Single Shooting and ESDIRK Methods for adjoint-based optimization of an oil reservoir

Capolei, Andrea; Völcker, Carsten; Frydendall, Jan; Jørgensen, John Bagterp

Published in:
Proceedings of the 17th Nordic Process Control Workshop

Publication date:
2012

Document Version
Publisher’s PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Single Shooting and ESDIRK Methods for adjoint-based optimization of an oil reservoir

Andrea Capolei * Carsten Völcker * Jan Frydendall * John Bagterp Jørgensen *

* Department of Informatics and Mathematical Modeling & Center for Energy Resources Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark. (e-mail: {acap,cv,jf,jb}@imm.dtu.dk).

Abstract: Conventional recovery techniques enable recovery of 10–50% of the oil in an oil field. Advances in smart well technology and enhanced oil recovery techniques enable significant larger recovery. To realize this potential, feedback model-based optimal control technologies are needed to manipulate the injections and oil production such that flow is uniform in a given geological structure. Even in the case of conventional water flooding, feedback based optimal control technologies may enable higher oil recovery than with conventional operational strategies. The optimal control problems that must be solved are large-scale problems and require specialized numerical algorithms. In this paper, we combine a single shooting optimization algorithm based on sequential quadratic programming (SQP) with explicit singly diagonally implicit Runge-Kutta (ESDIRK) integration methods and a continuous adjoint method for sensitivity computation. We demonstrate the procedure on a water flooding example with conventional injectors and producers.

Keywords: Optimal Control, Optimization, Numerical Methods, Oil Reservoir

1. INTRODUCTION

The growing demand for oil and the decreasing number of newly discovered significant oil fields require more efficient management of the existing oil fields. Oil fields are developed in two or three phases. In the primary phase, the reservoir pressure is large enough to make the oil flow to the production wells. In the secondary phase, water must be injected to maintain pressure and move the oil towards the producers. In some cases, a tertiary phase known as enhanced oil recovery is considered. Enhanced oil recovery includes technologies such as in situ combustion, surfactant flooding, polymer flooding, and steam flooding (Thomas, 2008). After the secondary phase, typically the oil recovery is somewhere between 10% and 50% (Chen, 2007; Jansen, 2011).

Optimal control technology and Nonlinear Model Predictive Control have been suggested for improving the oil recovery of the secondary phase (Jansen et al., 2008). In such applications, the controller adjusts the water injection rates and the bottom hole well pressures to maximize oil recovery or a financial measure such as net present value. In the oil industry, this control concept is also known as closed-loop reservoir management (Jansen et al., 2009). The controller in closed-loop reservoir management consists of a state estimator for history matching and an optimizer that solves a constrained optimal control problem for the production optimization. The main difference of the closed-loop reservoir management system from a traditional Nonlinear Model Predictive Controller (Binder et al., 2001) is the large state dimension ($10^6$ is not unusual) of an oil reservoir model. The size of the problem dictates that the ensemble Kalman filter is used for state estimation (history matching) and that single shooting optimization algorithms compute gradient based on adjoints (Jansen, 2011; Jørgensen, 2007; Sarma et al., 2005; Suwartadi et al., 2011; Völcker et al., 2011).

In this paper, we propose a high order temporal integration method (Explicit Singly Diagonally Implicit Runge-Kutta, ESDIRK) for forward computation of the initial value problem and for backward solution of the associated continuous-time adjoint. Conventional practice by commercial reservoir simulators is limited to the use of first order temporal implicit or semi-implicit integrators for the initial value problem and for backward solution of the associated continuous-time adjoint. Conventional practice by commercial reservoir simulators is limited to the use of first order temporal implicit or semi-implicit integrators for the initial value problem and the adjoints. Völcker et al. (2010a,b, 2009) introduce high order ESDIRK methods in two phase reservoir simulation. The high order scheme allows larger steps and therefore faster solution of the reservoir model equations. To compute the gradient of the objective function in a single shooting optimization method, Völcker et al. (2011) propose a method based on adjoints for the discretized equations. Cao et al. (2002) and Jansen (2011) provide an overview of gradient computation using the adjoint. Brouwer and Jansen (2004) and Sarma et al. (2005) explain and demonstrate gradient computation by the adjoint equations based on the implicit Euler discretization. Kourounis et al. (2010) suggest the

* This research project is financially supported by the Danish Research Council for Technology and Production Sciences. FTP Grant no. 274-06-0284
contiguous-time high order adjoint equations for gradient computation in production optimization. Nadarajah and Jameson (2007) compare gradients computed by discrete and contiguous adjoints for problems arising in aerodynamics. They conclude that the gradients computed from contiguous adjoints are accurate enough to be used in optimization algorithms. Since computation of gradients based on contiguous time adjoints is faster than gradients based on discrete adjoints, this conclusion implies that the gradient computations can be accelerated by using the contiguous time adjoint equations.

The novel contribution in this paper is an extension of the adjoint based optimization method suggested by Völker et al. (2011) to include gradient computation based on the continuous-time adjoint equation. Using a conventional oil field as case study, we demonstrate the new single-shooting optimization algorithm based on ESDIRK integration of the initial value problem and ESDIRK integration of the continuous-time adjoint equation. The case study illustrates the potential of optimal control for production optimization of water flooded oil reservoirs by maximizing the net present value. We do a parameter study to illustrate the sensitivity of the optimal solution to the discount factor.

The paper is organized as follows. Section 2 states the general constrained optimal control problem using a novel representation of the system dynamics. The ESDIRK algorithm for solution of the differential equation systems is described in Section 3, while Section 4 presents the continuous adjoint method. Section 5 describes the numerical case study and discusses the sensitivity of the optimal solution to the discount factor in the net present value. Conclusions are presented in Section 6.

2. OPTIMAL CONTROL PROBLEM

In this section, we present the continuous-time constrained optimal control problem and its transcription by the single shooting method to a finite dimensional constrained optimization problem. First we present the continuous-time optimal control problem. Then we parameterize the control function using piecewise constant basis functions, and finally we convert the problem into a constrained optimization problem using the single shooting method.

Consider the continuous-time constrained optimal control problem in the Bolza form

\[
\min_{x(t), u(t)} J = \Phi(x(t_b)) + \int_{t_a}^{t_b} \Phi(x(t), u(t)) dt
\]

subject to

\[
x(t_a) = x_0
\]

\[
\frac{dx(t)}{dt} = f(x(t), u(t)) \quad t \in [t_a, t_b]
\]

\[
u(t) \in U(t)
\]

\[
x(t) \in \mathbb{R}^n_x \text{ is the state vector and } u(t) \in \mathbb{R}^n_u \text{ is the control vector.}
\]

The time interval \( I = [t_a, t_b] \) as well as the initial state, \( x_0 \), are assumed to be fixed. (1c) represents the dynamic model and includes systems described by index-1 differential algebraic equations (DAE). (1d) represents constraints on the input values, e.g. \( u_{\text{min}} \leq u(t) \leq u_{\text{max}} \), \( c(u(t)) \geq 0 \), and some constraints related to rate of movement that are dependent on the input parametrization.

Path constraints \( \eta(x(t), u(t)) \geq 0 \) may render the optimization problem infeasible. For this reason and due to computational efficiency considerations when computing the sensitivities by the adjoint method (Capolei and Jørgensen, 2012; Jørgensen, 2007), we include these constraints as soft constraints using the following smooth approximation

\[
\chi_i(x(t), u(t)) = \frac{1}{2} \left( \sqrt{\eta_i(x(t), u(t)^2 + \beta_i^2) - \eta_i(x(t), u(t))} \right)
\]

to the exact penalty function \( \max(0, -\eta_i(x(t))) \) for \( i \in \{1, \ldots, n_\eta\} \). With this approximation of the path constraints, the resulting stage cost, \( \Phi(x(t), u(t)) \), used in (1a) consist of the inherent stage cost, \( \tilde{\Phi}(x(t), u(t)) \), and terms penalizing violation of the path constraints (2)

\[
\Phi(x, u) = \tilde{\Phi}(x, u) + \|\chi(x, u)\|_1 Q_1 + \frac{1}{2} \|\chi(x, u)\|_2^2 Q_2
\]

2.1 Discretization

Control Parametrization Let \( T_i \) denote the sample time such that an equidistant mesh can be defined as

\[
t_a = t_0 < \ldots < t_s < \ldots < t_N = t_b \]

with \( t_j = t_a + j T_i \) for \( j = 0, 1, \ldots, N \). We use a piecewise constant representation of the control function on this equidistant mesh, i.e. we approximate the control vector on every subinterval \([t_j, t_{j+1}]\) by the zero-order-hold parametrization

\[
u(t) = u_j, \quad t \in [t_j, t_{j+1}]
\]

such that (1) can be approximated with the finite dimensional constrained optimization problem

\[
\min_{\{u_k\}_{k=0}^{N-1}} \psi = \psi(\{u_k\}_{k=0}^{N-1}, x_0)
\]

s.t. \( u_{\text{min}} \leq u_k \leq u_{\text{max}} \quad k \in \mathcal{N} \)

\( \Delta u_{\text{min}} \leq \Delta u_k \leq \Delta u_{\text{max}} \)

\( c_k(u_k) \geq 0 \)

with \( \mathcal{N} = \{0, 1, \ldots, N - 1\} \).
In this section, we describe our implementation of the ESDIRK method for the computation of \( \psi(t) \) in (7). Computation of \( \psi(t_{k=0}^{N-1}, x_0) \) consists of two major operations: 1) For each integration step we first compute the model states \( x(t) \) solving the initial value problem (1c), 2) and then we compute, using the same quadrature points, the value of the Lagrange term

\[
\tilde{\psi}(t) := \int_{t_a}^{t} \Phi(x(t), u(t)) dt \quad t_a \leq t \leq t_b.
\]  

(9)

in the cost function (1a). Let \( t_n \) denote the integration times chosen by the step size controller in the integrator. Each integration step size, \( h_n \), is chosen such that it is smaller than or equal to the sample time, \( T \). Therefore, one sample interval contains many integration steps. The numerical solution of the IVP (1c) by an s-stage, stiffly accurate, Runge-Kutta ESDIRK method with an embedded error estimator, may in each integration step \([t_n, t_{n+1}]\) be denoted (Cappozi and Jørgensen, 2012; Völcker et al., 2010a)

\[
T_i = \tilde{t}_n, \quad T_i = \tilde{t}_n + c_i h_n \quad (10a)
\]

\[
X_i = x_n \quad (10b)
\]

\[
\phi_i(\{X_i\}_{j=1}^{i-1}, u) = g(x_i) + h_n \sum_{j=1}^{i-1} a_{ij} f(X_j, u) \quad (10c)
\]

\[
g(x_i) = \phi_i(\{X_i\}_{j=1}^{i-1}, u) + h_n \gamma f(x_i, u) \quad (10d)
\]

\[
x_{i+1} = X_n \quad (10e)
\]

\[
e_i + 1 = h_n \sum_{j=1}^{s} d_j f(X_j, u) \quad (10f)
\]

with \( i = 2, \ldots, s, X_i \) denotes the numerical solution at time \( T_i \) for \( i \in \{1, \ldots, s\} \). \( x_{n+1} \) is the numerical solution at time \( t_{n+1} = \tilde{t}_n + h_n \). \( e_{n+1} \) is the estimated error of the numerical solution, i.e. \( ||e_{n+1}|| \approx ||g(x_{n+1}) - g(x(t_{n+1}))|| \). Subsequent to solution of (10), we compute the numerical solution of the cost function (9)

\[
\tilde{\psi}(t_{n+1}) = \tilde{\psi}(\tilde{t}_n) + h_n \sum_{j=1}^{s} b_j \Phi(X_j, u) \quad (11)
\]

When \( t_{n+1} = t_b \), we add the Mayer term of (1a) such that

\[
\psi(\{u_k\}_{k=0}^{N-1}; x_0) = \psi(t_b) = \tilde{\psi}(t_b) + \tilde{\Phi}(x(t_b)) \quad (12)
\]

The main computational effort in the ESDIRK method is solution of the implicit equations (10d) using Newton based method. (10d) is solved by sequential solution of

\[
R_i(X_i) := g(x_i) - h_n \gamma f(x_i, u) - \phi_i(\{X_i\}_{j=1}^{i-1}, u) = 0
\]

for \( i = 2, \ldots, s \). (13) is solved using an inexact Newton method. Each iteration in the inexact Newton method for solution of (13) may be denoted

\[
M \Delta X_i^{[l]} = -R_i(X_i) \quad (14a)
\]

\[
X_i^{[l+1]} = X_i^{[l]} + \Delta X_i^{[l]} \quad (14b)
\]

The iteration matrix, \( M \), is an approximation

\[
M \approx J(X_i^{[l]}) \quad (15)
\]

to the Jacobian of the residual function

\[
J_i(X_i) = \frac{\partial R_i}{\partial X_i}(X_i) = \frac{\partial \tilde{\Phi}}{\partial X_i}(X_i) - h_n \gamma \frac{\partial f}{\partial X}(X_i, u) \quad (16)
\]

The iteration matrix, \( M \), and its LU factorization is updated adaptively by monitoring the convergence rate of the inexact Newton iterations. Convergence of the inexact Newton iteration is measured by

\[
\|R_i(X_i^{[l]})\| = \max_{j \in 1, \cdots, s} \max \{\|\tilde{\psi}_j\|, \|\tilde{\Phi}_j\|, \|g_j(X_i^{[l]})\|\} < \tau \quad (17)
\]

where \( \tilde{\psi} \) is the absolute tolerance and \( rtol \) is the relative tolerance. Steps are accepted if this measure of the residual is smaller than \( \tau \approx 0.1 \). In case of divergence or slow convergence, the iterations are terminated, the step size, \( h \), is decreased and the Jacobian of the iteration matrix is re-evaluated and factorized. As explained in e.g. Völcker et al. (2010b) and Capozi and Jørgensen (2012), the step size controller adjust the temporal step sizes such that the error estimate satisfies a norm similar to the norm used in (17).

4. CONTINUOUS ADJOINT METHOD

Gradient based methods such as sequential quadratic programming (SQP) methods for solution of (8) require the gradient of the objective function (7) with respect to the control vector parameters, i.e. \( \partial \psi / \partial u_k \) for \( k = 0, 1, \ldots, N-1 \). In this section, we describe a continuous-time adjoint based method for computation of these gradients.

**Proposition 1.** (Gradients based on Continuous Adjoints). Consider the function \( \psi = \psi(\{u_k\}_{k=0}^{N-1}; x_0) \) defined by (7).

The gradients, \( \partial \psi / \partial u_k \), may be computed as

\[
\frac{\partial \psi}{\partial u_k} = \int_{t_k}^{t_{k+1}} \left( \frac{\partial \tilde{\Phi}}{\partial u_k} - \lambda^T \frac{\partial f}{\partial x} \right) \, dt \quad k = 0, 1, \ldots, N-1
\]

in which \( x(t) \) is computed by solution of (1b)-(1c) and \( \lambda(t) \) is computed by solution of the adjoint equations

\[
\frac{d \lambda^T}{dt} \frac{\partial f}{\partial x} + \lambda^T \frac{\partial f}{\partial x} - \frac{\partial \Phi}{\partial x} = 0 \quad (18a)
\]

\[
\frac{\partial \Phi}{\partial x}(x(t_k)) + \lambda^T(t_k) \frac{\partial g}{\partial x}(x(t_k)) = 0 \quad (18b)
\]

**Proof.** See Appendix A.

**Remark 2.** (Computation using ESDIRK). \( x(t) \) is computed using the ESDIRK method applied to (1b)-(1c) and integration forwards. This solution is stored. The same ESDIRK method is applied for computation of \( \lambda(t) \) by solving (19) integrating backwards in time.

**Remark 3.** (Gradients Computed by Continuous Adjoints). The gradients computed using the continuous adjoints are not the exact gradients, \( \partial \psi / \partial u_k \), when the involved differential equations and integrals are computed by discretization using the ESDIRK method. However, they can be made sufficiently precise for the optimizer such that they do not affect the convergence (Nadarajah and Jamieson, 2007). The advantage of the continuous adjoint equations (19) is that they can be solved faster than the adjoint equations for the discretized system (10)-(12).

5. PRODUCTION OPTIMIZATION FOR A CONVENTIONAL OIL FIELD

In this section, we apply our algorithm for constrained optimal control problems to production optimization in a
conventional horizontal oil field that can be modeled as two phase flow in a porous medium (Chen, 2007; Völcker et al., 2009). The reservoir size is 450 m x 450 m x 10 m. By spatial discretization this reservoir is divided into 25 x 25 x 1 grid blocks. The configuration of injection wells and producers as well as the permeability field is illustrated in Fig. 1. As indicated in Fig. 1, the four injectors are located in the corners of the field, while the single producer is located in the center of the field. The specification of the two phase oil model consists of the injector (i ∈ I) and the producer (i ∈ P) location, the permeability parameters indicated in Fig. 1, and the parameters listed in Table 1. The initial reservoir pressure is 400 atm everywhere in the reservoir. The initial water saturation is 0.1 everywhere in the reservoir. This implies that initially the reservoir has a uniform oil saturation of 0.9.

The inherent discounted stage cost function (see (4))

\[ \Phi(t) = \Phi(x(t), u(t)) = \frac{1}{(1 + b)^{t/365}} \sum_{j \in P} (r_0(1 - f_w) - f_w r_w) q_j(t) \]  

(20)

accounts for the value of the oil produced minus the processing cost of the produced water. In this cost function, we have neglected the processing cost of injected water as well as the effect of pressure on injecting water. b is the discount factor. The fractional flow of water, \( f_w = \lambda_w / (\lambda_w + \lambda_o) \), indicates the relative flow of water. \( \lambda_w = \rho_w k k_{rw} / \mu_w \)

and \( \lambda_o = \rho_o k k_{ro} / \mu_o \) are the water and oil mobilities, respectively. In the problems considered, we do not have any cost-to-go terms, i.e. \( \Phi(t) = 0 \). Neither do we have any path constraints (2). Therefore, maximizing the net present value of the oil field corresponds to minimization of

\[ J(t_b) = -NPV(t_b) = \int_{t_0}^{t_b} \Phi(x(t), u(t)) dt \]  

(21)

with \( \Phi(x(t), u(t)) = \Phi(x(t), u(t)). \) The optimizer maximizes the net present value by manipulating the injection of water at the injectors and by manipulation of the total fluid production (oil and water) at the producers. Hence, the manipulated variable at time period \( k \in \mathcal{N} \) is \( q_{w,i,k} \) which indicates the relative flow of water. \( \lambda \)

(8d). By the total injection constraint, the optimization problem because the water injected at injectors and the production at the producers must both be positive and because each production facility has a maximum flow capacity. In the considered problem we have

\[ 0 \leq q_{w,i,k} \leq q_{\text{max}} \quad i \in I, \quad k \in \mathcal{N} \]  

(22a)

\[ 0 \leq q_{i,k} \leq q_{\text{max}} \quad i \in P, \quad k \in \mathcal{N} \]  

(22b)

The maximum flow capacity, \( q_{\text{max}} \), is the same for all injectors and producers in this case study. The rate of change for all injectors and producers are \( q_{i,k} - q_{i,k-1} \leq 5 \) for \( i \in I \cup P \) and \( k \in \mathcal{N} \). Since the injection of oil is zero, \( \rho_{o,i,k} = 0 \) for \( i \in I \), we get \( q_{w,i,k} - q_{w,i,k-1} \leq 5 \) for \( i \in I \) and \( k \in \mathcal{N} \). This leads to the rate of movement constraints (8c). In addition we use a voidage replacement constraint (Brouwer and Jansen, 2004; Jansen, 2011)

\[ \sum_{i \in I} q_{i,k} = \sum_{i \in I} q_{w,i,k} = \sum_{i \in P} q_{i,k} \quad k \in \mathcal{N} \]  

(23)

and enforce a constant total injection, \( \sum_{i \in I} q_{w,i,k} = q_{\text{max}} \) for \( k \in \mathcal{N} \). This translates into constraints of the type (8d). By the total injection constraint, the optimization problem reduces to a problem of redistributing the flows among the injectors.

The prediction and control horizon is \( t_h = 4270 \) days and the sampling period is \( T_s = 35 \). Hence the prediction and control horizon corresponds to \( N = 122 \) periods. With a total injection at each time period of \( q_{\text{max}} = 100 \) m³/day, these specifications corresponds to injection of 1.05 pore volume during operation of the reservoir. The predictor horizon is optimal in the reference case for a total injection of 100 m³/day.

The optimal water injection rates computed by solution of the constrained optimal control problem (1) for different discount factors, \( b \), are illustrated in Fig. 2. In addition, a base case with constant and equal water injection rates is illustrated. It is evident that the optimal injection rates are very sensitive to the discount factor, \( b \). The corresponding cumulative oil and water production are plotted in Fig. 3. Independent of the discount factor value, the optimized strategies produce more oil than the base case. For the high discount factor case, \( b = 0.12 \), less oil is recovered than in

---

**Table 1. Parameters for the two phase model and the discounted state cost function (20).**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \phi )</td>
<td>Porosity</td>
<td>0.2</td>
<td>-</td>
</tr>
<tr>
<td>( c_r )</td>
<td>Rock compressibility</td>
<td>0</td>
<td>Pa⁻¹</td>
</tr>
<tr>
<td>( \rho_o )</td>
<td>Oil density (400 atm)</td>
<td>800</td>
<td>kg/m³</td>
</tr>
<tr>
<td>( \rho_w )</td>
<td>Water density (400 atm)</td>
<td>1000</td>
<td>kg/m³</td>
</tr>
<tr>
<td>( c_o )</td>
<td>Oil compressibility</td>
<td>10⁻⁵</td>
<td>1/atm</td>
</tr>
<tr>
<td>( c_w )</td>
<td>Water compressibility</td>
<td>10⁻³</td>
<td>1/atm</td>
</tr>
<tr>
<td>( \mu_o )</td>
<td>Dynamic oil viscosity</td>
<td>1</td>
<td>1/atm</td>
</tr>
<tr>
<td>( \mu_w )</td>
<td>Dynamic water viscosity</td>
<td>1</td>
<td>1/atm</td>
</tr>
</tbody>
</table>

**Log10(K) (Darcy):**

-1.5
-1
-0.5
0
0.5
1

![Permeability Field](image_url)

**Fig. 1.** The permeability field and the location of wells. A circle indicates the location of an injector and a cross indicates the location of a producer.
In this paper, we solve constrained optimal control problems using a single shooting method based on a quasi-Newton implementation of Powell’s sequential quadratic programming (SQP) algorithm. The system of differential equations are formulated in a novel way to ensure mass conservation and the resulting initial value problem (1c) is solved with tailored ESDIRK integration methods. We also introduce a high order continuous adjoint system for efficient computation of the gradients. The algorithm is implemented in Matlab.

6. CONCLUSIONS
The resulting algorithm is tested on a production optimization problem for an oil reservoir with two phase flow. For all cases considered, the dynamic optimization increase the net present value of the oil field and give increased oil production. However, the optimal injection rates are very sensitive to the discount factor.

REFERENCES


Appendix A. PROOF OF PROPOSITION 1.

The idea in the proof stems from Cao et al. (2002). Define

\[ G(x, \dot{x}, u) = \frac{d}{dt} g(x(t)) - f(x(t), u(t)) \]  

and introducing the Lagrange multiplier, \( \lambda(t) \), to define the augmented objective function as

\[ J_A = J + \int_{t_k}^{t_{k+1}} \lambda(T(t)) G(x, \dot{x}, u) dt \]  

(A.2a)

(A.1) implies that the derivative of the augmented objective function, \( J_A \), can be expressed as

\[ \frac{dJ}{du_k} = \frac{dJ}{du} = \frac{\partial \Phi}{\partial x} \frac{\partial x}{\partial u_k} \bigg|_{t_k} + \int_{t_k}^{t_{k+1}} \left( \frac{\partial \Phi}{\partial u} \frac{\partial u(t)}{\partial u_k} + \frac{\partial \Phi}{\partial \dot{x}} \frac{\partial \dot{x}(t)}{\partial u_k} \right) dt + \int_{t_k}^{t_{k+1}} \lambda(T(t)) \left( \frac{\partial G}{\partial u} \frac{\partial u(t)}{\partial u_k} + \frac{\partial G}{\partial \dot{x}} \frac{\partial \dot{x}(t)}{\partial u_k} \right) dt \]  

(A.3)

Integrating by part

\[ \int_{t_k}^{t_{k+1}} \lambda(T(t)) \frac{\partial G}{\partial \dot{x}} \frac{\partial \dot{x}}{\partial u_k} dt = \left[ \lambda \frac{\partial G}{\partial \dot{x}} \frac{\partial \dot{x}}{\partial u_k} \right]_{t_k}^{t_{k+1}} - \int_{t_k}^{t_{k+1}} \frac{d}{dt} \left( \lambda \frac{\partial G}{\partial \dot{x}} \frac{\partial \dot{x}}{\partial u_k} \right) dt \]  

(A.4)

and using \( \frac{\partial x}{\partial u_k} (t_k) = 0 \) in our case, we can rearrange equation (A.3) as

\[ \frac{dJ}{du_k} = \int_{t_k}^{t_{k+1}} \left( \frac{\partial \Phi}{\partial u} + \lambda \frac{\partial G}{\partial \dot{x}} \right) dt + \int_{t_k}^{t_{k+1}} \left( \frac{\partial \Phi}{\partial \dot{x}} + \lambda \frac{\partial G}{\partial \dot{x}} \right) \frac{\partial \dot{x}}{\partial u_k} dt \]  

(A.5)

This expression gives the derivative \( dJ/du_k \) for any value (not just the optimal one) of \( \lambda(t) \). We choose \( \lambda(t) \) such that it satisfies

\[ \frac{\partial \Phi}{\partial \dot{x}} + \lambda \frac{\partial G}{\partial \dot{x}} = 0 \]  

(A.6a)

\[ \frac{\partial \Phi}{\partial \dot{x}} + \lambda \frac{\partial G}{\partial \dot{x}} \bigg|_{t_k}^{t_{k+1}} = 0 \]  

(A.6b)

and gives a simple expression for evaluation of \( dJ/du_k \)

\[ \frac{\partial J}{\partial u_k} = \frac{dJ}{du} + \int_{t_k}^{t_{k+1}} \left( \frac{\partial \Phi}{\partial u} - \lambda \frac{\partial f}{\partial u} \right) dt \]  

(A.6)

(A.1) implies

\[ \frac{\partial G}{\partial x} = \frac{d}{dt} \left( \frac{\partial g}{\partial x} \right) - \frac{\partial f}{\partial x} \]  

(A.9a)

\[ \frac{\partial G}{\partial \dot{x}} = \frac{\partial g}{\partial \dot{x}} \]  

(A.9b)