J Optim Theory Appl (2013) 159:1–40
DOI 10.1007/s10957-013-0299-3

First- and Second-Order Optimality Conditions for Optimal Control Problems of State Constrained Integral Equations

J. Frédéric Bonnans · Constanza de la Vega · Xavier Dupuis

Received: 24 May 2012 / Accepted: 4 March 2013 / Published online: 21 March 2013

Abstract This paper deals with optimal control problems of integral equations, with initial–final and running state constraints. The order of a running state constraint is defined in the setting of integral dynamics, and we work here with constraints of arbitrary high orders. First-order necessary conditions of optimality are given by the description of the set of Lagrange multipliers. Second-order necessary conditions are expressed by the nonnegativity of the supremum of some quadratic forms. Second-order sufficient conditions are also obtained in the case where these quadratic forms are of Legendre type.

Keywords Optimal control · Integral equations · State constraints · Second-order optimality conditions

1 Introduction

The dynamics in the optimal control problems we consider in this paper is given by an integral equation. Such equations, sometimes called nonlinear Volterra integral equations, belong to the family of equations with memory and thus are found in many

Communicated by Hans Joseph Pesch.

J.F. Bonnans · X. Dupuis
Inria Saclay & CMAP, Ecole Polytechnique, 91128 Palaiseau, France

J.F. Bonnans
e-mail: frederic.bonnans@inria.fr

X. Dupuis
e-mail: xavier.dupuis@cmap.polytechnique.fr

C. de la Vega (✉)
IMAS-CONICET & Departamento de Matemática, UBA, Buenos Aires, Argentina
e-mail: csfvega@dm.uba.ar

models. Among the fields of application of these equations are population dynamics in biology and growth theory in economy: see [1] or its translation in [2] for one of the first use of integral equations in ecology in 1927 by Volterra, who contributed earlier to their theoretical study [3]; in 1976, Kamien and Muller [4] modeled the capital replacement problem by an optimal control problem with an integral state equation. First-order optimality conditions for such problems were known under the form of a maximum principle since Vinokurov's paper [5] in 1967 (translated in 1969 [6–8]), whose proof has been questioned by Neustadt and Warga [9] in 1970. Maximum principles have then been provided by Bakke [10], Carlson [11], or more recently by de la Vega [12] for an optimal terminal time control problem. First-order optimality conditions for control problems of the more general family of equations with memory are obtained by Carlier and Tahraoui [13].

None of the previously cited articles considers what we will call "running state constraints." That is what Bonnans and de la Vega did in [14], where they provide Pontryagin's principle, i.e., first-order optimality conditions. In this work we are particularly interested in second-order necessary conditions in presence of running state constraints. Such constraints drive to optimization problems with inequality constraints in the infinite-dimensional space of continuous functions. Thus second-order necessary conditions on a so-called *critical cone* will contain an extra term, as it has been discovered in 1988 by Kawasaki [15] and generalized in 1990 by Cominetti [16] in an abstract setting. It is possible to compute this extra term in the case of state constrained optimal control problems; this is done by Páles and Zeidan [17] and Bonnans and Hermant [18, 19] in the framework of ODEs.

Our strategy here is different and follows [20], with the differences that we work with integral equations and that we add initial–final state constraints that lead to nonunique Lagrange multipliers. The idea was already present in [15] and is closely related to the concept of extended polyhedricity [21]: the extra term mentioned above vanishes if we write second-order necessary conditions on a subset of the critical cone, the so-called *radial critical cone*. This motivates introducing an auxiliary optimization problem, the *reduced problem*, for which under some assumptions the radial critical cone is dense in the critical cone. Optimality conditions for the reduced problem are relevant for the original problem, and the extra term now appears as the derivative of a new constraint in the reduced problem. We will devote a lot of effort to the proof of the density result, and we will mention a flaw in [20] concerning this proof.

The paper is organized as follows. We set the optimal control problem, define Lagrange multipliers, and work on the notion of order of a running state constraint in our setting in Sect. 2. The reduced problem is introduced in Sect. 3, followed by first-order necessary conditions and second-order necessary conditions on the radial critical cone. The main results are presented in Sect. 4. After some specific assumptions, we state and prove the technical Lemma 4.1, which is then used to strengthen the first-order necessary conditions already obtained and to get the density result that we need. With this density result, we obtain second-order necessary conditions on the critical cone. Second-order sufficient conditions are also given in this section. Some of the technical aspects are postponed to Appendix.

Notations We denote by h_t the value of a function h at time t if h depends only on t, and by $h_{i,t}$ its ith component if h is vector-valued. To avoid confusion, we denote partial derivatives of a function h of (t,x) by $D_t h$ and $D_x h$, and we keep the symbol D without any subscript for the differentiation w.r.t. all variables. We identify the dual space of $\mathbb{R}^n$ with the space $\mathbb{R}^{n*}$ of n-dimensional horizontal vectors. Generally, we denote by X^* the dual space of a topological vector space X. We denote by $|\cdot|$ both the Euclidean norm on finite-dimensional vector spaces and the cardinal of a finite set, and by $\|\cdot\|_s$ and $\|\cdot\|_{q,s}$ the standard norms on the Lebesgue spaces L^s and the Sobolev spaces $W^{q,s}$, respectively.

2 Optimal Control of State Constrained Integral Equations

2.1 Setting

We consider an optimal control problem with running and initial–final state constraints of the following type:

$$(P) \qquad \min_{(u,y)\in U\times Y} \int_0^T \ell(u_t, y_t)\,\mathrm{d}t + \phi(y_0, y_T) \tag{1}$$

$$\text{subject to} \quad y_t = y_0 + \int_0^t f(t,s,u_s,y_s)\,\mathrm{d}s, \quad t \in [0,T], \tag{2}$$

$$g(y_t) \le 0, \quad t \in [0,T], \tag{3}$$

$$\Phi^E(y_0, y_T) = 0, \tag{4}$$

$$\Phi^I(y_0, y_T) \le 0, \tag{5}$$

where

$$U := L^\infty\big([0,T];\mathbb{R}^m\big) \quad \text{and} \quad Y := W^{1,\infty}\big([0,T];\mathbb{R}^n\big)$$

are the *control space* and *state space*, respectively.

The data are $\ell\colon \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$, $\phi\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $f\colon \mathbb{R}\times\mathbb{R}\times\mathbb{R}^m\times\mathbb{R}^n \to \mathbb{R}^n$, $g\colon \mathbb{R}^n \to \mathbb{R}^r$, $\Phi^E: \mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}^{s_E}$, $\Phi^I: \mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}^{s_I}$, and $T>0$. We make the following assumption:

(A0) $\ell, \phi, f, g, \Phi^E, \Phi^I$ are of class C^∞, and f is Lipschitz.

Remark 2.1

1. We set τ as the symbol for the first variable of f. Observe that if $D_\tau f \equiv 0$, we recover an optimal control problem of a state constrained ODE. More generally, if $D^d_{\tau^d} f \equiv 0$, then the integral equation (2) can be written as a system of controlled differential equations by adding $d-1$ state variables.
2. The running cost ℓ and the running state constraints g appear in some applications as functions of (t,u,y) and (t,y), respectively. It fits our framework if ℓ and g are of class C^∞ w.r.t. all variables by adding a state variable, but the case where they are not regular w.r.t. t is not treated here.

We call a *trajectory* a pair $(u, y) \in U \times Y$ which satisfies the *state equation* (2). Under assumption (A0) it can be shown by standard contraction arguments that for any $(u, y_0) \in U \times \mathbb{R}^n$, the state equation (2) has a unique solution y in Y, denoted by $y[u, y_0]$. Moreover, the map $\Gamma : U \times \mathbb{R}^n \to Y$ defined by $\Gamma(u, y_0) := y[u, y_0]$ is of class C^∞.

2.2 Lagrange Multipliers

The dual space of the space of vector-valued continuous functions $C([0, T]; \mathbb{R}^r)$ is the space of finite vector-valued Radon measures $M([0, T]; \mathbb{R}^{r*})$ under the pairing

$$\langle \mu, h \rangle := \int_{[0,T]} \mathrm{d}\mu_t\, h_t = \sum_{1 \le i \le r} \int_{[0,T]} h_{i,t}\, \mathrm{d}\mu_{i,t}.$$

We define $BV([0, T]; \mathbb{R}^{n*})$, the space of vector-valued functions of bounded variation, as follows: let I be an open set that contains $[0, T]$; then

$$BV\big([0, T]; \mathbb{R}^{n*}\big) := \big\{h \in L^1\big(I; \mathbb{R}^{n*}\big) : Dh \in M\big(I; \mathbb{R}^{n*}\big), \operatorname{supp}(Dh) \subset [0, T]\big\},$$

where Dh is the distributional derivative of h; if h is of bounded variation, we denote it by $\mathrm{d}h$. For $h \in BV([0, T]; \mathbb{R}^{n*})$, there exist $h_{0_-}, h_{T_+} \in \mathbb{R}^{n*}$ such that

$$\begin{aligned} h &= h_{0_-} \quad \text{a.e. on }]-\infty, 0[\cap I, \\ h &= h_{T_+} \quad \text{a.e. on }]T, +\infty[\cap I. \end{aligned} \tag{6}$$

Conversely, we can identify any measure $\mu \in M([0, T]; \mathbb{R}^{r*})$ with the derivative of a function of bounded variation, denoted again by μ, such that $\mu_{T_+} = 0$. This motivates the notation $\mathrm{d}\mu$ for any measure in the sequel, setting implicitly $\mu_{T_+} = 0$. See Appendix A.1 for more details.

Let

$$M := M\big([0, T]; \mathbb{R}^{r*}\big), \qquad P := BV\big([0, T]; \mathbb{R}^{n*}\big).$$

For $p \in P$, we define the *Hamiltonian* $H[p] : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ by

$$H[p](t, u, y) := \ell(u, y) + p_t f(t, t, u, y) + \int_t^T p_s D_\tau f(s, t, u, y)\, \mathrm{d}s \tag{7}$$

and, for $\Psi \in \mathbb{R}^{s*}$, the *end points Lagrangian* $\Phi[\Psi] : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ by

$$\Phi[\Psi](y_1, y_2) := \phi(y_1, y_2) + \Psi \Phi(y_1, y_2), \tag{8}$$

where $s := s_E + s_I$ and $\Phi := (\Phi^E, \Phi^I)$. We also denote $K := \{0\}_{s_E} \times \mathbb{R}^{s_I}_-$, so that (4)–(5) can be rewritten as $\Phi(y_0, y_T) \in K$. The *normal cone* to K at a point $\bar{\Phi}$, denoted by $N_K(\bar{\Phi})$ and defined as the polar cone of the tangent cone $T_K(\bar{\Phi})$, has here the following characterization: $\Psi \in N_K(\bar{\Phi})$ iff

$$\bar{\Phi} \in K, \quad \Psi_i \ge 0, \quad \Psi_i \bar{\Phi}_i = 0, \quad i = s_E + 1, \ldots, s_E + s_I. \tag{9}$$

Given a trajectory (u, y) and $(\mathrm{d}\eta, \Psi) \in M \times \mathbb{R}^{s*}$, the *adjoint state* p, whenever it exists, is defined as the solution in P of

$$\begin{cases} -\mathrm{d}p_t = D_y H[p](t, u_t, y_t)\,\mathrm{d}t + \mathrm{d}\eta_t g'(y_t), \\ (-p_{0_-}, p_{T_+}) = D\Phi[\Psi](y_0, y_T). \end{cases} \tag{10}$$

Note that $\mathrm{d}\eta_t\, g'(y_t) = \sum_{i=1}^r \mathrm{d}\eta_{i,t}\, g_i'(y_t)$. The adjoint state does not exist in general, but when it does, it is unique. More precisely, we have the following:

Lemma 2.1 *There exists a unique solution in P of the adjoint state equation with final condition only* (*i.e.*, *without initial condition*):

$$\begin{cases} -\mathrm{d}p_t = D_y H[p](t, u_t, y_t)\,\mathrm{d}t + \mathrm{d}\eta_t\, g'(y_t), \\ p_{T_+} = D_{y_2}\Phi[\Psi](y_0, y_T). \end{cases} \tag{11}$$

Proof The contraction argument is given in Appendix A.1. □

We can now define Lagrange multipliers for optimal control problems in our setting:

Definition 2.1 The triple $(\mathrm{d}\eta, \Psi, p) \in M \times \mathbb{R}^{s*} \times P$ is a *Lagrange multiplier* associated with $(\bar{u}, \bar{y})$ iff

$$p \text{ is the adjoint state associated with } (\bar{u}, \bar{y}, \mathrm{d}\eta, \Psi), \tag{12}$$

$$\mathrm{d}\eta \geq 0, \quad g(\bar{y}) \leq 0, \quad \int_{[0,T]} \mathrm{d}\eta_t\, g(\bar{y}_t) = 0, \tag{13}$$

$$\Psi \in N_K\big(\Phi(\bar{y}_0, \bar{y}_T)\big), \tag{14}$$

$$D_u H[p](t, \bar{u}_t, \bar{y}_t) = 0 \quad \text{for a.a. } t \in [0, T]. \tag{15}$$

2.3 Linearized State Equation

For $s \in [1, \infty]$, let

$$V_s := L^s\big([0, T]; \mathbb{R}^m\big), \qquad Z_s := W^{1,s}\big([0, T]; \mathbb{R}^n\big).$$

Given a trajectory (u, y) and $(v, z_0) \in V_s \times \mathbb{R}^n$, we consider the *linearized state equation* in Z_s

$$z_t = z_0 + \int_0^t D_{(u,y)} f(t, s, u_s, y_s)(v_s, z_s)\,\mathrm{d}s. \tag{16}$$

It is easily shown that there exists a unique solution $z \in Z_s$ of (16), called the *linearized state* associated with the trajectory (u, y) and direction (v, z_0) and denoted by $z[v, z_0]$ (keeping in mind the nominal trajectory).

Lemma 2.2 *There exist $C>0$ and $C_s>0$ for any $s\in[1,\infty]$ (depending on (u,y)) such that, for all $(v,z_0)\in V_s\times\mathbb{R}^n$ and all $t\in[0,T]$,*

$$\big|z[v,z_0]_t\big|\le C\bigg(|z_0|+\int_0^t|v_s|\,\mathrm{d}s\bigg), \tag{17}$$

$$\big\|z[v,z_0]\big\|_{1,s}\le C_s\big(|z_0|+\|v\|_s\big). \tag{18}$$

Proof of Eq. (17) comes from Gronwall's lemma, and (18) follows from (17). □

For $s=\infty$, the linearized state equation arises naturally: let $(u,y_0)\in U\times\mathbb{R}^n$ and $y:=\Gamma(u,y_0)\in Y$. We consider the linearized state associated with the trajectory (u,y) and a direction $(v,z_0)\in U\times\mathbb{R}^n$. Then

$$z[v,z_0]=D\Gamma(u,y_0)(v,z_0). \tag{19}$$

Similarly, we can define the *second-order linearized state* $z^2[v,z_0]$ as the unique solution in $Z_{s/2}$ of

$$z_t^2=\int_0^t\Big(D_yf(t,s,u_s,y_s)z_s^2+D^2_{(u,y)^2}f(t,s,u_s,y_s)\big(v_s,z[v,z_0]_s\big)^2\Big)\,\mathrm{d}s \tag{20}$$

for $(v,z_0)\in V_s\times\mathbb{R}^n$ and $s\in[2,\infty]$. If $s=\infty$, then

$$z^2[v,z_0]=D^2\Gamma(u,y_0)(v,z_0)^2. \tag{21}$$

2.4 Running State Constraints

The running state constraints g_i, $i=1,\dots,r$, are considered along trajectories (u,y). They produce functions of one variable, $t\mapsto g_i(y_t)$, which belong to $W^{1,\infty}([0,T])$ a priori and satisfy

$$\frac{\mathrm{d}}{\mathrm{d}t}g_i(y_t)=g_i'(y_t)\bigg(f(t,t,u_t,y_t)+\int_0^tD_\tau f(t,s,u_s,y_s)\,\mathrm{d}s\bigg). \tag{22}$$

There are two parts in this derivative:

- $t\mapsto g_i'(y_t)f(t,t,u_t,y_t)$, where u appears pointwise.
- $t\mapsto g_i'(y_t)\int_0^tD_\tau f(t,s,u_s,y_s)\mathrm{d}s$, where u appears in an integral.

Below we will distinguish these two behaviors and set $\tilde u$ as the symbol for the pointwise variable, u for the integral variable (similarly for y). If there is no dependance on $\tilde u$, one can again differentiate (22) w.r.t. t. This motivates the definition of a notion of total derivative that always "forgets" the dependence on $\tilde u$. Let us do that formally.

First, we need a set which is stable under operations such as in (22), so that it will contain the derivatives of any order. It is also of interest to know how the functions we consider depend on $(u,y)\in U\times Y$. To answer this double issue, we define the following commutative ring:

$$S:=\bigg\{h:h(t,\tilde u,\tilde y,u,y)=\sum_\alpha a_\alpha(t,\tilde u,\tilde y)\prod_\beta\int_0^tb_{\alpha,\beta}(t,s,u_s,y_s)\,\mathrm{d}s\bigg\}, \tag{23}$$

where $(t,\tilde{u},\tilde{y},u,y)\in\mathbb{R}\times\mathbb{R}^m\times\mathbb{R}^n\times U\times Y$, a_α and $b_{\alpha,\beta}$ are real functions of class C^∞, the sum and the products are finite, and an empty product is equal to 1. The following is straightforward:

Lemma 2.3 *Let $h\in S$, $(u,y)\in U\times Y$. There exists $C>0$ such that, for a.a. $t\in[0,T]$ and for all $(\tilde{v},\tilde{z},v,z)\in\mathbb{R}^m\times\mathbb{R}^n\times U\times Y$,*

$$\left|D_{(\tilde{u},\tilde{y},u,y)}h(t,u_t,y_t,u,y)(\tilde{v},\tilde{z},v,z)\right|\leq C\left(|\tilde{v}|+|\tilde{z}|+\int_0^t\big(|v_s|+|z_s|\big)\,\mathrm{d}s\right).$$

Next, we define the derivation $D^{(1)}\colon S\longrightarrow S$ as follows (recall that we set τ as the symbol for the first variable of f or b):

1. For $h\colon(t,\tilde{u},\tilde{y})\in\mathbb{R}\times\mathbb{R}^m\times\mathbb{R}^n\mapsto a(t,\tilde{u},\tilde{y})\in\mathbb{R}$,

$$\begin{aligned}&\big(D^{(1)}h\big)(t,\tilde{u},\tilde{y},u,y)\\&\quad:=D_t a(t,\tilde{u},\tilde{y})+D_{\tilde{y}}a(t,\tilde{u},\tilde{y})\left(f(t,t,\tilde{u},\tilde{y})+\int_0^t D_\tau f(t,s,u_s,y_s)\,\mathrm{d}s\right).\end{aligned}$$

2. For $h\colon(t,u,y)\in\mathbb{R}\times U\times Y\mapsto\int_0^t b(t,s,u_s,y_s)\,\mathrm{d}s\in\mathbb{R}$,

$$\big(D^{(1)}h\big)(t,\tilde{u},\tilde{y},u,y):=b(t,t,\tilde{u},\tilde{y})+\int_0^t D_\tau b(t,s,u_s,y_s)\,\mathrm{d}s.$$

3. For any $h_1,h_2\in S$,

$$\begin{aligned}\big(D^{(1)}(h_1+h_2)\big)&=\big(D^{(1)}h_1\big)+\big(D^{(1)}h_2\big),\\\big(D^{(1)}(h_1h_2)\big)&=\big(D^{(1)}h_1\big)h_2+h_1\big(D^{(1)}h_2\big).\end{aligned}$$

It is clear that $D^{(1)}h\in S$ for any $h\in S$. The following formula, which is easily checked for $h=a(t,\tilde{u},\tilde{y})$ and $h=\int_0^t b(t,s,u_s,y_s)\,\mathrm{d}s$, will be used for any $h\in S$:

$$\begin{aligned}\big(D^{(1)}h\big)(t,u_t,y_t,u,y)&=D_t h(t,u_t,y_t,u,y)+D_{\tilde{y}}h(t,u_t,y_t,u,y)f(t,t,u_t,y_t)\\&\quad+D_{\tilde{y}}h(t,u_t,y_t,u,y)\int_0^t D_\tau f(t,s,u_s,y_s)\,\mathrm{d}s.\end{aligned}\tag{24}$$

Let us now highlight two important properties of $D^{(1)}$. First, it is a notion of total derivative:

Lemma 2.4 *Let $h\in S$ be such that $D_{\tilde{u}}h\equiv 0$, $(u,y)\in U\times Y$ be a trajectory, and*

$$\varphi\colon t\mapsto h(t,u_t,y_t,u,y).\tag{25}$$

Then $\varphi\in W^{1,\infty}([0,T])$, and

$$\frac{\mathrm{d}\varphi}{\mathrm{d}t}(t)=\big(D^{(1)}h\big)(t,u_t,y_t,u,y).\tag{26}$$

Proof We write h as in (23). If $D_{\tilde{u}} h \equiv 0$, then for any $u_0 \in \mathbb{R}^m$,

$$\varphi(t) = h(t, u_0, y_t, u, y) \tag{27}$$

$$= \sum_\alpha a_\alpha(t, u_0, y_t) \prod_\beta \int_0^t b_{\alpha,\beta}(t, s, u_s, y_s)\,\mathrm{d}s. \tag{28}$$

By (28), $\varphi \in W^{1,\infty}([0, T])$, and, by (27),

$$\begin{aligned}\frac{\mathrm{d}\varphi}{\mathrm{d}t}(t) &= D_t h(t, u_0, y_t, u, y) + D_{\tilde{y}} h(t, u_0, y_t, u, y)\dot{y}_t \\ &= D_t h(t, u_t, y_t, u, y) + D_{\tilde{y}} h(t, u_t, y_t, u, y)\dot{y}_t\end{aligned}$$

since $D_{\tilde{u}} D_t h \equiv D_t D_{\tilde{u}} h \equiv 0$ and $D_{\tilde{u}} D_{\tilde{y}} h \equiv 0$. Using the expression of $\dot{y}_t$ and (24), we recognize (26). □

Second, it satisfies a principle of commutation with the linearization:

Lemma 2.5 *Let h and (u, y) be as in Lemma 2.4. Let $s \in [1, \infty]$, $(v, z_0) \in V_s \times \mathbb{R}^n$, $z := z[v, z_0] \in Z_s$, and*

$$\psi : t \mapsto D_{(\tilde{y},u,y)} h(t, u_t, y_t, u, y)(z_t, v, z). \tag{29}$$

Then $\psi \in W^{1,s}([0, T])$, and

$$\frac{\mathrm{d}\psi}{\mathrm{d}t}(t) = D_{(\tilde{u},\tilde{y},u,y)}\big[\big(D^{(1)} h\big)(t, u_t, y_t, u, y)\big](v_t, z_t, v, z). \tag{30}$$

Proof Using $D_{\tilde{u}} D_{(\tilde{y},u,y)} h \equiv 0$, we have

$$\begin{aligned}\psi(t) &= D_{(\tilde{y},u,y)} h(t, u_0, y_t, u, y)(z_t, v, z) \\ &= \sum_\alpha D_{\tilde{y}} a_\alpha(t, u_0, y_t) z_t \prod_\beta \int_0^t b_{\alpha,\beta}\,\mathrm{d}s \\ &\quad + \sum_{\alpha,\beta} a_\alpha(t, u_0, y_t) \int_0^t D_{(u,y)} b_{\alpha,\beta}(t, s, u_s, y_s)(v_s, z_s)\,\mathrm{d}s \prod_{\beta' \neq \beta} \int_0^t b_{\alpha,\beta'}\,\mathrm{d}s.\end{aligned}$$

This implies that $\psi \in W^{1,s}([0, T])$ and that

$$\begin{aligned}\frac{\mathrm{d}\psi}{\mathrm{d}t}(t) &= D^2_{t,(\tilde{y},u,y)} h(t, u_t, y_t, u, y)(z_t, v, z) \\ &\quad + D^2_{\tilde{y},(\tilde{y},u,y)} h(t, u_t, y_t, u, y)\big(\dot{y}_t, (z_t, v, z)\big) + D_{\tilde{y}} h(t, u_t, y_t, u, y)\dot{z}_t.\end{aligned}$$

On the other hand, we differentiate $D^{(1)} h$ w.r.t. $(\tilde{u}, \tilde{y}, u, y)$ using (24). Then with the expressions of $\dot{y}_t$ and $\dot{z}_t$, we get relation (30). □

The same principle is true at the second-order:

Lemma 2.6 *Let h and (u, y) be as in Lemma 2.4. Let $s \in [2, \infty]$, $(v, z_0) \in V_s \times \mathbb{R}^n$, $z := z[v, z_0], \in Z_s$, $z^2 := z^2[v, z_0] \in Z_{s/2}$, and*

$$\phi\colon t \mapsto D^2_{(\tilde{y},u,y)^2} h(t, u_t, y_t, u, y)(z_t, v, z)^2 + D_{(\tilde{y},y)} h(t, u_t, y_t, u, y)\big(z_t^2, z^2\big). \quad (31)$$

Then $\phi \in W^{1,s/2}([0, T])$, and

$$\frac{\mathrm{d}\phi}{\mathrm{d}t}(t) = D^2_{(\tilde{u},\tilde{y},u,y)^2}\big[\big(D^{(1)}h\big)(t, u_t, y_t, u, y)\big](v_t, z_t, v, z)^2 + D_{(\tilde{y},y)}\big[\big(D^{(1)}h\big)(t, u_t, y_t, u, y)\big]\big(z_t^2, z^2\big). \quad (32)$$

Proof We apply the definitions and the results of this section to a problem where the control variables are (u, v), the state variables are (y, z), and the dynamics is given by

$$\begin{cases} y_t = y_0 + \int_0^t f(t, s, u_s, y_s)\,\mathrm{d}s, \\ z_t = z_0 + \int_0^t D_{(u,y)} f(t, s, u_s, y_s)(v_s, z_s)\,\mathrm{d}s. \end{cases} \quad (33)$$

Note that the linearized dynamics at (u, v, y, z) in the direction $(v, 0, z_0, 0)$ is given by

$$\begin{cases} z_t = z_0 + \int_0^t D_{u,y)} f(t, s, u_s, y_s)(v_s, z_s)\,\mathrm{d}s, \\ z_t^2 = \int_0^t \big(D_y f(t, s, u_s, y_s) z_s^2 + D^2(u, y)^2 f(t, s, u_s, y_s)(v_s, z_s)^2\big)\,\mathrm{d}s. \end{cases} \quad (34)$$

Let H be defined by

$$H(t, \tilde{u}, \tilde{v}, \tilde{y}, \tilde{z}, u, v, y, z) := D_{(\tilde{u},\tilde{y},u,y)} h(t, \tilde{u}, \tilde{y}, u, y)(\tilde{v}, \tilde{z}, v, z). \quad (35)$$

If $D_{\tilde{u}} h \equiv 0$, then $D_{(\tilde{u},\tilde{v})} H \equiv 0$, and

$$D_{(\tilde{y},\tilde{z},u,v,y,z)} H(t, u_t, v_t, y_t, z_t, u, v, y, z)\big(z_t, z_t^2, v, 0, z, z^2\big) = D^2_{(\tilde{y},u,y)^2} h(t, u_t, y_t, u, y)(z_t, v, z)^2 + D_{(\tilde{y},y)} h(t, u_t, y_t, u, y)\big(z_t^2, z^2\big). \quad (36)$$

By Lemma 2.5, the time derivative of this function is

$$D_{(\tilde{u},\tilde{v},\tilde{y},\tilde{z},u,v,y,z)}\big[\big(D^{(1)}H\big)(t, u_t, v_t, y_t, z_t, u, v, y, z)\big]\big(v_t, 0, z_t, z_t^2, v, 0, z, z^2\big), \quad (37)$$

and by Lemma 2.4, the definition of H, and Lemma 2.5 again, we get successively

$$\big(D^{(1)}H\big)(t, u_t, v_t, y_t, z_t, u, v, y, z) = \frac{\mathrm{d}}{\mathrm{d}t} H(t, u_t, v_t, y_t, z_t, u, v, y, z),$$

$$
\begin{aligned}
&= \frac{\mathrm{d}}{\mathrm{d}t} D_{(\tilde{y},u,y)} h(t,u_t,y_t,u,y)(z_t,v,z), \\
&= D_{(\tilde{u},\tilde{y},u,y)}\big[\big(D^{(1)}h\big)(t,u_t,y_t,u,y)\big](v_t,z_t,v,z). \qquad (38)
\end{aligned}
$$

Then Eq. (37) becomes

$$
\begin{aligned}
&D^2_{(\tilde{u},\tilde{y},u,y)^2}\big[\big(D^{(1)}h\big)(t,u_t,y_t,u,y)\big](v_t,z_t,v,z)^2 \\
&\quad + D_{(\tilde{y},y)}\big[\big(D^{(1)}h\big)(t,u_t,y_t,u,y)\big]\big(z_t^2,z^2\big), \qquad (39)
\end{aligned}
$$

and Lemma 2.6 is proved. □

Finally, we define the order of a running state constraint g_i. Let $g_i^{(0)} := g_i$ and $g_i^{(j+1)} := D^{(1)} g_i^{(j)}$. Note that $g_i \in S$, so $g_i^{(j)} \in S$ for all $j \geq 0$. Moreover, if we write $g_i^{(j)}$ as in (23), the a_α and $b_{\alpha,\beta}$ are combinations of derivatives of f and g_i.

Definition 2.2 The *order* of the constraint g_i is the greatest positive integer q_i such that

$$
D_{\tilde{u}} g_i^{(j)} \equiv 0 \quad \text{for all } j = 0, \dots, q_i - 1.
$$

We have a result similar to Lemma 9 in [18], but now for integral dynamics and up to the second-order. Let $(u,y) \in U \times Y$ be a trajectory, $(v,z_0) \in V_s \times \mathbb{R}^n$, $z := z[v,z_0] \in Z_s$, and $z^2 := z^2[v,z_0] \in Z_{s/2}$ for some $s \in [2,\infty]$.

Lemma 2.7 *Let g_i be of order at least $q_i \in \mathbb{N}$. Then*

$$
t \mapsto g_i(y_t) \in W^{q_i,\infty}\big([0,T]\big), \qquad (40)
$$

$$
t \mapsto g_i'(y_t) z_t \in W^{q_i,s}\big([0,T]\big), \qquad (41)
$$

$$
t \mapsto g_i''(y_t)(z_t)^2 + g_i'(y_t) z_t^2 \in W^{q_i,s/2}\big([0,T]\big), \qquad (42)
$$

and for $j = 1, \dots, q_i$,

$$
\frac{\mathrm{d}^j}{\mathrm{d}t^j} g_i(y)\bigg|_t = g_i^{(j)}(t,u_t,y_t,u,y), \qquad (43)
$$

$$
\frac{\mathrm{d}^j}{\mathrm{d}t^j} g_i'(y) z\bigg|_t = D_{(\tilde{u},\tilde{y},u,y)} g_i^{(j)}(t,u_t,y_t,u,y)(v_t,z_t,v,z), \qquad (44)
$$

$$
\begin{aligned}
\frac{\mathrm{d}^j}{\mathrm{d}t^j}\big(g_i''(y_t)(z_t)^2 + g_i'(y_t) z_t^2\big)\bigg|_t &= D^2_{(\tilde{u},\tilde{y},u,y)^2} g_i^{(j)}(t,u_t,y_t,u,y)(v_t,z_t,v,z)^2 \\
&\quad + D_{(\tilde{y},y)} g_i^{(j)}(t,u_t,y_t,u,y)\big(z_t^2,z^2\big). \qquad (45)
\end{aligned}
$$

Proof The result follows from Lemmas 2.4, 2.5, 2.6, by induction on j. Observe in particular that by Definition 2.2, formulas (43)–(45) depend neither on u_t nor on v_t for $j = 1, \dots, q_i - 1$. □

Example 2.1

1. The classical example of a state constraint of order q is $y_t \le 0$ for all $t \in [0,T]$ where $y_t^{(q)} = u_t$ for a.a. $t \in (0,T)$. This higher-order controlled differential equation can be written as a system of controlled differential equations, and the notion of order of state constraints for ODEs applies. It is interesting to note that this equation can be reduced to the following scalar integral equation:

$$y_t = \int_0^t \frac{(t-s)^{q-1}}{(q-1)!} u_s \, \mathrm{d}s, \qquad t \in [0,T]. \tag{46}$$

   Then the constraint $y_t \le 0$ for all $t \in [0,T]$ gives

$$g^{(0)}(t,\tilde{u},\tilde{y},u,y) = \tilde{y}, \tag{47}$$

$$g^{(j)}(t,\tilde{u},\tilde{y},u,y) = \int_0^t \frac{(t-s)^{q-1-j}}{(q-1-j)!} u_s \, \mathrm{d}s, \qquad j = 1,\dots,q-1, \tag{48}$$

$$g^{(q)}(t,\tilde{u},\tilde{y},u,y) = \tilde{u}. \tag{49}$$

   Thus, we find again that the constraint is of order q.
2. We consider the following variant of the previous example:

$$y_t = \int_0^t \frac{(t-s)^{q-1}}{(q-1)!} f(t,s) u_s \, \mathrm{d}s, \qquad t \in [0,T]. \tag{50}$$

   If f is not polynomial in t, then this integral equation cannot be in general reduced to a system of ODEs (see Remark 2.1.1), and the constraint $y_t \le 0$ for all $t \in [0,T]$ is still of order q.

3 Weak Results

3.1 A First Abstract Formulation

The optimal control problem (P) can be rewritten as an abstract optimization problem on (u, y_0). The most naive way to do that is the following equivalent formulation:

$$(P) \qquad \min_{(u,y_0)\in \mathcal{U}\times\mathbb{R}^n} J(u,y_0) \tag{51}$$

$$\text{subject to} \quad g\big(y[u,y_0]\big) \in C\big([0,T];\mathbb{R}^r_-\big), \tag{52}$$

$$\Phi\big(y_0, y[u,y_0]_T\big) \in K, \tag{53}$$

where

$$J(u,y_0) := \int_0^T \ell\big(u_t, y[u,y_0]_t\big)\,\mathrm{d}t + \phi\big(y_0, y[u,y_0]_T\big), \tag{54}$$

and $\Phi=(\Phi^E,\Phi^I)$, $K=\{0\}_{s_E}\times\mathbb{R}^{s_I}_{-}$. In order to write optimality conditions for this problem, we first compute its Lagrangian

$$L(u,y_0,\mathrm{d}\eta,\Psi):=J(u,y_0)+\int_{[0,T]}\mathrm{d}\eta_t\,g\big(y[u,y_0]_t\big)+\Psi\Phi\big(y_0,y[u,y_0]_T\big),\qquad(55)$$

where $(u,y_0,\mathrm{d}\eta,\Psi)\in\mathcal{U}\times\mathbb{R}^n\times\mathcal{M}\times\mathbb{R}^{s*}$ (see the beginning of Sect. 2.2). A *Lagrange multiplier* at (u,y_0) in this setting is any $(\mathrm{d}\eta,\Psi)$ such that

$$D_{(u,y_0)}L(u,y_0,\mathrm{d}\eta,\Psi)\equiv 0,\qquad(56)$$

$$(\mathrm{d}\eta,\Psi)\in N_{C([0,T];\mathbb{R}^r_-)\times K}\big(g(y),\Phi(y_0,y_T)\big),\qquad(57)$$

where $N_{C([0,T];\mathbb{R}^r_-)\times K}(g(y),\Phi(y_0,y_T))$ is the normal cone to $C([0,T];\mathbb{R}^r_-)\times K$ at $(g(y),\Phi(y_0,y_T))$. We have the following characterization:

$$(\mathrm{d}\eta,\Psi)\in N_{C([0,T];\mathbb{R}^r_-)\times K}\big(g(y),\Phi(y_0,y_T)\big)\qquad(58)$$

iff

$$\begin{aligned}&g_i(y)\le 0,\quad \mathrm{d}\eta_i\ge 0,\quad \int_0^T g_i(y_t)\,\mathrm{d}\eta_{i,t}=0,\quad i=1,\dots,r,\\&\Psi\in N_K\big(\Phi(y_0,y_T)\big)\quad\text{(see (9))}.\end{aligned}\qquad(59)$$

Definition (56)–(57) has to be compared to Definition 2.1:

Lemma 3.1 *The couple* $(\mathrm{d}\eta,\Psi)$ *is a Lagrange multiplier of the abstract problem* (51)–(53) *at* $(\bar u,\bar y_0)$ *iff* $(\mathrm{d}\eta,\Psi,p)$ *is a Lagrange multiplier of the optimal control problem* (1)–(5) *associated with* $(\bar u,y[\bar u,\bar y_0])$, *where* p *is the unique solution of* (11).

Proof Using the Hamiltonian (7) and the end-point Lagrangian (8), we have

$$\begin{aligned}L(u,y_0,\mathrm{d}\eta,\Psi)&=\int_0^T H[p](t,u_t,y_t)\,\mathrm{d}t+\int_{[0,T]}\mathrm{d}\eta_t\,g(y_t)+\Phi[\Psi](y_0,y_T)\\&\quad-\int_0^T\bigg(p_t f(t,t,u_t,y_t)+\int_t^T p_s D_\tau f(s,t,u_t,y_t)\,\mathrm{d}s\bigg)\mathrm{d}t\end{aligned}\qquad(60)$$

for $y=y[u,y_0]$ and any $p\in\mathcal{P}$. Moreover,

$$\begin{aligned}&\int_0^T\bigg(p_t f(t,t,u_t,y_t)+\int_t^T p_s D_\tau f(s,t,u_t,y_t)\,\mathrm{d}s\bigg)\mathrm{d}t\\&\quad=\int_0^T p_t\bigg(f(t,t,u_t,y_t)+\int_0^t D_\tau f(t,s,u_s,y_s)\,\mathrm{d}s\bigg)\mathrm{d}t\\&\quad=\int_{[0,T]}p_t\dot y_t\,\mathrm{d}t=-\int_{[0,T]}\mathrm{d}p_t\,y_t+p_{T_+}y_T-p_{0_-}y_0\end{aligned}\qquad(61)$$

by the formula of integration by parts (156) of the Appendix A.1. Then

$$L(u, y_0, \mathrm{d}\eta, \Psi) = \int_0^T H[p](t, u_t, y_t)\,\mathrm{d}t + \int_{[0,T]} \big(\mathrm{d}p_t y_t + \mathrm{d}\eta_t\, g(y_t)\big) + p_{0_-} y_0 - p_{T_+} y_T + \Phi[\Psi](y_0, y_T)$$

for any $p \in P$. We fix $(\bar{u}, \bar{y}_0, \mathrm{d}\eta, \Psi)$, differentiate L w.r.t. (u, y_0) at this point, and choose p as the unique solution of (11). Then

$$D_{(u,y_0)} L(\bar{u}, \bar{y}_0, \mathrm{d}\eta, \Psi)(v, z_0) = \int_0^T D_u H[p](t, \bar{u}_t, \bar{y}_t) v_t\,\mathrm{d}t + \big(p_{0_-} + D_{y_1}\Phi[\Psi](\bar{y}_0, \bar{y}_T)\big) z_0$$

for all $(v, z_0) \in U \times \mathbb{R}^n$. It follows that (56) is equivalent to (12) and (15). It is obvious that (57) is equivalent to (13)–(14). □

Second, we need a qualification condition.

Definition 3.1 We say that $(\bar{u}, \bar{y})$ is *qualified* iff

1. $\begin{cases} (v, z_0) \mapsto D\Phi^E(\bar{y}_0, \bar{y}_T)(z_0, z[v, z_0]_T) \\ U \times \mathbb{R}^n \to \mathbb{R}^{s_E} \end{cases}$ is onto,
2. there exists $(\bar{v}, \bar{z}_0) \in U \times \mathbb{R}^n$ such that, with $\bar{z} = z[\bar{v}, \bar{z}_0]$,

$$\begin{cases} D\Phi^E(\bar{y}_0, \bar{y}_T)(\bar{z}_0, \bar{z}_T) = 0, \\ D\Phi_i^I(\bar{y}_0, \bar{y}_T)(\bar{z}_0, \bar{z}_T) < 0, \quad i \in \{i : \Phi_i^I(\bar{y}_0, \bar{y}_T) = 0\}, \\ g_i'(\bar{y}_t)\bar{z}_t < 0 \quad \text{on } \{t : g_i(\bar{y}_t) = 0\},\ i = 1, \dots, r. \end{cases}$$

Remark 3.1

1. This condition is equivalent to Robinson's constraint qualification (introduced in [22], Definition 2) for the abstract problem (51)–(53) at $(\bar{u}, \bar{y}_0)$; see the discussion that follows Definition 3.4 and Definition 3.5 in [15] for a proof of the equivalence.
2. It is sometimes possible to give optimality conditions without qualification condition by considering an auxiliary optimization problem (see, e.g., the proof of Theorem 3.50 in [21]). Nevertheless, observe that if $(\bar{u}, \bar{y})$ is feasible but not qualified because (i) does not hold, then there exists a *singular Lagrange multiplier* of the form $(0, \Phi^E, 0)$. One can see that second-order necessary conditions become pointless since $-(0, \Phi^E, 0)$ is a singular Lagrange multiplier too. In this perspective, we only consider qualified solutions.

Finally, we derive the following first-order necessary optimality conditions:

Theorem 3.1 *Let $(\bar{u}, \bar{y})$ be a qualified local solution of (P). Then the set of associated Lagrange multipliers is nonempty, convex, and weakly $*$ compact.*

Proof Since the abstract problem (51)–(53) is qualified, we get the result for the set $\{(\mathrm{d}\eta, \Psi)\}$ of Lagrange multipliers in this setting (Theorem 4.1 in [23]). We conclude with Lemma 3.1 and the fact that

$$M \times \mathbb{R}^{s*} \longrightarrow M \times \mathbb{R}^{s*} \times P$$
$$(\mathrm{d}\eta, \Psi) \longmapsto (\mathrm{d}\eta, \Psi, p)$$

is affine continuous (it is obvious from the proof of Lemma 2.1). □

We will prove a stronger result in Sect. 4, relying on another abstract formulation, the so-called *reduced problem*. The main motivation for the reduced problem, as mentioned in the introduction, is actually to satisfy an *extended polyhedricity condition* (see Definition 3.52 in [21]), in order to easily get second-order necessary conditions (see Remark 3.47 in the same reference).

3.2 The Reduced Problem

In the sequel, *we fix a feasible trajectory* $(\bar{u}, \bar{y})$, i.e., which satisfies (2)–(5), and denote by Λ the set of associated Lagrange multipliers (Definition 2.1). We need some definitions.

Definition 3.2 An *arc* is a maximal interval, relatively open in $[0, T]$, denoted by $]\tau_1, \tau_2[$, such that the set of active running state constraints at time t is constant for all $t \in]\tau_1, \tau_2[$. It includes intervals of the form $[0, \tau[$ or $]\tau, T]$. If τ does not belong to any arc, we say that τ is a *junction time*.

Consider an arc $]\tau_1, \tau_2[$. It is a *boundary arc* for the constraint g_i if the latter is active on $]\tau_1, \tau_2[$; otherwise it is an *interior arc* for g_i.

Consider an interior arc $]\tau_1, \tau_2[$ for g_i. If $g_i(\tau_2) = 0$, then τ_2 is an *entry point* for g_i; if $g_i(\tau_1) = 0$, then τ_1 is an *exit point* for g_i. If τ is an entry point and an exit point, then it is a *touch point* for g_i.

Consider a touch point τ for g_i. We say that τ is *reducible* iff $\frac{\mathrm{d}^2}{\mathrm{d}t^2} g_i(\bar{y}_t)$, defined in a weak sense, is a function for t close to τ, continuous at τ, and

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2} g_i(\bar{y}_t)\bigg|_{t=\tau} < 0.$$

Remark 3.2 Let g_i be of order at least 2, and τ be a touch point for g_i. By Lemma 2.7, τ is reducible iff $t \mapsto g_i^{(2)}(t, \bar{u}_t, \bar{y}_t, \bar{u}, \bar{y})$ is continuous at τ and $g_i^{(2)}(\tau, \bar{u}_\tau, \bar{y}_\tau, \bar{u}, \bar{y}) < 0$. Note that the continuity holds if $\bar{u}$ is continuous at τ or if g_i is of order at least 3.

The interest of reducibility will appear with the next lemma. For $\tau \in [0, T]$, $\varepsilon > 0$ (to be fixed), and any function $x : [0, T] \to \mathbb{R}$, $x \in W^{2,\infty}$, we define $\mu_\tau(x)$ by

$$\mu_\tau(x) := \max\bigl\{x_t : t \in [\tau - \varepsilon, \tau + \varepsilon] \cap [0, T]\bigr\}. \tag{62}$$

Thus, we get a functional $\mu_\tau : W^{2,\infty}([0, T]) \to \mathbb{R}$.

Lemma 3.2 *Let g_i be of order at least* 2 (*i.e.*, $D_{\tilde{u}} g_i^{(1)} \equiv 0$), *and hence by Lemma* 2.7 $g_i(\bar{y}) \in W^{2,\infty}$. *Let τ be a reducible touch point for g_i. Then for $\varepsilon > 0$ small enough, μ_τ is C^1 in a neighborhood of $g_i(\bar{y})$ and twice Fréchet differentiable at $g_i(\bar{y})$, with first and second derivatives at $g_i(\bar{y})$ given by*

$$D\mu_\tau\big(g_i(\bar{y})\big)x = x_\tau, \tag{63}$$

$$D^2\mu_\tau\big(g_i(\bar{y})\big)(x)^2 = -\frac{(\frac{\mathrm{d}}{\mathrm{d}t} x_t|_\tau)^2}{\frac{\mathrm{d}^2}{\mathrm{d}t^2} g_i(\bar{y}_t)|_\tau}, \tag{64}$$

for any $x \in W^{2,\infty}([0,T])$.

Proof We apply Lemma 23 of [18] to $g_i(\bar{y})$ that belongs to $W^{2,\infty}([0,T])$ and satisfies the required hypotheses at τ by definition of a reducible touch point. □

Remark 3.3 We can write (63) and (64) for $x = g_i'(\bar{y})z[v, z_0]$ or $x = g_i''(\bar{y})(z[v, z_0])^2 + g_i'(\bar{y})z^2[v, z_0]$, $(v, z_0) \in U \times \mathbb{R}^n$, since by Lemma 2.7 they belong to $W^{2,\infty}([0,T])$. Moreover, we have

$$D^2\mu_\tau\big(g_i(\bar{y})\big)\big(g_i'(\bar{y})z\big)^2 = -\frac{(D_{(\tilde{y},u,y)} g_i^{(1)}(\tau, \bar{y}_\tau, \bar{u}, \bar{[}y])(z_\tau, v, z))^2}{g_i^{(2)}(\tau, \bar{u}_\tau, \bar{y}_\tau, \bar{u}, \bar{y})} \tag{65}$$

for $z = z[v, z_0]$, $(v, z_0) \in U \times \mathbb{R}^n$.

In view of these results, we distinguish running state constraints of order 1. Without loss of generality, we suppose that

- g_i is of order 1 for $i = 1, \dots, r_1$,
- g_i is of order at least 2 for $i = r_1 + 1, \dots, r$,

where $0 \le r_1 \le r$. We make now the following assumption:

(A1) There are finitely many junction times, and for $i = r_1 + 1, \dots, r$, all touch points for g_i are reducible.

For $i = 1, \dots, r_1$, we consider the contact sets of the constraints

$$I_i := \big\{t \in [0,T] : g_i(\bar{y}_t) = 0\big\}. \tag{66}$$

For $i = r_1 + 1, \dots, r$, we remove the touch points from the contact sets:

$$T_i := \text{the set of (reducible) touch points for } g_i, \tag{67}$$

$$I_i := \big\{t \in [0,T] : g_i(\bar{y}_t) = 0\big\} \setminus T_i. \tag{68}$$

For $i = 1, \dots, r$ and $\varepsilon \ge 0$, we denote

$$I_i^\varepsilon := \big\{t \in [0,T] : \mathrm{dist}(t, I_i) \le \varepsilon\big\}. \tag{69}$$

Assumption (A1) implies that I_i^ε has finitely many connected components for any $\varepsilon \geq 0$ ($1 \leq i \leq r$) and that T_i is finite ($r_1 < i \leq r$). Let $N := \sum_{r_1 < i \leq r} |T_i|$.

Now *we fix* $\varepsilon > 0$ small enough (so that Lemma 3.2 holds) and define

$$G_1(u, y_0) := \big(g_i\big(y[u, y_0]\big)\big|_{I_i^\varepsilon}\big)_{1 \leq i \leq r}, \qquad K_1 := \prod_{i=1}^{r} C\big(I_i^\varepsilon, \mathbb{R}_-\big), \tag{70}$$

$$G_2(u, y_0) := \big(\mu_\tau\big(g_i\big(y[u, y_0]\big)\big)\big)_{\tau \in T_i, r_1 < i \leq r}, \qquad K_2 := \mathbb{R}_-^N, \tag{71}$$

$$G_3(u, y_0) := \Phi\big(y_0, y[u, y_0]_T\big), \qquad K_3 := K. \tag{72}$$

Note that Lemma 3.2 does not enable us to consider touch points for constraints of order 1 in G_2, since we want the later to be regular enough. This is not a problem; we treat them with the boundary arcs in G_1, and we will see that an extended polyhedricity condition (Lemma 4.5) is satisfied.

Recall that J has been defined by (54); the *reduced problem* is the following abstract optimization problem:

$$(P_R) \qquad \min_{(u, y_0) \in U \times \mathbb{R}^n} J(u, y_0) \qquad \text{subject to} \quad \begin{cases} G_1(u, y_0) \in K_1, \\ G_2(u, y_0) \in K_2, \\ G_3(u, y_0) \in K_3. \end{cases}$$

Remark 3.4 We had fixed $(\bar{u}, \bar{y})$ as a feasible trajectory; then $(\bar{u}, \bar{y}_0)$ is feasible for (P_R). Moreover, $(\bar{u}, \bar{y})$ is a local solution of (P) iff $(\bar{u}, \bar{y}_0)$ is a local solution of (P_R), and the qualification condition at $(\bar{u}, \bar{y})$ (Definition 3.1) is equivalent to Robinson's constraints qualification for (P_R) at $(\bar{u}, \bar{y}_0)$ (using Lemma 3.2).

Thus, it is of interest for us to write optimality conditions for (P_R).

3.3 Optimality Conditions for the Reduced Problem

The *Lagrangian* of (P_R) is

$$\begin{aligned} &L_R(u, y_0, \mathrm{d}\rho, \nu, \Psi) \\ &\quad := J(u, y_0) + \sum_{1 \leq i \leq r} \int_{I_i^\varepsilon} g_i\big(y[u, y_0]_t\big)\, \mathrm{d}\rho_{i,t} \\ &\qquad + \sum_{\substack{\tau \in T_i \\ r_1 < i \leq r}} \nu_{i,\tau} \mu_\tau\big(g_i\big(y[u, y_0]\big)\big) + \Psi \Phi\big(y_0, y[u, y_0]_T\big), \end{aligned} \tag{73}$$

where $u \in U$, $y_0 \in \mathbb{R}^n$, $\mathrm{d}\rho \in \prod_{i=1}^r M(I_i^\varepsilon)$, $\nu \in \mathbb{R}^{N*}$, $\Psi \in \mathbb{R}^{s*}$.

As before, a measure on a closed interval is denoted by $\mathrm{d}\mu$ and is identified with the derivative of a function of bounded variation, which is null on the right of the interval.

A *Lagrange multiplier* of (P_R) at $(\bar{u}, \bar{y}_0)$ is any $(\mathrm{d}\rho, \nu, \Psi)$ such that

$$D_{(u,y_0)} L_R(\bar{u}, \bar{y}_0, \mathrm{d}\rho, \nu, \Psi) = 0, \tag{74}$$

$$\mathrm{d}\rho_i \geq 0, \quad g_i(\bar{y})|_{I_i^\varepsilon} \leq 0, \quad \int_{I_i^\varepsilon} g_i(\bar{y}_t)\,\mathrm{d}\rho_{i,t} = 0, \quad i = 1, \ldots, r, \tag{75}$$

$$\nu_{i,\tau} \geq 0, \quad \mu_\tau\big(g_i(\bar{y})\big) \leq 0, \quad \nu_{i,\tau}\mu_\tau\big(g_i(\bar{y})\big) = 0, \quad \tau \in T_i,\ i = r_1 + 1, \ldots, r, \tag{76}$$

$$\Psi \in N_K\big(\Phi(\bar{y}_0, \bar{y}_T)\big). \tag{77}$$

We denote by Λ_R the set of Lagrange multipliers of (P_R) at $(\bar{u}, \bar{y}_0)$. The first-order necessary conditions for (P_R) are the same as in Theorem 3.1:

Lemma 3.3 *Let $(\bar{u}, \bar{y}_0)$ be a qualified local solution of (P_R). Then Λ_R is nonempty, convex, and weakly $*$ compact.*

Given $(\mathrm{d}\rho, \nu) \in \prod_{i=1}^r M(I_i^\varepsilon) \times \mathbb{R}^{N*}$, we define $\mathrm{d}\eta \in M$ by

$$\mathrm{d}\eta_i := \begin{cases} \mathrm{d}\rho_i & \text{on } I_i^\varepsilon,\ i = 1, \ldots, r, \\ \sum_{\tau \in T_i} \nu_{i,\tau}\delta_\tau & \text{elsewhere},\ i = r_1 + 1, \ldots, r. \end{cases} \tag{78}$$

Conversely, given $\mathrm{d}\eta \in M$, we define $(\mathrm{d}\rho, \nu) \in \prod_{i=1}^r M(I_i^\varepsilon) \times \mathbb{R}^{N*}$ by

$$\begin{cases} \mathrm{d}\rho_i := \mathrm{d}\eta_i|_{I_i^\varepsilon} & i = 1, \ldots, r, \\ \nu_{i,\tau} := \mathrm{d}\eta_i\big(\{\tau\}\big) & \tau \in T_i,\ i = r_1 + 1, \ldots, r. \end{cases} \tag{79}$$

In the sequel we use these definitions to identify $(\mathrm{d}\rho, \nu)$ and $\mathrm{d}\eta$, and we denote

$$[\eta_{i,\tau}] := \mathrm{d}\eta_i\big(\{\tau\}\big). \tag{80}$$

Recall that Λ is the set of Lagrange multipliers associated with $(\bar{u}, \bar{y})$ (Definition 2.1). We have a result similar to Lemma 3.1:

Lemma 3.4 *The triple $(\mathrm{d}\rho, \nu, \Psi) \in \Lambda_R$ iff $(\mathrm{d}\eta, \Psi, p) \in \Lambda$, with p the unique solution of* (11).

Proof With the identification between $(\mathrm{d}\rho, \nu)$ and $\mathrm{d}\eta$ given by (78) and (79), it is clear that (75)–(76) are equivalent to (13). Let these relations be satisfied by $(\mathrm{d}\rho, \nu, \Psi)$ and $(\mathrm{d}\eta, \Psi)$. Then, in particular,

$$\begin{aligned} &\operatorname{supp}(\mathrm{d}\eta_i) = \operatorname{supp}(\mathrm{d}\rho_i) \subset I_i, \quad i = 1, \ldots, r_1, \\ &\operatorname{supp}(\mathrm{d}\eta_i) = \operatorname{supp}(\mathrm{d}\rho_i) \cup \operatorname{supp}\Big(\sum \nu_{i,\tau}\delta_\tau\Big) \subset I_i \cup T_i, \quad i = r_1 + 1, \ldots, r. \end{aligned} \tag{81}$$

We claim that in this case (74) is equivalent to (12) and (15). Indeed, as in the proof of Lemma 3.1, we have

$$
\begin{aligned}
L_R(u, y_0, \mathrm{d}\rho, \nu, \Psi) = {} & \int_{[0,T]} \big(H[p](t, u_t, y_t)\,\mathrm{d}t + \mathrm{d}p_t y_t\big) + p_{0_-} y_0 - p_{T_+} y_T \\
& + \sum_{1 \le i \le r} \int_{I_i} g_i(y_t)\,\mathrm{d}\eta_{i,t} + \sum_{\substack{\tau \in T_i \\ r_1 < i \le r}} [\eta_{i,\tau}] \mu_\tau\big(g_i(y)\big) \\
& + \Phi[\Psi](y_0, y_T)
\end{aligned} \tag{82}
$$

for any $p \in P$ and $y = y[u, y_0]$. Let us differentiate (say, for $i > r_1$)

$$
\int_{I_i} g_i(y_t)\,\mathrm{d}\eta_{i,t} + \sum_{\tau \in T_i} [\eta_{i,\tau}] \mu_\tau\big(g_i(y)\big) \tag{83}
$$

w.r.t. (u, y_0) at $(\bar{u}, \bar{y}_0)$ in the direction (v, z_0) and use (63), (80), (81); we get

$$
\int_{I_i} g_i'(\bar{y}_t) z_t\,\mathrm{d}\eta_{i,t} + \sum_{\tau \in T_i} [\eta_{i,\tau}] D\mu_\tau\big(g_i(\bar{y})\big)\big(g_i'(\bar{y}) z\big) = \int_{[0,T]} g_i'(\bar{y}_t) z_t\,\mathrm{d}\eta_{i,t}, \tag{84}
$$

where $z = z[v, z_0]$. Let us now differentiate similarly the whole expression (82) of L_R; we get

$$
\begin{aligned}
& \int_0^T D_u H[p](t, \bar{u}_t, \bar{y}_t) v_t\,\mathrm{d}t + \int_{[0,T]} \big(D_y H[p](t, \bar{u}_t, \bar{y}_t)\,\mathrm{d}t + \mathrm{d}p_t + \mathrm{d}\eta_t g'(\bar{y}_t)\big) z_t \\
& \quad + \big(p_{0_-} + D_{y_1} \Phi[\Psi](\bar{y}_0, \bar{y}_T)\big) z_0 + \big(-p_{T_+} + D_{y_2} \Phi[\Psi](\bar{y}_0, \bar{y}_T)\big) z_T.
\end{aligned} \tag{85}
$$

Fixing p as the unique solution of (11) in (85) gives

$$
\begin{aligned}
D_{(u,y_0)} L_R(\bar{u}, \bar{y}_0, \mathrm{d}\rho, \nu, \Psi)(v, z_0) = {} & \int_0^T D_u H[p](t, \bar{u}_t, \bar{y}_t) v_t\,\mathrm{d}t \\
& + \big(p_{0_-} + D_{y_1} \Phi[\Psi](\bar{y}_0, \bar{y}_T)\big) z_0.
\end{aligned}
$$

It is now clear that (74) is equivalent to (12) and (15). □

For the second-order optimality conditions, we need to evaluate the Hessian of L_R. For $\lambda = (\mathrm{d}\eta, \Psi, p) \in \Lambda$, $(v, z_0) \in U \times \mathbb{R}^n$, and $z = z[v, z_0] \in Y$, we denote

$$
\begin{aligned}
J[\lambda](v, z_0) := {} & \int_0^T D^2_{(u,y)^2} H[p](t, \bar{u}_t, \bar{y}_t)(v_t, z_t)^2\,\mathrm{d}t + D^2 \Phi[\Psi](\bar{y}_0, \bar{y}_T)(z_0, z_T)^2 \\
& + \sum_{1 \le i \le r} \int_{I_i} g_i''(\bar{y}_t)(z_t)^2\,\mathrm{d}\eta_{i,t} \\
& + \sum_{\substack{\tau \in T_i \\ r_1 < i \le r}} [\eta_{i,\tau}] \big[g_i''(\bar{y}_\tau)(z_\tau)^2 + D^2 \mu_\tau\big(g_i(\bar{y})\big)\big(g_i'(\bar{y}) z\big)^2\big].
\end{aligned} \tag{86}
$$

In view of (65) and (81), we can also write

$$\begin{aligned} J[\lambda](v, z_0) = &\int_0^T D^2_{(u,y)^2} H[p](t, \bar{u}_t, \bar{y}_t)(v_t, z_t)^2 \,\mathrm{d}t + D^2\Phi[\Psi](\bar{y}_0, \bar{y}_T)(z_0, z_T)^2 \\ &+ \int_{[0,T]} \mathrm{d}\eta_t\, g''(\bar{y}_t)(z_t)^2 \\ &- \sum_{\substack{\tau\in T_i \\ r_1<i\le r}} [\eta_{i,\tau}] \frac{(D_{(\tilde{y},u,y)} g_i^{(1)}(\tau, \bar{y}_\tau, \bar{u}, \bar{y})(z_\tau, v, z))^2}{g_i^{(2)}(\tau, \bar{u}_\tau, \bar{y}_\tau, \bar{u}, \bar{y})}. \end{aligned} \tag{87}$$

Lemma 3.5 *Let* $(\mathrm{d}\rho, \nu, \Psi) \in \Lambda_R$. *Let* $(\mathrm{d}\eta, \Psi, p) \in \Lambda$ *be given by Lemma 3.4 and denoted by* λ. *Then, for all* $(v, z_0) \in U \times \mathbb{R}^n$,

$$D^2_{(u,y_0)^2} L_R(\bar{u}, \bar{y}_0, \mathrm{d}\rho, \nu, \Psi)(v, z_0)^2 = J[\lambda](v, z_0). \tag{88}$$

Proof We will use (82) and (83) from the previous proof. First, we differentiate (83) twice w.r.t. (u, y_0) at $(\bar{u}, \bar{y}_0)$ in the direction (v, z_0). Let $z = z[v, z_0]$ and $z^2 = z^2[v, z_0]$, defined by (20); we get, for $i > r_1$,

$$\begin{aligned} &\int_{I_i} \big(g_i''(\bar{y}_t)(z_t)^2 + g_i'(\bar{y}_t) z_t^2\big)\, \mathrm{d}\eta_{i,t} \\ &\quad + \sum_{\tau\in T_i} [\eta_{i,\tau}]\big[D^2\mu_\tau\big(g_i(\bar{y})\big)\big(g_i'(\bar{y})z\big)^2 + D\mu_\tau\big(g_i(\bar{y})\big)\big(g_i''(\bar{y})(z)^2 + g_i'(\bar{y})z^2\big)\big] \\ &= \int_{I_i} g_i''(\bar{y}_t)(z_t)^2 \,\mathrm{d}\eta_{i,t} + \int_{[0,T]} g_i'(\bar{y}_t) z_t^2 \,\mathrm{d}\eta_{i,t} \\ &\quad + \sum_{\tau\in T_i} [\eta_{i,\tau}]\big[D^2\mu_\tau\big(g_i(\bar{y})\big)\big(g_i'(\bar{y})z\big)^2 + g_i''(\bar{y}_\tau)(z_\tau)^2\big], \end{aligned}$$

where we have used Remark 3.3, (63), and (81). Second, we differentiate L_R twice using (82), and then we fix p as the unique solution of (11). The result follows as in the proof of Lemma 3.4. □

Suppose that $\Lambda \ne \emptyset$ and let $\bar{\lambda} = (\mathrm{d}\bar{\eta}, \bar{\Psi}, \bar{p}) \in \Lambda$. We define the *critical* L^2 *cone* as the set C_2 of $(v, z_0) \in V_2 \times \mathbb{R}^n$ such that

$$\begin{cases} g_i'(\bar{y})z \le 0 & \text{on } I_i, \\ g_i'(\bar{y})z = 0 & \text{on } \operatorname{supp}(\mathrm{d}\bar{\eta}_i) \cap I_i, \end{cases} \quad i = 1, \dots, r, \tag{89}$$

$$\begin{cases} g_i'(\bar{y}_\tau) z_\tau \le 0, \\ [\bar{\eta}_{i,\tau}] g_i'(\bar{y}_\tau) z_\tau = 0, \end{cases} \quad \tau \in T_i,\ i = r_1 + 1, \dots, r, \tag{90}$$

$$\begin{cases} D\Phi(\bar{y}_0, \bar{y}_T)(z_0, z_T) \in T_K\big(\Phi(\bar{y}_0, \bar{y}_T)\big), \\ \bar{\Psi} D\Phi(\bar{y}_0, \bar{y}_T)(z_0, z_T) = 0, \end{cases} \tag{91}$$

where $z = z[v, z_0] \in Z_2$. Then the *critical cone* for (P_R) (see Proposition 3.10 in [21]) is the set

$$C_\infty := C_2 \cap \big(U \times \mathbb{R}^n\big),$$

and the *cone of radial critical directions* for (P_R) (see Definition 3.52 in [21]) is the set

$$C_\infty^R := \big\{(v, z_0) \in C_\infty : \exists \bar{\sigma} > 0 : g_i(\bar{y}) + \bar{\sigma} g_i'(\bar{y})z \le 0 \text{ on } I_i^\varepsilon,\ i = 1, \dots, r\big\},$$

where $z = z[v, z_0] \in Y$. These three cones do not depend on the choice of $\bar{\lambda}$.

In view of Lemma 3.5, the second-order necessary conditions for (P_R) can be written as follows:

Lemma 3.6 *Let $(\bar{u}, \bar{y}_0)$ be a qualified local solution of (P_R). Then, for any $(v, z_0) \in C_\infty^R$, there exists $\lambda \in \Lambda$ such that*

$$J[\lambda](v, z_0) \ge 0. \tag{92}$$

Proof Corollary 5.1 in [15]. □

4 Strong Results

Recall that $(\bar{u}, \bar{y})$ is a feasible trajectory that has been fixed to define the reduced problem at the beginning of Sect. 3.2.

4.1 Extra Assumptions and Consequences

We were so far under assumptions (A0)–(A1). We make now some extra assumptions, which will imply a partial qualification of the running state constraints, as well as the density of C_∞^R in a larger critical cone.

(A2) Each running state constraint g_i, $i = 1, \dots, r$, is of finite order q_i.

Notations Given a subset $J \subset \{1, \dots, r\}$, say $J = \{i_1 < \cdots < i_l\}$, we define

$$G_J^{(q)} : \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^n \times U \times Y \to \mathbb{R}^{|J|} \quad \text{by}$$

$$G_J^{(q)}(t, \tilde{u}, \tilde{y}, u, y) := \begin{pmatrix} \bar{g}_{i_1}^{(q_{i_1})}(t, \tilde{u}, \tilde{y}, u, y) \\ \vdots \\ \bar{g}_{i_l}^{(q_{i_l})}(t, \tilde{u}, \tilde{y}, u, y)T \end{pmatrix}. \tag{93}$$

For $\varepsilon_0 \ge 0$ and $t \in [0, T]$, let

$$I_t^{\varepsilon_0} := \big\{1 \le i \le r : t \in I_i^{\varepsilon_0}\big\}, \tag{94}$$

$$M_t^{\varepsilon_0} := D_{\tilde{u}} G_{I_t^{\varepsilon_0}}^{(q)}(t, \bar{u}_t, \bar{y}_t, \bar{u}, \bar{y}) \in M_{|I_t^{\varepsilon_0}|, m}(\mathbb{R}). \tag{95}$$

(A3) There exist $\varepsilon_0, \gamma > 0$ such that, for all $t \in [0, T]$,

$$\left|\left(M_t^{\varepsilon_0}\right)^* \xi\right| \geq \gamma |\xi| \quad \forall \xi \in \mathbb{R}^{|I_t^{\varepsilon_0}|}. \tag{96}$$

(A4) The initial condition satisfies $g(\bar{y}_0) < 0$, and the final time T is not an entry point (i.e., there exists $\tau < T$ such that the set I_t^0 of active constraints at time t is constant for $t \in]\tau, T]$).

Remark 4.1

1. We do not assume that $\bar{u}$ is continuous, as was done in [20].
2. Assumption (A3) says that $M_t^{\varepsilon_0}$ is onto, uniformly w.r.t. t. Note that each constraint is considered only in a neighborhood of its contact set. Note also that in the case of one running state constraint ($r = 1$) of order q and if $\bar{u}$ is continuous, assumption (A3) is equivalent to

$$\frac{\partial g^{(q)}}{\partial \tilde{u}}(t, \bar{u}_t, \bar{y}_t, \bar{u}, \bar{y}) \neq 0 \quad \forall t \in I. \tag{97}$$

   See Appendix A.2 for Example A.1, where this assumption is discussed.
3. Recall that ε has been fixed to define the reduced problem. Without loss of generality, we suppose that $2\varepsilon_0 < \min\{|\tau - \tau'| : \tau, \tau' \text{ distinct junction times}\}$ and $\varepsilon < \varepsilon_0 < \min\{\tau : \tau \text{ junction times}\}$. We omit it in the notation $M_t^{\varepsilon_0}$.
4. In some cases, we can treat the case where T is an entry point, say for the constraint g_i:

   - if $1 \leq i \leq r_1$ (i.e., if $q_i = 1$), then what follows works similarly.
   - if $r_1 < i \leq r$ (i.e., if $q_i > 1$) and $\frac{\mathrm{d}}{\mathrm{d}t} g_i(\bar{y}_t)|_{t=T} > 0$, then we can replace in the reduced problem the running state constraint $g_i(y[u, y_0])|_{[T-\varepsilon, T]} \leq 0$ by the final state constraint $g_i(y[u, y_0]_T) \leq 0$.

5. By assumption (A1) we can write

$$[0, T] = J_0 \cup \cdots \cup J_\kappa, \tag{98}$$

   where J_l ($l = 0, \ldots, \kappa$) are the maximal intervals in $[0, T]$ such that $I_t^{\varepsilon_0}$ is constant (say equal to I_l) for $t \in J_l$. We order $J_0, \ldots, J_\kappa$ in $[0, T]$. Observe that for any $l \geq 1$, $\overline{J_{l-1}} \cap \overline{J_l} = \{\tau \pm \varepsilon_0\}$ with τ a junction time.

For $s \in [1, \infty]$, we denote

$$W^{(q),s}([0, T]) := \prod_{i=1}^{r} W^{q_i,s}([0, T]), \qquad W^{(q),s}(I^\varepsilon) := \prod_{i=1}^{r} W^{q_i,s}(I_i^\varepsilon), \tag{99}$$

and for

$$\varphi = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_r \end{pmatrix} \in W^{(q),s}([0, T]), \qquad \varphi|_{I^\varepsilon} := \begin{pmatrix} \varphi_1|_{I_1^\varepsilon} \\ \vdots \\ \varphi_r|_{I_r^\varepsilon} \in W^{(q),s}(I^\varepsilon). \end{pmatrix}$$

Using Lemma 2.7, we define, for $s \in [1, \infty]$ and $z_0 \in \mathbb{R}^n$,

$$\begin{aligned} A_{s,z_0} \colon V_s &\longrightarrow W^{(q),s}\big([0,T]\big) \\ v &\longmapsto g'(\bar{y})z[v, z_0]. \end{aligned} \tag{100}$$

We give now the statement of a lemma in two parts, which will be of great interest for us (particularly in Sect. 4.3.3). The proof is technical and can be skipped at a first reading. It is given in the next section.

Lemma 4.1

(a) *Let* $s \in [1, \infty]$ *and* $z_0 \in \mathbb{R}^n$. *Let* $\bar{b} \in W^{(q),s}(I^\varepsilon)$. *Then there exists* $v \in V_s$ *such that*

$$(A_{s,z_0} v)|_{I^\varepsilon} = \bar{b}. \tag{101}$$

(b) *Let* $z_0 \in \mathbb{R}^n$. *Let* $(\bar{b}, \bar{v}) \in W^{(q),2}(I^\varepsilon) \times V_2$ *be such that*

$$(A_{2,z_0} \bar{v})|_{I^\varepsilon} = \bar{b}. \tag{102}$$

Let $b^k \in W^{(q),\infty}(I^\varepsilon)$, $k \in \mathbb{N}$, *be such that* $b^k \xrightarrow{W^{(q),2}(I^\varepsilon)} \bar{b}$. *Then there exist* $v^k \in U$, $k \in \mathbb{N}$, *such that* $v^k \xrightarrow{L^2} \bar{v}$ *and*

$$\big(A_{\infty,z_0} v^k\big)|_{I^\varepsilon} = b^k. \tag{103}$$

4.2 A Technical Proof

In this section we prove Lemma 4.1. The proofs of (a) and (b) are very similar; in both cases we proceed in $\kappa + 1$ steps using the decomposition (98) of $[0, T]$. At each step, we will use the following two lemmas, proved in Appendixes A.3 and A.2, respectively.

The first one uses only assumption (A1) and the definitions that follow.

Lemma 4.2 *Let* $t_0 := \tau \pm \varepsilon_0$ *where* τ *is a junction time.*

(a) *Let* $s \in [1, \infty]$ *and* $z_0 \in \mathbb{R}^n$. *Let* $(\bar{b}, v) \in W^{(q),s}(I^\varepsilon) \times V_s$ *be such that*

$$(A_{s,z_0} v)|_{I^\varepsilon} = \bar{b} \quad \text{on } [0, t_0]. \tag{104}$$

Then we can extend $\bar{b}$ *to* $\tilde{b} \in W^{(q),s}([0, T])$ *in such a way that*

$$\tilde{b} = A_{s,z_0} v \quad \text{on } [0, t_0]. \tag{105}$$

(b) *Let* $z_0 \in \mathbb{R}^n$. *Let* $(\bar{b}, \bar{v}) \in W^{(q),2}(I^\varepsilon) \times V_2$ *be such that*

$$(A_{2,z_0} \bar{v})|_{I^\varepsilon} = \bar{b}. \tag{106}$$

Let $(b^k, v^k) \in W^{(q),\infty}(I^\varepsilon) \times U$, $k \in \mathbb{N}$, *be such that*

$$\big(b^k, v^k\big) \xrightarrow{W^{(q),2} \times L^2} (\bar{b}, \bar{v}), \quad \text{and} \tag{107}$$

$$\big(A_{\infty,z_0} v^k\big)|_{I^\varepsilon} = b^k \quad \text{on } [0, t_0]. \tag{108}$$

Then we can extend b^k *to* $\tilde{b}^k \in W^{(q),\infty}([0,T])$, $k \in \mathbb{N}$, *in such a way that*

$$\tilde{b}^k \xrightarrow{W^{(q),2}([0,T])} A_{2,z_0}\bar{v}, \quad \textit{and} \tag{109}$$

$$\tilde{b}^k = A_{\infty,z_0} v^k \quad \textit{on } [0,t_0]. \tag{110}$$

The second lemma relies on assumption (A3).

Lemma 4.3 *Let* $s \in [1,\infty]$ *and* $z_0 \in \mathbb{R}^n$. *Let* l *be such that* $I_l \neq \emptyset$. *For* $t \in J_l$, *we denote*

$$\begin{cases} M_t := D_{\tilde{u}} G_{I_l}^{(q)}(t,\bar{u}_t,\bar{y}_t,\bar{u},\bar{y}) \in L(\mathbb{R}^m, \mathbb{R}^{|I_l|}), \\ N_t := D_{(\tilde{y},u,y)} G_{I_l}^{(q)}(t,\bar{u}_t,\bar{y}_t,\bar{u},\bar{y}) \in L(\mathbb{R}^{n*} \times U^* \times Y^*, \mathbb{R}^{|I_l|}). \end{cases} \tag{111}$$

(a) *Let* $(\bar{h}, v) \in L^s(J_l;\mathbb{R}^{|I_l|}) \times V_s$. *Then there exists* $\tilde{v} \in V_s$ *such that*

$$\begin{cases} \tilde{v} = v \quad \textit{on } J_0 \cup \cdots \cup J_{l-1}, \\ M_t\tilde{v}_t + N_t(z[\tilde{v},z_0]_t, \tilde{v}, z[\tilde{v},z_0]) = \bar{h}_t \quad \textit{for a.a. } t \in J_l. \end{cases} \tag{112}$$

(b) *Let* $(\bar{h}, \bar{v}) \in L^s(J_l;\mathbb{R}^{|I_l|}) \times V_s$ *be such that*

$$M_t\bar{v}_t + N_t(z[\bar{v},z_0]_t, \bar{v}, z[\bar{v},z_0]) = \bar{h}_t \quad \textit{for a.a. } t \in J_l. \tag{113}$$

Let $(h^k, v^k) \in L^\infty(J_l;\mathbb{R}^{|I_l|}) \times U$, $k \in \mathbb{N}$, *be such that* $(h^k, v^k) \xrightarrow{L^s \times L^s} (\bar{h}, \bar{v})$. *Then there exists* $\tilde{v}^k \in U$, $k \in \mathbb{N}$, *such that* $\tilde{v}^k \xrightarrow{L^s} \bar{v}$ *and*

$$\begin{cases} \tilde{v}^k = v^k \quad \textit{on } J_0 \cup \cdots \cup J_{l-1}, \\ M_t\tilde{v}^k_t + N_t(z[\tilde{v}^k,z_0]_t, \tilde{v}^k, z[\tilde{v}^k,z_0]) = h^k_t \quad \textit{for a.a. } t \in J_l. \end{cases} \tag{114}$$

Proof of Lemma 4.1 In the sequel we omit z_0 in the notations.

(a) Let $\bar{b} \in W^{(q),s}(I^\varepsilon)$. We need to find $v \in V_s$ such that

$$g_i'(\bar{y})z[v] = \bar{b}_i \quad \text{on } I_i^\varepsilon,\ i = 1,\ldots,r. \tag{115}$$

Since

$$v = v' \quad \text{on } [0,t] \quad \Longrightarrow \quad z[v] = z[v'] \quad \text{on } [0,t],$$

let us construct $v^0,\ldots,v^\kappa \in V_s$ such that, for all l,

$$\begin{cases} v^l = v^{l-1} \quad \text{on } J_0 \cup \cdots \cup J_{l-1}, \\ g_i'(\bar{y})z[v^l] = \bar{b}_i \quad \text{on } I_i^\varepsilon \cap J_l,\ i = 1,\ldots,r, \end{cases}$$

and $v := v^\kappa$ will satisfy (115).

By assumption (A4), $J_0 = [0, \tau_1 - \varepsilon_0[$, where τ_1 is the first junction time, and then $I_i^\varepsilon \cap J_0 = \emptyset$ for all i; we choose $v^0 := 0$.

Suppose that we have $v^0, \dots v^{l-1}$ for some $l \geq 1$ and let us construct v^l. We apply Lemma 4.2 (a) to $(\bar{b}, v^{l-1})$ with $\{t_0\} = \overline{J_{l-1}} \cap \overline{J_l}$, and we get $\tilde{b} \in W^{(q),s}([0,T])$. Since $I_i^\varepsilon \cap J_l = \emptyset$ if $i \notin I_l$, it is now enough to find v^l such that

$$\begin{cases} v^l = v^{l-1} & \text{on } J_0 \cup \dots \cup J_{l-1}, \\ g_i'(\bar{y})z[v^l] = \tilde{b}_i & \text{on } J_l,\ i \in I_l. \end{cases} \tag{116}$$

Suppose that $v^l = v^{l-1}$ on $J_0 \cup \dots \cup J_{l-1}$. Then $g_i'(\bar{y})z[v^l] = \tilde{b}_i$ on J_{l-1}, and it follows that

$$g_i'(\bar{y})z\bigl[v^l\bigr] = \tilde{b}_i \quad \text{on } J_l \tag{117}$$

$$\Updownarrow$$

$$\frac{\mathrm{d}^{q_i}}{\mathrm{d}t^{q_i}} g_i'(\bar{y})z\bigl[v^l\bigr] = \frac{\mathrm{d}^{q_i}}{\mathrm{d}t^{q_i}} \tilde{b}_i \quad \text{on } J_l. \tag{118}$$

And by Lemma 2.7, (118) is equivalent to

$$D_{\tilde{u}} g_i^{(q_i)}(t, \bar{u}_t, \bar{y}_t, \bar{u}, \bar{y}) v_t^l + D_{(\tilde{y},u,y)} g_i^{(q_i)}(t, \bar{u}_t, \bar{y}_t, \bar{u}, \bar{y}) \bigl(z\bigl[v^l\bigr]_t, v^l, z\bigl[v^l\bigr]\bigr) = \tilde{b}_i^{(q_i)}(t) \tag{119}$$

for a.a. $t \in J_l$.

If $I_l = \emptyset$, we choose $v^l := v^{l-1}$. Otherwise, say $I_l = \{i_1 < \dots < i_p\}$ and define on J_l

$$\bar{h} := \begin{pmatrix} \tilde{b}_{i_1}^{(q_{i_1})} \\ \vdots \\ \tilde{b}_{i_p}^{(q_{i_p})} \end{pmatrix} \in L^s\bigl(J_l; \mathbb{R}^{|I_l|}\bigr).$$

Then (116) is equivalent to

$$\begin{cases} v^l = v^{l-1} & \text{on } J_0 \cup \dots \cup J_{l-1}, \\ M_t v_t^l + N_t\bigl(z\bigl[v^l\bigr]_t, v^l, z\bigl[v^l\bigr]\bigr) = \bar{h}_t & \text{for a.a. } t \in J_l. \end{cases} \tag{120}$$

Applying Lemma 4.3 (a) to (h, v^{l-1}), we get $\tilde{v}$ such that (120) holds; we choose $v^l := \tilde{v}$.

(b) We follow a similar scheme to the one of the proof of (a).

Let $(\bar{b}, \bar{v}) \in W^{(q),2}(I^\varepsilon) \times V_2$ be such that

$$g_i'(\bar{y})z[\bar{v}] = \bar{b}_i \quad \text{on } I^\varepsilon,\ i = 1, \dots, r.$$

Let $b^k \in W^{(q),\infty}(I^\varepsilon)$, $k \in \mathbb{N}$, be such that $b^k \xrightarrow{W^{(q),2}} \bar{b}$. Let us construct $v^{k,0}, \dots, v^{k,\kappa} \in U$, $k \in \mathbb{N}$, such that for all l, $v^{k,l} \xrightarrow[k\to\infty]{L^2} \bar{v}$, and

$$\begin{cases} v^{k,l} = v^{k,l-1} & \text{on } J_0 \cup \dots \cup J_{l-1}, \\ g_i'(\bar{y})z\bigl[v^{k,l}\bigr] = b_i^k & \text{on } I_i^\varepsilon \cap J_l,\ i \in I_l. \end{cases}$$

We will conclude the proof by defining $v^k := v^{k,\kappa}$, $k \in \mathbb{N}$.

We choose, for $v^{k,0}$, the truncation of $\bar{v}$, $k \in \mathbb{N}$ (see Definition A.1 in Appendix A.2).

Suppose that we have $v^{k,0}, \ldots, v^{k,l-1}$, $k \in \mathbb{N}$, for some $l \geq 1$ and let us construct $v^{k,l}$, $k \in \mathbb{N}$. We apply Lemma 4.2 (b) to $(b^k, v^{k,l-1})$ with $\{t_0\} = \overline{J_{l-1}} \cap \overline{J_l}$, and we get, for $k \in \mathbb{N}$, $\tilde{b}^k \in W^{(q),\infty}([0,T])$. In particular,

$$\tilde{b}^k \xrightarrow{W^{(q),2}} \tilde{b}, \tag{121}$$

where $\tilde{b} := g'(\bar{y})z[\bar{v}] \in W^{(q),2}([0,T])$. And it is now enough to find $v^{k,l}$, $k \in \mathbb{N}$, such that $v^{k,l} \xrightarrow[k\to\infty]{L^2} \bar{v}$ and

$$\begin{cases} v^{k,l} = v^{k,l-1} & \text{on } J_0 \cup \cdots \cup J_{l-1}, \\ g_i'(\bar{y})z\big[v^{k,l}\big] = \tilde{b}_i^k & \text{on } J_l,\ i \in I_l. \end{cases} \tag{122}$$

If $I_l = \emptyset$, we choose $v^{k,l} = v^{k,l-1}$, $k \in \mathbb{N}$. Otherwise, say $I_l = \{i_1 < \cdots < i_p\}$ and define on J_l

$$\bar{h} := \begin{pmatrix} \tilde{b}_{i_1}^{(q_{i_1})} \\ \vdots \\ \tilde{b}_{i_p}^{(q_{i_p})} \end{pmatrix} \in L^2\big(J_l; \mathbb{R}^{|I_l|}\big), \qquad h^k := \begin{pmatrix} (\tilde{b}_{i_1}^k)^{(q_{i_1})} \\ \vdots \\ (\tilde{b}_{i_p}^k)^{(q_{i_p})} \end{pmatrix} \in L^\infty\big(J_l; \mathbb{R}^{|I_l|}\big).$$

We have

$$M_t \bar{v}_t + N_t\big(z[\bar{v}]_t, \bar{v}, z[\bar{v}]\big) = \bar{h}_t \quad \text{for a.a. } t \in J_l,$$

and (122) is equivalent to

$$\begin{cases} v^{k,l} = v^{k,l-1} & \text{on } J_0 \cup \cdots \cup J_{l-1}, \\ M_t v_t^{k,l} + N_t\big(z\big[v^{k,l}\big]_t, v^{k,l}, z\big[v^{k,l}\big]\big) = h_t^k & \text{for a.a. } t \in J_l. \end{cases} \tag{123}$$

By (121), $h^k \xrightarrow{L^2} \bar{h}$, and by assumption, $v^{k,l-1} \xrightarrow[k\to\infty]{L^2} \bar{v}$. Applying Lemma 4.3 (b) to $(h^k, v^{k,l-1})$, we get $\tilde{v}^k$, $k \in \mathbb{N}$, such that $\tilde{v}^k \xrightarrow{L^2} \bar{v}$ and (123) holds; we choose $v^{k,l} = \tilde{v}^k$, $k \in \mathbb{N}$. □

4.3 Necessary Conditions

Recall that we are under assumptions (A0)–(A4).

4.3.1 Structure of the Set of Lagrange Multipliers

Recall that we denote by Λ the set of Lagrange multipliers associated with $(\bar{u}, \bar{y})$ (Definition 2.1). We consider the projection map

$$\begin{aligned} \pi : M \times \mathbb{R}^{s*} \times P &\longrightarrow \quad \mathbb{R}^{N*} \times \mathbb{R}^{s*} \\ (\mathrm{d}\eta, \Psi, p) &\longmapsto \big(\big([\eta_{i,\tau}]\big)_{\tau,i}, \Psi\big) \end{aligned}$$

where $\tau \in T_i$, $i = r_1 + 1, \ldots, r$. A consequence of Lemma 4.1 (a) is the following:

Lemma 4.4 $\pi|_\Lambda$ *is injective.*

Proof We will use the fact that one of the constraints, namely G_1, has a surjective derivative. For $\mathrm{d}\rho \in \prod_{i=1}^r M(I_i^\varepsilon)$, we define $F_\rho \in (W^{(q),\infty}(I^\varepsilon))^*$ by

$$F_\rho(\varphi) := \sum_{1\le i\le r} \int_{I_i^\varepsilon} \varphi_{i,t} \mathrm{d}\rho_{i,t} \quad \text{for all } \varphi \in W^{(q),\infty}\big(I^\varepsilon\big).$$

Since by Lemma 2.7, $DG_1(\bar{u}, \bar{y}_0)(v, z_0) \in W^{(q),\infty}(I^\varepsilon)$ for all $(v, z_0) \in U \times \mathbb{R}^n$, we have

$$\begin{aligned}\big\langle \mathrm{d}\rho, DG_1(\bar{u}, \bar{y}_0)(v, z_0)\big\rangle &= \big\langle F_\rho, DG_1(\bar{u}, \bar{y}_0)(v, z_0)\big\rangle \\ &= \big\langle \big(DG_1(\bar{u}, \bar{y}_0)\big)^* F_\rho, (v, z_0)\big\rangle.\end{aligned}$$

Then differentiating L_R, defined by (73), w.r.t. (u, y_0) we get

$$\begin{aligned}&D_{(u,y_0)} L_R(\bar{u}, \bar{y}_0, \mathrm{d}\rho, \nu, \Psi) \\ &\quad = DJ(\bar{u}, \bar{y}_0) + DG_1(\bar{u}, \bar{y}_0)^* F_\rho + DG_2(\bar{u}, \bar{y}_0)^* \nu + DG_3(\bar{u}, \bar{y}_0)^* \Psi. \quad (124)\end{aligned}$$

Let $(\mathrm{d}\eta, \Psi, p), (\mathrm{d}\eta', \Psi', p') \in \Lambda$ be such that $\pi((\mathrm{d}\eta, \Psi, p)) = \pi((\mathrm{d}\eta', \Psi', p'))$. Let by Lemma 3.4 $(\mathrm{d}\rho, \nu), (\mathrm{d}\rho', \nu')$ be such that $(\mathrm{d}\rho, \nu, \Psi), (\mathrm{d}\rho', \nu', \Psi') \in \Lambda_R$. Then $(\nu, \Psi) = (\nu', \Psi')$, and by the definition of Λ_R,

$$D_{(u,y_0)} L_R(\bar{u}, \bar{y}_0, \mathrm{d}\rho, \nu, \Psi) = D_{(u,y_0)} L_R\big(\bar{u}, \bar{y}_0, \mathrm{d}\rho', \nu, \Psi\big) = 0.$$

Then, by (124), $DG_1(\bar{u}, \bar{y}_0)^* F_\rho = DG_1(\bar{u}, \bar{y}_0)^* F_{\rho'}$, and, as a consequence of Lemma 4.1 (a), $DG_1(\bar{u}, \bar{y}_0)^*$ is injective on $(W^{(q),\infty}(I^\varepsilon))^*$. Then $F_\rho = F_{\rho'}$, and by density of $W^{(q),\infty}(I^\varepsilon)$ in $\prod C(I_i^\varepsilon)$, we get $\mathrm{d}\rho = \mathrm{d}\rho'$. Together with $\nu = \nu'$, this implies $\mathrm{d}\eta = \mathrm{d}\eta'$ and then $(\mathrm{d}\eta, \Psi, p) = (\mathrm{d}\eta', \Psi', p')$. □

As a corollary, we get a refinement of Theorem 3.1:

Theorem 4.1 *Let* $(\bar{u}, \bar{y})$ *be a qualified local solution of* (P). *Then* Λ *is nonempty, convex, of finite dimension, and compact.*

Proof Let $\Lambda_\pi := \pi(\Lambda)$. By Theorem 3.1, Λ is nonempty, convex, and weakly $*$ compact, and Λ_π is nonempty, convex, of finite dimension, and compact (π is linear continuous, and its values lie in a finite-dimensional vector space). By Lemma 4.4, $\pi|_\Lambda \colon \Lambda \to \Lambda_\pi$ is a bijection. We claim that its inverse

$$\begin{aligned} m : \qquad \Lambda_\pi \qquad &\longrightarrow \qquad \Lambda \\ \big(\big([\eta_{i,\tau}]\big)_{\tau,i}, \Psi\big) &\longmapsto (\mathrm{d}\eta, \Psi, p)\end{aligned}$$

is the restriction of a continuous affine map. Since $\Lambda = m(\Lambda_\pi)$, the result follows. For the claim, using the convexity of both Λ_π and Λ, the linearity of π, and its injectivity when restricted to Λ, we get that m preserves convex combinations of elements from Λ_π. Thus, we can extend it to an affine map on the affine subspace of $\mathbb{R}^{N*} \times \mathbb{R}^{s*}$ spanned by Λ_π. Since this subspace is of finite dimension, the extension of m is continuous. □

4.3.2 Second-Order Conditions on a Large Critical Cone

Recall that for $\lambda \in \Lambda$, $J[\lambda]$ has been defined on $\mathcal{U} \times \mathbb{R}^n$ by (86) or (87).

Remark 4.2 The form J is quadratic w.r.t. (v, z_0) and affine w.r.t. λ. By Lemmas 2.2, 2.3, and 2.7, $J[\lambda]$ can be extended continuously to $\mathcal{V}_2 \times \mathbb{R}^n$ for any $\lambda \in \Lambda$. We obtain the so-called *Hessian of Lagrangian*

$$J[\lambda]: \mathcal{V}_2 \times \mathbb{R}^n \longrightarrow \mathbb{R}, \tag{125}$$

which is jointly continuous w.r.t. λ and (v, z_0).

The critical L^2 cone C_2 has been defined by (89)–(91). Let the *strict critical L^2 cone* be the set

$$C_2^S := \left\{(v, z_0) \in C_2 : g_i'(\bar{y})z = 0 \text{ on } I_i,\ i = 1, \dots, r\right\},$$

where $z = z[v, z_0] \in \mathcal{Z}_2$.

Theorem 4.2 *Let $(\bar{u}, \bar{y})$ be a qualified local solution of (P). Then, for any $(v, z_0) \in C_2^S$, there exists $\lambda \in \Lambda$ such that*

$$J[\lambda](v, z_0) \geq 0. \tag{126}$$

The proof is based on the following density lemma, announced in the introduction and proved in the next section:

Lemma 4.5 *$C_\infty^R \cap C_2^S$ is dense in C_2^S for the $L^2 \times \mathbb{R}^n$ norm.*

Proof of Theorem 4.2 Let $(v, z_0) \in C_2^S$. By Lemma 4.5, there exists a sequence $(v^k, z_0^k) \in C_\infty^R \cap C_2^S$, $k \in \mathbb{N}$, such that

$$\left(v^k, z_0^k\right) \longrightarrow (v, z_0).$$

By Lemma 3.6 there exists a sequence $\lambda^k \in \Lambda$, $k \in \mathbb{N}$, such that

$$J\left[\lambda^k\right]\left(v^k, z_0^k\right) \geq 0. \tag{127}$$

By Theorem 4.1, Λ is strongly compact; then there exists $\lambda \in \Lambda$ such that, up to a subsequence,

$$\lambda^k \longrightarrow \lambda.$$

We conclude by passing to the limit in (127), thanks to Remark 4.2. □

4.3.3 Density Result

In this section we prove Lemma 4.5, using Lemma 4.1 (b). A result similar to Lemma 4.5 is stated, in the framework of ODEs, as Lemma 5 in [20], but the proof given there is wrong. Indeed, the costates in the optimal control problems of steps (a) and (c) are actually not of bounded variation, and thus the solutions are not essentially bounded. It has to be highlighted that in Lemma 4.1 (b) we get a sequence of essentially bounded v^k.

Proof of Lemma 4.5 We define one more cone,

$$C_\infty^{R+} = \big\{(v, z_0) \in C_\infty^R \cap C_2^S : \exists \delta > 0 : g_i'(\bar{y})z[v, z_0] = 0 \text{ on } I_i^\delta,\ i = 1, \dots, r\big\},$$

and we show actually that C_∞^{R+} is dense in C_2^S.

To this end, we consider the following two normed vector spaces:

$$X_\infty^+ := \big\{(v, z_0) \in U \times \mathbb{R}^n : \exists \delta > 0 : g_i'(\bar{y})z[v, z_0] = 0 \text{ on } I_i^\delta,\ i = 1, \dots, r\big\},$$
$$X_2 := \big\{(v, z_0) \in V_2 \times \mathbb{R}^n : g_i'(\bar{y})z[v, z_0] = 0 \text{ on } I_i,\ i = 1, \dots, r\big\}.$$

Observe that C_∞^{R+} and C_2^S are defined as the same polyhedral cone by (90)–(91), respectively in X_∞^+ and X_2. In view of Lemma 1 in [24], it is then enough to show that X_∞^+ is dense in X_2.

We will need the following lemma, proved in Appendix A.3:

Lemma 4.6 *Let $\bar{b}_i \in W^{(q_i),2}(I_i^\varepsilon)$ be such that*

$$\bar{b}_i = 0 \quad \text{on } I_i. \tag{128}$$

Then there exists $b_i^\delta \in W^{(q_i),\infty}(I_i^\varepsilon)$, $\delta \in]0, \varepsilon[$, such that $b_i^\delta \xrightarrow[\delta \to 0]{W^{(q_i),2}} \bar{b}_i$ and

$$b_i^\delta = 0 \quad \text{on } I_i^\delta. \tag{129}$$

Going back to the proof of Lemma 4.5, let $(\bar{v}, \bar{z}_0) \in X_2$ and $\bar{b} := (A_{2,\bar{z}_0}\bar{v})|_{I^\varepsilon}$. We consider a sequence $\delta_k \searrow 0$ and, for $i = 1, \dots, r$, $b_i^k := b_i^{\delta_k} \in W^{(q_i),\infty}(I_i^\varepsilon)$ given by Lemma 4.6. Applying Lemma 4.1 b) to b^k, we get v^k, $k \in \mathbb{N}$. We have $(v^k, \bar{z}_0) \in X_\infty^+$ and $(v^k, \bar{z}_0) \longrightarrow (\bar{v}, \bar{z}_0)$. The proof is completed. □

4.4 Sufficient Conditions

We still are under assumptions (A0)–(A4).

Definition 4.1 A quadratic form Q over a Hilbert space X is a *Legendre form* iff it is weakly lower semi-continuous and satisfies the following property: if $x^k \rightharpoonup x$ weakly in X and $Q(x^k) \to Q(x)$, then $x^k \to x$ strongly in X.

Theorem 4.3 *Suppose that for any* $(v, z_0) \in C_2$, *there exists* $\lambda \in \Lambda$ *such that* $J[\lambda]$ *is a Legendre form and*

$$J[\lambda](v, z_0) > 0 \quad \textit{if } (v, z_0) \neq 0. \tag{130}$$

Then $(\bar{u}, \bar{y})$ *is a local solution of* (P) *satisfying the following quadratic growth condition*: *there exist* $\beta > 0$ *and* $\alpha > 0$ *such that*

$$J(u, y_0) \geq J(\bar{u}, \bar{y}_0) + \frac{1}{2}\beta\big(\|u - \bar{u}\|_2 + |y_0 - \bar{y}_0|\big)^2 \tag{131}$$

for any trajectory (u, y) *feasible for* (P) *and such that* $\|u - \bar{u}\|_\infty + |y_0 - \bar{y}_0| \leq \alpha$.

Remark 4.3 Let $\lambda = (\mathrm{d}\eta, \Psi, p) \in \Lambda$. The *strengthened Legendre–Clebsch condition*

$$\exists \bar{\alpha} > 0 : D^2_{uu} H[p](t, \bar{u}_t, \bar{y}_t) \geq \bar{\alpha} I_m \quad \text{for a.a. } t \in [0, T] \tag{132}$$

is satisfied iff $J[\lambda]$ is a Legendre form (it can be proved by combining Theorem 11.6 and Theorem 3.3 in [25]).

Proof of Theorem 4.3 (i) Let us assume that (130) holds but (131) does not. Then there exists a sequence of feasible trajectories (u^k, y^k) such that

$$\begin{cases} \big(u^k, y^k_0\big) \xrightarrow{L^\infty \times \mathbb{R}^n} (\bar{u}, \bar{y}_0), \quad \big(u^k, y^k_0\big) \neq (\bar{u}, \bar{y}_0), \\ J\big(u^k, y^k_0\big) \leq J(\bar{u}, \bar{y}_0) + o\big(\big\|u^k - \bar{u}\big\|_2 + \big|y^k_0 - \bar{y}_0\big|\big)^2. \end{cases} \tag{133}$$

Let $\sigma_k := \|u^k - \bar{u}\|_2 + |y^k_0 - \bar{y}_0|$ and $(v^k, z^k_0) := \sigma_k^{-1}(u^k - \bar{u}, y^k_0 - \bar{y}_0) \in \mathcal{U} \times \mathbb{R}^n$. There exists $(\bar{v}, \bar{z}_0) \in \mathcal{V}_2 \times \mathbb{R}^n$ such that, up to a subsequence,

$$\big(v^k, z^k_0\big) \rightharpoonup (\bar{v}, \bar{z}_0) \quad \text{weakly in } \mathcal{V}_2 \times \mathbb{R}^n.$$

(ii) We claim that $(\bar{v}, \bar{z}_0) \in C_2$.

Let $z^k := z[v^k, z^k_0] \in \mathcal{Y}$ and $\bar{z} := z[\bar{v}, \bar{z}_0] \in \mathcal{Z}_2$. We derive from the compact embedding $\mathcal{Z}_2 \subset C([0, T]; \mathbb{R}^n)$ that, up to a subsequence,

$$z^k \to \bar{z} \quad \text{in } C\big([0, T]; \mathbb{R}^n\big). \tag{134}$$

Moreover, it is classical (see, e.g., the proof of Lemma 20 in [18]) that

$$J\big(u^k, y^k_0\big) = J(\bar{u}, \bar{y}_0) + \sigma_k DJ(\bar{u}, \bar{y}_0)\big(v^k, z^k_0\big) + o(\sigma_k), \tag{135}$$

$$g\big(y^k\big) = g(\bar{y}) + \sigma_k g'(\bar{y}) z^k + o(\sigma_k), \tag{136}$$

$$\Phi\big(y^k_0, y^k_T\big) = \Phi(\bar{y}_0, \bar{y}_T) + \sigma_k D\Phi(\bar{y}_0, \bar{y}_T)\big(z^k_0, z^k_T\big) + o(\sigma_k). \tag{137}$$

It follows that

$$DJ(\bar{u}, \bar{y}_0)(\bar{v}, \bar{z}_0) \leq 0, \tag{138}$$

$$\begin{cases} g_i'(\bar{y})\bar{z} \leq 0 & \text{on } I_i,\ i = 1, \ldots, r_1, \\ g_i'(\bar{y})\bar{z} \leq 0 & \text{on } I_i \cup T_i,\ i = r_1 + 1, \ldots, r. \end{cases} \tag{139}$$

$$D\Phi(\bar{y}_0, \bar{y}_T)\big(\bar{z}_0, z[\bar{v}, \bar{z}_0]_T\big) \in T_K\big(\Phi(\bar{y}_0, \bar{y}_T)\big), \tag{140}$$

using (133) for (138) and the fact that $(\bar{u}, \bar{y})$ and (u^k, y^k) are feasible for (139) and (140). By Lemma 3.1, given $\bar{\lambda} = (\mathrm{d}\bar{\eta}, \bar{\Psi}, \bar{p}) \in \Lambda$, we have

$$DJ(\bar{u}, \bar{y}_0)(\bar{v}, \bar{z}_0) + \int_{[0,T]} \mathrm{d}\bar{\eta}_t g'(\bar{y}_t) + \bar{\Psi} D\Phi(\bar{y}_0, \bar{y}_T)(\bar{z}_0, \bar{z}_T) = 0.$$

Together with Definition 2.1 and (138)–(140), this implies that each of the three terms is null, i.e., $(\bar{v}, \bar{z}_0) \in C_2$.

(iii) Then by (130) there exists $\bar{\lambda} \in \Lambda$ such that $J[\bar{\lambda}]$ is a Legendre form and

$$0 \leq J[\bar{\lambda}](\bar{v}, \bar{z}_0). \tag{141}$$

In particular, $J[\bar{\lambda}]$ is weakly lower semi-continuous. Then

$$J[\bar{\lambda}](\bar{v}, \bar{z}_0) \leq \liminf_k J[\bar{\lambda}]\big(v^k, z_0^k\big) \leq \limsup_k J[\bar{\lambda}]\big(v^k, z_0^k\big). \tag{142}$$

We claim that

$$\limsup_k J[\bar{\lambda}]\big(v^k, z_0^k\big) \leq 0. \tag{143}$$

Indeed, similarly to (135)–(137), one can show that, $\bar{\lambda}$ being a multiplier,

$$L_R\big(u^k, y_0^k, \bar{\lambda}\big) - L_R(\bar{u}, \bar{y}_0, \bar{\lambda}) = \frac{1}{2}\sigma_k^2 D^2_{(u,y_0)^2} L_R(\bar{u}, \bar{y}_0, \bar{\lambda})\big(v^k, z_0^k\big)^2 + o\big(\sigma_k^2\big). \tag{144}$$

Since $L_R(u^k, y_0^k, \bar{\lambda}) - L_R(\bar{u}, \bar{y}_0, \bar{\lambda}) \leq J(u^k, y_0^k) - J(\bar{u}, \bar{y}_0)$, we derive from (133), (144), and Lemma 3.5 that

$$J[\bar{\lambda}]\big(v^k, z_0^k\big) \leq o(1). \tag{145}$$

(iv) We derive from (141), (142), and (143) that

$$J[\bar{\lambda}]\big(v^k, z_0^k\big) \longrightarrow 0 = J[\bar{\lambda}](\bar{v}, \bar{z}_0).$$

By (130), $(\bar{v}, \bar{z}_0) = 0$. Then $(v^k, z_0^k) \longrightarrow (\bar{v}, \bar{z}_0)$ strongly in $V_2 \times \mathbb{R}^n$ by the definition of a Legendre form. We get a contradiction with the fact that $\|v^k\|_2 + |z_0^k| = 1$ for all k. □

In view of Theorems 4.2 and 4.3, it appears that, under an extra assumption of the type of strict complementarity on the running state constraints, we can state no-gap second-order optimality conditions. We denote by $\mathrm{ri}(\Lambda)$ the relative interior of Λ (see Definition 2.16 in [21]).

Corollary 4.1 *Let* $(\bar{u}, \bar{y})$ *be a qualified feasible trajectory for* (P). *We assume that* $C_2^S = C_2$ *and that for any* $\lambda \in \operatorname{ri}(\Lambda)$, *the strengthened Legendre–Clebsch condition* (132) *holds. Then* $(\bar{u}, \bar{y})$ *is a local solution of* (P) *satisfying the quadratic growth condition* (131) *iff for any* $(v, z_0) \in C_2 \setminus \{0\}$, *there exists* $\lambda \in \Lambda$ *such that*

$$J[\lambda](v, z_0) > 0. \tag{146}$$

Proof Suppose that (146) holds for some $\lambda \in \Lambda$; then it holds for some $\lambda \in \operatorname{ri}(\Lambda)$ too, and now $J[\lambda]$ is a Legendre form. By Theorem 4.3, there is locally quadratic growth.

Conversely, suppose that (131) holds for some $\beta > 0$ and let

$$J_\beta(u, y_0) := J(u, y_0) - \frac{1}{2}\beta\big(\|u - \bar{u}\|_2 + |y_0 - \bar{y}_0|\big)^2.$$

Then $(\bar{u}, \bar{y}_0)$ is a local solution of the following optimization problem:

$$\min_{(u,y_0)\in U\times\mathbb{R}^n} \quad J_\beta(u, y_0) \qquad \text{subject to} \quad G_i(u, y_0) \in K_i, \quad i = 1, 2, 3.$$

This problem has the same Lagrange multipliers as the reduced problem (write that the respective Lagrangian is stationary at $(\bar{u}, \bar{y}_0)$), the same critical cones, and its Hessian of Lagrangian is

$$J_\beta[\lambda](v, z_0) = J[\lambda](v, z_0) - \beta\big(\|v\|_2 + |z_0|\big)^2.$$

Theorem 4.2 applied to this problem gives (146). □

Remark 4.4 A sufficient condition (not necessary a priori) to have $C_2^S = C_2$ is the existence of $(\mathrm{d}\bar{\eta}, \bar{\Psi}, \bar{p}) \in \Lambda$ such that

$$\operatorname{supp}(\mathrm{d}\bar{\eta}_i) = I_i, \quad i = 1, \dots, r.$$

5 Concluding Remarks

Our main result in this paper is the statement of second-order necessary conditions on a large critical cone. This result is obtained by density, under some assumptions on the running state constraints and their contact sets. The density technique might be adapted to mixed control-state constraints.

These necessary conditions turn out to be no-gap optimality conditions if a strict complementarity condition and a strengthened Legendre–Clebsch condition hold. It has to be noted that the latter would be satisfied if we could state second-order optimality conditions involving Pontryagin multipliers, as we intend to do in a future work.

An extension of the results presented here to other classes of equations with memory, such as delay differential equations, should also be possible.

Acknowledgements We thank the two anonymous referees for their useful remarks. The research leading to these results has received funding from the EU 7th Framework Programme (FP7-PEOPLE-2010-ITN), under GA No. 264735-SADCO. The first and third authors also thank the MMSN (Modélisation Mathématique et Simulation Numérique) Chair (EADS, Inria, and Ecole Polytechnique) for its support.

Appendix

A.1 Functions of Bounded Variation

The main reference here is [26], Sect. 3.2. Recall that with the definition of $BV([0,T];\mathbb{R}^{n*})$ given at the beginning of Sect. 2.2, for $h \in BV([0,T];\mathbb{R}^{n*})$, there exist $h_{0_-}, h_{T_+} \in \mathbb{R}^{n*}$ such that (6) holds.

Lemma A.1 *Let $h \in BV([0,T];\mathbb{R}^{n*})$. Let h^l, h^r be defined for all $t \in [0,T]$ by*

$$h_t^l := h_{0_-} + \mathrm{d}h\big([0,t[\big), \tag{147}$$

$$h_t^r := h_{0_-} + \mathrm{d}h\big([0,t]\big). \tag{148}$$

Then they are both in the same equivalence class of h, h^l is left continuous, h^r is right continuous, and, for all $t \in [0,T]$,

$$h_t^l = h_{T_+} - \mathrm{d}h\big([t,T]\big), \tag{149}$$

$$h_t^r = h_{T_+} - \mathrm{d}h\big(]t,T]\big). \tag{150}$$

Proof Theorem 3.28 in [26]. □

The identification between measures and functions of bounded variation that we mention at the beginning of Sect. 2.2 relies on the following:

Lemma A.2 *The linear map*

$$(c,\mu) \longmapsto \big(h: t \mapsto c - \mu\big([t,T]\big)\big) \tag{151}$$

is an isomorphism between $\mathbb{R}^{r*} \times M([0,T];\mathbb{R}^{r*})$ and $BV([0,T];\mathbb{R}^{r*})$, whose inverse is

$$h \longmapsto (h_{T_+}, \mathrm{d}h). \tag{152}$$

Proof Theorem 3.30 in [26]. □

Let us now prove Lemma 2.1:

Proof of Lemma 2.1 By (149), a solution in P of (11) is any $p \in L^1([0,T];\mathbb{R}^{n*})$ such that, for a.e. $t \in [0,T]$,

$$p_t = D_{y_2}\Phi[\Psi](y_0,y_T) + \int_t^T D_y H[p](s,u_s,y_s)\,\mathrm{d}s + \int_{[t,T]} \mathrm{d}\eta_s\, g'(y_s). \tag{153}$$

We define $\Theta: L^1([0,T];\mathbb{R}^{n*}) \to L^1([0,T];\mathbb{R}^{n*})$ by

$$\Theta(p)_t := D_{y_2}\Phi[\Psi](y_0,y_T) + \int_t^T D_y H[p](s,u_s,y_s)\,\mathrm{d}s + \int_{[t,T]} \mathrm{d}\eta_s\, g'(y_s) \tag{154}$$

for a.e. $t \in [0, T]$, and we show that Θ has a unique fixed point. Let $C > 0$ be such that $\|D_y f\|_\infty, \|D^2_{y,\tau} f\|_\infty \le C$ along (u, y). Then

$$\begin{aligned}\left|\Theta(p_1)_t - \Theta(p_2)_t\right| &= \left|\int_t^T \big(D_y H[p_1](s, u_s, y_s) - D_y H[p_2](s, u_s, y_s)\big)\,\mathrm{d}s\right| \\ &\le C \int_t^T \left[\left|p_1(s) - p_2(s)\right| + \int_s^T \left|p_1(\theta) - p_2(\theta)\right| \mathrm{d}\theta\right] \mathrm{d}s \\ &= C \int_t^T \left[\left|p_1(s) - p_2(s)\right| + \int_t^s \left|p_1(s) - p_2(s)\right| \mathrm{d}\theta\right] \mathrm{d}s \\ &\le C(1+T) \int_t^T \left|p_1(s) - p_2(s)\right| \mathrm{d}s.\end{aligned}$$

We consider the family of equivalent norms on $L^1([0, T]; \mathbb{R}^{n*})$

$$\|v\|_{1,K} := \left\|t \mapsto e^{-K(T-t)} v(t)\right\|_1 \quad (K \ge 0). \tag{155}$$

Then

$$\begin{aligned}\left\|\Theta(p_1) - \Theta(p_2)\right\|_{1,K} &\le C(1+T) \int_0^T \int_t^T e^{-K(T-t)} \left|p_1(s) - p_2(s)\right| \mathrm{d}s\,\mathrm{d}t \\ &= C(1+T) \int_0^T e^{-K(T-s)} \left|p_1(s) - p_2(s)\right| \left[\int_0^s e^{K(t-s)}\,\mathrm{d}t\right] \mathrm{d}s \\ &\le \frac{C(1+T)}{K} \|p_1 - p_2\|_{1,K}.\end{aligned}$$

For K big enough, Θ is a contraction on $L^1([0, T]; \mathbb{R}^{n*})$ for $\|\cdot\|_{1,K}$; its unique fixed point is the unique solution of (11). □

Another useful result is the following integration by parts formula:

Lemma A.3 *Let $h, k \in BV([0, T])$. Then $h^l \in L^1(\mathrm{d}k)$, $k^r \in L^1(\mathrm{d}h)$, and*

$$\int_{[0,T]} h^l\,\mathrm{d}k + \int_{[0,T]} k^r\,\mathrm{d}h = h_{T_+} k_{T_+} - h_{0_-} k_{0_-}. \tag{156}$$

Proof Let $\Omega := \{0 \le y \le x \le T\}$. Since $\chi_\Omega \in L^1(\mathrm{d}h \otimes \mathrm{d}k)$, we have by Fubini's theorem (Theorem 7.27 in [27]) and Lemma A.1 that $h^l \in L^1(\mathrm{d}k)$, $k^r \in L^1(\mathrm{d}h)$, and we can compute $\mathrm{d}h \otimes \mathrm{d}k(\Omega)$ in two different ways:

$$\begin{aligned}\mathrm{d}h \otimes \mathrm{d}k(\Omega) &= \int_{[0,T]} \int_{[y,T]} \mathrm{d}h_x\,\mathrm{d}k_y \\ &= \int_{[0,T]} \big(h_{T_+} - h^l_y\big)\,\mathrm{d}k_y \\ &= h_{T_+}(k_{T_+} - k_{0_-}) - \int_{[0,T]} h^l_y\,\mathrm{d}k_y,\end{aligned}$$

$$\begin{aligned} \mathrm{d}h \otimes \mathrm{d}k(\Omega) &= \int_{[0,T]} \int_{[0,x]} \mathrm{d}k_y \, \mathrm{d}h_x \\ &= \int_{[0,T]} k_x^r \, \mathrm{d}h_x - k_{0_-}(h_{T_+} - h_{0_-}). \end{aligned}$$

□

A.2 The Hidden Use of Assumption (A3)

We use assumption (A3) to prove Lemma 4.3 (and then Lemma 4.1, ...) through the following:

Lemma A.4 *Recall that* $M_t := D_{\tilde{u}} G^{(q)}_{I^{\varepsilon_0}(t)}(t, \bar{u}_t, \bar{y}_t, \bar{u}, \bar{y}) \in M_{|I_t^{\varepsilon_0}|,m}(\mathbb{R})$, $t \in [0,T]$. *Then* $M_t M_t^*$ *is invertible and* $|(M_t M_t^*)^{-1}| \leq \gamma^{-2}$ *for all* $t \in [0,T]$.

Proof For any $x \in \mathbb{R}^{|I^{\varepsilon_0}(t)|}$,

$$\langle M_t M_t^* x, x \rangle = |M_t^* x|^2 \geq \gamma^2 |x|^2.$$

Then $M_t M_t^* x = 0$ implies $x = 0$, and the invertibility follows.

Let $y \in \mathbb{R}^{|I^{\varepsilon_0}(t)|}$ and $x := (M_t M_t^*)^{-1} y$.

$$|y||x| \geq \langle y, x \rangle = \langle M_t M_t^* x, x \rangle = |M_t^* x|^2 \geq \gamma^2 |x|^2.$$

For $y \neq 0$, we have $x \neq 0$; dividing the previous inequality by $|x|$, we get

$$\gamma^2 \left| (M_t M_t^*)^{-1} y \right| \leq |y|.$$

The result follows. □

Before we prove Lemma 4.3, we define the truncation of an integrable function:

Definition A.1 Given any $\phi \in L^s(J)$ ($s \in [1, \infty[$ and J interval), we will call the *truncation of* ϕ the sequence $\phi^k \in L^\infty(J)$ defined for $k \in \mathbb{N}$ and a.a. $t \in J$ by

$$\phi_t^k := \begin{cases} \phi_t & \text{if } |\phi_t| \leq k, \\ k \frac{\phi_t}{|\phi_t|} & \text{otherwise.} \end{cases}$$

Observe that $\phi^k \xrightarrow[k \to \infty]{L^s} \phi$.

Proof of Lemma 4.3 In the sequel we omit z_0 in the notations.

(i) Let $v \in V_s$. We claim that v satisfies

$$M_t v_t + N_t\big(z[v]_t, v, z[v]\big) = h_t \quad \text{for a.a. } t \in J_l \tag{157}$$

iff there exists $w \in L^s(J_l; \mathbb{R}^m)$ such that (v, w) satisfies

$$\begin{cases} M_t w_t = 0, \\ v_t = M_t^* (M_t M_t^*)^{-1} \big(h_t - N_t\big(z[v]_t, v, z[v]\big)\big) + w_t, \end{cases} \quad \text{for a.a. } t \in J_l. \tag{158}$$

Clearly, if (v, w) satisfies (158), then v satisfies (157). Conversely, suppose that v satisfies (157). With Lemma A.4 in mind, we define $\alpha \in L^s(J_l; \mathbb{R}^{|I_l|})$ and $w \in L^s(J_l; \mathbb{R}^m)$ by

$$\alpha := \left(MM^*\right)^{-1} Mv,$$
$$w := \left(I_m - M^*\left(MM^*\right)^{-1} M\right)v.$$

Then

$$\begin{cases} Mw = 0, \\ v = M^*\alpha + w, \end{cases} \quad \text{on } J_l. \tag{159}$$

We derive from (157) and (159) that

$$M_t M_t^* \alpha_t + N_t\left(z[v]_t, v, z[v]\right) = h_t \quad \text{for a.a. } t \in J_l.$$

Using again Lemma A.4 and (159), we get (158).

(ii) Given $(v, h, w) \in V_s \times L^s(J_l; \mathbb{R}^{|I_l|}) \times L^s(J_l; \mathbb{R}^m)$, there exists a unique $\tilde{v} \in V_s$ such that

$$\begin{cases} \tilde{v} = v \quad \text{on } J_0 \cup \cdots \cup J_{l-1} \cup J_{l+1} \cup \cdots \cup J_\kappa, \\ \tilde{v}_t = M_t^*\left(M_t M_t^*\right)^{-1}\left(h_t - N_t\left(z[\tilde{v}]_t, \tilde{v}, z[\tilde{v}]\right)\right) + w_t \quad \text{for a.a. } t \in J_l, \end{cases} \tag{160}$$

Indeed, one can define a mapping from V_s to V_s, using the right-hand side of (160). Then it can be shown, as in the proof of Lemma 2.1, that this mapping is a contraction for a well-suited norm, using Lemmas 2.2, 2.3, and A.4. The existence and uniqueness follow. Moreover, a version of the contraction mapping theorem with parameter (see, e.g., Théorème 21-5 in [28]) shows that $\tilde{v}$ depends continuously on (v, h, w).

(iii) Let us prove (a): let $(\bar{h}, v) \in L^s(J_l; \mathbb{R}^{|I_l|}) \times V_s$, and let $w := 0$. Let $\tilde{v} \in V_s$ be the unique solution of (160) for $(v, \bar{h}, w)$. Then $\tilde{v}$ is a solution of (112) by (i).

(iv) Let us prove (b): let $(\bar{h}, \bar{v}) \in L^s(J_l; \mathbb{R}^{|I_l|}) \times V_s$ as in the statement, and let $\bar{w}$ be given by (i). Then $\bar{v}$ is the unique solution of (160) for $(\bar{v}, \bar{h}, \bar{w})$.

Let $(h^k, v^k) \in L^\infty(J_l; \mathbb{R}^{|I_l|}) \times U$, $k \in \mathbb{N}$, be such that $(h^k, v^k) \xrightarrow{L^s \times L^s} (\bar{h}, \bar{v})$, and let $w^k \in L^\infty(J_l; \mathbb{R}^m)$, $k \in \mathbb{N}$, be the truncation of $\bar{w}$. It is obvious from Definition A.1 that

$$M_t w_t^k = 0 \quad \text{for a.a. } t \in J_l.$$

Let $\tilde{v}^k \in U$ be the unique solution of (160) for (v^k, h^k, w^k), $k \in \mathbb{N}$. Then, by the uniqueness and continuity in (ii),

$$\tilde{v}^k \xrightarrow{L^s} \bar{v}, \tag{161}$$

and $\tilde{v}^k$ is a solution of (114) by (i). □

We finish this section with an example where assumption (A3) can be satisfied or not.

Example A.1 We consider the scalar Example 2.1.2 with $q=1$ and $f(t,s)=f(2t-s)$:

$$y_t = \int_0^t f(2t-s)u_s\,\mathrm{d}s, \quad t \in [0,T], \tag{162}$$

where f is a continuous function and is not a polynomial, and the trajectory $(\bar{u},\bar{y})=(0,0)$. Then

$$M_t = f(t) \in M_{1,1}(\mathbb{R}),$$

and (A3) is satisfied iff

$$f(t) \neq 0 \quad \forall t \in [0,T].$$

A.3 Approximations in $W^{q,2}$

We will prove in this section Lemmas 4.2 and 4.6. First, we give the statement and the proof of a general result:

Lemma A.5 *Let $\hat{x} \in W^{q,2}([0,1])$. For $j=0,\dots,q-1$, we denote*

$$\begin{cases} \hat{\alpha}_j := \hat{x}^{(j)}(0), \\ \hat{\beta}_j := \hat{x}^{(j)}(1), \end{cases} \tag{163}$$

and we consider $\alpha_j^k, \beta_j^k \in \mathbb{R}^q$, $k \in \mathbb{N}$, such that $(\alpha_j^k, \beta_j^k) \longrightarrow (\hat{\alpha}_j, \hat{\beta}_j)$. Then there exists $x^k \in W^{q,\infty}([0,1])$, $k \in \mathbb{N}$, such that $x^k \xrightarrow{W^{q,2}} \hat{x}$ and, for $j=0,\dots,q-1$,

$$\begin{cases} \left(x^k\right)^{(j)}(0) = \alpha_j^k, \\ \left(x^k\right)^{(j)}(1) = \beta_j^k. \end{cases} \tag{164}$$

Proof Given $u \in L^2([0,1])$, we define $x_u \in W^{q,2}([0,1])$ by

$$x_u(t) := \int_0^t \int_0^{s_1} \cdots \int_0^{s_{q-1}} u(s_q)\,\mathrm{d}s_q\,\mathrm{d}s_{q-1}\cdots\mathrm{d}s_1, \quad t \in [0,1].$$

Then $x_u^{(q)} = u$ and, for $j=0,\dots,q-1$,

$$x_u^{(j)}(1) = \gamma_j \quad \Longleftrightarrow \quad \langle a_j, u\rangle_{L^2} = \gamma_j,$$

where $a_j \in C([0,1])$ is defined by

$$a_j(t) := \frac{(1-t)^{q-1-j}}{(q-1-j)!}, \quad t \in [0,1].$$

Indeed, a straightforward induction shows that

$$x_u^{(j)}(1) = \int_0^1 \int_0^{s_{j+1}} \cdots \int_0^{s_{q-1}} u(s_q)\,\mathrm{d}s_q\,\mathrm{d}s_{q-1}\cdots\mathrm{d}s_{j+1}.$$

Then integrations by parts give the expression of the a_j. Note that the a_j ($j = 0, \dots, q-1$) are linearly independent in $L^2([0,1])$. Then

$$A: \quad \mathbb{R}^q \longrightarrow L^2\big([0,1]\big)$$
$$\begin{pmatrix} \lambda_0 \\ \vdots \\ \lambda_{q-1} \end{pmatrix} \longmapsto \sum_{j=0}^{q-1} \lambda_j a_j$$

is such that A^*A is invertible (here A^* is the adjoint operator), and

$$x_u^{(j)}(1) = \gamma_j, \quad j = 0, \dots, q-1 \iff A^*u = (\gamma_0, \dots, \gamma_{q-1})^T. \tag{165}$$

Going back to the lemma, let $\hat{u} := \hat{x}^{(q)} \in L^2([0,1])$. Observe that

$$\hat{x}(t) = \sum_{l=0}^{q-1} \frac{\hat{\alpha}_l}{l!} t^l + x_{\hat{u}}(t), \quad t \in [0,1],$$

and that $A^*\hat{u} = (\hat{\gamma}_0, \dots, \hat{\gamma}_{q-1})^T$, where

$$\hat{\gamma}_j := \hat{\beta}_j - \sum_{l=j}^{q-1} \frac{\hat{\alpha}_l}{(l-j)!}, \quad j = 0, \dots, q-1.$$

Then we consider, for $k \in \mathbb{N}$, the truncation (Definition A.1) $\hat{u}^k \in L^\infty([0,1])$ of $\hat{u}$ and

$$\gamma_j^k := \beta_j^k - \sum_{l=j}^{q-1} \frac{\alpha_l^k}{(l-j)!}, \quad j = 0, \dots, q-1, \tag{166}$$
$$\gamma^k := \big(\gamma_0^k, \dots, \gamma_{q-1}^k\big)^T,$$
$$u^k := \hat{u}^k + A\big(A^*A\big)^{-1}\big(\gamma^k - A^*\hat{u}^k\big),$$
$$x^k(t) := \sum_{l=0}^{q-1} \frac{\alpha_l^k}{l!} t^l + x_{u^k}(t), \quad t \in [0,1]. \tag{167}$$

It is clear that $u^k \in L^\infty([0,1])$ (by the definition of A); then $x^k \in W^{q,\infty}([0,T])$. Since $A^*u^k = \gamma^k$ and in view of (165), (166), and (167), (164) is satisfied. Finally, $\gamma_j^k \longrightarrow \hat{\gamma}_j$, for $j = 1$ to $q-1$; then $\gamma^k \longrightarrow A^*\hat{u}$ and $u^k \longrightarrow \hat{u}$. □

We can also prove the following:

Lemma A.6 *Let $\hat{x} \in W^{q,2}([0,1])$ be such that $\hat{x}^{(j)}(0) = 0$ for $j = 0, \dots, q-1$. Then for $\delta > 0$, there exists $x^\delta \in W^{q,\infty}([0,1])$ such that $x^\delta \xrightarrow[\delta \to 0]{W^{q,2}} \hat{x}$ and*

$$x^\delta = 0 \quad \text{on } [0,\delta]. \tag{168}$$

Proof We consider $u^\delta \in L^\infty([0,1])$, $\delta > 0$, such that $u^\delta = 0$ on $[0,\delta]$ and $u^\delta \xrightarrow[\delta\to 0]{L^2} \hat{u} := \hat{x}^{(q)}$. Then we define $x^\delta := x_{u^\delta}$ (see the previous proof). □

Now the proof of Lemma 4.6 is straightforward.

Proof of Lemma 4.6 We observe that $\bar{b}_i = 0$ on I_i implies that $\bar{b}_i^{(j)} = 0$ at the end points of I_i for $j = 0, \dots, q_i - 1$ (note that with the definition (68), if one component of I_i is a singleton, then $q_i = 1$). Then the conclusion follows with Lemma A.6 applied on each component of $I_i^\varepsilon \setminus I_i$. □

Finally, we use Lemma A.5 to prove Lemma 4.2.

Proof of Lemma 4.2 In the sequel we omit z_0 in the notations. We define a *connection in* $W^{q,\infty}$ between ψ_1 at t_1 and ψ_2 at t_2 as any $\psi \in W^{q,\infty}([t_1,t_2])$ such that

$$\begin{cases} \psi^{(j)}(t_1) = \psi_1^{(j)}(t_1), \\ \psi^{(j)}(t_2) = \psi_2^{(j)}(t_2), \end{cases} \quad j = 0, \dots, q-1.$$

(a) We define $\tilde{b}_i$ on $[0,t_0]$ by $\tilde{b}_i := g_i'(\bar{y})z[v]$, $i = 1, \dots, r$. We need to explain how we define $\tilde{b}_i$ on $]t_0,T]$, using $\bar{b}_i$ and connections, to have $\tilde{b}_i \in W^{q_i,s}([0,T])$ and $\tilde{b}_i = \bar{b}_i$ on each component of $I_i^\varepsilon \cap]t_0,T]$. The construction is slightly different whether $t_0 \in I_i^\varepsilon$ or not, i.e., whether $i \in I_{t_0}^\varepsilon$ or not. Note that by the definition of ε_0 and of t_0, I_t^ε is constant for t in a neighborhood of t_0. We now distinguish the two cases just mentioned:

1. $i \in I_{t_0}^\varepsilon$: We denote by $[t_1,t_2]$ the connected component of I_i^ε such that $t_0 \in]t_1,t_2[$. We derive from (105) that $\tilde{b}_i = \bar{b}_i$ on $[t_1,t_0]$. Then we define $\tilde{b}_i := \bar{b}_i$ on $]t_0,t_2]$.
   If I_i^ε has another component in $]t_2,T]$, we denote the first one by $[t_1',t_2']$. Let ψ be a connection in $W^{q_i,\infty}$ between $\tilde{b}_i$ at t_2 to $\bar{b}_i$ at t_1'. We define $\tilde{b}_i := \psi$ on $]t_2,t_1'[$, $\tilde{b}_i := \bar{b}_i$ on $[t_1',t_2']$, and so forth on $]t_2',T]$.
   If I_i^ε has no more component, we define $\tilde{b}_i$ on what is left as a connection in $W^{q_i,\infty}$ between $\bar{b}_i$ and $g_i'(\bar{y})z[v]$ at T.
2. $i \notin I_{t_0}^\varepsilon$: If I_i^ε has a component in $[t_0,T]$, we denote the first one by $[t_1,t_2]$. Note that $t_1 - t_0 \geq \varepsilon_0 - \varepsilon > 0$. We consider a connection in $W^{q_i,\infty}$ between $\tilde{b}_i$ at t_0 and $\bar{b}_i$ at t_1 and continue as in 1.
   If I_i^ε has no component in $[t_0,T]$, we do as in 1.

(b) For all $k \in \mathbb{N}$, we apply (a) to (b^k, v^k), and we get $\tilde{b}^k$. We just need to explain how we can get, for $i = 1, \dots, r$,

$$\tilde{b}_i^k \xrightarrow[k\to\infty]{W^{q_i,2}} g_i'(\bar{y})z[\bar{v}].$$

By construction we have

$$\text{on } [0, t_0], \quad \tilde{b}_i^k = g_i'(\bar{y})z\big[v^k\big] \longrightarrow g_i'(\bar{y})z[\bar{v}],$$
$$\text{on } I_i^\varepsilon, \quad \tilde{b}_i^k = b_i^k \longrightarrow \bar{b}_i = g_i'(\bar{y})z[\bar{v}].$$

Then it is enough to show that every connection which appears when we apply (a) to (b^k, v^k), for example, $\psi_i^k \in W^{q_i,\infty}([t_1, t_2])$, can be chosen in such a way that

$$\psi_i^k \longrightarrow g_i'(\bar{y})z[\bar{v}] \quad \text{on } [t_1, t_2].$$

This is possible by Lemma A.5. □

References

1. Volterra, V.: Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. Memoria R. comitato talassografico italiano, vol. 131 (1927)
2. Scudo, F.M., Ziegler, J.R.: The Golden Age of Theoretical Ecology, 1923–1940. Lecture Notes in Biomathematics, vol. 22, pp. 1923–1940. Springer, Berlin (1978). A collection of works by V. Volterra, V.A. Kostitzin, A.J. Lotka, and A.N. Kolmogoroff
3. Volterra, V.: Leçons sur les équations intégrales et les équations intégro-différentielles. Collection de monographies sur la théorie des fonctions. Gauthier-Villars, Paris (1913). Leçons professées à la Faculté des sciences de Rome en 1910
4. Kamien, M.I., Muller, E.: Optimal control with integral state equations. Rev. Econ. Stud. **43**(3), 469–473 (1976)
5. Vinokurov, V.R.: Optimal control of processes describable by integral equations. I. Izv. Vysš. Učebn. Zaved., Mat. **62**(7), 21–33 (1967)
6. Vinokurov, V.R.: Optimal control of processes described by integral equations. I. SIAM J. Control **7**, 324–336 (1969)
7. Vinokurov, V.R.: Optimal control of processes described by integral equations. II. SIAM J. Control **7**, 337–345 (1969)
8. Vinokurov, V.R.: Optimal control of processes described by integral equations. III. SIAM J. Control **7**, 346–355 (1969)
9. Neustadt, L.W., Warga, J.: Comments on the paper "Optimal control of processes described by integral equations. I" by V.R. Vinokurov. SIAM J. Control **8**, 572 (1970)
10. Bakke, V.L.: A maximum principle for an optimal control problem with integral constraints. J. Optim. Theory Appl. **13**, 32–55 (1974)
11. Carlson, D.A.: An elementary proof of the maximum principle for optimal control problems governed by a Volterra integral equation. J. Optim. Theory Appl. **54**(1), 43–61 (1987)
12. de la Vega, C.: Necessary conditions for optimal terminal time control problems governed by a Volterra integral equation. J. Optim. Theory Appl. **130**(1), 79–93 (2006)
13. Carlier, G., Tahraoui, R.: On some optimal control problems governed by a state equation with memory. ESAIM Control Optim. Calc. Var. **14**(4), 725–743 (2008)
14. Bonnans, J.F., de la Vega, C.: Optimal control of state constrained integral equations. Set-Valued Var. Anal. **18**(3–4), 307–326 (2010)
15. Kawasaki, H.: An envelope-like effect of infinitely many inequality constraints on second-order necessary conditions for minimization problems. Math. Program. **41**(1, Ser. A), 73–96 (1988)
16. Cominetti, R.: Metric regularity, tangent sets, and second-order optimality conditions. Appl. Math. Optim. **21**(3), 265–287 (1990)
17. Páles, Z., Zeidan, V.: Optimal control problems with set-valued control and state constraints. SIAM J. Optim. **14**(2), 334–358 (2003). (Electronic)
18. Bonnans, J.F., Hermant, A.: No-gap second-order optimality conditions for optimal control problems with a single state constraint and control. Math. Program. **117**(1–2, Ser. B), 21–50 (2009)
19. Bonnans, J.F., Hermant, A.: Second-order analysis for optimal control problems with pure state constraints and mixed control-state constraints. Ann. Inst. Henri Poincaré, Anal. Non Linéaire **26**(2), 561–598 (2009)

20. Bonnans, J.F., Hermant, A.: Revisiting the analysis of optimal control problems with several state constraints. Control Cybern. **38**(4A), 1021–1052 (2009)
21. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer Series in Operations Research. Springer, New York (2000)
22. Robinson, S.M.: Stability theory for systems of inequalities. II. Differentiable nonlinear systems. SIAM J. Numer. Anal. **13**(4), 497–513 (1976)
23. Zowe, J., Kurcyusz, S.: Regularity and stability for the mathematical programming problem in Banach spaces. Appl. Math. Optim. **5**(1), 49–62 (1979)
24. Dmitruk, A.V.: Jacobi type conditions for singular extremals. Control Cybern. **37**(2), 285–306 (2008)
25. Hestenes, M.R.: Applications of the theory of quadratic forms in Hilbert space to the calculus of variations. Pac. J. Math. **1**, 525–581 (1951)
26. Ambrosio, L., Fusco, N., Pallara, D.: Functions of Bounded Variation and Free Discontinuity Problems. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York (2000)
27. Folland, G.B.: Real Analysis, 2nd edn. Pure and Applied Mathematics (New York). Wiley, New York (1999). Modern techniques and their applications, A Wiley-Interscience Publication
28. Choquet, G.: Cours d'Analyse. Tome II: Topologie. Espaces Topologiques et Espaces Métriques. Fonctions Numériques. Espaces Vectoriels Topologiques. Masson et Cie, Éditeurs, Paris (1969). Deuxième édition, revue et corrigée