Switching Time Optimization for Bang–Bang and Singular Controls: Variational Derivatives and Applications

G. Vossen [1]

Abstract. In this report, we will present an appendix for the paper (Ref. 14) by the author. The results of this paper could not be included into (Ref. 14) due to page restrictions. Section numbering is taken over from (Ref. 14).

Some remarks on switching time optimization for controls without feedback representation will be presented in Subsection 3.4. Proofs for variational derivatives which are used in switching time optimization for bang–bang and singular controls are given in Section 4. A detailed elaboration of switching optimization for the Goddard Problem will be presented in Subsection 5.2.

3.4 Controls Without Feedback Representation

If instead of Assumption 3.1 the control can only be determined as a function $u^i(t,x,\lambda)$ depending also on the adjoint variable λ, a modified approach can be used to accomplish switching time optimization. We will formulate an augmented induced optimization problem involving the optimization vector

$$z := (x_0^T, \lambda_0^T, t_1, \ldots, t_d, t_{d+1})^T \in \mathbb{R}^{2n+d+1}$$

where λ_0 denotes the initial value $\lambda(0)$ of the adjoint variable. We denote by $x(\cdot,z)$ and $\lambda(\cdot,z)$ the absolutely continuous solution of the coupled initial value problem

$$\begin{aligned} x(0) &= x_0, & \dot{x}(t) &= f(t, x(t), u(t, x(t), \lambda(t))), \\ \lambda(0) &= \lambda_0, & \dot{\lambda}(t) &= H_x(t, x(t), u(t, x(t), \lambda(t)), \lambda(t)) \end{aligned}$$

where $u(t,x,\lambda)$ is piecewisely defined by the functions $u^i(t,x,\lambda)$ in each interval J_i, $i = 1,\ldots,d$. Here, the second differential equation is motivated by the adjoint differential equation (7), (Ref. 14) which holds along the optimal trajectory. Then, the augmented induced problem additionally includes the transversality conditions (8) and (9), (Ref. 14) as constraints for $\lambda(\cdot,z)$ at $t=0$ and $t=t_f$:

$$\begin{aligned} \min \quad & G(z) := g(x_0, x(t_f,z), t_f) \\ \text{s. t.} \quad & \Phi(z) := \begin{pmatrix} \phi(x_0, x(t_f,z), t_f) \\ \lambda_0 + l_{x_0}(x_0, x(t_f,z), t_f, \rho_0, \rho) \\ \lambda(t_f,z) - l_{x_f}(x_0, x(t_f,z), t_f, \rho_0, \rho) \end{pmatrix} = 0. \end{aligned} \tag{1}$$

[1] Lehr- und Forschungsanstalt für Nichtlineare Dynamik
RWTH Aachen
Steinbachstraße 15
52074 Aachen
Germany

As some initial values $\lambda_i(0)$ may be given by the transversality condition (8), (Ref. 14), these values should, analogously to Remark 3.3, (Ref. 14), be eliminated from the induced problem. This approach was used by Vossen/Rehbock/Siburian (Ref. 32).

If the control cannot be determined as a function $u^i(t,x,\lambda)$ since the control does not appear in any derivative of the switching function (i.e., the singular control is of order $q=\infty$), the problem cannot be reduced to a finite dimensionional one by this approach. As an extension to (21), (Ref. 14), one idea is to optimize the discretized control along all singular arcs where the function $u^i(t,x,\lambda)$ is unknown. This approach was described by Büskens et al. (Ref. 33).

4 Variational Derivatives in the Induced Problem

Obviously, the verification of optimality conditions (24), respectively, (25), (Ref. 14) and hence, the calculation of variational derivatives of the Lagrangian function with respect to the optimization vector z requires the calculation of variational derivatives of the state $x(\cdot,z)$ with respect to z.

4.1 Derivatives of the Function h

In the following, we will need the derivatives of the function h in (20), (Ref. 14) with respect to the states. The first derivative is given by

$$h_x(t,x) = f_x(t,x,u(t,x)) + f_u(t,x)u_x(t,x) \tag{2}$$

and, with arbitrary vectors $\mu,\nu \in \mathbb{R}^n$, the second derivative can be written as

$$\begin{aligned}\mu^T h_{xx}(t,x)\nu = \mu^T f_{xx}(t,x,u(t,x))\nu + \mu^T f_{xu}(t,x)(u_x(t,x)\nu)\\ + (u_x(t,x)\mu)^T f_{ux}(t,x)\nu + f_u(t,x)(\mu^T u_{xx}(t,x)\nu).\end{aligned} \tag{3}$$

Here, the derivative $\mu^T h_{xx}\nu$ is a column vector with the components $\mu^T(h_k)_{xx}\nu$, $k=1,\ldots,n$. The derivatives f_{xx}, f_{xu}, f_{ux} und u_{xx} are understood in an appropiate way. Note that the term including f_{uu} was deleted in (3) as u appears only linearly in f.

For the calculations of the Lagrangian derivatives, the following result will be useful.

Proposition 4.1. *Along a trajectory $\hat{\mathcal{T}} = (\hat{x},\hat{u})$ which satisfies the necessary optimality conditions (7)–(12), (Ref. 14) of the minimum principle, the following holds for all $t\in[0,\hat{t}_f]$:*

$$\lambda(t)h_x(t) = H_x(t), \tag{4}$$

$$\lambda(t)\big(\mu^T h_{xx}(t)\nu\big) = \mu^T\big(H_{xx}(t) + H_{xu}(t)u_x(t) + (u_x(t))^T H_{ux}(t)\big)\nu \tag{5}$$

for arbitrary column vectors μ, $\nu\in\mathbb{R}^n$.

Proof. Due to (2), we obtain

$$\lambda h_x = \lambda(f_x + f_u u_x) = (\lambda f)_x + (\lambda f)_u u_x = H_x + H_u u_x = H_x$$

for all $t \in [0, \hat{t}_f]$. The last equality arises from the following fact. Consider a component $\hat{u}_k$, $1 \leq k \leq m$, for some $t \in [0, \hat{t}_f[$. For sufficiently small $\epsilon > 0$, $\hat{u}_k$ is either singular or bang–bang on $[t, t+\epsilon]$. Hence, $H_{u_k}(t) = \sigma_k(t) \equiv 0$ or $(u_k)_x(t, x(t)) \equiv 0$ holds along the interval $[t, t+\epsilon]$. As this is true for all $k = 1, \ldots, m$, we obtain $H_u(t) u_x(t) = 0$ for all $t \in [0, \hat{t}_f[$. Considering the interval $[t-\epsilon, t]$ for $t = \hat{t}_f$, the representation (4) is proved. To verify equation (5), we obtain in view of (3)

$$\begin{aligned}
\lambda\Big(\mu^T h_{xx}\nu\Big) &= \sum_{j=1}^{n} \Big(\lambda_j \mu^T (h_j)_{xx}\nu\Big) \\
&= \mu^T H_{xx}\nu + \mu^T H_{xu}(u_x\nu) + (u_x\mu)^T H_{ux}\nu + H_u \mu^T u_{xx}\nu \\
&= \mu^T (H_{xx} + H_{xu}u_x + u_x^T H_{ux})\nu
\end{aligned}$$

for all $t \in [0, \hat{t}_f]$. As in the proof for (4), the last equality holds due to

$$H_u \mu^T u_{xx}\nu = \sum_{k=1}^{m} \Big(\mu^T H_{u_k} (u_k)_{xx}\nu\Big) = 0$$

as we have $H_{u_k}(t) \equiv 0$ if $\hat{u}_k(t)$ is singular and $(u_k)_{xx}(t, x(t)) \equiv 0$ if $\hat{u}_k(t)$ is bang–bang. □

4.2 Variational Derivatives of the States

In a first step, formulas for variational derivatives of the states are given. These representations will be used to calculate variational derivatives of the Lagrangian function which are essential for the verification of first and second–order optimality conditions in the induced problem (20), (Ref. 14).

4.2.1 First–Order Variational Derivatives

The functions

$$v^i(t,z) := \frac{\partial x}{\partial (x_0)_i}(t,z) \qquad i = 1, \ldots, n, \tag{6}$$

$$y^i(t,z) := \frac{\partial x}{\partial t_i}(t,z), \qquad i = 1, \ldots, d, \tag{7}$$

$$y^f(t,z) := \frac{\partial x}{\partial t_f}(t,z), \tag{8}$$

are called first–order variational derivatives of the states. We shall use the abbreviations $\hat{v}^i(t) := v^i(t, \hat{z})$, $\hat{y}^i(t) := y^i(t, \hat{z})$ and $\hat{y}^f(t) := y^f(t, \hat{z})$. The following result is well-known from the theory of ODEs.

Proposition 4.2. *The function* $v^i(t,z)$, $1 \le i \le n$, *is the solution of the IVP*

$$v^i(0,z) = e_i, \qquad \dot{v}^i(t,z) = h_x(t, x(t,z))v^i(t,z), \tag{9}$$

where e_i *is the* i*-th unit vector.*

Proposition 4.3. *The function* $y^i(t,z)$, $1 \le i \le d$, *satisfies* $y^i(t,z) \equiv 0$ *on* $[0, t_i[$ *and for* $t \ge t_i$ *it is the solution of the IVP*

$$y^i(t_i, z) = -[\dot{x}]^i = -[h]^i, \qquad \dot{y}^i(t,z) = h_x(t, x(t,z))y^i(t,z), \quad t \ge t_i. \tag{10}$$

Proof. Variation of a switching time t_i will change the solution of IVP (19), (Ref. 14) only in the interval $[t_i, t_f]$. Hence, we have $y^i(t,z) \equiv 0$ on $[0, t_i[$. Furthermore, the solution of IVP (19), (Ref. 14) can be written as

$$x(t,z) = x(t_i^-, z) + \int_{t_i^+}^{t} h(s, x(s,z))ds, \qquad t \ge t_i.$$

Differentiating this equation with respect to t_i, we obtain

$$y^i(t,z) = \dot{x}(t_i^-, z) - \dot{x}(t_i^+, z) + \int_{t_i^+}^{t} h_x(s, x(s,z))y^i(s,z)\, ds \tag{11}$$

which yields (10). □

The following obvious result is given for the purpose of completeness.

Proposition 4.4. *The function* y^f *satisfies* $y^f(t,z) \equiv 0$ *on* $[0, t_f[$ *and for* $t = t_f$

$$y^f(t_f, z) = \dot{x}(t_f, z) = h(t_f, x(t_f, z)). \tag{12}$$

4.2.2 Second–Order Variational Derivatives

For the computation of second–order variational derivatives of the states, we will calculate only the entries $\partial^2 x/(\partial z_j \partial z_i)$ for $i \le j$ as the matrices $\partial^2(x_l)/\partial z^2$ are symmetric for $l = 1, \ldots, n$. We use the following notations for the second–order derivatives:

$$v^{ij}(t,z) := \frac{\partial^2 x}{\partial (x_0)_j \partial (x_0)_i}(t,z) = \frac{\partial v^i}{\partial (x_0)_j}(t,z), \qquad 1 \le i \le j \le d, \tag{13}$$

$$w^{ij}(t,z) := \frac{\partial^2 x}{\partial t_j \partial (x_0)_i}(t,z) = \frac{\partial v^i}{\partial t_j}(t,z), \qquad 1 \le i \le n, \quad 1 \le j \le d, \tag{14}$$

$$v^{if}(t,z) := \frac{\partial^2 x}{\partial t_f \partial (x_0)_i}(t,z) = \frac{\partial v^i}{\partial t_f}(t,z), \qquad 1 \le i \le n, \tag{15}$$

$$y^{ij}(t,z) := \frac{\partial^2 x}{\partial t_j \partial t_i}(t,z) = \frac{\partial y^i}{\partial t_j}(t,z), \qquad 1 \le i \le j \le d, \tag{16}$$

$$y^{if}(t,z) := \frac{\partial^2 x}{\partial t_f \partial t_i}(t,z) = \frac{\partial y^i}{\partial t_f}(t,z), \qquad 1 \le i \le d, \tag{17}$$

$$y^{ff}(t,z) := \frac{\partial^2 x}{\partial t_f \partial t_f}(t,z) = \frac{\partial y^f}{\partial t_f}(t,z), \tag{18}$$

respectively, $\hat{v}^{ij}(t) := v^{ij}(t,\hat{z})$, $\hat{w}^{ij}(t) := w^{ij}(t,\hat{z})$ etc. Also the second–order variational derivatives can be computed via certain IVPs. For notational convenience we shall omit all arguments of the variations in the ODEs.

Proposition 4.5. *The function $v^{ij}(t,z)$, $1 \le i \le j \le d$, is the solution of the IVP*

$$v^{ij}(0,z) = 0, \qquad \dot{v}^{ij} = h_x v^{ij} + (v^i)^T h_{xx} v^j. \tag{19}$$

Proof. By (9), the function $v^i(t,z)$ satisfies

$$v^i(t,z) = e_i + \int_0^{t_f} h_x(s, x(s,z)) v^i(t,s)\, ds.$$

Differentiating this equation with respect to $(x_0)_j$, $j \ge i$, yields

$$v^{ij}(t,z) = \int_0^{t_f} \Big(h_x(s,x(s,z)) v^{ij}(t,s) + (v^i(t,s))^T h_{xx}(s,x(s,z)) v^j(t,s) \Big)\, ds$$

which proves (19). □

Proposition 4.6. *The function $w^{ij}(t,z)$, $1 \le i \le n$, $1 \le j \le d$, satisfies $w^{ij}(t,z) \equiv 0$ on $[0, t_j[$ and for $t \ge t_j$ it is the solution of the IVP*

$$w^{ij}(t_j,z) = -[h_x]^j v^i(t_j,z), \qquad \dot{w}^{ij} = h_x w^{ij} + (v^i)^T h_{xx} y^j. \tag{20}$$

Proof. As in Proposition 4.3, we have $w^{ij} \equiv 0$ on $[0, t_j[$. The solution of the IVP (9) can be written as

$$v^i(t,z) = v^i(t_j^-, z) + \int_{t_j^+}^{t_f} h_x(s, x(s,z)) v^i(s,z)\, ds.$$

Differentiating this equation with respect to t_j yields

$$\begin{aligned} w^{ij}(t,z) &= \dot{v}^i(t_j^-, z) - \dot{v}^i(t_j^+, z) \\ &\quad + \int_{t_j^+}^{t_f} \left(h_x(s, x(s,z)) w^{ij}(s,z) + (v^i(s,z))^T h_{xx}(s, x(s,z)) y^j(s,z) \right) ds \end{aligned}$$

which proves (20) in view of the representation (9) for $\dot{v}^i$. □

The following result is analogous to (12).

Proposition 4.7. *The function $v^{if}(t,z)$, $1 \le i \le n$, satisfies*

$$v^{if}(t_f, z) = \dot{v}^i(t_f, z) = h_x(t_f, x(t_f, z)) v^i(t_f, z). \tag{21}$$

Proposition 4.8. *The function $y^{ij}(t,z)$, $1 \le i \le j \le d$, satisfies $y^{ij}(t,z) \equiv 0$ on $[0, t_j[$ and for $t \ge t_j$ it is the solution of the following IVP:*

(a) For $i = j$ we have

$$\begin{aligned} y^{ii}(t_i, z) &= -[h_t]^i - [h_x]^i h^{i-} - h_x^{i+} y^i(t_i, z), \\ \dot{y}^{ii} &= h_x y^{ii} + (y^i)^T h_{xx} y^i. \end{aligned} \tag{22}$$

(b) For $i < j$ the IVP is given by

$$y^{ij}(t_j, z) = -[h_x]^j y^i(t_j, z), \qquad \dot{y}^{ij} = h_x y^{ij} + (y^i)^T h_{xx} y^j. \tag{23}$$

Proof. As in Proposition 4.3, we have $y^{ij} \equiv 0$ on $[0, t_j[$. Result (b) can be proved in complete analogy to Proposition 4.6. In the proof for (a), differentiation of equation (11) with respect to t_i yields

$$y^{ii}(t,z) = -[h_t]^i - [h_x]^i \dot{x}^{i-} - (\dot{y}^i)^{i+} + \int_{t_i^+}^{t} \left(h_x y^{ii} + (y^i)^T h_{xx} y^i \right) ds.$$

In view of (10), IVP (22) is proved. □

The next two results are analogous to (12).

Proposition 4.9. *The function* $y^{if}(t,z)$, $1 \le i \le d$, *satisfies*

$$y^{if}(t_f,z) = \dot{y}^i(t_f,z) = h_x(t_f,x(t_f,z))y^i(t_f,z). \tag{24}$$

Proposition 4.10. *The function* $y^{ff}(t,z)$ *satisfies*

$$y^{ff}(t_f,z) = \dot{h}(t_f,x(t_f,z)) = h_t(t_f,x(t_f,z)) + h_x(t_f,x(t_f,z))h(t_f,x(t_f,z)). \tag{25}$$

Remark 4.1. If the control is bang–bang in $[0,t_f]$, all terms involving u_x and u_{xx} vanish in (2) and (3). Hence, we have $h_x = f_x$ and $h_{xx} = f_{xx}$ in this case and the representations of the variational derivatives are identical to those given by Osmolovskii/Maurer (Ref. 9) for bang–bang controls.

4.3 Variational Derivatives of the Lagrangian

Consider a trajectory $\hat{\mathcal{T}} = (\hat{x},\hat{u})$ which satisfies the necessary optimality conditions (7)–(12), (Ref. 14) of the minimum principle.

4.3.1 First–Order Variational Derivatives

We will now calculate explicit representations for the first–order variational derivatives of the Lagrangian with respect to the optimization vector z, i.e., the free initial states, the switching times and the free final time.

Proposition 4.11. *For* $i = 1,\dots,n$ *the following holds:*

$$\frac{\partial}{\partial (x_0)_i}\mathcal{L}(\hat{z},\rho_0,\rho) = 0 \tag{26}$$

Proof. Applying the chain rule and using (23), (Ref. 14) as well as the transversality conditions (8) and (9), (Ref. 14), the first–order variational derivatives of the Lagrangian with respect to free initial values $(x_0)_i$ of the states are given by

$$\begin{aligned}\frac{\partial}{\partial (x_0)_i}\mathcal{L}(\hat{z},\rho_0,\rho) &= l_{(x_0)_i}(\hat{x}_b,\hat{t}_f,\rho_0,\rho) + l_{x_f}(\hat{x}_b,\hat{t}_f,\rho_0,\rho)\hat{v}^i(\hat{t}_f)\\ &= -\lambda_i(0) + \lambda(\hat{t}_f)\hat{v}^i(\hat{t}_f).\end{aligned} \tag{27}$$

Together with (9), the last term can be written as

$$\lambda_i(0) + \int_0^{\hat{t}_f} \frac{d}{dt}(\lambda \hat{v}^i)\,dt. \tag{28}$$

Let us transform the integrand. In view of (7), (Ref. 14), (9) and (4), we obtain

$$\frac{d}{dt}(\lambda\hat{v}^i) = \dot{\lambda}\hat{v}^i + \lambda\dot{\hat{v}}^i = (-H_x + \lambda h_x)\hat{v}^i = 0. \tag{29}$$

Substituting (28) and (29) into (27) yields (26). □

Proposition 4.12. *For $i = 1, \ldots, d$, we have*

$$\frac{\partial}{\partial t_i}\mathcal{L}(\hat{z}, \rho_0, \rho) = 0. \tag{30}$$

Proof. Equation (23), (Ref. 14) together with transversality condition (9), (Ref. 14) yields

$$\frac{\partial}{\partial t_i}\mathcal{L}(\hat{z}, \rho_0, \rho) = l_{x_f}(\hat{x}_b, \hat{t}_f, \rho_0, \rho)\hat{y}^i(\hat{t}_f) = \lambda(\hat{t}_i)\hat{y}^i(\hat{t}_i) + \int_{\hat{t}_i}^{\hat{t}_f} \frac{d}{dt}(\lambda\hat{y}^i)dt.$$

Due to (7), (Ref. 14), (10) and (4), the integrand satisfies

$$\frac{d}{dt}(\lambda\hat{y}^i) = \dot{\lambda}\hat{y}^i + \lambda\dot{\hat{y}}^i = (-H_x + \lambda h_x)\hat{y}^i = 0.$$

Furthermore, (10) implies that the first term can be written as

$$\lambda(\hat{t}_i)\hat{y}^i(\hat{t}_i) = -\lambda(\hat{t}_i)[h]^i = -\sigma(\hat{t}_i)[\hat{u}]^i = 0$$

as per definition of a switching time, cf., Remark 2.1, (Ref. 14). This proves (30). □

Proposition 4.13. *If the final time t_f is free, the following holds:*

$$\frac{\partial}{\partial t_f}\mathcal{L}(\hat{z}, \rho_0, \rho) = 0. \tag{31}$$

Proof. Equation (23), (Ref. 14) together with transversality condition (9), (Ref. 14) and equation (11), (Ref. 14) yields

$$\begin{aligned}\frac{\partial}{\partial t_f}\mathcal{L}(\hat{z}, \rho_0, \rho) &= l_{x_f}(\hat{x}_b, \hat{t}_f, \rho_0, \rho)\hat{y}^f(\hat{t}_f) + l_{t_f}(\hat{x}_b, \hat{t}_f, \rho_0, \rho) \\ &= (\lambda f)(\hat{t}_f) - H(\hat{t}_f) = 0.\end{aligned}$$

This proves (31). □

At this point, we will summarize our results.

Lemma 4.1. *Let $\hat{\mathcal{T}} = (\hat{x}, \hat{u})$ be a trajectory which satisfies the necessary conditions (7)–(12), (Ref. 14) of the minimum principle. Then, the first–order variational derivatives of the Lagrangian vanish, i.e.,*

$$\frac{\partial}{\partial z}\mathcal{L}(\hat{z}, \rho_0, \rho) = 0.$$

In other words, Lemma 4.1 says the following.

Corollary 4.1. *Let $\hat{\mathcal{T}} = (\hat{x}, \hat{u})$ be a trajectory which satisfies the necessary conditions (7)–(12), (Ref. 14) of the minimum principle. Then, the necessary conditions (24), (Ref. 14) in the induced optimization problem (21), (Ref. 14) are fulfilled.*

4.3.2 Second–Order Variational Derivatives

We will now present explicit representations for the second–order variational derivatives of the Lagrangian with respect to the optimization vector z. Due to symmetry of the matrix $\mathcal{L}_{zz}$, we will only investigate the derivatives $\mathcal{L}_{z_i z_j}$ for $i \leq j$. After presenting all results, we will give comments on how to prove the representations. For notational convenience, we will drop all arguments in the endpoint Lagrangian l and its partial derivatives which will be evaluated at $(\hat{x}_b, \hat{t}_f, \rho_0, \rho)$.

Proposition 4.14. *For $1 \leq i \leq j \leq n$ we have*

$$\begin{aligned}\frac{\partial^2}{\partial(x_0)_j \partial(x_0)_i}\mathcal{L}(\hat{z},\rho_0,\rho) &= l_{(x_0)_i(x_0)_j} + l_{(x_0)_i x_f}\hat{v}^j(\hat{t}_f)\\ &\quad + (\hat{v}^i(\hat{t}_f))^T\big(l_{x_f(x_0)_j} + l_{x_f x_f}\hat{v}^j(\hat{t}_f)\big)\\ &\quad + \int_0^{\hat{t}_f} (\hat{v}^i)^T(H_{xx} + H_{xu}u_x + (u_x)^T H_{ux})\hat{v}^j dt.\end{aligned} \tag{32}$$

Proposition 4.15. *For $i = 1,\ldots,n$ and $j = 1,\ldots,d$ the following holds:*

$$\begin{aligned}\frac{\partial^2}{\partial t_j \partial(x_0)_i}\mathcal{L}(\hat{z},\rho_0,\rho) &= l_{(x_0)_i x_f}\hat{y}^j(\hat{t}_f) + \hat{v}^i(\hat{t}_f) l_{x_f x_f}\hat{y}^j(\hat{t}_f)\\ &\quad + \int_{\hat{t}_j}^{\hat{t}_f} (\hat{v}^i)^T(H_{xx} + H_{xu}u_x + (u_x)^T H_{ux})\hat{y}^j dt.\end{aligned} \tag{33}$$

Proposition 4.16. *For $i = 1,\ldots,n$ we obtain*

$$\begin{aligned}\frac{\partial^2}{\partial t_f \partial(x_0)_i}\mathcal{L}(\hat{z},\rho_0,\rho) &= l_{(x_0)_i t_f} + l_{(x_0)_i x_f}\hat{y}^f(\hat{t}_f)\\ &\quad + \hat{v}^i(\hat{t}_f)^T\big(l_{x_f t_f} + l_{x_f x_f}\hat{y}^f(\hat{t}_f)\big) + H_x(\hat{t}_f)\hat{v}^i(\hat{t}_f).\end{aligned} \tag{34}$$

Proposition 4.17. *The variational derivatives with respect to the switching times are given as follows.*

(a) For $i = 1,\ldots,d$ we have

$$\begin{aligned}\frac{\partial^2}{\partial t_i \partial t_i}\mathcal{L}(\hat{z},\rho_0,\rho) &= D^i(H) - [H_x]^i\hat{y}^i(\hat{t}_i) + \hat{y}^i(\hat{t}_f)^T l_{x_f x_f}\hat{y}^i(\hat{t}_f)\\ &\quad + \int_{\hat{t}_i}^{\hat{t}_f} (\hat{y}^i)^T(H_{xx} + H_{xu}u_x + u_x^T H_{ux})\hat{y}^i\, dt.\end{aligned} \tag{35}$$

(b) For $1 \le i < j \le d$ we get

$$\frac{\partial^2}{\partial t_j \partial t_i}\mathcal{L}(\hat{z},\rho_0,\rho) = -[H_x]^j \hat{y}^i(\hat{t}_j) + \hat{y}^j(\hat{t}_f)^T l_{x_f x_f} \hat{y}^i(\hat{t}_f) + \int\limits_{\hat{t}_j}^{\hat{t}_f} (\hat{y}^j)^T (H_{xx} + H_{xu}u_x + u_x^T H_{ux}) \hat{y}^i \, dt. \tag{36}$$

Proposition 4.18. *For $i = 1, \dots, d$ we obtain*

$$\frac{\partial^2}{\partial t_f \partial t_i}\mathcal{L}(\hat{z},\rho_0,\rho) = (\hat{y}^i(\hat{t}_f))^T \left(l_{x_f x_f} \hat{y}^f(\hat{t}_f) + l_{x_f t_f} \right) + (H_x \hat{y}^i)(\hat{t}_f). \tag{37}$$

Proposition 4.19. *The following holds:*

$$\frac{\partial^2}{\partial t_f \partial t_f}\mathcal{L}(\hat{z},\rho_0,\rho) = (\hat{y}^f(\hat{t}_f))^T \left(l_{x_f x_f} \hat{y}^f(\hat{t}_f) + l_{x_f t_f} \right) + l_{t_f x_f} \hat{y}^f(\hat{t}_f) + l_{t_f t_f} + (H_t + H_x f)(\hat{t}_f). \tag{38}$$

The proofs for Propositions 4.14-4.19 are very similar. As an example, we will give the proof for Proposition 4.17 below. The structure in each proof is as follows. Applying the chain rule to calculate the derivative of $\mathcal{L}$, one obtains one term including a second–order variational derivative of the state. This term can be transformed into an integral term in which the second–order variational derivative can then be deleted by using the IVP representations calculated in the previous paragraph.

Proof. For both cases (a) and (b), i.e. $1 \le i \le j \le d$, we obtain

$$\mathcal{L}_{t_i t_j} = \frac{\partial}{\partial t_j}\left(l_{x_f} y^i \right)|_{z=\hat{z}} = \hat{y}^i(\hat{t}_f) l_{x_f x_f} \hat{y}^j(\hat{t}_f) + l_{x_f} \hat{y}^{ij}(\hat{t}_f).$$

In view of transversality condition (9), (Ref. 14), the last term can be written as

$$l_{x_f} \hat{y}^{ij}(\hat{t}_f) = \lambda(\hat{t}_f) \hat{y}^{ij}(\hat{t}_f) = \lambda(\hat{t}_j) \hat{y}^{ij}(\hat{t}_j) + \int\limits_{\hat{t}_j}^{\hat{t}_f} \frac{d}{dt}(\lambda \hat{y}^{ij}) \, dt \tag{39}$$

Using the adjoint differential equation (7), (Ref. 14), ODE (23) for y^{ij} and equations (4) and (5), we have

$$\begin{aligned} \frac{d}{dt}(\lambda \hat{y}^{ij}) &= -H_x \hat{y}^{ij} + \lambda h_x \hat{y}^{ij} + \lambda (\hat{y}^i)^T h_{xx} \hat{y}^j \\ &= (\hat{y}^i)^T (H_{xx} + H_{xu} u_x + (u_x)^T H_{ux}) \hat{y}^j \end{aligned}$$

for the integrand. We point out that equation (4) was essential to delete the second–order variational derivative y^{ij} from the integrand. For the first term in (39) we will consider the cases $i = j$ and $i < j$ separately.

(a) For $i = j$, the initial condition in (22) for $\hat{y}^{ii}(\hat{t}_i)$ together with (4) yields

$$\begin{aligned}\lambda(\hat{t}_i)\hat{y}^{ii}(\hat{t}_i) &= \lambda(\hat{t}_i)(-[h_t]^i - [h_x]^i h^{i-} - h_x^{i+}\hat{y}^i(\hat{t}_i)) \\ &= -[H_t]^i - [H_x]^i h^{i-} - H_x^{i+}\hat{y}^i(\hat{t}_i).\end{aligned}$$

Due to the initial condition (10) for $\hat{y}^i(\hat{t}_i)$, we have $H_x^{i-}\hat{y}^i(\hat{t}_i) + H_x^{i-}[h]^i = 0$ which can be added to the last equation. Together with (14), (Ref. 14), this yields

$$\begin{aligned}\lambda(\hat{t}_i)\hat{y}^{ii}(\hat{t}_i) &= -[H_t]^i - [H_x]^i h^{i-} - H_x^{i+}\hat{y}^i(\hat{t}_i) + H_x^{i-}[h]^i + H_x^{i-}\hat{y}^i(\hat{t}_i) \\ &= D^i(H) - [H_x]^i\hat{y}^i(\hat{t}_i).\end{aligned}$$

(b) In the case $i < j$, using initial condition (23) for $\hat{y}^{ij}(\hat{t}_i)$ and (4), we obtain

$$\lambda(\hat{t}_j)\hat{y}^{ij}(\hat{t}_j) = -\lambda(\hat{t}_j)[h_x]^j\hat{y}^i(\hat{t}_j) = -[\lambda h_x]^j\hat{y}^i(\hat{t}_j) = -[H_x]^j\hat{y}^i(\hat{t}_j).$$

Hence, (35) and (36) are proved. □

We note that the proofs for Propositions 4.16, 4.18 and 4.19 are simpler as we have given direct representations of the corresponding second–order variational derivatives instead of IVPs in the last paragraph. Hence, after applying the chain rule, the second–order terms can directly be replaced instead of using the integral approach.

We will summarize all results for the second–order variational derivatives of the Lagrangian function.

Lemma 4.2. *Let $\hat{\mathcal{T}} = (\hat{x}, \hat{u})$ be a trajectory which satisfies the necessary conditions (7)–(12), (Ref. 14) of the minimum principle. Then, the second–order variational derivatives of the Lagrangian are given by (32)–(38) and hence, depend only on first–order but not on second–order variational derivatives of the states.*

4.4 Variational Derivatives of the Function Φ

Concluding the variational computations in this section, we will now calculate the variational derivatives of the function Φ which are essential for the verification of second–order sufficient conditions (25), (Ref. 14) if the induced problem involves constraints. Applying the chain rule, we obtain

$$\begin{aligned}\frac{\partial}{\partial(x_0)_i}\Phi(\hat{z}) &= \phi_{(x_0)_i}(\hat{x}_b, \hat{t}_f) + \phi_{x_f}(\hat{x}_b, \hat{t}_f)\hat{v}^i(\hat{t}_f), \qquad i = 1, \ldots, n, \\ \frac{\partial}{\partial t_i}\Phi(\hat{z}) &= \phi_{x_f}(\hat{x}_b, \hat{t}_f)\hat{y}^i(\hat{t}_f), \qquad i = 1, \ldots, d, \\ \frac{\partial}{\partial t_f}\Phi(\hat{z}) &= \phi_{x_f}(\hat{x}_b, \hat{t}_f)\hat{y}^f(\hat{t}_f) + \phi_{t_f}(\hat{x}_b, \hat{t}_f).\end{aligned} \tag{40}$$

5.2 Goddard Problem

We consider the following optimal control problem with three state variables h, v and m, a scalar control u and free final time t_f. This model can also be found in Bryson/Ho (Ref. 15) and Maurer (Ref. 16). We note that we have taken over the notations from the references as the notations therein are suitable to the meaning of the occuring functions and parameters (see below). There shall be no confusion with the notations used before.

$$
\begin{array}{ll}
\max & h(t_f) \\
\text{s. t.} & \dot{h} = v, \quad \dot{v} = \frac{1}{m}(cu - D(v,h)) - g(h), \quad \dot{m} = -u, \\
& h(0) = h_0, \quad v(0) = v_0, \quad m(0) = m_0, \quad m(t_f) = m_f, \\
& 0 \le u(t) \le u^{\max} \qquad \forall\, t \in [0, t_f].
\end{array}
\tag{41}
$$

Here, h denotes the height, v the velocity and m the mass of a rocket which shall be controlled to a maximal height at the end of the time horizon. The initial mass m_0 consists of the rocket mass m_f and the mass of the initial amount of fuel in the rocket. The dynamics involve the drag function $D(v,h)$ and the gravity function $g(h)$ which are defined as

$$
D(v,h) = \alpha v^2 \exp(-\beta h), \qquad g(h) = g_0 \frac{r_0^2}{(r_0 + h)^2}.
$$

The data are taken from Maurer (Ref. 16):

$$
\begin{array}{c}
\alpha = 0.01227, \quad \beta = 0.000145, \quad g_0 = 9.81, \quad r_0 = 6.371 \cdot 10^6, \quad c = 2060, \\
m_0 = 214.839, \quad m_f = 67.9833, \quad u^{\max} = 9.52551, \quad h_0 = v_0 = 0, \quad t_f \text{ free.}
\end{array}
$$

The Hamiltonian and the switching function are given by

$$
\begin{aligned}
H(x,\lambda,u) &= \lambda_h v + \lambda_v \Big(\frac{1}{m}(cu - D(v,h)) - g(h) \Big) - \lambda_m u, \\
\sigma(x,\lambda) &= \frac{c\lambda_v}{m} - \lambda_m.
\end{aligned}
$$

Using the solver IPOPT, we obtain the following optimal control structure:

$$
u(t) = \begin{cases} u^{\max}, & 0 \le t \le t_1, \\ u^{\text{sing}}(x(t)), & t_1 \le t \le t_2, \\ 0, & t_2 \le t \le t_f. \end{cases}
$$

where, as it is shown in Maurer (Ref. 16), the singular control of order $q = 1$ can be obtained in the feedback form

$$
u^{\text{sing}}(h,v,m) = \frac{D}{c} + m \frac{(c-v)D_h + (D_v + cD_{vv})g + c(mg_h - D_{vh}v)}{D + 2cD_v + c^2 D_{vv}}.
$$

Hence, the induced optimization problem involving the optimization vector $z = (t_1, t_2, t_f)^T$, respectively, $\tilde{z} = (\zeta_1, \zeta_2, \zeta_3)^T$ is given by

$$\begin{array}{ll} \min & h(t_f, z) \\ \text{s. t.} & m(t_f, z) - m_f = 0. \end{array}$$

NUDOCCCS provides a solution with switching times $t_1 = 4.11526$, $t_2 = 46.04061$ and the final time $t_f = 212.90299$. The maximal height is $h(t_f) = 161445.136$ and the jumps of u are given by $[u]^1 = -7.39657$, $[u]^2 = -4.18838$. We depict the control with the switching function in Figure 1 and the states in Figure 2.

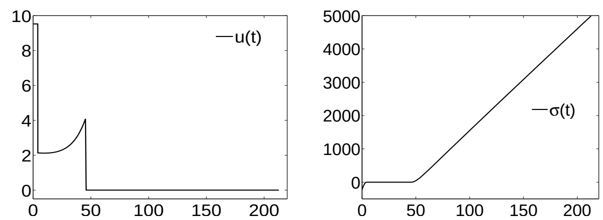

Figure 1: Optimal control u and switching function σ for the Goddard Problem

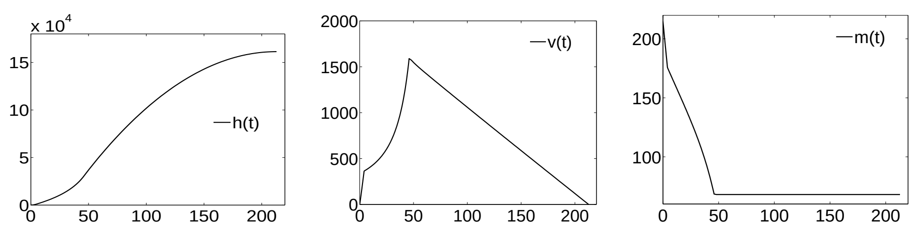

Figure 2: Optimal states h, v, m for the Goddard Problem

Finally, we will verify the SSC for this solution z. Figures 3 and 4 show the first–order variational derivatives of the states with respect to t_1 and t_2, respectively, where y^i satisfies IVP (10) with initial condition $y^i(t_i) = (0, -c[u]^i/m(t_i), [u]^i)^T$, $i = 1, 2$.

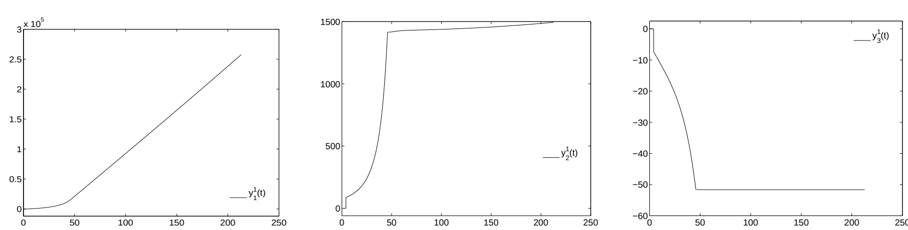

Figure 3: First–order variational derivatives y_1^1, y_2^1, y_3^1 for the Goddard Problem

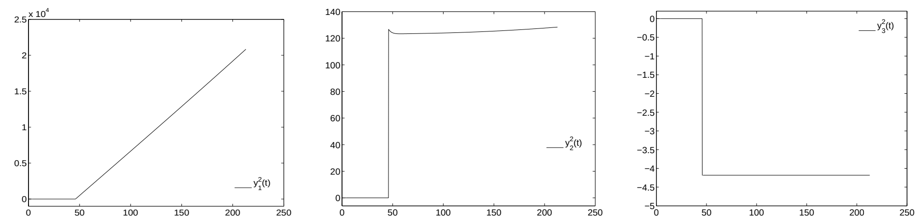

Figure 4: First–order variational derivatives y_1^2, y_2^2, y_3^2 for the Goddard Problem

In view of (37), (38) and transversality condition (9), (Ref. 14), we obtain the following entries in the Hessian matrix of the Lagrangian:

$$\mathcal{L}_{t_i t_f} = (H_x y^i)(t_f) = -y_2^i(t_f), \quad i = 1, 2, \qquad \mathcal{L}_{t_f t_f} = (H_x f)(t_f) = g(t_f),$$

whereas the entries $L_{t_i t_j}$, $1 \le i \le j \le 2$, are given by (35) and (36). Here, $D^1(H) = D^2(H) = 0$ holds due to Corollary 2.2, (Ref. 14). Furthermore, we have

$$\Phi_{t_i} = (\phi_{x_f} y^i)(t_f) = y_3^i(t_f), \quad i = 1, 2, \qquad \Phi_{t_f} = (\phi_{x_f} y^f)(t_f) = 0.$$

Note that $\Phi_{t_2} = [u]^2 = -u^{\text{sing}}(t_2^-)$ holds as we have $\dot{y}_3^2 \equiv 0$ along the interval J_3. Hence, we obtain

$$\mathcal{L}_{zz} = \begin{pmatrix} -207556.986340 & -16491.426723 & -1494.323147 \\ -16491.426723 & -1235.204985 & -128.567507 \\ -1494.323147 & -128.567507 & 9.331096 \end{pmatrix},$$
$$\Phi_z = (-51.635009, -4.186749, 0).$$

Obviously, z is normal but the matrix $\mathcal{L}_{zz}$ is not positive definite on $\mathbb{R}^3$. However, the reduced Hessian matrix defined in (26), (Ref. 14) is given by

$$H_{\text{red}} = N^T \mathcal{L}_{zz} N = \begin{pmatrix} 74.082672 & -7.378291 \\ -7.378291 & 9.331096 \end{pmatrix}$$

and hence, positive definite on $\mathbb{R}^2$ with eigenvalues 74.9128 and 8.500999. Therefore, the switching times and the final time are optimal due to the SSC (25), (27), (Ref. 14). We conclude with the remark that NUDOCCCS provides the matrices $\tilde{\mathcal{L}}_{\tilde{z}\tilde{z}}$ and $\tilde{\Phi}_{\tilde{z}}$ which lead to similar matrices $\mathcal{L}_{zz}$ and Φ_z by using formulas (31), (Ref. 14):

$$\mathcal{L}_{zz} = \begin{pmatrix} -206969.265855 & -16492.177383 & -1492.002048 \\ -16492.177383 & -1239.847916 & -128.833926 \\ -1492.002048 & -128.833926 & 9.327778 \end{pmatrix},$$
$$\Phi_z = (-51.560409, -4.195724, 0).$$

The reduced Hessian is obtained as

$$H_{\text{red}} = \begin{pmatrix} 73.2412 & -7.39793 \\ -7.39793 & 9.32778 \end{pmatrix}$$

with eigenvalues 74.0863 and 8.48265. In comparison to our method, the maximal relative difference of the matrix entries is 1.14%, the maximal relative difference of the eigenvalues is 1.10%.

References

[1] MAURER, H., First and second–order sufficient optimality conditions in mathematical programming and optimal control, *Mathematical Programming Study*, **10**, pp. 163–177, 1981.

[2] PICKENHAIN, S., Sufficiency conditions for weak local minima in multidimensional optimal control problems with mixed control-state restrictions, *Zeitschrift für Analysis und ihre Anwendungen*, **11**, pp. 559–568, 1992.

[3] ZEIDAN, V., The Riccati equation for optimal control problems with mixed state-control constraints: necessity and sufficiency, *SIAM Journal on Control and Optimization*, **32**, 1297–1321, 1994.

[4] BOLTYANSKII, V. G., Sufficient conditions for optimality and the justification of the dynamic programming method, *SIAM Journal on Control and Optimization*, **4**, 326–361, 1966.

[5] PICCOLI, B. and SUSSMANN, H. J., Regular synthesis and sufficiency conditions for optimality, *SIAM Journal on Control and Optimization*, **39**, 359–410, 2001.

[6] NOBLE, J. and SCHÄTTLER, H., Sufficient conditions for relative minima of broken extremals in optimal control theory, *Journal of Mathematical Analysis and Applications*, **269**, No. 1, 98–128, 2002.

[7] AGRACHEV, A. A., STEFANI, G. and ZEZZA, P. L., Strong optimality for a bang–bang trajectory, *SIAM Journal on Control and Optimization*, **41**, No. 4, pp. 991–1014, 2002.

[8] OSMOLOVSKII, N. P. and MAURER, H., Equivalence of second–order optimality conditions for bang–bang control problems, part 1, *Control and Cybernetics*, **34**, Part 3, pp. 927–950, 2005.

[9] OSMOLOVSKII, N. P. and MAURER, H., Equivalence of second–order optimality conditions for bang–bang control problems, part 2, *Control and Cybernetics*, **36**, Part 1, pp. 5–46, 2007.

[10] DMITRUK, A. V., Quadratic order conditions of a local minimum for singular extremals in a general optimal control problem, *Proc. of Symposia in Pure Math*, American Math. Society, **64**, pp. 163–198, 1999.

[11] POGGIOLINI, L. and STEFANI, G., Minimum time optimality for a bang-singular arc: second–order sufficient conditions, Proceedings of the 44th IEEE Conference on Decision and Control, 2005 and 2005 European Control Conference, CDC-ECC '05, Seville, Spain, pp. 1433–1438, 2005.

[12] VOSSEN, G., *Numerische Lösungsmethoden, hinreichende Optimalitätsbedingungen und Sensitivitätsanalyse für optimale bang–bang und singuläre Steuerungen*, PhD Thesis, Universität Münster, Germany, 2006.

[13] PONTRYAGIN, L. S., BOLTYANSKII, V. G., GAMKRELIDZE, R. V. and MISHCHENKO, E. F., *The Mathematical Theory of Optimal Processes*, Fizmatgiz, Moscow, 1961. English translation: Pergamon Press, New York, 1964.

[14] VOSSEN, G., Switching Time Optimization for Bang–Bang and Singular Controls, to appear in *Journal of Optimization Theory and Applications*, 2009

[15] BRYSON, A. E. and HO, Y. C., *Applied Optimal Control*, Wiley, New York, 1975.

[16] MAURER, H., Numerical solution of singular control problems using multiple shooting techniques, *Journal of Optimization Theory and Applications*, **18**, No. 2, pp. 235–257, 1976.

[17] WÄCHTER, A. and BIEGLER, L. T., On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, *Mathematical Programming*, **106**, No. 1, pp. 25–57, 2006.

[18] BÜSKENS, C., *Optimierungsmethoden und Sensitivitätsanalyse für optimale Steuerprozesse mit Steuer- und Zustandsbeschränkungen*, PhD Thesis, Universität Münster, Germany, 1998.

[19] MILYUTIN, A. A. and OSMOLOVSKII, N. P., *Calculus of Variations and Optimal Control*, Transl. Math. Monogr., **180**, AMS, Providence, RI., 1998.

[20] SZYMKAT, M. and KORYTOWSKI, A., Method of monotone structural evolution for control and state constrained optimal control problems. Proceedings of the European Control Conference ECC 2003, University of Cambridge, U.K., September 1-4, 2003.

[21] KELLEY, H. J., KOPP, R. E. and MOYER, H. G., Singular extremals, *Topics in Optimization*, Edited by G. Leitmann, Academic Press, New York, pp. 63–101, 1967.

[22] BELL, D. J. and JACOBSON, D. H., *Singular Optimal Control Problems*, Academic Press, London et al., 1975.

[23] FRASER-ANDREWS, G., Finding candidate singular optimal controls: a state of the art survey, *Journal of Optimization Theory and Applications*, **60**, No. 2, pp. 173–190, 1989.

[24] MCDANELL, J. P. and POWERS, W. F., Necessary conditions for joining singular and nonsingular subarcs, *SIAM Journal on Control*, **9**, No. 2, pp. 161–173, 1971.

[25] KRENER, A. J., The high order maximal principle and its application to singular extremals, *SIAM Journal on Control and Optimization*, **15**, No. 2, pp. 256–293, 1977.

[26] VOSSEN, G. and MAURER, H., On L^1–minimization in optimal control and applications to robotics, *Optimal Control Applications and Methods*, **27**, pp. 301–321, 2006.

[27] CHYBA, M., SUSSMANN, H., MAURER, H. and VOSSEN, G., *Underwater vehicles: The minimum time problem*, Proceedings of the 43rd IEEE Conference on Decision and Control, CDC 04, The Bahamas, IEEE Control Society, pp. 1370–1375, 2004.

[28] FIACCO, A. V. and MCCORMICK, G. P., *Nonlinear Programming*, John Wiley, New York, 1968.

[29] KAYA, C. Y. and NOAKES, J. L. Computational method for time-optimal switching control, *Journal of Optimization Theory and Applications*, **117**, pp. 69–92, 2003.

[30] MAURER, H., BÜSKENS, C., KIM, J.-H. R. and KAYA, C. Y., Optimization methods for the verification of second–order sufficient conditions for bang–bang controls, *Optimal Control Applications and Methods*, **26**, pp. 129–156, 2005.

[31] LEE, H. W. J., TEO, K. L., JENNINGS, L. S. and REHBOCK, V., Control Parametrization Enhancing Technique for Time Optimal Control Problems, *Dynamic Systems and Applications*, **6**, Vol. 2, 243-261, 1997.

[32] VOSSEN, G., REHBOCK, V. and SIBURIAN, A., Numerical solution methods for singular control with multiple state dependent forms, *Optimization Methods and Software*, **22**, No. 4, pp. 551-559, 2007.

[33] BÜSKENS, C., PESCH, H. J. and WINDERL, S., Real-time solutions to bang–bang and singular optimal control problems, *Online Optimization of Large Scale Problems*, Edited by M. Grötschel, S.O. Krumke and J. Rambau), Springer-Verlag, Berlin, pp. 129–142, 2001.

[34] JAMES, E. M., Time optimal control and the Van der Pol Oscillator, *Journal of Applied Mathematics*, **13**, No. 1, pp. 67–81, 1974.

[35] STOER, J. and BULIRSCH, R., *Introduction to Numerical Analysis*, Springer, New York, 1993.