

THE NATURE OF ADJOINT VARIABLES AND THEIR ROLE IN OPTIMAL PROBLEMS

O. K. CROSSER

University of Missouri-Rolla
Rolla, MO 65401

ADJOINT VARIABLES ARE frequently arbitrarily introduced into the textbook discussion of optimal or extremal theory. For example, Bryson and Ho [1] "adjoin" them to the optimization problem, Denn [2] introduces them as a "convenience," and Leitman [11] regards them as a transformation to a "useful" vector space basis. Only Jackson [10] has shown that they are desirable as a general transformation from one set of variables which appear naturally during the formulation of the problem to the set of interest in the solution search problems. Adjoint variables are the sensitivity coefficients in optimal search problems. Adjoint variables exist because the coefficient matrix of every system (of describing equations) has a transpose, and there are, therefore, two independent solutions to the homogeneous form of the system.

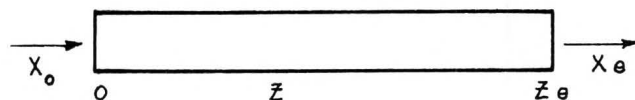
However, it was the late Professor F. M. J. Horn who in 1958 most directly presented the fundamental nature of the adjoint variables and their role in optimal reactors in chemical engi-



O. K. Crosser received his PhD. from Rice University in 1955 and is currently professor of chemical engineering at the University of Missouri-Rolla. His primary interests include optimization of preliminary plant design and fixed bed separations.

neering [5]. The original papers [6, 7, 8] and the more complete elaboration in his thesis were in German, with results published in English [9] by 1967. Publications about Pontryagin's Principle [8, 10] became the standard literature reference, and the directness of Horn's approach became less available for the beginning student to appreciate. Furthermore, this appreciation or understanding of adjoint variables makes much of Horn's later work in optimal chemical reactors—effect of bypassing, cyclical operation of non-linear processes—much easier to follow.

This demonstration makes use of the example presented in detail in appendix I.11 of his Thesis [5]. One asks for the optimal temperature profile for a plug flow reactor with several independent chemical reactions. The set of independent chemical reactions is any set of the smallest number of time dependent stoichiometric equations sufficient to define all reaction compositions uniquely in time. The proper interpretation of independent is also clearly discussed in this thesis, although a more formidable presentation is now available [12]. We suppose a tubular plug flow reactor with several chemical reactions and arbitrary kinetics (Arrhenius)



$$dx_i/dz = v_i(x^1, x^2, \dots, x^m, T), \quad (i = 1 \text{ to } m) \quad (1)$$

and presume that the objective function has the form

$$M = M(x^{1e}, x^{2e}, \dots, x^{me}) = M(x_e) \quad (2)$$

M depends only upon the exit composition (extents) x_e and a straightforward solution to the problem would be to assume a temperature profile, calculate the exit composition vector to give

M, then presume another temperature profile and continue to adapt the profile in some beneficial way until an extreme in M was obtained. Suppose we had two such solutions (we use x for the vector of extents of the independent reactions and V for the vector of reaction rates, and the super dot to imply differentiation with respect to z)

$$x_1 \rightarrow \dot{x}_1 = V(x_1, T_1) \quad \text{and} \quad x_2 \rightarrow \dot{x}_2 = V(x_2, T_2) \quad (3)$$

so that for sufficiently small differences between T_1 and T_2 a first difference (perturbation) is sufficient. Then

$$\dot{x}_2 - \dot{x}_1 = [\partial V_i / \partial x_j](x_2 - x_1) + (\partial V_i / \partial T)(T_2 - T_1)$$

or

$$\dot{y} = [\partial V_i / \partial x_j]y + (\partial V_i / \partial T)\bar{T} \quad (4a)$$

where y stands for the perturbation in x caused by the perturbation \bar{T} in T . We will also have the perturbed response

$$\begin{aligned} \bar{M} = M(x_{2e}) - M(x_{1e}) &= \sum_{k=1}^m \frac{\partial M}{\partial x_{ke}} y_k(z_e) \\ &= (\partial M / \partial x_{ie})' y(z_e) \end{aligned} \quad (4b)$$

Since both x_2 and x_1 are zero at z equal zero, $y(0)$ is zero. Note that the matrix, $[\partial V_i / \partial x_j]$ and vector, $(\partial V_i / \partial T)$ are functions of z only, so that the system of Eqs. (4a,b) is a set of linear differential equations in which the coefficients are functions only of the independent variable z . y is the response of the system to \bar{T} . What we desire is to solve Eqs. (4a,b) subject to the restriction that M be an extreme, so that it is necessary that

$$dM = (\partial M / \partial x_{ie})' dx = 0 \quad (5)$$

$(\partial M / \partial x_{ie})'$ is the transpose of the vector of partial derivatives of M with respect to x_{ie} , that is, x at the end of the reactor. A system like (4) is usually solved by Variation of Parameters (Boyce & DePrima [3] or Hochstadt [4]), finding first the homogeneous (complementary) solutions. The form of these solutions is more conveniently manipulated if we use the solution matrix rather than the solution vector (in contrast with usual forms for systems with constant coefficients). Therefore

$$\dot{y}_c = [B] y_c \quad \text{where} \quad B_{ij} = \partial V_i / \partial x_j \quad (6)$$

has the homogeneous solution matrix $[Y]$ such that

$$[\dot{Y}] = [B][Y] \quad (7)$$

and we take $[Y(z_e)]$ to be $[I]$ the identity matrix. Any other boundary condition on $[Y]$ may be obtained directly from this one. Using the Variation of Parameters we suppose $y = [Y]c$ and hope to find the vector c to fit the inhomogeneous part, which is the second term of Eq. (4)

$$\dot{y} = \{[Y]\dot{c}\} = \{\dot{Y}\}c + [Y]\dot{c} = [B][Y]c + (\partial V_i / \partial T)\bar{T} \quad (8)$$

Substituting from Eq. (6) we have

$$[Y]\dot{c} = (\partial V_i / \partial T)\bar{T} + \dot{c} = [Y]^{-1}(\partial V_i / \partial T)\bar{T} \quad (9)$$

and we see that the vector of the particular solutions c is directly related to the temperature profile, \bar{T} . Now, these functions depend only on z , and we intend to keep the same inlet temperature but to alter the shape of the profile. Therefore any one of the particular solutions c must have the property

$$c_k(z_e) = \int_0^{z_e} \dot{c}_k(z) dz \quad \text{with} \quad c_k(0) = 0$$

and since

$$y = [Y]c \quad \text{and} \quad [Y] = [I] \quad \text{at} \quad z = z_e$$

then

$$y(z_e) = c(z_e) \quad (10)$$

Then from Eq. 4b, using t as a dummy variable and recalling that $(\partial M / \partial x_{ie})$ does not depend on z

$$\begin{aligned} \bar{M} &= \int_0^{z_e} \dot{\bar{M}} dt = \sum_{k=1}^m (\partial M / \partial x_{ke}) c_k(z_e) \\ &= \int_0^{z_e} \sum_{k=1}^m (\partial M / \partial x_{ke}) c_k(t) dt \end{aligned}$$

then

$$\dot{\bar{M}} = \sum_{k=1}^m (\partial M / \partial x_{ke}) \dot{c}_k(z_e) = (\partial M / \partial x_{ie})' \dot{c} \quad (11)$$

(Note that $\dot{\bar{M}}$ is the derivative, with respect to the independent variable of the response \bar{M} to the perturbation \bar{T} .)

We now have to solve simultaneously $m + 1$ linear equations involving \dot{c} . There are m independent chemical reactions, and Eq. (11) for $\dot{\bar{M}}$. Since these equations must be linearly de-

Thus, the differential equations for the adjoint variables and their corresponding conditions at the end of the reactor show that the influence of the exit extents upon the value of the objective function can be obtained for any entering conditions to the reactor by integrating their adjoint variable differential equations from the end of the reactor to the entrance.

pendent, their determinant must be zero

$$\begin{vmatrix} \dot{M} & (\partial M / \partial x_{ie})' \\ (\partial v_i / \partial T) \bar{T} & [Y] \end{vmatrix} = 0 \quad \text{or} \quad \begin{vmatrix} \dot{M} & u' \\ w & [Y] \end{vmatrix} = 0$$

using vectors u and w for notational convenience, then

$$\dot{M} |Y| + \begin{vmatrix} 0 & u' \\ w & [Y] \end{vmatrix} = 0 \quad \text{and} \quad \dot{M} = - \frac{\begin{vmatrix} 0 & u' \\ w & [Y] \end{vmatrix}}{|Y|} \quad (12)$$

As we expand the numerator determinant of Eq. (12) first about w , the first column (deleting the i th row in Y) and then about u' in the first row (deleting the j th column in Y), we will obtain the cofactors of the elements Y_{ij} , in Y , which we label a_{ij} and Eq. (12) can be written

$$\dot{M} = - \frac{\begin{vmatrix} 0 & u' \\ w & Y \end{vmatrix}}{|Y|} = \frac{w' [a_{ij}] u}{|Y|} = \frac{w [\text{Adj}(Y')] u}{|Y|} = w' \lambda \quad (13)$$

because $[a_{ij}]$ is the adjoint matrix of the transpose of $[Y]$, (a sign change occurs as the $i + 1$ index in the determinant decreases to the i index for w). The adjoint variables, λ , are defined by Eq. (13) and

$$\lambda = [Y']^{-1} u + [Y'] \lambda = u$$

Since

$$\dot{u} = (\partial M / \partial x_{ie}) = 0$$

(for the extreme in M)

$$\{[Y'] \lambda\} = 0 = [\dot{Y}'] \lambda + [Y'] \dot{\lambda} + \dot{\lambda} = - [Y']^{-1} [\dot{Y}'] \lambda$$

But

$$[\dot{Y}'] = [B][Y] + [\dot{Y}'] \lambda = \{[B][Y]\}' = [Y'] [B']$$

then

$$\dot{\lambda} = - [Y']^{-1} [Y'] [B'] \lambda \quad \text{and} \quad \dot{\lambda} = - [B'] \lambda$$

These are the differential equations for the adjoint variables. The boundary conditions of $[Y] = [I]$ at $z = z_e$ imply (for $m = 2$ for clarity) because I_{ij} is the unit ij cofactor from the identity

$$\lambda_1(z_e) = \frac{\begin{vmatrix} \partial M / \partial x_1 & \partial M / \partial x_2 \\ 0 & 1 \end{vmatrix}}{|I|}$$

$$\lambda_2(z_e) = - \frac{\begin{vmatrix} \partial M / \partial x_1 & \partial M / \partial x_2 \\ 1 & 0 \end{vmatrix}}{|I|}$$

Hence

$$\lambda(z_e) = \begin{vmatrix} u' \\ I_{ij} \end{vmatrix} = u = (\partial M / \partial x_{ie})$$

matrix.

Thus, the differential equations for the adjoint variables and their corresponding conditions at the end of the reactor show that the influence of the exit extents upon the value of the objective function can be obtained for any entering conditions to the reactor by integrating their adjoint variable differential equations from the end of the reactor to the entrance. These functions, therefore, explain how the optimal result is affected by changing the values of the extents of reaction at any point along the reactor such as the entrance. Since there is a direct correspondence between length in a plug flow reactor and time, it is equally clear how the adjoint variables apply to time optimization as well.

The adjoint variables are therefore nothing more than the additional homogeneous solution for the linear perturbation. Had the problem been cast in the form of time optimal control, they would have indicated the switching functions; introduced with an Hamiltonian or Lagrange multiplier problem, they would have been the corresponding multipliers [13].

The thing to see is that all of these structures rely essentially only on a Cramer's rule for solving a dependent set of linear equations and that the adjoint variables appear naturally as the added homogeneous solutions to the transpose of the system coefficient matrix, and they show how temperature changes along the reactor affect the objective function, which depends on the conversion at the exit from the reactor.

ACKNOWLEDGMENTS

One of the original reviewers of this article observed that the subject of this paper is contained in modern control theory texts. It is a pleasure to suggest to students that *Linear Systems* by Thomas Kailath (Prentice-Hall 1980) is an excellent reference with good examples and exercises. The most directly relevant part is section 9.1 pp 598-606 and example 9.1-3 p. 605, but there are many other items of interest throughout the entire text.

The University of Missouri-Rolla awarded the Faculty Sabbatical during which this note was written. D. W. T. Rippin and his Systems Engineering Group of ETH Zurich provided the affectionate welcome and gentle scholarly support. Don MacElroy offered a most helpful suggestion toward the end of the work.

LITERATURE CITED

1. Bryson, A. E., Y. Ho, *Applied Optimal Control*, Halsted Press, New York (1975) [chapter 2, and pp 47-50, pp 149-150].
2. Denn, M., *Optimization by Variational Methods*, McGraw-Hill, New York (1969), [pp 102-109].
3. DiPrima, R. C., W. E. Boyce, *Elementary Differential Equations and Boundary Value Problems*, John Wiley & Sons, New York (1977).
4. Hochstadt, H., *Differential Equations*, Dover Press, New York (1975).
5. Horn, R. "Optimalprobleme bei kontinuierlichen chemischen Prozessen," Thesis, Tech. Hochsch. Wien, Ostereich (1958).
6. Horn, F., Discussion of "Optimum Temperature Sequences in Reactors," K. G. Denbigh, *Chem. Eng. Sci.*, Special Supplement, 8, 131 (1958).
7. Horn, F., U. Troltenier, "uber den Optimalen Temperatur verlauf im Reaktionsrohr," *Chem. Ing. Tech.*, 32, 382 (1960).
8. Horn, G., "Adjungierte Variable und Maximum prinzip in der Theorie Chemischer Reaktoren," *Osterreichische Chemiker-Zeitung*, 6, 186 (1967).
9. Horn, F., "Mathematical Models in the Design and Development of Chemical Reactors," *Ber. Buns. Ges.*, 74, 81-89, (1970) (in English).
10. Jackson, R., "Optimization of Chemical Reactors With Respect to Flow Configuration," *J. Theo. App. Opt.*, 2, 240-259 (1968).
11. Leitman, G., *An Introduction to Optimal Control*, McGraw-Hill, New York (1966), [p 26, Chapter 2].
12. Smith, W. R., R. W. Missen, *Chemical Reaction Equilibrium Analysis*, John Wiley & Sons, New York (1982), [Chapter 2—also see the reference by these authors in *Chem. Eng. Educ.* which contains a review of the literature to 1976].
13. Report: SEG/R/128 (83), Systems Engineering Group, Tech. Chem. Labor, ETH, Zurich.

POSITIONS AVAILABLE

Use CEE's reasonable rates to advertise. Minimum rate
1/2 page \$60; each additional column inch \$25.

MICHIGAN STATE UNIVERSITY

Chemical Engineering . . . Tenure system faculty positions. Doctorate in Chemical Engineering or closely related field. A strong commitment to teaching and the ability to develop a quality research program is expected. Preference will be given to candidates with research interests in the areas of Biochemical Engineering, Surface Science, Solid State Phenomena, or Polymeric Materials. However, applicants with outstanding credentials and research interests in other fields related to Chemical Engineering are encouraged to apply. Teaching and/or industrial experience desirable but not essential. Michigan State University is an affirmative action-equal opportunity employer and welcomes applications from women and minority groups. Send applications and names of references to Chairperson, Faculty Search Committee, Department of Chemical Engineering, Michigan State University, East Lansing, Michigan 48824-1226.

NOMENCLATURE

		Determinant
[]	Matrix (square)
()	Vector (column)
()'	Transposed vector (row)
z		Length of plug flow reactor
x		Extent of reaction
V		Vector of reaction rates
y		Perturbation in x
T		Temperature
M		Objective function to be optimized
m		Number of independent chemical reactions
B		Coefficient matrix from partial derivatives of rates V
Y		Matrix of homogeneous solutions to Eq. 4a
I		Identity matrix
c		Vector of particular solutions for Eq. 4a
u'		Row vector ($\partial M / \partial x_{ie}$)'
w		Column vector ($\partial v_i / \partial T$) \bar{T}
λ		Vector of adjoint variables

Subscripts

o	Entrance to reactor
e	Exit from reactor
i,j,k	Row column indexes

Superscripts

i	Index to independent chemical reactions (1 to m)
.	Differentiation with respect to length
-1	Inverse matrix
'	Transpose
—	Indicates perturbation value