ELSEVIER

Journal of Computational and Applied Mathematics 124 (2000) 361–371

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.nl/locate/cam

Optimal control

R.W.H. Sargent *

Centre for Process Systems Engineering, Imperial College, Prince Consort Road, London SW7 2BY, UK

Received 20 August 1999; received in revised form 25 November 1999

Abstract

This paper gives a brief historical survey of the development of the theory of the calculus of variations and optimal control, and goes on to review the different approaches to the numerical solution of optimal control problems.

Keywords: Historical review; Optimal control theory; Calculus of variations; Numerical solutions

1. Introduction

Optimal control theory is an outcome of the calculus of variations, with a history stretching back over 360 years, but interest in it really mushroomed only with the advent of the computer, launched by the spectacular successes of optimal trajectory prediction in aerospace applications in the early 1960s.

Fortunately, Goldstine [27] has written an excellent treatise on the early history, and there have been three later publications [10,59,64] carrying the story up to the present day. We therefore give only an outline of the main steps in the historical development of the theory, then focus on the development of numerical techniques for solution.

2. A brief history of the theory

Some geometrical optimization problems were known and solved in classical times, such as the line representing the shortest distance between two points, or the "isoperimetric problem": the shape of the plane curve of given length enclosing the largest area. However our story really begins with Galileo, who in 1638 posed two shape problems: the shape of a heavy chain suspended between

* Tel.: +0171-594-6604; fax: +0171-594-6606.
E-mail address: r.w.h.sargent@ic.ac.uk (R.W.H. Sargent).

0377-0427/00/$ - see front matter
PII: S0377-0427(00)00418-0

two points (the catenary), and the shape of a wire such that a bead sliding along it under gravity traverses the distance between its end-points in minimum time (the brachistochrone). Later, in 1662, Fermat postulated the principle that light always chooses the path through a sequence of optical media such that it traverses them in minimum time.

Galileo's conjectures on the solutions of his two problems were incorrect, and Newton in 1685 was the first to solve a shape problem — the nose shape of a projectile providing minimum drag — though he did not publish the result until 1694.

In 1696 Johann Bernoulli challenged his contemporaries to solve the brachistochrone problem by the end of the year. Five mathematicians responded to the challenge: Johann's elder brother Jakob, Leibnitz, l'Hopital, Tschirnhaus and Newton. Bernoulli published all their solutions, together with his own, in April 1697.

The competition aroused interest in this type of problem and there followed a period of activity by a number of mathematicians. The resulting ideas were collected in a book [25] published in 1744 by Euler, a student of Bernoulli working in Basel, who remarked "nothing at all takes place in the universe in which some rule of maximum or minimum does not appear".

In essence, Euler formulated the problem in general terms as one of finding the curve $x(t)$ over the interval $a \leqslant t \leqslant b$, with given values $x(a), x(b)$, which minimizes

$$J = \int_a^b L(t, x(t), \dot{x}(t))\,\mathrm{d}t \tag{1}$$

for some given function $L(t, x, \dot{x})$, where $\dot{x} \equiv \mathrm{d}x/\mathrm{d}t$, and he gave a necessary condition of optimality for the curve $x(\cdot)$

$$\frac{\mathrm{d}}{\mathrm{d}x} L_{\dot{x}}(t, x(t), \dot{x}(t)) = L_x(t, x(t), \dot{x}(t)), \tag{2}$$

where the suffix x or $\dot{x}$ implies the partial derivative with respect to x or $\dot{x}$.

Up to this point the solution techniques had been essentially geometric, but in a letter to Euler in 1755, Lagrange described an analytical approach, based on perturbations or "variations" of the optimal curve and using his "undetermined multipliers", which led directly to Euler's necessary condition, now known as the "Euler–Lagrange equation". Euler enthusiastically adopted this approach, and renamed the subject "the calculus of variations".

The Euler–Lagrange equation, based on first-order variations, yields only a stationarity condition, and it was Legendre in 1786 (see [33]), who studied the second variation and produced a second-order necessary condition of optimality. Legendre derived it for the scalar case, but it was later extended to the vector case by Clebsch and is now known as the Legendre–Clebsch condition

$$L_{\dot{x}\dot{x}}(t, x(t), \dot{x}(t)) \geqslant 0, \quad t \in [a, b], \tag{3}$$

interpreted as requiring the matrix to be nonnegative definite along the optimal trajectory.

Meanwhile Hamilton [28], through his "principle of least action", had been reformulating the equations of mechanics as a variational principle. He introduced the function, now known as the "Hamiltonian function"

$$H(t, y, x) = \langle y, \dot{x} \rangle - L(t, x, \dot{x}), \tag{4}$$

where

$$y(t, x, \dot{x}) = L_{\dot{x}}(t, x, \dot{x}) \tag{5}$$

and in (4) $\dot{x}$ is obtained as a function of (t, y, x) by solving Eq. (5). It follows immediately that

$$\dot{x}(t) = H_y(t, y(t), x(t)), \quad \dot{y}(t) = -H_x(t, y(t), x(t)) \tag{6}$$

if and only if the Euler–Lagrange Eq. (2) is satisfied.

Much later Caratheodory [16] showed that (5) is indeed always solvable for $\dot{x}$ in a neighbourhood defined by weak variations (those for which perturbations δx and $\delta\dot{x}$ are both small) of a regular optimal trajectory (one for which the strict Legendre–Clebsch condition holds). Under these conditions H is well defined and twice continuously differentiable, and $H_{yy}(t, y, x)$ is positive definite if and only if $L_{\dot{x}\dot{x}}(t, x, \dot{x})$ is positive definite.

Hamilton expressed his principle in terms of a pair of partial differential equations, but in 1838 Jacobi showed that it could be more compactly written in terms of what is now known as the Hamilton–Jacobi equation:

$$\phi_t(t, x) + H(t, \phi_x(t, x), x(t)) = 0. \tag{7}$$

If $\phi(t, x)$ is a twice continuously differentiable solution of this equation, then Eq. (6) defines a regular optimal trajectory, provided that

(a) The strict Legendre–Clebsch condition (i.e. $H_{yy}(t, y(t), x(t)) > 0$) is satisfied.
(b) There are no points conjugate [1] to a along the trajectory (the Jacobi condition).

The next step was taken by Weierstrass [66], who considered strong variations (with δx small but no restriction on $\delta\dot{x}$). He considered the special case where $L(t, x, \dot{x})$ is a positively homogeneous function not depending explicitly on t, but with no loss of generality, since any $L(t, x, \dot{x})$ can be transformed to this form. He introduced the "excess function":

$$E(t, x, \dot{x}, u) = L(t, x, u) - L(t, x, \dot{x}) - (u - \dot{x})L_{\dot{x}}(t, x, \dot{x}) \tag{8}$$

with $x(t), \dot{x}(t)$ evaluated along the optimal trajectory.

Again Caratheodory [16] showed that $E > 0$ if and only if $H_{yy} > 0$, thus confirming the sufficiency of the Hamilton–Jacobi solution even under strong variations.

Given the basic problem (1), it was natural to require $x(\cdot)$ to be differentiable on (a, b), and to consider minimization over all such curves, but the development of measure theory allowed the interpretation of (1) as a Lebesgue integral and the relaxation of $x(\cdot)$ to be absolutely continuous on (a, b), incidently providing a closure property for the family of functions J defined for different $\dot{x}(\cdot)$.

Caratheodory was well aware of the need to establish existence of optimal trajectories, and established this through the rather strong sufficient conditions for the existence of the requisite solutions of the Hamilton–Jacobi equation. However, the new viewpoint allowed Tonelli [61] to address directly the problem of existence of optimal trajectories for problem (1). His proof required the convexity of the function $L(t, x, \cdot)$ for all t, x, and a growth condition of the type

$$|L(t, x, \dot{x})| \geqslant \alpha.\|\dot{x}\|^2 - \beta \tag{9}$$

for some positive constants α and β.

[1] Conjugacy is defined with respect to solutions of an associated partial differential equation.

The next step was to consider the restriction of the class of admissible functions $\dot{x}(\cdot)$ to a subset of R^n, specifically so that they also satisfy the set of equations

$$g(t,x(t),\dot{x}(t)) = 0, \quad t \in [a,b], \tag{10}$$

which we would today recognize as a general set of differential–algebraic equations. However, sufficient conditions were imposed to ensure that there exist functions $\dot{x}(t) = f(t,x(t))$ satisfying (10) with sufficient degrees of freedom to ensure the existence of neighbouring functions also satisfying (10).

The resulting problem is known as the problem of Lagrange, since it was solved by the use of Lagrange multipliers, and the complete solution can be found in [6,42].

This set the scene for parameterizing the degrees of freedom implicit in the Lagrange problem by considering constraints of the form

$$\dot{x}(t) = f(t,x(t),u(t)), \tag{11}$$

where the parameters $u(t)$ or "controls" can be chosen at each $t \in (a,b)$, possibly restricted to some fixed subset $\Omega \subset \mathbb{R}^m$, yielding the "optimal control problem":

Find $u(\cdot)$ on (a,b) to minimize

$$J = \int_a^b L(t,x(t),\dot{x}(t))\,\mathrm{d}t$$

subject to

$$\dot{x}(t) = f(t,x(t),u(t)), \quad t \in (a,b), \quad u(t) \in \Omega \subset \mathbb{R}^m, \quad t \in (a,b), \quad x(a) \text{ and } x(b) \text{ given.} \tag{12}$$

The necessary conditions of optimality for this problem were established by Pontryagin [54] in his famous "maximum principle", which can be expressed in the form

$$\begin{aligned} &\dot{x}(t) = H'_y(t,y(t),x(t),\dot{x}(t)), \qquad \dot{y}(t) = -H'_x(t,y(t),x(t),\dot{x}(t)),\\ &H'(t,y(t),x(t),\dot{x}(t)) = \max_{u\in\Omega} H'(t,y(t),x(t),u). \end{aligned} \tag{13}$$

The function H' in (13) is not the classical Hamiltonian, but what Clarke [18] later termed the "pseudo-Hamiltonian", still defined by (4) but with $\dot{x}$ as an independent argument instead of being defined as a solution of (5). However, it is easy to see that the Hamilton equations (6) still hold for the pseudo-Hamiltonian if and only if the Euler–Lagrange equation (2) holds, though it is now more natural to express the Legendre–Clebsch condition in the form

$$H'_{\dot{x}\dot{x}}(t,y(t),x(t),\dot{x}(t)) \leqslant 0, \quad t \in (a,b). \tag{14}$$

In fact, Pontryagin et al. [54] considered several variants and extensions of the basic problem (12), and from this point there was an increasing avalanche of publications.

Many workers tackled the problem of pure state inequalities along the trajectory of the form

$$g(t,x(t)) \geqslant 0, \quad t \in (a,b) \tag{15}$$

and this work is well summarized in the recent review in [29], which also describes some open questions remaining in this area. More recently, Sargent [56] has presented necessary conditions of optimality for systems described by a mixed set of general differential–algebraic equations and inequalities.

Roxin [55] extended the classical Tonelli existence theorem to more general problems like (12), and at about the same time Warga [65] showed that convexity of the "extended velocity set" could be dropped by extending the class of admissible controls to include "relaxed" or chattering controls, which allow rapid oscillation between two points of this set, thus effectively replacing the set by its convex hull.

Recently, Bell et al. [4] extended Roxin's result to infinite horizon problems, using the "strong optimality criterion", in the ordinary sense that the integral in (12) remains finite as $b \to \infty$, and is minimized by the optimal control. However, in many cases the integral can become infinite, and successive weakenings of the optimality criteria have been proposed.

Overtaking optimality:

$$\limsup_{b\to\infty} [J_b(\hat{x},\hat{u}) - J_b(x,u)] \leqslant 0.$$

Weakly overtaking optimality:

$$\liminf_{b\to\infty} [J_b(\hat{x},\hat{u}) - J_b(x,u)] \leqslant 0.$$

Finite optimality:

$$J_b(\hat{x},\hat{u}) - J_b(x,u) \leqslant 0$$

for all (x,u) defined on (a,b) such that $x(b) = \hat{x}(b)$, for all $b > a$.

All these optimality criteria are treated in some detail in the book in [17].

Two more general formulations of the problem have been much studied, mainly by Rockafellar, Clarke, Loewen, Mordukhovich and Vinter:

(a) *The differential inclusion problem*:

$$\text{Minimize } J = l(x(a),x(b)) + \int_a^b L(t,x(t),\dot{x}(t))\,\mathrm{d}t$$

over absolutely continuous functions $x(\cdot)$ on (a,b) subject to:

$$\dot{x}(t) \in F(t,x(t)), \quad t \in (a,b) \text{ a.e., } (x(a),x(b)) \in C \subset \mathbb{R}^{2n}.$$

Here F is a multifunction, mapping $[a,b] \times \mathbb{R}^n$ into subsets of $\mathbb{R}^n$.

(b) *The generalized Bolza problem*:

$$\text{Minimize } J = l(x(a),x(b)) + \int_a^b L(t,x(t),\dot{x}(t))\,\mathrm{d}t$$

over all absolutely continuous functions $x(\cdot)$ on $[a,b]$.

This deceptively simple form can subsume a wide variety of problems by allowing l and L to take infinite values and defining them appropriately, for example by setting l or L to infinity whenever their arguments fail to satisfy the constraints.

Study of these problems is accompanied by a relaxation of the differentiability conditions necessary in the classical results, requiring the techniques of nonsmooth analysis. To state specific results would take us too far into technicalities, and reference should be made to the textbooks in [18,19,64], and to the survey paper in [20] for further details.

However, in two recent papers [36,37] Loewen and Rockafellar give the latest results on necessary conditions for the generalized Bolza problem, first for fixed values of a and b, and in the later paper for variable values.

Another important issue has been the quest for conditions under which the sufficient conditions obtained via the Hamilton–Jacobi equation are also necessary. Here Clarke and Vinter [21] have shown in the context of the differential inclusion problem that the weakest condition is "local calmness", which is implied by "strong normality", so that the existence of a generalized solution of the Hamilton–Jacobi equation is both necessary and sufficient for optimality for all "reasonable" problems.

Finally, Zeidan [67] has given a complete treatment of second-order necessary and sufficient conditions for problem (12) with additional mixed state-control inequalities along the trajectory. The paper includes some interesting insights on the concept of strong normality, and a connection between the Jacobi condition and solutions of a Riccati equation.

3. Numerical solution techniques

Compared with the intricacies and subtleties of the theoretical development, the history of numerical solution techniques is relatively straightforward. There are essentially three approaches to solve these problems:

1. Solution of the two-point boundary value problem given by the necessary conditions, with solution of the local Hamiltonian optimization problem at each time-step.
2. Complete discretization of the problem, converting it into a finite-dimensional nonlinear program.
3. Finite parameterization of the control trajectory, again converting the problem into a nonlinear program, but with the objective and constraint functions evaluated by integration of the system equations, and their gradients with respect to the control parameters by integration of the adjoint equations or sensitivity equations.

The early numerical methods tackled problems without control or end-point constraints, which of course still yielded a two-point boundary value problem, with given initial values for the system equations and given terminal values for the adjoint system. Bryson [12] and Breakwell [8] used the "shooting method" for solving this problem, guessing the unknown initial values of the adjoint variables, integrating both system and adjoint equations forward, and re-estimating the initial guesses from residuals at the end-point.

Difficulties arose from the extreme sensitivity, even instability, of the solutions to the initial guesses, which led to the "multiple-shooting" technique of Bulirsch and his co-workers [14,58], who subdivided the time interval and re-estimated starting values for each subinterval from the mismatches. An alternative approach, due to Miele and co-workers [43,44] was quasilinearization, in which the system equations were linearized about the current trajectory and integrated forward in time (along with the adjoint equations) for sets of initial conditions spanning the space of the unknown initial values. A linear combination of these solutions was then computed to satisfy the conditions at both initial and final times, and the system re-linearized about the new trajectory.

At almost the same time as these early attempts, Kelley [31] proposed a gradient method, essentially the simplest member the class of control parameterization techniques. He estimated control values on a closely spaced fixed grid, and used this grid to integrate the system equations. He then

integrated the adjoint system backwards from the known values at the final time and used the results to obtain the gradients with respect to the control values and hence a correction to the estimated control values.

Of course steepest descent methods have slow final convergence rate, and to speed this up several workers used techniques based on second variations [9,30,32]. At the same time others used numerical acceleration techniques from the parallel developments in nonlinear programming: Lasdon et al. [34] incorporated conjugate gradients while Pollard and Sargent [53] used quasi-Newton approximations. The latter authors also pointed out that the parameterization of the controls need not be tied to the integration step and used piecewise-constant or piecewise-linear controls on a coarser grid, yielding a much smaller optimization problem.

In all these methods, control and terminal constraints were at first dealt with using penalty functions, but control constraints (usually simple bounds) were later treated by projection. A special difficulty arises with end-point equality constraints, because there is then a corresponding number of undetermined terminal values for the adjoint variables and of course the terminal $x(b)$ values from the integration do not necessarily satisfy the terminal equality constraints. Again, the problem was dealt with by penalty functions, but not very effectively. Obviously, the quasilinearization approach provides a general technique for dealing with an arbitrary mix of initial and final equations and inequalities, and Bryson and Ho [11] describe a similar technique for generating a set of adjoint systems, subsequently linearly combined to reduce the residuals of the end-point equality constraints. They also describe "min-H" algorithms which determine the correction to the controls by performing a local constrained optimization of the Hamiltonian at each time-step of the shooting methods. In fact, this textbook [11] provides an excellent summary of the state of the art at the end of the 1960s, except that it completely ignores the development of the complete discretization approach.

Of course, nonlinear programming techniques were also in their infancy over this period, so complete discretization was late on the scene. Early proposals were made in [41,52], and the state of the art was probably summed up by the textbook of Canon et al. [15].

From this point, all three approaches had their followers and were improved using advances in the enabling technologies of nonlinear programming and integration of ordinary differential equations and differential–algebraic equations.

Early methods of complete discretization used finite differences, but Tsang et al. [62] introduced collocation methods, while Biegler and co-workers [22,38,60] developed this approach, showing how to incorporate error measures as constraints to decide on the number and placing of additional nodes in the "collocation on finite elements" method. A parallel development was carried out in [2,5] in the aerospace arena, and recently, Dennis et al. [23] described an interior-point SQP method using a trust-region algorithm.

Sargent and co-workers [57,49,63,51] developed the control parameterization approach, extending it to deal with multistage systems, high-index DAEs and state inequalities, while Bock and his school [7,35] have similarly developed a hybrid approach, with some elements of all three approaches. He starts with a control parameterization, then generates a finite-dimensional nonlinear programme by integrating the system independently over each control subinterval, as in multiple shooting.

Bulirsch and his school [14,58,50,13] and Maurer [39,40,1] have continued to develop the multiple shooting approach to solving the two-point boundary value problem, particularly in relation to dealing with inequality state constraints along the trajectory, while Miele and co-workers

have pursued the quasilinearization approach [43–47] linking it with multiple shooting (which they call the "multipoint" approach). Dixon and Bartholomew-Biggs [24] proposed an adjoint-control transformation to help with the sensitivity of the shooting technique, and recently, Fraser-Andrews [26] has combined this with multiple shooting.

Efforts to deal with state inequalities along the trajectory were hampered in the early days by inadequate understanding of the theory, but it is still a formidable problem because of the possibility of discontinuities in the adjoint variables at the "junction-points", where the inequalities become active or cease to be active. The complete discretization approach appears to deal with the problem, since these constraints become ordinary nonlinear constraints of the nonlinear programme, but the discretization destroys the hidden additional constraints arising from differentiations of the original constraints for high-index problems, so a fine discretization is necessary to be sure of satisfying them. Pantelides et al. [51] achieved some success in the control parameterization approach by converting these constraints into end-point constraints by integrating the violations, coupled with a finite set of interior-point constraints, but strictly the problem is still nonsmooth, and the two-point boundary value approach remains the only safe method.

Maurer [39,40,1] has pioneered the judicious use of some pre-analysis of the problem to determine the structure of junction-points, coupled with multiple shooting, and the power of this approach is well demonstrated by the successful solution of the difficult "wind-shear" problem by Bulirsch et al. [50]. However, such solutions require real expertise.

Very recently, Bell and Sargent [3] have sought to avoid the difficulties by conversion of the inequalities to equalities using slack variables and elimination of the resulting bounds by an interior-point approach, converting the optimal control problem into a smooth two-point boundary problem for a DAE system. Preliminary results indicate surprising success in obtaining close approximation of this smooth problem to the original nonsmooth one.

Finally, although the emphasis in this survey has been on optimal control for general nonlinear systems, it would not be complete without a mention of the spectacular successes of the use of discrete-time linear models with quadratic objective function in on-line control applications. This formulation particularly suits the process industries, and requires the solution of a small finite-dimensional quadratic programme on-line, which can easily be solved within the requisite time. Fortunately, a comprehensive review of the relevant issues has recently been published [48], to which the reader is referred for further details.

4. Concluding comments

As already noted, this survey has been limited to optimal control of general nonlinear systems, essentially described by ordinary differential equations or differential–algebraic equations. Although there has been a recent upsurge of interest in control of distributed-parameter systems, it would widen the scope too far to attempt to include this area. Similarly, little has been said of the extensive work on linear systems, and nothing at all on the effects of uncertainty. Each of these would require a separate review of at least the same length. Even so the scope is still enormous and many facets have been left unexplored. It is virtually impossible to keep abreast of all developments, and the selection of topics must inevitably reflect a personal viewpoint.

References

[1] D. Augustin, H. Maurer, Stability and sensitivity analysis of optimal control problems subject to pure state constraints, presented at the 16th International Symposium on Mathematical Programming, Lausanne, August, 1997.

[2] T. Bauer, J.T. Betts, W. Hallman, W.P. Huffman, K. Zondervan, Solving the optimal control problem using a nonlinear programming technique, Parts I and II, AIAA Papers 84-2037 and 84-2038, 1984.

[3] M.L. Bell, R.W.H. Sargent, Optimal control of inequality-constrained DAE systems, Centre for Process Systems Engineering Report No. 98-B , Imperial College, 1998. Comput. Chem. Eng., accepted for publication.

[4] M.L. Bell, R.W.H. Sargent, R.B. Vinter, Existence of optimal controls for continuous-time infinite horizon problems, Int. J. Control 68 (1997) 887–896.

[5] J.T. Betts, W.P. Huffman, Application of sparse nonlinear programming to trajectory optimization, J. Guidance Control Dynamics 15 (1) (1992) 198–206.

[6] G.A. Bliss, The problem of Lagrange in the calculus of variations, Am. J. Math. 52 (1930) 673–744.

[7] H.G. Bock, K.J. Plitt, A multiple-shooting algorithm for direct solution of optimal control problems, Proceedings of the Nineth IFAC World Congress, Budapest, Pergamon Press, Oxford, 1984.

[8] J.V. Breakwell, The optimization of trajectories, SIAM J. 7 (1959) 215–247.

[9] J.V. Breakwell, J. Speyer, A.E. Bryson, Optimization and control of nonlinear systems using the second variation, SIAM J. Control Ser. A 1 (2) (1963) 193–223.

[10] A.E. Bryson Jr., Optimal control – 1950 to 1985, IEEE Control Systems (1996) 26–33.

[11] A.E. Bryson, Y-C. Ho, Applied Optimal Control, Blaisdell, Waltham, MA, 1969.

[12] A.E. Bryson, S.E. Ross, Optimum rocket trajectories with aerodynamic drag, Jet Propulsion (1958).

[13] R. Bulirsch, F. Montrone, H. J. Pesch, Abort landing in the presence of a windshear as a minimax optimal control problem: Part 1: Necessary conditions, Part 2: Multiple shooting and homotopy, Technische Universität Munchen, Report No.210, 1990.

[14] R. Bulirsch, Die Mehrzielmethode zur numerischen Losung von nichtlinearen Randwertproblemen und Aufgaben der optimalen Steuerung, Report of the Carl-Cranz Gesellschaft, DLR, Oberpfaffenhofen, 1971.

[15] M.D. Canon, C.D. Cullum, E. Polak, Theory of Optimal Control and Mathematical Programming, McGraw-Hill, New York, 1970.

[16] C. Caratheodory, Variationsrechnung und partielle Differentialgleichungen erste Ordnung, Teubner, Berlin, 1935 (Translation: Calculus of Variations and Partial Differential Equations of the First Order, R. Dean, Holden-Day, San Francisco, 1967).

[17] D.A. Carlson, A. Haurie, Infinite Horizon Optimal Control, Springer, Berlin, 1987.

[18] F.H. Clarke, Optimization and Non-smooth Analysis, Wiley, New York, 1983 (Reprinted by SIAM, Philadelphia, PA, 1990).

[19] F.H. Clarke, Methods of Dynamic and Non-smooth Optimization, SIAM, Philadelphia, 1989.

[20] F.H. Clarke, Y.S. Ledyaev, R.J. Stern, P.R. Wolenski, Qualitative properties of trajectories of control systems: a survey, J. Dynamics Control Systems 1 (1995) 1–48.

[21] F.H. Clarke, R.B. Vinter, Local optimality conditions and Lipschitzian solutions to the Hamilton–Jacobi equation, SIAM J. Control Optim. 21 (1983) 856–870.

[22] J.E. Cuthrell, L.T. Biegler, On the optimization of differential–algebraic process systems, A.I.C.h.E. J. 33 (1987) 1257–1270.

[23] J.E. Dennis, M. Heinkenschloss, L.N. Vicente, Trust-region interior-point algorithms for a class of nonlinear programming problems, SIAM J. Control Optim. 36 (1998) 1750–1794.

[24] L.C.W. Dixon, M.C. Bartholomew-Biggs, Adjoint-control transformations for solving practical optimal control problems, Optim. Control Appl. Methods 2 (1981) 365–381.

[25] L. Euler, Methodus Inveniendi Lineas Curvas Maximi Minimive Propriatate Gaudientes, sive Solutio Problematis Isoperimetrici Latissimo Sensu Accepti, Bousquent, Lausannae and Genevae, 1744.

[26] G. Fraser-Andrews, A multiple-shooting technique for optimal control, J. Optim. Theory Appl. 102 (1999) 299–313.

[27] H.H. Goldstine, A History of the Calculus of Variations from the 17th to the 19th Century, Springer, New York, 1981.

[28] W.R. Hamilton, Mathematical Papers, Cunningham Memoir No. XIII, Cambridge University Press, Cambridge, 1931.

[29] R.F. Hartl, S.P. Sethi, R.G. Vickson, A survey of the maximum principles for optimal control problems with state constraints, SIAM Rev. 37 (2) (1995) 181–218.
[30] D.H. Jacobson, D.Q. Mayne, Differential Dynamic Programming, Elsevier, New York, 1970.
[31] H.J. Kelley, Gradient theory of optimal flight paths, J. Amer. Rocket Soc. 30 (1960).
[32] H.J. Kelley, R.E. Kopp, G. Moyer, A trajectory optimization technique based upon the theory of the second variation, presented at the AIAA Astrodynamics Conference, New Haven, CT, 1963.
[33] J.L. Lagrange, Théorie des Fonctions Analytiques, Paris, 1797.
[34] L.S. Lasdon, S.K. Mitter, A.D. Waren, The conjugate gradient methods for optimal control problems, IEEE Trans. Automat Control AC-12 (1967) 132–138.
[35] D. Leineweber, H.G. Bock, J.P. Schloder, J.V. Gallitzendorfer, A. Schafer, P. Jansohn, A boundary value problem approach to the optimization of chemical processes described by DAE models, University of Heidelberg, Preprint 97-14 (SFB359), 1997.
[36] P.D. Loewen, R.T. Rockafellar, New necessary conditions for the generalized problem of Bolza, SIAM J. Control Optim. 34 (1996) 1496–1511.
[37] P.D. Loewen, R.T. Rockafellar, Bolza problems with general time constraints, SIAM J. Control Optim. 35 (1997) 2050–2069.
[38] J.S. Logsdon, L.T. Biegler, Accurate solution of differential–algebraic optimization problems, Ind. Eng. Chem. Res. 28 (1989) 1628–1639.
[39] H. Maurer, Optimale Steuerprozesse mit Zustandsbeschrankungen, Habilitations-schrift, University of Wurzburg, 1976.
[40] H. Maurer, W. Gillessen, Application of multiple shooting to the numerical solution of optimal control problems with bounded state variables, Computing 15 (1975) 105–126.
[41] D.Q. Mayne, A second-order gradient method for determining optimal trajectories of nonlinear discrete-time systems, Int. J. Control 3 (1) (1966) 85–95.
[42] E.J. McShane, On multipliers for Lagrange problems, Amer. J. Math. 61 (1939) 809–819.
[43] A. Miele, Method of particular solutions for linear two-point boundary-value problems, J. Optim. Theory Appl. 2 (4) (1968).
[44] A. Miele, Gradient methods in optimal control theory, in: M. Avriel, M.J. Rijckaert, D.J. Wilde (Eds.), Optimization and Design, Prentice-Hall, Englewood Cliffs, NJ, 1973, pp. 323–343.
[45] A. Miele, R.R. Iyer, General technique for solving nonlinear two-point boundary-value problems via the method of particular solutions, J. Optim. Theory Appl. 5 (1970) 382–399.
[46] A. Miele, T. Wang, Parallel computation of two-point boundary-value problems via particular solutions, J. Optim. Theory Appl. 79 (1993) 5–29.
[47] A. Miele, K.H. Well, J.L. Tietze, Multipoint approach to the two-point boundary-value problem, J. Math. Anal. Appl. 44 (1973) 625–642.
[48] M. Morari, J.H. Lee, Model predictive control: past, present and future, Comput. Chem. Eng. 23 (1999) 667–682.
[49] K.R. Morison, R.W.H. Sargent, Optimization of multi-stage processes described by differential–algebraic equations, in: J.P. Hennart (Ed.), Numerical Analysis, Springer, Berlin, 1986.
[50] H.J. Oberle, W. Grimm, BNDSCO – a program for the numerical solution of optimal control problems, Internal Report No. 515-89/22, Institute for Flight Systems Dynamics, DLR, Oberpfaffenhofen, 1989.
[51] C.C. Pantelides, R.W.H. Sargent, V.S. Vassiliadis, Optimal control of multi-stage systems described by high-index differential–algebraic equations, in: R. Bulirsch, D. Kraft (Eds.), Computational Optimal Control, Birkhauser, Basel, 1994.
[52] E. Polak, On primal and dual methods for solving discrete optimal control problems, in: L.A. Zadeh, L.W. Neustadt, A.V. Balakrishnan (Eds.), Computing Methods in Optimization Problems, Vol. 2, Academic Press, New York, 1969, pp. 317–331.
[53] G.M. Pollard, R.W.H. Sargent, Off-line computation of optimum controls for a plate distillation column, Automatica 6 (1970) 59–76.
[54] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, E.F. Mischenko, The Mathematical Theory of Optimal Processes (Translation by L.W. Neustadt), Wiley, New York, 1962.
[55] E. Roxin, The existence of optimal controls, Michigan Math. J. 9 (1962) 109–119.

[56] R.W.H. Sargent, Necessary conditions for optimal control of inequality-constrained DAE systems, Centre for Process Systems Engineering Report No. B98-17, Imperial College, 1998.
[57] R.W.H. Sargent, G.R. Sullivan, The development of an efficient optimal control package, in: J. Stoer (Ed.), Proceedings of the Eighth IFIP Conference on Optimization Techniques, Part 2, Springer, Heidelberg, 1978.
[58] J. Stoer, R. Bulirsch, Introduction to Numerical Analysis, Springer, New York, 1980.
[59] H.J. Sussmann, J.C. Willems, 300 years of optimal control: from the brachystochrone to the maximum principle, IEEE Control Systems (1997) 32–44.
[60] P. Tanartkit, L.T. Biegler, A nested simultaneous approach for dynamic optimization problems, Comput. Chem. Eng. 20 (1996) 735–741.
[61] L. Tonelli, Sull' existenza del minimo in problemi di calcolo delle variationi, Ann. R. Scuola Normale Sup. Pisa, Sci. Fis. Mat. 1 (1931) 89–100.
[62] T.H. Tsang, D.M. Himmelblau, T.F. Edgar, Optimal control via collocation and nonlinear programming, Int. J. Control 21 (1975) 763–768.
[63] V.S. Vassiliadis, R.W.H. Sargent, C.C. Pantelides, Solution of a class of multi-stage dynamic optimization problems: 1, problems without path constraints, 2, problems with path constraints, I and EC Res. 33 (1994) 2111–2122, 2123–2133.
[64] R.B. Vinter, Optimal Control, Birkhauser, Boston, 1999.
[65] J. Warga, Relaxed variational problems, J. Math. Anal. Appl. 4 (1962) 111–128.
[66] K. Weierstrass, Werke, Bd. 7, Vorlesungen uber Variationsrechnung, Akademisches Verlagsgesellschaft, Leipzig, 1927.
[67] V. Zeiden, The Riccati equation for optimal control problems with mixed state-control constraints: necessity and sufficiency, SIAM J. Control Optim. 32 (1994) 1297–1321.