb{T}_1}, \mathbb{R}^m)}, \|u_{\varepsilon_k, \Pi}^R(r, \alpha)\|_{\mathbb{R}^m})$$

$$\leq \max(R, \|u_{\varepsilon_k}^R(r)\|_{\mathbb{R}^m} + \alpha\|y - u_{\varepsilon_k}^R(r)\|_{\mathbb{R}^m}),$$

for every $\alpha \in \Xi$. Moreover, one has

$$\|u_{\varepsilon_k, \Pi}^R(\cdot, \alpha) - u^*\|_{\mathrm{L}_{\mathbb{T}_1}^1([a, b^*)_{\mathbb{T}_1}, \mathbb{R}^m)} \leq \|u_{\varepsilon_k, \Pi}^R(\cdot, \alpha) - u_{\varepsilon_k}^R\|_{\mathrm{L}_{\mathbb{T}_1}^1([a, b^*)_{\mathbb{T}_1}, \mathbb{R}^m)} + \|u_{\varepsilon_k}^R - u^*\|_{\mathrm{L}_{\mathbb{T}_1}^1([a, b^*)_{\mathbb{T}_1}, \mathbb{R}^m)}$$

$$\leq \alpha\mu_1(r)\|y - u_{\varepsilon_k}^R(r)\|_{\mathbb{R}^m} + \sqrt{\varepsilon_k}.$$

Therefore $(u_{\varepsilon_k, \Pi}^R(\cdot, \alpha), \bar{q}_{a,\varepsilon_k}^R) \in \mathrm{E}_R^{\Omega,0}(u^*, \bar{q}_a^*)$ for every $\alpha \in \Xi$ sufficiently small and every k sufficiently large. It then follows from (21) that

$$-\sqrt{\varepsilon_k} \|u_{\varepsilon_k, \Pi}^R(\cdot, \alpha) - u_{\varepsilon_k}^R\|_{\mathrm{L}_{\mathbb{T}_1}^1([a, b^*)_{\mathbb{T}_1}, \mathbb{R}^m)} \leq J_{\varepsilon_k}^R(u_{\varepsilon_k, \Pi}^R(\cdot, \alpha), \bar{q}_{a,\varepsilon_k}^R) - J_{\varepsilon_k}^R(u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R),$$

and thus

$$-\sqrt{\varepsilon_k}\, \mu_1(r)\|y - u_{\varepsilon_k}^R(r)\|_{\mathbb{R}^m} \leq \frac{J_{\varepsilon_k}^R(u_{\varepsilon_k, \Pi}^R(\cdot, \alpha), \bar{q}_{a,\varepsilon_k}^R)^2 - J_{\varepsilon_k}^R(u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R)^2}{\alpha(J_{\varepsilon_k}^R(u_{\varepsilon_k, \Pi}^R(\cdot, \alpha), \bar{q}_{a,\varepsilon_k}^R) + J_{\varepsilon_k}^R(u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R))}.$$

31

Using Proposition 3, we infer that

$$\lim_{\alpha \to 0} \frac{J^R_{\varepsilon_k}(u^R_{\varepsilon_k,\Pi}(\cdot,\alpha),\bar{q}^R_{a,\varepsilon_k})^2 - J^R_{\varepsilon_k}(u^R_{\varepsilon_k},\bar{q}^R_{a,\varepsilon_k})^2}{\alpha}$$
$$= 2\big(q^0(b^*,u^R_{\varepsilon_k},\bar{q}^R_{a,\varepsilon_k}) - q^{0*}(b^*) + \varepsilon_k\big)^+ w^0_\Pi(b^*,u^R_{\varepsilon_k},\bar{q}^R_{a,\varepsilon_k})$$
$$+ 2\Big\langle g(q^R_{a,\varepsilon_k},q(b^*,u^R_{\varepsilon_k},\bar{q}^R_{a,\varepsilon_k})) - \mathrm{P}_\mathrm{S}\big(g(q^R_{a,\varepsilon_k},q(b^*,u^R_{\varepsilon_k},\bar{q}^R_{a,\varepsilon_k}))\big),$$
$$\frac{\partial g}{\partial q_2}(q^R_{a,\varepsilon_k},q(b^*,u^R_{\varepsilon_k},\bar{q}^R_{a,\varepsilon_k}))\, w_\Pi(b^*,u^R_{\varepsilon_k},\bar{q}^R_{a,\varepsilon_k})\Big\rangle_{\mathbb{R}^j}.$$

Since $J^R_{\varepsilon_k}(u^R_{\varepsilon_k,\Pi}(\cdot,\alpha),\bar{q}^R_{a,\varepsilon_k})$ converges to $J^R_{\varepsilon_k}(u^R_{\varepsilon_k},\bar{q}^R_{a,\varepsilon_k})$ as α tends to 0, using (22) and (23) it follows that

$$-\sqrt{\varepsilon_k}\,\mu_1(r)\|y - u^R_{\varepsilon_k}(r)\|_{\mathbb{R}^m} \leq -\psi^{0R}_{\varepsilon_k}w^0_\Pi(b^*,u^R_{\varepsilon_k},\bar{q}^R_{a,\varepsilon_k})$$
$$- \Big\langle \Big(\frac{\partial g}{\partial q_2}(q^R_{a,\varepsilon_k},q(b^*,u^R_{\varepsilon_k},\bar{q}^R_{a,\varepsilon_k}))\Big)^\top \psi^R_{\varepsilon_k}, w_\Pi(b^*,u^R_{\varepsilon_k},\bar{q}^R_{a,\varepsilon_k})\Big\rangle_{\mathbb{R}^n}.$$

By letting k tend to $+\infty$ and using Lemma 11, the lemma follows. $\qquad\square$

We define the sets

$$A = \big\{t \in [a,b^*)_{\mathbb{T}_1} \mid u^R_{\varepsilon_k}(t) \text{ does not converge to } u^*(t) \text{ when } k \text{ tends to } +\infty\big\},$$
$$A_k = \big\{t \in [a,b^*)_{\mathbb{T}_1} \mid t \notin \mathscr{L}_{[a,b^*)_{\mathbb{T}}}(\bar{f}(\cdot,\bar{q}(\cdot,u^R_{\varepsilon_k},\bar{q}^R_{a,k}),u^{R\Phi}_{\varepsilon_k}))\big\}, \quad k \in \mathbb{N}.$$

Lemma 19. *We have $\mu_{\Delta_1}\big(A \cup \bigcup_{k\in\mathbb{N}} A_k\big) = 0$.*

Proof. Since $\mu_{\Delta_1}(A) = 0$, it suffices to prove that $\mu_{\Delta_1}(A_k) = 0$ for every $k \in \mathbb{N}$. Let $k \in \mathbb{N}$. We set

$$B_k = \big\{t \in [a,b^*)_{\mathbb{T}} \mid t \notin \mathscr{L}_{[a,b^*)_{\mathbb{T}}}(\bar{f}(\cdot,\bar{q}(\cdot,u^R_{\varepsilon_k},\bar{q}^R_{a,k}),u^{R\Phi}_{\varepsilon_k}))\big\}.$$

We know that $\mu_\Delta(B_k) = 0$. Hence, $B_k \subset \mathrm{RD}$ and consequently, $\mu_\Delta(B_k) = \mu_L(B_k) = 0$. Since $A_k \subset B_k$, $\mu_L(A_k) = 0$. To conclude, it suffices to prove that $\mu_{\Delta_1}(A_k) = \mu_L(A_k)$. To see this, let us prove that $A_k \subset \mathrm{RD}_1$. By contradiction, let us assume that there exists $t \in A_k \cap \mathrm{RS}_1$. Since $t \in A_k \subset \mathrm{RD}$, we conclude that $u^{R\Phi}_{\varepsilon_k}$ is constant on $[t,\min(\sigma_1(t),b^*))_{\mathbb{T}}$, where $(t,\min(\sigma_1(t),b^*))_{\mathbb{T}} \neq \emptyset$. As a consequence, $\bar{f}(\cdot,\bar{q}(\cdot,u^R_{\varepsilon_k},\bar{q}^R_{a,k}),u^{R\Phi}_{\varepsilon_k})$ is continuous on $[t,\min(\sigma_1(t),b^*))_{\mathbb{T}}$ and consequently $t \in \mathscr{L}_{[a,b^*)_{\mathbb{T}}}(\bar{f}(\cdot,\bar{q}(\cdot,u^R_{\varepsilon_k},\bar{q}^R_{a,k}),u^{R\Phi}_{\varepsilon_k}))$. This leads to a contradiction. $\qquad\square$

We define the set of Lebesgue times $t \in [a,b^*)_{\mathbb{T}_1}$ by

$$\mathscr{L}^{R,1}_{[a,b^*)_{\mathbb{T}}} = [a,b^*)_{\mathbb{T}_1} \setminus \Big(A \cup \bigcup_{k\in\mathbb{N}} A_k\Big).$$

Note that $\mu_{\Delta_1}(\mathscr{L}^{R,1}_{[a,b^*)_{\mathbb{T}}}) = \mu_{\Delta_1}([a,b^*)_{\mathbb{T}_1})$ and then $\mu_{\Delta_1}(\mathscr{L}^{R,1}_{[a,b^*)_{\mathbb{T}}} \cap \mathrm{RD}_1) = \mu_{\Delta_1}([a,b^*)_{\mathbb{T}_1} \cap \mathrm{RD}_1)$.

Lemma 20. *For every $s \in \mathscr{L}^{R,1}_{[a,b^*)_{\mathbb{T}}} \cap \mathrm{RD}_1$ and for every $z \in \Omega \cap \overline{\mathrm{B}}_{\mathbb{R}^m}(0,R)$, considering the needle-like variation $\Pi = (s,z)$ as defined in Section 4.2.3, one has*

$$\psi^{0R}w^0_\Pi(b^*,u^*,\bar{q}^*_a) + \Big\langle \Big(\frac{\partial g}{\partial q_2}(q^*_a,q^*(b^*))\Big)^\top \psi^R, w_\Pi(b^*,u^*,\bar{q}^*_a)\Big\rangle_{\mathbb{R}^n} \leq 0, \qquad (25)$$

where the variation vector $\bar{w}_\Pi = (w_\Pi,w^0_\Pi)^\top$ is defined by (18) (replacing f with $\bar{f}$).

Proof. For every $k \in \mathbb{N}$ and any $\beta \in \mathcal{V}_1^{s,b^*}$, we recall that $u_{\varepsilon_k,\mathrm{II}}^R(\cdot, \beta) \in \mathrm{L}_{\mathbb{T}_1}^\infty([a,b^*)_{\mathbb{T}_1}, \Omega)$ and

$$\|u_{\varepsilon_k,\mathrm{II}}^R(\cdot, \beta)\|_{\mathrm{L}_{\mathbb{T}_1}^\infty([a,b^*)_{\mathbb{T}_1}, \mathbb{R}^m)} \leq \max(\|u_{\varepsilon_k}^R\|_{\mathrm{L}_{\mathbb{T}_1}^\infty([a,b^*)_{\mathbb{T}_1}, \mathbb{R}^m)}, \|z\|_{\mathbb{R}^m}) \leq R,$$

and

$$\|u_{\varepsilon_k,\mathrm{II}}^R(\cdot, \beta) - u^*\|_{\mathrm{L}_{\mathbb{T}_1}^1([a,b^*)_{\mathbb{T}_1}, \mathbb{R}^m)} \leq \|u_{\varepsilon_k,\mathrm{II}}^R(\cdot, \beta) - u_{\varepsilon_k}^R\|_{\mathrm{L}_{\mathbb{T}_1}^1([a,b^*)_{\mathbb{T}_1}, \mathbb{R}^m)} + \|u_{\varepsilon_k}^R - u^*\|_{\mathrm{L}_{\mathbb{T}_1}^1([a,b^*)_{\mathbb{T}_1}, \mathbb{R}^m)}$$
$$\leq 2R\beta + \sqrt{\varepsilon_k}.$$

Therefore $(u_{\varepsilon_k,\mathrm{II}}^R(\cdot, \beta), \bar{q}_{a,\varepsilon_k}^R) \in \mathrm{E}_R^{\Omega,0}(u^*, \bar{q}_a^*)$ for every β sufficiently small and every k sufficiently large. It then follows from (21) that

$$-\sqrt{\varepsilon_k}\|u_{\varepsilon_k,\mathrm{II}}^R(\cdot, \beta) - u_{\varepsilon_k}^R\|_{\mathrm{L}_{\mathbb{T}_1}^1([a,b^*)_{\mathbb{T}_1}, \mathbb{R}^m)} \leq J_{\varepsilon_k}^R(u_{\varepsilon_k,\mathrm{II}}^R(\cdot, \beta), \bar{q}_{a,\varepsilon_k}^R) - J_{\varepsilon_k}^R(u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R),$$

and thus

$$-2R\sqrt{\varepsilon_k} \leq \frac{J_{\varepsilon_k}^R(u_{\varepsilon_k,\mathrm{II}}^R(\cdot, \beta), \bar{q}_{a,\varepsilon_k}^R)^2 - J_{\varepsilon_k}^R(u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R)^2}{\beta(J_{\varepsilon_k}^R(u_{\varepsilon_k,\mathrm{II}}^R(\cdot, \beta), \bar{q}_{a,\varepsilon_k}^R) + J_{\varepsilon_k}^R(u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R))}.$$

Using Proposition 4, we infer that

$$\lim_{\beta \to 0} \frac{J_{\varepsilon_k}^R(u_{\varepsilon_k,\mathrm{II}}^R(\cdot, \beta), \bar{q}_{a,\varepsilon_k}^R)^2 - J_{\varepsilon_k}^R(u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R)^2}{\beta}$$
$$= 2\left(q^0(b^*, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R) - q^{0*}(b^*) + \varepsilon_k\right)^+ w_{\mathrm{II}}^0(b^*, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R)$$
$$+ 2\Big\langle g(q_{a,\varepsilon_k}^R, q(b^*, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R)) - \mathrm{P}_\mathrm{S}\big(g(q_{a,\varepsilon_k}^R, q(b^*, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R))\big),$$
$$\frac{\partial g}{\partial q_2}(q_{a,\varepsilon_k}^R, q(b^*, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R)) w_{\mathrm{II}}(b^*, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R)\Big\rangle_{\mathbb{R}^j}.$$

Since $J_{\varepsilon_k}^R(u_{\varepsilon_k,\mathrm{II}}^R(\cdot, \beta), \bar{q}_{a,\varepsilon_k}^R)$ converges to $J_{\varepsilon_k}^R(u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R)$ as β tends to 0, using (22) and (23) it follows that

$$-2R\sqrt{\varepsilon_k} \leq -\psi_{\varepsilon_k}^{0R} w_{\mathrm{II}}^0(b^*, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R) - \Big\langle \Big(\frac{\partial g}{\partial q_2}(q_{a,\varepsilon_k}^R, q(b^*, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R))\Big)^\top \psi_{\varepsilon_k}^R, w_{\mathrm{II}}(b^*, u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R)\Big\rangle_{\mathbb{R}^n}.$$

By letting k tend to $+\infty$, and using Lemma 14, the lemma follows. $\qquad\square$

Lemma 21. *For every $\bar{q}_a \in \mathbb{R}^n \times \{0\}$, considering the variation of initial point as defined in Section 4.2.4, one has*

$$\psi^{0R} w_{\bar{q}_a}^0(b^*, u^*, \bar{q}_a^*) + \Big\langle \Big(\frac{\partial g}{\partial q_2}(q_a^*, q^*(b^*))\Big)^\top \psi^R, w_{\bar{q}_a}(b^*, u^*, \bar{q}_a^*)\Big\rangle_{\mathbb{R}^n}$$
$$\leq -\Big\langle \Big(\frac{\partial g}{\partial q_1}(q_a^*, q^*(b^*))\Big)^\top \psi^R, q_a\Big\rangle_{\mathbb{R}^n}, \quad (26)$$

where the variation vector $\bar{w}_{\bar{q}_a} = (w_{\bar{q}_a}, w_{\bar{q}_a}^0)^\top$ is defined by (19) (replacing f with $\bar{f}$).

Proof. For every $k \in \mathbb{N}$ and every $\gamma \geq 0$, one has

$$\|\bar{q}_{a,\varepsilon_k}^R + \gamma\bar{q}_a - \bar{q}_a^*\|_{\mathbb{R}^{n+1}} \leq \gamma\|\bar{q}_a\|_{\mathbb{R}^{n+1}} + \|\bar{q}_{a,\varepsilon_k}^R - \bar{q}_a^*\|_{\mathbb{R}^{n+1}} \leq \gamma\|\bar{q}_a\|_{\mathbb{R}^{n+1}} + \sqrt{\varepsilon_k}.$$

Therefore $(u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R + \gamma\bar{q}_a) \in \mathrm{E}_R^{\Omega,0}(u^*, \bar{q}_a^*)$ for every γ sufficiently small and every k sufficiently large. It then follows from (21) that

$$-\sqrt{\varepsilon_k}\|\bar{q}_{a,\varepsilon_k}^R + \gamma\bar{q}_a - \bar{q}_{a,\varepsilon_k}^R\|_{\mathbb{R}^{n+1}} \leq J_{\varepsilon_k}^R(u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R + \gamma\bar{q}_a) - J_{\varepsilon_k}^R(u_{\varepsilon_k}^R, \bar{q}_{a,\varepsilon_k}^R),$$

and thus

$$-\sqrt{\varepsilon_k}\,\|\bar{q}_a\|_{\mathbb{R}^{n+1}} \le \frac{J^R_{\varepsilon_k}(u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k} + \gamma\bar{q}_a)^2 - J^R_{\varepsilon_k}(u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k})^2}{\gamma(J^R_{\varepsilon_k}(u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k} + \gamma\bar{q}_a) + J^R_{\varepsilon_k}(u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k}))}.$$

Using Proposition 5, we infer that

$$
\begin{aligned}
&\lim_{\gamma\to 0} \frac{J^R_{\varepsilon_k}(u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k} + \gamma\bar{q}_a)^2 - J^R_{\varepsilon_k}(u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k})^2}{\gamma} \\
&= 2\big(q^0(b^*, u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k}) - q^{0*}(b^*) + \varepsilon_k\big)^+ w^0_{\bar{q}_a}(b^*, u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k}) \\
&\quad + 2\Big\langle g(q^R_{a,\varepsilon_k}, q(b^*, u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k})) - \mathrm{P}_{\mathrm{S}}\big(g(q^R_{a,\varepsilon_k}, q(b^*, u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k}))\big), \\
&\qquad \frac{\partial g}{\partial q_1}(q^R_{a,\varepsilon_k}, q(b^*, u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k}))\, q_a + \frac{\partial g}{\partial q_2}(q^R_{a,\varepsilon_k}, q(b^*, u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k}))\, w_{\bar{q}_a}(b^*, u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k})\Big\rangle_{\mathbb{R}^j}.
\end{aligned}
$$

Since $J^R_{\varepsilon_k}(u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k} + \gamma\bar{q}_a)$ converges to $J^R_{\varepsilon_k}(u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k})$ as γ tends to 0, using (22) and (23) it follows that

$$
\begin{aligned}
-\sqrt{\varepsilon_k}\,\|\bar{q}_a\|_{\mathbb{R}^{n+1}} \le\; & -\psi^{0R}_{\varepsilon_k} w^0_{\bar{q}_a}(b^*, u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k}) - \Big\langle \Big(\frac{\partial g}{\partial q_1}(q^R_{a,\varepsilon_k}, q(b^*, u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k}))\Big)^\top \psi^R_{\varepsilon_k}, q_a \Big\rangle_{\mathbb{R}^n} \\
& - \Big\langle \Big(\frac{\partial g}{\partial q_2}(q^R_{a,\varepsilon_k}, q(b^*, u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k}))\Big)^\top \psi^R_{\varepsilon_k}, w_{\bar{q}_a}(b^*, u^R_{\varepsilon_k}, \bar{q}^R_{a,\varepsilon_k}) \Big\rangle_{\mathbb{R}^n}.
\end{aligned}
$$

By letting k tend to $+\infty$, and using Lemma 17, the lemma follows. $\qquad\square$

At this step, we have obtained in the previous lemmas the three inequalities (24), (25) and (26), valuable for any $R > \|u^*\|_{\mathrm{L}^\infty_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \mathbb{R}^m)}$. Recall that $|\psi^{0R}|^2 + \|\psi^R\|^2_{\mathbb{R}^j} = 1$ and that $-\psi^R \in \mathcal{O}_{\mathrm{S}}[g(q^*_a, q^*(b^*))]$. Then, considering a sequence of real numbers R_ℓ converging to $+\infty$ as ℓ tends to $+\infty$, we infer that there exist $\psi^0 \le 0$ and $\psi \in \mathbb{R}^j$ such that ψ^{0R_ℓ} converges to ψ^0 and ψ^{R_ℓ} converges to ψ as ℓ tends to $+\infty$, and moreover $|\psi^0|^2 + \|\psi\|^2_{\mathbb{R}^j} = 1$ and $-\psi \in \mathcal{O}_{\mathrm{S}}[g(q^*_a, q^*(b^*))]$.

We set $\mathscr{L}^1_{[a,b^*)_{\mathbb{T}}} = \bigcap_{\ell\in\mathbb{N}} \mathscr{L}^{R_\ell,1}_{[a,b^*)_{\mathbb{T}}}$. Note that $\mu_{\Delta_1}(\mathscr{L}^1_{[a,b^*)_{\mathbb{T}}}) = \mu_{\Delta_1}([a,b^*)_{\mathbb{T}_1})$ and consequently $\mu_{\Delta_1}(\mathscr{L}^1_{[a,b^*)_{\mathbb{T}}} \cap \mathrm{RD}_1) = \mu_{\Delta_1}([a,b^*)_{\mathbb{T}_1} \cap \mathrm{RD}_1)$. Taking the limit in ℓ in (24), (25) and (26), we get the following lemma.

Lemma 22. *For every $r \in [a,b^*)_{\mathbb{T}_1} \cap \mathrm{RS}_1$, and every $y \in \mathrm{D}^\Omega_{\mathrm{stab}}(u^*(r))$, one has*

$$\psi^0 w^0_\Pi(b^*, u^*, \bar{q}^*_a) + \Big\langle \Big(\frac{\partial g}{\partial q_2}(q^*_a, q^*(b^*))\Big)^\top \psi, w_\Pi(b^*, u^*, \bar{q}^*_a) \Big\rangle_{\mathbb{R}^n} \le 0, \tag{27}$$

where the variation vector $\bar{w}_\Pi = (w_\Pi, w^0_\Pi)^\top$ associated with the needle-like variation $\Pi = (r,y)$ of u is defined by (17) (replacing f with $\bar{f}$).

For every $s \in \mathscr{L}^1_{[a,b^*)_{\mathbb{T}}} \cap \mathrm{RD}_1$ and every $z \in \Omega$, one has

$$\psi^0 w^0_{\mathrm{II}}(b^*, u^*, \bar{q}^*_a) + \Big\langle \Big(\frac{\partial g}{\partial q_2}(q^*_a, q^*(b^*))\Big)^\top \psi, w_{\mathrm{II}}(b^*, u^*, \bar{q}^*_a) \Big\rangle_{\mathbb{R}^n} \le 0, \tag{28}$$

where the variation vector $\bar{w}_{\mathrm{II}} = (w_{\mathrm{II}}, w^0_{\mathrm{II}})^\top$ associated with the needle-like variation $\mathrm{II} = (s,z)$ of u is defined by (18) (replacing f with $\bar{f}$);

For every $\bar{q}_a \in \mathbb{R}^n \times \{0\}$, one has

$$
\begin{aligned}
\psi^0 w^0_{\bar{q}_a}(b^*, u^*, \bar{q}^*_a) + \Big\langle \Big(\frac{\partial g}{\partial q_2}(q^*_a, q^*(b^*))\Big)^\top \psi, w_{\bar{q}_a}(b^*, u^*, \bar{q}^*_a) \Big\rangle_{\mathbb{R}^n} & \\
\le -\Big\langle \Big(\frac{\partial g}{\partial q_1}(q^*_a, q^*(b^*))\Big)^\top \psi, q_a \Big\rangle_{\mathbb{R}^n}, & \tag{29}
\end{aligned}
$$

where the variation vector $\bar{w}_{\bar{q}_a} = (w_{\bar{q}_a}, w^0_{\bar{q}_a})^\top$ associated with the variation $\bar{q}_a$ of the initial point $\bar{q}^*_a$ is defined by (19) *(replacing f with $\bar{f}$).*

This result concludes the application of the Ekeland variational principle. The last step of the proof consists of deriving the PMP from these inequalities.

4.3.3   Proof of Remark 11

In this subsection, we prove the formulation of the PMP mentioned in Remark 11. Note that we do not prove, at this step, the transversality condition on the final time.

We define $\bar{p} = (p, p^0)^\top$ as the unique solution on $[a, b^*]_\mathbb{T}$ of the backward shifted linear Δ-Cauchy problem

$$\bar{p}^\Delta(t) = -\left(\frac{\partial \bar{f}}{\partial \bar{q}}(t, \bar{q}^*(t), u^{*\Phi}(t)) \right)^\top \bar{p}^\sigma(t), \quad \bar{p}(b^*) = \left(\left(\frac{\partial g}{\partial q_2}(q^*_a, q^*(b^*)) \right)^\top \psi, \psi^0 \right)^\top. \tag{30}$$

The existence and uniqueness of $\bar{p}$ are ensured by [15, Theorem 6]. Since $\bar{f}$ does not depend on q^0, it is clear that p^0 is constant with $p^0 = \psi^0 \leq 0$.

Right-dense points.   Let $s \in \mathscr{L}^1_{[a,b^*)_\mathbb{T}} \cap \mathrm{RD}_1$ and $z \in \Omega$. Since the function $\langle \bar{p}, \bar{w}_\Pi(\cdot, u^*, \bar{q}^*_a) \rangle_{\mathbb{R}^{n+1}}$ is absolutely continuous and satisfies $\langle \bar{p}, \bar{w}_\Pi(\cdot, u^*, \bar{q}^*_a) \rangle^\Delta_{\mathbb{R}^{n+1}} = 0$ Δ-a.e. on $[s, b^*)_\mathbb{T}$ from the Leibniz formula (3), this function is constant on $[s, b^*]_\mathbb{T}$. It thus follows from (28) that

$$\langle \bar{p}(s), \bar{w}_\Pi(s, u^*, \bar{q}^*_a) \rangle_{\mathbb{R}^{n+1}} = \langle \bar{p}(b^*), \bar{w}_\Pi(b^*, u^*, \bar{q}^*_a) \rangle_{\mathbb{R}^{n+1}}$$
$$= \psi^0 w^0_\Pi(b^*, u^*, \bar{q}^*_a) + \left\langle \left(\frac{\partial g}{\partial q_2}(q^*_a, q^*(b^*)) \right)^\top \psi, w_\Pi(b^*, u^*, \bar{q}^*_a) \right\rangle_{\mathbb{R}^n} \leq 0,$$

and since $\bar{w}_\Pi(s, u^*, \bar{q}^*_a) = \bar{f}(s, \bar{q}^*(s), z) - \bar{f}(s, \bar{q}^*(s), u^*(s))$, we finally get

$$\langle \bar{p}(s), \bar{f}(s, \bar{q}^*(s), z) \rangle_{\mathbb{R}^{n+1}} \leq \langle \bar{p}(s), \bar{f}(s, \bar{q}^*(s), u^*(s)) \rangle_{\mathbb{R}^{n+1}}.$$

Since this inequality holds for every $z \in \Omega$, we have obtained the maximization condition

$$H(s, q^*(s), p(s), p^0, u^*(s)) = \max_{z \in \Omega} H(s, q^*(s), p(s), p^0, z).$$

Right-scattered points.   Let $r \in [a, b^*)_{\mathbb{T}_1} \cap \mathrm{RS}_1$ and $y \in \mathrm{D}^\Omega_{\mathrm{stab}}(u^*(r))$. Since $\langle \bar{p}, \bar{w}_\Pi(\cdot, u^*, \bar{q}^*_a) \rangle_{\mathbb{R}^{n+1}}$ is absolutely continuous and satisfies $\langle \bar{p}, \bar{w}_\Pi(\cdot, u^*, \bar{q}^*_a) \rangle^\Delta_{\mathbb{R}^{n+1}} = 0$ Δ-a.e. on $[\sigma^*_1(r), b^*)_\mathbb{T}$ from the Leibniz formula (3), this function is constant on $[\sigma^*_1(r), b^*]_\mathbb{T}$. It thus follows from (27) that

$$\langle \bar{p}(\sigma^*_1(r)), \bar{w}_\Pi(\sigma^*_1(r), u^*, \bar{q}^*_a) \rangle_{\mathbb{R}^{n+1}} = \langle \bar{p}(b^*), \bar{w}_\Pi(b^*, u^*, \bar{q}^*_a) \rangle_{\mathbb{R}^{n+1}}$$
$$= \psi^0 w^0_\Pi(b^*, u^*, \bar{q}^*_a) + \left\langle \left(\frac{\partial g}{\partial q_2}(q^*_a, q^*(b^*)) \right)^\top \psi, w_\Pi(b^*, u^*, \bar{q}^*_a) \right\rangle_{\mathbb{R}^n} \leq 0.$$

We recall that $\bar{w}_\Pi(\sigma^*_1(r), u^*, \bar{q}^*_a) = \bar{h}_\Pi(\sigma^*_1(r), u^*, \bar{q}^*_a)$ where the variation vector $\bar{h}_\Pi = (h_\Pi, h^0_\Pi)^\top$ associated with the needle-like variation $\Pi = (r, y)$ of u is defined by (16) (replacing f by $\bar{f}$). Since $\bar{h}_\Pi(r, u^*, \bar{q}^*_a) = 0$, it follows that

$$\int_{[r, \sigma^*_1(r))_\mathbb{T}} \langle \bar{p}, \bar{h}_\Pi(\cdot, u^*, \bar{q}^*_a) \rangle^\Delta_{\mathbb{R}^{n+1}}(\tau) \, \Delta\tau = \langle \bar{p}(\sigma^*_1(r)), \bar{h}_\Pi(\sigma^*_1(r), u^*, \bar{q}^*_a) \rangle_{\mathbb{R}^{n+1}} \leq 0.$$

Using the Leibniz formula (3) in the integrand and using (16) (replacing f by $\bar{f}$) and (30), we finally get

$$\left\langle \int_{[r, \sigma^*_1(r))_\mathbb{T}} \frac{\partial H}{\partial u}(\tau, q^*(\tau), p^\sigma(\tau), p^0, u^*(r)) \, \Delta\tau \, , \, y - u^*(r) \right\rangle_{\mathbb{R}^m} \leq 0.$$

Transversality conditions. The transversality condition on the adjoint vector p at the final time b^* has been obtained by definition (note that $-\psi \in \mathcal{O}_S[g(q_a^*, q^*(b^*))]$ as mentioned previously). Let us now establish the transversality condition on the adjoint vector p at the initial time a. Let $\bar{q}_a \in \mathbb{R}^n \times \{0\}$. With the same arguments as before, we prove that the function $\langle \bar{p}, \bar{w}_{\bar{q}_a}(\cdot, u^*, \bar{q}_a^*) \rangle_{\mathbb{R}^{n+1}}$ is constant on $[a, b^*]_{\mathbb{T}}$. It thus follows from (29) that

$$
\begin{aligned}
\langle \bar{p}(a), \bar{w}_{\bar{q}_a}(a, u^*, \bar{q}_a^*) \rangle_{\mathbb{R}^{n+1}} &= \langle \bar{p}(b^*), \bar{w}_{\bar{q}_a}(b^*, u^*, \bar{q}_a^*) \rangle_{\mathbb{R}^{n+1}} \\
&= \psi^0 w_{\bar{q}_a}^0(b^*, u^*, \bar{q}_a^*) + \left\langle \left(\frac{\partial g}{\partial q_2}(\bar{q}_a^*, \bar{q}^*(b^*)) \right)^\top \psi, w_{\bar{q}_a}(b^*, u^*, \bar{q}_a^*) \right\rangle_{\mathbb{R}^n} \\
&\le - \left\langle \left(\frac{\partial g}{\partial q_1}(q_a^*, q^*(b^*)) \right)^\top \psi, q_a \right\rangle_{\mathbb{R}^n},
\end{aligned}
$$

and since $\bar{w}_{\bar{q}_a}(a, u^*, \bar{q}_a^*) = \bar{q}_a = (q_a, 0)$, we finally get

$$
\left\langle p(a) + \left(\frac{\partial g}{\partial q_1}(q_a^*, q^*(b^*)) \right)^\top \psi, q_a \right\rangle_{\mathbb{R}^n} \le 0.
$$

Since this inequality holds for every $\bar{q}_a \in \mathbb{R}^n \times \{0\}$, the left-hand equality of (8) follows.

4.3.4 End of the proof

In this subsection, we conclude the proof of Theorem 3. Note that we prove the fourth item of Theorem 3 in Remark 30 below. To conclude the proof of Theorem 3, we will use the result claimed in Remark 11. Let us separate two cases.

Firstly, let us assume that g is submersive at $(q_a^*, q^*(b^*))$. In that case, we have just to prove that the couple (p, p^0) is not trivial. Let us assume that p is trivial. Then $p(a) = p(b^*) = 0$ and, from the transversality conditions on the adjoint vector, ψ belongs to the kernels of $(\partial_{q_1} g(q_a^*, q^*(b^*)))^\top$ and $(\partial_{q_2} g(q_a^*, q^*(b^*)))^\top$. It follows that ψ belongs to the orthogonal of the image of the differential of g at $(q_a^*, q^*(b^*))$. Since g is submersive at $(q_a^*, q^*(b^*))$, it implies that $\psi = 0$. Since the couple (ψ, p^0) is not trivial, we conclude that $p^0 \ne 0$ and then (p, p^0) is not trivial. This concludes the proof of Theorem 3 in this first case.

Secondly, let us assume that g is not necessarily submersive at $(q_a^*, q^*(b^*))$. In that case, one has to note that if q^* is an optimal solution of $(\mathbf{OSDCP})^{\mathbb{T}}_{\mathbb{T}_1}$ associated with the function g and with the closed convex S, then q^* is also an optimal solution of $(\mathbf{OSDCP})^{\mathbb{T}}_{\mathbb{T}_1}$ associated with the function $\tilde{g}$ defined by $\tilde{g}(q_1, q_2) = (q_1, q_2)$ (which is submersive at any point) and with the closed convex $\tilde{S} = \{q_a^*\} \times \{q^*(b^*)\}$. Then, we get back to the above first case, but with a different function g and a different closed convex S that would impact only the third item of Theorem 3.

Remark 30. Throughout this remark, we assume that all asumptions of the fourth item of Theorem 3 are satisfied. As mentioned in Remark 5, Theorem 3 can be easily extended (without the fourth item, for now) to the framework of dynamical systems on time scales with several sampled-data controls with different sets of controlling times. To prove the transversality condition on the final time, we use the *multiscale version* (without the fourth item) of Theorem 3.

Let $\delta > 0$ be such that $[b^* - \delta, b^* + \delta] \subset \mathbb{T}$ and such that (f, f^0) is of class C^1 on $[b^* - \delta, b^* + \delta] \times \mathbb{R}^n \times \mathbb{R}^m$. Obviously, the trajectory q^*, associated with b^* and u^*, is an optimal solution of

the optimal sampled-data control problem, with free final time $b \in (b^* - \delta, b^* + \delta)$, defined by

$$(\mathbf{O})_{[b^*-\delta, b^*+\delta]} \quad \begin{cases} \min \int_{b^*-\delta}^{b} f^0(\tau, q(\tau), u^\Phi(\tau)) \, d\tau, \\ \dot{q}(t) = f(t, q(t), u^\Phi(t)), \quad \text{for a.e. } t \in [b^*-\delta, b), \\ u \in \mathrm{L}^\infty_{\mathbb{T}_1}(\mathbb{T}_1, \Omega), \\ (q(b^*-\delta), g(q^*(a), q(b))) \in \{q^*(b^*-\delta)\} \times \mathrm{S}. \end{cases}$$

We set $\rho^*(s) = \delta(s-1) + b^*$, $x^*(s) = q^* \circ \rho^*(s)$ and $w^*(s) = \delta$, for every $s \in [0,1]$. With the change of variable $x = q \circ \rho$ and with $\dot{\rho} = w$, it is clear that the augmented trajectory (ρ^*, x^*), associated with the augmented control (w^*, u^*), is an optimal solution of the optimal sampled-data control problem, with fixed final time $s = 1$, defined by

$$(\mathbf{O})_{[0,1]} \quad \begin{cases} \min \int_0^1 w(\tau) f^0(\rho(\tau), x(\tau), u^{\Phi \circ \rho}(\tau)) \, d\tau, \\ (\dot{\rho}(t), \dot{x}(t)) = (w(t), w(t) f(\rho(t), x(t), u^{\Phi \circ \rho}(t))), \quad \text{for a.e. } t \in [0,1), \\ (w, u) \in \mathrm{L}^\infty([0,1], [\delta/2, +\infty)) \times \mathrm{L}^\infty_{\mathbb{T}_1}(\mathbb{T}_1, \Omega), \\ (\rho(0), x(0)) = (b^*-\delta, q^*(b^*-\delta)), \\ (\rho(1), g(q^*(a), x(1))) \in [b^*-\delta, b^*+\delta] \times \mathrm{S}. \end{cases}$$

In this new optimal sampled-data control problem, the sampled-data controls w and u are defined on different time scales. Its Hamiltonian $\tilde{H}$ is $\tilde{H}(\rho, x, p_\rho, p_x, p^0, w, u) = p_\rho w + \langle p_x, wf(\rho, x, u) \rangle_{\mathbb{R}^n} + p^0 w f^0(\rho, x, u)$. Applying the *multiscale version* (without the fourth item) of Theorem 3, since w^* takes its values in the interior of the constraint set $[\delta/2, +\infty)$, it follows in particular that

$$\frac{\partial \tilde{H}}{\partial w}(\rho^*(s), x^*(s), p_\rho(s), p_x(s), p^0, w^*(s), u^{*, \Phi \circ \rho^*}(s)) = 0,$$

for almost every $s \in [0,1)$. Moreover, since $\rho^*(1) = b^*$ belongs to the interior of $(b^*-\delta, b^*+\delta)$, we can select (p_ρ, p_x) such that $p_\rho(1) = 0$ (usual transversality condition on the adjoint vector). It follows that

$$\langle p_x(s), f(\rho^*(s), x^*(s), u^{*, \Phi \circ \rho^*}(s)) \rangle_{\mathbb{R}^n} + p^0 f^0(\rho^*(s), x^*(s), u^{*, \Phi \circ \rho^*}(s)) = -p_\rho(s),$$

for almost every $s \in [0,1]$. We have thus proved that the function $t \mapsto H(t, q^*(t), p(t), p^0, u^{*\Phi}(t))$ coincides, almost everywhere on $[b^*-\delta, b^*)$, with a continuous function vanishing at $t = b^*$.

Remark 31. Another method in order to prove the fourth item of Theorem 3 consists of considering variations of the final time b in a neighborhood of b^* and to modify accordingly the functional of Section 4.3.2 to which Ekeland's variational principle is applied.

4.4 Proof of Theorem 2

The following proof can be read independently of the rest of Section 4. It is inspired from [50, Chap. 6]. We only treat the case where the final time $b \in \mathbb{T}$ is fixed. The proof can be easily adapted to the case of a free final time.

Let us consider a sequence $(q_k)_{k \in \mathbb{N}}$ of $\mathcal{M}$, associated with sampled-data controls $u_k \in \mathrm{L}^\infty_{\mathbb{T}_1}(\mathbb{T}_1, \Omega)$, minimizing the cost considered in $(\mathbf{OSDCP})^{\mathbb{T}}_{\mathbb{T}_1}$. It follows from the assumptions that the sequence $(f(\cdot, q_k, u_k^\Phi), f^0(\cdot, q_k, u_k^\Phi))_{k \in \mathbb{N}}$ is bounded in $\mathrm{L}^2_{\mathbb{T}}([a,b)_{\mathbb{T}}, \mathbb{R}^{n+1})$. Hence a subsequence (that we do not

relabel) converges in the weak topology of $L^2_{\mathbb{T}}([a,b)_{\mathbb{T}}, \mathbb{R}^{n+1})$ to a function (F, F^0). Moreover, a subsequence of $(q_k(a))_{k \in \mathbb{N}}$ (that we do not relabel) converges in $\mathbb{R}^n$ and we denote the limit by $x(a)$. We define the absolutely continuous functions $x(t) = x(a) + \int_{[a,t)_{\mathbb{T}}} F(\tau) \Delta \tau$ and $x^0(t) = \int_{[a,t)_{\mathbb{T}}} F^0(\tau) \Delta \tau$. In particular, note the pointwise convergence of $(q_k)_{k \in \mathbb{N}}$ to x on $[a,b]_{\mathbb{T}}$. Since g is continuous and since S is closed, we have $g(x(a), x(b)) \in$ S. Note that $x^0(b)$ is equal to the infimum of admissible costs. To conclude the proof, we have to prove the existence of $u \in L^\infty_{\mathbb{T}_1}(\mathbb{T}_1, \Omega)$ such that $x(t) = x(a) + \int_{[a,t)_{\mathbb{T}}} f(\tau, x(\tau), u^\Phi(\tau)) \Delta \tau$ and $x^0(t) = \int_{[a,t)_{\mathbb{T}}} f^0(\tau, x(\tau), u^\Phi(\tau)) \Delta \tau$.

The sequence $(f(\cdot, x, u_k^\Phi), f^0(\cdot, x, u_k^\Phi))_{k \in \mathbb{N}}$ is bounded in $L^2_{\mathbb{T}}([a,b)_{\mathbb{T}}, \mathbb{R}^{n+1})$. Hence, a subsequence (that we do not relabel) converges in the weak topology of $L^2_{\mathbb{T}}([a,b)_{\mathbb{T}}, \mathbb{R}^{n+1})$ to some $(\tilde{F}, \tilde{F}^0)$. It follows from the Lebesgue dominated convergence theorem and from the global Lipschitz continuity of f in q on $[a,b]_{\mathbb{T}} \times \overline{\mathrm{B}}_{\mathbb{R}^n}(0, M) \times \Omega$ that $\int_{[a,b)_{\mathbb{T}}} \varphi(\tau) F(\tau) \Delta \tau = \int_{[a,b)_{\mathbb{T}}} \varphi(\tau) \tilde{F}(\tau) \Delta \tau$ for every $\varphi \in L^2_{\mathbb{T}}([a,b)_{\mathbb{T}}, \mathbb{R})$. Hence $F = \tilde{F}$ (and similarly $F^0 = \tilde{F}^0$) Δ-a.e. on $[a,b)_{\mathbb{T}}$.

We define the set $\mathcal{W}$ of functions $\Theta \in L^2_{\mathbb{T}}([a,b)_{\mathbb{T}}, \mathbb{R}^{n+1})$ such that $\Theta(t) \in \mathcal{W}(t, x(t))$ for Δ-a.e. $t \in [a,b)_{\mathbb{T}}$. It follows from the assumptions that $\mathcal{W}$ is a closed (convex) subset of $L^2_{\mathbb{T}}([a,b)_{\mathbb{T}}, \mathbb{R}^{n+1})$ with its usual topology, and thus with its weak topology as well (see [19]). We infer that $(F, F^0) = (\tilde{F}, \tilde{F}^0) \in \mathcal{W}$. Hence, for Δ-a.e. $t \in [a,b)_{\mathbb{T}}$, there exists $v(t) \in \Omega$ such that $(F(t), F^0(t)) = (f(t, x(t), v(t)), f^0(t, x(t), v(t)))$. Note that v can be selected μ_Δ-measurable on $[a,b)_{\mathbb{T}}$.[5]

Since $\mathrm{RS}_1 \cap [a,b)_{\mathbb{T}}$ is at most countable, using a diagonal argument, there exists a subsequence of $(u_k)_{k \in \mathbb{N}}$ (that we do not relabel) such that $u_k(r)$ converges to some $u_r \in \Omega$ for every $r \in \mathrm{RS}_1 \cap [a,b)_{\mathbb{T}}$. It follows from the Lebesgue dominated convergence theorem that $x(t) = x(a) + \int_{[a,t)_{\mathbb{T}}} f(\tau, x(\tau), u^\Phi(\tau)) \Delta \tau$ and $x^0(t) = \int_{[a,t)_{\mathbb{T}}} f^0(\tau, x(\tau), u^\Phi(\tau)) \Delta \tau$, where u is defined by $u(t) = u_r$ if $t = r \in \mathrm{RS}_1 \cap [a,b)_{\mathbb{T}}$ and $u(t) = v(t)$ if $t \in \mathrm{RD}_1 \cap [a,b)_{\mathbb{T}}$. It is easy to prove that $u \in L^\infty_{\mathbb{T}_1}([a,b)_{\mathbb{T}_1}, \Omega)$ (the μ_{Δ_1}-measurability of u on $[a,b)_{\mathbb{T}_1}$ is established as in the proof of Proposition 1). This concludes the proof of Theorem 2.

References

[1] R.P. Agarwal and M. Bohner. Basic calculus on time scales and some of its applications. *Results Math.*, 35(1-2):3–22, 1999.

[2] R.P. Agarwal, M. Bohner, and A. Peterson. Inequalities on time scales: a survey. *Math. Inequal. Appl.*, 4(4):535–557, 2001.

[3] R.P. Agarwal, V. Otero-Espinar, K. Perera, and D.R. Vivero. Basic properties of Sobolev's spaces on time scales. *Adv. Difference Equ.*, Art. ID 38121, 14, 2006.

[4] A.A. Agrachev and Y.L. Sachkov. *Control theory from the geometric viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004.

[5] B. Aulbach and L. Neidhart. Integration on measure chains. In *Proceedings of the Sixth International Conference on Difference Equations*, 239–252, Boca Raton, FL, 2004. CRC.

[6] F.M. Atici, D.C. Biles, and A. Lebedinsky. An application of time scales to economics. *Math. Comput. Modelling*, 43(7-8):718–726, 2006.

---

[5]Indeed, similarly to the proof of Proposition 1, one has just to consider the extension function $\tilde{v}$ defined on $[a,b)$ by $\tilde{v}(t) = v(t)$ if $t \in \mathbb{T}$ and by $\tilde{v}(t) = v(r)$ if $t \in (r, \sigma(r))$, where $r \in \mathrm{RS}$. Then, the function $\tilde{v}$ can be selected μ_L-measurable on $[a,b)$ from [43, Lemma 3A p. 161], which implies the μ_Δ-measurability of v on $[a,b)_{\mathbb{T}}$ from [23].

[7] Z. Bartosiewicz and D.F.M. Torres. Noether's theorem on time scales. *J. Math. Anal. Appl.*, 342(2):1220–1226, 2008.

[8] M. Bohner. Calculus of variations on time scales. *Dynam. Systems Appl.*, 13(3-4):339–349, 2004.

[9] M. Bohner and G.S. Guseinov. Double integral calculus of variations on time scales. *Comput. Math. Appl.*, 54(1):45–57, 2007.

[10] M. Bohner and A. Peterson. *Dynamic equations on time scales. An introduction with applications.* Birkhäuser Boston Inc., Boston, MA, 2001.

[11] M. Bohner and A. Peterson. *Advances in dynamic equations on time scales.* Birkhäuser Boston Inc., Boston, MA, 2003.

[12] V.G. Boltyanskiĭ. *Optimal control of discrete systems.* John Wiley & Sons, New York-Toronto, Ont., 1978.

[13] B. Bonnard, M. Chyba, *The role of singular trajectories in control theory.* Springer Verlag, 2003.

[14] B. Bonnard, L. Faubourg, and E. Trélat. *Mécanique céleste et contrôle des véhicules spatiaux,* volume 51 of *Mathématiques & Applications (Berlin) [Mathematics & Applications].* Springer-Verlag, Berlin, 2006.

[15] L. Bourdin, E. Trélat. General Cauchy-Lipschitz theory for Delta-Cauchy problems with Carathéodory dynamics on time scales. *J. Difference Equ. Appl.*, 20(4):526-547, 2014.

[16] L. Bourdin, E. Trélat. Pontryagin maximum principle for finite dimensional nonlinear optimal control problems on time scales. *SIAM J. Control Optim.*, 51(5):3781-3813, 2013.

[17] L. Bourdin. Nonshifted calculus of variations on time scales with nabla-differentiable sigma. *J. Math. Anal. Appl.*, 411(2):543-554, 2014.

[18] A. Bressan and B. Piccoli. *Introduction to the mathematical theory of control,* volume 2 of *AIMS Series on Applied Mathematics.* Springfield, MO, 2007.

[19] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations.* Springer, New York, 2011.

[20] J.a.e. Bryson and Y.C. Ho. *Applied optimal control.* Hemisphere Publishing Corp. Washington, D. C., 1975. Optimization, estimation, and control, Revised printing.

[21] F. Bullo, A.D. Lewis, *Geometric control of mechanical systems. Modeling, analysis, and design for simple mechanical control systems.* Texts in Applied Mathematics, 49, Springer-Verlag, New York, 2005.

[22] A. Cabada and D.R. Vivero. Criterions for absolute continuity on time scales. *J. Difference Equ. Appl.*, 11(11):1013–1028, 2005.

[23] A. Cabada and D.R. Vivero. Expression of the Lebesgue Δ-integral on time scales as a usual Lebesgue integral: application to the calculus of Δ-antiderivatives. *Math. Comput. Modelling*, 43(1-2):194–207, 2006.

[24] M.D. Canon, J.C.D. Cullum, and E. Polak. *Theory of optimal control and mathematical programming.* McGraw-Hill Book Co., New York, 1970.

[25] L. Cesari, *Optimization – theory and applications. Problems with ordinary differential equations.* Applications of Mathematics, 17, Springer-Verlag, New York, 1983.

[26] I. Ekeland. On the variational principle. *J. Math. Anal. Appl.*, 47:324–353, 1974.

[27] L. Fan and C. Wang. *The discrete maximum principle: a study of multistage systems optimization.* John Wiley & Sons, New York, 1964.

[28] R.A.C. Ferreira and D.F.M. Torres. Higher-order calculus of variations on time scales. In *Mathematical control theory and finance*, pages 149–159. Springer, Berlin, 2008.

[29] J.G.P. Gamarra and R.V. Solvé. Complex discrete dynamics from simple continuous population models. *Bull. Math. Biol.*, 64:611–620, 2002.

[30] R.V. Gamkrelidze. Discovery of the maximum principle. In *Mathematical events of the twentieth century*, pages 85–99. Springer, Berlin, 2006.

[31] G.S. Guseinov. Integration on time scales. *J. Math. Anal. Appl.*, 285(1):107–127, 2003.

[32] H. Halkin. A maximum principle of the Pontryagin type for systems described by nonlinear difference equations. *SIAM J. Control*, 4:90–111, 1966.

[33] M.R. Hestenes. *Calculus of variations and optimal control theory.* Robert E. Krieger Publishing Co. Inc., Huntington, N.Y., 1980. Corrected reprint of the 1966 original.

[34] S. Hilger. *Ein Maßkettenkalkül mit Anwendungen auf Zentrumsmannigfaltigkeiten.* PhD thesis, Universität Würzburg, 1988.

[35] R. Hilscher and V. Zeidan. Calculus of variations on time scales: weak local piecewise C_{rd}^1 solutions with variable endpoints. *J. Math. Anal. Appl.*, 289(1):143–166, 2004.

[36] R. Hilscher and V. Zeidan. First-order conditions for generalized variational problems over time scales. *Comput. Math. Appl.*, 62(9):3490–3503, 2011.

[37] R. Hilscher and V. Zeidan. Weak maximum principle and accessory problem for control problems on time scales. *Nonlinear Anal.*, 70(9):3209–3226, 2009.

[38] R. Hilscher and V. Zeidan. Time scale embedding theorem and coercivity of quadratic functionals. *Analysis (Munich)*, 28(1):1–28, 2008.

[39] J.M. Holtzman. Convexity and the maximum principle for discrete systems. *IEEE Trans. Automatic Control*, AC-11:30–35, 1966.

[40] J.-B. Hiriart-Urruty. *La commande optimale pour les débutants.* Opuscules. Ellipses, Paris, 2008.

[41] J.M. Holtzman and H. Halkin. Discretional convexity and the maximum principle for discrete systems. *SIAM J. Control*, 4:263–275, 1966.

[42] V. Jurdjevic, *Geometric control theory.* Cambridge Studies in Advanced Mathematics, 52, Cambridge University Press, 1997.

[43] E. B. Lee and L. Markus, *Foundations of optimal control theory.* John Wiley, New York, 1967.

[44] R.M. May. Simple mathematical models with very complicated dynamics. *Nature*, 261:459–467, 1976.

[45] B.S. Mordukhovich. *Variational analysis and generalized differentiation, I: Basic theory, II: Applications.* Volumes 330 and 331 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 2006.

[46] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mishchenko. The mathematical theory of optimal processes. Interscience Publishers John Wiley & Sons, Inc. New York-London, 1962.

[47] H. Schättler, U. Ledzewicz, *Geometric optimal control, theory, methods and examples.* Interdisciplinary Applied Mathematics, Vol. 38, Springer, 2012.

[48] A. Seierstad and K. Sydsæter. *Optimal control theory with economic applications.* North-Holland Publishing Co., Amsterdam, Volume 24, 1987.

[49] S.P. Sethi and G.L. Thompson. *Optimal control theory. Applications to management science and economics.* Kluwer Academic Publishers, Boston, MA, second edition, 2000.

[50] E. Trélat. *Contrôle optimal, théorie & applications.* Mathématiques Concrètes. Vuibert, Paris, 2005.

[51] E. Trélat. *Optimal control and applications to aerospace: some results and challenges.* J. Optim. Theory Appl. 154(3):713–758, 2012.