

# Chapter 1

## Constraint Splitting and Projection Methods for Optimal Control of Double Integrator

Heinz H. Bauschke and Regina S. Burachik and C. Yalçın Kaya

**Abstract** We consider the minimum-energy control of a car, which is modelled as a point mass sliding on the ground in a fixed direction, and so it can be mathematically described as the double integrator. The control variable, representing the acceleration or the deceleration, is constrained by simple bounds from above and below. Despite the simplicity of the problem, it is not possible to find an analytical solution to it because of the constrained control variable. To find a numerical solution to this problem we apply three different projection-type methods: (i) Dykstra's algorithm, (ii) the Douglas–Rachford (DR) method and (iii) the Aragón Artacho–Campoy (AAC) algorithm. To the knowledge of the authors, these kinds of (projection) methods have not previously been applied to continuous-time optimal control problems, which are infinite-dimensional optimization problems. The problem we study in this article is posed in infinite-dimensional Hilbert spaces. Behaviour of the DR and AAC algorithms are explored via numerical experiments with respect to their parameters. An error analysis is also carried out numerically for a particular instance of the problem for each of the algorithms.

**Key words:** optimal control, Dykstra projection method, Douglas-Rachford method, Aragón Artacho–Campoy algorithm, linear quadratic optimal control, control constraints, numerical methods

**AMS 2010 Subject Classification:** 49M27, 65K10, 90C20

---

Heinz H. Bauschke  
Department of Mathematics, University of British Columbia, Kelowna, B.C. V1V 1V7, Canada,  
e-mail: heinz.bauschke@ubc.ca

Regina S. Burachik  
School of Information Technology and Mathematical Sciences, University of South Australia,  
Mawson Lakes, S.A. 5095, Australia, e-mail: regina.burachik@unisa.edu.au

C. Yalçın Kaya  
School of Information Technology and Mathematical Sciences, University of South Australia,  
Mawson Lakes, S.A. 5095, Australia, e-mail: yalcin.kaya@unisa.edu.au

## 1.1 Introduction

In this paper, we provide (to the best of our knowledge also first) application of various best approximation algorithms to solve a continuous-time optimal control problem. Operator splitting methods were applied previously to discrete-time optimal control problems [19, 26], which are finite-dimensional problems. In [26], for example, the state difference equations comprise the constraint  $\mathcal{A}$ , and the box constraints on the state and control variables comprise  $\mathcal{B}$ . The condition of belonging to the sets  $\mathcal{A}$  and  $\mathcal{B}$  are then appended to the objective function via indicator functions. The original objective function that is considered in [26] is quadratic in the state and control variables. In the next step in [26], the new objective function is split into its quadratic and convex parts and the Douglas-Rachford splitting method is applied to solve the problem.

In the current paper, we deal with continuous-time optimal control problems, which are infinite-dimensional optimization problems that are set in Hilbert spaces. After splitting the constraints of the problem, we apply Dykstra's algorithm [11], the Douglas-Rachford (DR) method [6, 9, 17, 18, 25, 29], and the Aragón Artacho-Campoy (AAC) algorithm [3], all of which solve the underlying best approximation problem.

The exposure of the current paper is more in the style of a tutorial. We pose the problem of minimum-energy control of a simplified model of a car, amounting to the double integrator, where the control variable has simple lower and upper bounds and the initial and terminal state variables are specified. We split the constraints into two,  $\mathcal{A}$  and  $\mathcal{B}$ , representing respectively the state differential equations (the double integrator) along with their boundary conditions and the constraints on the control variable. We define two subproblems, one subject to  $\mathcal{A}$ , and the other one subject to  $\mathcal{B}$ . We take advantage of the relatively simple form of the optimal control problem and derive analytical expressions for the optimality conditions and implement these in defining the projections onto  $\mathcal{A}$  and  $\mathcal{B}$ .

The solutions of these subproblems provide the projections of a given point in the control variable space onto the constraint sets  $\mathcal{A}$  and  $\mathcal{B}$ , respectively, in some optimal way. By performing these projections in the way prescribed by the above-listed algorithms, we can ensure convergence to a solution of the original optimal control problem,

Note that while the minimum-energy control of the double integrator without any constraints on the control variable can be solved analytically, the same problem with (even simple bound, i.e., box) constraints on the control variable can in general be solved only numerically. This problem should be considered within the framework of control-constrained linear-quadratic optimal control problems for which new numerical methods are constantly being developed—see for example [1, 12] and the references therein.

The current paper is a prototype for future applications of projection methods to solving more general optimal control problems. Indeed, the minimum-energy control of double integrator is a special case of linear quadratic optimal control

problems; so, with the reporting of the current study, an extension to more general problems will be imminent.

The paper is organized as follows. In Section 2, we state the control-constrained minimum-energy problem for the double integrator, and write down the optimality conditions. We provide the analytical solution for the unconstrained problem. For the control-constrained case, we briefly describe the standard numerical approach and consider an instance of the problem which we use in the numerical experiments in the rest of the paper. We define the constraint sets  $\mathcal{A}$  and  $\mathcal{B}$ . In Section 3, we provide the expressions for the projections onto  $\mathcal{A}$  and  $\mathcal{B}$ . We describe the algorithms in Section 4 and in the beginning of Section 5. In the remaining part of Section 5, we present numerical experiments to study parametric behaviour of the algorithms as well as the errors in the state and control variables with each algorithm. In Section 6, we provide concluding remarks and list some open problems.

## 1.2 Minimum-Energy Control of Double Integrator

We consider the minimum-energy control of a car, with a constrained control variable. Consider the car as a point unit mass, moving on a frictionless ground in a fixed line of action. Let the position of the car at time  $t$  be given by  $y(t)$  and the velocity by  $\dot{y}(t) := (dy/dt)(t)$ . By Newton's second law of motion,  $\ddot{y}(t) = u(t)$ , where  $u(t)$  is the summation of all the external forces applied on the car, in this case the force simply representing the acceleration and deceleration of the car. This differential equation model is referred to as the *double integrator* in system theory literature, since  $y(t)$  can be obtained by integrating  $u(t)$  twice.

**Optimal Control Problem** Suppose that the total force on the car, i.e., the acceleration or deceleration of the car, is constrained by a magnitude of  $a > 0$ . Let  $x_1 := y$  and  $x_2 := \dot{y}$ . Then the problem of minimizing the energy of the car, which starts at a position  $x_1(0) = s_0$  with a velocity  $x_2(0) = v_0$  and finishes at some other position  $x_1(1) = s_f$  with velocity  $x_2(1) = v_f$ , within one unit of time, can be posed as follows.

$$(P) \begin{cases} \min & \frac{1}{2} \int_0^1 u^2(t) dt \\ \text{subject to} & \dot{x}_1(t) = x_2(t), \quad x_1(0) = s_0, \quad x_1(1) = s_f, \\ & \dot{x}_2(t) = u(t), \quad x_2(0) = v_0, \quad x_2(1) = v_f, \quad |u(t)| \leq a. \end{cases}$$

Here, the functions  $x_1$  and  $x_2$  are referred to as the *state variables* and  $u$  the *control variable*. As a first step in writing the conditions of optimality for this optimization problem, define the Hamiltonian function  $H$  for Problem (P) simply as

$$H(x_1, x_2, u, \lambda_1, \lambda_2) := \frac{1}{2} u^2 + \lambda_1 x_2 + \lambda_2 u, \quad (1.1)$$

where  $\lambda(t) := (\lambda_1(t), \lambda_2(t)) \in \mathbf{R}^2$  is the *adjoint variable* (or *costate*) *vector* such that (see [21])

$$\dot{\lambda}_1 = -\partial H / \partial x_1 \quad \text{and} \quad \dot{\lambda}_2 = -\partial H / \partial x_2. \quad (1.2)$$

Equations in (1.2) simply reduce to

$$\lambda_1(t) = c_1 \quad \text{and} \quad \lambda_2(t) = -c_1 t - c_2, \quad (1.3)$$

where  $c_1$  and  $c_2$  are real constants. Let  $x(t) := (x_1(t), x_2(t)) \in \mathbf{R}^2$  denote the state variable vector.

**Maximum Principle** If  $u$  is an optimal control for Problem (P), then there exists a continuously differentiable vector of adjoint variables  $\lambda$ , as defined in (1.2), such that  $\lambda(t) \neq 0$  for all  $t \in [0, t_f]$ , and that, for a.e.  $t \in [0, t_f]$ ,

$$u(t) = \underset{v \in [-a, a]}{\operatorname{argmin}} H(x, v, \lambda(t)), \quad (1.4)$$

i.e.,

$$u(t) = \underset{v \in [-a, a]}{\operatorname{argmin}} \frac{1}{2} v^2 + \lambda_2(t) v; \quad (1.5)$$

see e.g. [21]. Condition (1.5) implies that the optimal control is given by

$$u(t) = \begin{cases} -\lambda_2(t), & \text{if } -a \leq \lambda_2(t) \leq a, \\ a, & \text{if } \lambda_2(t) \leq -a, \\ -a, & \text{if } \lambda_2(t) \geq a. \end{cases} \quad (1.6)$$

From (1.6), we can also conclude that the optimal control  $u$  for Problem (P) is continuous.

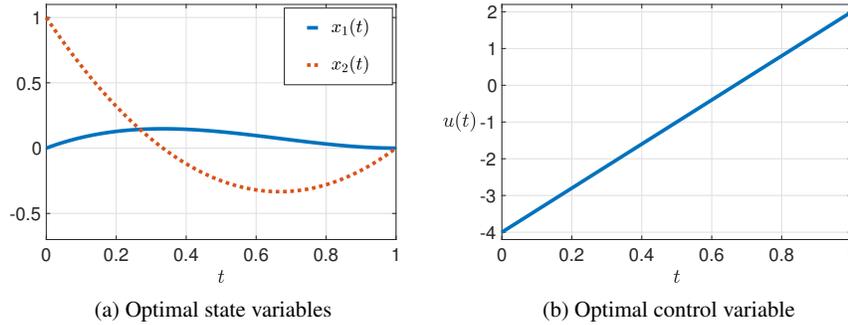
When  $a$  is large enough, the control constraint does not become active, so the optimal control is simply  $-\lambda_2$ , and it is a straightforward classroom exercise to find the analytical solution as

$$\begin{aligned} u(t) &= c_1 t + c_2, \\ x_1(t) &= \frac{1}{6} c_1 t^3 + \frac{1}{2} c_2 t^2 + v_0 t + s_0, \\ x_2(t) &= \frac{1}{2} c_1 t^2 + c_2 t + v_0, \end{aligned}$$

for all  $t \in [0, 1]$ , where

$$\begin{aligned} c_1 &= -12(s_f - s_0) + 6(v_0 + v_f), \\ c_2 &= 6(s_f - s_0) - 2(2v_0 + v_f). \end{aligned}$$

The solution of an instance of Problem (P), with  $s_0 = 0$ ,  $s_f = 0$ ,  $v_0 = 1$ ,  $v_f = 0$ , and large  $a$ , say  $a = 9$ , is depicted in Figure 1.1. Note that, for all  $t \in [0, 1]$ ,



**Fig. 1.1** Solution of Problem (P) with large  $a$  (so that  $u(t)$  is unconstrained),  $s_0 = 0$ ,  $s_f = 0$ ,  $v_0 = 1$ ,  $v_f = 0$

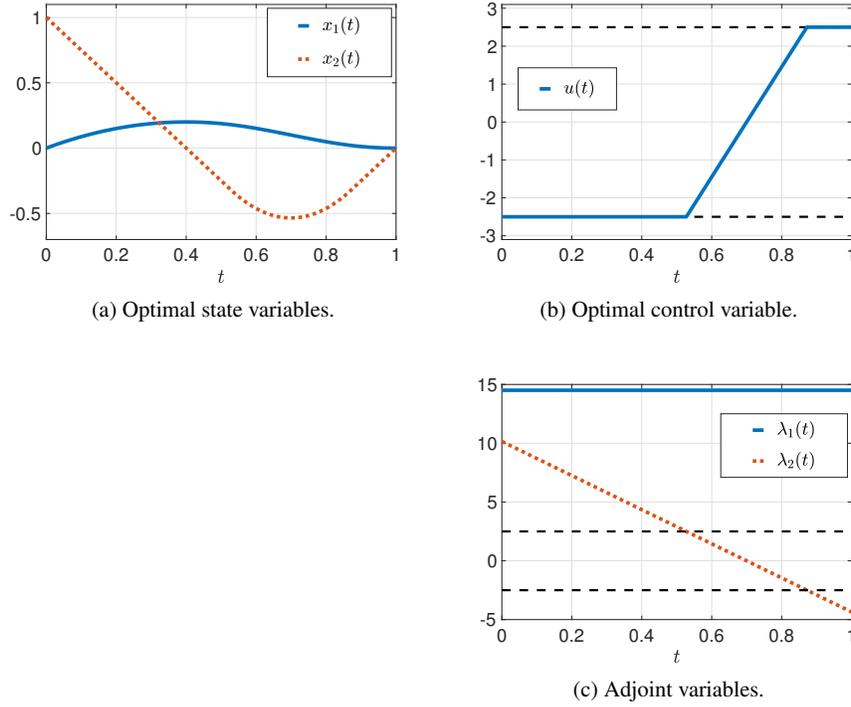
$\lambda_2(t) = -u(t) = -6t + 4$  and  $\lambda_1(t) = c_1 = 6$ . The graphs of  $\lambda_1$  and  $\lambda_2$  are not displayed for this particular instance.

When  $a$  is not so large, say  $a = 2.5$ , as we will consider next so that the control constraint becomes active, it is usually not possible to find an analytical solution, i.e., a solution has to be found numerically, as described below.

**Numerical Approach** A straightforward and popular numerical approach to solving Problem (P) is to discretize Problem (P) over a partition of the time horizon  $[0, 1]$  and then use some finite-dimensional optimization software to get a *discrete* (finite-dimensional) *solution* for the state and control variables  $x(t)$  and  $u(t)$ . The discrete solution is an approximation of the continuous-time solution. This approach is often referred to as the *direct method* or the *(first-)discretize-then-optimize* approach. A survey and discussion of Euler discretization of linear-quadratic optimal control problems and convergence of their discretized solutions to their continuous-time solutions can be found in [12, Section 5].

Figure 1.2 depicts the discrete solution of Problem (P) with the instance where  $a = 2.5$ ,  $s_0 = 0$ ,  $s_f = 0$ ,  $v_0 = 1$ ,  $v_f = 0$ . The solution was obtained by pairing up the optimization modelling language AMPL [20] and the finite-dimensional optimization software Ipopt [30]. The number of discretization nodes was taken to be 2000. The multipliers of the (Euler approximation of the) state differential equation constraints are provided by Ipopt when it finds an optimal solution to the discretized (finite-dimensional) problem. These multipliers have been plotted in Figure 1.2(c). It should be noted that the graph of the adjoint variable  $\lambda_2(t)$  given in Figure 1.2(c) verifies the graph of the optimal control  $u(t)$  in Figure 1.2(b) via the optimal control rule in (1.6). In Figures 1.2(b) and (c), the bounds  $\pm 2.5$  have been marked by horizontal dashed lines for ease of viewing.

*Remark 1.* If  $a$  is too small, there will obviously be no solution to Problem (P). For the particular instance of the problem considered here, the critical value of  $a$ , below which there exists no solution, is somewhere between 2.414 and 2.415, as our



**Fig. 1.2** Solution of direct discretization of Problem (P), with  $a = 2.5$ ,  $s_0 = 0$ ,  $s_f = 0$ ,  $v_0 = 1$ ,  $v_f = 0$

numerical experiments show (not reported in detail here). At this critical value, the only feasible solution is bang–bang, i.e.,  $u(t)$  switches once from  $-a$  to  $a$  at around  $t = 0.71$ . It should be noted that, in this case, the optimal control in (1.6) requires the adjoint variable  $\lambda_2$  to switch from a value  $\alpha \geq a$  to another value  $\beta \leq -a$ , i.e., be discontinuous, which is not allowed by the maximum principle. In this paper, we only consider the case when  $a$  is strictly greater than its critical value so that the maximum principle can be applied.

**Function Spaces** For the numerical methods, we consider projection/reflection methods in Hilbert spaces. The spaces associated with Problem (P) are set up as follows. Let  $q \in \mathbf{N}$  and  $L^2(0, 1; \mathbf{R}^q)$  be the Banach space of Lebesgue measurable functions

$$\begin{aligned} z : [0, 1] &\rightarrow \mathbf{R}^q \\ t &\mapsto (z_1(t), \dots, z_q(t)), \end{aligned}$$

with finite  $L^2$  norm. Namely, define

$$\|z\|_2 := \left( \sum_{i=1}^q \|z_i\|_2^2 \right)^{1/2},$$

where

$$\|z_i\|_2 := \left( \int_0^1 |z_i(t)|^2 dt \right)^{1/2},$$

for  $i = 1, \dots, q$ , with  $|\cdot|$  the modulus or absolute value. In other words,

$$L^2(0, 1; \mathbf{R}^q) := \{z: [0, 1] \rightarrow \mathbf{R}^q : \|z\|_2 < \infty\}.$$

Furthermore,  $W^{1,2}(0, 1; \mathbf{R}^q)$  is the Sobolev space of absolutely continuous functions, namely

$$W^{1,2}(0, 1; \mathbf{R}^q) = \{z \in L^2(0, 1; \mathbf{R}^q) \mid \dot{z} = dz/dt \in L^2(0, 1; \mathbf{R}^q)\},$$

endowed with the norm

$$\|z\|_{W^{1,2}} := \left( \sum_{i=1}^q [\|z_i\|_2^2 + \|\dot{z}_i\|_2^2] \right)^{1/2}.$$

In Problem (P), the state variable  $x \in W^{1,2}(0, 1; \mathbf{R}^2)$  and the control variable  $u \in L^2(0, 1; \mathbf{R})$ .

**Constraint Splitting** Next, we split the constraints of Problem (P) into two subsets,  $\mathcal{A}$  and  $\mathcal{B}$ . The subset  $\mathcal{A}$  collects together all the feasible control functions satisfying only the dynamics of the car. The subset  $\mathcal{B}$ , on the other hand, collects all the control functions whose values are constrained by  $-a$  and  $a$ .

$$\begin{aligned} \mathcal{A} &:= \{u \in L^2(0, 1; \mathbf{R}) \mid \exists x \in W^{1,2}(0, 1; \mathbf{R}^2) \text{ which solves} \\ &\quad \dot{x}_1(t) = x_2(t), \quad x_1(0) = s_0, \quad x_1(1) = s_f, \\ &\quad \dot{x}_2(t) = u(t), \quad x_2(0) = v_0, \quad x_2(1) = v_f, \quad \forall t \in [0, 1]\}, \quad (1.7) \end{aligned}$$

$$\mathcal{B} := \{u \in L^2(0, 1; \mathbf{R}) \mid -a \leq u(t) \leq a, \text{ for all } t \in [0, 1]\}. \quad (1.8)$$

The rationale behind this sort of splitting is as follows: The problem of minimizing the energy of the car subject to only  $\mathcal{A}$  or only  $\mathcal{B}$  is much easier to solve – in fact, the solutions can be analytically written up in each case. If, for some given  $u$ , a solution exists to the two-point boundary-value problem (TPBVP) in (1.7) then that solution is unique by the linearity of the TPBVP [5, 28]. Note that a control solution  $u$  as in (1.7) exists by the (Kalman) controllability of the double integrator – see [27]. So the set  $\mathcal{A}$  is nonempty. Note that the constraint set  $\mathcal{A}$  is an *affine subspace* and  $\mathcal{B}$  a *box*.

### 1.3 Projections

All of the projection methods that we will consider involve projections onto the sets  $\mathcal{A}$  and  $\mathcal{B}$ . The projection onto  $\mathcal{A}$  from a current iterate  $u^-$  is the point  $u$  solving the following problem.

$$(P1) \begin{cases} \min & \frac{1}{2} \int_0^1 (u(t) - u^-(t))^2 dt \\ \text{subject to} & u \in \mathcal{A}. \end{cases}$$

In (P1), we minimize the squared  $L^2$ -norm distance between  $u^-$  and  $u$ . The projection onto  $\mathcal{B}$  from a current iterate  $u^-$  is similarly the point  $u$  solving the following problem.

$$(P2) \begin{cases} \min & \frac{1}{2} \int_0^1 (u(t) - u^-(t))^2 dt \\ \text{subject to} & u \in \mathcal{B}. \end{cases}$$

**Proposition 1 (Projection onto  $\mathcal{A}$ ).** *The projection  $P_{\mathcal{A}}$  of  $u^- \in L^2(0, 1; \mathbf{R})$  onto the constraint set  $\mathcal{A}$ , as the solution of Problem (P1), is given by*

$$P_{\mathcal{A}}(u^-)(t) = u^-(t) + c_1 t + c_2, \quad (1.9)$$

for all  $t \in [0, 1]$ , where

$$c_1 = 12(x_1(1) - s_f) - 6(x_2(1) - v_f), \quad (1.10)$$

$$c_2 = -6(x_1(1) - s_f) + 2(x_2(1) - v_f), \quad (1.11)$$

and  $x_1(1)$  and  $x_2(1)$  are obtained by solving the initial value problem

$$\dot{x}_1(t) = x_2(t), \quad x_1(0) = s_0, \quad (1.12)$$

$$\dot{x}_2(t) = u^-(t), \quad x_2(0) = v_0, \quad (1.13)$$

for all  $t \in [0, 1]$ .

*Proof.* The Hamiltonian function for Problem (P1) is

$$H_1(x_1, x_2, u, \lambda_1, \lambda_2, t) := \frac{1}{2} (u - u^-)^2 + \lambda_1 x_2 + \lambda_2 u,$$

where the adjoint variables  $\lambda_1$  and  $\lambda_2$  are defined as in (1.2), with  $H$  replaced by  $H_1$ , and the subsequent solutions are given as in (1.3). The optimality condition for Problem (P1) is akin to that in (1.4) for Problem (P) and, owing to the fact that the control  $u$  is now unconstrained, can more simply be written as

$$\frac{\partial H_1}{\partial u}(x, u, \lambda, t) = 0,$$

which yields the optimal control as  $u(t) = u^-(t) - \lambda_2(t)$ , i.e.

$$u(t) = u^-(t) + c_1 t + c_2, \quad (1.14)$$

for all  $t \in [0, 1]$ . We need to show that  $c_1$  and  $c_2$  are found as in (1.10)–(1.11). Using (1.14) in (1.7) yields the following time-varying, linear two-point boundary-value problem.

$$\dot{x}_1(t) = x_2(t), \quad x_1(0) = s_0, \quad x_1(1) = s_f, \quad (1.15)$$

$$\dot{x}_2(t) = u^-(t) + c_1 t + c_2, \quad x_2(0) = v_0, \quad x_2(1) = v_f, \quad (1.16)$$

for all  $t \in [0, 1]$ . In other words, Problem (P1) is reduced to solving Equations (1.15)–(1.16) for the unknown parameters  $c_1$  and  $c_2$ . Once  $c_1$  and  $c_2$  are found, the projected point  $u$  in (1.14) is found. Since Equations (1.15)–(1.16) are linear in  $x_1$  and  $x_2$ , a *simple shooting technique* [5, 28] provides the solution for  $c_1$  and  $c_2$  in just one iteration. The essence of this technique is that the initial-value problem (IVP)

$$\frac{\partial z_1(t, c)}{\partial t} = z_2(t, c), \quad z_1(0, c) = s_0, \quad (1.17)$$

$$\frac{\partial z_2(t, c)}{\partial t} = u^-(t) + c_1 t + c_2, \quad z_2(0, c) = v_0, \quad (1.18)$$

for all  $t \in [0, 1]$ , is solved repeatedly, so as to make the *discrepancy at  $t = 1$*  vanish. Namely, we seek a parameter  $c := (c_1, c_2)$  such that  $z_1(1, c) - s_f = 0$  and  $z_2(1, c) - v_f = 0$ . The procedure is as follows. For a given  $c$ , there exists a unique solution  $z(t, c) := (z_1(t, c), z_2(t, c))$  of (1.17)–(1.18). Define the *near-miss* (vector) *function*  $\varphi : \mathbf{R}^2 \rightarrow \mathbf{R}^2$  as follows:

$$\varphi(c) := \begin{bmatrix} z_1(1, c) - s_f \\ z_2(1, c) - v_f \end{bmatrix}. \quad (1.19)$$

The Jacobian of the near-miss function is

$$J_\varphi(c) := \begin{bmatrix} \frac{\partial z_1(1, c)}{\partial c_1} & \frac{\partial z_1(1, c)}{\partial c_2} \\ \frac{\partial z_2(1, c)}{\partial c_1} & \frac{\partial z_2(1, c)}{\partial c_2} \end{bmatrix}$$

The shooting method looks for a pair  $c$  such that  $\varphi(c) := 0$  (i.e., a pair  $c$  such that the terminal boundary conditions are met). Expanding  $\varphi$  about, say,  $\bar{c} = 0$ , and discarding the terms of order 2 or higher, we obtain

$$\varphi(c) \approx \varphi(0) + J_\varphi(0) c.$$

Substituting  $\varphi(c) = 0$  in the above expression, replacing “ $\approx$ ” with “ $=$ ”, and re-arranging, gives the single (Newton) iteration of the shooting method:

$$c = -[J_\varphi(0)]^{-1}\varphi(0). \quad (1.20)$$

The components  $(\partial z_i / \partial c_j)(1, c)$ ,  $i, j = 1, 2$ , of  $J_\varphi(c)$ , can be obtained by solving the variational equations for (1.15)–(1.16) with respect to  $c_1$  and  $c_2$ , i.e., by solving the following system for  $(\partial z_i / \partial c_j)(\cdot, c)$ :

$$\begin{aligned} \frac{\partial}{\partial t} \left( \frac{\partial z_1}{\partial c_1} \right) (t, c) &= \frac{\partial z_2}{\partial c_1} (t, c), & \frac{\partial z_1}{\partial c_1} (0, c) &= 0, \\ \frac{\partial}{\partial t} \left( \frac{\partial z_1}{\partial c_2} \right) (t, c) &= \frac{\partial z_2}{\partial c_2} (t, c), & \frac{\partial z_1}{\partial c_2} (0, c) &= 0, \\ \frac{\partial}{\partial t} \left( \frac{\partial z_2}{\partial c_1} \right) (t, c) &= t, & \frac{\partial z_2}{\partial c_1} (0, c) &= 0, \\ \frac{\partial}{\partial t} \left( \frac{\partial z_2}{\partial c_2} \right) (t, c) &= 1, & \frac{\partial z_2}{\partial c_2} (0, c) &= 0. \end{aligned}$$

Elementary calculations lead to the following solution of the above system:

$$\frac{\partial z}{\partial c} (t, c) = \begin{bmatrix} t^3/6 & t^2/2 \\ t^2/2 & t \end{bmatrix},$$

which is independent of  $c$ . Hence,

$$J_\varphi(0) = \frac{\partial z}{\partial c} (1, 0) = \begin{bmatrix} 1/6 & 1/2 \\ 1/2 & 1 \end{bmatrix},$$

with inverse:

$$\left[ \frac{\partial z}{\partial c} (1, 0) \right]^{-1} = [J_\varphi(0)]^{-1} = \begin{bmatrix} -12 & 6 \\ 6 & -2 \end{bmatrix}. \quad (1.21)$$

Setting  $(x_1(\cdot), x_2(\cdot)) := (z_1(\cdot, 0), z_2(\cdot, 0))$ , the IVP (1.17)–(1.18) becomes (1.12)–(1.13). Then substitution of (1.19) and (1.21) with  $c = 0$  into Equation (1.20), and expanding out, yield (1.10)–(1.11). The proof is complete.

**Proposition 2 (Projection onto  $\mathcal{B}$ ).** *The projection  $P_{\mathcal{B}}$  of  $u^- \in L^2(0, 1; \mathbf{R})$  onto the constraint set  $\mathcal{B}$ , as the solution of Problem (P2), is given by*

$$P_{\mathcal{B}}(u^-)(t) = \begin{cases} u^-(t), & \text{if } -a \leq u^-(t) \leq a, \\ -a, & \text{if } u^-(t) \leq -a, \\ a, & \text{if } u^-(t) \geq a, \end{cases} \quad (1.22)$$

for all  $t \in [0, 1]$ .

*Proof.* The expression (1.22) is the straightforward solution of Problem (P2).

## 1.4 Best Approximation Algorithms

In this section, we discuss best approximation algorithms. In the following,

$$X \text{ is a real Hilbert space} \quad (1.23)$$

with inner product  $\langle \cdot, \cdot \rangle$ , induced norm  $\| \cdot \|$ . We also assume that

$$A \text{ is a closed affine subspace of } X, \text{ and } B \text{ is a nonempty closed convex subset of } X. \quad (1.24)$$

Given  $z \in X$ , our aim is to find

$$P_{A \cap B}(z), \quad (1.25)$$

the projection of  $z$  onto the intersection  $A \cap B$  which we assume to be *nonempty*. We also assume that we are able to compute the projectors onto the constraints  $P_A$  and  $P_B$ .

Many algorithms are known which could be employed to find  $P_{A \cap B}(z)$ ; here, however, we focus on three simple methods that do not require a product space set-up as some of those considered, in, e.g., [6, 7, 13, 14].

In the next section, we will numerically test these algorithms when  $X = L^2(0, 1; \mathbf{R})$ ,  $A = \mathcal{A}$ ,  $B = \mathcal{B}$ , and  $z = 0$ .

### 1.4.1 Dykstra's Algorithm

We start with Dykstra's algorithm (see [11]), which operates as follows<sup>1</sup>: Set  $a_0 := z$  and  $q_0 := 0$ . Given  $a_n, q_n$ , where  $n \geq 0$ , update

$$b_n := P_B(a_n + q_n), \quad a_{n+1} := P_A(b_n), \quad \text{and} \quad q_{n+1} := a_n + q_n - b_n. \quad (1.26)$$

It is known that both  $(a_n)_{n \in \mathbb{B}}$  and  $(b_n)_{n \in \mathbb{N}}$  converge *strongly* to  $P_{A \cap B}(z)$ .

### 1.4.2 Douglas–Rachford Algorithm

Given  $\beta > 0$ , we specialize the Douglas–Rachford algorithm (see [17], [25] and [18]) to minimize the sum of the two functions  $f(x) = \iota_B(x) + \frac{\beta}{2} \|x - z\|^2$  and  $g := \iota_A$  which have respective proximal mappings (see [6, Proposition 24.8(i)])

<sup>1</sup> In the general case, there is also an auxiliary sequence  $(p_n)$  associated with  $A$ ; however, because  $A$  is an affine subspace, it is not needed in our setting.

$P_f(x) = P_B\left(\frac{1}{1+\beta}x + \frac{\beta}{1+\beta}z\right)$  and  $P_g = P_A$ . Set  $\lambda := \frac{1}{1+\beta} \in ]0, 1[$ . It follows that the Douglas–Rachford operator  $T := \text{Id} - P_f + P_g(2P_f - \text{Id})$  turns into

$$Tx = x - P_B(\lambda x + (1 - \lambda)z) + P_A\left(2P_B(\lambda x + (1 - \lambda)z) - x\right). \quad (1.27)$$

Now let  $x_0 \in X$  and given  $x_n \in X$ , where  $n \geq 0$ , update

$$b_n := P_B(\lambda x_n + (1 - \lambda)z), \quad x_{n+1} := Tx_n = x_n - b_n + P_A(2b_n - x_n). \quad (1.28)$$

Then it is known (see [29] or [9]) that  $(b_n)_{n \in \mathbb{N}}$  converges weakly to  $P_{A \cap B}(z)$ . Note that (1.28) simplifies to

$$x_{n+1} := x_n - P_B(\lambda x_n) + P_A(2P_B(\lambda x_n) - x_n) \quad \text{provided that } z = 0. \quad (1.29)$$

### 1.4.3 Aragón Artacho–Campoy Algorithm

The Aragón Artacho–Campoy (AAC) Algorithm was recently presented in [3]; see also [2, 4]. Given two fixed parameters  $\alpha$  and  $\beta$  in  $]0, 1[$ , define

$$\begin{aligned} Tx &= (1 - \alpha)x \\ &\quad + \alpha \left( 2\beta \left( P_A \left( 2\beta (P_B(x+z) - z) - x + z \right) - z \right) + x + 2\beta (z - P_B(x+z)) \right) \\ &= x + 2\alpha\beta \left( P_A \left( 2\beta (P_B(x+z) - z) - x + z \right) - P_B(x+z) \right). \end{aligned} \quad (1.30)$$

Now let  $x_0 \in X$  and given  $x_n \in X$ , where  $n \geq 0$ , update

$$b_n := P_B(x_n + z), \quad (1.31)$$

and

$$x_{n+1} := Tx_n = x_n + 2\alpha\beta \left( P_A \left( 2\beta (b_n - z) - x_n + z \right) - b_n \right). \quad (1.32)$$

By [3, Theorem 4.1(iii)], the sequence  $(b_n)_{n \in \mathbb{N}}$  converges strongly to  $P_{A \cap B}(z)$  provided that<sup>2</sup>  $z - P_{A \cap B}(z) \in (N_A + N_B)(P_{A \cap B}z)$ . Note that (1.32) simplifies to

$$x_{n+1} := Tx_n = x_n + 2\alpha\beta \left( P_A(2\beta P_B x_n - x_n) - P_B x_n \right) \quad \text{provided that } z = 0. \quad (1.33)$$

---

<sup>2</sup> It appears that this constraint qualification is not easy to check in our setting.

## 1.5 Numerical Implementation

### 1.5.1 The Algorithms

In this section, we gather the algorithms considered abstractly and explain how we implemented them.

We start with Dykstra's algorithm from Subsection 1.4.1.

#### Algorithm 1 (Dykstra)

- Step 1 (*Initialization*) Choose the initial iterates  $u^0 = 0$  and  $q^0 = 0$ . Choose a small parameter  $\varepsilon > 0$ , and set  $k = 0$ .
- Step 2 (*Projection onto  $\mathcal{B}$* ) Set  $u^- = u^k + q^k$ . Compute  $\tilde{u} = P_{\mathcal{B}}(u^-)$  by using (1.22).
- Step 3 (*Projection onto  $\mathcal{A}$* ) Set  $u^- := \tilde{u}$ . Compute  $\hat{u} = P_{\mathcal{A}}(u^-)$  by using (1.9).
- Step 4 (*Update*) Set  $u^{k+1} := \hat{u}$  and  $q^{k+1} := u^k + q^k - \tilde{u}$ .
- Step 5 (*Stopping criterion*) If  $\|u^{k+1} - u^k\|_{L^\infty} \leq \varepsilon$ , then return  $\tilde{u}$  and stop. Otherwise, set  $k := k + 1$  and go to Step 2.

Next is the Douglas–Rachford method from Subsection 1.4.2.

#### Algorithm 2 (DR)

- Step 1 (*Initialization*) Choose a parameter  $\lambda \in ]0, 1[$  and the initial iterate  $u^0$  arbitrarily. Choose a small parameter  $\varepsilon > 0$ , and set  $k = 0$ .
- Step 2 (*Projection onto  $\mathcal{B}$* ) Set  $u^- = \lambda u^k$ . Compute  $\tilde{u} = P_{\mathcal{B}}(u^-)$  by using (1.22).
- Step 3 (*Projection onto  $\mathcal{A}$* ) Set  $u^- := 2\tilde{u} - u^k$ . Compute  $\hat{u} = P_{\mathcal{A}}(u^-)$  by using (1.9).
- Step 4 (*Update*) Set  $u^{k+1} := u^k + \hat{u} - \tilde{u}$ .
- Step 5 (*Stopping criterion*) If  $\|u^{k+1} - u^k\|_{L^\infty} \leq \varepsilon$ , then return  $\tilde{u}$  and stop. Otherwise, set  $k := k + 1$  and go to Step 2.

Finally, we describe the Aragón Artacho–Campoy algorithm from Subsection 1.4.3.

#### Algorithm 3 (AAC)

- Step 1 (*Initialization*) Choose the initial iterate  $u^0$  arbitrarily. Choose a small parameter  $\varepsilon > 0$ , two parameters<sup>3</sup>  $\alpha$  and  $\beta$  in  $]0, 1[$ , and set  $k = 0$ .
- Step 2 (*Projection onto  $\mathcal{B}$* ) Set  $u^- = u^k$ . Compute  $\tilde{u} = P_{\mathcal{B}}(u^-)$  by using (1.22).
- Step 3 (*Projection onto  $\mathcal{A}$* ) Set  $u^- = 2\beta\tilde{u} - u^k$ . Compute  $\hat{u} = P_{\mathcal{A}}(u^-)$  by using (1.9).
- Step 4 (*Update*) Set  $u^{k+1} := u^k + 2\alpha\beta(\hat{u} - \tilde{u})$ .
- Step 5 (*Stopping criterion*) If  $\|u^{k+1} - u^k\|_{L^\infty} \leq \varepsilon$ , then return  $\tilde{u}$  and stop. Otherwise, set  $k := k + 1$  and go to Step 2.

<sup>3</sup> Aragón Artacho and Campoy recommend  $\alpha = 0.9$  and  $\beta \in [0.7, 0.8]$ ; see [3, End of Section 7].

We provide another version of each of Algorithms 1–3, as Algorithms 1b–3b, in Appendix A. In Algorithm 1b, we *monitor* the sequence of iterates which are the projections onto set  $\mathcal{A}$ , instead of monitoring the projections onto set  $\mathcal{B}$  in Algorithm 1. On the other hand, in Algorithms 2b–3b, the order in which the projections are done is reversed: the first projection is done onto the set  $\mathcal{A}$  and the second projection onto  $\mathcal{B}$ .

Although the order of projections will not matter in view of the existing results stating that convergence is achieved under any order – see [8, Proposition 2.5(i)], the order does make a difference in early iterations (as well as in the number of iterations required for convergence of Algorithms 2 and 2b, as we will elaborate on later). If our intent is to stop the algorithm early so that we can use the current iterate as an initial guess in more accurate computational optimal control algorithms, which can find the junction times with a high precision (see [22–24]), then it is desirable to implement Algorithms 1–3 above, rather than Algorithms 1b–3b, because any iterate of Algorithms 1–3 will satisfy the constraints on the control variable, while that of Algorithms 1b–3b will in general not.

### 1.5.2 Numerical Experiments

In what follows, we study the working of Algorithms 1–3 for an instance of Problem (P). Suppose that the car is initially at a reference position 0 and has unit speed. It is desired that the car come back to the reference position and be at rest after one unit of time; namely that  $s_0 = 0$ ,  $s_f = 0$ ,  $v_0 = 1$ ,  $v_f = 0$ . For these boundary conditions, no solution exists if one takes the control variable bound  $a = 2.4$  or smaller but a solution does exist for  $a = 2.5$ . So, we use  $a = 2.5$ . In the ensuing discussions, we use the *stopping tolerance*  $\varepsilon = 10^{-8}$  unless otherwise stated.

**Discretization** Algorithms 1–3, as well as 1b–3b, carry out iterations with functions. For computations, we consider discrete approximations of the functions over the partition  $0 = t_0 < t_1 < \dots < t_N = 1$  such that

$$t_{i+1} = t_i + h, \quad i = 0, 1, \dots, N,$$

$h := 1/N$  and  $N$  is the number of subdivisions. Let  $u_i$  be an approximation of  $u(t_i)$ , i.e.,  $u_i \approx u(t_i)$ ,  $i = 0, 1, \dots, N-1$ ; similarly,  $x_{1,i} \approx x_1(t_i)$  and  $x_{2,i} \approx x_2(t_i)$ , or  $x_i := (x_{1,i}, x_{2,i}) \approx x(t_i)$ ,  $i = 0, 1, \dots, N$ . In other words, the functions  $u$ ,  $x_1$  and  $x_2$  are approximated by the  $N$ -dimensional array  $u_h$ , with components  $u_i$ ,  $i = 0, 1, \dots, N-1$ , and the  $(N+1)$ -dimensional arrays  $x_{1,h}$  and  $x_{2,h}$ , with components  $x_{1,i}$  and  $x_{2,i}$ ,  $i = 0, 1, \dots, N$ , respectively. We define a discretization  $P_{\mathcal{A}}^h$  of the projection  $P_{\mathcal{A}}$  as follows.

$$P_{\mathcal{A}}^h(u_h^-)(t) = u_h^- + c_1 t_h + c_2, \quad (1.34)$$

where  $t_h = (0, t_1, \dots, t_N)$ ,

$$c_1 = 12(x_{1,N} - s_f) - 6(x_{2,N} - v_f), \quad (1.35)$$

$$c_2 = -6(x_{1,N} - s_f) + 2(x_{2,N} - v_f), \quad (1.36)$$

and  $x_{1,N}$  and  $x_{2,N}$  are obtained from the Euler discretization of (1.12)–(1.13): Given  $x_{1,0} = s_0$  and  $x_{2,0} = v_0$ ,

$$x_{1,i+1} = x_{1,i} + hx_{2,i}, \quad (1.37)$$

$$x_{2,i+1} = x_{2,i} + hu_i^-(t), \quad (1.38)$$

for  $i = 0, 1, \dots, N - 1$ .

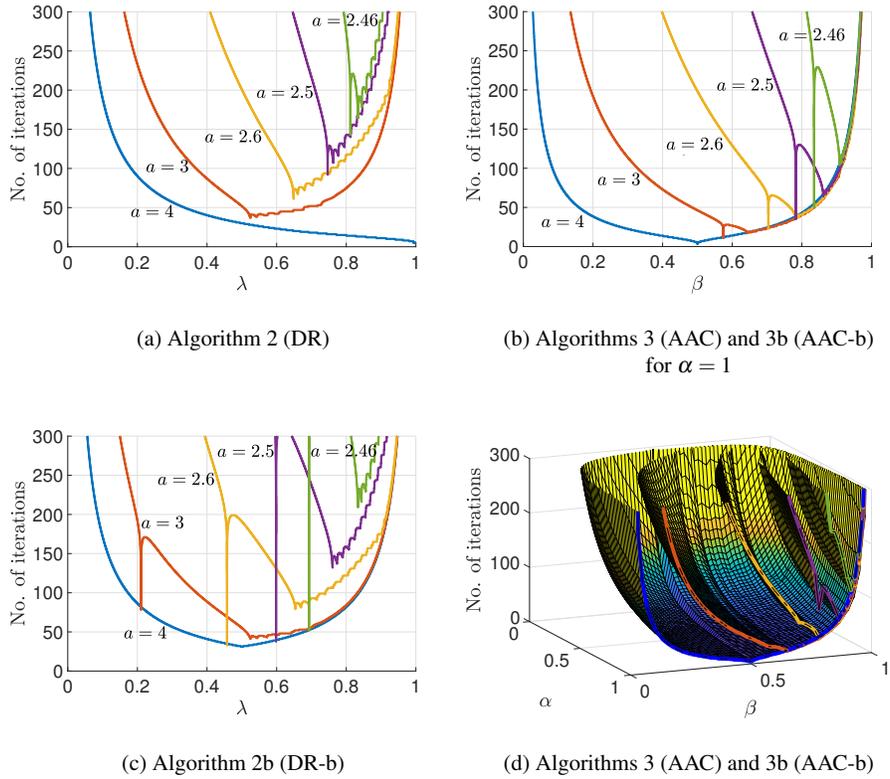
The discretization  $P_{\mathcal{B}}^h$  of the projection  $P_{\mathcal{B}}$  can be defined in a straightforward manner, by simply replacing  $u^-$  in (1.22) with the discrete components  $u_i^-$  of  $u_h^-$ .

**Parametric Behaviour** Obviously, the behaviour of Algorithms 2 and 2b, the Douglas–Rachford method, depend on the parameter  $\lambda$ , and the behaviour of Algorithms 3 and 3b on the two parameters  $\alpha$  and  $\beta$ . Figure 1.3 displays the dependence of the number of iterations it takes to converge on these parameters, for various values of  $a$ . The dependence for a given value of  $a$  appears to be continuous, albeit the presence of downward *spikes*.

The graphs for Algorithms 2 and 2b, shown in parts (a) and (c) of Figure 1.3, respectively, differ significantly from one another. The bound  $a = 4$  corresponds to the case when the control constraint becomes active only at  $t = 0$  – see Figure 1.1. In other words, when  $a > 0$  the optimal control variable is truly unconstrained. When  $a = 4$ , the best value of  $\lambda$  is 1 for Algorithm 2, yielding the solution in just 6 iterations. For Algorithm 2b, the best value for  $\lambda$  is 0.5, as can be seen in (c), producing the solution in 30 iterations. Going back to Algorithm 2, with decreasing values of  $a$ , the values of  $\lambda$  minimizing the number of iterations shift to the right. For example, the minimum number of iterations is 91, with  $a = 2.5$  and  $\lambda = 0.7466$  (found by a refinement of the graph).

As for Algorithm 2b, the minimizer for  $a = 2.5$  is  $\lambda = 0.5982766$  and the corresponding minimum number of iterations is 38. This is a point where a downward spike occurs and so the number of iterations is very sensitive to changes in  $\lambda$ . For example, the rounded-off value of  $\lambda = 0.598277$  results in 88 iterations instead of 38, and  $\lambda = 0.55$  yields 444 iterations for convergence. The number of iterations is less sensitive to the local minimizer  $\lambda = 0.7608$ , which results in 132 iterations. It is interesting to note that the graph with  $a = 4$  appears to be an envelope for the number of iterations for all  $\lambda \in ]0, 1[$ .

The graphs for Algorithms 3 and 3b, the Aragón Artacho–Campoy algorithm, are indistinguishable to one’s eye; therefore we only display the one in Figure 1.3(b). Part (d) of Figure 1.3 shows surface plots of the number of iterations versus the algorithmic parameters  $\alpha$  and  $\beta$ , for the same values of  $a$  as in the rest of the graphs in the figure. It is interesting to observe that the surfaces look to be cascaded with (roughly) the outermost surface being the one corresponding to  $a = 4$ . The surface plot suggests that for minimum number of iterations, one must have  $\alpha = 1$ . Although



**Fig. 1.3** Numerical experiments with  $s_0 = 0$ ,  $s_f = 0$ ,  $v_0 = 1$ ,  $v_f = 0$ .

theory requires  $\alpha < 1$ ,  $\alpha = 1$  seems to cause no concerns in this particular instance; so, we set  $\alpha = 1$  for the rest of the paper. The cross-sectional curves at  $\alpha = 1$  are shown with much more precision in part (b) of the figure. The *spikes* that are observed in part (d) can also be seen in the graph in part (b).

In fact, the first observation one has to make here is that, for  $a = 4$ , convergence can be achieved in merely one iteration, with  $\beta = 0.5$ . This is quite remarkable, compared with Algorithms 2 and 2b. The graphs in (b) appear to be enveloped as well by the graph for  $a = 4$ , as in part (c). For the values of  $a$  other than 4, the globally minimum number of iterations seems to be achieved at a downward spike, which as a result is very sensitive to changes in  $\beta$ . For example, for  $a = 2.5$ , the optimal  $\beta$  value is 0.78249754 for a minimum 35 iterations. A rounded-off  $\beta = 0.782$  results in 111 iterations, and  $\beta = 0.7$  yields 243 iterations. Sensitivity at the local minimizer  $\beta = 0.8617$  giving 64 iterations is far less: Choosing  $\beta = 0.8$  or 0.9 results in 128 or 90 iterations, respectively. It is interesting to note that, as in the case of Algorithms 2 and 2b, the graphs in Figure 1.3(b) are approximately enveloped by the graph/curve drawn for  $a = 4$ .

**Behaviour in Early Iterations** Figure 1.4(a)–(c) illustrates the working of all three algorithms for the same instance. All three algorithms converge to the optimal solution, with the stopping tolerance of  $\varepsilon = 10^{-8}$ . The optimal values of the algorithmic parameters,  $\lambda = 0.7466$  for Algorithm 2, and  $\alpha = 1$  and  $\beta = 0.8617$  for Algorithm 3, have been used. The third, fifth and fifteenth iterates, as well as the solution curve, are displayed for comparisons of behaviour. At least for the given instance of the problem, it is fair to say from Figure 1.4(c) that Algorithm 3 gets closer to the solution much more quickly than the others in the few initial iterations—see the third and fifth iterates. It also achieves convergence in a smaller number of iterations (64 as opposed to 530 and 91 iterations of the Algorithms 1 and 2, respectively).

**Error Analysis via Numerical Experiments** For a fixed value of  $N$ , Algorithms 1–3 converge only to some approximate solution of the original Problem. Therefore, the question as to how the algorithms behave as the time partition is refined, i.e.,  $N$  is increased, needs to be investigated. For the purpose of a numerical investigation, we define, in the  $k$ th iteration, the following errors. Suppose that the pair  $(u^*, x^*)$  is the optimal solution of Problem (P) and  $(u_h^k, x_h^k)$  an approximate solution of Problem (P) in the  $k$ th iteration of a given algorithm. Define

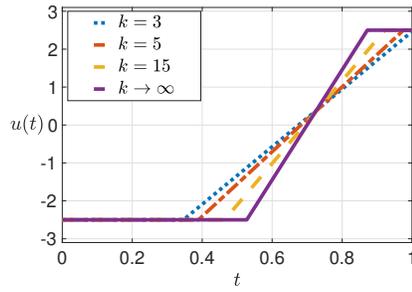
$$\sigma_u^k := \max_{0 \leq i \leq N-1} |u_i^k - u^*(t_i)| \quad \text{and} \quad \sigma_x^k := \max_{0 \leq i \leq N} \|x_i^k - x^*(t_i)\|_\infty,$$

where  $\|\cdot\|_\infty$  is the  $\ell_\infty$ -norm in  $\mathbf{R}^2$ . For large  $N$ , these expressions are reminiscent of the  $L^\infty$ -norm, and therefore they will be referred to as the  $L^\infty$ -error.

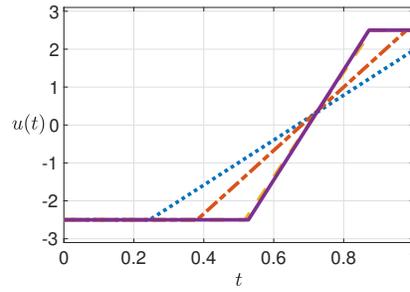
For  $(u^*, x^*)$  in the error expressions, we have used the discretized (approximate) solution obtained for the Euler-discretized Problem (P) utilizing the Ipopt–AMPL suite, with  $N = 10^6$  and the tolerance set at  $10^{-14}$ .

For  $N = 2000$ , these errors are depicted in Figure 1.4(d) and (e). From the graphs it is immediately clear that no matter how much smaller the stopping tolerance is selected, the best error that is achievable with  $N = 2000$  is around  $10^{-2}$  for the control variable and around  $10^{-3}$  for the state variable vector. In fact, the graphs also tell that perhaps a much smaller stopping threshold than  $10^{-8}$  would have achieved the same approximation to the continuous-time solution of Problem (P). By just looking at the graphs, one can see that Algorithm 1 could have been run just for about 300 iterations instead of 530, and Algorithms 2 and 3 could have been run for about 50 iterations to achieve the best possible approximation with  $N = 2000$ .

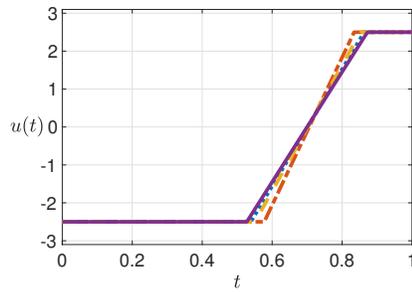
In Figure 1.5, we depict the same errors for  $N = 10^3$  (parts (a) and (b)),  $N = 10^4$  (in parts (c) and (d)) and  $N = 10^5$  (in parts (e) and (f)). It is observed that, with a ten-fold increase in  $N$  (which is a ten-fold decrease in  $h$ ) the errors in both  $u$  and  $x$  are reduced by ten-folds, implying that the error (both in  $x$  and in  $u$ ) depends on the stepsize  $h$  linearly. This is in line with the theory of Euler-discretization of optimal control problems; see, for example, [15, 16]. Furthermore, even for very large values of  $N$ , it can be seen from these graphs that a stopping threshold slightly smaller than  $10^{-8}$  would suffice to get even more stringent error levels, such as around  $10^{-4}$  for the control variable and around  $10^{-5}$  for the state variable vector. A larger stopping threshold would obviously result in smaller number of iterations.



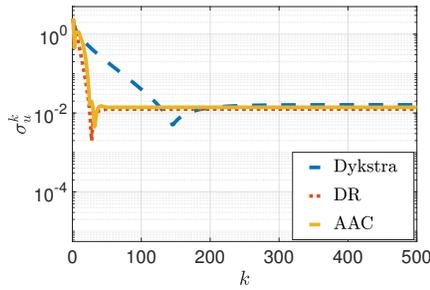
(a) Algorithm 1 (Dykstra)  
stops after 530 iterations.



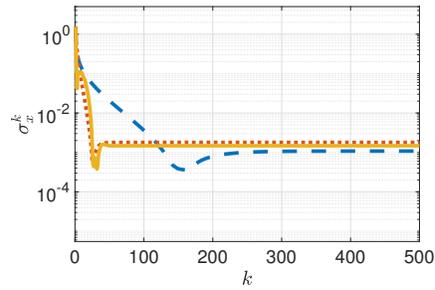
(b) Algorithm 2 (DR,  $\lambda = 0.7466$ )  
stops after 91 iterations.



(c) Algorithm 3 (AAC,  $\alpha = 1$ ,  $\beta = 0.8617$ )  
stops after 64 iterations.

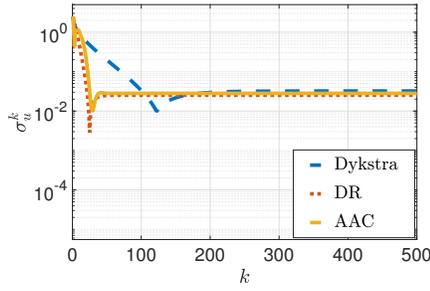


(d)  $L^\infty$ -error in control by Algorithms 1–3.

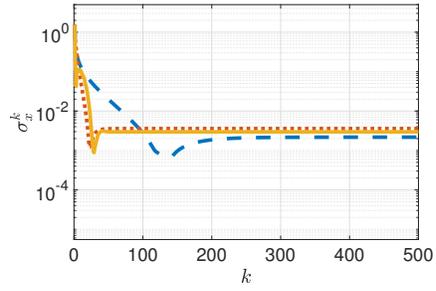


(e)  $L^\infty$ -error in states by Algorithms 1–3.

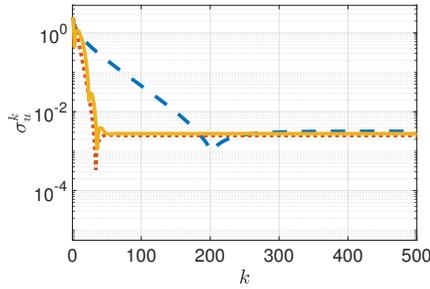
**Fig. 1.4** Numerical experiments with  $a = 2.5$ ,  $s_0 = 0$ ,  $s_f = 0$ ,  $v_0 = 1$ ,  $v_f = 0$ , and the number of discretization subintervals  $N = 2000$ . The graphs in (a)–(c) show approximations of the optimal control function with Algorithms 1–3, after  $k = 3, 5, 15$  iterations, with  $\varepsilon = 10^{-8}$ . All algorithms are observed to converge to the optimal solution indicated by  $k \rightarrow \infty$ , in various rates. The semi-log graphs in (d) and (e) show the  $L^\infty$  errors in the state and control variables, respectively, in each iteration of the three algorithms.



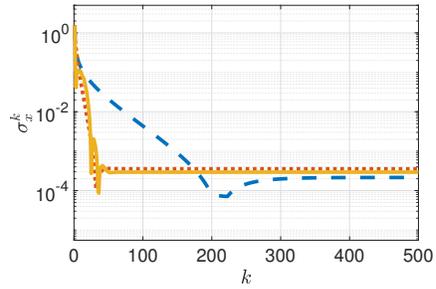
(a)  $L^\infty$ -error in control with  $N = 10^3$ .



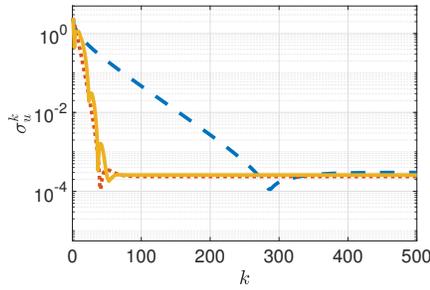
(b)  $L^\infty$ -error in states with  $N = 10^3$ .



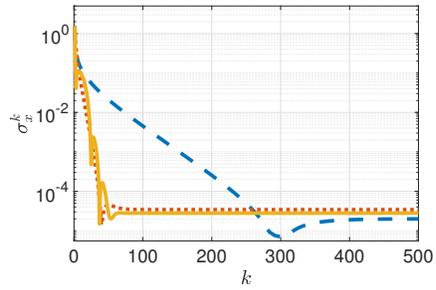
(c)  $L^\infty$ -error in control with  $N = 10^4$ .



(d)  $L^\infty$ -error in states with  $N = 10^4$ .



(e)  $L^\infty$ -error in control with  $N = 10^5$ .



(f)  $L^\infty$ -error in states with  $N = 10^5$ .

**Fig. 1.5** Numerical experiments with  $a = 2.5$ ,  $s_0 = 0$ ,  $s_f = 0$ ,  $v_0 = 1$ ,  $v_f = 0$ . The semi-log graphs show the  $L^\infty$  errors in the state and control variables, respectively, in each iteration of the three algorithms, with various  $N$  from coarse ( $N = 1000$ ) to fine ( $N = 100000$ ).

**Table 1.1** Least errors that can be achieved by Algorithms 1–3 and Ipopt, with  $\varepsilon = 10^{-8}$ .

$N$	Dykstra	DR	AAC	Ipopt
$10^3$	$3.2 \times 10^{-2}$	$2.5 \times 10^{-2}$	$2.8 \times 10^{-2}$	$3.2 \times 10^{-2}$
$10^4$	$3.2 \times 10^{-3}$	$2.5 \times 10^{-3}$	$2.8 \times 10^{-3}$	$7.7 \times 10^{-3}$
$10^5$	$3.0 \times 10^{-4}$	$2.4 \times 10^{-4}$	$2.6 \times 10^{-4}$	$1.6 \times 10^{-2}$

(a)  $L^\infty$ -error in control,  $\sigma_u^k$ .

$N$	Dykstra	DR	AAC	Ipopt
$10^3$	$2.2 \times 10^{-3}$	$3.6 \times 10^{-3}$	$3.0 \times 10^{-3}$	$2.2 \times 10^{-3}$
$10^4$	$2.1 \times 10^{-4}$	$3.6 \times 10^{-4}$	$2.9 \times 10^{-4}$	$2.3 \times 10^{-4}$
$10^5$	$2.0 \times 10^{-5}$	$3.4 \times 10^{-5}$	$2.8 \times 10^{-5}$	$8.7 \times 10^{-5}$

(b)  $L^\infty$ -error in states,  $\sigma_x^k$ .

Table 1.1 displays the values of the errors, separately in  $u$  and  $x$ , after the stopping criteria with  $\varepsilon = 10^{-8}$  was satisfied, for each of the three algorithms. A precise 10-fold reduction in error with a 10-fold increase in  $N$  can be verified with these numbers, as discussed in the previous paragraph. We have added the experiments we have carried out with Ipopt, version 3.12, an interior point optimization software [30], which solved the direct Euler-discretization of Problem (P), with the same values of  $N$  and the same tolerance  $10^{-8}$ . Ipopt, running with linear solver MA57, was paired up with the optimization modelling language AMPL [20]. The same 10-fold decreases in error cannot be observed with Ipopt, unless one sets the tolerance for Ipopt to be much smaller than  $10^{-8}$ , say  $10^{-14}$  (which also means longer computational times). With the tolerance set at  $10^{-14}$ , the error values with Ipopt becomes pretty much the same as those with Dykstra (still with  $\varepsilon = 10^{-8}$ ), which is interesting to note.

As we pointed out earlier, the same errors listed in Table 1.1 can be achieved with bigger stopping thresholds. For  $N = 10^3, 10^4, 10^5$ , respectively: with  $\varepsilon = 10^{-6}, 10^{-6}, 10^{-7}$ , Algorithm 1 converges in 281, 359 and 454 iterations; with  $\varepsilon = 10^{-5}, 10^{-5}, 10^{-7}$ , Algorithm 2 converges in 65, 50 and 101 iterations; with  $\varepsilon = 10^{-4}, 10^{-5}, 10^{-6}$ , Algorithm 3 converges in 49, 60 and 70 iterations.

In Table 1.2, the CPU times (in seconds) each algorithm takes, with the respective  $\varepsilon$  values listed above, are tabulated. Note that Algorithms 1–3 have been coded and run on Matlab, 64-bit (maci64) version R2017b. All software, including AMPL and Ipopt, were run on MacBook Pro, with operating system macOS Sierra version 10.12.6, processor 3.3 GHz Intel Core i7 and memory 6 GB 2133 MHz LPDDR3. In Table 1.2, the CPU times for Ipopt are listed with the tolerance  $10^{-14}$ , since with only this fine tolerance it is possible to obtain the same order of the error magnitudes as those obtained by Algorithms 1–3. With  $\varepsilon = 10^{-8}$ , the CPU times for Ipopt are

**Table 1.2** CPU times [sec] taken by Algorithms 1–3 and Ipopt. For  $N = 10^3, 10^4, 10^5$ , respectively:  $\varepsilon = 10^{-6}, 10^{-6}, 10^{-7}$  for Algorithm 1,  $\varepsilon = 10^{-5}, 10^{-5}, 10^{-7}$  for Algorithm 2, and  $\varepsilon = 10^{-4}, 10^{-5}, 10^{-6}$  for Algorithm 3, have been used. The tolerance for Ipopt was set as  $10^{-14}$ .

$N$	Dykstra	DR	AAC	Ipopt
$10^3$	0.03	0.01	0.01	0.08
$10^4$	0.16	0.05	0.05	0.71
$10^5$	1.6	0.41	0.28	7.3

0.06, 0.45 and 4.4 seconds, respectively, which are significantly higher than the times taken by Algorithms 1–3, in addition to worse errors.

Numerical observations suggest two joint winners: Algorithms 2 and 3, i.e., the Douglas–Rachford method and the Aragón Artacho–Campoy algorithm, in both accuracy and speed.

## 1.6 Conclusion and Open Problems

We have applied three well-known projection methods to solve an optimal control problem, i.e., control-constrained minimum-energy control of double integrator. We have derived the projectors for the optimal control problem and demonstrated that they can be used in Dykstra’s algorithm, the Douglas–Rachford (DR) method and the Aragón Artacho–Campoy (AAC) algorithm, effectively. We carried out extensive numerical experiments for an instance of the problem and concluded that the DR and AAC algorithms (Algorithms 2 and 3) were jointly the most successful. We also made comparisons with the standard discretization approach, only to witness the benefit of using projection methods.

It is interesting to note that when we apply alternating projections, we also seem to converge to  $P_{\mathcal{A} \cap \mathcal{B}}(0)$  even though this is not supported by existing theory.

To the best of authors’ knowledge, the current paper constitutes the first of its kind which involves projection methods and continuous-time optimal control problems. It can be considered as a prototype for future studies in this direction. Some of the possible directions are listed as follows.

- The setting we have introduced could be extended to general control-constrained linear-quadratic problems.
- We have used some discretization of the projector as well as the associated IVP in (1.34)–(1.38). This might be extended to problems in more general form. On the other hand, for the particular problem we have dealt with in the present paper, one might take into account the fact that if  $u^-(t)$  is piecewise linear then its projection is piecewise linear. This might simplify further the expressions given in Proposition 1.

- Although theory for projection methods can in principle vouch convergence only for convex problems, it is well-known that the DR method can be successful for nonconvex problems, see, for example, [10]. It would be interesting to extend the formulations in the current paper to nonconvex optimal control problems.
- For a certain value of an algorithmic parameter, Figure 3 exhibits downward spikes. It would be interesting to see if this phenomenon is also observed in other control-constrained optimal control problems, as well as under other stopping criteria.

## Appendix

### Algorithm 1b (Dykstra-b)

Steps 1–4 (*Initialization*) Do as in Steps 1–4 of Algorithm 1.

Step 5 (*Stopping criterion*) If  $\|u^{k+1} - u^k\|_{L^\infty} \leq \varepsilon$ , then return  $u^{k+1}$  and stop. Otherwise, set  $k := k + 1$  and go to Step 2.

### Algorithm 2b (DR-b)

Step 1 (*Initialization*) Choose a parameter  $\lambda \in ]0, 1[$  and the initial iterate  $u^0$  arbitrarily. Choose a small parameter  $\varepsilon > 0$ , and set  $k = 0$ .

Step 2 (*Projection onto  $\mathcal{A}$* ) Set  $u^- = \lambda u^k$ . Compute  $\tilde{u} = P_{\mathcal{A}}(u^-)$  by using (1.9).

Step 3 (*Projection onto  $\mathcal{B}$* ) Set  $u^- := 2\tilde{u} - u^k$ . Compute  $\hat{u} = P_{\mathcal{B}}(u^-)$  by using (1.22).

Step 4 (*Update*) Set  $u^{k+1} := u^k + \hat{u} - \tilde{u}$ .

Step 5 (*Stopping criterion*) If  $\|u^{k+1} - u^k\|_{L^\infty} \leq \varepsilon$ , then return  $\tilde{u}$  and stop. Otherwise, set  $k := k + 1$  and go to Step 2.

### Algorithm 3b (AAC-b)

Step 1 (*Initialization*) Choose the initial iterate  $u^0$  arbitrarily. Choose a small parameter  $\varepsilon > 0$ , two parameters  $\alpha$  and  $\beta$  in  $]0, 1[$ , and set  $k = 0$ .

Step 2 (*Projection onto  $\mathcal{A}$* ) Set  $u^- = u^k$ . Compute  $\tilde{u} = P_{\mathcal{A}}(u^-)$  by using (1.9).

Step 3 (*Projection onto  $\mathcal{B}$* ) Set  $u^- = 2\beta\tilde{u} - u^k$ . Compute  $\hat{u} = P_{\mathcal{B}}(u^-)$  by using (1.22).

Step 4 (*Update*) Set  $u^{k+1} := u^k + 2\alpha\beta(\hat{u} - \tilde{u})$ .

Step 5 (*Stopping criterion*) If  $\|u^{k+1} - u^k\|_{L^\infty} \leq \varepsilon$ , then return  $\tilde{u}$  and stop. Otherwise, set  $k := k + 1$  and go to Step 2.

## References

1. Alt, W., Kaya, C. Y., Schneider, C.: Dualization and discretization of linear-quadratic control problems with bang–bang solutions. *EURO J. Comput. Optim.* **4**, 47–77 (2016)

2. Alwadani, S., Bauschke, H. H., Moursi, W. M., Wang, X.: On the asymptotic behaviour of the Aragon Artacho-Campoy algorithm. <https://arxiv.org/abs/1805.11165v1>.
3. Aragón Artacho, F. J., Campoy, R.: A new projection method for finding the closest point in the intersection of convex sets. *Comput. Optim. Appl.* **69**, 99–132 (2018)
4. Aragón Artacho, F. J., Campoy, R.: Computing the resolvent of the sum of maximally monotone operators with the averaged alternating modified reflections algorithm. <https://arxiv.org/abs/1805.09720v1>.
5. Ascher, U. M., Mattheij, R. M. M., Russell, R. D.: *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*. SIAM Publications, Philadelphia (1995)
6. Bauschke, H. H., Combettes, P. L.: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Second edition. Springer (2017)
7. Bauschke, H. H., Koch, V. R.: Projection methods: Swiss Army knives for solving feasibility and best approximation problems with halfspaces. *Infinite Products of Operators and Their Applications*, 1–40 (2012)
8. Bauschke, H. H., Moursi, W. M.: On the order of the operators in the Douglas–Rachford algorithm. *Optimization Letters* **10**, 447–455 (2016)
9. Bauschke, H. H., Moursi, W. M.: On the Douglas–Rachford algorithm. *Math. Program., Ser. A* **164**, 263–284 (2017)
10. Borwein, J. M., Sims, B.: The Douglas-Rachford Algorithm in the absence of convexity. In: Bauschke H., Burachik R., Combettes P., Elser V., Luke D., Wolkowicz H. (eds) *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. Springer Optimization and Its Applications, vol 49, pp.93-109. Springer, New York, NY.
11. Boyle J. P., Dykstra R. L.: A method for finding projections onto the intersection of convex sets in Hilbert spaces. In: *Advances in Order Restricted Statistical Inference*, vol. 37, Lecture Notes in Statistics, 28–47. Springer (1986)
12. Burachik, R. S., Kaya, C. Y., Majeed, S. N.: A duality approach for solving control-constrained linear-quadratic optimal control problems. *SIAM J. Control Optim.* **52**, 1771–1782 (2014)
13. Combettes, P. L.: A block-iterative surrogate constraint splitting method for quadratic signal recovery. *IEEE Trans. Sig. Proc.* **51**, 2432–2442 (2003)
14. Combettes, P. L.: Iterative construction of the resolvent of a sum of maximal monotone operators. *J. Convex Anal.* **16**, 727–748 (2009)
15. Dontchev, A. L., Hager, W. W.: The Euler approximation in state constrained optimal control. *Math. Comp.* **70**, 173–203 (2001)
16. Dontchev, A. L., Hager, W. W., Malanowski, K.: Error bound for Euler approximation of a state and control constrained optimal control problem. *Numer. Funct. Anal. Optim.* **21**, 653–682 (2000)
17. Douglas, J., Rachford, H. H.: On the numerical solution of heat conduction problems in two and three space variables. *Trans. Amer. Math. Soc.* **82**, 421–439 (1956)
18. Eckstein, J., Bertsekas, D. P.: On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program., Ser. A* **55**, 293–318 (1992)
19. Eckstein, J., Ferris, M. C.: Operator-splitting methods for monotone affine variational inequalities, with a parallel application to optimal control. *INFORMS J. Comput.* **10**, 218–235 (1998)
20. Fourer, R., Gay, D. M., Kernighan, B. W.: *AMPL: A Modeling Language for Mathematical Programming*, Second Edition. Brooks/Cole Publishing Company / Cengage Learning (2003)
21. Hestenes, M. R.: *Calculus of Variations and Optimal Control Theory*. John Wiley & Sons, New York (1966)
22. Kaya, C. Y., Lucas, S. K., Simakov, S. T.: Computations for bang–bang constrained optimal control using a mathematical programming formulation. *Optim. Contr. Appl. Meth.* **25**, 295–308 (2004)
23. Kaya, C. Y., Noakes, J. L.: Computations and time-optimal controls, *Opt. Cont. Appl. Meth.* **17**, 171–185 (1996)
24. Kaya, C. Y., Noakes, J. L.: Computational algorithm for time-optimal switching control. *J. Optim. Theory App.* **117**, 69–92 (2003)
25. Lions, P.-L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. *SIAM J. Numer. Anal.* **16**, 964–979 (1979)

26. O'Donoghue, B., Stathopoulos, G., Boyd, S.: A splitting method for optimal control. *IEEE Trans. Contr. Sys. Tech.* **21**, 2432–2442 (2013)
27. Rugh, W. J.: *Linear System Theory*, 2nd Edition. Pearson (1995)
28. Stoer, J., Bulirsch, R.: *Introduction to Numerical Analysis*, 2nd Edition. Springer-Verlag, New York (1993)
29. Svaiter, B. F.: On weak convergence of the Douglas-Rachford method. *SIAM J. Control Optim.* **49**, 280–287 (2011)
30. Wächter, A., Biegler, L. T.: On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. *Math. Progr.* **106**, 25–57 (2006)