Second-Order Necessary Conditions of Optimality for Measure Driven Control Systems*

Aram Arutyunov[1], Vladimir Dykhta[2], Fernando Pereira[3]

[1]arutun@orc.ru, Differential Equations and Functional Analysis Department
Peoples Friendship Univ. of Russia, Moscow, Russia

[2]dykhta@isea.ru, Department of Mathematics,
Baikal State University of Economy and Law, Irkutsk, Russia

[3]flp@fe.up.pt, Institute for Systems and Robotics,
Faculdade de Engenharia da Universidade do Porto, Porto, Portugal

Abstract. In this communication, we discuss necessary conditions of optimality for impulsive control problems. That is, problems whose control space includes measures besides the conventional class of measurable functions. More precisely, we present second-order necessary conditions of optimality for control problems with equality and inequality endpoint state constraints and control constraints. These enable the selection of informative multipliers provided by the local maximum principle even when the optimal control process is abnormal.

Key Words. Impulsive Control, Nondegeneracy, Necessary Conditions of Optimality.

1 Introduction

In this article, we present non degenerate first-order and second-order necessary conditions of optimality for measure driven dynamic control systems with control constraints, and equality and inequality endpoint state constraints. This result is essentially contained in [4], and is part of the effort of migrating the results in [3, 1] for the impulsive control context. In this respect, it can be regarded as an extension of the conditions obtained in [2, 4] since besides the dependence of the singular dynamics on the state variable, we also have control constraints. One key feature of these necessary conditions of optimality is the fact that they do not degenerate, i.e., they remain informative even for abnormal control processes, in spite of being derived in the absence of a priori normality assumptions.

In order to clarify the nature of these conditions, let us first introduce the basic issue of nondegeneracy in the context of a simple static nonlinear optimization problem. Given $x \in R^n$, and the C^2 functions $f : R^n \to R$ and $g : R^n \to R^m$, consider the problem

$$\text{Minimize } f(x) \text{ s.t. } g(x) = 0.$$

*The first and second authors were, respectively, partially supported by Russian Foundation for Basic Research grants N 02-01-00334 and N 02-01-00869, and the third author was partially supported by the Fundacao para a Ciencia e Tecnologia via the Praxis XXI CorDyAL project. All the authors acknowledge the support of INVOTAN mobility grants.

Let x^* be a solution and denote by L, the Lagrangian, defined by $\lambda_0 f(x) + \lambda' g(x)$ for some $\lambda_0 \in R$ and $\lambda \in R^m$.
Clearly, if the operator $g_x(x^*)$ fails to be onto, then, besides the nonuniquess of the multipliers associated with a given reference control process, it is possible to find a nontrivial noninformative multiplier, i.e., $(\lambda_0, \lambda) \neq 0$ with $\lambda_0 = 0$, such that $L_x(x^*, \lambda_0, \lambda) = 0$.
However, if we consider a subset of multipliers for which the dimension of the subspace of maximum dimension in Ker $g_x(x^*)$ where the quadratic form $v' L_{xx}(x^*, \lambda_0, \lambda) v$ is negative definite is not greater than codim Im $g_x(x^*)$, then, for some $(\lambda_0, \lambda) \in \Lambda_a$, we may guarantee that $v' L_{xx}(x^*, \lambda_0, \lambda) v \geq 0$ on Ker $g_x(x^*)$.

In the next section, we present the optimal impulsive control problem to which this idea will be used to derive nondegenerate second order conditions and discuss the assumptions on its data. Refer to [4] for a detailed proof for the case without constraints on the conventional control variable. Then, in the third section, the optimality conditions are stated and, finally, a small example is presented.

2 Problem Statement

In order to state the optimal control problem, let t_0, and t_1, with $t_0 < t_1$, be given and consider the functions $f : [t_0, t_1] \times R^n \times R^m \to R^n$, $G : [t_0, t_1] \times R^n \to R^{n \times q}$, $f_0 : R^n \times R^n \to R$, $g : R^n \times R^n \to R^{d(g)}$, $h : R^n \times R^n \to R^{d(h)}$, and $M : R \times R^m \to R^{d(M)}$. Here, $d(f)$ the dimension of the range space of the function f. Then, our problem is:

$$
\begin{aligned}
(P) \text{ Minimize} \quad & f_0(x(t_0), x(t_1)) \\
\text{subject to} \quad & \\
& dx(t) = f(t, x(t), u(t))dt + G(t, x(t))d\mu(t), \; t \in [t_0, t_1], \\
& g(x(t_0), x(t_1)) \leq 0, \qquad h(x(t_0), x(t_1)) = 0, \\
& u \in \mathcal{U} := \{u \in L^m_\infty[t_0, t_1] \; : M(t, u) = 0 \; \mathcal{L} \text{ a.e.}\}, \\
& d\mu \in C^*([t_0, t_1]; K), \; K \subset R^q.
\end{aligned}
$$

The definition of trajectory, x associated with a control pair (x_0, u, μ) has been presented in [4]. It is a function of bounded variation satisfying, $\forall t \in (t_0, t_1]$:

$$
x(t) \;\; = \;\; x_0 + \int_{t_0}^{t} [f(s, x(s), u(s)) + G(s, x(s)) w_{ac}(s)] ds + \int_{t_0}^{t} G(s, x(s)) d\mu_{sc}(s) + \sum_{s_i \leq t} \Delta x(s_i)
$$

where $\Delta x(s_i) = \xi_i(1) - \xi_i(0)$, $\xi_i(0) = x(s_i^-)$, $\dot{\xi}_i(s) = G(s_i, \xi_i(s))\mu_{sa}(\{s_i\})$, $[0,1]$-a.e., and w_{ac}, $d\mu_{sc}$, and $d\mu_{sa} := \sum_{s_i < t} \mu(\{s_i\})\delta_{s_i}(t)$, are, respectively, the absolutely continuous, the singular continuous, and the atomic components of the control measure.

We say that the admissible process (x_0^*, u^*, μ^*) is a local minimizer of the problem (P) if there exists $\varepsilon > 0$ and, for any finite-dimensional subspace $\mathbf{R} \subset L^m_\infty[t_0, t_1]$, $\varepsilon_{\mathbf{R}} > 0$ such that the process (x_0^*, u^*, μ^*) yields a minimum to the problem (P) with the additional constraints $\|a - a^*\| < \varepsilon$, $\|d\mu - d\mu^*\|_{C^*([t_0,t_1];R^q)} < \varepsilon$, $u \in \mathcal{U}_\varepsilon(u^*)$ defined as the set of controls u satisfying $\|u - u^*\|_{L^m_\infty[t_0,t_1]} < \varepsilon_{\mathbf{R}}$, $u \in \mathbf{R}$, and $u(t) \in U(t)$.
As can be seen in [4], our result requires the following assumptions on the data:

- The cone $K \subset R^q$ is convex, closed, and pointed.
- Smoothness - f_0, g, h, and G are C^2, and f is C^2 w.r.t. (x, u) for all $t \in [t_0, t_1]$.

- Functions f and G and their first and second order derivatives are bounded on any bounded subset and measurable w.r.t. t.
- The matrix-valued function G satisfies the Frobenius condition.
- The function M is regular, i.e., $\forall V \subset R^m$, $\exists \epsilon > 0$ such that, for a.a. $t \in [t_0, t_1]$, $\det(M_u(t,u)M_u^T(t,u)) \geq \epsilon \in V$ s.t. $|M(t,u)| \leq \epsilon$.

3 Local Maximum Principle

Let $\psi \in R^n$, $\lambda = (\lambda_0, \lambda_g, \lambda_h) \in R^1 \times R^{d(g)} \times R^{d(h)}$ and denote by x_0 and x_1 the initial and final endpoints of the state trajectory. The Pontryagin function H, and the endpoint Lagrangian l^λ are, respectively, given by:

$$\begin{aligned} H &= H_f + H_G \text{ where } H_f(t,x,\psi,u) = \langle \psi, f(t,x,u) \rangle, \; H_G(t,x,\psi,v) = \langle \psi, G(t,x)v \rangle, \\ l^\lambda(a) &= \lambda_0 f_0(a) + \langle \lambda_g, g(a) \rangle + \langle \lambda_h, h(x_0,x_1) \rangle, \; a = (x_0, x_1). \end{aligned}$$

A control process (x_0^*, u^*, μ^*) satisfies the local maximum principle if there exists $\lambda \neq 0$, such that $\lambda_0 \geq 0$, $\lambda_g \geq 0$, $\langle \lambda_g, g(a^*) \rangle = 0$, a vector function $\psi \in BV^n[t_0, t_1]$, solution to the adjoint system

$$\begin{cases} -d\psi(t) = H_{fx}(t)dt + H_{x_v}(t)d\mu^*(t), \\ -\psi(t_1) = l_{x_1}^\lambda(a^*), \end{cases} \tag{1}$$

and a vector function $m \in L_\infty^{d(M)}[t_0, t_1]$ which satisfy the following conditions:

$$\begin{aligned} \psi(t_0) &= l_{x_0}^\lambda(a^*), & (2) \\ H_u(t) - M_u(t)m(t) &= 0 \quad \mathcal{L}\text{-a.e.}, & (3) \\ \langle H_v(t), v \rangle &\leq 0 \quad \forall (t,v) \in [t_0,t_1] \times K, & (4) \\ \langle H_v(t), \omega^*(t) \rangle &= 0 \quad d\mu_c^*\text{-a.e.}, & (5) \\ \langle H_v(s_i, z^{s_i}(\tau), q^{s_i}(\tau)), \mu_a(\{s_i\}) \rangle &= 0 \quad \forall \tau \in [0,1] \quad \forall s_i \in S^* & (6) \end{aligned}$$

where S^* is the support of the atomic component of $d\mu^*$ and $\omega^*(t) = \frac{d\mu_c^*(t)}{d|\mu_c^*(t)|}$ is the Radon Nicodym derivative of its singular continuous component with respect to its total variation measure. The jump in the adjoint variable, solution to equation (1), at time s_i is given by $q^{s_i}(1) - \psi(s_i^-)$ with

$$\frac{dq^{s_i}(\tau)}{d\tau} = H_{x_v}(s_i, z^{s_i}(\tau), q^{s_i}(\tau))\mu_a(\{s_i\}), \quad q^{s_i}(0) = \psi(s_i^-) \tag{7}$$

being, as before $z^{s_i}(\tau)$ the corresponding solution to the singular dynamics with $x(s_i^-) = x^*(s_i^-)$.

We will denote by Λ the set of all normalized Lagrange multipliers λ satisfying the local maximum principle (i.e., $\|\lambda\| = 1$). If the reference control process is not normal then it may happen that these conditions hold for multipliers with $\lambda_0 = 0$. In the next section, we present a second-order condition which ensures that that will not be the case if the reference control process is a local minimum.

Given the complexity of the various formulae to be presented in the next section, we adopt the following short notation. When some arguments of a given function are missing, i.e., $H(t)$, $H(t,u)$, or $H_v(t)$, this means that the function is considered being evaluated along the considered reference process. The dot over the function label means the total derivative with

respect to time. An argument variable appearing in sub index means that a partial derivative is being considered, e.g., $H_{x_v}(t) = \frac{\partial^2 H}{\partial v \partial x}(t)$.
Furthermore, we will use the total derivative w.r.t. time along the solution to the following ordinary differential system

$$\left\{ \begin{array}{rcl} \dot{x} & = & F(t,x,u,v) \\ -\dot{\psi} & = & H_x(t,x,\psi,u,v) \\ \dot{w} & = & v, \qquad v \in K \end{array} \right. \tag{8}$$

where $F(t,x,u,v) = f(t,x,u)+G(t,x)v$, and w (w^*) is the right continuous function of bounded variation associated with μ (μ^*)[1]. We also let $\bar{g} = \text{col}(f_0, g)$, $\delta a = (\delta x(t_0), \delta x(t_1))$, and $\delta w_1 = \delta w(t_1)$.

4 Second-order Necessary Conditions of Optimality

The second order conditions are cast in terms of the positive semidefiniteness of a certain quadratic form for each variation in a certain cone, the cone of critical variations. Furthermore, in order to simplify the statement of the result, we consider only control measures without the singular continuous component.

$\mathcal{K}_{cr}$ is the cone of all critical variations, i.e., the triples $(\delta x_0, \delta u, \delta w) \in R^n \times L^m_\infty[t_0,t_1] \times BV^q(S^*)$[2] whose state trajectory variation, $\delta x \in BV^n(S^*)$, satisfies the following conditions:

$$\begin{array}{l} \langle \bar{g}_a(a^*), \delta a\rangle + \langle \bar{g}_{x_1}(a^*), G(t_1)\delta w_1\rangle \le 0, \\ \langle h_a(a^*), \delta a\rangle + \langle h_{x_1}(a^*), G(t_1)\delta w_1\rangle = 0, \\ \dot{\delta x} = F_x(t)\delta x + F_u(t)\delta u - (\dot{H}_v)^T_\psi(t)\delta w, \quad t \notin S^* \\ \delta u(t) M_u(t) = 0 \quad \mathcal{L} - \text{a.e.} \\ d(\delta w) \in \mathcal{K} + \text{Lin}\{d\mu^*\}, \ \delta w(t_0) = 0 \end{array}$$

where $\forall s_i \in S^*$, $\delta x(s_i) = \delta q(1; s_i, \mu^*(\{s_i\}))$ being $\delta q(\tau; s_i, \mu^*(\{s_i\})) := \delta q^{s_i}(\tau)$ the solution to the system

$$\frac{d(\delta q^{s_i}(\tau))}{d\tau} = H_{G\psi x}(s_i, z^{s_i}(\tau), \mu^*(\{s_i\}))\delta q^{s_i}(\tau), \ \delta q^{t_0}(0) = \delta x_0, \ \delta q^{s_i}(0) = \delta x(s_i^-), \quad s_i > t_0$$

and the function $z^{s_i}(\tau)$ is the solution to the system defining the singular dynamics.

For any $\lambda \in \Lambda$ define the quadratic form $\Omega^\lambda(\delta x_0, \delta u, \delta w)$ by

$$\delta a^T l^\lambda_{aa}(a^*)\delta a + Q^\lambda_1(\delta a, \delta w_1) - \int_{t_0}^{t_1} Q^\lambda(\delta x, \delta u, \delta w)(t)dt + \int_{t_0}^{t_1} \delta u^T(t)\langle M(t), m(t)\rangle_{uu}\delta u(t)dt$$

$$\begin{array}{rcl} \text{where} \quad Q^\lambda(\delta x, \delta u, \delta w) & = & \delta u^T H^\lambda_{uu}\delta u + 2\delta x^T H^\lambda_{xu}\delta u - 2\delta w^T(\dot{H}^\lambda_v)_u \delta u - \delta w^T(\ddot{H}^\lambda_v)_v \delta w \\ & & -2\delta w^T(\dot{H}^\lambda_v)_x \delta x + \delta x^T H^\lambda_{xx}\delta x \\ \text{and} \quad Q^\lambda_1(\delta x(\cdot), \delta w_1) & = & 2\delta x(t_0)^T l^\lambda_{x_0 x_1}(a^*)G(t_1)\delta w_1 - 2\delta x(t_1)^T H^\lambda_{xv}(t_1)\delta w_1 \\ & & +\delta w_1^T G^T(t_1)(l^\lambda_{x_1 x_1}(a^*)G(t_1) - H^\lambda_{xv}(t_1))\delta w_1 \\ & & -\sum_{s \in S^*}[\delta x^T(s)\Psi^\lambda(s)\delta x(s) - \delta x^T(s^-)\Psi^\lambda(s^-)\delta x(s^-)]. \end{array}$$

[1]Remark that, under the Frobenius condition, this derivative does not depend on v
[2]Denote by $BV^n(S^*)$ the set of n-dimensional vector functions of bounded variation whose singular variations are supported on the zero Lebesgue measure set S^*.

Here, $\Psi^\lambda(t) = -Z^T(1;t)\int_0^1 Z^{-1T}(\tau;t)H_{Gxx}(\tau;t)Z^{-1}(\tau;t)d\tau Z(1;t) \in BV^{n\times n}(S^*)$ where $H_{Gxx}(\tau;t) = H_{Gxx}(t, z^t(\tau), q^t(\tau), w^*(t))$ and the $n \times n$ matrix $Z(\tau;t)$ satisfies the linear differential equation

$$-\frac{dZ}{d\tau} = ZH_{G\psi x}(t, z^t(\tau), w^*(t)), \qquad Z(0;t) = I,$$

with $z^t(\tau)$, and $q^t(\tau)$ being the solutions to the respective singular limiting systems.

Now, let π is the matrix of the orthogonal projection from R^k onto the linear subspace $N = K \cap (-K)$ and consider the following modified variational equation for $t \notin S^*$:

$$\dot{\delta x} = F_x(t)\delta x + F_u(t)\delta u - (\dot{H}_v)_\psi(t)\pi\delta w, \tag{9}$$

with $\delta x(0) = \delta x_0 \in R^n$, $\delta w \in L^k_\infty$, $\delta u \in L^m_\infty$, together with the considered jump conditions and linearized constraints.

Let $L = (g, h)$ and denote by $\mathcal{K}_\pi$ the linear subspace of $R^n \times L^m_\infty \times L^k_\infty \times R^k$ of all tuples $(\delta x_0, \delta u, \delta w, c) \in R^n \times L^m_\infty \times L^k_\infty \times R^k$ satisfying the corresponding linearized constraints and $\mathcal{A}(\delta x(0), \delta u, \delta w, c) = 0$ where the linear operator $\mathcal{A}$ is defined by

$$\mathcal{A}(\delta x(0), \delta u, \delta w, c) = L_{x_0}(a^*)\delta x_0 + L_{x_1}(a^*)\delta x_1 + L_{x_1}(a^*)G(t_1)\pi c. \tag{10}$$

Here, δx is the corresponding solution to (9), linearized constraints, and jump dynamics.

Let $d = \text{codim}\,(Im\mathcal{A})$, and define the quadratic form Ω^λ_a on $R^n \times L^m_\infty \times L^k_\infty \times R^k$ obtained from Ω^λ by formally replacing δw_1 by c. Let $\Lambda_a(x^*, u^*, w^*)$ be the subset of $\Lambda(x^*, u^*, w^*)$ for which the index of the form Ω^λ_a on the subspace $\mathcal{K}_\pi$ is not greater then d.

Theorem. (Necessary conditions of optimality.)
If the control process (x^*, u^*, w^*) is a local optimum to problem (P). Then, $\Lambda_a \neq \emptyset$ and, for any $(\delta x_0, \delta u, \delta w) \in \mathcal{K}_{cr}$, we have

$$\max_{\lambda\in\Lambda_a} \Omega^\lambda(\delta x_0, \delta u, \delta w) \geq 0.$$

Most of the proof of this result appears in [4]. It consists in applying a certain a nonlinear transformation, [5], to the initial problem, so that the new one is such that the impulsive dynamics do not depend on x, in applying a variety of the first and second order necessary conditions of optimality derived in [1], and, then, in expressing these in terms of the data of the original problem.

5 Example

Let us consider the following example from [4]:

$$\begin{aligned}
\text{Minimize} \quad & \langle \zeta, (x_1(1), \ldots, x_k(1)) \rangle \\
\text{s.t.} \quad & dx_i = f_i(x,t)dt + dw_i, \quad i = \overline{1,k}, \quad w = \text{col}(w_1, \ldots, w_k) \\
& dx_n = f_n(x,t)dt + \langle Q\text{col}(x_1, \ldots, x_k), dw \rangle, \\
& t \in [0,1], \quad x(0) = 0, \quad x_n(1) = 0, \quad K = R^k
\end{aligned}$$

where $n \geq 5$, $k = n-1$, $x = \text{col}(x_1, \ldots, x_n) \in R^n$, Q is a symmetric $k \times k$ matrix such that the index of each of the matrices Q and $(-Q)$ is not less than 2, $\zeta \in R^k$ is a given nonzero vector, and, for $i = 1, \ldots, n$, the f_i's are arbitrary given smooth functions such that $f_i(0,t) \equiv 0$, $f_{ix}(0,t) \equiv 0$, and $f_{nxx}(0,t) \equiv 0$. Because of the symmetry of Q, it can be easily shown that the

Frobenius condition holds. We consider the reference admissible control process $(0,0,0)$ and show that it is not locally optimal.

Fix any $\lambda \in \Lambda$. From (4), we obtain, for $\psi(\cdot) = \psi^\lambda(\cdot) = (\psi_1(\cdot), \dots \psi_n(\cdot))$, $\psi_i(t) \equiv 0$, $i = 1, \dots, k$, and from (1), we have $\psi_n(t) \equiv \psi_{n,0} =$ const. Hence, by using (1), (2), and $\zeta \neq 0$, we obtain $\Lambda = \{\lambda : \lambda_0 = 0, \lambda_{2,i} = 0, i = \overline{1, n-1}, \lambda_{2,n} = -\lambda_{2,n+1}\}$ and, consequently, Λ consists of only two vectors $\bar{\bar{\lambda}} = -\bar{\lambda}$ and $\bar{\lambda} = \frac{1}{\sqrt{2}}(0, \dots, 0, 1, -1)$ and $\psi_{n,0} = \pm\frac{1}{\sqrt{2}}$.

It can be easily shown that $d = 1$ and $\Omega_a^\lambda(\delta w) = \psi_{n,0} \int_0^1 \langle Q\text{col}(\delta x_1, \dots, \delta x_k), \delta w \rangle dt$. Hence, $\Omega_a^\lambda(\delta w) = \frac{1}{2}\psi_{n,0}\langle Q\text{col}(\delta x_1(1), \dots, \delta x_k(1)), \text{col}(\delta x_1(1), \dots, \delta x_k(1))\rangle$. This implies that, for any $\psi_{n,0} = \pm 1$ the index of the function Ω_a^λ is not less than 2. So $\Lambda_a = \emptyset$ and consequently the process $(0,0,0)$ is not optimal. Also notice that this process is abnormal and $\max_{\lambda \in \Lambda} \Omega^\lambda(\delta w) \geq 0$, $\forall\, \delta w$ (because of $\bar{\lambda}$, $\bar{\bar{\lambda}} \in \Lambda$) and the last inequality is not useful.

References

[1] A. Arutyunov. *Optimality Conditions: Abnormal and Degenerate Problems.* Kluwer Academic Publishers, 2000.

[2] A. Arutyunov, V. Jacimovic, and F. Lobo Pereira. Second order necessary conditions for optimal impulsive control problems. In *Proceedings European Control Conference 2001*, pages 2759–2763, Porto, Portugal, 2001.

[3] A.V. Arutyunov. Second order necessary conditions in optimal control problems. *Doklady Math.*, 61(2):158–161, 2000.

[4] F.L. Pereira A.V. Arutyunov, V. Dykhta. Necessary conditions for impulsive nonlinear optimal control problems without apriori normality assumptions. *Journal Optimization Theory and Applications*, to appear in 2004.

[5] V.A. Dykhta and O.N. Samsonyuk. *Optimal impulse control with applications.* Nauka, Moscow, 2000.

[6] B.M. Miller. The optimality conditions in a problem of control of a system that can be described with a differential equation with measure. *Autom. Telemekh.*, 6:752–761, 1982.

[7] A. V. Sarychev. Optimization of generalized controls in a nonlinear time optimal problem. *Differential Equations*, 27(5):539–550, 1991.

[8] S. Zavalischin and A. Sesekin. *Impulsive processes: models and applications.* Nauka, Moscow, 1991.