2014

Reformulations for control systems and optimization problems with impulses

Jacob Blanton
Louisiana State University and Agricultural and Mechanical College

REFORMULATIONS FOR CONTROL SYSTEMS AND OPTIMIZATION PROBLEMS WITH IMPULSES

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Mathematics

by
Jacob R. Blanton
B.S., University of Central Florida, 2005
M.S., Louisiana State University, 2006
May 2014

This dissertation is dedicated to the memory of Corporal Aaron M. Griner, U.S. Army, whose only outstanding weakness of character was inimitable loyalty to those he cared for.

Acknowledgments

I would like to express my gratitude to my advisor Peter Wolenski for his patient guidance and encouragement throughout my time as a graduate student. I would also like to thank my committee members Charles Delzell, Richard Litherland, Frank Neubrander, Stephen Shipman, and Yejun Wu for their support, understanding, and advice. I would like to give particular thanks to Frank Neubrander for being perpetually helpful in much of the administrative aspects of my graduate study, and particularly for his happy cooperation within its final stages.

I would like to acknowledge LSU Mathematics, all its staff and faculty, for providing the opportunity to pursue my goals in such a helpful and friendly environment.

I would like to express my warmest gratitude to my parents who have always supported and encouraged my academic pursuits and to my sister who has shared this lifelong support with me. I am lucky to have too many friends to concisely list here. Their talents and virtues have inspired and often contributed directly to my education, but their warm companionships have graced it with purpose.

Contents

Abstract

This dissertation studies two different techniques for analyzing control systems whose dynamics include impulses, or more specifically, are measure-driven. In such systems, the state trajectories will have discontinuities corresponding to the atoms of the Borel measure driving the dynamics, and these discontinuities require further definition in order for the control system to be treated with the broad range of results available to non-impulsive systems. Both techniques considered involve a reparameterization of the system variables including state, time, and controls.

The first method is that of the graph completion, which provides an explicit reparameterization of the time and state variables. The reparameterization is continuous, which allows for the analysis of the system within classical control theory, yet it retains enough information about the discontinuous, or impulsive, trajectories that the results of such analyses may be interpreted for the original impulsive system. We utilize this reparameterization to formulate equivalent solution concepts between impulsive differential inclusions and impulsive differential equations. We also demonstrate that the graph completion is generally equivalent to a solution concept established for a neural spiking model, and make use of a specific such model as a numerical example.

The second method considered is similar to the graph completion but differs in that it utilizes implicit reparameterizations of all variables considered as families of functions which meet continuity and other requirements. This is particularly beneficial to optimal control problems as the choices of controls, impulsive and non-impulsive, may be varied within the optimization problem and analysis thereof. Necessary conditions for optimal control problems of Mayer form with fixed end time have been established under this reparameterization technique, and we extend these necessary conditions in a general context to a Mayer problem with free end time. Corollary to this, we deduce necessary conditions for a Bolza problem and a minimum time problem for impulsive control systems. Much of these results are obtained through reformulation techniques.

Chapter 1
Introduction

What happens when a ball tracing a smooth trajectory through the air is impacted by a racquet, or when a relay on an electronic control board is instantaneously engaged or disengaged? In what way are the population dynamics of a school of fish affected when a fraction of their population has been swiftly netted?

The theory of control is at this point a fully-developed field of mathematics, capable of providing insight and modelling to many types of problems which occur throughout nearly all of science and engineering, but these questions represent one of the many boundaries of this body of knowledge. Impulsive control theory is an active area of research seeking to provide fuller descriptions or more accurate mathematical models of phenomena involving shocks like those described above. One popular way of doing so is to reformulate a problem with impulses as one without, perform the analyses already available to non-impulsive control systems, and then translate these results back to the impulsive case. This and similar reformulations will be the paradigm of the present work.

The next chapter will introduce the general control system and provide some of the basic definitions from real, functional, and non-smooth analysis. It will also outline the classical cases of the optimal control problems studied later within the impulsive context.

We introduce in the third chapter a technique for handling impulses in control dynamics via time reparameterizations, called graph completions. This concept has its originations in the work [17] where the impulses enter the system as the product of a Radon measure with a state-independent term of the dynamics. The analysis of this case relies on the novel idea of a time reparameterization. An extension of this work was made in [13] to handle impulses that manifest as a product of a Radon measure and a state-dependent term via the concept of graph completions. Observe that it is not immediately clear how to define the product between the measure and state variable at an atom of the measure since the atom induces a jump discontinuity in the state at that point, and this definition is precisely what the graph completion introduced in [13] provides. Further work on graph completions is made in [5, 6] where different assumptions on the vector fields for impulsive systems are exploited.

The results of this dissertation in Chapter 3 also include the verification of a solution to an impulsive control differential inclusion as a solution to an impulsive control differential equation. This result hinges on a selection assumption similar to that provided by the Filippov Lemma. The differential inclusion solution concept used is that defined in [23]. An application of the graph completion to a neural spiking model provides a general result on defining solutions to such a model along with specific examples of this application.

Chapter 4 studies the necessary conditions for optimal control problems established in [2], and extends these results from a Mayer problem of fixed end-time to one of free end-time. Also, necessary conditions for a Bolza problem and a minimum time problem are deduced.

We conclude the dissertation in Chapter 5 with a summary of novel contributions and suggestions for future work ranging from immediate to remote.

Chapter 2
Preliminaries

In this chapter we introduce the basic concept of a control system without impulses, referred to throughout as the classical case or a general control system. We also study optimization problems of non-impulsive control systems and the Calculus of Variations and examine necessary conditions for each in the form of a Pontryagin Maximum Principle. This chapter should serve as background and motivation for the following chapters, but it is merely a summary and references are provided for further exploration of the topics described.

2.1 The control system

The central object to the theory of control is the system

$$\dot{x}(t) = f(t, x(t), u(t)), \qquad x(0) = \bar{x}, \tag{2.1}$$

where t is the scalar-valued time ranging on some interval $[0, T]$, x is the vector-valued *state* of the system, and u is the vector-valued *control*. It is common to refer to the function f as the *dynamics* or *right-hand side* of the system and to the second equality as the *initial condition*. As a whole, (2.1) is called a *control system*. By considering a control system whose right-hand side is independent of u, we get $f(t, x, u) = f(t, x)$, the right-hand side of a dynamical system, which suggests that in a general sense the theory of control subsumes the theory of dynamical systems.

What is more useful, is to realize that once a control function $u(\cdot)$ has been chosen, (2.1) is simply a differential equation and therefore more or less amenable to the extensive theory available in that subject. We sometimes refer to this differential equation and initial condition as a *Cauchy problem*. Remaining informal, we recall from the theory of differential equations that a *solution* or *trajectory* of (2.1) is a function $x(\cdot)$ which satisfies the Cauchy problem. This immediately raises questions of existence and uniqueness of such solutions, which have satisfactory answers within existing theory. In all that follows, the assumptions taken on the dynamics of systems under consideration will be sufficient to guarantee whatever degree of existence and uniqueness is appropriate for the problem, and these issues will be considered settled despite being unmentioned. Corresponding to each system considered there will be at least one solution definition. For more on solutions of ordinary differential equations we suggest [22], and for more on existence and uniqueness of trajectories for control systems we suggest [4] and [20].

It is informative, at least intuitively, to view the system (2.1) as the integral

$$x(t) = \bar{x} + \int_0^t f(\tau, x(\tau), u(\tau))d\tau \tag{2.2}$$

for time $t \in [0,T]$ and a chosen control function $u(\cdot)$. In this light, the fundamental theorem of calculus for Lebesgue integrals (cf. [14]) suggests what regularity the dynamics should have, namely that the composition of functions $f(t,x(\cdot),u(\cdot))$ be at least measurable, if not integrable. It is desirable in practical applications to take control functions which are measurable, often even piecewise continuous, and having values in some compact set, so in order for the integrand of (2.2) to be measurable it is sufficient to take f to be Lebesgue measurable in t and continuous in u. Partly for existence and uniqueness purposes, it is common to take f to be Lipschitz continuous in x, and altogether this is enough to imply that the integral of (2.2) exists. With this existence, the fundamental theorem of calculus yields absolute continuity of the function $x(\cdot)$; recall that the set of absolutely continuous functions is the most general set of functions satisfying the fundamental theorem. This is why it is common in classical control problems to consider solutions from the class of absolutely continuous functions. We sometimes refer to an absolutely continuous function which satisfies the Cauchy problem (2.1) almost everywhere as a *Caratheodory solution*.

Thus, the integral equation indicates needed regularity assumptions for the control system dynamics, and it indicates the class of functions in which solutions to the system will be found. In what follows, this integral perspective will play a key role in generalizing the notion of system and solution beyond that of the form (2.1) and its described solution.

2.2 Differential inclusions and nonsmooth analysis

It is convenient to consider the system (2.1) under a set of control functions u having values in a compact set $U \subset \mathbb{R}^m$ as the *differential inclusion*

$$\dot{x}(t) \in F(t,x(t)), \quad \text{a.e. } t \in [0,T], \tag{2.3}$$

with

$$x(0) = \bar{x},$$

where the object $F : \mathbb{R} \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a *multifunction*, a map from $\mathbb{R} \times \mathbb{R}^n$ into subsets of $\mathbb{R}^n$. The right hand side of (2.3) may be thought of as the set of possible velocities of the control system (2.1) and written as

$$F(t,x) = \{v : v = f(t,x,u) \text{ for some } u \in U\}.$$

Multifunctions in their own right, have a rich theory, cf. [3]. We will only be concerned with their basic forms of regularity, which can be referred to without providing exact definitions. It is obvious that a solution $x(\cdot)$ to the system (2.1), for a measurable control $u \in U$, satisfies the differential inclusion

$$\dot{x}(t) \in \{f(t,x(t),u(t))\}, \quad \text{for a.e. } t \in [0,T]$$

with $x(0) = \bar{x}$. Thus, we see that a trajectory of a control system provides a trajectory to a corresponding differential inclusion. For the reverse to be true, a trajectory of a differential inclusion to be a trajectory to a corresponding differential equation, an issue of selecting a measurable control u must be resolved. The well-known Filippov Lemma (cf. [10, 12]) provides that for a measurable multifunction F and a solution x of the differential inclusion (2.3), there exists a measurable $\hat{u}$ such that x satisfies (2.1) with right-hand side $f(t, x(t), \hat{u}(t))$. One of the results of this paper will be to provide this latter direction in the context of impulsive systems under graph completions.

Differential inclusions are particularly critical in optimal control problems, where it is generally true that problems with smooth objective functions can have nonsmooth solutions (cf. [12]). This lack of conventional differentiability and the practical desire to consider a broad class of optimal control problems necessitates the theory of nonsmooth analysis. We investigate the fundamental concepts of two separate but comparable theories which are required in the present work.

Clarke normal cones. We use the following definitions and facts regarding Clarke normal cones.

Definition 2.1. *Let $S \subset \mathbb{R}^n$ be nonempty and closed, and take $x \in S$. A vector $\zeta \in \mathbb{R}^n$ is a* ***proximal normal*** *to the set S at the point x if and only if there exists $\sigma = \sigma(x, \zeta) \geq 0$ such that*

$$\langle \zeta, u - x \rangle \leq \sigma |u - x|^2 \quad \forall u \in S \tag{2.4}$$

The set $N_S^P(x)$ of all such ζ defines the proximal normal cone to S at x.

Definition 2.2. *We define the* ***limiting normal cone*** *$N_S^L(x)$ to S at x by means of a closure operation applied to N_S^P:*

$$N_S^L(x) = \{\zeta = \lim_{i \to \infty} \zeta_i : \zeta_i \in N_S^P(x_i),\ x_i \to x,\ x_i \in S\}. \tag{2.5}$$

Recall that epi$\{f\}$, the epigraph of f, is the set $\{(x, \eta) : f(x) \leq \eta\}$.

Definition 2.3. *Given a lower semicontinuous function $f : \mathbb{R}^k \to \mathbb{R} \cup \{+\infty\}$ and a point $x \in \mathbb{R}^k$ such that $f(x) < \infty$, we define the* ***limiting subdifferential*** *of f at x, written $\partial f(x)$, to be*

$$\partial f(x) := \{\xi : (-1, \xi) \in N_{epi\{f\}}(f(x), x)\}. \tag{2.6}$$

We note that if f is Lipschitz continuous in a neighborhood of x, co$\partial\{f(x)\}$ coincides with the *generalized gradient* of f at x, where co is the convex hull operation. A direct definition of the generalized gradient, a proof of the equivalence claimed, and further details regarding nonsmooth analysis can be found in [10, 11, 12] among others.

Mordukhovich normal cones. The second and comparable theory we refer to involves the Mordukhovich normal cone. As our definition of the Mordukhovich normal cone for finite-dimensional space, we will take the characterization given by Theorem 1.6 in [16].

Definition 2.4. *For a closed set K in a finite-dimensional space X, we define the* Euclidean projector *of x on P as*

$$P_K(x) = \{w \in K : |x - w| = \operatorname{dist}(x, K)\},$$

where $\operatorname{dist}(x, K) := \inf_{y \in K} |x-y|$. *The* Mordukovhich normal cone *at $\bar{x} \in K$ is defined by*

$$N_K(\bar{x}) := \limsup \operatorname{cone}(x - P_K(x)) := \bigcap_{\epsilon>0} \operatorname{cl}\left[\bigcup_{x \in \bar{x}+B_X} \operatorname{cone}(x - P_K(x))\right],$$

where cl *and* cone *denote the closure and conic hull respectively. As a convention, we take $N_K(x) = \emptyset$ if $x \notin K$.*

Two more facts from [16] will be useful. First from proposition 1.5, the Mordukhovich normal cone of a convex subset K of a finite-dimensional space X is

$$N_K(\bar{x}) = \{y \in \mathbb{R}^m | \langle y, x - \bar{x}\rangle \leq 0 \quad \forall x \in K\}. \tag{2.7}$$

And second from proposition 1.3, if $\bar{x} = (\bar{x}_1, \bar{x}_2) \in K_1 \times K_2 \subset X_1 \times X_2$, then

$$N_{K_1 \times K_2}(\bar{x}) = N_{K_1}(\bar{x}_1) \times N_{K_2}(\bar{x}_2). \tag{2.8}$$

Let K be a convex subset of the finite-dimensional space X and take a scalar $\alpha > 0$. If for some $v \in X$ and $\bar{x} \in K$ we have $\alpha \cdot v \in N_K(\bar{x})$, then (2.7) implies for all $x \in K$

$$\langle \alpha \cdot v, x - \bar{x}\rangle \leq 0$$

which by homogeneity of the inner product is equivalent to

$$\langle v, x - \bar{x}\rangle \leq 0.$$

Thus,

$$\alpha \cdot v \in N_K(\bar{x}) \Leftrightarrow v \in N_K(\bar{x}), \quad \text{whenever} \quad \alpha > 0. \tag{2.9}$$

Also, it is a simple exercise to work out that $N_{\{\bar{x}\}}(x) \neq \emptyset$ if and only if $x = \bar{x}$, and

$$N_{\{\bar{x}\}}(\bar{x}) = X. \tag{2.10}$$

There is quite a bit of overlap between the theories of Mordukhovich and Clarke, which we need not consider. For more on the comparison between the two see [16].

2.3 Classical optimal control theory and calculus of variations

This section provides an outline of classical optimal control, by which we mean optimal control problems with impulse-free dynamics and continuous trajectories. Closely related to these problems is the well-known calculus of variations, which we reference here again as a body of problems and techniques precluding impulses. We follow [4] in presenting the Pontryagin Maximum Principle for the Mayer problem, fixed and free end time, and then deriving from this maximum principles for the Bolza problem, minimum time problem, and the calculus of variations. This suggests that in a sense, the necessary conditions of optimal control are a generalization of that of the calculus of variations problem. In fact, the proof of the first theorem we present below on the Mayer problem uses a construction known as a "needle variation" which is analogous to the variations studied in the Calculus of Variations.

Of particular note, this section serves as an outline and classical guide to the main results presented below in the chapter on optimal control of impulsive systems via measure-adjoint functions.

2.3.1 The Mayer problem with fixed end time and with free end time

We consider the optimal control problem in Mayer form

$$\max_{u \in \mathcal{U}} \ \phi_0(x(T, u)) \tag{2.11}$$

subject to

$$\dot{x} = f(t, x(t), u(t)), \quad x(0) = \bar{x}. \tag{2.12}$$

For a given set $U \subset \mathbb{R}^m$, the family of admissible controls is defined as

$$\mathcal{U} = \{u : [0, T] \to U; u \text{ measurable}\}. \tag{2.13}$$

In this case the terminal time T is fixed, and we put further constraints on the problem by requiring that the terminal state $x(T)$ satisfies

$$x(T) \in S = \{x \in \mathbb{R}^n : \phi_i(x) = 0, \ i = 1, ..., k\}. \tag{2.14}$$

We make the following assumptions for the classical case:

(H) The set $\Omega \subset \mathbb{R} \times \mathbb{R}^n$ is open, the function $f = f(t, x, u)$ is continuous on $\Omega \times U$ and continuously differentiable with respect to x. The functions $\phi_0, \phi_1, ..., \phi_k : \mathbb{R}^n \to \mathbb{R}$ are all continuously differentiable.

To an admissible control $u(\cdot)$ there corresponds a trajectory $x(\cdot)$ of (2.12), and the pair $(x(\cdot), u(\cdot))$ is referred to as a process. A process $(x^*(\cdot), u^*(\cdot))$ is called a maximizing process, or more generally an optimal process, if $x^*(\cdot)$ satisfies the maximality condition (2.11). Finding such an optimal process is a non-trivial problem, and there are a few techniques available to obtain such a process. In the present work we consider the use of necessary conditions in the form of the Pontryagin Maximum Principle to identify or narrow down the search for an optimal process.

Theorem 2.5. *Consider the optimal control problem (2.11)-(2.14), under the assumptions (H). Let $(x^*(\cdot), u^*(\cdot))$ be an optimal process, and assume the gradients $\nabla\phi_0, ..., \nabla\phi_k$ are linearly independent at the terminal point $x^*(T)$. Then there exists a nontrivial, absolutely continuous vector function $p : [0,T] \to \mathbb{R}^n$ which satisfies the equations*

$$\dot{p}(t) = -p(t) \cdot D_x f(t, x^*(t), u^*(t)), \tag{2.15}$$

$$p(t) \cdot f(t, x^*(t), u^*(t)) = \max_{\omega \in U} \{p(t) \cdot f(t, x^*(t), \omega)\} \tag{2.16}$$

for almost every time $t \in [0,T]$, together with the terminal conditions

$$p(T) = \sum_{i=0}^{k} \lambda_i \nabla\phi_i(x^*(T)) \tag{2.17}$$

for some constants $\lambda_0, ..., \lambda_k$, with $\lambda_0 \geq 0$.

We do not prove the theorem, but refer to [4] for a thorough treatment of the matter. Of more interest to our present consideration, is the extension of the Pontryagin Maximum Principle (PMP) for the given Mayer problem with terminal constraints and fixed end time to a Mayer problem with terminal constraints but a free or variable end time. The statement of such a problem is

$$\max_{u \in \mathcal{U}} \phi_0(T, x(T, u)) \tag{2.18}$$

for the control system described by

$$\dot{x}(t) = f(t, x(t), u(t)), \quad x(0) = (\bar{x}), \quad u(t) \in \mathbf{U} \quad a.e., \tag{2.19}$$

where the terminal time and the terminal point are subject to the constraints

$$\phi_i(T, x(T, u)) = 0, \quad i = 1, ..., k. \tag{2.20}$$

Note that an optimal policy of (2.18)-(2.20) is now a pair (T^*, u^*), where $u^* : [0, T^*] \to \mathbf{U}$ is measurable and the corresponding trajectory $x^*(\cdot)$ yields the maximum in (2.18) among all those satisfying (2.20). Here is the PMP for this problem.

Theorem 2.6. *Consider the optimal control problem (2.18)-(2.20), under the assumptions (H). Let $(x^*(\cdot), u^*(\cdot))$ be an optimal process for the problem where $u^* : [0, T^*] \to \mathbf{U}$ is a bounded control for the problem. Assume that f is continuously differentiable with respect to both t and x, and that the vectors $\nabla\phi_i = (\frac{\partial\phi_i}{\partial t}, \frac{\partial\phi_i}{\partial x_1}, ..., \frac{\partial\phi_i}{\partial x_n})$, $i = 1, ..., k$ are linearly independent at the point $(T^*, x^*(T^*))$. Then there exists a nontrivial, absolutely continuous row-vector $p(\cdot)$ such that*

$$\dot{p}(t) = -p(t) \cdot D_x f(t, x^*(t), u^*(t)), \tag{2.21}$$

$$p(t) \cdot f(t, x^*(t), u^*(t)) = \max_{\omega \in U} \{p(t) \cdot f(t, x^*(t), \omega)\} \tag{2.22}$$

at almost every time $t \in [0, T^*]$. Moreover, there exist constants $\lambda_0, ..., \lambda_k$ with $\lambda_0 \geq 0$ such that

$$(p_1, ..., p_n)(T^*) = \sum_{i=0}^{k} \lambda_i (\frac{\partial \phi_i}{\partial x_1}, ..., \frac{\partial \phi_i}{\partial x_n})(T^*, x^*(T^*)) \neq (0, ..., 0), \tag{2.23}$$

$$\max_{\omega \in \mathbf{U}} p(T^*) \cdot f(T^*, x^*(T^*), \omega) = -\sum_{i=0}^{k} \lambda_i \frac{\partial \phi_i}{\partial t}(T^*, x^*(T^*)). \tag{2.24}$$

Finally, the function $t \mapsto p(t) \cdot \dot{x}^*(t)$ in (2.22) coincides a.e. with an absolutely continuous function satisfying

$$\frac{d}{dt}\{p(t) \cdot f(t, x^*(t), u^*(t))\} = p(t) \cdot D_t f(t, x^*(t), u^*(t)). \tag{2.25}$$

This theorem is proved essentially by forming an auxiliary optimization problem in which an additional state, representing a reparameterization of time, is added to the problem statement and the previous fixed end time PMP is applied to this auxiliary problem. We sketch the proof here, and refer to [4] for a full proof.

We intend to apply Theorem 2.5 to an auxiliary problem in $n+1$ space variables with fixed terminal time T^*. Set $\mathbf{x} = (x_0, x) \in \mathbb{R}^{n+1}$, $\mathbf{u} = (u_0, u) \in \mathbb{R}^{m+1}$ and consider the problem

$$\max_{\mathbf{u}} \phi_0(\mathbf{x}(T^*, \mathbf{u}), \tag{2.26}$$

for the $(n+1)$-dimensional system

$$\begin{cases} \dot{x}_0(\tau) = u_0(\tau) \\ \dot{x}(\tau) = u_0(\tau) f(x_0(\tau), x(\tau), u(\tau)), \end{cases} \tag{2.27}$$

$$(x_0, x)(0) = (0, \bar{x}) \tag{2.28}$$

subject to the constraints

$$\phi_i(\mathbf{x}(T^*)) = 0, \quad i = 1, ..., k, \tag{2.29}$$

$$u_0(\tau) \in \left[\frac{1}{2}, 2\right], \quad u(\tau) \in \mathbf{U} \quad \text{for a.e. } \tau \in [0, T^*]. \tag{2.30}$$

From the chain rule, we get $\frac{dx}{dx_0} = \frac{dx}{d\tau}\frac{d\tau}{dx_0} = f(x_0, x, u)$, so x_0 plays the role of t from the original dynamics. It can also be shown that if $u^* : [0, T^*] \to \mathbf{U}$ is optimal for the original problem (2.18)-(2.20), then $\mathbf{u}^* = (1, u^*)$ is optimal for (2.26)-(2.30). This relies on the strict monotonicity, and thus invertibility, of x_0 implied by $\dot{x}_0(\tau) = u_0(\tau) \in [1/2, 2]$, which allows construction of an optimal control for the original problem based on an optimal control of the auxiliary problem. Details of this argument are in [4].

We now apply Theorem 2.5 to the optimal control $\mathbf{u}^* = (1, u^*)$ for the problem (2.18)-(2.20), and recalling that $x_0 = t$, we obtain the existence of an absolutely continuous adjoint vector $\mathbf{p} = (p_0, p)$ with the properties

$$\begin{cases} \dot{p}_0(t) = -p(t) \cdot D_t f(t, x^*(t), u^*(t)), \\ \dot{p}(t) = -p(t) \cdot D_x f(t, x^*(t), u^*(t)), \end{cases} \tag{2.31}$$

$$p_0(t) \cdot 1 + p(t) \cdot f(t, x^*(t), u^*(t)) = \max_{\frac{1}{2} \le \omega_0 \le 2,\ \omega \in \mathbf{U}} \{p_0(t)\omega_0 + p(t) \cdot \omega_0 f(t, x^*(t), \omega)\}, \tag{2.32}$$

$$(p_0(T^*), ..., p_n(T^*)) = \sum_{i=0}^{k} \lambda_i \left(\frac{\partial \phi_i}{\partial t}, \frac{\partial \phi_i}{\partial x_1}, ..., \frac{\partial \phi_i}{\partial x_n} \right) (T^*, x^*(T^*)), \tag{2.33}$$

for some constants $\lambda_0, ..., \lambda_k$ with $\lambda_0 \ge 0$. Since the maximum in (2.32) is attained when $\omega_0 \equiv 1$ and $\omega = u^*(t)$, we must have

$$\left. \frac{d}{d\omega_0} \right|_{(\omega_0, \omega) = (1, u^*(t))} [p_0(t)\omega_0 + p(t) \cdot \omega_0 f(t, x^*(t), \omega)] = 0, \tag{2.34}$$

which implies

$$p_0(t) = -p(t) \cdot f(t, x^*(t), u^*(t)) \quad \text{a.e.} \tag{2.35}$$

Since p_0 is absolutely continuous, from (2.35) and then from the first equation in (2.31), we derive

$$\frac{d}{dt}\{p(t) \cdot f(t, x^*(t), u^*(t))\} = \frac{d}{dt}\{-p_0(t)\} = p(t) \cdot D_t f(t, x^*(t), u^*(t)), \tag{2.36}$$

which yields (2.25).

The adjoint linear equation (2.21) follows from (2.31).

The maximality condition (2.22) follows from (2.32).

The terminal condition (2.23) follows from (2.33).

The identities (2.35) and (2.22) each hold almost everywhere and together imply

$$-p_0(t) = \max_{\omega \in \mathbf{U}} p(t) \cdot f(t, x^*(t), \omega) \qquad \forall t \in [0, T^*],$$

since the two sides are continuous. Taking this identity with that of (2.33), at $t = T^*$ yields (2.24), and thus completes the sketch of the proof.

2.3.2 The minimum time problem, Bolza problem, and calculus of variations

With the aid of the last theorem, we can establish maximum principles for other optimization problems.

Minimum time problem. If we take the function $\phi_0(T, x) = -T$ as the payoff function in problem (2.18)-(2.20) and define the target set as

$$S(T) := \{x : \phi_i(T, x) = 0,\ i = 1, ..., k\},$$

then we obtain the minimum time problem, which in words is the problem of reaching the set S in a minimal amount of time. Thus, the minimum time problem is a special case of the free end-time Mayer problem, and the maximum principle of Theorem 2.6 may be applied to it.

Notice the fact $\max -\phi_0 = -\min \phi_0$ has been used here. It is a common procedure to treat a minimization problem as a maximization problem through this reformulation, and more importantly it implies that the Pontryagin Maximum Principle is applicable to both types of problems. In the present work, discussions of necessary conditions for optimality will implicitly use this reformulation.

Bolza problem. The Bolza problem is given by

$$\min_{u\in\mathcal{U}} \int_0^T L(t, x(t), u(t))dt \tag{2.37}$$

subject to

$$\dot{x}(t) = f(t, x(t), u(t)), \quad x(0) = (\bar{x}), \quad u(t) \in \mathbf{U} \quad a.e., \tag{2.38}$$

with terminal constraints

$$\phi_i(T, x(T, u)) = 0, \quad i = 1, ..., k. \tag{2.39}$$

The scalar-valued function L, commonly called the *Lagrangian*, is assumed to be continuous in all variables and continuously differentiable with respect to t, x. With this assumption, we can define the auxiliary variable

$$x_{n+1}(t) = \int_0^t L(s, x(s), u(s))ds,$$

and thereby reformulate the Bolza problem as a Mayer problem. Indeed, the cost (2.37) is now given by

$$\min_{u\in\mathcal{U}} x_{n+1}(T),$$

subject to the terminal constraints (2.39) and the $n+1$-dimensional dynamics

$$\begin{cases} \dot{x}_i(t) = f_i(t, x(t), u(t)) & i = 1, ..., n \\ \dot{x}_{n+1}(t) = L(t, x(t), u(t)) \end{cases}$$

with initial condition $(x_1, ..., x_n, x_{n+1}) = (\bar{x}_1, ..., \bar{x}_n, 0)$.

An application of Theorem (2.6) to this problem establishes the PMP for the Bolza problem which is described in the following theorem.

Theorem 2.7. *Let f and L be continuous in all variables and continuously differentiable with respect to t, x. Let the bounded control $u^* : [0, T^*] \to \mathbf{U}$ be optimal for the problem (2.37)-(2.39) and assume that the vectors $(\frac{\partial \phi_i}{\partial t}, \frac{\partial \phi_i}{\partial x_1}, ... \frac{\partial \phi_i}{\partial x_n})(T^*, x^*(T^*)),\ i = 1, ..., k$, are linearly independent. Then there exist a nontrivial adjoint vector $p =$*

$(p_1, ..., p_n)$ and constants $\lambda_0, ..., \lambda_k$ with $\lambda_0 \geq 0$ such that, for almost every $t \in [0, T^*]$,

$$\dot{p}(t) = -p(t) \cdot D_x f(t, x^*(t), u^*(t)) - \lambda_0 \nabla_x L(t, x^*(t), u^*(t)),$$

$$\begin{aligned} p(t) \cdot f(t, x^*(t), u^*(t)) + \lambda_0 L(t, x^*(t), u^*(t)) = \\ \min_{u \in \mathbf{U}} \{p(t) \cdot f(t, x^*(t), u) + \lambda_0 L(t, x^*(t), u)\}, \end{aligned}$$

$$\begin{aligned} \frac{d}{dt}\{p(t) \cdot f(t, x^*, u^*) + \lambda_0 L(t, x^*, u^*)\} = \\ p(t) \cdot \frac{\partial f(t, x^*, u^*)}{\partial t} + \lambda_0 \frac{\partial L(t, x^*, u^*)}{\partial t}, \end{aligned}$$

$$(p_1, ..., p_n)(T^*) = \sum_{i=1}^{k} \lambda_i \left(\frac{\partial \phi_i}{\partial x_1}(T^*, x^*(T^*)), ..., \frac{\partial \phi_i}{\partial x_n}(T^*, x^*(T^*))\right),$$

$$\begin{aligned} \min_{u \in \mathbf{U}} \{p(T^*) \cdot f(T^*, x^*(T^*), u) + \lambda_0 L(T^*, x^*(T^*), u)\} \\ = -\sum_{i=1}^{k} \lambda_i \frac{\partial \phi_i}{\partial t}(T^*, x^*(T^*)). \end{aligned}$$

Calculus of Variations problem. The last classical problem we consider is the standard problem in the calculus of variations. It is given by

$$\min_{x(\cdot)} \int_0^T L(t, x(t), \dot{x}(t)) dt, \tag{2.40}$$

subject to

$$x(0) = \bar{x}, \quad x(T) = \bar{y}, \tag{2.41}$$

for $\bar{x}, \bar{y} \in \mathbb{R}^n$.

The first order necessary conditions for this problem, called the Euler-Lagrange and Weierstrass necessary conditions, can be derived from Theorem 2.7, and are given in the following theorem.

Theorem 2.8. *Assume that L is continuously differentiable in all variables $t, x, \dot{x}$. Let $x^*(\cdot)$ be a Lipschitz continuous function which attains the minimum for the problem (2.40)-(2.41). Then*

- *The function $t \mapsto \frac{\partial L}{\partial \dot{x}}(t, x^*(t), \dot{x}^*(t))$ coincides a.e. with an absolutely continuous function, such that*

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{x}}(t, x^*, \dot{x}^*)\right] = \frac{\partial L}{\partial x}(t, x^*, \dot{x}^*).$$

- *The function $t \mapsto L(t, x^*(t), \dot{x}^*(t)) - \frac{\partial L}{\partial \dot{x}}(t, x^*(t), \dot{x}^*(t)) \cdot \dot{x}_i^*$ coincides a.e. with an absolutely continuous function such that*

$$\frac{d}{dt}\left[L(t, x^*, \dot{x}^*) - \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{x}_i}(t, x^*, \dot{x}^*) \cdot \dot{x}_i^*\right] = \frac{\partial L}{\partial t}(t, x^*, \dot{x}^*).$$

- *For almost every $t \in [0,T]$ and every $\omega \in \mathbb{R}^n$,*

$$L(t, x^*(t), \omega) \geq L(t, x^*(t), \dot{x}^*(t)) + \frac{\partial L(t, x^*(t), \dot{x}^*(t))}{\partial \dot{x}} \cdot (\omega - \dot{x}^*(t)).$$

2.4 Fundamental definitions and results

We provide notation and basic definitions and facts to be used throughout the following.

2.4.1 Notation and definitions

Let us list here notations and conventions to be used.

$C([a,b], \mathbb{R}^n)$ denotes the vector space of continuous $\mathbb{R}^n$-valued functions on $[a,b]$ with supremeum norm, and $C^*([a,b], \mathbb{R}^n)$ its topological dual.

$C^m([a,b], \mathbb{R}^n)$ denotes the collection of $\mathbb{R}^n$-valued functions on $[a,b]$ which have continuous m^{th} derivative.

$C^+([a,b], \mathbb{R}^n) \subset C^*([a,b], \mathbb{R}^n)$ is the cone of functionals taking nonnegative values on nonnegative functions.

$AC([a,b], \mathbb{R}^n)$ is the space of absolutely continuous $\mathbb{R}^n$-valued functions on $[a,b]$.

$BV^+([a,b], \mathbb{R}^n)$ denotes the vector space of $\mathbb{R}^n$-valued functions on $[a,b]$, of bounded variation, which are continuous from the right on (a,b). The Borel measure associated with some $x \in BV^+([a,b], \mathbb{R}^n)$ is denoted dx. We use a similar notation for $BV^-([a,b], \mathbb{R}^n)$, the class of functions of bounded variation which are left-continuous. It will sometimes be necessary to restrict either of these classes further to being right continuous at 0.

The weak* topology on $BV^+([a,b], \mathbb{R}^n)$ refers to the weak* topology on $(\mathbb{R}^n \times C([a,b], \mathbb{R}^n))^*$ under the isomorphism

$$x \to (x(0), dx). \tag{2.42}$$

Consequently, "$x_i \to x$ (weakly*)" means that $x_i(0) \to x(0)$ and $dx_i \to dx$ (weakly* in $C^*([a,b], \mathbb{R}^n)$).

$\mathcal{L}$ denotes the Lebesgue subsets of $[a,b]$, $\mathcal{B}$ the Borel subsets in $\mathbb{R}^k$, and $\mathcal{L} \times \mathcal{B}$ the product σ-field. The notation $\mathcal{B}([a,b])$ may in places be used to denote the collection of Borel subsets of $[a,b]$; clarification will be provided when necessary.

2.4.2 Functions of bounded variation and regular Borel measures

Here we recall some important properties of functions of bounded variation and introduce the associated notation.

Definition 2.9. *A function $u : [0,T] \to \mathbb{R}^m$ is said to be of* bounded total variation *or just* bounded variation *on the subinterval $[a,b] \subseteq [0,T]$ if there exists a constant $C \geq 0$ such that, for each finite set of points $t_0, ..., t_k$ satisfying*

$$a = t_0 < t_1 < ... < t_k = b,$$

the inequality

$$\sum_{j=1}^{k} |u(t_j) - u(t_{j-1})| \leq C$$

holds, where $|\cdot|$ *denotes the Euclidean norm. The least* C *which satisfies the above condition is called the variation of* u *on* $[a, b]$*, and it is denoted by* $V_a^b(u)$*. The number* $V_0^T(u)$ *will be called the* total variation *of* u*.*

As it will be used throughout the following, we mention that the left and right hand limits of a function u at a point t will be written as $u(t^-) = \lim_{s \to t^-} u(s)$ and $u(t^+) = \lim_{s \to t^+} u(s)$, respectively.

In order to simplify the analysis to follow, we will in places work with the subset of functions of bounded total variation which are one-side continuous. For instance, recall $BV^-([0, T], \mathbb{R}^m)$ defined above as the set of all functions of bounded variation on $[0, T]$ taking values in $\mathbb{R}^m$ which are left continuous on $(0, T]$ and right continuous at 0. Recall that functions which have bounded variation on an interval $[a, b]$ have only jump-type discontinuities on that interval and their one-sided limits exist at all points in the interval [14, 15], so the results which follow these restrictions can be extended with proper modifications to the broader class of functions of bounded total variation. This is shown below. Let us point out that this choice is arbitrary and in the present work and other literature, [19, 23, 24] for example, the subset of right continuous functions in $BV([0, T], \mathbb{R}^m)$ is sometimes chosen instead.

Recall that for every $u \in BV([0, T], \mathbb{R}^m)$ the distributional derivative $\dot{u}$ is an $\mathbb{R}^m$-valued, regular Borel measure on $(0, T)$. If $u \in BV^-([0, T], \mathbb{R}^m)$, then $\dot{u}$ is characterized by the equality

$$\dot{u}([t_1, t_2)) = u(t_2) - u(t_1)$$

for every subinterval $[t_1, t_2) \subseteq (0, T)$. For every $t \in (0, T)$,

$$\dot{u}(\{t\}) = \triangle u(t) := u(t^+) - u(t^-).$$

The integral of a function $f : [0, T] \to \mathbb{R}$ with respect to the measure $\dot{u}$ on a Borel subset E of $(0, T)$ may be denoted by

$$\int_E f\dot{u},$$

or by

$$\int_E f du,$$

where du is as defined in section 2.4.1.

Lastly, recall that if $u \in BV^-([0, T], \mathbb{R}^m)$, then, for every subinterval $[t_1, t_2) \subseteq (0, T)$, we have

$$|\dot{u}|([t_1, t_2)) = V_{t_1}^{t_2}(u),$$

where $|\dot{u}|$ is the total variation of the measure $\dot{u}$. Furthermore,

$$|\dot{u}|((0,T)) = V_0^T(u),$$

for every $u \in BV^-([0,T],\mathbb{R}^m)$.

Since we can associate the map $u \in BV([0,T],\mathbb{R}^m)$ with the left continuous map $u^- \in BV^-([0,T],\mathbb{R}^m)$ defined by

$$u^-(t) = u(t^-), \quad t \in (0,T], \quad u^-(0) = u(0^+)$$

and show that $u^- = u$ almost everywhere, we have that $\dot{u}^- = \dot{u}$ as measures on $(0,T)$. Thus, our choice of left continuous functions does not result in a loss of generality.

Chapter 3

Graph completion methods for impulsive control problems

The theory of graph completions has its origins in the work [17], where a time reparameterization is used to handle impulses. Further development of this concept was made in the work [13], where the time reparameterization concept is extended to control dynamics which involve the multiplication of a state-dependent term with a vector-valued impulse. The key insight in this step is the concept of a graph completion which we will consider below. Other developments of the graph completion occur in [5, 6], where the cases of commutative and noncommutative vector fields for impulsive systems are considered separately. A solution concept for impulsive differential inclusions is set forth in [18], of which the companion paper [19] established a maximum principle for a fixed end-time Mayer problem. An equivalent solution concept for the impulsive inclusion is provided by [23], which the current work uses to provide a relationship between solutions of impulsive differential inclusions and those of impulsive differential equations analogous to the discussion in section 2.2.

As an example and a source of new results, we apply graph completion solution techniques to the neural spiking model of [8].

3.1 Impulsive control differential equations

Consider the Cauchy problem

$$\begin{cases} \dot{x}(t) = f(x(t)) + \sum_{i=1}^{m} g_i(x(t))\dot{u}^i(t), \quad t \in [0,T], \\ x(0) = \bar{x}, \end{cases} \tag{3.1}$$

where the $m+1$ vector fields $f, g_1, ..., g_m$ from $\mathbb{R}^n$ into $\mathbb{R}^n$ are assumed to be globally bounded and continuously differentiable on $\mathbb{R}^n$ and the control $u = (u_1, ..., u_m) : [0,T] \to \mathbb{R}^m$ is vector-valued. Notice u enters into the equation only through its derivative $\dot{u}$. If u is absolutely continuous on $[0,T]$, then (3.1) admits a unique absolutely continuous solution in the sense of Caratheodory.

Allowing u to be a function of bounded variation on $[0,T]$ and such that u is continuous at 0 and T has the effect of introducing impulsive dynamics into (3.1) as the term $\dot{u}$ must then be interpreted as a distributional derivative or a measure which potentially has atoms $t \in (0,T)$. This will in general cause trajectories of the system to have jumps. The main difficulty faced in [13] is handling, and in particular defining, the multiplication of $\dot{u}_i$ and g_i at an atom t of $\dot{u}_i$ when g_i depends on the state variable x. The authors there introduce a solution concept in the sense of measure to (3.1) which includes a rule for evolving g_i along the jumps produced by atoms of $\dot{u}_i$ and thereby defines their product in a natural way. Then an auxiliary system is formed in the augmented state variable y through the use of a function φ, called a *graph completion*, which connects or fills in the jumps of the function u in a Lipschitz continuous manner. This auxiliary system turns out to have a unique absolutely continuous solution in the sense of Caratheodory, and moreover, this solution equals that of the measure equation for almost every $t \in [0,T]$ through the inverse, or psuedo-inverse, of the time component φ_0 of the graph completion, that is, $x = y(\varphi_0^{-1})$.

In what follows, we provide the details to the above summary stating needed results of [13] but leaving the proofs to be consulted therein.

3.1.1 A solution in the sense of measure

We now define a solution to (3.1) in the sense of measure. The first task is to handle the multiplication of g_i and $\dot{u}_i$ at the atoms of $\dot{u}_i$.

Let $(z,p) \in \mathbb{R}^n \times \mathbb{R}^m$. For every $i = 1, ..., m$, define the function $\tilde{g}_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ as

$$\tilde{g}_i(z,p) = \int_0^1 g_i(\exp(\sigma \sum_{j=1}^m p^j g_j)z)\ d\sigma, \tag{3.2}$$

where $\exp(\sigma \sum_{j=1}^m p^j g_j)z$ denotes the value at time $s = \sigma$ of the solution to the Cauchy problem

$$\frac{dw}{ds} = \sum_{j=1}^m p^j g_j(w(s)), \quad w(0) = z. \tag{3.3}$$

Now observe

$$\begin{aligned} \sum_{i=1}^m p^i \tilde{g}_i(z,p) &= \int_0^1 \sum_{i=1}^m p^i g_i(\exp(\sigma \sum_{j=1}^m p^j g_j)z)\ d\sigma \\ &= \exp(\sigma \sum_{i=1}^m p^i g_i)z \Big|_0^1 \\ &= \exp(\sum_{i=1}^m p^i g_i)z - z \end{aligned}$$

by definition (3.2), then by the fact that $\exp(\cdot)$ solves (3.3), and lastly by evaluation. Thus we have

$$\sum_{i=1}^m p^i \tilde{g}_i(z,p) = \exp(\sum_{i=1}^m p^i g_i)z - z, \tag{3.4}$$

which will be a useful fact in what follows.

Now consider for $t \in (0,T)$ the Cauchy problem,

$$\dot{x} = f(x) + \sum_{i=1}^m \tilde{g}_i(x(t^-), \dot{u}(\{t\}))\dot{u}^i, \quad x(0^+) = \bar{x}, \tag{3.5}$$

and observe that (3.5) admits a Caratheodory solution if the measure $\dot{u}$ has no atoms. We can now define a solution to this problem.

Definition 3.1. *Let $u \in BV^-([0,T],\mathbb{R}^m)$. A solution of (3.5) is a map $x \in BV^-([0,T],\mathbb{R}^n)$ which satisfies (3.5) in the sense of measures on $(0,T)$. That is, x satisfies*

$$\int_B \dot{x} = \int_B f(x(t))dt + \sum_{i=1}^m \int_B \tilde{g}_i(x(t^-),\dot{u}(\{t\}))\dot{u}^i$$

for every Borel subset B of $(0,T)$.

3.1.2 Solutions via graph completions

We can view the problem given by (3.1) from a different perspective by associating with it the $n+1$-dimensional Cauchy problem

$$\begin{cases} \frac{dy_0(s)}{ds} = \frac{d\varphi_0(s)}{ds}, & y_0(0)=0, \qquad s\in[0,1] \\ \frac{dy(s)}{ds} = f(y(s))\frac{d\varphi_0(s)}{ds} + \sum_{i=1}^m g_i(y(s))\frac{d\varphi_i(s)}{ds}, & y(0)=\bar{x}, \end{cases} \tag{3.6}$$

where $\tilde{y}=(y_0,y):[0,1]\to[0,T]\times\mathbb{R}^n$, φ_0 is a non-decreasing map from $[0,1]$ onto $[0,T]$, and $\varphi=(\varphi_1,...,\varphi_m)$ is an m-dimensional control from $[0,1]$ into $\mathbb{R}^m$.

For $u\in C^1([0,T],\mathbb{R}^m)$, (3.1) has a unique solution on $[0,T]$. If $\varphi:[0,1]\to\mathbb{R}^m$ is any C^1 reparametrization of the graph of u, meaning $\varphi_0\in C^1([0,1])$, $\varphi_0(0)=0$, $\varphi_0(1)=T$, $d\varphi_0(s)/ds\geq 0$, and $u_i(\varphi_0(s))=\varphi_i(s)$ for every $s\in[0,1]$, then it is simple to show that the solution $\tilde{y}=(y_0,y)$ of (3.6) corresponding to φ satisfies $x(\varphi_0(s))=y(s)$ for every $s\in[0,1]$. Thus

$$x(t)=y(\varphi_0^{-1}(t)) \quad \forall t\in[0,T],$$

so long as φ_0 is strictly increasing.

If u does not possess this regularity, say we let u be a function of bounded total variation on $[0,T]$, then the above argument no longer suffices. Nonetheless, it serves as guidance and motivation for the construction of a Lipschitz continuous function, φ in the above, which reparametrizes the graph of u in such a way that (3.6) has classical solutions. Moreover, this graph completion will allow us to establish a relationship between these classical solutions and the solutions of (3.5). We begin with the precise definition of a graph completion.

Definition 3.2. (Graph Completion). *Consider any function $u:[0,T]\to\mathbb{R}^m$. A Lipschitz continuous path $\varphi=(\varphi_0,\varphi_1,...,\varphi_m):[0,S]\to[0,T]\times\mathbb{R}^m$ is a* graph completion *of u if*

- $\varphi(0)=(0,u(0)),\ \varphi(S)=(T,u(T)),$
- $\varphi_0(s_1)\leq\varphi_0(s_2)$ *for all* $0\leq s_1\leq s_2\leq S,$
- *for each $t\in[0,T]$ there exists some s such that $\varphi(s)=(t,u(t))$.*

Notice that in the above definition and in other places the notation for the graph completion (φ_0, φ) may be compressed as $\varphi = (\varphi_0, ..., \varphi_m)$ for convenience. It will be apparent when this is the case, and unless otherwise indicated φ should be assumed to be only the m-vector portion of the graph completion.

The path φ must provide a continuous parameterization of the graph of u, and in particular, it must form an arc across any jump discontinuities that u may have. The lemma below, from [4], indicates the class of functions for which a graph completion can be constructed.

Lemma 3.3. *A graph completion of u exists if and only if u has bounded total variation.*

Proof. Step 1. Let $\varphi = (\varphi_0, ..., \varphi_m)$ be a graph completion of u. For any finite sequence $0 = t_0 < t_1 < ... < t_k = T$ we can choose parameter values $0 = s_0 < s_1 < ... < s_k = S$ such that $\varphi_0(s_j) = t_j$. We then have

$$\sum_j |u(t_j) - u(t_{j-1})| \leq \sum_j |\varphi(s_j) - \varphi(s_{j-1})| \leq \int_0^S |\dot{\varphi}(s)| \, ds.$$

Taking the supremum over all increasing sequences $t_0 < t_1 < ... < t_k, \ k \geq 1$, we obtain

$$V_0^T(u) \leq \int_0^S |\dot{\varphi}(s)| ds < \infty$$

since φ is Lipschitz continuous.

Step 2. Conversely, assume that the control function $u(\cdot)$ has bounded total variation. This means u is continuous almost everywhere with at most countably many points of jump discontinuities (cf. [15]). Moreover, it admits left and right limits $u(\tau^-)$, $u(\tau^+)$ at every time τ. We will construct a graph-completion of u by bridging each of its jumps with a straight segment. For each $\tau \in [0, T]$, consider the total variation of u restricted to the subinterval $[0, \tau]$

$$V_0^\tau(u) = \sup_{0 \leq t_0 < t_1 < ... < t_N \leq \tau} \sum_{j=1}^N |u(t_j) - u(t_{j-1})|.$$

Set $S = T + V_0^T(u)$ and define the path $\varphi : [0, S] \to [0, T] \times \mathbb{R}^m$ as follows.

The map $t \mapsto t + V_0^t(u)$ is strictly increasing, given $s \in [0, S]$ there exists exactly one $\tau \in [0, T]$ such that $\tau + V_0^{\tau^-}(u) \leq s \leq \tau + V_0^{\tau^+}(u)$. We consider the following cases:

- If $s = \tau + V_0^\tau(u)$, set $\varphi(s) = (\tau, u(\tau))$. This happens if u is continuous at τ.
- If $\tau + V_0^{\tau^-}(u) \leq s \leq \tau + V_0^\tau(u)$, say $s = \theta[\tau + V_0^\tau(u)] + (1 - \theta)[\tau + V_0^{\tau^-}(u)]$ for some $\theta \in [0, 1]$, set

$$\varphi(s) = (\tau, \theta u(\tau) + (1 - \theta) u(\tau^-)).$$

- If $\tau + V_0^\tau(u) \le s \le \tau + V_0^{\tau^+}(u)$, say $s = \theta[\tau + V_0^{\tau^+}(u)] + (1-\theta)[\tau + V_0^\tau(u)]$ for some $\theta \in [0,1]$, set

$$\varphi(s) = (\tau, \theta u(\tau^+) + (1-\theta)u(\tau)).$$

It is simple to check that the above contruction satisfies all conditions required by the definition of graph completion. In particular, observe that the map $s \mapsto \varphi(s)$ is Lipschitz continuous with constant $L = 1$.

□

It is important to notice that trajectories of the auxiliary system (3.6) determined by a given graph completion φ of u will depend on the path by which φ connects the graph of u and that distinct graph completions will yield distinct trajectories. However, the trajectory determined by a given graph completion is independent of the time parametrization (cf. [4]). The dependence of trajectories on graph completions permits some degree of freedom in deciding the behavior of trajectories at jumps. It is our purpose to establish solutions of (3.6) which coincide through the inverse of φ_0 (at least where this inverse exists) with solutions of (3.5), so we introduce the canonical graph completion.

Definition 3.4. (Canonical Graph Completion). *Let u belong to $BV^-([0,T], \mathbb{R}^m)$, and set*

$$\eta(t) := \frac{t + V_0^t(u)}{T + V_0^T(u)}, \qquad t \in [0,T], \tag{3.7}$$

where $V_0^t(u)$ is the total variation of the function u on the interval $[0,t]$. Call η the reparameterization function*. The* canonical graph completion *φ of u is defined by*

$$\varphi(s) := \begin{cases} (t, u(t)) & \text{if } s = \eta(t), \\ (t, u(t) + \frac{s-\eta(t)}{\triangle\eta(t)} \triangle u(t)) & \text{if } s \in (\eta(t), \eta(t^+)), \end{cases} \tag{3.8}$$

where $\triangle u(t) := u(t^+) - u(t^-)$.

The canonical graph completion φ of u is the Lipschitz continuous map which reparametrizes the graph of u by connecting each jump in the graph with the shortest line segment connecting the points on either side of the jump. Some easily verified properties of φ which will later be useful are as follows:

- $\triangle\eta(t) = \frac{|\triangle u(t)|}{T + V_0^T(u)}$;

- $\frac{d\varphi}{ds}(s) = (0, \frac{\triangle u(t)}{\triangle \eta(t)}), \quad \forall s \in (\eta(t), \eta(t^+))$;

- φ is Lipschitz continuous with constant $T + V_0^T(u)$.

We now state without proof one of the main results Theorem 2.2 of [13], which provides the link between solutions of the system (3.5) and solutions of the auxiliary system (3.6). The proof of this theorem is in the mentioned article, and it is worth noting that part of this proof will be extended to the proof of Theorem 3.11 of the present work.

Theorem 3.5. *Let u belong to $BV^-([0,T],\mathbb{R}^m)$. Then a map $x \in BV([0,T],\mathbb{R}^n)$ is a solution of (3.5) if and only if there exists a solution $\tilde{y} = (y_0, y)$ of (3.6) corresponding to the canonical graph completion φ such that*

$$x(t) = y(\eta(t)) \tag{3.9}$$

for almost every $t \in (0,T)$.

The proof of this theorem requires the use of Volpert's averaged superposition along with a related theorem, which we state for later use in our own analysis. For more on this concept see [21], and for the proof of the theorem see [1].

Let $A : \mathbb{R}^p \to \mathbb{R}^q$ be a bounded Borel function, and let v belong to $BV([0,T],\mathbb{R}^p)$. The function $\hat{A}(v) : [0,T] \to \mathbb{R}^q$ defined by

$$\hat{A}(v)(t) := \int_0^1 A(v(t^-) + \sigma(v(t^+) - v(t^-))d\sigma = \int_0^1 A(\sigma v(t^+) + (1-\sigma)v(t^-))d\sigma$$

is called the average superposition of A and v.

Observe that $\hat{A}(v)(t) = A(v(t))$ for each t at which v is continuous.

Theorem 3.6. *Let $\psi : [0,1] \to \mathbb{R}^n$ be a Lipschitz continuous function and let $z \in BV([0,T],[0,1])$. If the map α is defined by*

$$\alpha(t) := \psi(z(t)) \quad \forall t \in [0,T],$$

then

- *$\alpha \in BV([0,T],\mathbb{R}^n)$;*
- *the identity of measures $\dot{\alpha} = \hat{\psi}_*(z)\dot{z}$ holds, where ψ_* denotes any Borel function coinciding with the derivative $d\psi/ds$ almost everywhere with respect to Lebesgue measure.*

3.1.3 Example: A one-dimensional system with impulse

We consider a one dimensional impulsive differential equation and determine its solution via the canonical graph completion technique described above. The equation with initial value is

$$\frac{dx}{dt} = cx\delta(t - t_*), \qquad x(0) = x_0, \tag{3.10}$$

where $c > 0$, $t \in [0,T]$, and $t_* \in (0,T)$. This is the initial value problem first considered in [8] with the time interval here being finite, a necessity for the graph completion, and with state variable x replacing the variable u. We will see that the finiteness of the time interval will ultimately make no difference in the solution established through the graph completion.

Recall that the Dirac δ-function is the generalized function with the following properties:

$$\delta(t) = \begin{cases} +\infty & t = 0, \\ 0 & t \neq 0, \end{cases}$$

and

$$\int_{-\infty}^{t} \delta(\tau)d\tau = \begin{cases} 0 & t \leq 0, \\ 1 & t > 0. \end{cases} \tag{3.11}$$

We have chosen δ to have a left-continuous integral to be congruent with the class of impulse functions already considered, but again this is merely a matter of choice and right-continuity could be chosen just as well. The right hand side of (3.11) is often denoted

$$H(t) := \begin{cases} 0 & t \leq 0, \\ 1 & t > 0, \end{cases} \tag{3.12}$$

which is the left-continuous version of the Heaviside function.

For the impulse function used in [8], $\delta(t-t_*)$ for $t_* \in (0,T)$, we derive the graph completion

$$\varphi(s) = (\varphi_0(s), \varphi_1(s)) = \begin{cases} (s(T+1), 0) & \text{if } 0 \leq s \leq \frac{t_*}{T+1}, \\ (\frac{t_*}{T+1}, (T+1)s - t_*) & \text{if } \frac{t_*}{T+1} < s < \frac{t_*+1}{T+1}, \\ (s(T+1)-1, 1) & \text{if } \frac{t_*+1}{T+1} \leq s \leq 1, \end{cases} \tag{3.13}$$

which has derivative

$$\dot{\varphi}(s) = (\dot{\varphi}_0(s), \dot{\varphi}_1(s)) = \begin{cases} (T+1, 0) & \text{if } 0 \leq s \leq \frac{t_*}{T+1}, \\ (0, T+1) & \text{if } \frac{t_*}{T+1} < s < \frac{t_*+1}{T+1}, \\ (T+1, 0) & \text{if } \frac{t_*+1}{T+1} \leq s \leq 1. \end{cases} \tag{3.14}$$

We form the auxiliary system for this problem according to (3.6) to get

$$\begin{cases} \dot{y}_0(s) = \dot{\varphi}_0(s) \\ \dot{y}(s) = cy(s)\dot{\varphi}_1(s), \quad s \in [0,1] \end{cases} \tag{3.15}$$

with initial condition

$$(y_0(0), y(0)) = (0, x_0).$$

This initial value problem has solution

$$(y_0(s), y(s)) = \begin{cases} ((T+1)s, x_0) & 0 \le s \le \frac{t_*}{T+1} \\ (t_*, x_0 e^{c(T+1)(s-\frac{t_*}{T+1})}) & \frac{t_*}{T+1} \le s \le \frac{t_*+1}{T+1} \\ ((T+1)s - 1, x_0 e^c) & \frac{t_*+1}{T+1} \le s \le 1. \end{cases}$$

According to Theorem 3.5, we regain a solution x of bounded variation to (3.10) in the sense of measure from the relation $x(t) = y(\eta(t))$. Hence,

$$x(t) = \begin{cases} x_0, & t \in [0, t_*] \\ x_0 e^c, & t \in (t_*, T], \end{cases} \tag{3.16}$$

which matches the solution to the example considered in [8] by letting the final time T be arbitrarily large.

3.2 Application to a neural spiking model

We turn our attention to the neural spiking model presented by [8]. Electrochemical signals are passed between two neurons at a junction called a synapse in the form of chemical neurotransmitters which cross the synapse from the presynaptic neuron to the postsynaptic neuron. The neurotransmitter is stored in vesicles, and when the presynaptic neuron fires, some of the vesicles filled with neurotransmitter move to the edge of the neuron to release neurotransmitter into the synapse. The model considers the vesicles to be in one of three states: available, active, and recovering with population fractions denoted by x, y, and z, respectively. There is a phenomenon whereby impulses sent in sufficiently rapid succession actually facilitate a series of spikes each with a signal or voltage spike larger than the previous one. This effect will be represented by the state variable u. The full model is given as the following system of impulsive ODEs:

$$\begin{aligned} \frac{dx}{dt} &= \frac{z}{\tau_{\text{rec}}} - \delta(t - t_*)xu, \\ \frac{dy}{dt} &= -\frac{y}{\tau_{\text{in}}} + \delta(t - t_*)xu, \\ \frac{dz}{dt} &= \frac{y}{\tau_{\text{in}}} - \frac{z}{\tau_{\text{rec}}}, \\ \frac{du}{dt} &= -\frac{u}{\tau_{\text{facil}}} + \delta(t - t_*)k(1 - u). \end{aligned} \tag{3.17}$$

This system is analyzed in [8] through the use of a sequence of continuous functions which converge to the δ function. Plots of the active vesicle state are given for some specific parameter values along with a plot of a numerical solution to the system for certain parameter values. We establish comparable results through the use of the graph completion techniques described above and provide the corresponding plots obtained from our analysis. The authors in [8] first consider the decoupled equation for u, so we follow the same steps to establish a thorough comparison.

3.2.1 Result: A solution to the model via the canonical graph completion

Facilitated state. We first consider the fourth equation in the system of differential equations given above, taking the parameter $\tau_{\text{facil}} = 1$. Observe that this equation decouples from the remaining system, and as we will show yields an explicit solution which may then be substituted back into the whole system. The initial value problem is

$$\begin{cases} \frac{du}{dt} = -u + k(1-u)\delta(t-t_*) \\ u(0) = \bar{u},\ t \in [0,T]. \end{cases} \tag{3.18}$$

We use the canonical graph completion φ of δ along with the auxiliary system formulation given by (3.6) to write

$$\begin{cases} \dot{U}_0(s) = \dot{\varphi}_0(s) \\ \dot{U}(s) = -U(s)\dot{\varphi}_0(s) + k(1-U(s))\dot{\varphi}_1(s) \\ (U_0,U)(0) = (0,\bar{u}), \quad s \in [0,1], \end{cases} \tag{3.19}$$

which, with the explicit values of $\dot{\varphi}$ from (3.14), is

$$(\dot{U}_0(s), \dot{U}(s)) = \begin{cases} (T+1, -(T+1)U(s)) & s \in [0, \frac{t_*}{T+1}] \\ (0, k(T+1)(1-U(s))) & s \in [\frac{t_*}{T+1}, \frac{t_*+1}{T+1}] \\ (T+1, -(T+1)U(s)) & s \in [\frac{t_*+1}{T+1}, 1]. \end{cases} \tag{3.20}$$

This system has solution

$$(U_0(s), U(s)) = \begin{cases} ((T+1)s,\ \bar{u}e^{-(T+1)s}) & s \in [0, \frac{t_*}{T+1}] \\ (t_*,\ 1 + e^{-k((T+1)s-t_*)}(\bar{u}e^{-t_*}-1)) & s \in [\frac{t_*}{T+1}, \frac{t_*+1}{T+1}] \\ ((T+1)s-1,\ e^{(t_*+1)-(T+1)s}(1+(\bar{u}e^{-t_*}-1)e^{-k})) & s \in [\frac{t_*+1}{T+1}, 1]. \end{cases}$$

Once again we use the rule $u(t) = U(\eta(t))$ to recover, for the original system in u, the solution

$$u(t) = \begin{cases} \bar{u}e^{-t*} & t \leq t_* \\ (1-(1-\bar{u}e^{-t_*})e^{-k})e^{-(t-t_*)} & t > t_*. \end{cases}$$

By letting T be arbitrarily large, this solution matches the solution given in [8], and it follows that the same jump condition $u(t_*^+) = u(t_*^-) + (1 - e^{-k})(1 - u(t_*^-))$ is obtained from this solution.

Full model. Again taking φ to be the canonical graph completion of δ and appealing to the formulation of (3.6), we obtain for (3.17) the auxiliary system

$$\begin{aligned}
\frac{dX_0}{ds} &= \dot{\varphi}_0(s) \\
\frac{dX}{ds} &= \frac{Z(s)}{\tau_{\text{rec}}}\dot{\varphi}_0(s) - X(s)U(s)\dot{\varphi}_1(s) \\
\frac{dY}{ds} &= -\frac{Y(s)}{\tau_{\text{in}}}\dot{\varphi}_0(s) + X(s)U(s)\dot{\varphi}_1(s) \\
\frac{dZ}{ds} &= \left(\frac{Y(s)}{\tau_{\text{in}}} - \frac{Z(s)}{\tau_{\text{rec}}}\right)\dot{\varphi}_0(s) \\
\frac{dU}{ds} &= \frac{-U(s)}{\tau_{\text{facil}}}\dot{\varphi}_0(s) + k(1 - U(s))\dot{\varphi}_1(s).
\end{aligned} \tag{3.21}$$

The values of $(\dot{\varphi}_0, \dot{\varphi})(s)$ may be substituted into (3.21), and the resulting system may be solved numerically. From this we recover the numerical solution to the original system in time t through the rule $x(t) = X(\eta(t))$. Figure 3.1 is a plot of this numerical solution for the original system in time t.

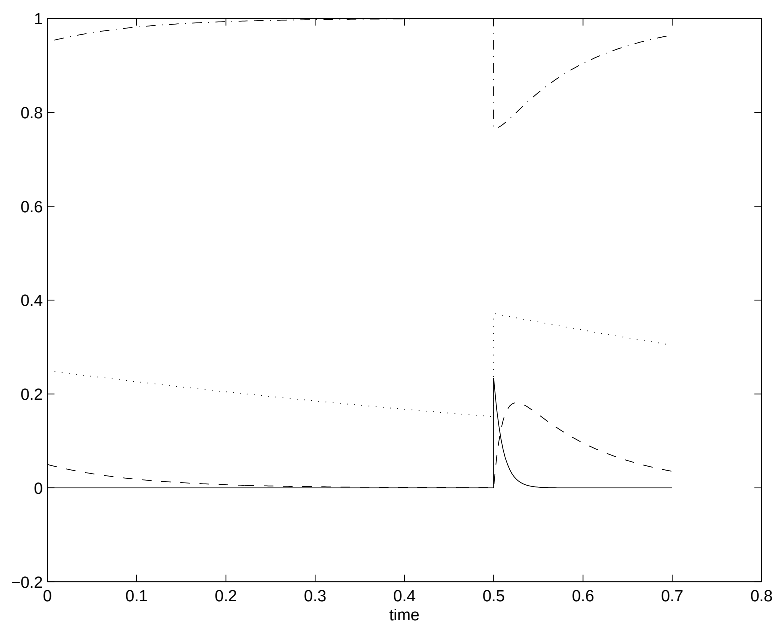

FIGURE 3.1. A plot of the solution to the neural spiking model (3.17) in the original states x, y, z, u and time t regained from the numerical solution to the reparameterized system (3.21). The dotted and dashed line is x, the solid line is y, the dashed line is z, and the dotted line is u. Parameter values are $\tau_{in} = 0.01$,$\tau_{rec} = 0.1$,$\tau_{facil} = 1.0$, and $k = 0.3$.

Thus, we see the flexibility of the reparameterization technique, for we have considered at least one case where the canonical graph completion is applied to a system to yield an analytical solution and a case where the canonical graph completion is applied to a system to yield a numerical solution.

3.2.2 Result: The canonical graph completion solution as a solution to the general jump-specified impulsive system

Consider the problem

$$\begin{cases} \frac{dx}{dt} = f(t,x) + g(t,x)\delta(t-t_*), & t \in [0,T], \\ x(0) = \bar{x}, \end{cases} \tag{3.22}$$

where $t_* \in (0,T)$, f and g are continuously differentiable functions with $g(t_*,\cdot)$ nowhere zero, and the first line is interpreted as an equality between measures.

The result Proposition 5.1 of [8] demonstrates that this system, when viewed as a perturbation theory problem through the use of continuous δ_ϵ sequences converging to δ as $\epsilon \to 0$, can be equivalently recast in the jump-specified form of an impulsive differential equation as

$$\frac{dx}{dt} = f(t,x), \quad t \neq t_*, \tag{3.23}$$

with the implicit jump condition

$$G_*(x(t_*^+)) - G_*(x(t_*^-)) = 1, \tag{3.24}$$

where

$$G_*(x) = \int \frac{dx}{g(t_*,x)}.$$

It is natural to consider whether the reparameterization methods from [13] and described above also yield a solution satisfying this jump-specified form, and thus provide an alternate means of obtaining such a solution. It turns out that this is the case if we consider the auxiliary system formed by the canonical graph completion corresponding to the δ distribution.

In order to apply these techniques, we must first remove the explicit time dependence of f and g by adding the new state variable $x_0 = t$ and the equation $\dot{x}_0(t) = 1$ with $x_0(0) = 0$. Then (3.22) becomes

$$\begin{cases} \dot{x}_0(t) = 1, & (x_0(0), x(0)) = (0, \bar{x}), \\ \dot{x}(t) = f(x_0(t), x(t)) + g(x_0(t), x(t))\delta(t-t_*), \end{cases} \tag{3.25}$$

which by (3.14) and (3.6) has auxiliary system

$$(\dot{y}_0(s), \dot{y}(s)) = \begin{cases} (T+1, (T+1)f(y_0(s), y(s))) & s \in [0, \frac{t_*}{T+1}] \\ (0, (T+1)g(y_0(s), y(s))) & s \in (\frac{t_*}{T+1}, \frac{t_*+1}{T+1}) \\ (T+1, (T+1)f(y_0(s), y(s))) & s \in [\frac{t_*+1}{T+1}, 1]. \end{cases} \tag{3.26}$$

Note that the reparameterized equation for the state x_0 coincides with that of the reparameterized time state y_0, so to simiplify notation the redundant equation is not listed above and the state y_0 will act as both the reparameterization of the state x_0 and the reparameterized time state y_0. Also note that the auxiliary system decouples and we immediately obtain $y_0(s)$ explicitly as

$$y_0(s) = \begin{cases} (T+1)s, & s \in [0, \frac{t_*}{T+1}] \\ t_*, & s \in (\frac{t_*}{T+1}, \frac{t_*+1}{T+1}) \\ (T+1)s - 1, & s \in [\frac{t_*+1}{T+1}, 1], \end{cases} \tag{3.27}$$

so (3.26) becomes

$$(\dot{y}_0(s), \dot{y}(s)) = \begin{cases} (T+1, (T+1)f((T+1)s, y(s))) & s \in [0, \frac{t_*}{T+1}] \\ (0, (T+1)g(t_*, y(s))) & s \in (\frac{t_*}{T+1}, \frac{t_*+1}{T+1}) \\ (T+1, (T+1)f((T+1)s-1, y(s))) & s \in [\frac{t_*+1}{T+1}, 1]. \end{cases} \tag{3.28}$$

We can now state a proposition relating a solution of (3.28) to a solution of (3.23), (3.24). (The details of the initial condition at $\bar{x}$ are omitted for simplicity in the following but can easily be verified within the proof.)

Proposition 3.7. *If the function $y(\cdot)$ is a solution to the y-component of the system (3.28), then the function $x(t) := y(\eta(t))$ satisfies the conditions of the jump-specified form of the impulsive differential equation given by (3.23) and (3.24), where $\eta(t)$ is as defined in (3.7).*

Proof. Let $y(\cdot)$ be a solution to the second state component of (3.28), so

$$\dot{y}(s) = \begin{cases} (T+1)f((T+1)s, y(s)) & s \in [0, \frac{t_*}{T+1}] \\ (T+1)g(t_*, y(s)) & s \in (\frac{t_*}{T+1}, \frac{t_*+1}{T+1}) \\ (T+1)f((T+1)s-1, y(s)) & s \in [\frac{t_*+1}{T+1}, 1]. \end{cases} \tag{3.29}$$

Define $x(t) := y(\eta(t)$, and observe in this case that

$$\eta(t) = \begin{cases} \frac{t}{T+1}, & t \leq t_* \\ \frac{t+1}{T+1}, & t > t_*. \end{cases} \tag{3.30}$$

Thus, η has derivative $\dot{\eta}(t) = \frac{1}{T+1}$ for $t \neq t_*$. The reparameterized times $s \in [0, \frac{t_*}{T+1}]$ and $s \in [\frac{t_*+1}{T+1}, 1]$ correspond to the non-impulsive times $t < t_*$ and $t > t_*$, respectively, and $s = \eta(t)$ for these times. Therefore,

$$\begin{aligned} \dot{x}(t) &= \dot{y}(\eta(t))\dot{\eta}(t) \\ &= \frac{1}{T+1}\dot{y}(\eta(t)), \end{aligned}$$

for which (3.29) and (3.30) give

$$\dot{x}(t) = \begin{cases} f(t, y(\frac{t}{T+1})) & , t < t_* \\ f(t, y(\frac{t+1}{T+1})) & , t > t_*. \end{cases} \tag{3.31}$$

But with (3.30) and the definition of $x(t)$ we may rewrite (3.31) as

$$\dot{x}(t) = f(t, x(t)) \quad t \neq t_*,$$

which satisfies (3.23).

Now we consider the time $t = t_*$, which has corresponding reparameterized times $s \in (\frac{t_*}{T+1}, \frac{t_*+1}{T+1})$ with $y(\frac{t_*}{T+1}) = x(t_*^-)$ and $y(\frac{t_*+1}{T+1}) = x(t_*^+)$ as can be seen from (3.27). For these times, (3.28) gives

$$\frac{dy}{ds} = (T+1)g(t_*, y). \tag{3.32}$$

Hence

$$\begin{aligned} \int_{x(t_*-)}^{x(t_*+)} \frac{dx}{g(t_*, x)} &= \int_{y(\frac{t_*}{T+1})}^{y(\frac{t_*+1}{T+1})} \frac{dy}{g(t_*, y)} \\ &= \int_{\frac{t_*}{T+1}}^{\frac{t_*+1}{T+1}} (T+1)ds \\ &= 1, \end{aligned}$$

which is equivalent to (3.24), so the proof is complete.

□

Remark 3.8. The canonical graph completion is used and explicitly mentioned in the above, but in such a case where the impulsive measure is scalar-valued the canonical graph completion provides the unique path by which the graph of the measure's distribution function is reparameterized as a Lipschitz function. Note once more that a trajectory of the auxiliary system formed with a given graph completion depends only on this path and not on the way the path is parameterized. The canonical graph completion simply provides a convenient method for constructing and handling graph completions.

3.3 Impulsive control differential inclusions

As in [23], we consider the measure driven differential inclusion

$$\begin{cases} dx \in F(x(t))dt + G(x(t))\mu(dt), \\ x(0^-) = \bar{x}, \quad t \in [0, T] \end{cases} \tag{3.33}$$

for μ a vector-valued, regular Borel measure taking values in the compact set $K \subset \mathbb{R}^m$ on $[0, T]$. We let $c > 0$ and make the following assumptions on the multifunctions F and G in accordance with [23]

(H1) The multifunction F has closed graph and convex values, and satisfies

$$f \in F(x) \to |f| \leq c(1 + |x|) \quad \forall x \in \mathbb{R}^n.$$

(H2) The multifunction G has closed graph and closed, convex values, and satisfies

$$g \in G(x) \rightarrow ||g|| \leq c(1+|x|) \quad \forall x \in \mathbb{R}^n.$$

Observe that if μ is absolutely continuous with respect to Lebesgue measure, then (3.33) reduces to an ordinary differential inclusion which has an extensive theory. If we are to allow the trajectory $x(\cdot)$ to be a function of bounded variation, then we will need to provide additional framework similar to that provided in the differential equations case above.

Recall that the distribution function $u(\cdot) : [0,T] \rightarrow \mathbb{R}^m$ of μ is given by $u(t) = \mu([0,t])$. Just as above, we can construct the canonical graph completion (φ_0, φ) of u, so for every $t \in [0,T]$ there exists an $s \in [0,1]$ such that $(\varphi_0(s), \varphi(s)) = (t, u(t))$. Note that the total variation in the sense of measure of μ is equal to the total variation in the function sense of its distribution function u. Hence, we may use either of these values to compute the graph completion. For more on using the total variation of the measure μ for the graph completion see [18, 19]. Observe that

$$\varphi_0(s) = t \quad \Leftrightarrow \quad \eta(t-) \leq s \leq \eta(t+),$$

where η is the reparameterization function from (3.7). The left and righthand limits $\eta(t-)$ and $\eta(t+)$ are equal if and only if t is not an atom of μ. Let $\mathrm{Ds}(\mu)$ be the set of atoms of μ, and let $\mathcal{I}$ be the at most countable index set of $\mathrm{Ds}(\mu)$. Let $I_i := [s_i^-, s_i^+] := \varphi_0^{-1}(t_i)$ be the time consumed during the jump. Since φ_0 is Lipschitz continuous, it is differentiable almost everywhere on $[0,1]$. Let supp $\mu_\sigma \subseteq [0,T]$ denote the closed support of μ_σ, the continuous singular component of the decomposition of the measure μ (for more on this decomposition see [14]). Define $\Gamma := \eta(\text{supp } \mu_\sigma) \subseteq [0,1]$, and set $\tilde{\Gamma} := \Gamma \cup (\bigcup_{i \in \mathcal{I}} I_i)$.

For the regular Borel measure μ from above, we can consider the three-tuple

$$X_\mu := (x(\cdot), \varphi(\cdot), \{y_i(\cdot)\}_{i \in \mathcal{I}}) \tag{3.34}$$

with the following constituents: $x(\cdot) : [0,T] \rightarrow \mathbb{R}^n$ is of bounded variation with its points of discontinuity a subset of μ's atoms $\mathrm{Ds}(\mu)$, $\varphi(\cdot) : [0,1] \rightarrow \mathbb{R}^m$ is the canonical graph completion of μ's distribution function $u(\cdot)$, and $\{y_i(\cdot)\}_{i \in \mathcal{I}}$ is a collection of Lipschitz functions, each defined on the nondegenerate interval $\mathcal{I}_i := [s_i^-, s_i^+]$ and satisfying $y_i(s_i^\pm) = x(t_i^\pm)$.

In [23], a modification to differential inclusions of the Bressan-Rampazzo definition of solutions to controlled differential equations is provided. We can now state that definition.

Definition 3.9. *Consider a three-tuple X_μ as in (3.34), and let*

$$y(s) = \begin{cases} x(t) & s \notin \bigcup_{i \in \mathcal{I}} I_i, \quad t = \varphi_0(s) \\ y_i(s) & s \in I_i. \end{cases}$$

Then X_μ *is a* reparameterized solution *of (3.33) provided* $y(\cdot)$ *is Lipschitz on* $[0,1]$ *and satisfies*

$$\begin{cases} \dot{y}(s) \in F(y(s))\dot{\varphi}_0(s) + G(y(s))\dot{\varphi}(s) \quad a.e.\ s \in [0,1] \\ y(0) = \bar{x}. \end{cases} \tag{3.35}$$

It will now be our purpose to show that this reparameterized solution to the differential inclusion (3.33) coincides with the reparameterized solution to an appropriately modified system of differential equations of a form similar to (3.1). In other words, we will provide a link between solutions of impulsive differential inclusions and that of impulsive differential equations which parallels the link for the non-impulsive, classic case.

3.4 Result: obtaining a system of impulsive differential equations from an impulsive differential inclusion

Before we can establish the link between the differential inclusion and the differential equation we must extend the generality of equation (3.1) to make it comparable to the inclusion (3.33). Namely, consider the system

$$\begin{cases} \dot{x}(t) = f((x(t), v(t)) + \sum_{i=1}^{m} g_i(x(t), w(t)) \cdot \dot{u}^i(t), \\ x(0) = \bar{x}, \qquad t \in [0,T], \end{cases} \tag{3.36}$$

where the maps $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ are continuously differentiable and globally bounded, $u \in BV^-([0,T], \mathbb{R}^m)$, and v, w are each measurable functions on $[0,T]$ taking values in the compact set $U \subset \mathbb{R}^m$.

Note that the term $\sum_{i=1}^{m} g_i \cdot \dot{u}^i$ in the differential equation (3.36) can be expressed as the matrix product

$$\sum_{i=1}^{m} g_i \cdot \dot{u}^i = \begin{bmatrix} g_{1,1} & g_{1,2} & \cdots & g_{1,m} \\ g_{2,1} & g_{2,2} & \cdots & g_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n,1} & g_{n,2} & \cdots & g_{n,m} \end{bmatrix} \begin{bmatrix} \dot{u}^1 \\ \dot{u}^2 \\ \vdots \\ \dot{u}^m \end{bmatrix},$$

where the arguments of the functions have been disregarded for neatness. Moreover, we note that the term $G \cdot d\mu$ appearing in the above differential inclusions is an analogous matrix product with the only difference being that the elements of G are subsets of $\mathbb{R}$. This remark should be sufficient to allow transition clearly from one notation to the other in what follows.

We adapt the procedure shown in section 3.1.1 above to define a solution in the sense of measure. Let $(z, w, p) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$. For every $i = 1, ..., m$, define the function $\tilde{g}_i : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^n$ as

$$\tilde{g}_i(z,w;p) = \int_0^1 g_i(\exp(\sigma \sum_{j=1}^m p^j g_j)(z,w))\ d\sigma, \tag{3.37}$$

where $\exp(\sigma \sum_{j=1}^m p^j g_j)(z,w)$ denotes the value at time $s=\sigma$ of the solution to the Cauchy problem

$$\frac{d\omega}{ds} = \sum_{j=1}^m p^j g_j(\omega(s)), \quad \omega(0) = (z,w). \tag{3.38}$$

Extending the identity (3.4), we have for every $(z,w,p) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^m$, the identity

$$\sum_{i=1}^m p^i \tilde{g}_i(z,w;p) = \exp(\sum_{i=1}^m p^i g_i)(z,w) - (z,w). \tag{3.39}$$

We now have, corresponding to the system (3.36), the Cauchy problem

$$\dot{x} = f(x,v) + \sum_{i=1}^m \tilde{g}_i(x(t^-), w(t^-); \dot{u}(\{t\}))\dot{u}^i, \quad x(0^+) = \bar{x}, \tag{3.40}$$

on $(0,T)$ and for which we can define a solution in the sense of measure.

Definition 3.10. *Let $u \in BV^-([0,T],\mathbb{R}^m)$. A solution of (3.40) is a map $x \in BV^-([0,T],\mathbb{R}^n)$ which satisfies (3.40) in the sense of measures on $(0,T)$. That is, x satisfies*

$$\int_B \dot{x} = \int_B f(x(t),v(t))dt + \sum_{i=1}^m \int_B \tilde{g}_i(x(t^-), w(t^-); \dot{u}(\{t\}))\dot{u}^i$$

for every Borel subset B of $(0,T)$.

We turn our attention to the differential inclusion (3.33) and its reparametrized counterpart (3.35) to mention that the latter inclusion, having a Lipschitz continuous solution y, is eligible to have a selection theorem like Filippov's lemma, mentioned in section 2.2, applied to it after obvious modification (in particular, viewing the sum of two multifunctions as a single multifunction). Such a selection theorem will provide for the existence of a pair of functions $(v,w) : [0,1] \to U \times U$, called selections and whose regularity depends on the regularity of the multifunctions F and G. In some sense, the selection inherits the regularity of the multifunction. We do not investigate the minimal regularity of the given multifunctions which guarantees the needed regularity of (v,w), but accept that such regularity is attainable for a sufficiently regular choice of multifunction. We do note that the pair (v,w) must be measurable functions and that in light of (3.40), w must have a well-defined left-hand limit at all $s \in [0,1]$. Measurability of the multifunction will induce measurability of the selection. As for a well-defined left-hand limit of w, it is sufficient, though not necessary, to take a continuous w having values in the compact set $U \subset \mathbb{R}^n$ to meet this criterion. Such a continuous selection is guaranteed if the multifunction under consideration is continuous, but as the theory of multifunctions is rich it may well be possible to guarantee a one-sided continuous selection from a less regular multifunction. For more details on regularity and selections of multifunctions see [3] and [9].

Theorem 3.11. *Let μ be a regular Borel measure with distribution function u an element of $BV^-([0,T],\mathbb{R}^m)$, and let φ be the canonical graph completion of u with reparameterization function η as defined in (3.7). Assume the three-tuple X_μ from (3.34) yields a reparameterized solution $y(\cdot)$ of (3.33), and that there exists a measurable selection $(v,w) : [0,1] \to U \times U$ such that $w(s^-)$ is well-defined for all $s \in [0,1]$ and such that the map $y(\cdot)$ satisfies*

$$\frac{dy(s)}{ds} = f((y(s),v(s))\frac{d\varphi_0(s)}{ds} + \sum_{i=1}^{m} g_i(y(s),w(s))\frac{d\varphi_i(s)}{ds}, \quad y(0) = \bar{x}. \tag{3.41}$$

Then the function $x(t) := y(\eta(t))$ is a solution of (3.40) under the controls $v(\eta(\cdot))$ and $w(\eta(\cdot))$.

Proof. By definition, y is a solution of the integral equation

$$y(s) = \int_0^s (f(y(\xi),v(\xi))\varphi_*^0(\xi) + \sum_{i=1}^{m} g_i(y(\xi),w(\xi))\varphi_*^i)\ d\xi,$$

where φ_*^0 and φ_*^i are Borel functions coinciding with the derivatives $d\varphi^0/ds$ and $d\varphi^i/ds$ almost everywhere with respect to Lebesgue measure. Since f and the functions g_i are bounded, and the canonical graph completion φ is Lipschitz continuous, y is also Lipschitz continuous. Note that η is an element of $BV^-([0,T],[0,1])$. Hence, Theorem 3.6 implies

$$\dot{x} = \hat{y}_*(\eta)\dot{\eta}, \tag{3.42}$$

where

$$y_*(s) = f(y(s),v(s))\varphi_*^0(s) + \sum_{i=1}^{m} g_i(y(s),w(s))\varphi_*^i(s).$$

Recall $\mathrm{Ds}(\mu)$ is the set of atoms of μ so $[0,T]\backslash\mathrm{Ds}(\mu)$ is the set of points at which the distribution function u (hence η) is continuous. Let $A \subset [0,T]\backslash\mathrm{Ds}(\mu)$. We have the identity of measures

$$\dot{x}(A) = \int_A \hat{y}_*(\eta)\dot{\eta},$$

and, by the continuity of η at every $t \in A$,

$$\hat{y}_*(\eta)(t) = y_*(\eta(t)) = f(y(\eta(t)),v(\eta(t)))\varphi_*^0(\eta(t)) + \sum_{i=1}^{m} g_i(y(\eta(t)),w(\eta(t)))\varphi_*^i(\eta(t)).$$

Then

$$\dot{x}(A) = \int_A f(y(\eta), v(\eta))\varphi^0_*(\eta)\dot{\eta} + \sum_{i=1}^m \int_A g_i(y(\eta), w(\eta))\varphi^i_*(\eta)\dot{\eta}.$$

Since $\varphi^0(\eta(t)) = t$ for every $t \in A$, Theorem 3.6 implies that $\varphi^0_*(\eta)\dot{\eta}$ coincides with Lebesgue measure dt on A. Similarly, since $\varphi^i(\eta(t)) = u^i(t)$, we have $\varphi^i_*(\eta)\dot{\eta} = \dot{u}^i$ as measures on A, for $i = 1, ..., m$. Therefore,

$$\int_A \dot{x} = \int_A f(x(t), v(\eta(t))) + \sum_{i=1}^m \int_A g_i(x, w)\dot{u}^i.$$

Note for every $t \in [0, T]\backslash \mathrm{Ds}(\mu)$ and for all $i = 1, ..., m$ we have $g_i(x(t), w(\eta(t)) \equiv \tilde{g}_i(x(t^-), w(\eta(t^-)); 0) = \tilde{g}_i(x(t), w(\eta(t)); \dot{u}(\{t\}))$, where $\tilde{g}$ is as above. Thus, the above equality can be written in the form

$$\int_A \dot{x} = \int_A f(x(t), v(\eta(t)))\ dt + \sum_{i=1}^m \int_A \tilde{g}_i(x, w(\eta), \dot{u}(\{t\}))\dot{u}^i. \tag{3.43}$$

Now let $t \in \mathrm{Ds}(\mu)$, and observe that (3.42) gives

$$\dot{x}(\{t\}) = \hat{y}_*(\eta)(t)\dot{\eta}(\{t\}). \tag{3.44}$$

Recall from the canonical graph completion discussion that $\dot{\varphi}(s) = (0, \frac{\Delta u(t)}{\Delta \eta(t)})$ for every $s \in (\eta(t), \eta(t^+))$, which with equation (3.41) implies for every $\sigma \in (0, \dot{\eta}(\{t\}))$

$$y(\eta(t) + \sigma) = \exp(\sigma \sum_{i=1}^m \frac{\dot{u}^i(\{t\})}{\dot{\eta}(\{t\})} g_i) y(\eta(t)). \tag{3.45}$$

Therefore

$$\begin{aligned}
\hat{y}_*(\eta)(t) &= \int_0^1 \frac{dy}{ds}(\eta(t) + \sigma\dot{\eta}(\{t\}))\ d\sigma \\
&= \frac{1}{\dot{\eta}(\{t\})} \int_{\eta(t)}^{\eta(t)+\dot{\eta}(t)} \frac{dy}{ds}(s)\ ds \\
&= \frac{1}{\dot{\eta}(\{t\})} [\exp\left(\sum_{i=1}^m \dot{u}^i(\{t\}) g_i\right) y(\eta(t)) - y(\eta(t))].
\end{aligned}$$

This equation with (3.44) and the fact that $y(\eta(t)) = x(t)$ gives

$$\int_{\{t\}} \dot{x} = \exp\left(\sum_{i=1}^m \dot{u}^i(\{t\}) g_i\right) x(t) - x(t).$$

By (3.39), the right-hand side coincides with the vector

$$\sum_{i=1}^m \tilde{g}_i(x(t), w(t); \dot{u}(\{t\}))\dot{u}^i(\{t\}).$$

Furthermore, $\int_{\{t\}} f(x(t), v(\eta(t))\ dt = 0$, so

$$\int_{\{t\}} \dot{x} = \int_{\{t\}} f(x(t), v(\eta(t)))\ dt + \sum_{i=1}^{m} \int_{\{t\}} \tilde{g}_i(x(t), w(\eta(t)); \dot{u}(\{t\}))\dot{u}^i. \tag{3.46}$$

Since we have (3.43) and (3.46) for every $A \subset [0, T]\backslash\mathrm{Ds}(\mu)$ and for every $t \in \mathrm{Ds}(\mu)$, respectively, the map $x(t) = y(\eta(t))$ is a solution of (3.40) according to Definition 3.10. □

Chapter 4

Necessary conditions for optimal control problems with impulses

In this chapter we examine necessary conditions in the form of the Pontryagin Maximum Principle for optimal impulsive control problems. We begin with the maximum principle derived in [18] where the system considered is driven by an impulsive measure which is nonnegative and scalar-valued and the optimization is of fixed end-time Mayer form. Furthermore, the function g of the product $g \cdot \mu(dt)$, which provides the impulsive portion of the dynamics, may only depend on the time and state t, x. However, we will see that this maximum principle accommodates nonsmooth dynamics.

The second maximum principle we study, from [2], will be in regard to a fixed end-time Mayer problem subject to dynamics whose impulsive term is vector-valued and generally has values in $\mathbb{R}^k$, that is, it may have negative components. Moreover, the function g of the impulsive dynamics may generally depend on the control u, this being made possible by a more sophisticated definition of the impulsive control which, among other things, provides for a continuous evolution along system jumps similar to that described in section 3.1.1. The results of [2] are preceded by the results of [18], so such extensions of generality are expected. However, the maximum principle, and possibly the solution definition, of the second problem does not necessarily accommodate nonsmooth dynamics like the first. Such an extension of generality can be noted as a point of future work.

In what follows, we make a comparison between the necessary conditions mentioned above. We extend the necessary conditions of the fixed end-time Mayer problem given by [2] to those of a free end-time Mayer problem, a Bolza problem, and a minimum time problem. Such extensions are made in the existing literature on case-to-case basis, whereas the present work provides these extensions in full generality. We comment on the challenges and possibility of extending these conditions to an impulsive Calculus of Variations problem. Such an extension is a desirable link in the reformulation of optimal control problems as Calculus of Variations problems (and vice versa) since techniques for solving each problem may then be interchanged. This is discussed for the nonimpulsive case in [10].

4.1 Necessary conditions for the optimal impulsive control problem via graph completions

We now turn our attention to the problem

$$
(\mathrm{P}) \quad \begin{cases} \text{Minimize} \;\; h(x(0), x(1)) \\ \text{subject to} \;\; dx(t) = f(t, x(t), u(t))dt + g(t, x(t))\mu(dt), \quad t \in [0,1], \\ (x(0), x(1)) \in C, \\ u(t) \in U_t, \quad \mathcal{L} - \text{a.e. } t \in [0,1], \text{ and } \mu \geq 0, \end{cases} \tag{4.1}
$$

Necessary conditions for this problem are established in [19], while the companion paper [18] provides the background concerning the solution definitions of this possibly impulsive system along with the development of the reparameterization techniques utilized. The authors of these papers extend the notions of solutions and reparameterizations of impulsive systems of control differential equations given in [13] to those involving differential inclusions. The outstanding advantage of the differential inclusion formulation lies in its ability to handle the nonsmooth dynamics often present in optimization problems, which in turn provides a means of finding, or at least narrowing searches for, optimal processes for a wider range of problems.

In regard to the problem (P), we adopt for this section the definitions and notation used in [19] in order to summarize the results provided by that work. Let $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $f : [0,1] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, and $g : [0,1] \times \mathbb{R}^n \to \mathbb{R}^n$ be given functions. U is a Borel subset of $[0,1] \times \mathbb{R}^m$, U_t denotes the section $\{x : (t, x) \in U\}$, and C is a closed subset of $\mathbb{R}^n \times \mathbb{R}^n$. We will take a *control policy* to be a pair (u, μ), where the conventional control component u is a Lebesgue measurable function satisfying $u(t) \in U_t$ a.e. with respect to Lebesgue measure and the impulsive control μ is a scalar-valued, nonnegative, regular Borel measure. Thus, a *process* in this context will be a triple (x, u, μ), where x is the *state trajectory* resulting from the choice of control policy (u, μ).

4.1.1 Change of variables

Similar to the reparameterization set forth in [13], the authors of [18, 19] address the distribution function F for a given measure $\mu \in C^+((0,1), \mathbb{R})$ through the relationship

$$
F(t) := \begin{cases} \int_{[0,t]} \mu(ds), & t \in (0,1] \\ 0 & \text{if } t = 0. \end{cases} \tag{4.2}
$$

Define the *reparameterization function* η corresponding to μ as

$$
\eta(t) := \begin{cases} (t + \int_{[0,t]} \mu(d\tau))/(1 + \mu([0,1])), & t \in (0,1] \\ 0, & t = 0. \end{cases} \tag{4.3}
$$

Clearly η is a strictly increasing $BV^+((0,1),\mathbb{R})$ function. Define the continuous, nondecreasing function $\varphi_0 : [0,1] \to [0,1]$ to be

$$\varphi_0(s) := \sup\{t \in [0,1] : \eta(t) \geq s\} \quad \forall s \in [0,1]. \tag{4.4}$$

Let $\{t_i\}$ be an enumeration of $\mathrm{Ds}(\mu)$, the set of μ's atoms, and let $S_i = [\sigma_i', \sigma_i'']$ be the subintervals $S_i := \theta^{-1}(\{t_i\})$ for $i = 1, 2, ...$ Now we may define the function $\varphi : [0,1] \to \mathbb{R}^+$ to be

$$\varphi(s) := \begin{cases} F(\varphi_0(s)) & \text{if } s \in [0,1] \setminus \bigcup_{i=1}^{\infty} S_i \\ F(t_i^-) + \frac{s-\sigma_i'}{\sigma_i''-\sigma_i'}(F(t_i) - F(t_i^-)) & \text{if } s \in S_i, \quad i = 1, 2, ... \end{cases} \tag{4.5}$$

If $t_i = 0$ for some i, then $F(t_i^-)$ and $F(t_i)$ are interpreted in the above formula as $F(0)$ and $F(0^+)$, respectively.

It can be shown that the function $(\varphi_0, \varphi) : [0,1] \to [0,1] \times \mathbb{R}^+$ thus constructed, is the canonical graph completion of the measure μ. Therefore, it has the properties listed below Definition 3.2. We see some of these conditions and more in the following proposition from [18].

Proposition 4.1. *Let (φ_0, φ) be the graph completion of $\mu \in C^+((0,1),\mathbb{R})$. Then*

- *φ_0 and φ are Lipschitz continuous, nonnegative, nondecreasing functions and*

$$\dot{\varphi}_0(s) + \dot{\varphi}(s) = 1 + \mu([0,1]) \quad \mathcal{L} - a.e.$$

- *For any Borel measurable function h which is μ integrable and any Borel set $T \subset [0,1]$ we have*

$$\int_{\varphi_0^{-1}(T)} h(\varphi_0(s))\dot{\varphi}(s)ds = \int_T h(\tau)\mu(d\tau).$$

- *For any $\mathcal{L}$-integrable function g and Borel set $S \subset [0,1]$, $\varphi_0(S)$ is also a Borel set and*

$$\int_S g(\varphi_0(s))\dot{\varphi}_0(s)ds = \int_{\varphi_0(S)} g(\tau)d\tau. \tag{4.6}$$

- *Let $\{\mu_i\}$ be a sequence of elements in $C^+((0,1),\mathbb{R})$, and let $\{(\varphi_0^i, \varphi_i)\}$ be the corresponding graph completions. Suppose that $\mu_i \to \mu$ weakly*. Then $(\varphi_0^i, \varphi_i) \to (\varphi_0, \varphi)$ uniformly and $(\dot{\varphi}_0^i, \dot{\varphi}_i) \to (\dot{\varphi}_0, \dot{\varphi})$ weakly in L^1.*

4.1.2 Measure-driven differential inclusions

In order to study the necessary conditions of [19], we must make precise the notion of robust solutions of measure-driven differential inclusions of the form

$$\begin{cases} dx(t) & \in \; F_1(t, x(t))dt + F_2(t, x(t))\mu(dt), \quad t \in [0,1] \\ x(0) & = \; \bar{x} \end{cases} \tag{4.7}$$

where $F_1 : [0,1]\times\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ and $F_2 : [0,1]\times\mathbb{R}^n \rightrightarrows \mathbb{R}^n$ are multifunctions mapping points in $[0,1]\times\mathbb{R}^n$ to subsets of $\mathbb{R}^n$.

The multifunction $\tilde{F}_2 : [0,1]\times\mathbb{R}^n\times[0,\infty) \rightrightarrows \mathbb{R}^n$ will be used in the definition of robust solution and is defined as

$$\tilde{F}_2(t,v;\alpha) := \{\alpha^{-1}[\xi(1)-\xi(0)] : \xi \in AC((0,1),\mathbb{R}^n),\ \dot{\xi}(\sigma) \in \alpha F_2(t,\xi(\sigma)) \text{ a.e.}, \\ \text{and } \xi(0) = v\}$$

whenever $\alpha > 0$, and $\tilde{F}_2(t,v;0) := F_2(t,v)$.

Definition 4.2. *We say that a function $x \in BV^+((0,1),\mathbb{R}^n)$ is a* robust solution *to (4.7) if there exist an $\mathcal{L}$-integrable function ϕ_1 and μ-integrable function ϕ_2 such that*

$$\phi_1 \in F_1(t,x(t)), \quad \mathcal{L}\text{-a.e.}, \tag{4.8}$$

$$\phi_2 \in \tilde{F}_2(t,x(t^-);\mu(\{t\})), \quad \mu\text{-a.e.} \tag{4.9}$$

and

$$x(t) = x(0) + \int_0^t \phi_1(\tau)d\tau + \int_{[0,t]} \phi_2(\tau)\mu(d\tau) \quad \forall\ t \in (0,1].$$

The reparameterization of μ by means of the graph completion $(\varphi_0(\cdot),\varphi(\cdot))$ and reparameterization function η results in a conventional differential inclusion. The following proposition, stated and proved in [19], describes this.

Proposition 4.3. *Suppose that the data for the measure-driven differential inclusion (4.7) satisfies the following:*

- *F_1 has values-closed sets and is $\mathcal{L}\times\mathcal{B}$ measurable and*
- *F_2 has values-closed sets and is Borel measurable.*

Fix a measure $\mu \in C^+((0,1),\mathbb{R})$ and an initial state $\bar{x}$. We have the following:

(i) Suppose that $x(\cdot) \in BV^+([0,1],\mathbb{R}^n)$ is a robust solution to (4.7). Then there exists a solution $y(\cdot) \in AC([0,1],\mathbb{R}^n)$ to

$$\begin{cases} \dot{y}(s) &= F_1(\varphi_0(s),y(s))\dot{\varphi}_0(s) + F_2(\varphi_0(s),y(s))\dot{\varphi}(s), \quad s \in [0,1] \\ y(0) &= \bar{x} \end{cases} \tag{4.10}$$

for which

$$x(t) = y(\eta(t)) \quad \forall t \in [0,1]. \tag{4.11}$$

Conversely,

(ii) suppose that $y \in AC([0,1],\mathbb{R}^n)$ is a solution to (4.10). Then there exists a robust solution $x(\cdot) \in BV^+([0,1],\mathbb{R}^n)$ to (4.7) for which (4.11) is satisfied.

This proposition is the measure-driven differential inclusion analogue of Theorem 3.5.

4.1.3 Necessary conditions: a maximum principle for the Mayer problem with fixed end time

In order to obtain the optimality conditions derived in [18] for problem (P), the following hypotheses are required:

(H1) There exists a constant $K_f(\cdot) \in L^1$ such that

$$|f(t,x,u) - f(t,y,u)| \leq K_f(t)|x-y| \text{ for } (x,u) \in \mathbb{R}^n \times \mathbb{R}^m \text{ and } t \in [0,1].$$

(H2) $f(\cdot, x, \cdot)$ is $\mathcal{L} \times \mathcal{B}$-measurable.

(H3) $g(\cdot,\cdot)$ is continuous and there exists a constant K_g such that

$$|g(t,x) - g(t,y)| \leq K_g|x-y| \; \forall x,y \in \mathbb{R}^n, t \in [0,1].$$

(H4) $U \in \mathbb{R}^{1+m}$ is a Borel set.

After an intermediate theorem regarding processes (x,u,μ) which generate boundary points of some reachable set of the system with dynamics given in (P), the authors of [18] establish the following maximum principle for (P).

Theorem 4.4. *Let $(\bar{x}(\cdot), \bar{u}(\cdot), \bar{\mu}(\cdot))$ be a minimizing process for (P). Assume that h is locally Lipschitz continuous, that C is a closed subset, and that hypotheses (H1)-(H4) are satisfied. Then there exist $\lambda \geq 0$ and $p \in BV^+([0,1],\mathbb{R}^n)$ such that $||p(\cdot)||_{L^\infty} + \lambda > 0$ and $(\bar{x}(\cdot), p(\cdot))$ is a robust solution of the MDI*

$$d\begin{bmatrix} \bar{x}(t) \\ p(t) \end{bmatrix} \in \begin{bmatrix} f(t,\bar{x}(t),\bar{u}(t)) \\ -p(t)\cdot co\partial_x f(t,\bar{x}(t),\bar{u}(t)) \end{bmatrix} dt + \begin{bmatrix} g(t,\bar{x}(t)) \\ -p(t)\cdot \overline{\partial_x} g(t,\bar{x}(t)) \end{bmatrix} \bar{\mu}(dt).$$

Furthermore,

$$(p(0), -p(1)) \in N_C(\bar{x}(0),\bar{x}(1)) + \lambda\partial h(\bar{x}(0),\bar{x}(1)), \tag{4.12}$$

$$p(t)\cdot f(t,\bar{x}(t),\bar{u}(t)) = \max_{u\in U_t}\{p(t)\cdot f(t,\bar{x}(t),u)\} \quad a.e.\ t \in [0,1], \tag{4.13}$$

$$p(t)\cdot g(t,\bar{x}(t)) \leq 0 \quad \forall t \in [0,1], \tag{4.14}$$

$$p(t)\cdot g(t,\bar{x}(t)) = 0 \quad \mu - a.e.\ on\ [0,1]. \tag{4.15}$$

Corresponding to every atom t of $\bar{\mu}$, there exists a solution $\begin{bmatrix}\xi_t(\cdot) \\ \alpha_t(\cdot)\end{bmatrix}$ to

$$\begin{bmatrix} \dot{\xi}_t(s) \\ \dot{\alpha}_t(s) \end{bmatrix} \in \bar{\mu}(\{t\}) \begin{bmatrix} g(t,\xi_t(s)) \\ -\alpha_t(s)\cdot\overline{\partial_x} g(t,\xi_t(s)) \end{bmatrix} \quad on \quad [0,1] \tag{4.16}$$

which satisfies

$$(\xi_t(0), \alpha_t(0)) = (\bar{x}(t^-), p(t^-)), \quad (\xi_t(1),\alpha_t(1)) = (\bar{x}(t), p(t)), \tag{4.17}$$
$$\alpha_t(s)\cdot g(t,\xi_t(s)) \geq 0 \quad \forall s \in [0,1]. \tag{4.18}$$

We defer to [18] for the proof of this theorem. Note that the end time for the problem (P) is the fixed time $t = 1$, so that (P) represents a fixed end time problem.

4.2 Necessary conditions for the optimal impulsive control problem via measure-adjoint functions

More recent work in optimal impulsive control problems, [2], has established maximum conditions for problems whose jump dynamics include a dependence on a conventional control u and whose jumps are induced by a vector-valued measure μ. It is also of interest to note that the authors of the referenced paper are able to suppress the notational burden of the auxiliary system used in past papers, [19, 18, 13, 5, 6, 23, 24], by establishing the necessary conditions via integrals, as we will see in what follows. We will see that this suppression of the auxiliary system has advantages and disadvantages.

4.2.1 Necessary conditions for the impulsive Mayer problem with fixed end time

Consider the problem:

$$(P) \qquad \begin{cases} \min \varphi(x(0), x(T)) \\ dx = f(t,x,u)dt + g(t,x,u)d\vartheta, \quad t \in [0,T], \\ p = (x(0), x(T)) \in S, \end{cases}$$

where

- $[0,T]$ is a fixed time interval,
- S is a closed subset of $\mathbb{R}^{2n}$,
- $\varphi(\cdot,\cdot)$ is the objective function to be minimized,
- $\vartheta = (\mu, \nu, \{u_\tau, v_\tau\})$ is referred to as the impulsive control.

Let us point out that the problem above does not include mixed constraints as the problem considered in [2] does and that the necessary conditions we cite from this work will reflect this lack of mixed constraints. The mixed constraints considered in [2] are given by $R(t,x,u) \in C$ for $t \in [0,T]$ and some closed, convex subset C of $\mathbb{R}^r$. By taking $C = \mathbb{R}^r$ and $R \equiv 0 \in R^r$, these constraints are trivially met by the dynamics of (P), and we are then able to deduce the simplifications which yield the maximum principle we present below.

The impulsive control ϑ is a new object consisting of three components, together comprising implicit information about the jumps. This object eliminates the need for an explicit graph completion in the analysis of necessary conditions for problems with vector-valued measures and whose dynamics during jumps may depend on conventional controls. The first component μ is a vector-valued Borel measure with range a closed, convex cone $K \subset \mathbb{R}^k$. The second component ν is the variation of the impulsive control ϑ, which is defined as a scalar-valued Borel measure such that

$\nu \geq |\mu|$, where $|\mu|$ denotes the total variation of μ. Recall that the total variation of a vector-valued measure μ is the sum of the total variations of all the components, $|\mu| = \sum_{i=1}^{k} |\mu^i|$. We may also refer to the variation of the impulsive control as $|\vartheta|$, so $|\vartheta| = \nu$. The third component $\{u_\tau, v_\tau\}$ consists of an infinite family of measurable functions defined on $[0, 1]$ and depending on the real parameter $\tau \in [0, T]$. We provide the exact properties of this family, the definition of the impulsive control, and the solution concept for the dynamics of (P) later in this section.

The function $u : \mathbb{R} \to \mathbb{R}^m$ is the conventional control, or usual control, and it is assumed to be measurable and essentially bounded with respect to the usual Lebesgue measure l and with respect to the Lebesgue-Stieltjes measure $|\vartheta|$. Additionally, for the compact set $U \subset \mathbb{R}^m$, we require that an admissible control u have values in U almost everywhere. Let $\mathcal{U}$ be the set of admissible controls so described. We clarify as in [2] that the term "Lebesgue-Stieltjes" refers to the fact that the Borel measure $|\vartheta|$ is completed up to sets of measure zero, and any Borel measure μ on the σ-algebra $\mathcal{B}([0, T])$ of Borel subsets of $[0, T]$ can be uniquely extended by the Lebesgue extension of measure [14]. This extension is itself complete, and it is the Lebesgue-Stieltjes measure generated by μ. Also, a set A is measurable with respect to both the measures l and $|\vartheta|$ if and only if it is measurable with respect to the Lebesgue-Stieltjes measure $l + |\vartheta|$, since A can be represented as the union of Borel and zero measure sets. From this, we deduce that $u(\cdot)$ is measurable with respect to $l + |\vartheta|$.

The functions in (P) have the forms

$$\begin{aligned}
&\varphi : \mathbb{R}^{2n} \to \mathbb{R}^1,\\
&f : \mathbb{R}^1 \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n,\\
&g : \mathbb{R}^1 \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n \times \mathbb{R}^k,
\end{aligned}$$

and satisfy the following hypotheses:

(H) The function f and its partial derivatives in (x, u) are measurable in t for each (x, u) with respect to Lebesgue measure and are continuously differentiable in (x, u) for almost all t uniformly in t. The function g is assumed to be continuous, and it is continuously differentiable in (x, u). Both functions f and g, as well as their partials in (x, u) are bounded on any bounded set. The function φ is assumed to be continuously differentiable.

We can now provide the concepts of solution and impulsive control to (P). Take K a nonempty closed convex cone in $\mathbb{R}^k$, and consider a Borel vector-valued measure μ such that range$(\mu) \subset K$. Define $V(\mu)$ to be the set of scalar-valued nonnegative Borel measures ν such that $\exists \mu_i$: range$(\mu_i) \subset K$ and $(\mu_i, |\mu_i|) \xrightarrow{w} (\mu, \nu)$. The convergence $\xrightarrow{w}$ means convergence of each component of μ_i in the weak* topology $C^*([0, T])$, the space dual to the space of continuous functions $C([0, T])$. For example, if K is conained in one of the orthants, then $V(\mu) = \{|\mu|\}$ a singleton. If K is a half-space or the whole space, then $V(\mu) = \{\nu \in C^*([0, T]) : \nu \geq |\mu|\}$.

Observe that $|\mu| \in V(\mu)$, so $V(\mu) \neq \emptyset$ and also $\nu \geq |\mu|\ \forall \nu \in V(\mu)$. Now consider an arbitrary scalar-valued measure $\nu \in V(\mu)$, a number $\tau \in [0,T]$, and a measurable vector-valued function $v_\tau : [0,1] \to K$ such that

- $\sum_{i=1}^{k} |v_\tau^j(s)| = \nu(\{\tau\})$ a.e. $s \in [0,1]$;
- $\int_0^1 v_\tau^j(s)ds = \mu^j(\{\tau\}), \quad j = 1, ..., k.$

Observe that $\mu(\{\tau\})$ is a vector in K which is only nonzero whenever τ is an atom of μ. A family of vector-valued functions $\{v_\tau\}$ depending on the real parameter τ is said to be *adjoint* to a vector-valued measure (μ, ν) if, for every τ, conditions 1. and 2. hold.

Definition 4.5. *The* impulsive control *of problem* (P) *is a triple* $\vartheta = (\mu, \nu, \{u_\tau, v_\tau\})$*, where* $\nu \in V(\mu)$*,* v_τ *is a family of functions adjoint to* (μ, ν) *and* $\{u_\tau\}$ *is any essentially bounded family of measurable vector-valued functions defined on the closed interval* $[0,1]$ *and taking values in* $\mathbb{R}^m$*. The measure* ν *is the* variation *of the impulsive control* ϑ *and is denoted by* $|\vartheta|$*.*

Now for a given $\tau \in [0,T]$, we consider the impulsive control $\vartheta = (\mu, \nu, \{u_\tau, v_\tau\})$ and a given vector $x \in \mathbb{R}^n$. Denote by $\chi_\tau(\cdot) = \chi_\tau(\cdot, x)$ the solution to the system

$$\begin{cases} \dot{\chi}_\tau(s) = g(\tau, \chi_\tau(s), u_\tau(s))v_\tau(s), \quad s \in [0,1], \\ \chi_\tau(0) = x. \end{cases} \tag{4.19}$$

A *solution* to the differential equation of (P) corresponding to the coupled control (u, ϑ) with $u, u_\tau \in \mathcal{U}$ for all $\tau \in \mathrm{Ds}(\vartheta)$ and initial point $\bar{x}$, is a function $x(t)$ of bounded variation on the interval $[0,T]$ satisfying $x(0) = \bar{x}$ and, for every $t \in (0,T]$,

$$\begin{aligned} x(t) = \bar{x} &+ \int_0^t f(\tau, x(\tau), u(\tau))d\tau + \int_{[0,t]} g(\tau, x(\tau), u(\tau))d\mu_{\text{c.p.}} \\ &+ \sum_{\tau \leq t} [\chi_\tau(1, x(\tau^-)) - x(\tau^-)]. \end{aligned} \tag{4.20}$$

Observe that the term for the jump dynamics, $g \cdot d\mu$, splits into the continuous part $\mu_{\text{c.p.}}$ and the discrete part of the measure which corresponds to the summation over the countably many atoms of μ, $Ds(\vartheta)$.

Denote by H the Hamiltonian function

$$H(t, x, u, \psi) := \langle f(t, x, u,), \psi \rangle,$$

and by Q the vector-valued function $Q(t, x, u, \psi) := g^{\mathrm{Tr}}(t, x, u)\psi$, where g^{Tr} is the transpose of the matrix g. We are now ready to state the necessary conditions derived in [2].

Theorem 4.6. *Let* $(\hat{x}, \hat{u}, \hat{\vartheta})$ *be an optimal process for problem* (P)*, and let hypothesis (H) hold. Then there exist a number* $\lambda \geq 0$*, a vector-valued function* ψ

of bounded variation, and, for every point $\tau \in Ds(\hat{\vartheta})$, there exists an absolutely continuous vector-valued function σ_τ defined on the closed interval $[0,1]$ such that

$$\begin{gathered} \lambda + |\psi(t)| \neq 0 \quad \forall\, t \in [0,T], \\ \lambda + |\sigma_\tau(s)| \neq 0 \quad \forall s \in [0,1], \quad \forall \tau \in Ds(\hat{\vartheta}); \end{gathered} \tag{4.21}$$

$$\begin{aligned} \psi(t) = \psi(0) &- \int_0^t \frac{\partial H}{\partial x}(\tau, \hat{x}(\tau), \hat{u}(\tau), \psi(\tau)) d\tau \\ &- \int_{[0,t]} \frac{\partial}{\partial x} \langle Q(\tau, \hat{x}(\tau), \hat{u}(\tau), \psi(\tau)), d\hat{\mu}_{c.p.} \rangle \\ &+ \sum_{\tau \in Ds(\hat{\vartheta}) : \tau \leq t} [\sigma_\tau(1) - \psi(\tau^-)], \qquad \forall t \in (0,T]; \end{aligned} \tag{4.22}$$

$$\begin{cases} \frac{d\hat{\chi}_\tau(s)}{ds} = g(\tau, \hat{\chi}_\tau(s), \hat{u}_\tau(s)) \hat{v}_\tau(s) \\ \frac{d\sigma_\tau(s)}{ds} = -\frac{\partial}{\partial x} \langle Q(\tau, \hat{\chi}_\tau(s), \hat{u}_\tau(s), \sigma_\tau(s)), \hat{v}_\tau(s) \rangle \\ \hat{\chi}_\tau(0) = \hat{x}(\tau^-), \quad \sigma_\tau(0) = \psi(\tau^-), \quad s \in [0,1]; \end{cases} \tag{4.23}$$

$$(\psi(0), -\psi(T)) \in \lambda \frac{\partial \varphi}{\partial p}(\hat{p}) + N_S(\hat{p}), \tag{4.24}$$

$$\max_{u \in U} H(t, \hat{x}(t), u, \psi(t)) = H(t, \hat{x}(t), \hat{u}(t), \psi(t)) \quad a.e.\ t \in [0,T]; \tag{4.25}$$

$$\begin{cases} \max_{u \in U} Q(t, \hat{x}(t), u, \psi(t)) \in N_K(0) \quad \forall t \in [0,T] \\ \max_{u \in U} Q(\tau, \hat{\chi}_\tau(s), u, \sigma_\tau(s)) \in N_K(0) \quad \forall s \in [0,1], \quad \forall \tau \in Ds(\hat{\vartheta}), \\ \int\limits_{[0,T]} \langle g^{Tr}(t, \hat{x}(t), \hat{u}(t)) \psi(t)), d\hat{\mu_{c.p.}} \rangle + \\ \qquad \sum_{\tau \in Ds(\vartheta)} \int\limits_{[0,1]} \langle g^{Tr}(s, \hat{x}(s), \hat{u}(s)) \sigma_\tau(s)), \hat{v}_\tau(s) \rangle \geq 0. \end{cases} \tag{4.26}$$

Let us review the formulas of the above theorem. The conditions of (4.21) represent the non-triviality condition. The equation (4.22) is the adjoint equation for the regular, non-impulsive portion of the trajectory with ψ being the adjoint arc. The conditions of (4.23) represent the adjoint equations of the impulsive portion of the trajectory with the family σ_τ being the adjoint arcs along the jump evolutions. The transversality condition is given by (4.24). The maximum condition for the non-impulsive portion of the trajectory is given by (4.25), and the maximum condition of the impulsive part of the trajectory is given by (4.26).

Remark 4.7. Omitting arguments here for convenience, we discuss the terms $\frac{\partial H}{\partial x}$ and $\frac{\partial}{\partial x} \langle Q, d\mu_{\text{c.p.}} \rangle$ of the above theorem. These terms may be rewritten by definition as $\frac{\partial}{\partial x} \langle f, \psi \rangle$ and $\frac{\partial}{\partial x} \langle g^{Tr} \cdot \psi, d\mu_{\text{c.p.}} \rangle$, respectively. In the present case, such a notation is appropriate as it allows for the matrix and vector products to be written in a well-defined manner, and its clarity rests on the fact that only the functions f and g depend on x. However, the theorem of the next section will involve a partial with respect to time t of f and g, which with the present convention would introduce some ambiguity as other functions in the product depend on t. We set forth here the alternate notation f_x to be the vector whose components are the partial derivative with respect to the one-dimensional variable x of each corresponding component of the vector function f, and similarly, g_x to be the matrix whose components are the partial derivative of each corresponding component of the matrix function g.

4.2.2 Main result: necessary conditions for the impulsive Mayer problem with free end time

We now turn our attention to a Mayer problem whose terminal time T is unfixed or allowed to vary according to what is optimal for the problem. The necessary conditions we derive for this problem are established by reformulating the free end time problem as a fixed end time problem and applying Theorem 4.6. The resulting theorem and its corollaries constitute the main results of this work.

The free end time problem under consideration is

$$(P_v) \qquad \begin{cases} \min \varphi(T, x(T), x(0)) \\ dx(t) = f(t, x(t), u(t))dt + g(t, x(t), u(t))d\vartheta, \quad t \in [0, T], \\ p = (x(T), x(0)) \in C_f \times C_{in}, \end{cases}$$

where $C_f \times C_{in} \in \mathbb{R}^n \times \mathbb{R}^n$ is closed, $u \in \mathcal{U}$ is the conventional control, and ϑ is the impulsive control. Note $\mathcal{U}$ and ϑ are as described in the previous section. The value function φ now depends on the variable terminal time T, so a process for this problem is denoted $x(T, u, \vartheta)$ where $x(\cdot)$ is the trajectory resulting from the policy (T, u, ϑ) and (u, ϑ) are admissible controls applied to the problem for times $t \in [0, T]$.

As in the previous section, the impulsive control ϑ is of the form $(\mu, \nu, \{u_\tau, v_\tau\})$ where the measures μ and ν are assumed to be regular Borel measures with μ taking values in some non-empty closed convex cone $K \subset \mathbb{R}^k$, u_τ is a family of measurable essentially bounded functions on $[0, 1]$ taking values in $\mathbb{R}^m$, and v_τ is a family of vector-valued functions adjoint to (μ, ν). In the fixed end time problem it was tacitly assumed that the measure μ was a regular Borel measure on the interval $[0, T]$ for a fixed T. We are presently faced with the technical choice of assuming that the measure μ has values in K and is a regular Borel measure on $\mathbb{R}^1$. Another option, more appropriate for actual application of the maximum principle, is to assume that μ satisfies this criteria on an interval $[0, T_F]$ such that the optimal T of (P_v) lies in $(0, T_F]$. Observe that either of these options are feasible in the context of measure adjoint functions, whereas the dependence of the graph completion on a known end time T, apparent in the formula for the reparameterization function (3.7), restricts the graph completion option to the latter option.

The data of this problem are the same as for problem (P), although we require a slightly stronger hypothesis:

(H_v) The functions f, g, and φ are continuously differentiable in all arguments.

This stronger hypothesis essentially guarantees more regularity of the functions in the time variable which is intuitively acceptable in light of the fact that problem (P_v) depends on a final time T which is varying rather than fixed.

Theorem 4.8. *Let hypothesis (H_v) be satisfied and let $\hat{x}(\hat{T}, \hat{u}, \hat{\vartheta})$ be an optimal process for (P_v). Then there exists a number $\lambda \geq 0$, a vector-valued function ψ of bounded variation, and for every point $\tau \in Ds(\hat{\vartheta})$, there exists an absolutely continuous vector-valued function σ_τ defined on the interval* $[0, 1]$ *such that*

$$\begin{aligned} \psi(t) = \psi(0) &- \int_0^t \frac{\partial H}{\partial x}(\tau, \hat{x}(\tau), \hat{u}(\tau), \psi(\tau)) d\tau \\ &- \int_{[0,t]} \frac{\partial}{\partial x} \langle Q(\tau, \hat{x}(\tau), \hat{u}(\tau), \psi(\tau)), d\hat{\mu}_{c.p.} \rangle \\ &+ \sum_{\tau \in Ds(\hat{\vartheta}) : \tau \leq t} [\sigma_\tau(1) - \psi(\tau^-)], \qquad \forall t \in (0, \hat{T}]; \end{aligned} \tag{4.27}$$

$$\begin{cases} \frac{d\hat{\chi}_\tau(s)}{ds} = g(\tau, \hat{\chi}_\tau(s), \hat{u}_\tau(s)) \hat{v}_\tau(s) \\ \frac{d\sigma_\tau(s)}{ds} = -\frac{\partial}{\partial x} \langle g(\tau, \hat{\chi}_\tau(s), \hat{u}_\tau(s)) \sigma_\tau(s)), \hat{v}_\tau(s) \rangle \\ \hat{\chi}_\tau(0) = \hat{x}(\tau^-), \quad \sigma_\tau(0) = \psi(\tau^-), \quad s \in [0, 1]; \end{cases} \tag{4.28}$$

$$(-\psi(\hat{T}), \psi(\hat{0})) \in \lambda \frac{\partial \varphi}{\partial p}(\hat{T}, \hat{p}) + N_{C_f \times C_{in}}(\hat{p}) \tag{4.29}$$

$$\max_{u \in U} H(t, \hat{x}(t), u, \psi(t)) = H(t, \hat{x}(t), \hat{u}(t), \psi(t)) \quad a.e.\ t \in [0, \hat{T}]; \tag{4.30}$$

$$\begin{cases} \max_{u \in U} Q(t, \hat{x}(t), u, \psi(t)) \in N_K(0) \quad \forall t \in [0, \hat{T}] \\ \max_{u \in U} Q(\tau, \hat{\chi}_\tau(s), u, \sigma_\tau(s)) \in N_K(0) \quad \forall s \in [0, 1], \quad \forall \tau \in Ds(\hat{\vartheta}), \\ \int\limits_{[0,\hat{T}]} \langle g^{Tr}(\tau, \hat{x}(\tau), \hat{u}(\tau)) \psi(\tau), d\hat{\mu}_{c.p.} \rangle + \\ \qquad \sum\limits_{\tau \in Ds(\vartheta)} \int\limits_{[0,1]} \langle g^{Tr}(\tau, \hat{\chi}_\tau(\xi), \hat{u}_\tau(\xi)) \sigma_\tau(\xi), \hat{v}_\tau(\xi) \rangle \geq 0; \end{cases} \tag{4.31}$$

$$\psi(t)\cdot f(t,\hat{x}(t),\hat{u}(t)) = \int_0^t \langle f_t(\tau,\hat{x}(\tau),\hat{u}(\tau)),\psi(\tau)\rangle d\tau + \int_{[0,t]} \langle g_t^{Tr}(\tau,\hat{x}(\tau),\hat{u}(\tau))\sigma_\tau(\tau), d\mu_{c.p.}\rangle + \sum_{\tau\in Ds(\hat{\vartheta}):\tau\le t} \int_{[0,1]} \langle g_t^{Tr}(\tau,\hat{\chi}_\tau(s),\hat{u}_\tau(s))\sigma_\tau(s),\hat{v}_\tau(s)\rangle ds, \quad a.e.\ t\in[0,\hat{T}]. \tag{4.32}$$

Furthermore, if $\hat{T}\notin Ds(\vartheta)$, then we also have

$$\lambda\frac{\partial\varphi}{\partial t}(\hat{p}) = \langle f(\hat{T},\hat{x}(\hat{T}),\hat{u}(\hat{T})),\psi(\hat{T})\rangle. \tag{4.33}$$

Notice that we have used the partial derivative notation of Remark 4.7 for condition (4.32).

Proof. The theorem is proved by reformulating problem (P_v) as a fixed end time problem resembling problem (P) of this section, relating a maximum of (P_v) to that of (P), and applying the maximum principle of Theorem 4.6 and simplifying the results.

Step 1. Let us form the augmented state variable $\mathbf{x} = (x_0, x)\in\mathbb{R}^{1+n}$ and the augmented conventional control $\mathbf{u} = (u_0, u)\in\mathbb{R}^{1+m}$. The impulsive control $\vartheta = (\mu,\nu,\{u_\tau,v_\tau\})$ will only require its jump-time family of conventional controls u_τ to be augmented to $\mathbf{u}_\tau = (u_\tau^0, u_\tau)\in\mathbb{R}^{1+m}$, which then forms the augmented impulsive control $\boldsymbol{\vartheta} = (\mu,\nu,\{\mathbf{u}_\tau, v_\tau\})$.

Now consider the problem

$$(P_v')\qquad \begin{cases} \min\varphi(\mathbf{x}(\hat{T}),\mathbf{x}(0)) \text{ subject to}\\ dx_0(s) = u_0(s)ds\\ dx(s) = f(x_0(s),x(s),u(x_0(s)))u_0(s)ds\\ \qquad\quad + g(x_0(s),x(s),u(x_0(s)))u_0(s)d\vartheta,\\ \mathbf{p} = (\mathbf{x}(\hat{T}),\mathbf{x}(0))\in([0,2\hat{T}]\times C_f)\times(\{0\}\times C_{in}),\quad s\in[0,\hat{T}], \end{cases}$$

where u_0 is taken to be measurable with respect to Lebesgue measure l and the measure $|\vartheta|$ and such that $u_0(s)\in[1/2,2]$ for a.e. $s\in[0,\hat{T}]$. Also, u and u_τ are supposed to take values in the compact set $U\subset\mathbb{R}^m$. Observe the first measure differential equation of the dynamics has right hand side $u_0(s)ds$, an l-measurable function against Lebesgue measure. Thus, the function $x_0(t) = \int_0^t u_0(s)ds$ is absolutely continuous, and furthermore given that $1/2\le u_0(s)\le 2$ for a.e. $s\in[0,\hat{T}]$, x_0 is an increasing function, and therefore invertible, which takes values in a closed subset of $[0,2\hat{T}]$. The latter deduction is consistent with the endpoint condition on

$\mathbf{x}(T)$ in (P_v'). The fact that x_0 has an inverse will be used below in the second step of the proof.

We point out that the value function φ of (P_v') resembles that of problem (P) in that it depends only on the final and initial values of the state $\mathbf{x}$. The set used to define the endpoint conditions $([0, 2\hat{T}] \times C_f) \times (\{0\} \times C_{in})$ is closed since $C_f \times C_{in}$ is closed in accordance with the statement of problem (P_v).

In conjunction with the augmented states, we define the functions $\mathbf{f} : \mathbb{R} \times \mathbb{R}^{n+1} \times \mathbb{R}^{m+1} \to \mathbb{R}^{n+1}$ and $\mathbf{g} : \mathbb{R} \times \mathbb{R}^{n+1} \times \mathbb{R}^{m+1} \to \mathbb{R}^{n+1} \times \mathbb{R}^{k}$ as

$$\mathbf{f}(\mathbf{x}(s), \mathbf{u}(s)) = \begin{bmatrix} u_0(s) \\ u_0(s) f(x_0(s), x(s), u(x_0(s)) \end{bmatrix}$$

and

$$\mathbf{g}(\mathbf{x}(s), \mathbf{u}(s)) = \begin{bmatrix} \vec{0} \\ u_0(s) g(x_0(s), x(s), u(x_0(s)) \end{bmatrix},$$

where we note that $\mathbf{f}$ is an $(n+1)$-vector and $\mathbf{g}$ is an $(n+1) \times k$-matrix whose top row is all zeros. Observe that since f and g are each continuously differentiable in all arguments, as functions of the augmented variables, $\mathbf{f}$ and $\mathbf{g}$ are readily seen to be continuously differentiable in all augmented variables. Thus, hypothesis (H_v) is met by the dynamics of (P_v'). The functions H and Q defined in the previous section are analogously extended to the augmented functions $\mathbf{H}$ and $\mathbf{Q}$ taking arguments among $\mathbf{x}$, $\mathbf{u}$, $\mathbf{u}_\tau$ $\boldsymbol{\psi}$, and $\boldsymbol{\sigma}_\tau$.

We may now express, in terms of the augmented functions and variables defined, the dynamics of (P_v') as the single equation

$$d\mathbf{x}(s) = \mathbf{f}(\mathbf{x}(s), \mathbf{u}(s))ds + \mathbf{g}(\mathbf{x}(s), \mathbf{u}(s))d\boldsymbol{\vartheta} \quad s \in [0, \hat{T}], \tag{4.34}$$

which demonstrates that the augmented system is indeed in the form of the dynamics of problem (P).

Lastly, observe that x_0 in the dynamics of (P_v') plays the role of the time variable, similar to what was shown in the sketch of the proof of Theorem 2.6.

Step 2. Let $(\hat{T}, \hat{u}, \hat{\vartheta})$ be a minimizing policy for (P_v). We show that the augmented policy $(\hat{\mathbf{u}}, \hat{\boldsymbol{\vartheta}}) = ((1, \hat{u}), (\hat{\mu}, \hat{\nu}, \{(1, \hat{u}_\tau), \hat{v}_\tau\}))$, where $u_0 \equiv 1$ and $u_{\tau,0} \equiv 1$, is optimal for problem (P_v'). Suppose for contradiction, that

$$(\mathbf{w}, \boldsymbol{\omega}) := ((w_0, w), (\mu_\omega, \nu_\omega, \{(w_\tau^0, w_\tau), z_\tau\}))$$

is another admissible control such that

$$\varphi(\mathbf{x}(\hat{T}, \mathbf{w}, \boldsymbol{\omega}), \mathbf{x}(0, \mathbf{w}, \boldsymbol{\omega})) < \varphi(\hat{T}, x(\hat{T}, \hat{u}, \hat{\vartheta}), x(0, \hat{u}, \hat{\vartheta})). \tag{4.35}$$

Since x_0 is invertible, we may take its inverse function $\tau(t)$ and construct the control

$$(u^{\#}, \vartheta^{\#})(t) = (v, \omega)(\tau(t)),$$

where the notation $(u, \vartheta)(t)$ is shorthand for applying a conventional control u and an impulsive control ϑ to the dynamics of (P_v) for time t.

Applying this control to the dynamics of (P_v) with terminal time $T = x_0(\hat{T})$, we obtain by (4.35)

$$\varphi(T, x(T, u^{\#}, \vartheta^{\#}), x(0, u^{\#}, \vartheta^{\#})) < \varphi(\hat{T}, x(\hat{T}, \hat{u}, \hat{\vartheta}), x(0, \hat{u}, \hat{\vartheta})),$$

which violates the optimality of $(\hat{T}, \hat{u}, \hat{\vartheta})$.

Step 3. We are now in a position to apply Theorem 4.6 to problem (P_v') for the optimal process $(\hat{\mathbf{x}}, \hat{\mathbf{u}}, \hat{\boldsymbol{\vartheta}})$. Before doing so, we point out some important facts and notational conventions to simplify the analysis.

- We state the implied conditions of the theorem in terms of $(n+1)$-vectors with the 0^{th}-component expanded from the original n-vectors of (P_v). In some places, it is more convenient to leave the expressions in the augmented or bold form, in which case the terms will be expanded upon investigation.

- We will use the partial derivative notation introduced in Remark 4.7 wherever such clarity is needed.

- Since $\hat{u}_0 \equiv 1$, the optimal trajectory of the reparameterized time x_0 is the identity function $\hat{x}_0(s) = s$, which reduces the reparameterized time scale $[0, x_0(T)]$ back to the original time scale $[0, T]$. This being the case, some statements will initially involve $\hat{x}_0(\cdot)$ to show the immediate implication of Theorem 4.6, whereas other statements may be immediately reduced to t. This is done in order to provide sufficient details without cluttering the proof, and ultimately all instances of $\hat{x}_0(\cdot)$ will be reduced to t.

The theorem implies the existence of a number $\lambda \geq 0$, a vector-valued function $\boldsymbol{\psi} = (\psi_0, \psi) \in \mathbb{R}^{1+n}$ of bounded variation, and for every $\tau \in \mathrm{Ds}(\hat{\vartheta})$ an absolutely continuous vector-valued function $\boldsymbol{\sigma}_\tau = (\sigma_\tau^0, \sigma_\tau) \in \mathbb{R}^{1+n}$ defined on $[0,1]$ such that

$$\begin{aligned}
\begin{bmatrix} \psi_0(s) \\ \psi(s) \end{bmatrix} = \begin{bmatrix} \psi_0(0) \\ \psi(0) \end{bmatrix} &- \int_0^s \begin{bmatrix} \frac{\partial \mathbf{H}}{\partial x_0}(\hat{\mathbf{x}}(\tau), \hat{\mathbf{u}}(\tau), \boldsymbol{\psi}(\tau)) \\ \frac{\partial \mathbf{H}}{\partial x}(\hat{\mathbf{x}}(\tau), \hat{\mathbf{u}}(\tau), \boldsymbol{\psi}(\tau)) \end{bmatrix} d\tau \\
&- \int_{[0,s]} \begin{bmatrix} \frac{\partial}{\partial x_0} \langle \mathbf{Q}(\hat{\mathbf{x}}(\tau), \hat{\mathbf{u}}(\tau), \boldsymbol{\psi}(\tau)), d\hat{\mu}_{\text{c.p.}} \rangle \\ \frac{\partial}{\partial x} \langle \mathbf{Q}(\hat{\mathbf{x}}(\tau), \hat{\mathbf{u}}(\tau), \boldsymbol{\psi}(\tau)), d\hat{\mu}_{\text{c.p.}} \rangle \end{bmatrix} \\
&+ \sum_{\tau \in \mathrm{Ds}(\hat{\vartheta}) : \tau \leq \hat{x}_0(s)} \begin{bmatrix} \sigma_\tau^0(1) - \psi_0(\tau^-) \\ \sigma_\tau(1) - \psi(\tau^-) \end{bmatrix}, \qquad \forall s \in (0, \hat{T}];
\end{aligned} \tag{4.36}$$

$$\begin{cases}
\frac{d}{d\xi} \begin{bmatrix} \hat{\chi}_\tau^0(\xi) \\ \hat{\chi}_\tau(\xi) \end{bmatrix} &= \begin{bmatrix} 0 \\ g(\tau, \hat{\chi}_\tau(\xi), \hat{u}_\tau(\xi)) \cdot \hat{v}_\tau(\xi) \end{bmatrix} \\
\frac{d}{d\xi} \begin{bmatrix} \sigma_\tau^0(\xi) \\ \sigma_\tau(\xi) \end{bmatrix} &= \begin{bmatrix} -\frac{\partial}{\partial x_0} \langle \mathbf{Q}(\tau, \hat{\boldsymbol{\chi}}_\tau(\xi), \hat{\mathbf{u}}_\tau(\xi), \boldsymbol{\sigma}_\tau(\xi)), \hat{v}_\tau(\xi) \rangle \\ -\frac{\partial}{\partial x} \langle \mathbf{Q}(\tau, \hat{\boldsymbol{\chi}}_\tau(\xi), \hat{\mathbf{u}}_\tau(\xi), \boldsymbol{\sigma}_\tau(\xi)), \hat{v}_\tau(\xi) \rangle \end{bmatrix} \\
\begin{bmatrix} \hat{\chi}_\tau^0(0) \\ \hat{\chi}_\tau(0) \end{bmatrix} &= \begin{bmatrix} \hat{x}_0(\tau^-) \\ \hat{x}(\tau^-) \end{bmatrix}, \quad \begin{bmatrix} \sigma_\tau^0(0) \\ \sigma_\tau(0) \end{bmatrix} = \begin{bmatrix} \psi_0(\tau^-) \\ \psi(\tau^-) \end{bmatrix}, \quad \xi \in [0,1];
\end{cases} \tag{4.37}$$

$$\left(\begin{bmatrix}-\psi_0(\hat{T})\\ -\psi(\hat{T})\end{bmatrix},\begin{bmatrix}\psi_0(0)\\ \psi(0)\end{bmatrix}\right)\in\lambda\begin{bmatrix}\frac{\partial\varphi}{\partial p_0}(\hat{\mathbf{p}})\\ \frac{\partial\varphi}{\partial p}(\hat{\mathbf{p}})\end{bmatrix}+\begin{bmatrix}N_{[0,2\hat{T}]\times\{0\}}(\hat{p}_0)\\ N_{C_f\times C_{in}}(\hat{p})\end{bmatrix}, \tag{4.38}$$

$$\max_{\mathbf{u}\in[\frac{1}{2},2]\times U}\left\langle\begin{bmatrix}u_0\\ u_0\cdot f(\hat{x}_0(s),\hat{x}(s),u)\end{bmatrix},\begin{bmatrix}\psi_0(s)\\ \psi(s)\end{bmatrix}\right\rangle=\left\langle\begin{bmatrix}1\\ f(\hat{x}_0(s),\hat{x}(s),u(s))\end{bmatrix},\begin{bmatrix}\psi_0(s)\\ \psi(s)\end{bmatrix}\right\rangle \quad \text{a.e. } s\in[0,\hat{T}]; \tag{4.39}$$

$$\begin{cases}\max_{\mathbf{u}\in[\frac{1}{2},2]\times U}\begin{bmatrix}0 & u_0\cdot g^{Tr}(\hat{x}_0(s),\hat{x}(s),u)\end{bmatrix}\begin{bmatrix}\psi_0(s)\\ \psi(s)\end{bmatrix}\in N_K(0) \quad \forall s\in[0,\hat{T}]\\ \max_{\mathbf{u}\in[\frac{1}{2},2]\times U}\begin{bmatrix}0 & u_0\cdot g^{Tr}(\tau,\hat{\chi}(\xi),u)\end{bmatrix}\begin{bmatrix}\sigma^0_\tau(\xi)\\ \sigma_\tau(\xi)\end{bmatrix}\in N_K(0) \quad \forall\xi\in[0,1], \quad \forall\tau\in\mathrm{Ds}(\hat{\vartheta}),\\ \displaystyle\int_{[0,\hat{T}]}\langle g^{\mathrm{Tr}}(\hat{x}_0(s),\hat{x}(s),\hat{u}(s))\psi(s),d\hat{\mu_{\text{c.p.}}}\rangle+\\ \qquad\displaystyle\sum_{\tau\in\mathrm{Ds}(\vartheta)}\int_{[0,1]}\langle g^{\mathrm{Tr}}(\tau,\hat{x}(\xi),\hat{u}(\xi))\sigma_\tau(\xi),\hat{v}_\tau(\xi)\rangle\geq 0.\end{cases} \tag{4.40}$$

We first expand the ψ_0 component of the non-jump adjoint equation represented by (4.36) to get for all $s\in(0,\hat{T}]$

$$\begin{aligned}\psi_0(s)=\psi_0(0)&-\int_0^s\frac{\partial}{\partial x_0}\left\langle\begin{bmatrix}1\\ f(\tau,\hat{x}(\tau),\hat{u}(\tau))\end{bmatrix},\begin{bmatrix}\psi_0(\tau)\\ \psi(\tau)\end{bmatrix}\right\rangle d\tau\\ &-\int_{[0,s]}\frac{\partial}{\partial x_0}\left\langle\begin{bmatrix}0 & g^{Tr}(\tau,\hat{x}(\tau),\hat{u}(\tau))\end{bmatrix}\begin{bmatrix}\psi_0(\tau)\\ \psi(\tau)\end{bmatrix},d\mu_{\text{c.p.}}\right\rangle\\ &+\sum_{\tau\in\mathrm{Ds}(\hat{\vartheta}):\tau\leq\hat{x}_0(s)}\sigma^0_\tau(1)-\psi_0(\tau^-),\end{aligned}$$

which, taking into account that $\frac{\partial}{\partial x_0}[1]=\frac{\partial}{\partial x_0}[0]=0$, simplifies to

$$\begin{aligned}\psi_0(s)=\psi_0(0)&-\int_0^s\frac{\partial}{\partial x_0}\langle f(\tau,\hat{x}(\tau),\hat{u}(\tau)),\psi(\tau)\rangle\,d\tau\\ &-\int_{[0,s]}\frac{\partial}{\partial x_0}\langle g^{Tr}(\tau,\hat{x}(\tau),\hat{u}(\tau))\psi(\tau),d\mu_{\text{c.p.}}\rangle\\ &+\sum_{\tau\in\mathrm{Ds}(\hat{\vartheta}):\tau\leq\hat{x}_0(s)}\sigma^0_\tau(1)-\psi_0(\tau^-).\end{aligned} \tag{4.41}$$

The term $\begin{bmatrix}0 & g^{Tr}(\tau,\hat{x}(\tau),\hat{u}(\tau))\end{bmatrix}$ appearing in the second integral of each equation is the $k\times(n+1)$-matrix whose first column is all zeros and whose remaining columns are formed by the $k\times n$-matrix g^{Tr}.

We do the same for the n-vector, ψ, of (4.36) to get

$$\begin{aligned}\psi(s) = \psi(0) &- \int_0^s \frac{\partial}{\partial x}\left\langle f(\tau,\hat{x}(\tau),\hat{u}(\tau)),\psi(\tau)\right\rangle d\tau \\ &- \int_{[0,s]} \frac{\partial}{\partial x}\left\langle g^{Tr}(\tau,\hat{x}(\tau),\hat{u}(\tau))\psi(\tau), d\mu_{\text{c.p.}}\right\rangle \\ &+ \sum_{\tau\in \mathrm{Ds}(\hat{\vartheta}):\tau\le\hat{x}_0(s)} \sigma_\tau(1)-\psi(\tau^-).\end{aligned} \tag{4.42}$$

Recalling the definitions of the functions H and Q and replacing the dummy variable s with t, we see that (4.42) yields condition (4.27) of the theorem.

Similarly, we expand the adjoint equations of the jump dynamics represented by (4.37) to get for all $\xi \in [0,1]$

$$\frac{d\hat{\chi}^0_\tau(\xi)}{d\xi} = 0 \tag{4.43}$$

for the χ^0_τ component,

$$\frac{d\hat{\chi}_\tau(\xi)}{d\xi} = g(\hat{\chi}^0_\tau(\tau),\hat{\chi}_\tau(\xi),\hat{u}_\tau(\xi))\cdot\hat{v}_\tau(\xi) \tag{4.44}$$

for the χ_τ vector,

$$\begin{aligned}\frac{d\sigma^0_\tau(\xi)}{d\xi} &= -\frac{\partial}{\partial x_0}\langle\begin{bmatrix}0 & g^{Tr}(\hat{\chi}_0(\xi),\hat{\chi}_\tau(\xi),\hat{u}_\tau(\xi))\end{bmatrix}\begin{bmatrix}\sigma^0_\tau(\xi))\\ \sigma_\tau(\xi))\end{bmatrix},\hat{v}_\tau(\xi)\rangle \\ &= -\frac{\partial}{\partial x_0}\langle g^{Tr}(\hat{\chi}_0(\xi),\hat{\chi}_\tau(\xi),\hat{u}_\tau(\xi))\cdot\sigma_\tau(\xi)),\hat{v}_\tau(\xi)\rangle\end{aligned} \tag{4.45}$$

for the 0^{th}-component σ^0_τ, and

$$\begin{aligned}\frac{d\sigma_\tau(\xi)}{d\xi} &= -\frac{\partial}{\partial x}\langle\begin{bmatrix}0 & g^{Tr}(\hat{\chi}_0(\xi),\hat{\chi}_\tau(\xi),\hat{u}_\tau(\xi))\end{bmatrix}\begin{bmatrix}\sigma^0_\tau(\xi))\\ \sigma_\tau(\xi))\end{bmatrix},\hat{v}_\tau(\xi)\rangle \\ &= -\frac{\partial}{\partial x}\langle g^{Tr}(\hat{\chi}_0(\xi),\hat{\chi}_\tau(\xi),\hat{u}_\tau(\xi))\cdot\sigma_\tau(\xi)),\hat{v}_\tau(\xi)\rangle\end{aligned} \tag{4.46}$$

for the n-vector σ_τ.

According to the initial conditions of (4.37) and the differential equation (4.43), for each $\tau \in Ds(\vartheta)$, we see that $\hat{\chi}^0_\tau$ is the constant function $\hat{\chi}^0_\tau \equiv \hat{x}_0(\tau^-)$. Since $\hat{x}_0$ is the identity function, we have $\hat{\chi}^0_\tau = \hat{x}_0(\tau^-) = \tau$. Using this fact in equations (4.44) and (4.46) yields the two differential equations of (4.28), and the initial conditions corresponding to each of these equations in (4.37) provide the initial conditions of (4.28).

Next we rewrite the maximum condition given by (4.39) as

$$\begin{aligned}&\max_{\mathbf{u}\in[\frac{1}{2},2]\times U}\{u_0\cdot\psi_0(s)+\langle u_0\cdot f(\hat{x}_0(s),\hat{x}(s),u),\psi(s)\rangle\} \\ &\qquad = 1\cdot\psi_0(s)+\langle f(\hat{x}_0(s),\hat{x}(s),\hat{u}(s)),\psi(s)\rangle \quad \text{a.e. } s\in[0,\hat{T}].\end{aligned} \tag{4.47}$$

This equation tells us $\hat{u}_0 \equiv 1$ and $\hat{u}(\cdot)$ are maximal for the function $u_0 \cdot \psi_0(s) + \langle u_0 \cdot f(\hat{x}_0(s), \hat{x}(s), u), \psi(s)\rangle$, so the derivative with respect to u_0 of this function evaluated at $(u_0, u(\cdot)) = (1, \hat{u}(\cdot))$ must be zero. This implies

$$\psi_0(t) = -\langle f(t, \hat{x}(t), \hat{u}(t)), \psi(t)\rangle, \quad \text{for a.e. } t \in [0, \hat{T}] \tag{4.48}$$

after reducing $\hat{x}_0(s)$ to the original t time.

Observe that integrating (4.45) and applying the initial condition $\sigma_\tau^0(0) = \psi_0(\tau^-)$ provides

$$\sigma_\tau^0(s) = \psi_0(\tau^-) - \int_{[0,s]} \langle g^{Tr}(\hat{\chi}_0(\xi), \hat{\chi}_\tau(\xi), \hat{u}_\tau(\xi)) \cdot \sigma_\tau(\xi)), \hat{v}_\tau(\xi)\rangle. \tag{4.49}$$

Reducing the time in equation (4.41) back to t, combining this with equation (4.48) and substituting σ_τ^0 with the right hand side of (4.49) yields condition (4.32) of the theorem. Also, reducing the time in (4.47) back to t implies condition (4.30).

Observe that simplifying the multiplication with the 0 vector in each of the first two conditions of (4.40) gives

$$\max_{\mathbf{u} \in [\frac{1}{2}, 2] \times U} u_0 \cdot g^{Tr}(t, \hat{x}(t), u)\psi(t) \in N_K(0) \quad \forall t \in [0, \hat{T}] \tag{4.50}$$

and

$$\max_{\mathbf{u} \in [\frac{1}{2}, 2] \times U} u_0 \cdot g^{Tr}(\tau, \hat{\chi}(\xi), u)\sigma_\tau(\xi) \in N_K(0) \quad \forall \xi \in [0, 1], \quad \forall \tau \in \mathrm{Ds}(\hat{\vartheta}), \tag{4.51}$$

where in the former inclusion we have again reduced $\hat{x}_0$ to time t. Since each of the above maximums are taken over a $u_0 > 0$, we may apply (2.9) to each and by definition of Q get both maximum conditions of (4.31). Note that the multiplication by zero in the inequality of (4.40) was implicitly simplified in writing it, and by reducing $\hat{x}_0$ back to original time t, we obtain the inequality of (4.31).

Using (2.8), we may consider the n-vector component, ψ, of the inclusion (4.38) and obtain condition (4.29)

$$(-\psi(\hat{T}), \psi(0)) \in \frac{\partial \varphi}{\partial p}(\hat{T}, \hat{p}) + N_{C_f \times C_{in}}(\hat{p}),$$

since

$$\begin{aligned} \varphi(\hat{\mathbf{p}}) &= \varphi(\hat{x}_0(T), \hat{x}(T), \hat{x}_0(0), \hat{x}(0)) \\ &= \varphi(\hat{T}, \hat{x}(T), \hat{x}(0)) \\ &= \varphi(\hat{T}, \hat{p}). \end{aligned}$$

Note in the above that the second equality is due to φ's independence of the variable x_0^{in}, and that this independence also yields $\frac{\partial \varphi}{\partial x_0^{in}}(\hat{p}) = 0$.

With another appeal to (2.8) and arguing similarly, the 0-components of the transversality condition (4.38) can be expressed as

$$-\psi_0(\hat{T}) = \lambda \frac{\partial \varphi}{\partial x_0^f}(\hat{T}, \hat{p}) + N_{[0,2\hat{T}]}(\hat{x}_0(\hat{T}))$$

and

$$\psi_0(0)) = \lambda \frac{\partial \varphi}{\partial x_0^{in}}(\hat{T}, \hat{p}) + N_{\{0\}}(\hat{x}_0(0)).$$

Observe that the variable x_0^f is just the end-time variable T. Recall that $\hat{x}_0(\hat{T}) = \hat{T}$ and $\hat{x}_0(0) = 0$, and use (2.7) to get $N_{[0,2\hat{T}]}(\hat{x}_0(\hat{T})) = \emptyset$ and (2.10) to get $N_{\{0\}}(\hat{x}_0(0)) = \mathbb{R}$. Thus, the above inclusions yield respectively

$$-\psi_0(\hat{T}) = \lambda \frac{\partial \varphi}{\partial t}(\hat{p}) \tag{4.52}$$

and

$$\psi_0(0) \in \mathbb{R}. \tag{4.53}$$

Observe that (4.41) implies that ψ_0 is continuous at points $t \notin Ds(\vartheta)$. Thus, if $\hat{T} \notin Ds(\vartheta)$, then (4.48) holds at $\hat{T}$ which with (4.52) gives

$$\lambda \frac{\partial \varphi}{\partial t}(\hat{p}) = \langle f(\hat{T}, \hat{x}(\hat{T}), \hat{u}(\hat{T})), \psi(\hat{T}) \rangle.$$

Thus, we have condition (4.33).

□

Remark 4.9. Observe that the technique in the above proof of letting x_0 represent a new time scale which depends on u_0 is possible since the function g in the measure-adjoint minimization problem may depend on the control u and subsequently $\mathbf{g}$ may depend on (u_0, u). On the other hand, the function g of the graph completion minimization problem in section 4.1 may not depend on u, so another technique or deeper analysis is required to extend this PMP of the graph completion problem to one with a free end-time.

4.2.3 Corollaries: necessary conditions for additional optimization problems containing impulses

We now use the free end-time Mayer problem to derive necessary conditions for a Bolza problem with impulsive dynamics and impulsive Lagrangian. In turn, we will use the necessary conditions of the Bolza problem to derive necessary conditions for the minimum time problem whose objective function may involve penalties on the impulses of the system. Such a penalization is important for the minimum time problem in order to preclude trivial optimal processes which attain $T = 0$ as the minimum time by evolving only along jump dynamics. Last, we comment on the basic problem of the Calculus of Variations within the impulsive context. All of these derivations are described in classical form in section 2.3.2 above.

The Bolza problem. Consider the Bolza problem of the form

$$(P_B)\begin{cases}\min\limits_{u\in U}\int_0^T L(t,x(t),u(t))dt+\int_{[0,T]}\langle I(t,x(t),u(t)),d\vartheta\rangle\\ \text{subject to}\\ dx(t)=f(t,x(t),u(t))dt+g(t,x(t),u(t))d\vartheta,\quad \text{for a.e. } t\in[0,T]\\ p=(x(T),x(0))\in C_f\times C_{in},\end{cases}\tag{4.54}$$

where the functions f and g, the sets $C_f \times C_{in}$ and U, and the impulsive control ϑ satisfy all definitions and assumptions of section 4.2.2. The scalar-valued Lagrangian L is assumed to be continuous in all variables and continuously differentiable with respect to t, x. The end time T is variable, and the function $I : \mathbb{R}\times\mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R}^k\times\mathbb{R}$ is assumed to be continuously differentiable in all variables.

We follow the reformulation suggested in section 2.3.2 and construct the variable

$$x_{n+1}(t)=\int_0^t L(\tau,x(\tau),u(\tau))d\tau+\int_{[0,T]}\langle I(t,x(t),u(t)),d\vartheta\rangle$$

to form the auxiliary minimization problem

$$\min x_{n+1}(T)\tag{4.55}$$

over $u \in \mathcal{U}$ and admissible impulsive controls ϑ, subject to the auxiliary system

$$\begin{cases}dx(t)=f(t,x(t),u(t))dt+g(t,x(t),u(t))d\vartheta\\ dx_{n+1}(t)=L(t,x(t),u(t))dt+\langle I(t,x(t),u(t)),d\vartheta\rangle,\end{cases}\tag{4.56}$$

for almost every $t\in[0,T]$. The initial condition of the auxiliary system will be

$$\mathbf{p}=(\mathbf{x}(T),\mathbf{x}(0))\in(C_f\times\mathbb{R})\times(C_{in}\times\{0\}),\tag{4.57}$$

where we adopt the bold notation of the previous section as the augmented $(n+1)$-dimensional state variable $\mathbf{x}=(x,x_{n+1})$. It is straightforward to verify that the dynamics of the augmented system inherit the regularity of the functions f, g, L, and I as demonstrated for equation (4.34) in the above proof. Thus, we may apply the maximum principle of Theorem 4.8 to obtain the necessary conditions for problem (4.55)-(4.57) in the form of the next corollary.

Corollary 4.10. *Let $(\hat{T},\hat{u},\hat{\vartheta})$ be an optimal policy for the problem (P_B) with corresponding optimal trajectory $\hat{x}$. Let (H_v) be satisfied, and let the Lagrangian L be continuous in all variables and continuously differentiable with respect to t, x, and let I be continuously differentiable in all variables. Then there exists a number $\lambda \geq 0$, a vector-valued function ψ of bounded variation, and for every point in $\tau \in Ds(\hat{\vartheta})$, there exists an absolutely continuous vector-valued function σ_τ defined on the interval $[0,1]$ such that*

$$\psi(t) = \psi(0) - \int_0^t \frac{\partial}{\partial x}\langle f(\tau,\hat{x}(\tau),\hat{u}(\tau)),\psi(\tau)\rangle + \lambda L_x(\tau,\hat{x}(\tau),\hat{u}(\tau))\ d\tau$$
$$- \int_{[0,t]} \frac{\partial}{\partial x}\langle g^{Tr}(\tau,\hat{x}(\tau),\hat{u}(\tau))\psi(\tau) + \lambda I(\tau,\hat{x}(\tau),\hat{u}(\tau)), d\mu_{c.p.}\rangle \quad (4.58)$$
$$+ \sum_{\tau\in Ds(\vartheta):\tau\le t} [\sigma_\tau(1) - \psi(\tau^-)] \quad \forall t \in (0,\hat{T}];$$

$$\begin{cases} \frac{d\hat{\chi}_\tau(s)}{ds} = g(\tau,\hat{\chi}_\tau(s),\hat{u}_\tau(s))\hat{v}_\tau(s) \\ \frac{d\sigma_\tau(s)}{ds} = -\frac{\partial}{\partial x}\langle g^{Tr}(\tau,\hat{\chi}_\tau(s),\hat{u}_\tau(s))\sigma_\tau(s) + \lambda I(\tau,\hat{\chi}_\tau(s),\hat{u}_\tau(s)),\hat{v}_\tau(s)\rangle \\ (\hat{\chi}_\tau(0) = x(\tau^-), \quad \sigma_\tau(0) = \psi_\tau(\tau^-), \quad s\in[0,1]; \end{cases} \quad (4.59)$$

$$(-\psi(\hat{T}),\psi(0)) \in N_{C_f\times C_{in}}(\hat{p}); \quad (4.60)$$

$$\max_{u\in U}\{\langle f(t,\hat{x}(t),u),\psi(t)\rangle + \lambda L(t,\hat{x}(t),u)\} =$$
$$\langle f(t,\hat{x}(t),\hat{u}(t)),\psi(t)\rangle + \lambda L(t,\hat{x}(t),\hat{u}(t)) \quad (4.61)$$
$$a.e.\ t\in[0,\hat{T}];$$

$$\begin{cases} \max_{u\in U} g^{Tr}(t,\hat{x}(t),u)\psi(t) + \lambda I(t,\hat{x}(t),u) \in N_K(0) \quad \forall t\in[0,\hat{T}], \\ \max_{u\in U} g^{Tr}(\tau,\hat{\chi}_\tau(s),u)\sigma_\tau(s) + \lambda I(\tau,\hat{\chi}_\tau(s),u) \in N_K(0) \quad \forall s\in[0,1], \quad \forall\tau\in Ds(\vartheta), \\ \int_{[0,\hat{T}]}\langle g^{Tr}(\tau,\hat{x}(\tau),\hat{u}(\tau))\psi(\tau) + \lambda I(\tau,\hat{x}(\tau),\hat{u}(\tau)), d\hat{\mu}_{c.p.}\rangle + \\ \sum_{\tau\in Ds(\hat{\vartheta})} \int_{[0,1]}\langle g^{Tr}(\tau,\hat{\chi}_\tau(s),\hat{u}_\tau(s))\sigma_\tau(s) + \lambda I(\tau,\hat{\chi}_\tau(s),\hat{u}_\tau(s)),\hat{v}_\tau(s)\rangle ds \ge 0; \end{cases}$$
$$(4.62)$$

$$\psi(t)\cdot f(t,\hat{x}(t),\hat{u}(t)) + \lambda L(t,\hat{x}(t),\hat{u}(t)) =$$
$$\int_0^t \langle f_t(\tau,\hat{x}(\tau),\hat{u}(\tau)),\psi(\tau)\rangle + \lambda L_t(\tau,\hat{x}(\tau),\hat{u}(\tau))d\tau$$
$$+ \int_{[0,t]} \langle g_t^{Tr}(\tau,\hat{x}(\tau),\hat{u}(\tau))\psi(\tau) + \lambda I_t(\tau,\hat{x}(\tau),\hat{u}(\tau)), d\mu_{c.p.}\rangle$$
$$+ \sum_{\tau\in Ds(\hat{\vartheta}):\tau\le t} \int_{[0,1]} \langle g_t^{Tr}(\tau,\hat{\chi}_\tau(s),\hat{u}_\tau(s))\sigma_\tau(s) + \lambda I_t(\tau,\hat{\chi}_\tau(s),\hat{u}_\tau(s)),\hat{v}_\tau(s)\rangle ds,$$
$$a.e.\ t\in[0,\hat{T}]. \quad (4.63)$$

Furthermore, if $\hat{T} \notin Ds(\hat{\vartheta})$, *then*

$$\lambda\int_0^{\hat{T}} L_t(t,\hat{x}(t),\hat{u}(t))dt + \lambda\int_{[0,\hat{T}]} \langle I_t(t,\hat{x}(t),\hat{u}(t)), d\hat{\vartheta}\rangle = \langle f(\hat{T},\hat{x}(\hat{T}),\hat{u}(\hat{T})),\psi(\hat{T})\rangle,$$
$$(4.64)$$

where L_t is the partial derivative of L with respect to t and likewise for I_t.

Proof. The reformulation (4.55)-(4.57) satisfies the conditions of Theorem 4.8, which implies there exist a number $\lambda \geq 0$, a vector-valued function of bounded variation (ψ, ψ_{n+1}), and for every $\tau \in Ds(\vartheta)$ an absolutely continuous, vector-valued function $(\sigma_\tau, \sigma_\tau^{n+1})$ such that

$$\begin{aligned}\psi(t) = \psi(0) - &\int_0^t \frac{\partial}{\partial x}\langle f(\tau, \hat{x}(\tau), \hat{u}(\tau)), \psi(\tau)\rangle + L_x(\tau, \hat{x}(\tau), \hat{u}(\tau))\psi_{n+1}(\tau)\ d\tau \\ &- \int_{[0,t]} \frac{\partial}{\partial x}\langle g^{Tr}(\tau, \hat{x}(\tau), \hat{u}(\tau))\psi(\tau) + I(\tau, \hat{x}(\tau), \hat{u}(\tau))\psi_{n+1}(\tau), d\mu_{c.p.}\rangle \\ &+ \sum_{\tau \in Ds(\vartheta):\tau \leq t} [\sigma_\tau(1) - \psi(\tau^-)], \\ \psi_{n+1}(t) =& \psi_{n+1}(0) + \sum_{\tau \in Ds(\vartheta):\tau \leq t} [\sigma_\tau^{n+1}(1) - \psi_{n+1}(\tau^-)], \quad \forall t \in (0, \hat{T}];\end{aligned} \tag{4.65}$$

$$\begin{cases} \frac{d\hat{\chi}_\tau(s)}{ds} = g(\tau, \hat{\chi}_\tau(s), \hat{u}_\tau(s))\hat{v}_\tau(s) \\ \frac{d\hat{\chi}_\tau^{n+1}(s)}{ds} = \langle I(\tau, \hat{\chi}_\tau(s), \hat{u}_\tau(s)), \hat{v}_\tau(s)\rangle \\ \frac{d\sigma_\tau(s)}{ds} = -\frac{\partial}{\partial x}\langle g^{Tr}(\tau, \hat{\chi}_\tau(s), \hat{u}_\tau(s))\sigma_\tau(s) + I(\tau, \hat{\chi}_\tau(s), \hat{u}_\tau(s))\sigma_\tau^{n+1}(s), \hat{v}_\tau(s)\rangle \\ \frac{d\sigma_\tau^{n+1}(s)}{ds} = -\langle g_{x_{n+1}}^{Tr}(\tau, \hat{\chi}_\tau(s), \hat{u}_\tau(s))\sigma_\tau(s) + I_{x_{n+1}}(\tau, \hat{\chi}_\tau(s), \hat{u}_\tau(s))\sigma_\tau^{n+1}(s), \hat{v}_\tau(s)\rangle = 0 \\ (\hat{\chi}_\tau(0), \hat{\chi}_\tau^{n+1}(0)) = (x(\tau^-), x_{n+1}(\tau^-)) \\ (\sigma_\tau(0), \sigma_\tau^{n+1}(0)) = (\psi(\tau^-), \psi_{n+1}(\tau^-)), \quad s \in [0,1]; \end{cases} \tag{4.66}$$

$$\begin{aligned}(-\psi(\hat{T}), \psi(0)) &\in \lambda\frac{\partial \varphi}{\partial p}(\hat{\mathbf{p}}) + N_{C_f \times C_{in}}(\hat{p}) \\ (-\psi_{n+1}(\hat{T}), \psi_{n+1}(0)) &\in \frac{\partial \varphi}{\partial p_{n+1}}(\hat{\mathbf{p}}) + N_{\mathbb{R}\times\{0\}}(\hat{p}_{n+1});\end{aligned} \tag{4.67}$$

$$\begin{aligned}\max_{u \in U}\{&\langle f(t, \hat{x}(t), u), \psi(t)\rangle + L(t, \hat{x}(t), u)\psi_{n+1}(t)\} = \\ &\langle f(t, \hat{x}(t), \hat{u}(t)), \psi(t)\rangle + L(t, \hat{x}(t), \hat{u}(t))\psi_{n+1}(t) \\ &\text{for a.e. } t \in [0, \hat{T}];\end{aligned} \tag{4.68}$$

$$\begin{cases} \max_{u \in U} g^{Tr}(t, \hat{x}(t), u)\psi(t) + I(t, \hat{x}(t), u)\psi_{n+1}(t) \in N_K(0) \quad \forall t \in [0, \hat{T}], \\ \max_{u \in U} g^{Tr}(\tau, \hat{\chi}_\tau(s), u)\sigma_\tau(s) + I(\tau, \hat{\chi}_\tau(s), u)\sigma_\tau^{n+1}(s) \in N_K(0) \quad \forall s \in [0,1], \\ \qquad \forall \tau \in Ds(\vartheta), \\ \int_{[0,\hat{T}]} \langle g^{Tr}(\tau, \hat{x}(\tau), \hat{u}(\tau))\psi(\tau) + I(\tau, \hat{x}(\tau), \hat{u}(\tau))\psi_{n+1}(\tau), d\hat{\mu}_{c.p.}\rangle + \\ \sum_{\tau \in Ds(\hat{\vartheta})} \int_{[0,1]} \langle g^{Tr}(\tau, \hat{\chi}_\tau(s), \hat{u}_\tau(s))\sigma_\tau(s) + I(\tau, \hat{\chi}_\tau(s), \hat{u}_\tau(s))\sigma_\tau^{n+1}(s), \hat{v}_\tau(s)\rangle ds \geq 0; \end{cases} \tag{4.69}$$

$$\begin{aligned} \psi(t) \cdot f(t, \hat{x}(t), \hat{u}(t)) + \psi_{n+1}(t)L(t, \hat{x}(t), \hat{u}(t)) = \\ \int_0^t \langle f_t(\tau, \hat{x}(\tau), \hat{u}(\tau)), \psi(\tau)\rangle + \psi_{n+1}(\tau)L_t(\tau, \hat{x}(\tau), \hat{u}(\tau))d\tau \\ + \int_{[0,t]} \langle g_t^{\mathrm{Tr}}(\tau, \hat{x}(\tau), \hat{u}(\tau))\psi(\tau) + I_t(\tau, \hat{x}(\tau), \hat{u}(\tau))\psi_{n+1}(\tau), d\mu_{\text{c.p.}}\rangle \\ + \sum_{\tau \in \mathrm{Ds}(\hat{\vartheta}): \tau \leq t} \int_{[0,1]} \langle g_t^{Tr}(\tau, \hat{\chi}_\tau(s), \hat{u}_\tau(s))\sigma_\tau(s) + I_t(\tau, \hat{\chi}_\tau(s), \hat{u}_\tau(s))\sigma_\tau^{n+1}(s), \hat{v}_\tau(s)\rangle ds, \\ a.e.\ t \in [0, \hat{T}]. \end{aligned} \tag{4.70}$$

Observe that the conditions have been listed as $(n+1)$-vectors with the $(n+1)^{\text{st}}$ component written separately, as in the proof of Theorem 4.8. The notation L_x in condition (4.65) is taken to be the vector of partials $L_x = [L_{x_1}, ..., L_{x_n}]^{Tr}$; otherwise, the partial notation of Remark 4.7 is in effect. We see that the equation for ψ_{n+1} in (4.65) has been simplified due to the fact that the functions f and g are independent of the variable x_{n+1} and therefore $f_{x_{n+1}} \equiv 0$ and $g_{x_{n+1}} \equiv 0$. This comment also verifies the value of 0 on the right hand side of the fourth equation of (4.66).

For each $\tau \in Ds(\hat{\vartheta})$, condition (4.66) implies that the component σ_τ^{n+1} of the jump adjoint arc is constant. In particular,

$$\sigma_\tau^{n+1} \equiv \psi_{n+1}(\tau^-), \quad \forall \tau \in Ds(\hat{\vartheta}). \tag{4.71}$$

Therefore, the sum in the second equation of the adjoint equation (4.65) is 0, so

$$\psi_{n+1}(t) = \psi_{n+1}(0), \quad \forall t \in (0, \hat{T}]. \tag{4.72}$$

Note that the value function for the problem formulation (4.55)-(4.57) is given by $\varphi(\mathbf{p}) = x_{n+1}(T)$, so φ depends only on the variable endpoint $x_{n+1}(T)$ and partials with respect to all other variables are 0. The transversality conditions (4.67) can thus be split via (2.8) and written as

$$\begin{aligned} (-\psi(\hat{T}), \psi(0)) \in \lambda \cdot 0 + N_{C_f \times C_{in}}(\hat{p}), \\ \psi_{n+1}(0) \in \lambda \cdot 0 + N_{\{0\}}(0) = \mathbb{R}, \\ -\psi_{n+1}(\hat{T}) \in \lambda \cdot 1 + N_{\mathbb{R}}(\hat{x}_{n+1}(\hat{T})) = \lambda + \emptyset, \end{aligned}$$

and further simplified to

$$\begin{aligned} (-\psi(\hat{T}), \psi(0)) &\in N_{C_f \times C_{in}}(\hat{p}), \\ \psi_{n+1}(0) &\in \mathbb{R}, \\ -\psi_{n+1}(\hat{T}) &= \lambda. \end{aligned} \tag{4.73}$$

The second inclusion is a vacuous condition on ψ_{n+1}, but the equation of the third line and (4.72) imply that ψ_{n+1} is the constant function

$$\psi_{n+1}(t) = \lambda, \quad \forall t \in (0, \hat{T}]. \tag{4.74}$$

In turn, the identities (4.71) and (4.74) yield

$$\sigma_\tau^{n+1}(s) = \lambda, \quad \forall s \in [0,1], \quad \forall \tau \in Ds(\hat{\vartheta}). \tag{4.75}$$

The adjoint equation (4.58), the maximum condition (4.61), and the condition (4.70), follow from applying the substitutions given by (4.74) and (4.75) to (4.65), (4.68), and (4.70), respectively.

The jump adjoint equation (4.59) follows from (4.75) and the non-$(n+1)$ components of (4.66).

The transversality condition (4.60) is given directly by the first inclusion of (4.73).

The jump maximum conditions (4.62) are derived from (4.69) via (4.74) and (4.75).

Finally, suppose $\hat{T} \notin Ds(\vartheta)$. Then Theorem 4.8 implies

$$\lambda \frac{\partial \varphi}{\partial t}(\hat{p}) = \langle f(\hat{T}, \hat{x}(\hat{T}), \hat{u}(\hat{T})), \psi(\hat{T}) \rangle,$$

which by the definitions of φ and x_{n+1}, is equivalent to

$$\lambda \int_0^{\hat{T}} L_t(t, \hat{x}(t), \hat{u}(t)) dt + \lambda \int_{[0,\hat{T}]} \langle I_t(t, \hat{x}(t), \hat{u}(t)), d\hat{\vartheta} \rangle = \langle f(\hat{T}, \hat{x}(\hat{T}), \hat{u}(\hat{T})), \psi(\hat{T}) \rangle.$$

This last equation is condition (4.64). □

The Minimum Time problem. To form the minimal time problem, take $L \equiv 1$ for the non-impulsive portion of the Lagrangian in problem (P_B). We let I be a continuously differentiable function to be chosen as a penalty function for jumps in the dynamics. The impulsive minimum time problem is then given by

$$\min \int_0^T dt + \int_{[0,T]} \langle I(t, x(t), u(t)), d\vartheta \rangle \tag{4.76}$$

subject to

$$dx(t) = f(t, x(t), u(t))dt + g(t, x(t), u(t))d\vartheta, \quad \text{for a.e. } t \in [0, T] \tag{4.77}$$

and

$$p = (x(T), x(0)) \in C_f \times C_{in}, \tag{4.78}$$

where C_f is the target set. Necessary conditions for this problem may readily be obtained by the maximum principle of Corollary 4.10.

The Calculus of Variations problem. Similarly to section 2.3.2, we consider the Calculus of Variations problem

$$\min_{x(\cdot)} \int_0^T L(t, x(t), \dot{x}(t))dt + \int_{[0,T]} \langle I(t, x(t), \dot{x}(t)), d\vartheta \rangle, \tag{4.79}$$

subject to

$$x(0) = \bar{x}, \quad x(T) = \bar{y}, \tag{4.80}$$

for $\bar{x}, \bar{y} \in \mathbb{R}^n$. The functions of the Lagrangian depend on the derivative $\dot{x}$. However,in order for the problem to allow impulses we must allow x to be a function in $BV([0,T], \mathbb{R}^n)$, which raises the problem of a definition for $\dot{x}$.

In the context of the present work, this problem rests between two possible approaches. The first would be to consider the PMP of the graph completion method given in section 4.1.3, which provides an auxiliary system with an absolutely continuous trajectory y whose derivative is well-defined almost everywhere. However, we would need to extend the PMP for the fixed end-time Mayer problem to the free end-time Mayer problem and in turn the Bolza problem. Remark 4.9 points out the challenges associated with this task. The second approach points out a weakness in the maximum principles associated with the measure-adjoint problems, namely the lack of a graph completion to form an auxiliary system whose trajectory has a well-defined derivative. Both of these approaches are future problems resulting from the present work, with the measure-adjoint approach being preferred as it will allow for vector-valued impulses.

Chapter 5
Conclusions and future work

We have made a survey of the theory of impulsive control systems and made connections between problem formulations, as in the solution equivalence of impulsive differential inclusions and impulsive differential equations. The application of the graph completion and reparameterization to a neural spiking model provides a new type of analysis for such models which was shown to be generally equivalent to a technique in use. A comparison between Pontryagin-type maximum principles for optimal impulsive control problems was made, and the more recent maximum principle for the fixed end-time Mayer problem was extended generally to a free end-time Mayer problem, a Bolza problem, and a minimum time problem.

Future work ranges from immediate details to more remote topics including:

- Establish a nontriviality condition for the maximum principle of the free end-time Mayer problem, or describe the assumptions to guarantee nontriviality.
- Complete an extension of the PMP for the impulsive Bolza problem to an impulsive Calculus of Variations problem.
- Form a precise comparison between the maximum principles of the two impulsive fixed end-time Mayer problems and extend the measure-adjoint solution concept to accommodate nonsmooth dynamics.
- Apply the maximum principle of the impulsive Bolza problem to the neural spiking model to optimize impulse trains, and therefore nerve signal strength, over the parameters of the model.
- Investigate the Hamilton-Jacobi theory and other sufficient conditions for the impulsive optimal control problems studied in the present work.

References

[1] L. Ambrosio and G. Dal Maso, *A general chain rule for distributional derivatives*, Proc. Amer. Math. Soc., **108**, pp. 691-702, 1990.

[2] A.V. Arutyunov, D.Yu. Karamzin, and F.L. Pereira, *On constrained impulsive control problems*, Journal of Mathematical Sciences, **165**, No. 6, pp.654-688, 2010.

[3] J.P. Aubin and A. Cellina, *Differential Inclusions*, Springer, New York, 1984.

[4] A. Bressan and B. Piccoli, *Introduction to the Mathematical Theory of Control*, American Institute of Mathematical Sciences, USA, 2007.

[5] A. Bressan and F. Rampazzo, *Impulsive systems with commutative vector fields*, J. of Optimization Theory and Applications **71**, pp. 67-83, 1991.

[6] A. Bressan and F. Rampazzo, *Impulsive systems without commutativity assumptions*, J. of Optimization Theory and Applications **81**, pp. 435-457, 1994.

[7] P. Cannarsa and C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*, Birkhauser, Boston, 2004.

[8] A. Catlla, D. Schaeffer, T. Witelski, E. Monson, A. Lin, *On Spiking Models for Synaptic Activity and Impulsive Differential Equations*, SIAM Review, **50**, No. 3, pp. 553-569, 2008.

[9] V.V. Chistyakov, *Selections of Bounded Variation*, Journal of Applied Analysis, Vol. 10, No. 1, pp. 1-82, 2004.

[10] F.H. Clarke, *Optimization and Nonsmooth Analysis*, Canadian Mathematical Society, Wiley-Interscience, Toronto, 1983.

[11] F.H. Clarke, *Functional Analysis, Calculus of Variations, and Control Theory*, Springer, London, 2013.

[12] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Springer, New York, 1998.

[13] G. Dal Maso and F. Rampazzo, *On systems of ordinary differential equations with measures as controls*, Differential Integral Equations, v. 4, pp. 739-765, 1991.

[14] G. B. Folland, *Real Analysis, Modern Techniques and their Applications*, John Wiley, New York, 1984.

[15] A. N. Kolmogorov and S. V. Fomin, *Introductory Real Analysis*, Dover Publications, New York, 1970.

[16] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation I*, Springer-Verlag, Berlin, 2006.

[17] R. W. Rishel, *An extended Pontryagin principle for control systems whose control laws contain measures*, SIAM J. of Control, **3**, (1965) 191-205.

[18] G. N. Silva and R. B. Vinter, *Measure differential inclusions*, J. Math. Anal. Appl., 202, pp.727-746, 1996.

[19] G. N. Silva and R. B. Vinter, *Necessary conditions for optimal imulsive control systems*, SIAM J. of Control and Optimization, **35**, pp. 1829-1846, 1998.

[20] E. Sontag *Mathematical Control Theory: Deterministic Finite-dimensional Systems*, Second edition. Springer-Verlag, New York, 1998.

[21] A. I. Volpert, *The spaces BV and quasi-linear equations*, Math. USSR Sb 2, pp. 225-267, 1967.

[22] W. Walter *Ordinary Differential Equations*, Springer, New York, 1998.

[23] P. R. Wolenski and S. Zabic, *A differential solution concept for impulsive systems*, Differential Equations and Dynamical Systems, DCDIS Proceedings **2**, pp. 199-210, 1996.

[24] P. R. Wolenski and S. Zabic, *A sampling method and approximation results for for impulsive systems*, SIAM Journal on Control and Optimization, Volume 46, Issue 3, June 2007.

Vita

Jacob Blanton was born in October 1980 in Tampa, Florida. He completed his undergraduate studies in mathematics at the University of Central Florida in Orlando, Florida in 2005. He began his graduate studies in August 2005 at Louisiana State University, and in Fall 2006 he earned a Master of Science degree in mathematics. He is currently a candidate for the Doctor of Philosophy degree in mathematics to be awarded in Spring 2014.