AN ADAPTIVE FINITE ELEMENT METHOD FOR NONLINEAR OPTIMAL CONTROL PROBLEMS

KARIN KRAFT AND STIG LARSSON

ABSTRACT. Lagrange's method in the calculus of variations is applied to a nonlinear optimal control problem. The optimality conditions are discretized by a finite element method. The methodology of dual weighted residuals is used to derive an *a posteriori* error estimate. This is combined with Newton iterations in an adaptive multilevel method, which is implemented and tested on model problems.

1. INTRODUCTION

We consider optimal control problems with a nonlinear system of ordinary differential equations as state equations. The system of differential equations has boundary values at the initial and final times, and the final time is fixed. The goal functional is nonlinear, although in practice it is often quadratic. The optimality conditions are derived by Lagrange's method in the calculus of variations and the resulting differential/algebraic equations are discretized by a finite element method. The purpose is to investigate the potential of adaptive finite element methods for the numerical solution of this type of problem.

The finite element method has been widely used for spatial discretization of optimal control problems for partial differential equations [1, 8, 11, 12]. The use of finite element methods for temporal discretization is not as common but it has been used in [2, 3, 4, 5, 6].

The Lagrange framework requires the solution of a linearized adjoint system of the same size as the equations of state. In the previous work [3, 4, 6] all equations are merged into one large system, which is solved by an adaptive finite element method. The theoretical basis is the standard duality argument for proving *a posteriori* error estimates. This requires the solution of the linearized adjoint of the new system, thereby doubling the number of variables.

The dual weighted residuals methodology [1] for *a posteriori* error analysis is formulated within the Lagrange framework and is therefore well suited

1991 *Mathematics Subject Classification.* 65L60, 49K15.

Key words and phrases. adaptive finite element, a posteriori, error estimate, nonlinear, optimal control, multilevel, Newton method, dual weighted residuals.

Research supported by the Swedish Research Council (VR) and by the Swedish Foundation for Strategic Research (SSF) through GMMC, the Gothenburg Mathematical Modelling Centre.

for optimal control problems. No additional adjoint problem is introduced. The procedure provides a representation formula for the error in the goal functional. The formula can be expanded into an error estimator, which is an elementwise sum of dual weighted residuals. Each of these consist of a residual from the state equation multiplied by a weight computed from the adjoint solution and a residual from the adjoint equation multiplied by a weight from the state equation.

In our previous work [5] we developed this approach for optimal control problems with quadratic goal function and linear state equations. In the present work we extend this to fully nonlinear problems. Discretization by the finite element method results in an nonlinear algebraic problem, which is solved by Newton's method. The error estimator includes terms representing the approximate solution of the algebraic equations. The Newton iteration is combined with the adaptive refinement iteration into a multilevel adaptive method, which is implemented and tested on model problems.

The outline of the article is: In Section 2 the mathematical setting of the optimal control problem is done, the Lagrange framework is presented, the optimality conditions are derived and the Newton method for the solution of the nonlinear equations is described. In Section 3 the discretization is described and in Section 4 a representation formula for the error is proved and the computation of the resulting error estimate is discussed. The implementation of the multilevel adaptive finite element solver is described in Section 5. The last section contains two numerical examples and a discussion of the performance of the solver.

2. A Nonlinear optimal control problem

We consider optimal control problems of the form: Determine states $x(t) \in \mathbb{R}^d$ and controls $u(t) \in \mathbb{R}^m$, which

$$
\begin{aligned}
&\text{minimize} && \mathcal{J}(x,u) = l(x(0), x(T)) + \int_0^T L(x(t), u(t))\,\mathrm{d}t, \\
&\text{subject to} && \dot{x}(t) = f(x(t), u(t)), \quad 0 < t < T, \\
& && I_0 x(0) = x_0, \quad I_T x(T) = x_T.
\end{aligned} \tag{2.1}
$$

Here

$$
\begin{aligned}
l &\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}, \\
L &\colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}, \\
f &\colon \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d,
\end{aligned}
$$

are smooth functions and I_0 and I_T are binary diagonal matrices, and $x_0 \in R(I_0)$, $x_T \in R(I_T)$, where $R(A)$ denotes the range of a matrix A.

2.1. Lagrange framework.

In order to set the optimal control problem in the Lagrange framework we introduce some function spaces. Let $\mathcal{C}^k$ denote k times continuously differentiable functions and let H^1 denote functions with square integrable derivative. Further, $\mathcal{C}^1_{\mathrm{PW}}$ denotes piecewise continuously differentiable functions $[0,T] \to \mathbb{R}^d$; more precisely, functions that are $\mathcal{C}^1$ except at a finite number of points in $[0,T]$ and with left and right limits $w(t^-) = \lim_{s\downarrow t} w(s)$, $w(t^+) = \lim_{s\uparrow t} w(s)$ for all points $t \in [0,T]$.

We introduce the function spaces

$$\begin{aligned}
\mathcal{W} &= \mathbb{R}^d \times \mathcal{C}^1_{\mathrm{PW}}([0,T],\mathbb{R}^d) \times \mathbb{R}^d;\\
\dot{\mathcal{W}} &= R(I-I_0) \times \mathcal{C}^1_{\mathrm{PW}}([0,T],\mathbb{R}^d) \times R(I-I_T)\\
&= \Big\{ w \in \mathcal{W} : I_0 w(0^-) = 0,\ I_T w(T^+) = 0 \Big\};\\
\mathcal{V} &= H^1([0,T],\mathbb{R}^d);\\
\mathcal{U} &= H^1([0,T],\mathbb{R}^m).
\end{aligned}$$

The two factors $\mathbb{R}^d$ in $\mathcal{W}$ are used to accomodate the boundary values $w(0^-)$ and $w(T^+)$. These spaces are linear spaces. For some $\hat{x} \in \mathcal{W}$ such that $I_0\hat{x}(0^-) = x_0$ and $I_T\hat{x}(T^+) = x_T$, we also define the affine space

$$\hat{\mathcal{W}} = \hat{x} + \dot{\mathcal{W}} = \Big\{ w \in \mathcal{W} : w - \hat{x} \in \dot{\mathcal{W}} \Big\}.$$

The weak formulation of the state equation in (2.1) is: Given $u \in \mathcal{U}$ find $x \in \hat{\mathcal{W}}$ such that

$$\mathcal{F}(x,u;\varphi) = 0 \quad \forall \varphi \in \mathcal{V}. \tag{2.2}$$

Here, the functional

$$\mathcal{F} \colon \mathcal{W} \times \mathcal{U} \times \mathcal{V} \to \mathbb{R},$$

is defined by

$$\mathcal{F}(x,u;\varphi) = \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\dot{x} - f(x,u), \varphi)\,\mathrm{d}t + \sum_{n=0}^{N} ([x]_n, \varphi(t_n)), \tag{2.3}$$

where $(\cdot,\cdot)$ denotes the scalar product in $\mathbb{R}^d$, $[x]_n = x(t_n^+) - x(t_n^-)$, and the sum is taken over all points $\{t_n\}_{n=0}^N$ of discontinuity of x, and $t_0 = 0$, $t_N = T$. Although x is expected to be smooth, the functional $\mathcal{F}$ is defined for piecewise differentiable functions $x \in \mathcal{W}$ with $\dot{x}$ written as a weak derivative in order to admit also piecewise smooth finite element functions. We use the notation that functionals depend arbitrarily on the arguments before the semicolon and linearly on the arguments after the semicolon.

2.2. **Necessary conditions for optimality.** In order to derive the optimality conditions, we introduce the Lagrange functional

$$\mathcal{L}(x,u;z) = \mathcal{J}(x,u) + \mathcal{F}(x,u;z), \quad (x,u,z) \in \mathcal{W} \times \mathcal{U} \times \mathcal{V}, \tag{2.4}$$

where z is a Lagrange multiplier. In order to find minima we seek $(x,u,z) \in \hat{\mathcal{W}} \times \mathcal{U} \times \mathcal{V}$ such that

$$\mathcal{L}'(x,u;z,\varphi) = 0 \quad \forall \varphi \in \dot{\mathcal{W}} \times \mathcal{U} \times \mathcal{V}. \tag{2.5}$$

Taking partial derivatives of the Lagrangian in (2.4) gives

$$\mathcal{L}'_x(x,u;z,\varphi_x) = \mathcal{J}'_x(x,u;\varphi_x) + \mathcal{F}'_x(x,u;z,\varphi_x) = 0 \qquad \forall \varphi_x \in \dot{\mathcal{W}}, \tag{2.6a}$$

$$\mathcal{L}'_u(x,u;z,\varphi_u) = \mathcal{J}'_u(x,u;\varphi_u) + \mathcal{F}'_u(x,u;z,\varphi_u) = 0 \qquad \forall \varphi_u \in \mathcal{U}, \tag{2.6b}$$

$$\mathcal{L}'_z(x,u;z,\varphi_z) = \mathcal{F}(x,u;\varphi_z) = 0 \qquad \forall \varphi_z \in \mathcal{V}. \tag{2.6c}$$

Expanding (2.6a)–(2.6c) and using integration by parts in (2.6a) gives

$$\begin{aligned}&\int_0^T (\varphi_x, -\dot{z} - f'_x(x,u)^* z + L'_x(x,u))\,\mathrm{d}t \\ &\quad + (\varphi^+_{x,N}, z_N + l'_2(x^-_0, x^+_N)) \\ &\quad + (\varphi^-_{x,0}, -z_0 + l'_1(x^-_0, x^+_N)) = 0 \qquad \forall \varphi_x \in \dot{\mathcal{W}},\end{aligned} \tag{2.7a}$$

$$\int_0^T (\varphi_u, L'_u(x,u) - f'_u(x,u)^* z)\,\mathrm{d}t = 0 \qquad \forall \varphi_u \in \mathcal{U}, \tag{2.7b}$$

$$\sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} (\dot{x} - f(x,u), \varphi_z)\,\mathrm{d}t + \sum_{n=0}^{N} ([x]_n, \varphi_{z,n}) = 0 \qquad \forall \varphi_z \in \mathcal{V}, \tag{2.7c}$$

$$I_0 x^-_0 = x_0, \quad I_T x^+_T = x_T, \tag{2.7d}$$

where $x^\pm_n = x(t^\pm_n)$, $\varphi^\pm_{x,n} = \varphi_x(t^\pm_n)$, $\varphi_{z,n} = \varphi_z(t_n)$, and l'_1 and l'_2 denote derivative with respect to the first and the second variable. .

Here we have differentiated the term

$$f(x,u;z) := (f(x,u), z),$$

and identified the derivatives with matrices $f'_x(x,u)$, $f'_u(x,u)$ by means of the Riesz representation theorem:

$$\begin{aligned} f'_x(x,u;z,\varphi_x) &= (f'_x(x,u)\varphi_x, z) = (\varphi_x, f'_x(x,u)^* z), \\ f'_u(x,u;z,\varphi_u) &= (f'_u(x,u)\varphi_u, z) = (\varphi_u, f'_u(x,u)^* z). \end{aligned} \tag{2.8}$$

Similarly, we defined the vectors $L'_x(x,u)$, $l'_i(x^-_0, x^+_N)$, $i = 1,2$.

2.3. **Newton's method.** We use Newton's method to solve the nonlinear equations in (2.7). Given an approximate solution (x,u,z) it yields a new approximate solution $(\hat{x},\hat{u},\hat{z})$ by

$$(\hat{x},\hat{u},\hat{z}) = (x,u,z) + \alpha(\delta_x,\delta_u,\delta_z), \tag{2.9}$$

where $\alpha \in \mathbb{R}$ is a parameter and the increment $\delta = (\delta_x,\delta_u,\delta_z) \in \dot{\mathcal{W}} \times \mathcal{U} \times \mathcal{V}$ is the solution of

$$\mathcal{L}''(x,u;z,\varphi,\delta) = -\mathcal{L}'(x,u;z,\varphi) \quad \forall \varphi \in \dot{\mathcal{W}} \times \mathcal{U} \times \mathcal{V}. \tag{2.10}$$

To clarify the following equations we identify matrices $f''_{xx}(x,u;z)$, $f''_{xu}(x,u;z)$, $f''_{uu}(x,u;z)$ by further differentiation in (2.8):

$$\begin{aligned} f''_{xx}(x,u;z,\varphi_x,\psi_x) &= (f''_{xx}(x,u;z)\varphi_x,\psi_x) = (\varphi_x, f''_{xx}(x,u;z)\psi_x), \\ f''_{xu}(x,u;z,\varphi_x,\varphi_u) &= (f''_{xu}(x,u;z)\varphi_x,\varphi_u) = (\varphi_x, f''_{xu}(x,u;z)^*\varphi_u), \\ f''_{uu}(x,u;z,\varphi_u,\psi_u) &= (f''_{uu}(x,u;z)\varphi_u,\psi_u) = (\varphi_u, f''_{uu}(x,u;z)\psi_u), \end{aligned}$$

and similarly for derivatives of L and l.

Writing (2.10) explicitly yields the following equations: Find $(\delta_x, \delta_u, \delta_z) \in \dot{\mathcal{W}} \times \mathcal{U} \times \mathcal{V}$ satisfying

$$\begin{aligned}
&\int_0^T (\varphi_x, -\dot{\delta}_z - f_x'(x,u)^* \delta_z + (L_{xx}''(x,u) - f_{xx}''(x,u;z))\delta_x \\
&\qquad + (L_{xu}''(x,u) - f_{xu}''(x,u;z)^*)\delta_u)\,\mathrm{d}t \\
&\quad + (\varphi_{x,N}^+, \delta_{z,N} + l_{21}''(x_0^-, x_N^+)\delta_{x,0}^- + l_{22}''(x_0^-, x_N^+)\delta_{x,N}^+) \\
&\quad + (\varphi_{x,0}^-, -\delta_{z,0} + l_{11}''(x_0^-, x_N^+)\delta_{x,0}^- + l_{12}''(x_0^-, x_N^+)\delta_{x,N}^+) \\
&\ = -\int_0^T (\varphi_x, -\dot{z} - f_x'(x,u)^* z + L_x'(x,u))\,\mathrm{d}t \\
&\quad - (\varphi_{x,N}^+, z_N + l_2'(x_0^-, x_N^+)) - (\varphi_{x,0}^-, -z_0 + l_1'(x_0^-, x_N^+)) \qquad \forall \varphi_x \in \dot{\mathcal{W}}, \\
&\int_0^T (\varphi_u, (L_{uu}''(x,u) - f_{uu}''(x,u;z))\delta_u \\
&\qquad + (L_{ux}''(x,u) - f_{ux}''(x,u;z)^*)\delta_x - f_u'(x,u)^*\delta_z)\,\mathrm{d}t \\
&\ = -\int_0^T (\varphi_u, L_u'(x,u) - f_u'(x,u)^* z)\,\mathrm{d}t \qquad \forall \varphi_u \in \mathcal{U}, \\
&\sum_{n=1}^N \int_{t_{n-1}}^{t_n} (\dot{\delta}_x - f_x'(x,u)\delta_x - f_u'(x,u)\delta_u, \varphi_z)\,\mathrm{d}t + \sum_{n=0}^N ([\delta_x]_n, \varphi_{z,n}) \\
&\ = -\sum_{n=1}^N \int_{t_{n-1}}^{t_n} (\dot{x} - f(x,u), \varphi_z)\,\mathrm{d}t - \sum_{n=0}^N ([x]_n, \varphi_{z,n}) \qquad \forall \varphi_z \in \mathcal{V}.
\end{aligned}$$

3. Discretization

The optimality conditions and the Newton equations derived in the previous section are discretized by a finite element method and solved.

3.1. A finite element problem.

The equations in (2.6a)–(2.6c) are discretized by a finite element method based on the mesh $0 = t_0 < t_1 < t_2 < \ldots < t_N = T$ with steps $h_n = t_n - t_{n-1}$ and intervals $J_n = (t_{n-1}, t_n)$. With P^k denoting polynomials of degree k, we introduce some function spaces. Let $\mathcal{W}_h$ denote a space of discontinuous piecewise constant vector-valued functions, that is,

$$\begin{aligned}
\mathcal{W}_h &= \mathbb{R}^d \times \Big\{ w : w|_{J_n} \in P^0(J_n, \mathbb{R}^d),\ n = 1, \ldots, N \Big\} \times \mathbb{R}^d, \\
\dot{\mathcal{W}}_h &= R(I - I_0) \times \Big\{ w : w|_{J_n} \in P^0(J_n, \mathbb{R}^d),\ n = 1, \ldots, N \Big\} \times R(I - I_T) \\
&= \Big\{ w \in \mathcal{W}_h : I_0 w_0^- = 0,\ I_T w_N^+ = 0 \Big\}.
\end{aligned}$$

In the definition of $\hat{\mathcal{W}} = \hat{x} + \dot{\mathcal{W}}$ we may choose $\hat{x} \in \mathcal{W}_h$ and define

$$\hat{\mathcal{W}}_h = \hat{x} + \dot{\mathcal{W}}_h.$$

The spaces $\mathcal{V}_h$ and $\mathcal{U}_h$ consist of continuous piecewise linear functions and defined as

$$\mathcal{V}_h = \Big\{ v \in \mathcal{C}([0,T], \mathbb{R}^d) : v|_{J_n} \in P^1(J_n, \mathbb{R}^d) \Big\},$$
$$\mathcal{U}_h = \Big\{ v \in \mathcal{C}([0,T], \mathbb{R}^m) : v|_{J_n} \in P^1(J_n, \mathbb{R}^m) \Big\}.$$

Then we have $\mathcal{W}_h \subset \mathcal{W}$, $\dot{\mathcal{W}}_h \subset \dot{\mathcal{W}}$, $\hat{\mathcal{W}}_h \subset \hat{\mathcal{W}}$, $\mathcal{U}_h \subset \mathcal{U}$, and $\mathcal{V}_h \subset \mathcal{V}$.

The finite element problem now reads: Find $(x_h, u_h, z_h) \in \hat{\mathcal{W}}_h \times \mathcal{U}_h \times \mathcal{V}_h$ such that

$$\mathcal{J}_x'(x_h, u_h; \varphi_{x,h}) + \mathcal{F}_x'(x_h, u_h; z_h, \varphi_{x,h}) = 0 \quad \forall \varphi_{x,h} \in \dot{\mathcal{W}}_h, \tag{3.1a}$$
$$\mathcal{J}_u'(x_h, u_h; \varphi_{u,h}) + \mathcal{F}_u'(x_h, u_h; z_h, \varphi_{u,h}) = 0 \quad \forall \varphi_{u,h} \in \mathcal{U}_h, \tag{3.1b}$$
$$\mathcal{F}(x_h, u_h; \varphi_{z,h}) = 0 \quad \forall \varphi_{z,h} \in \mathcal{V}_h. \tag{3.1c}$$

3.2. **Newton's method.** We solve this nonlinear system by Newton's method. For a given approximate solution $(x_h, u_h, z_h) \in \hat{\mathcal{W}}_h \times \mathcal{U}_h \times \mathcal{V}_h$, find $\delta_h = (\delta_{x,h}, \delta_{u,h}, \delta_{z,h}) \in \dot{\mathcal{W}}_h \times \mathcal{U}_h \times \mathcal{V}_h$ such that

$$\mathcal{L}''(x_h, u_h; z_h, \varphi_h, \delta_h) = -\mathcal{L}'(x_h, u_h; z_h, \varphi_h) \quad \forall \varphi_h \in \dot{\mathcal{W}}_h \times \mathcal{U}_h \times \mathcal{V}_h. \tag{3.2}$$

Then set

$$(\hat{x}_h, \hat{u}_h, \hat{z}_h) = (x_h, u_h, z_h) + \alpha(\delta_{x,h}, \delta_{u,h}, \delta_{z,h}). \tag{3.3}$$

The equation (3.2) has the same form as the corresponding equations in Subsection 2.3. By using standard finite element basis functions we obtain a linear system of equations as follows. The piecewise constant basis functions $\{\phi_n\}_{n=0}^{N+1}$ are defined by $\phi_n(t) = 0$ except for

$$\begin{aligned} \phi_0(t) &= 1, \quad t < 0, \\ \phi_n(t) &= 1, \quad t_{n-1} < t < t_n, \\ \phi_{N+1}(t) &= 1, \quad t > t_N, \end{aligned}$$

and the piecewise linear basis functions $\{\varphi_n\}_{n=0}^{N}$ are defined by

$$\varphi_n(t) = \begin{cases} 0, & \text{if } t \notin J_n \cup J_{n+1}, \\ \frac{t - t_{n-1}}{t_n - t_{n-1}}, & \text{if } t \in J_n, \\ \frac{t - t_{n+1}}{t_n - t_{n+1}}, & \text{if } t \in J_{n+1}. \end{cases}$$

We make the Ansatz

$$\delta_{x,h}(t) = \sum_{i=0}^{N+1} \delta_{x,i} \phi_i(t), \quad \delta_{x,i} \in \mathbb{R}^d,$$
$$\delta_{u,h}(t) = \sum_{i=0}^{N} \delta_{u,i} \varphi_i(t), \quad \delta_{u,i} \in \mathbb{R}^m,$$
$$\delta_{z,h}(t) = \sum_{i=0}^{N} \delta_{z,i} \varphi_i(t), \quad \delta_{z,i} \in \mathbb{R}^d.$$

When inserted into (3.2) this gives rise to a linear symmetric system of equations of the form

$$\begin{bmatrix} A_{xx} & A_{ux}^* & A_{zx}^* \\ A_{ux} & A_{uu} & A_{zu}^* \\ A_{zx} & A_{zu} & 0 \end{bmatrix} \begin{bmatrix} \delta_x \\ \delta_u \\ \delta_z \end{bmatrix} = \begin{bmatrix} F_x \\ F_u \\ F_x \end{bmatrix}.$$

This is of saddle point form provided that

$$\begin{bmatrix} A_{xx} & A_{ux}^* \\ A_{xu} & A_{uu} \end{bmatrix}$$

is positive definite.

4. A posteriori error analysis

4.1. **A representation formula for the error.** The error can be measured in various ways. The dual weighted residuals methodology yields a representation formula for the error in the goal functional $\mathcal{J}$.

Theorem 4.1. *Let* $(x, u, z) \in \hat{\mathcal{W}} \times \mathcal{U} \times \mathcal{V}$ *be a solution to the optimality conditions in* (2.5) *and* $(\hat{x}_h, \hat{u}_h, \hat{z}_h) \in \hat{\mathcal{W}}_h \times \mathcal{U}_h \times \mathcal{V}_h$ *denote an approximate solution of the discrete problem in* (3.1). *Then the error in the objective functional* $\mathcal{J}$ *satisfies*

$$\mathcal{J}(x, u) - \mathcal{J}(\hat{x}_h, \hat{u}_h) = \tfrac{1}{2}\rho_x + \tfrac{1}{2}\rho_u + \tfrac{1}{2}\rho_z + \mathcal{F}(\hat{x}_h, \hat{u}_h, \hat{z}_h) + R, \tag{4.1}$$

with the residuals

$$\begin{aligned} \rho_x &= \mathcal{J}_x'(\hat{x}_h, \hat{u}_h; x - \hat{x}_h) + \mathcal{F}_x'(\hat{x}_h, \hat{u}_h; \hat{z}_h, x - \hat{x}_h), \\ \rho_u &= \mathcal{J}_u'(\hat{x}_h, \hat{u}_h; u - \hat{u}_h) + \mathcal{F}_u'(\hat{x}_h, \hat{u}_h; \hat{z}_h, u - \hat{u}_h), \\ \rho_z &= \mathcal{F}(\hat{x}_h, \hat{u}_h; z - \hat{z}_h), \end{aligned}$$

and the remainder

$$\begin{aligned} R = \tfrac{1}{2} \int_0^1 \Big(& J'''(x_h + s\hat{e}_x, u_h + s\hat{e}_u; \hat{e}, \hat{e}, \hat{e}) \\ & + \mathcal{F}'''(\hat{x}_h + s\hat{e}_x, \hat{u}_h + s\hat{e}_u; \hat{z}_h + s\hat{e}_z, \hat{e}, \hat{e}, \hat{e}) \Big) s(s-1)\,\mathrm{d}s, \end{aligned}$$

where $\hat{e} = (\hat{e}_x, \hat{e}_u, \hat{e}_z) = (x - \hat{x}_h, u - \hat{u}_h, z - \hat{z}_h) \in \dot{\mathcal{W}} \times \mathcal{U} \times \mathcal{V}$.

Proof. The definition of $\mathcal{L}$ in (2.4) and the fact that $\mathcal{F}(x, u; z) = 0$ from (2.6c) gives

$$
\begin{aligned}
&\mathcal{J}(x,u)-\mathcal{J}(\hat{x}_h,\hat{u}_h)\\
&\quad=\mathcal{L}(x,u;z)-\mathcal{F}(x,u;z)-\mathcal{L}(\hat{x}_h,\hat{u}_h;\hat{z}_h)+\mathcal{F}(\hat{x}_h,\hat{u}_h;\hat{z}_h)\\
&\quad=\mathcal{L}(x,u;z)-\mathcal{L}(\hat{x}_h,\hat{u}_h;\hat{z}_h)+\mathcal{F}(\hat{x}_h,\hat{u}_h;\hat{z}_h)\\
&\quad=\int_0^1\mathcal{L}'(\hat{x}_h+s\hat{e}_x,\hat{u}_h+s\hat{e}_u;\hat{z}_h+s\hat{e}_z,\hat{e})\,\mathrm{d}s+\mathcal{F}(\hat{x}_h,\hat{u}_h;\hat{z}_h)\\
&\quad=\tfrac{1}{2}\mathcal{L}'(\hat{x}_h,\hat{u}_h;\hat{z}_h,\hat{e})+\tfrac{1}{2}\mathcal{L}'(x,u;z,\hat{e})+R+\mathcal{F}(\hat{x}_h,\hat{u}_h;\hat{z}_h)\\
&\quad=\tfrac{1}{2}\mathcal{L}'(\hat{x}_h,\hat{u}_h;\hat{z}_h,\hat{e})+\mathcal{F}(\hat{x}_h,\hat{u}_h;\hat{z}_h)+R.
\end{aligned}
\tag{4.2}
$$

We used the trapezoidal rule with remainder R for the integral and $\mathcal{L}'(x,u;z,\hat{e})=0$ by (2.5). The remainder term R is readily computed from

$$\int_0^1 f(s)\,\mathrm{d}s=\tfrac{1}{2}(f(0)+f(1))+\tfrac{1}{2}\int_0^1 f''(s)s(s-1)\,\mathrm{d}s.$$

This completes the proof. □

4.2. **An *a posteriori* error estimate.** In order to develop the error formula into an error estimator we need to take further steps. The first is the following *a posteriori* error estimate.

Corollary 4.2. *We have*

$$|\mathcal{J}(x,u)-\mathcal{J}(\hat{x}_h,\hat{u}_h)|\le\tfrac{1}{2}(\hat{\rho}_x+\hat{\rho}_u+\hat{\rho}_z)+|\mathcal{F}(\hat{x}_h,\hat{u}_h;\hat{z}_h)|+|R|,\tag{4.3}$$

where (with $h_0=h_{N+1}=0$)

$$
\begin{aligned}
\hat{\rho}_x&=\sum_{n=1}^N\int_{J_n}\|x-\hat{x}_h\|\|L_x'(\hat{x}_h,\hat{u}_h)-\dot{\hat{z}}_h-f_x'(\hat{x}_h,\hat{u}_h)^*\hat{z}_h\|\,\mathrm{d}t\\
&\quad+\|x_N-\hat{x}_{h,N}^+\|\|l_2'(\hat{x}_{h,0}^-,\hat{x}_{h,N}^+)+\hat{z}_{h,N}\|\\
&\quad+\|x_0-\hat{x}_{h,0}^-\|\|l_1'(\hat{x}_{h,0}^-,\hat{x}_{h,N}^+)-\hat{z}_{h,0}\|\\
\hat{\rho}_u&=\sum_{n=1}^N\int_{J_n}\|u-\hat{u}_h\|\|L_u'(\hat{x}_h,\hat{u}_h)-f_u'(\hat{x}_h,\hat{u}_h)^*\hat{z}_h)\|\,\mathrm{d}t\\
\hat{\rho}_z&=\sum_{n=1}^N\Big|\int_{J_n}(\dot{\hat{x}}_h-f(\hat{x}_h,\hat{u}_h),z-\hat{z}_h)\,\mathrm{d}t\\
&\qquad+\frac{h_n}{h_{n-1}+h_n}([\hat{x}_h]_{n-1},z_{n-1}-\hat{z}_{h,n-1})\\
&\qquad+\frac{h_n}{h_n+h_{n+1}}([\hat{x}_h]_n,z_n-\hat{z}_{h,n})\Big|.
\end{aligned}
$$

Proof. The result follows directly from applying the triangle inequality in (4.1) using the expressions in (2.7) after a symmetric distribution of the jump terms over the mesh intervals according to a calculation of the form

$$\sum_{n=0}^{N} a_n = \sum_{n=0}^{N} \frac{h_n + h_{n+1}}{h_n + h_{n+1}} a_n = \sum_{n=1}^{N} \frac{h_n}{h_n + h_{n+1}} a_n + \sum_{n=0}^{N-1} \frac{h_{n+1}}{h_n + h_{n+1}} a_n$$
$$= \sum_{n=1}^{N} \Big(\frac{h_n}{h_n + h_{n+1}} a_n + \frac{h_n}{h_{n-1} + h_n} a_{n-1} \Big).$$

□

The functions x, u, z that appear in the error estimate are not computable. We compute them approximately by solving the necessary conditions for optimality on a much finer mesh to obtain x_{fine}, u_{fine}, z_{fine}. Then we replace x by x_{fine} and so on in the error estimate. We also replace $\hat{x}_h$ by x_h and so on. The remainder is formally cubic in $\hat{e}$ and is neglected.

In this way we obtain an error estimator in the form of an elementwise sum of dual weighted residuals. So, for example, we see that $\hat{\rho}_x$ is a sum of residuals from the adjoint equation weighted by $\|x - \hat{x}_h\|$ from the state equation, and $\hat{\rho}_z$ contains residuals from the state equation with weights from the adjoint equation. The term $|\mathcal{F}(\hat{x}_h, \hat{u}_h; \hat{z}_h)|$ measures how well the discretized state equation is satisfied and is computed similarly to $\hat{\rho}_z$.

The terms $\hat{\rho}_x$, $\hat{\rho}_u$, and $\hat{\rho}_z$ are associated with the discretization error, while the "algebraic residual" $|\mathcal{F}(\hat{x}_h, \hat{u}_h; \hat{z}_h)|$ is related to the error in the nonlinear equations solver, that is, the Newton iteration.

5. Description of the solver

In the following section the implementation of the solver is described.

5.1. **The Newton solver.** Solving the linear system (3.3) yields a search direction for one Newton iteration. In order to decide how far to go in this direction, a simple line search is performed. We compute $(x_h, u_h, z_h)_{\text{new}}$ for various $\alpha \in [0, 1]$ according to

$$(x_h, u_h, z_h)_{\text{new}} = (x_h, u_h, z_h)_{\text{old}} + \alpha(\delta_{x,h}, \delta_{u,h}, \delta_{z,h}), \tag{5.1}$$

and the α that gives the smallest right hand side of (3.2) is chosen.

5.2. **The adaptive solver.** An adaptive finite element solver based on the error estimate in Corollary (4.2) has been implemented. After solving the problem on a coarse mesh, the discretization part of the error estimate is computed. The intervals that give the largest contribution to the total error are refined. The procedure is iterated until the tolerance of the discretization error is reached.

5.3. **A multilevel algorithm.** Following [7] we combine the Newton loop and the refinement iteration into an adaptive multilevel algorithm.

The algorithm generates a sequence $(\hat{\mathcal{W}}_h \times \mathcal{U}_h \times \mathcal{V}_h)_0 \subset (\hat{\mathcal{W}}_h \times \mathcal{U}_h \times \mathcal{V}_h)_1 \subset \cdots \subset (\hat{\mathcal{W}}_h \times \mathcal{U}_h \times \mathcal{V}_h)_N$ of finite element spaces based on adaptively refined meshes, and a sequence of approximate solutions $(\hat{x}_h, \hat{u}_h, \hat{z}_h)_n \in (\hat{\mathcal{W}}_h \times \mathcal{U}_h \times \mathcal{V}_h)_n$. A solution $(\hat{x}_h, \hat{u}_h, \hat{z}_h)_0 \in (\hat{\mathcal{W}}_h \times \mathcal{U}_h \times \mathcal{V}_h)_0$ is computed on the coarsest mesh. This solution can either be achieved by iterating a fixed number of times or until the algebraic solution has reached a given tolerance. The coarse mesh is then refined using the adaptive algorithm in Subsection 5.2. The solution on the coarse mesh is extrapolated to the refined mesh and used as a starting guess for the Newton iteration on the refined mesh. This procedure is iterated until the total error, that is, the discretization plus algebraic error, has reached the desired tolerance. In practice, only a few Newton iterations on each mesh seem to be sufficient.

6. Numerical examples

In this section we present numerical examples which have been solved with the method derived in previous sections.

6.1. **A hyper-sensitive optimal control problem.** The following example is taken from [9]:

$$\text{Minimize} \qquad \int_0^T (x(t)^2 + u(t)^2)\,\mathrm{d}t$$

subject to

$$\begin{aligned} &\dot{x}(t) = -x^3(t) + u(t), \quad 0 < t < T, \\ &x(0) = 1, \quad x(T) = 1. \end{aligned}$$

The example is solved with $T = 25$ and the error tolerance 10^{-3}. The state and control can be found in Figure 6.1. The adaptive multilevel algorithm starts on a coarse mesh with 10 nodes and two Newton iterations are done on each mesh, except for the last, where the solver is iterated until tolerance of 10^{-12}. The adaptively refined mesh is in Figure 6.2. We can see that many nodes are inserted close to the boundaries while only a few nodes are needed in the middle of the time interval. In Figure 6.3 the performance of the adaptive solver is compared to the same solver using uniform refinement. With an adaptive refinement the number of nodes needed to reach a certain precision is substantially lower. The solution has been validated against PROPT [10] with very good agreement.

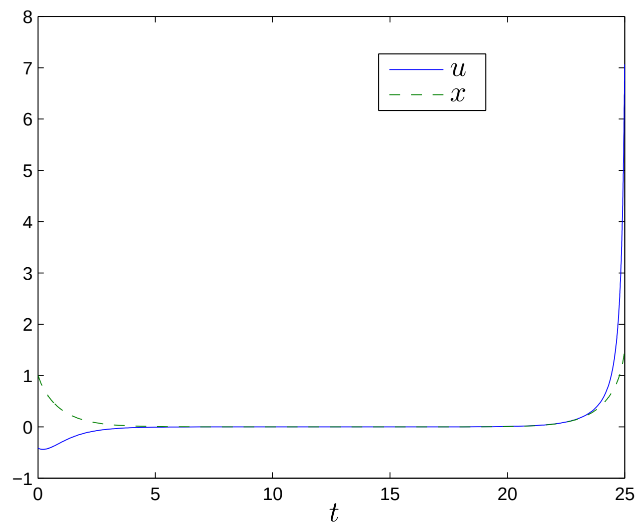

FIGURE 6.1. The optimal state and control for the hypersensitive optimal control problem.

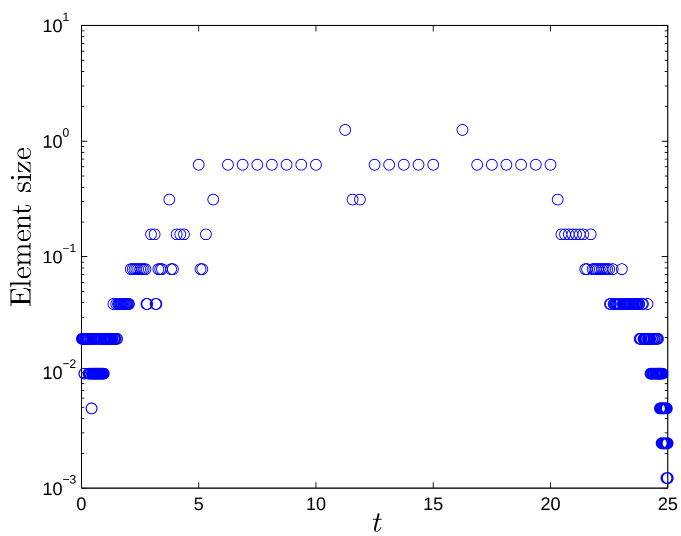

FIGURE 6.2. The adaptively refined mesh. The adaptive solver has inserted nodes at the ends of the interval.

6.2. **Rayleigh.** The following example is taken from the manual of PROPT [10].

$$\text{Minimize} \qquad \int_0^T (x_1(t)^2 + u(t)^2)\,\mathrm{d}t$$

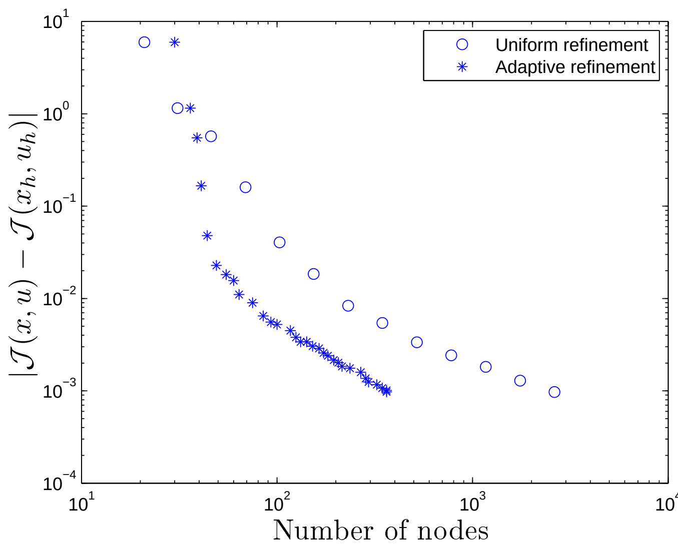

FIGURE 6.3. The error on an adaptively refined mesh compared to the error on a uniformly refined mesh.

subject to

$$\begin{bmatrix}\dot{x}_1(t)\\ \dot{x}_2(t)\end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1 + 1.4x_2 - 0.14x_2^3 + 4u\end{bmatrix}, \quad 0 < t < T,$$
$$\begin{bmatrix}x_1(0)\\ x_2(0)\end{bmatrix} = \begin{bmatrix}-5\\ -5\end{bmatrix}, \quad \begin{bmatrix}x_1(T)\\ x_2(T)\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}.$$

The result can be found in Figures 6.4–6.6. In Figure 6.4 the optimal states, computed with the adaptive finite element solver, are plotted together with the states from Propt. As we can see the solutions coincide. The optimal control is plotted in Figure 6.5. The refined mesh can be found in Figure 6.6. In this example the initial mesh consists of five nodes and the tolerance of the total error is 10^{-3}.

REFERENCES

1. R. Becker and R. Rannacher, *An optimal control approach to a posteriori error estimation in finite element methods*, Acta Numer. **10** (2001), 1–102.
2. A. Eriksson and A. Nordmark, *Temporal finite element formulation of optimal control in mechanisms*, Comput. Methods Appl. Mech. Engrg. **199** (2010), 1783–1792.
3. D. Estep, D. H. Hodges, and M. Warner, *Computational error estimation and adaptive error control for a finite element solution of launch vehicle trajectory problems*, SIAM J. Sci. Comput. **21** (1999), 1609–1631 (electronic).

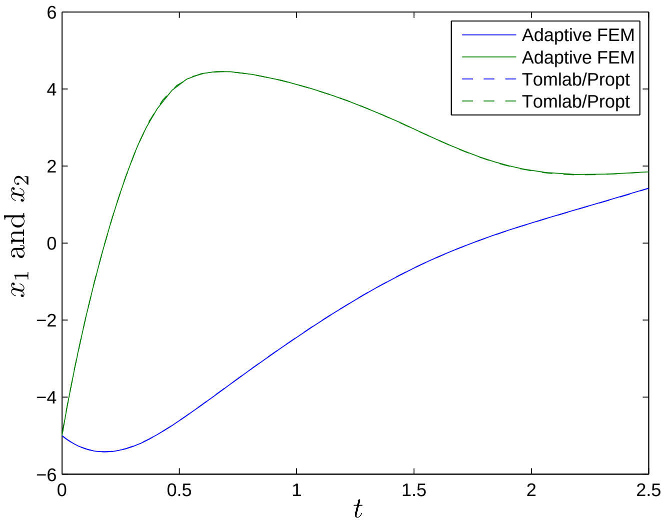

FIGURE 6.4. The optimal states for the Rayleigh problem.

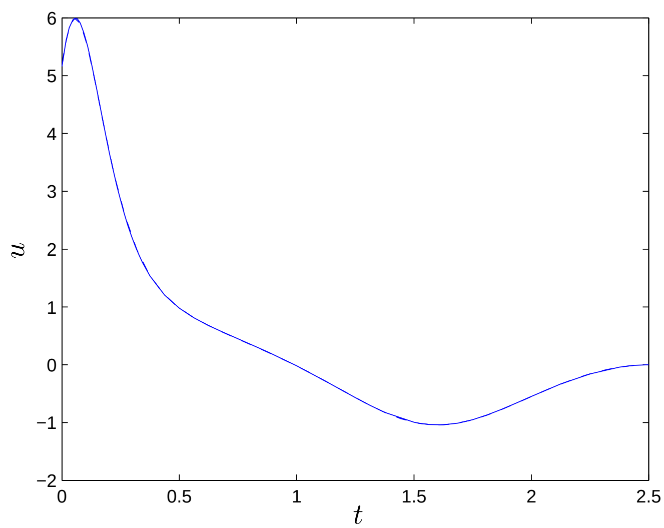

FIGURE 6.5. The optimal control for the Rayleigh problem.

4. D. J. Estep, D. H. Hodges, and M. Warner, *The solution of a launch vehicle trajectory problem by an adaptive finite-element method*, Comput. Methods Appl. Mech. Engrg. (2001), 4677–4690.
5. K. Kraft and S. Larsson, *The dual weighted residuals approach to optimal control of ordinary differential equations*, BIT **50** (2010), 587–607.

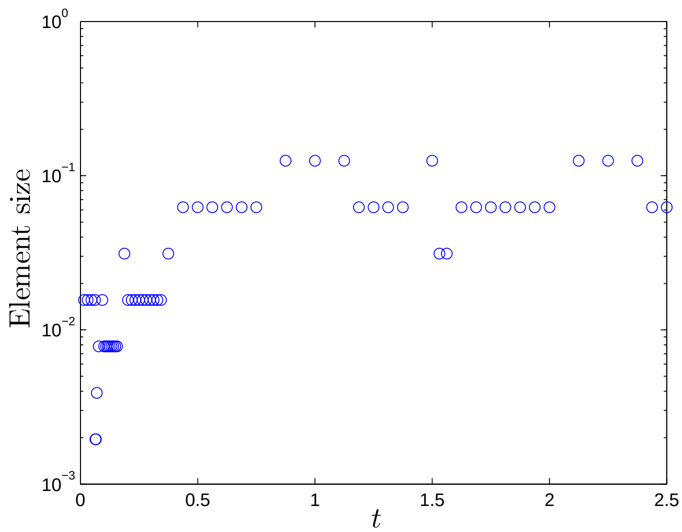

FIGURE 6.6. The adaptively refined mesh.

6. K. Kraft, S. Larsson, and M. Lidberg, *Using an adaptive FEM to determine the optimal control of a vehicle during a collision avoidance manoeuvre*, Proceedings of the 48th Scandinavian Conference on Simulation and Modeling (SIMS2007), Linköping University Electronic Press, 2007, `http://www.ep.liu.se/ecp/027/`.
7. M. G. Larson and A. J. Niklasson, *Adaptive multilevel finite element approximations of semilinear elliptic boundary value problems*, Numer. Math. **84** (1999), 249–274.
8. D. Meidner and B. Vexler, *Adaptive space-time finite element methods for parabolic optimization problems*, SIAM J. Control Optim. **46** (2007), 116–142.
9. A. V. Rao and K. D. Mease, *An eigenvector approximate dichotomic basis method for solving hyper-sensitive optimal control*, Optim. Control Appl. Meth. **21** (2000), 1–19.
10. P. Rutquist and M. M. Edvall, *PROPT-MATLAB Optimal Control Software*, `http://tomdyn.com/`.
11. B. Vexler and W. Wollner, *Adaptive finite elements for elliptic optimization problems with control constraints*, SIAM J. Control Optim. **47** (2008), 509–534.
12. R. Winther, *Initial value methods for parabolic control problems*, Math. Comp. **34** (1980), 115–125.

DEPARTMENT OF MATHEMATICAL SCIENCES, CHALMERS UNIVERSITY OF TECHNOLOGY AND UNIVERSITY OF GOTHENBURG, SE–412 96 GOTHENBURG, SWEDEN
E-mail address: `karin.kraft@chalmers.se`

DEPARTMENT OF MATHEMATICAL SCIENCES, CHALMERS UNIVERSITY OF TECHNOLOGY AND UNIVERSITY OF GOTHENBURG, SE–412 96 GOTHENBURG, SWEDEN
E-mail address: `stig@chalmers.se`