

ornl

**OAK RIDGE
NATIONAL
LABORATORY**

MARTIN MARIETTA

OPERATED BY
MARTIN MARIETTA ENERGY SYSTEMS, INC.
FOR THE UNITED STATES
DEPARTMENT OF ENERGY



3 4456 0069232 9

ORNL/TM-10058

**Deterministic Sensitivity Theory
for Nonlinear Systems:
The Forward and Adjoint Methods**

Dan G. Cacuci

OAK RIDGE NATIONAL LABORATORY
CENTRAL RESEARCH LIBRARY
CIRCULATION SECTION
4500 ROSS 173
LIBRARY LOAN COPY
DO NOT TRANSFER TO ANOTHER PERSON
If you wish someone else to see this
report, send in name with report and
the library will arrange a loan.

Printed in the United States of America. Available from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Road, Springfield, Virginia 22161
NTIS price codes—Printed Copy: A05 Microfiche A01

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Engineering Physics and Mathematics Division

**DETERMINISTIC SENSITIVITY THEORY FOR NONLINEAR SYSTEMS:
THE FORWARD AND ADJOINT METHODS**

Lecture Notes for the ISPRA Course "Data Uncertainties,
Sensitivities, Consistency, and Adjustment"

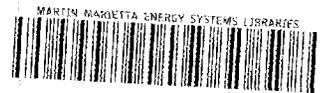
Dan G. Cacuci

Date Issued: August 1986

Research sponsored by
U.S. DOE Office of
Energy Research

*Joint Research Centre of the Commission of the European Communities, Ispra, Italy, April 14-18, 1986.

Prepared by the
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, Tennessee 37831
operated by
Martin Marietta Energy Systems, Inc.
Under Contract No. DE-AC05-84OR21400
for the
U.S. Department of Energy



3 4456 0069232 9

TABLE OF CONTENTS

I. INTRODUCTION	1
II. MATHEMATICAL BACKGROUND	5
II.A. VECTOR SPACE	5
II.B. NORMED VECTOR SPACE	6
II.C. BANACH SPACE (COMPLETE NORMED VECTOR SPACE)	6
II.D. HILBERT SPACE	7
II.E. SOME BASIC TERMINOLOGY OF OPERATOR THEORY	8
II.F. SOME ASPECTS OF DIFFERENTIAL CALCULUS IN VECTOR SPACES	10
III. RIGOROUS MATHEMATICAL FORMULATION OF THE FORWARD AND ADJOINT METHODS OF SENSITIVITY THEORY FOR NONLINEAR SYSTEMS	13
III.A. NONLINEAR SYSTEMS WITH OPERATOR-TYPE RESPONSES	13
III.A.1. Mathematical Representation of the Physical Problem	13
III.A.2. Sensitivity Theory	14
<i>III.A.2.a. The Forward Method</i>	15
<i>III.A.2.b. The Adjoint Method</i>	16
<i>III.A.2.b.(i). System Responses: Functionals</i>	18
<i>III.A.2.b.(ii). System Responses: Operators</i>	19
III.A.3. Comparative Discussion of Previous Approaches to Sensitivity Analysis	22
III.B. NONLINEAR SYSTEMS WITH RESPONSES DEFINED AT CRITICAL POINTS	26
III.B.1. Introduction	26
III.B.2. System Response: A Functional Defined at a Critical Point of a Function of the System's State Vector and Parameters	26
III.B.3. The Forward Method	28
III.B.4. The Adjoint Method	30
III.B.5. Discussion	33
III.B.6. Illustrative Application: Sensitivity Analysis of Extremum-Type Responses in Reactor Safety	35
<i>III.B.6.a. Problem Description</i>	35
<i>III.B.6.b. Sensitivity Analysis: Theory</i>	37
<i>III.B.6.c. Sensitivity Analysis: Results</i>	48

<i>III.B.6.d. Summary and Discussion of Results</i>	53
<i>III.B.6.e. Appendix A: Mathematical Representation of the MELT-IIIB Model</i>	54
<i>III.B.6.f. Appendix B: The G-differential and G-derivative of a Nonlinear Operator in the MELT-IIIB Model</i>	57
<i>III.B.6.g. Appendix C: Representation of the Adjoint System</i>	60
<i>III.B.6.h. Nomenclature</i>	63
III.C. NONLINEAR SYSTEMS WITH FEEDBACK	65
III.C.1. Introduction	65
III.C.2. Definition and Estimation of Sensitivity to Feedback	65
IV. SUMMARY AND CONCLUSIONS	72
V. REFERENCES	76

ABSTRACT

Concepts of nonlinear functional analysis are employed to investigate the mathematical foundations underlying sensitivity theory. This makes it possible not only to ascertain the limitations inherent in existing analytical approaches to sensitivity analysis, but also to rigorously formulate a considerably more general sensitivity theory for physical problems characterized by systems of nonlinear equations and by nonlinear functionals as responses. Two alternative formalisms, the Forward Method and the Adjoint Method, are developed in order to evaluate the sensitivity of the response to variations in the system parameters. The forward method is formulated in normed linear spaces, and the existence of the Gâteaux differentials of the operators appearing in the problem is shown to be both necessary and sufficient for its validity. This method is conceptually straightforward and can be advantageously used to assess the effects of relatively few parameter alterations on many responses. On the other hand, for problems involving many parameter alterations or a large data base and comparatively few functional-type responses, the alternative adjoint method is computationally more economical. However, it is shown that this method can be developed only under conditions that are more restrictive than those underlying the validity of the forward method. In particular, the requirement that operators acting on the state vector and on the system parameters must admit densely defined Gâteaux derivatives is shown to be of fundamental importance for the validity of this formalism. The present analysis significantly extends the scope of sensitivity theory and provides a basis for still further generalizations.

There are physical systems where a critical point of a function that depends on the system's state vector and parameters defines the location in phase-space where the response functional is evaluated. The Gâteaux differentials giving the sensitivities of both the functional and the critical point to changes in the system's parameters are obtained by alternative formalisms. The forward method is the simpler and more general, but may be prohibitively expensive for problems with large data bases. The adjoint method, although less generally applicable and requiring several adjoint calculations, is likely to be the only practical approach. Sensitivity theory is also extended to include treatment of general operators, acting on the system's state vector and parameters, as response. In this case, the forward method is the same as for functional responses, but the adjoint method is considerably different. The adjoint method requires expanding the indirect effect term, an element of a Hilbert space, in terms of elements of an orthonormal basis. Since as many calculations of adjoint functions are required as there are nonzero terms in this expansion, careful consideration of truncating the expansion is needed to assess the advantages of the adjoint method over the forward method.

The sensitivity theory for nonlinear systems with responses defined at critical points of a function of the system's state variables and parameters is applied to a protected transient with scram on high-power level in the Fast Flux Test Facility. The single-phase segment of the fast reactor safety code MELT-IIIB is used to model this transient. The response analyzed is the maximum fuel temperature in the hot channel. For the purposes of sensitivity analysis, a complete characterization of this response requires consideration of both the numerical value of the response at the maximum, and the location in phase space where the maximum occurs. This is because variations in the system parameters alter not only the value at this maximum but also alter the location of the maximum in phase space.

Expressions for the sensitivities of the numerical value of this maximum-type response and expressions for the sensitivities of the phase-space location at which the maximum occurs are derived in terms of adjoint functions. The adjoint systems satisfied by each of these adjoint functions are derived and solved. It is shown that the complete sensitivity analysis requires (a) the computation of as many adjoint functions as there are nonzero components of the maximum in phase space, and (b) the computation of one additional adjoint function for evaluating the sensitivities of the numerical value of the response. The same computer code can be used to calculate all the required adjoint functions; once these adjoint functions are available, the sensitivities to *all* possible variations in the system parameters are obtained by quadratures. The sensitivities obtained by this efficient method are used to predict both changes in the numerical values of the maximum fuel temperature and the new phase-space location at which the perturbed maximum occurs when the system parameters are varied. These predictions are shown to agree well with direct recalculations using perturbed parameter values.

Finally, we present an efficient method for calculating the sensitivity of a mathematical model's result to feedback. Feedback is defined in terms of an operator acting on the model's dependent variables. The sensitivity to feedback is defined as a functional derivative, and a method is presented to evaluate this derivative using adjoint functions. Typically, this method allows the individual effect to many different feedbacks to be estimated with a total additional computing time comparable to only one recalculation. It is anticipated that this method of estimating the effect of feedback will be useful for more complex models where extensive recalculations for each of a variety of different feedbacks is impractical.

I. INTRODUCTION

Modelling complex physical phenomena has led to the development of a variety of large computer code systems. The large number of physical effects treated in these codes and their high running costs make these codes ill-suited to applications of a parametric or survey nature. To address important and difficult tasks such as assessing confidence levels and uncertainties in calculated design parameters and determining the effects of changes in the input data on the results of complex calculations, it is essential to have a sensitivity analysis methodology that can efficiently treat the complex systems -- with many parameters -- encountered in practice.

The simplest and perhaps the most common procedure for sensitivity analysis of a code consists of varying selected input parameters, rerunning the code, and recording the corresponding changes in the results (i.e., responses) calculated by the code. The model parameters responsible for the largest relative changes in the responses are then classified to be the most important. For complex models, though, the large amount of computing time needed by such recalculations severely restricts the scope of this sensitivity analysis procedure. In practice, this means that the modeler can investigate only a few parameters that he judges *a priori* to be important.

A way of investigating sensitivities to more parameters is to consider simplified models obtained by developing fast-running approximations to complex processes. Although this makes rerunning less expensive, the parameters must still be selected *a priori*, and consequently important sensitivities may be missed. Also, it is difficult to demonstrate that the respective sensitivities of the simplified and complex models are the same.

To obtain as much information as possible from a limited number of recalculations, statistical techniques have been used to develop sensitivity and uncertainty analysis methods known as "response surface methods." The use of response surface methods involves (a) selection of a small number of model parameters that are thought by the modeler to be important, (b) a strategy to select design points, in the space of model parameters, at which the computer code calculates the responses of interest to sensitivity/uncertainty analysis, (c) recalculations using altered parameter values, (d) use of these recalculated results to construct "response surfaces," which are simple approximations representing the behavior of the response as a function of the chosen model parameters, (e) use of the response surface thus obtained to replace the original model for subsequent statistical studies (such as Monte Carlo and moment matching) to estimate sensitivities and uncertainty distributions for the responses.

The application of response surface methods to sensitivity/uncertainty analysis of computer codes is conceptually straightforward and requires relatively little developmental work. Therefore, the use of response surface methods has gained popularity in several application areas. But despite progress towards reducing the number of recalculations used to map the response surfaces, the response surface methods remain expensive and limited in scope since, in practice, (a) the number of model parameters is very large, so only a small subset can be selected for sensitivity/uncertainty analysis, (b) information about data importance is required *prior* to initiating the analysis, (c) the data importance is largely unknown, and a considerable probability of missing important effects exists, and (d) sensitivities can only be estimated, but not calculated exactly.

As the need systematic sensitivity/uncertainty analysis gained recognition, other sensitivity analysis techniques were developed -- most notably in conjunction with applications to chemical kinetics, system theory, and reactor physics and shielding. For example, three sensitivity analysis methods developed in conjunction with applications in chemical kinetics and system theory are the

Fourier Amplitude Sensitivity Test (acronym FAST), the direct method, and the Green's function method. The FAST technique gives statistical mean values and standard deviations for model responses, but requires a very large number of calculations (despite its acronym, it is not a fast running technique).

In contrast to the response surface method and the FAST technique, the direct method and Green's function method are deterministic (rather than statistical) methods. The direct method, for example, involves differentiation of the equations describing the model with respect to a parameter. The resulting set of equations is solved for the derivatives of all the model variables with respect to a parameter. The resulting set of equations is solved for the derivative of all the model variables with respect to that parameter. Note that the actual form of the differentiated equations depends on the parameter under consideration. Consequently, for each parameter a different set of equations must be solved to obtain the corresponding sensitivity.

For models that involve a large number of parameters and comparatively few responses, sensitivity analysis can be performed very efficiently by using deterministic methods based on adjoint functions. The use of adjoint functions for sensitivity and analyses appeared as early as the 1940's; these analyses were either based on perturbation theory or based on variational approaches. In reactor theory, for example, the first use of perturbation theory is attributed¹ to Wigner,² while the variational principles are considered³ to have evolved from works of Levine and Schwinger⁴ and Roussopolos.⁵ The scope of both the variational formulation and the perturbation theory approach has subsequently been generalized and extended (see e.g., Refs. 6-12). The great potential of adjoint-function based approaches to sensitivity analysis of several linear problems encountered in reactor theory has been demonstrated in the comprehensive reviews given by Stacey³ and Greenspan.¹² These successes have generated considerable interest in extending and applying such approaches to sensitivity analysis of several inherently nonlinear problems in other areas.¹³⁻¹⁷ Higher-order perturbation theories have also been proposed¹⁸ for sensitivity analysis of neutronics problems involving linear operators. Developments through 1979 in adjoint-operator based approaches to sensitivity and uncertainty analyses have been comprehensively reviewed in Ref. 19.

To date, several alternative theoretical approaches to adjoint-based sensitivity equations have evolved, the three prominent being:

1. variational approaches,^{3,7,9,10,14}
2. perturbation theory approaches, including "generalized perturbation theory,"^{6,8,12,13,16,18}
3. differential approaches.^{15,17}

All of these approaches have been focused on deriving expressions for the sensitivities of the system responses (i.e., system performance parameters) to changes in the input parameters. The system responses considered in these approaches have been particular forms of functionals, and the sensitivities have been defined as the derivatives of these responses with respect to the input parameters. However, the necessary and sufficient conditions underlying the validity of these approaches have not been rigorously analyzed. Consequently, questions have been raised²⁰ regarding the applicability of these approaches to sensitivity analysis of problems that are more complex than those treated so far; of practical interest are, for example, thermal-hydraulics problems involving discontinuous state functions and parameters.

More recently, Cacuci *et al.*²¹ have introduced and employed concepts of nonlinear functional analysis²²⁻²⁴ in an attempt to set sensitivity theory on a more rigorous mathematical foundation,

and to extend the scope of the theory. In addition, they have also presented a sensitivity theory formulation for a class of discretized nonlinear systems, and have enlarged the type of functionals considered as responses. Although, rigorous within explicitly stated limitations, their derivations repeatedly required the existence of the Frechet derivatives²²⁻²⁶ of the various operators, without providing an analysis of the motivations underlying the *necessity* of these requirements. Since operators that are not Frechet differentiable can also be encountered in practice, an investigation of the aforementioned question of necessity was undertaken by Cacuci,²⁸ who provided a detailed and rigorous investigation of the mathematical concepts underlying sensitivity theory.

These lecture notes present the main aspects underlying the rigorous mathematical formulation of deterministic sensitivity theory for nonlinear systems; the presentation of this formulation is based on the work by Cacuci²⁸ and employs concepts and methods of nonlinear functional analysis.²²⁻²⁶ As these lecture notes are intended to be self-contained, the functional analytic concepts underlying sensitivity theory are briefly reviewed in Sec. II.

The rigorous formulation of sensitivity theory for nonlinear systems is presented in Sec. III. This section is divided in three main parts. Thus, Sec. III.A. presents the sensitivity theory for nonlinear systems with operator-type responses. The system of nonlinear operator equations and the associated response, itself a general nonlinear operator, are introduced and described in Sec. III.A.1.; altogether, they are intended to be sufficiently general to include – as particular cases – the mathematical representation of a large number of problems in a wide variety of fields. The problem is formulated here in normed linear spaces over the scalar field of real numbers. This choice of space is sufficiently general for the purposes of this study:

1. it provides the framework for the clear exposition of the necessary and sufficient conditions underlying the sensitivity theory formalisms presented in Sec. III.A.2.,
2. it opens the possibility to establish the limitations inherent in the previous approaches (by direct comparison with the present approach),
3. it provides a basis for still further extensions of the theory.

The formulation of the sensitivity theory presented in Sec. III. is centered on evaluating the Gâteaux differential of the response; this quantity is considered to be "the most general measure of the sensitivity of a response to variations in the system parameters."²⁷ Consequently, Secs. III.A.2.a and III.A.2.b are devoted to the presentation of the two alternative methods for evaluating this Gâteaux differential. The conceptually and computationally straightforward method is labeled the Forward Method and is dealt with in Sec. III.A.2.a. The alternative method, labeled the Adjoint Method is presented in Sec. III.A.2.b. The motivation underlying the development of this Adjoint Method is well known,^{1-19,21} and this method is a great deal more economical to apply, *if possible*, to the broad class of practical problems characterized by large data bases and comparatively few responses. However, the present analysis also reveals the fact that the Adjoint Method can be formulated *only* under conditions that are more restrictive than those underlying the formulation of the Forward Method. The limitations inherent in the previous approaches^{1-19,21} to sensitivity theory are assessed in Sec. III.A.3. by examination of their underlying assumptions and by comparison to the formalisms presented in Sec. III.A.2.

Section III.B. presents the development of the sensitivity theory for nonlinear systems with responses defined at critical points of a function of the system's state vector and parameters. Such responses are characterized both by the numerical value at the maximum and by the position in phase-space where the maximum occurs. In this case, varying the system parameters alters not only the value at this maximum but also alters the position of the maximum in phase-space. The sensitivity theory presented in Sec. III.B. allows treatment of a general response comprising, as particular cases, the representation of maxima, minima, and saddle points.

Section III.B.6 presents the application of the sensitivity theory formulated in Secs. III.B.1 through III.B.5. to the single-phase modules of the MELT-IIIB fast reactor safety code, emphasizing, in particular, results regarding the sensitivity of the locations (in phase space) of two important responses — the maximum fuel temperature and the maximum normalized reactor power level. Section III.B.6.a. describes the mathematical representation of the physical problem and the responses. The sensitivity theory developments, including the derivation of the appropriate adjoint systems and the expression of sensitivities in terms of adjoint functions, are presented in Sec. III.B.6.b. The numerical results obtained for the sensitivities are discussed in Sec. III.B.6.c. Finally, the summary and conclusions presented in Sec. III.B.6.d. highlight the practical usefulness of applying the Adjoint Method to perform sensitivity analysis of realistic reactor safety problems.

Section III.C. presents sensitivity theory for nonlinear systems with feedback. Feedback occurs when quantities that are normally input in the mathematical model of the physical process (e.g., parameters or data) are allowed to depend on the model's output (i.e., the dependent variables). Such parameter variations can no longer be prescribed *a priori* since they depend on the model's output; therefore, the term **sensitivity to feedback** is defined in Sec. III.C.2. It is then shown that this sensitivity to feedback provides an estimate of the actual effect of feedback correct to first order in the strength of the feedback, and it is shown how this sensitivity can be efficiently obtained using the Adjoint Method. Finally, Sec. IV. summarizes and highlights the main points underlying deterministic sensitivity theory for nonlinear systems and discusses the potential of using functional analytic concepts to extend further the scope of this theory.

II. MATHEMATICAL BACKGROUND

II.A. VECTOR SPACE

The vector space axioms are suggested by the algebraic properties of vector addition and multiplication by scalars for three-dimensional Cartesian vectors.

II.A.1 Definition

Let V be a non-empty set, and suppose that any pair of elements $f, g \in V$ can be combined by an operation called addition to give an element $f + g$ in V . Assume that for any $f, g, h \in V$:

- i. $f + g = g + f$;
- ii. $f + (g + h) = (f + g) + h$;
- iii. there is a unique element 0 (called zero) in V such that $f + 0 = f$ for all $f \in V$;
- iv. for each $f \in V$ there is a unique element $(-f)$ in V such that $f + (-f) = 0$.

In the following the scalars will either be the real numbers R or the complex numbers C . Suppose any $f \in V$ and any scalar α can be combined to give an element αf in V , and assume that for any scalars α, β ,

1. $\alpha(f + g) = \alpha f + \alpha g$;
2. $(\alpha + \beta)f = \alpha f + \beta f$;
3. $(\alpha\beta)f = \alpha(\beta f)$;
4. $1 \times f = f$.

Then V is called a **vector space** (or a complex vector space) if the scalar field is C or a **real vector space** if it is R . The members f, g, h, \dots of V are known as **points, elements, or vectors** depending on which seems most appropriate in the context.

II.A.2. Definition

Let V be a vector space. A *finite* set $S = \{f_j\}_{j=1}^n$ of vectors in V is called **linearly dependent** iff (if and only if) there are scalars $\alpha_1, \dots, \alpha_n$ not all of which are zero such that $\sum \alpha_j f_j = 0$, otherwise S is said to be **linearly independent**. An arbitrary set S of vectors in V is **linearly independent** iff every finite non-empty subset of S is linearly independent; otherwise it is **linearly dependent**.

If there is a positive integer n such that V contains n but not $n+1$ linearly independent vectors, V is said to be **finite dimensional** with **dimension** n . V is **infinite dimensional** iff it is not finite dimensional. The finite set S of vectors in V is called a **basis** of V iff S is linearly independent and each element of V may be written as $\sum_1^n \alpha_j f_j$ for some $\alpha_1, \dots, \alpha_n \in C$ and $f_1, \dots, f_n \in S$ (of course n is the dimension of V).

II.A.3. Example

From the point of view of applications by far the most important vector spaces are those whose elements are functions. To illustrate the natural laws of combination, consider the set V of complex valued functions defined on an interval $[a, b]$. For $f, g \in V$ and $\alpha \in \mathbb{C}$, define new functions $f + g$ and αf by requiring the following relations to hold for all $x \in [a, b]$:

$$\begin{aligned}(f + g)(x) &= f(x) + g(x); \\ (\alpha f)(x) &= \alpha f(x).\end{aligned}$$

Of course $(f + g)(x), (\alpha f)(x)$ denote the values of the functions $(f + g), \alpha f$ respectively at x ; these laws of combination are described as pointwise addition and multiplication by scalars. It is easy to check that the vector space axioms are satisfied, and V is a (complex) vector space. The real valued functions similarly form a real vector space. Obviously V is infinite dimensional.

II.B. NORMED VECTOR SPACE

II.B.1 Definition

Let V be a vector space, and suppose that to each element $f \in V$ a non-negative number $\|f\|$ is assigned in such a way that for all $f, g \in V$:

- i. $\|f\| = 0$ iff $f = 0$;
- ii. $\|\alpha f\| = |\alpha| \|f\|$ for any scalar α ;
- iii. $\|f + g\| \leq \|f\| + \|g\|$ (the triangle inequality).

The quantity $\|f\|$ is called the **norm** of f , and V is known as a **normed vector space**.

II.C. BANACH SPACE (Complete Normed Vector Space)

Although it is possible to obtain meaningful generalizations of many of the concepts useful in finite dimensions to infinite dimensional normed vector spaces, analysis cannot be carried out satisfactorily in every space of this type. The reason is that the convergence of sequences – which is of fundamental importance in analysis – can pose problems which are extremely intractable. Roughly this is because sequences which "ought" to be convergent do not always turn out to be so. To make substantial progress with analysis, it is essential to restrict the space further. Out of the various possibilities, one strategy which has achieved considerable success is to impose the condition of "completeness" on the norm, and to study complete normed vector spaces (or Banach spaces as they will be called); this will be the course followed here. It will appear as we proceed that the assumption of completeness significantly simplifies the abstract analysis and at the same time is satisfied by a wide range of normed vector spaces. Completeness is indeed one of the most important concepts in functional analysis, and the contents of this section, which consist of a study of this property and related ideas and of the illustration of these ideas in specific spaces, are fundamental to the later development.

II.C.1. Definition

Let V be a normed vector space. A sequence (f_n) in V is said to be **Cauchy** iff

$$\lim_{m, n \rightarrow \infty} \|f_n - f_m\| = 0,$$

that is, iff for each $\epsilon > 0$ there is an n_0 such that $\|f_n - f_m\| < \epsilon$ whenever $m, n > n_0$.

II.C.2. Definition

A set S in a normed vector space V is said to be **complete** iff each Cauchy sequence in S converges to a point of S . V itself is known as a complete normed vector space or a **Banach space** iff it is complete.

II.D. HILBERT SPACE

Hilbert spaces are the simplest type of infinite dimensional normed spaces to play a significant role in functional analysis. Their relative simplicity is due to an additional structure – called an inner product – which is imposed on the space; a Banach space with an inner product is a Hilbert space. The inner product is itself a generalization of the scalar product of elementary Cartesian vector analysis. The scalar product is usually defined in terms of the components of the vector, but in accordance with the standard tactics in functional analysis, the algebraic properties of the scalar product are taken as axioms in the abstract context.

The presence of this additional algebraic structure much enriches the geometrical properties of the space. Most significantly, it is possible to define a notion of perpendicularity for two vectors, and the geometry corresponds in several fundamental respects with Euclidean geometry. The effect of the inner product on the analytical (as opposed to the geometric) properties is more subtle. Basically, as in Banach spaces, the main problems in Hilbert spaces are connected with their infinite dimensionality. However, in some respects there can be considerable simplification. This will be seen later in this section when bases of Hilbert space are considered. From the point of view of applications perhaps the most important fact is that a sensible definition of a self adjoint operator may be given in Hilbert space and a powerful body of theory based on this concept developed.

II.D.1. Definition

Let V be a vector space. An **inner product** is a complex valued function (\cdot, \cdot) on $V \times V$ such that for all $f, g, h \in V$ and $\alpha \in C$ the following hold:

- i. $(f, f) \geq 0$, and $(f, f) = 0$ iff $f = 0$;
- ii. $(f, g + h) = (f, g) + (f, h)$;
- iii. $(f, g) = \overline{(g, f)}$, where the bar denotes the complex conjugate;
- iv. $(\alpha f, g) = \alpha(f, g)$.

II.D.2. Definition

A space V equipped with an inner product is known as a **pre-Hilbert space** (the term *inner product space* is also used in the literature). If V is a real vector space, and the inner product is real valued, a **real pre-Hilbert space** is obtained.

II.D.3. Definition

A pre-Hilbert space which is complete with respect to the norm $\|f\| = (f, f)^{1/2}$ is called a **Hilbert space**. We shall denote this Hilbert space by H .

II.D.4. Definition

A set K of vectors in H is said to be **complete** (the reader should be warned that the terminology in this area of Hilbert space theory is not uniform in the literature) iff $(f, \phi) = 0$ for all $\phi \in K$ implies that $f = 0$. A countable set $K = \{\phi_n\}_{n=1}^{\infty}$ is called **orthonormal** iff $(\phi_n, \phi_m) = \delta_{nm}$ for all $m, n \geq 1$. The numbers (f, ϕ_n) are known as the **Fourier coefficients** of f (with respect to K), and the **Fourier series** of f is the formal series $\sum_n (f, \phi_n) \phi_n$.

II.D.5. Definition

An orthonormal set $K = \{\phi_n\}$ is called an **orthonormal basis** of H iff for every $f \in H$,

$$f = \sum_n (f, \phi_n) \phi_n \quad .$$

This expansion is of course the Fourier series of f (Definition II.D.4).

II.E. SOME BASIC TERMINOLOGY OF OPERATOR THEORY

Let V and W be vector spaces. Let A be a mapping defined on some subset $D(A)$ of V , and assume that A assigns to each element f of $D(A)$ a unique element Af in W (in the initial stages $D(A)$ will usually be the whole of V).

II.E.1. Definition

The set $D(A)$ (sometimes denoted just by D if there is only one mapping under discussion) is called the **domain** of A . For $f \in D(A)$, the element Af is known as the **image** of f . Likewise the image $A(S)$ of a set $S \subset D(A)$ is the set of the images of all the elements of S . In particular the image of $D(A)$ is called the **range** of A and will be written as $R(A)$. The **preimage** of a set $S_1 \subset W$ is the set $A^{-1}(S_1) = \{f : f \in D(A), Af \in S_1\}$.

II.E.2. Definition

A is called an **operator** or a **function** from V into W . The notation $A : S \rightarrow W$ will mean that A is an operator with domain S and range in W , and we say that A maps S into W .

The following points arising from these definitions should be noted. First, an operator is always single valued in that it assigns exactly one element of its range to each element in its domain. Second, the statement that A is an operator from V into W allows the possibility that $D(A)$ is a proper subset of V ; in contrast $A : V \rightarrow W$ means that $D(A) = V$. Lastly, although there is no strict distinction between "operator" and "function", it is customary to reserve "function" for the case when V and W are finite dimensional and to use "operator" otherwise. In view of its importance one particular type of operator is given a name of its own.

II.E.3. Definition

Let V be a complex (respectively real) vector space, and suppose that $W = C$ (respectively R). Then an operator from V into W is known as a **functional**.

II.E.4. Definition

Let A and A_1 be operators from V into W . A and A_1 are said to be **equal** iff $D(A) = D(A_1)$ and $A_1f = Af$ for all f in $D(A)$. A_1 is said to be an **extension** of A (written $A \subset A_1$), and A to be a **restriction** of A_1 , iff $D(A_1) \supset D(A)$ and $Af = A_1f$ for all $f \in D(A)$. The extension is described as **proper** iff $D(A_1) \neq D(A)$.

II.E.5. Definition

Let V and W be normed vector spaces, and suppose that A is an operator from V into W . A is said to be **continuous at the point** $f_0 \in D(A)$ iff one of the following pair of equivalent conditions holds:

- i. For each $\epsilon > 0$ there is a $\delta > 0$ such that $\|Af - Af_0\| < \epsilon$ if $f \in D(A)$ and $\|f - f_0\| < \delta$;
- ii. For every sequence (f_n) in $D(A)$ with limit f_0 , $\lim Af_n = Af_0$. A is said to be **continuous** iff it is continuous at every point of $D(A)$.

II.E.6. Definition

Let V and W be vector spaces, and let $D(L)$ be a linear subspace of V . An operator L from V into W with domain $D(L)$ is said to be **linear** iff

$$L(\alpha f + \beta g) = \alpha Lf + \beta Lg$$

for all $\alpha, \beta \in C$ (or R if V and W are real spaces) and all $f, g \in D(L)$. (The restriction that $D(L)$ be a linear subspace is obviously necessary if the definition is to make sense; note that $R(L)$ is also a linear subspace).

The space of **continuous linear functionals** defined on a Banach space B is called the **dual** of B and is denoted by B^* . For $f \in B$ and $f^* \in B^*$, $f^*(f)$ will denote the complex number assigned to f by the mapping f^* .

II.E.7. The Riesz Representation Theorem.

Corresponding to every element g^* of the dual H^* of a Hilbert space H , there is a unique element g of H such that $g^*(f) = (f, g)$ for all $f \in H$. Also $\|g^*\| = \|g\|$.

Note that Hilbert spaces are self-dual, i.e., $H^* = H$.

II.E.8. Definition.

Let H be a Hilbert space, and suppose that $L \in \mathcal{L}(H)$, i.e., L is an element of the (linear) space of linear operators from H into H . The relation

$$(Lf, g) = (f, L^*g)$$

required to hold for all $f, g \in H$ defines the bounded linear operator $L^* \in \mathcal{L}(H)$ called the (Hilbert space) **adjoint** of L .

II.E.8. Definition.

A **multi-index** α is an n -tuple $(\alpha_1, \dots, \alpha_n)$ of non-negative integers. We write $|\alpha| = \alpha_1 + \dots + \alpha_n$; this conflicts with the notation for the Euclidean distance in R^n , but the meaning will always be clear from the context. Multi-indices will be denoted by α and β .

A point in R^n will be $x = (x_1, \dots, x_n)$ with $|x|^2 = \sum x_j^2$ and $x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}$. We write $D_j = \partial/\partial x_j$ and $D^\alpha = D_1^{\alpha_1} \dots D_n^{\alpha_n}$. With these conventions the notation for a partial differential equation may be simplified by writing

$$\sum_{|\alpha| \leq m} p_\alpha D^\alpha = \sum_{j=0}^m \sum_{\alpha_1 + \dots + \alpha_n = j} p_{\alpha_1, \dots, \alpha_n} D_1^{\alpha_1} \dots D_n^{\alpha_n}.$$

II.E.9. Definition

Assume that for some multi indices α, β with $|\alpha| = |\beta| = m$, there exist complex valued variable coefficients $p_{\alpha\beta} \neq 0$, such that for all $\alpha, \beta, p_{\alpha\beta} \in C^\infty(\Omega)$. For $\phi \in C^{2m}$, define the forms

$$l\phi = \sum_{|\alpha|, |\beta| \leq m} (-1)^{|\alpha|} D^\alpha (p_{\alpha\beta} D^\beta \phi),$$

$$l_p \phi = (-1)^m \sum_{|\alpha| = |\beta| = m} D^\alpha (p_{\alpha\beta} D^\beta \phi).$$

l is called a **formal partial differential operator** of order $2m$; l_p is known as the **principal part** of l .

II.E.10. Definition.

The operator l^* , where

$$l^* \phi = \sum_{|\alpha|, |\beta| \leq m} (-1)^{|\alpha|} D^\alpha (\bar{p}_{\beta\alpha} D^\beta \phi), \quad (\phi \in C^{2m}),$$

is called the **formal adjoint** of l ; l is said to be **formally self-adjoint** iff $l = l^*$.

II.F. SOME ASPECTS OF DIFFERENTIAL CALCULUS IN VECTOR SPACES

II.F.1. Definition

Let X and Y be normed real spaces and U an open subset of X . Let $x_0 \in U$ and h be a fixed nonzero element in X . Since U is open there exists an interval $I = (-\tau, \tau)$ for some $\tau > 0$ such that if $t \in I$, then $x_0 + th \in U$. If the mapping $\Phi: I \rightarrow X$ defined by $\Phi(t) = F(x_0 + th)$ has a derivative at $t = 0$, then $\Phi'(0)$ is called the **Gâteaux variation** of F at x_0 with increment h and is denoted by $\delta F(x_0; h)$, i.e.,

$$\delta F(x_0; h) = \frac{d}{dt} F(x_0 + th) \Big|_{t=0} = \lim_{t \rightarrow 0} \frac{1}{t} \{F(x_0 + th) - F(x_0)\} .$$

Note that this equation may be used to define $\delta F(x_0; h)$ when X is any linear space, not necessarily normed. Note also that the operator $\delta F(x_0; h)$ is not necessarily linear nor continuous in h .

II.F.2. Theorem

Suppose that F has a Gâteaux variation at x_0 . A necessary and sufficient condition for $\delta F(x_0; h)$ to be linear and continuous in h is that F satisfies the following two conditions:

- i. to each h corresponds a $\delta(h)$ such that $|t| \leq \delta$ implies $\|F(x_0 + th) - F(x_0)\| \leq M \|th\|$, where M does not depend on h , and
- ii. $F(x_0 + th_1 + th_2) - F(x_0 + th_1) - F(x_0 + th_2) + F(x_0) = o(t)$.

II.F.3. Definition.

F has a **Gâteaux differential** at x_0 if $\delta F(x_0; \cdot)$ is linear and continuous. In this case $\delta F(x_0; \cdot)$ is denoted by $DF(x_0)$ or $F'(x_0)$ and is called the **Gâteaux derivative**.

Note that some authors refer to the variation $\delta F(x; h)$ as the Gâteaux differential, and then use the phrase "linear Gâteaux differential" whenever $\delta F(x; \cdot)$ is linear.

II.F.4 Definition.

Let X and Y be normed real linear spaces and let $L(X, Y)$ denote the space of all continuous linear operators on X to Y , with the usual norm. A map $F: U \rightarrow Y$, where U is an open subset of X , is said to be **Fréchet differentiable** at $x_0 \in U$ if there exists a continuous linear operator $L(x_0): X \rightarrow Y$ such that the following representation holds for every $h \in X$ with $x_0 + h \in U$,

$$F(x_0 + h) - F(x_0) = L(x_0)h + r(x_0; h) \quad , \quad (a)$$

where

$$\lim_{\|h\| \rightarrow 0} \frac{\|R(x_0; h)\|}{\|h\|} = 0 \quad . \quad (b)$$

The unique $L(x_0)h$ in (a) is called the **Fréchet differential** of F at x_0 and is denoted by $dF(x_0; h)$. The operator $F'(x_0) \in L(X, Y)$ defined by $h \rightarrow dF(x_0; h)$ is called the **Fréchet derivative** of F at x_0 ; we write $dF(x_0; h) = F'(x_0)h$.

II.F.5 Definition.

Let X be an open subset of the product space $\Pi = E_1 \times \dots \times E_n$. Let $F: X \rightarrow Y$. If there exists a bounded linear operator $L(x_1, \dots, x_n; \cdot)$ such that for all $h_i \in E_i$ with

$$(x_1, \dots, x_{i-1}, x_i + h_i, x_{i+1}, \dots, x_n) \in X \quad ,$$

$$F(x_1, \dots, x_{i-1}, x_i + h_i, \dots, x_n) - F(x_1, \dots, x_n) =$$

$$L(x_1, \dots, x_n; h_i) + R(x_1, \dots, x_n; h_i) \quad ,$$

where

$$\frac{\|R(x_1, \dots, x_n; h_i)\|}{\|h_i\|} \rightarrow 0 \text{ as } h_i \rightarrow 0 \text{ ,}$$

then $L(x_1, \dots, x_n; h_i)$ is called the **Fréchet partial differential** and is denoted by $d_i F(x_1, \dots, x_n; h_i)$. F is said to be **totally differentiable** if it is **Fréchet differentiable** considered as a mapping on $X \subset E_1 \times \dots \times E_n$ into Y , that is if there exists an $L(x; h)$, $x = (x_1, \dots, x_n) \in X$, $h = (h_1, \dots, h_n) \in \Pi$, which is linear and continuous in h such that

$$\lim_{h \rightarrow 0} \frac{\|F(x_1 + h_1, \dots, x_n + h_n) - F(x_1, \dots, x_n) - L(x_1, \dots, x_n; h_1, \dots, h_n)\|}{\|h_1\| + \dots + \|h_n\|} = 0 \text{ .}$$

$L(x_1, \dots, x_n; h_1, \dots, h_n)$ is called the **total Fréchet differential** of F and is denoted by $dF(x_1, \dots, x_n; h_1, \dots, h_n)$.

An operator $F: X \subset \Pi \rightarrow Y$ which is totally differentiable at x_1, \dots, x_n is partially differentiable with respect to each variable, and its total differential is the sum of the differentials with respect to each of the variables.

II.F.6 Definition.

Let X be an open subset of the product space $\Pi = E_1 \times \dots \times E_n$. Let $F: X \rightarrow Y$. If there exists a bounded linear operator $D_i F(x_1, \dots, x_n; \cdot): E_i \rightarrow Y$ such that

$$\begin{aligned} & F(x_1, \dots, x_{i-1}, x_i + h_i, x_{i+1}, \dots, x_n) - F(x_1, \dots, x_n) \\ &= D_i F(x_1, \dots, x_n; h_i) + R(x_1, \dots, x_n; h_i) \end{aligned}$$

where

$$\lim_{t \rightarrow 0} \frac{R(x_1, \dots, x_n; th_i)}{t} = 0 \text{ ,}$$

then $D_i F(x_1, \dots, x_n; h_i)$ is called the *Gâteaux partial differential*. The operator F is said to be *totally Gâteaux differentiable* at x if F , considered as a mapping on $X \subset \Pi$ into Y , is Gâteaux differentiable at x . This means that

$$\begin{aligned} & F(x_1 + h_1, \dots, x_n + h_n) - F(x_1, \dots, x_n) \\ &= L(x_1, \dots, x_n; h_1, \dots, h_n) + R(x_1, \dots, x_n; h_1, \dots, h_n) \end{aligned}$$

where L is a continuous linear operator in $h = (h_1, \dots, h_n)$, and

$$\lim_{t \rightarrow 0} t^{-1} R(x_1, \dots, x_n; th_1, \dots, th_n) = 0 \text{ .}$$

III. RIGOROUS MATHEMATICAL FORMULATION OF THE FORWARD AND ADJOINT METHODS OF SENSITIVITY THEORY FOR NONLINEAR SYSTEMS

III.A. NONLINEAR SYSTEMS WITH OPERATOR-TYPE RESPONSES

The material presented in Sec. III.A is largely based on the following works:

a. D. G. Cacuci, "Sensitivity Theory for Nonlinear Systems. I. Nonlinear Functional Analysis Approach," *J. Math. Phys.* **22**(12), 2794-2802 (1981).

b. D. G. Cacuci, "Sensitivity Theory for Nonlinear Systems. II. Extensions to Additional Classes of Responses," *J. Math. Phys.*, **22**(12), 2803-2812 (1981).

III.A.1. Mathematical Representation of the Physical Problem

Consider, for the sake of generality, that the physical problem under consideration is represented by the following system of K coupled nonlinear equations written in operator form as

$$\mathbf{N}[u(x), \alpha(x)] = \mathbf{Q}[\alpha(x), x] . \quad (1)$$

The quantities appearing in Eq. (1) are defined as follows:

1. $x = (x_1, \dots, x_J)$ is the phase-space position vector $x \in \Omega \subset R^J$, where Ω is a subset of the J -dimensional real vector space R^J ,

2. $u(x) = [u_1(x), \dots, u_K(x)]$ is the state vector; $u(x) \in E_u$, where E_u is a normed linear space over the scalar field Λ of real numbers,

3. $\alpha(x) = [\alpha_1(x), \dots, \alpha_I(x)]$ is the vector of system parameters; $\alpha \in E_\alpha$, where E_α is also a normed linear space. In practical applications, E_α may be one of the Hilbert spaces L_2 or l_2 ; occasionally, the components of α may simply be a set of real scalars, in which case E_α is R^I ,

4. $\mathbf{Q}[\alpha(x), x] = [\mathbf{Q}_1(\alpha, x), \dots, \mathbf{Q}_K(\alpha, x)]^T$ is a (column) vector whose elements represent inhomogeneous source terms (the symbol T denoted "transposition"); $\mathbf{Q} \in E_Q$, where E_Q is again a normed linear space. The components of \mathbf{Q} may be operators (rather than just functions) acting on $\alpha(x)$ and x , operators (rather than just functions) acting on $\alpha(x)$ and x ,

5. the components of the (column) vector $\mathbf{N} = [\mathbf{N}_1(u, \alpha), \dots, \mathbf{N}_K(u, \alpha)]^T$ are nonlinear operators acting, in general, not only on the state vector $u(x)$, but also on the vector of system parameters $\alpha(x)$.

In view of the definitions given above, \mathbf{N} represents the mapping $\mathbf{N} : S \subset E \rightarrow E_Q$, where $S = S_u \times S_\alpha$, $S_u \subset E_u$, $S_\alpha \subset E_\alpha$, and $E = E_u \times E_\alpha$. Note that an arbitrary element $e \in E$ is of the form $e = (u, \alpha)$. Even though in most practical applications E and E_Q will be Hilbert spaces (e.g., the space L_2 , the Sobolev spaces H^m), this restriction is not imposed at this stage for the sake of generality. In the same vein of generality, the components of \mathbf{N} are considered here to be defined in terms of operators such as differential, difference, integral, distributions, or infinite matrices. The domain S of \mathbf{N} is, of course, intimately related to the characteristics of these operators. Thus, if differential operators appear in Eq. (1), then a corresponding set of boundary and/or initial conditions — which is essential to define S — must also be given. This set can be represented as

$$[\mathbf{B}(u, \alpha) - \mathbf{A}(\alpha)]_{\partial\Omega} = 0 , \quad (2)$$

where \mathbf{A} and \mathbf{B} are operators and $\partial\Omega$ is the boundary of Ω ; the operator $\mathbf{A}[\alpha]$ represents all inhomogeneous boundary terms.

To be definite, $u(x)$ is considered to be the unique nontrivial solution of the physical problem described in Eqs. (1) and (2). This requirement is usually fulfilled (or assumed to be fulfilled when rigorous existence and uniqueness proofs are lacking) in most problems of practical interest. The following purposes are accomplished as consequences of imposing this requirement:

1. elimination from further consideration of those points in nonlinear problems where bifurcation (i.e., branching) of solutions occurs,
2. inclusion of the treatment of source-free problems as a special case of Eq. (1).

In this vein, Eq. (1) is considered to include any equality constraints that $u(x)$ might be required to satisfy. The specifications introduced so far are sufficiently general to allow Eqs. (1) and (2) to include, as particular cases, the mathematical modeling of a wide range of problems of practical interest in many diverse fields.

The system's response, (i.e., performance parameter) \mathbf{R} associated with the problem modeled by Eqs. (1) and (2) must also be specified. The most general type of system response, which includes phase-space dependent mappings of the system's state vector u and parameters α , is the operator

$$\mathbf{R}(e) : D \subseteq E \rightarrow E_R , \quad (3)$$

where E_R is a normed vector space.

III.A.2. Sensitivity Theory

The most general and fundamental concept for the definition of the sensitivity of a response to variations in the system parameters is the Gateaux (G)-differential. The G-differential $\mathcal{V}\mathbf{R}(e^0; h)$ of $\mathbf{R}(e)$ at e^0 with increment h , is defined as

$$\lim_{t \rightarrow 0} [\mathbf{R}(e^0 + th) - \mathbf{R}(e^0)]/t = \mathcal{V}\mathbf{R}(e^0; h) \quad (4)$$

for $t \in \Lambda$, and all (i.e., arbitrary) vectors $h \in E$; here, $h = (h_u, h_\alpha)$, since $E = E_u \times E_\alpha$.

The G-differential $\mathcal{V}\mathbf{R}(e^0; h)$ is related to the total variation $[\mathbf{R}(e^0 + h) - \mathbf{R}(e^0)]$ of \mathbf{R} at e^0 through the relationship

$$\mathbf{R}(e^0; h) - \mathbf{R}(e^0) = \mathcal{V}\mathbf{R}(e^0; h) + \Delta(h) , \quad (5)$$

where

$$\lim_{t \rightarrow 0} [\Delta(th)/t] = 0.$$

It is important to note that, in view of the properties of the G differential, \mathbf{R} need not be continuous in u and/or α for $V\mathbf{R}(e^0;h)$ to exist at $e^0 = (u^0, \alpha^0)$, and that $V\mathbf{R}(e^0;h)$ is not necessarily linear in h . It thus becomes apparent that by defining $V\mathbf{R}(e^0;h)$ to be *the sensitivity of the response \mathbf{R}* , the definitions of sensitivity encountered in the previously mentioned works¹⁻¹⁹ are considerably generalized and extended. With the present definition, the concept of sensitivity also becomes meaningful for certain types of physical problems and responses (e.g., involving discontinuities) which could not have been treated within the framework of the previous approaches.

Thus, the objective of sensitivity theory is to evaluate $V\mathbf{R}(e^0;h)$. To achieve this objective, two alternative formalisms — the "Forward Method" and the "Adjoint Method" are developed and discussed in the following.

III.A.2.a. The Forward Method

It is observed that, given the vector of "changes" h_α around the "base-case configuration" α^0 , the sensitivity $V\mathbf{R}(e^0;h)$ of $\mathbf{R}(e)$ at e^0 can be evaluated only after determining the vector h_u , since h_u and h_α are not independent. A relationship between h_u and h_α is obtained by taking the G-differentials of Eqs. (1) and (2). This gives

$$VN(e^0;h) - VQ(\alpha^0;h_\alpha) = 0, \quad (6)$$

and

$$\{VB(e^0;h) - VA(\alpha^0;h_\alpha)\}_{\delta\Omega} = 0. \quad (7)$$

respectively. Of course, the above system of equations — which will subsequently be referred to as the "forward sensitivity equations" — is meaningful if and only if the respective G-differentials of the operators \mathbf{N} , \mathbf{B} , \mathbf{Q} , and \mathbf{A} exist. Note again that these G-differentials need not necessarily be linear operators in either h_u or h_α , and that their existence does not require the operators \mathbf{N} , \mathbf{B} , \mathbf{Q} , and \mathbf{A} to be continuous in u or α at e^0 .

For a given vector of "changes" h_α around α^0 , one must be able to solve the system given in Eqs. (6) and (7) to obtain h_u ; otherwise, of course, it would be impossible to perform sensitivity analysis of the given physical system. [However, a detailed analysis of the conditions under which Eqs. (6) and (7) can be solved for h_u is not within the scope of this work.] Once h_u is determined, it can be employed, in turn, to evaluate the sensitivity $V\mathbf{R}(e^0;h)$ of $\mathbf{R}(e)$ at e^0 , for a given vector of "changes" h_α .

It should be noted here that the "Forward Method" is characterized in a fundamental sense by the fact that the solution h_u of the h_α -dependent "forward sensitivity equations" [viz., Eqs. (6) and (7)] is needed to evaluate $V\mathbf{R}(e^0;h)$. Consequently, from the standpoint of computational costs, the "Forward Method" is advantageous to employ only if, in the problem under consideration, the number of different responses of interest exceeds the number of input parameters. However, a large number of problems of practical interest are characterized by very large data bases (i.e., α has many components) and comparatively few responses. In such situations, it is not economical

to employ the "Forward Method" to answer all sensitivity questions that might arise in practice, since it becomes prohibitively expensive to repeatedly solve the h_α -dependent "forward sensitivity equations" to determine h_u for all possible vectors h_α . Hence, it is clearly desirable to devise (if possible) an alternative procedure to evaluate $\mathcal{V}\mathbf{R}(e^o;h)$, to avoid the necessity of repeatedly solving the "forward sensitivity equations."

III.A.2.b. The Adjoint Method

The practical motivation underlying the development of this alternative method for sensitivity analysis is to avoid the need for repeatedly solving Eqs. (6) and (7). This goal can be achieved if we can eliminate all unknown values of h_u from the expression of $\mathcal{V}\mathbf{R}(e^o;h)$. This elimination can be accomplished by constructing an *adjoint system* that is (i) uniquely defined, (ii) independent of the vectors h_u and h_α , and (iii) such that its solution can be used to eliminate all unknown values of h_u from the expression of $\mathcal{V}\mathbf{R}(e^o;h)$.

Adjoint operators can only be introduced uniquely for densely defined *linear operators* in Banach spaces. However, at this stage, $\mathcal{V}\mathbf{N}(e^o;h)$, $\mathcal{V}\mathbf{B}(e^o;h)$, and $\mathcal{V}\mathbf{R}(e^o;h)$ are not necessarily linear in h , and E is not necessarily complete. It follows that developing the Adjoint Method requires the introduction of restrictions in addition to those underlying the validity of the Forward Method.

There are several equivalent theorems giving necessary and sufficient conditions in order that a nonlinear operator $F(e)$ with domain in E and range in another normed linear space admit a G differential $\mathcal{V}F(e^o;h)$ at e^o that is linear in h . A set of such conditions is provided by Theorem II.F.2:

Theorem: the G-differential $\mathcal{V}F(e^o;h)$ of F at e^o is linear in $h \in E$ iff:

$F(e)$ satisfies a weak Lipschitz condition at e^o , and:

$$F(e^o + th_1 + th_2) - F(e^o + th_1) - F(e^o + th_2) + F(e^o) = o(t) . \quad (8)$$

Thus, $\mathcal{V}\mathbf{N}(e^o;h)$ and $\mathcal{V}\mathbf{B}(e^o;h)$ are linear in h if \mathbf{N} and \mathbf{B} satisfy, in turn, conditions identical to those stated in Eq. (8) for $F(e)$. For the purposes of subsequent derivations, $\mathcal{V}\mathbf{N}(e^o;h)$ and $\mathcal{V}\mathbf{B}(e^o;h)$ are henceforth considered to be linear in h , and denoted by $\mathcal{D}\mathbf{N}(e^o;h)$ and $\mathcal{D}\mathbf{B}(e^o;h)$, respectively. Recalling now that, in our case, $E = E_u \times E_\alpha$, it further follows that

$$\mathcal{D}\mathbf{N}(e^o;h) = \mathbf{N}'_u(e^o)h_u + \mathbf{N}'_\alpha(e^o)h_\alpha \quad (9)$$

and

$$\mathcal{D}\mathbf{B}(e^o;h) = \mathbf{B}'_u(e^o)h_u + \mathbf{B}'_\alpha(e^o)h_\alpha . \quad (10)$$

In the above expressions, $\mathbf{N}'_u(e^o)$ and $\mathbf{B}'_u(e^o)$ denote, respectively, the partial G-derivatives at e^o of \mathbf{N} and \mathbf{B} with respect to u , while $\mathbf{N}'_\alpha(e^o)$ and $\mathbf{B}'_\alpha(e^o)$ denote the partial G-derivatives at e^o of \mathbf{N} and \mathbf{B} with respect to α . Note that $\mathbf{N}'_u(e^o)$ and $\mathbf{B}'_u(e^o)$ are linear operators in h_u with domain in E_u and range in E_Q [i.e., $\mathbf{N}'_u(e^o), \mathbf{B}'_u(e^o) \in L(E_u, E_Q)$], and are independent of h_α ; similarly, $\mathbf{N}'_\alpha(e^o), \mathbf{B}'_\alpha(e^o) \in L(E_\alpha, E_Q)$, and are independent of h_u . The explicit representation of $\mathbf{N}'_u(e^o)$ and $\mathbf{N}'_\alpha(e^o)$ are matrices whose elements are the partial G-derivatives at e^o of the components of u

and the components of α . [The elements of the matrices representing $\mathbf{B}'_u(e^o)$ and $\mathbf{B}'_\alpha(e^o)$ are obtained in a similar manner.] For example, $\mathbf{N}'_u(e^o)$ is represented by the matrix

$$\mathbf{N}'_u(e^o) = [L_{ij}(e^o)]; \quad L_{ij}(e^o) = (N_i)'_u(e^o); \quad i, j = 1, \dots, K. \quad (11)$$

In view of Eqs. (9) and (10), the "forward sensitivity equations" [given in Eqs. (6) and (7)] become

$$\mathbf{N}'_u(e^o)h_u = VQ(\alpha^o; h_\alpha) - \mathbf{N}'_\alpha(e^o)h_\alpha \quad (12)$$

and

$$\{\mathbf{B}'_u(e^o)h_u\}_{\partial\Omega} = \{VA(\alpha^o; h_\alpha) - \mathbf{B}'_\alpha(e^o)h_\alpha\}_{\partial\Omega}. \quad (13)$$

Although $\mathbf{N}'_u(e^o), \mathbf{B}'_u(e^o) \in L(E_u, E_Q)$, further progress toward constructing the desired adjoint system can be made only if $\mathbf{N}'_u(e^o)$ is densely defined and the underlying normed linear spaces are complete. (Otherwise, of course, adjoint operators cannot be uniquely determined.) Since the lack of an inner product in a general Banach space gives rise to significant conceptual distinctions between the adjoint of a linear operator on a Banach space and the adjoint of a linear operator on a Hilbert space, the choice of space becomes important for subsequent derivations. To motivate the choice to be made here, it is recalled that all of the previous approaches to sensitivity theory made use of real inner products. Therefore, clarification of the conditions underlying the validity of these approaches is facilitated by the simplifying properties of Hilbert spaces. Specifically, the spaces E_u and E_Q are henceforth required to be real Hilbert spaces, denoted by H_u and H_Q , respectively. The inner products on H_u and H_Q are denoted by \langle, \rangle and $(,)$ respectively.

Since Hilbert spaces are self-dual, the following relationship holds for a vector $v \in H_Q$:

$$(v, \mathbf{N}'_u(e^o)h_u) = \langle L^*(e^o)v, h_u \rangle + \{P[h_u, v]\}_{\partial\Omega}. \quad (14)$$

In the above equation, the operator $L^*(e^o)$ is the $K \times K$ matrix

$$L^*(e^o) = [L_{ji}^*(e^o)], \quad i, j = 1, \dots, K \quad (15)$$

obtained by transposing the formal adjoints of the operators $L_{ij}(e^o)$, and $\{P[h_u, v]\}_{\partial\Omega}$ is the associated bilinear form evaluated on $\partial\Omega$. The domain of L^* is determined by selecting appropriate adjoint boundary conditions, represented here in operator form as

$$[\mathbf{B}^*(v; e^o) - \mathbf{A}^*(e^o)]_{\partial\Omega} = 0. \quad (16)$$

These boundary conditions are obtained by requiring that

1. they be independent of h_u, h_α , and G-derivatives with respect to α , and
2. the substitution of Eqs. (13) and (16) into the expression of $\{P[h_u, v]\}_{\partial\Omega}$ must cause all terms containing unknown values of h_u to vanish.

This selection of the adjoint boundary conditions reduces $\{P[h_u, v]\}_{\partial\Omega}$ to a quantity designated here by $\hat{P}[h_{\alpha, v}; e^o]$, where \hat{P} contains boundary terms involving only known values of $h_{\alpha, v}$, and (possibly) e^o . In general, \hat{P} does not automatically vanish as a result of these manipulations,³³ although it may do so in particular instances. Hence, Eq. (14) can also be written as

$$(v, N'_u(e^o)h_u) = \langle L^*(e^o)v, h_u \rangle + \hat{P}[h_{\alpha, v}; e^o] . \quad (17)$$

The above equation can be further transformed by recalling Eq. (12); then Eq. (17) becomes

$$\langle L^*(e^o)v, h_u \rangle = (v, VQ(\alpha^o; h_u) - N'_u(e^o)h_u) - \hat{P}[h_{\alpha, v}; e^o] . \quad (18)$$

At this stage in the development of the Adjoint Method, we first examine the special case when E_R is simply the field of reach scalars, denoted by Λ , so that the system's response reduces to the nonlinear functional (rather than operator) $R: D \rightarrow \Lambda$. This will facilitate the subsequent generalization to the operator case $R: D \rightarrow E_R$.

III.A.2.b.(i) System Responses: Functionals

When $R: D \rightarrow \Lambda$, the sensitivity $VR(e^o; h)$ also reduces to a functional that takes values in Λ . We now note that the right-hand side of Eq. (18) does not contain any values of h_u . Thus, if in the functional $VR(e^o; h)$ the h_u dependence could be separated from the h_α dependence, and the quantity containing this h_u dependence could be expressed in terms of the left-hand side of Eq. (18), then the construction of the Adjoint Method would be concluded. However, $\langle L^*(e^o)v, h_u \rangle$ is linear in h_u , while in general, $VR(e^o; h)$ is not. For $VR(e^o; h)$ to be linear in h (and, consequently, in h_u), it becomes apparent that $R(e)$ must be required to satisfy the same conditions as those required of $F(e)$ in Eq. (8). Then, the linear G differential $VR(e^o; h)$ is denoted by $DR(e^o; h)$, and can be expressed as

$$DR(e^o; h) = R'_u(e^o)h_u + R'_\alpha(e^o)h_\alpha , \quad (19)$$

where $R'_u(e^o)$ and $R'_\alpha(e^o)$ are, respectively, the partial G derivatives at e^o of $R(e)$ with respect to u and α .

As desired, the h_u dependence has been separated from the h_α dependence. Note here that, historically, quantities corresponding to the functions $R'_u(e^o)h_u$ and $R'_\alpha(e^o)h_\alpha$ have been referred to as the "indirect effect term" and the "direct effect term," respectively. This terminology reflects the fact that in the previous works¹⁻¹⁹ the response was considered to depend on α both "directly" and "indirectly" -- via the state vector u , i.e., the response was considered to be a mapping from the space of the input parameters into the real numbers. Although this interpretation of the response is in contradistinction with the concepts introduced and employed in this work, it is still convenient to continue to use this traditional terminology when referring to $R'_u(e^o)h_u$ and $R'_\alpha(e^o)h_\alpha$.

Since the functional $R'_u(e^o)h_u$ is linear in h_u and since Hilbert spaces are self-dual, the Riesz representation theorem ensures that there exists a unique vector $\nabla_u R(e^o) \in H_u$ such that

$$R'_u(e^o)h_u = \langle \nabla_u R(e^o), h_u \rangle, \quad h_u \in H_u. \quad (20)$$

At this stage, it can be required that the right-hand side of Eq. (20) and the left-hand side of Eq. (18) represent the same functional. Then, the Riesz representation theorem ensures that the relationship

$$L^*(e^o)v = \nabla_u R(e^o) \quad (21)$$

holds uniquely, where v satisfies the boundary conditions given in Eq. (16).

The construction of the desired adjoint system — consisting of Eqs. (21) and (16) — has thus been completed. Furthermore, the desired elimination of the unknown values of h_u from the expression giving the sensitivity $DR(e^o;h)$ of $R(e)$ at e^o to variations of h_α has also been accomplished, since in view of Eqs. (18)-(21),

$$DR(e^o;h) = R'_\alpha(e^o)h_\alpha + (v, VQ(e^o;h_\alpha) - N'_\alpha(e^o)h_\alpha) - \hat{P}[h_\alpha, v; e^o]. \quad (22)$$

Once the single calculation to determine the adjoint function v is performed, Eq. (22) provides the most efficient means to obtain the sensitivity $DR(e^o;h)$ of $R(e)$. However, it is important to reemphasize that Eq. (22) holds if and only if all the requirements imposed in this section on the various operators are satisfied.

III.A.2.b.(ii) System Responses: Operators

The analysis — presented in the previous section — of the necessary and sufficient conditions underlying the validity of the Adjoint Method for responses that are functionals also establishes the guidelines for treating operator responses; in this case, the sensitivity $VR(e^o;h)$ is itself an operator. From the developments presented in III.A.2.b.(i), the following guidelines emerge for developing the Adjoint Method for operator responses:

(G.1) isolate the h_u dependence of $VR(e^o;h)$ from the functional dependence of $VR(e^o;h)$ on the remaining quantities,

(G.2) express the quantity containing this h_u dependence in the form of linear combinations of functionals that are themselves linear in h_u ,

(G.3) employ the Adjoint Method for functionals to evaluate the functionals determined in item (G.2) above.

The development of these guidelines into a rigorous formalism will necessarily involve the use of adjoint operators. Since adjoint operators in Hilbert spaces are more convenient to deal with than adjoint operators in Banach spaces, the subsequent developments are facilitated by taking advantage of the simplifying geometrical properties of Hilbert spaces while still retaining sufficient generality for practical applications. In this vein, the spaces E_u, E_Q , and E_R are henceforth considered Hilbert spaces and denoted as $H_u(\Omega), H_Q(\Omega)$ and $H_R(\Omega_R)$, respectively. The elements of $H_u(\Omega)$ and $H_Q(\Omega)$ are, as before, vector functions defined on the open set $\Omega \subset R^J$ with the smooth boundary $\partial\Omega$. The elements of $H_R(\Omega_R)$ are vector or scalar functions defined on the open set $\Omega_R \subset R^m, 1 \leq m \leq J$, with a smooth boundary $\partial\Omega_R$. (Of course, if $J = 1$, then $\partial\Omega$ merely consists of two endpoints; similarly, if $m = 1$, then $\partial\Omega$ consists of two endpoints only.) The inner products on $H_u(\Omega), H_Q(\Omega)$, and $H_R(\Omega_R)$ are denoted by $[\cdot, \cdot], < \cdot, \cdot >$, and $\{ \cdot, \cdot \}$, respectively.

In view of the foregoing guidelines (G.1) and (G.2), it becomes apparent that further progress is possible only if $VR(e^o;h)$ is linear in h . Applying Theorem II.F.2 readily shows that $VR(e^o;h)$ is linear in h if and only if

$R(e)$ satisfies a weak Lipschitz condition at e^o , and

$$R(e^o + th_1 + th_2) - R(e^o + th_1) - R(e^o + th_2) + R(e^o) = o(t); h_1, h_2 \in H_u \times H_\alpha; t \in \Lambda. \quad (23)$$

In such a case, $VR(e^o;h)$ is denoted by $DR(e^o;h)$, and $R(e)$ admits a total G derivative at $e^o = (u^o, \alpha^o)$. It follows that the relationship

$$DR(e^o;h) = R'_u(e^o)h_u + R'_\alpha(e^o)h_\alpha \quad (24)$$

holds, where $R'_u(e^o)$ and $R'_\alpha(e^o)$ are the partial G derivatives at e^o of $R(e)$ with respect to u and α .

With the derivation of Eq. (24), the task outlined in guideline (G.1) has been completed, and Eq. (23) gives the necessary and sufficient conditions underlying this completion. Note also that $R'_u(e^o)$ is a linear operator from H_u into H_R , i.e., $R'_u(e^o) \in L(H_u(\Omega), H_R(\Omega_R))$. By analogy to the particular case when the response is a functional [cf. Eq. (19) *et seq.*], it is still convenient to refer to the quantities $R'_u(e^o)h_u$ and $R'_\alpha(e^o)h_\alpha$ appearing in Eq. (24) as the "indirect effect term" and the "direct effect term", respectively.

The direct effect term can be evaluated efficiently at this stage. To proceed with the evaluation of the indirect effect term, consider that the orthonormal set $\{\phi_k\}_{k \in K}$, where k runs through an index set K , is an orthonormal basis of $H_R(\Omega_R)$. Then, since $R'_u(e^o)h_u \in H_R(\Omega_R)$, it follows that

$$R'_u(e^o)h_u = \sum_{k \in K} \{R'_u(e^o)h_u, \phi_k\} \phi_k. \quad (25)$$

The notation $\sum_{k \in K}$ is used to signify that in the above sum only an at most countable number of elements are different from zero, and the series extended upon the nonzero elements converges unconditionally. According to customary terminology, the functionals $\{R'_u(e^o)h_u, \phi_k\}$ are called the Fourier coefficients (in this case, of $R'_u(e^o)h_u$) with respect to the basis $\{\phi_k\}$. These functionals are linear in h_u since $R(e)$ was required to satisfy the conditions stated in Eq. (23). Thus, the derivation of Eq. (25) has completed the task outlined in guideline (G.2).

To accomplish the task outlined in guideline (G.3), it is first recalled that the Adjoint Method for functionals required the indirect effect term to be represented as an inner product of h_u with an appropriately defined vector in H_u [cf. Eq. (20)]. This indicates that progress can be made here only if each of the functionals in Eq. (25) is expressed as an inner product of h_u with a uniquely defined vector in $H_u(\Omega)$ yet to be determined.

The construction of the aforementioned inner products can readily be accomplished with the help of the operator adjoint to $\mathbf{R}'_u(e^o)$. Since $\mathbf{R}'_u(e^o) \in L(H_u(\Omega), H_R(\Omega_R))$, and since Hilbert spaces are self-dual, the adjoint of $\mathbf{R}'_u(e^o)$ is the operator $\mathbf{M}(e^o) \in L(H_R(\Omega_R), H_u(\Omega))$ defined by means of relationship

$$\{\mathbf{R}'_u(e^o)h_u, \phi_k\} = [h_u, \mathbf{M}(e^o)\phi_k], \quad k \in K. \quad (26)$$

The operator $\mathbf{M}(e^o)$ is unique if $\mathbf{R}'_u(e^o)$ is densely defined.

The adjoint sensitivity formalism for functional can now be used to construct the adjoint system whose solution will subsequently enable the elimination of unknown values of h_u from the expression of each functional $[h_u, \mathbf{M}(e^o)\phi_k], k \in K$. To construct this system, the necessary and sufficient conditions underlying the validity of Eqs. (12) and (13) must be satisfied. Then, for every vector $z_k \in H_Q, k \in K$, the following relationship holds:

$$\langle z_k, \mathbf{N}'_u(e^o)h_u \rangle = [\mathbf{L}^*(e^o)z_k, h_u] + \{P(h_u; z_k)\}_{\partial\Omega}, \quad k \in K \quad (27)$$

where $\mathbf{L}^*(e^o)$ is the operator formally adjoint to $\mathbf{N}'_u(e^o)$, and $\{P(h_u; z_k)\}_{\partial\Omega}$ is the associated bilinear form evaluated on $\partial\Omega$. The adjoint boundary conditions which determine the domain of $\mathbf{L}^*(e^o)$ are obtained by requiring that they satisfy criteria analogous to the criteria satisfied by the adjoint boundary conditions given in Eq. (16). From this requirement and from the fact that Eqs. (27) and (15) are formally identical, it follows that the desired adjoint boundary conditions are formally identical to the boundary conditions given in Eq. (15) and can be expressed as

$$\{\mathbf{B}^*(z_k; e^o) - \mathbf{A}^*(e^o)\}_{\partial\Omega} = 0, \quad k \in K. \quad (28)$$

As before, selecting the adjoint boundary conditions given in Eq. (28) reduces the bilinear form $\{P(h_u; z_k)\}_{\partial\Omega}$ appearing in Eq. (27) to $\hat{P}(h_{\alpha}, z_k; e^o)$. In view of this and Eq. (12), Eq. (27) becomes

$$[\mathbf{L}^*(e^o)z_k, h_u] = \langle z_k, \mathbf{VQ}(\alpha^o; h_{\alpha}) - \mathbf{N}'_{\alpha}(e^o)h_{\alpha} \rangle - \hat{P}(h_{\alpha}, z_k; e^o), \quad k \in K. \quad (29)$$

Comparing the left-hand side of Eq. (29) with the right-hand side of Eq. (26) shows that

$$\mathbf{L}^*(e^o)z_k = \mathbf{M}(e^o)\phi_k, \quad k \in K. \quad (30)$$

This relationship holds uniquely in view of the Riesz representation theorem.

The construction of the desired adjoint system, consisting of Eq. (30) and the boundary conditions given in Eq. (28) has thus been completed. Furthermore, Eqs. (24), (25), (26), (29), and (30) can now be used to obtain the following expression for the sensitivity $DR(e^o;h)$ of $\mathbf{R}(e)$ at e^o :

$$DR(e^o;h) = \mathbf{R}'_u(e^o)h_u + \sum_{k \in K} [\langle z_k, VQ(\alpha^o;h_u) - N'_u(e^o)h_u \rangle - \hat{P}(h_u, z_k; e^o)] \phi_k . \quad (31)$$

This accomplishes the desired elimination of all unknown values of h_u from the expression giving the sensitivity of $\mathbf{R}(e)$ at e^o . Note that Eq. (31) includes the particular case of functional-type responses. In such a case, the summation $\sum_{k \in K}$ would only contain a single term, and the derivations presented in this section would reduce to those presented in the previous section.

To evaluate the sensitivity $DR(e^o;h)$ by means of Eq. (31), it is required to compute as many adjoint functions z_k from Eqs. (28) and (30) as there are nonzero terms in the representation of $\mathbf{R}'_u(e^o)h_u$ given in Eq. (25). Although the linear combination of basis elements ϕ_k given in Eq. (25) may, in principle, contain infinitely many terms, obviously only a finite number of the corresponding adjoint functions z_k can be calculated in practice. Therefore, special attention is required to select the Hilbert space $H_R(\Omega_R)$, a basis $\{\phi_k\}_{k \in K}$, and a notion of convergence to best suit the problem at hand. This selection is guided by the need to represent the indirect effect term $\mathbf{R}'_u(e^o)h_u$ as accurately as possible with the smallest number of basis elements; a related consideration is the viability of deriving bounds and/or asymptotic expressions for the remainder after truncating Eq. (25) to the first few terms.

III.A.3. Comparative Discussion of Previous Approaches to Sensitivity Analysis

In all of the works based on the differential^{15,17} and the generalized perturbation theory^{3,8,13,16,18,19} approaches to sensitivity analysis, the problems were *a priori* considered to depend explicitly and implicitly through the state functions on the system parameters. [The terminology "generalized perturbation theory" is customarily used in works on reactor theory³ to denote that the perturbation estimate obtained accounts not only for effects resulting directly from the alteration of the system parameters (i.e., "perturbation theory") but also for indirect effects arising from the changes in the state function (i.e., the dependent variable) due to the system alteration, without explicitly calculating the altered state function.] This would conceptually correspond to interpreting the problem under consideration (including the response) as a complicated mapping of a subset $D_a \subset E_a$ into the set Λ of real numbers.

Consequently, in order to obtain expressions for the sensitivity coefficients, the respective derivations must rely explicitly and/or implicitly on the existence and uniform continuity of the derivatives of the operators and the state functions with respect to the system parameters (and, possibly, with respect to the phase-space variables).

In the works dealing with nonlinear problems, it was further stated that the "differentiated equations"^{15,17} (obtained by formally differentiating the nonlinear operator equations and response with respect to an arbitrary input parameter) or, correspondingly, the "equations for the altered state functions"^{13,16} (obtained by formal first-order perturbation theory expansions around the "base-case configuration" of the state functions and input parameters) are linear. In fact, these equations correspond conceptually to our "forward sensitivity equations" given in Eqs. (6) and (7). This correspondence makes the conditions underlying the validity of the "differentiated equations" or the "equations for the altered state functions" become evident: as derived,^{13,15-17} these equations are rigorously valid only if the input parameters are real scalars, if the derivatives of the various

state functions with respect to these input parameters are uniformly continuous, and if all operators (including the response) appearing in the formulation of the problem under consideration admit Fréchet derivatives^{22,25} with respect to the state functions. It should also be mentioned that, in these works,¹⁵⁻¹⁷ the adjoint system was always assumed to exist, and was introduced in a heuristic manner with initially unspecified source terms. These source terms were subsequently identified with the "derivatives of the response with respect to the state functions" by making use of inner products. Again, linearity of this "response derivative"^{15,17} (or, correspondingly, linearity of the "response perturbation" with respect to the "perturbations in the state functions"^{13,16}) was implicitly assumed. Furthermore, the uniqueness of the end products (e.g., adjoint systems, sensitivities) was assumed but not actually demonstrated.

The variational approaches^{3,7,10,14} relied on constructing an appropriate variational functional, which was subsequently required to satisfy a stationarity condition for the base-case values of the state functions and system parameters. Expressions for the sensitivity coefficients then resulted from this requirement. In the earlier formulations [see, e.g., Ref. 3, p 6], an unspecified function appeared in the expression of the variational functional to be made stationary. This function was subsequently identified with the "adjoint" function that satisfied an "adjoint system" whose existence was *a priori* assumed. Significant advances were made (see, e.g., Stacey's review³) in modifying earlier variational principles by using Lagrange multipliers so that restrictions which are mathematically necessary to impose on the class of trial functions correspond to the physical conditions associated with the original problem and, just as important, so that the constraints are directly incorporated in the variational principle. Although considerable ingenuity is always required to construct an appropriate variational functional -- whose explicit form depends on the problem under consideration -- these variational approaches did not require (in principle) the existence of derivatives of the state functions with respect to the system parameters. In this sense, the assumptions underlying these variational approaches are less restrictive than the assumptions underlying the previously mentioned differential and generalized perturbation theory approaches. However, derivatives of the various operators with respect to the state functions and the system parameters were still needed. Although the exact nature of these derivatives (and, consequently, the necessary and sufficient conditions underlying their existence) were not generally analyzed, Stacey defines and employs a quantity referred to in his work³ as the "variation of a functional." In the light of the concepts of nonlinear functional analysis,^{22,25} it becomes apparent that his definition is in fact the definition of the Fréchet differential of that functional. This implies that the "functional derivatives" encountered in these variational approaches^{3,14} must be interpreted as Fréchet derivatives.

It is noted that these approaches¹⁻¹⁹ to sensitivity analysis were developed to analyze specific practical problems encountered in reactor physics, shielding, depletion, and heat transfer. These specific problems involved sufficiently well-behaved operators, and the parameters considered for sensitivity analysis were, in fact, real scalars. Therefore, even though the derivation underlying these approaches are mathematically not entirely rigorous, the end results are essentially correct.

In reformulating both the differential and the variational approaches to sensitivity analysis of nonlinear systems of equations, Cacuci *et al.*²¹ considered a typical nonlinear problem as a mapping defined on a product space corresponding to $E = E_u \times E_\alpha$ as defined in Sec. III.A.1. (Note, however, that these spaces were considered at the outset to be Hilbert spaces.) This completely eliminated the need for the existence of derivatives of the state vector with respect to the system parameters. In addition, the definition of sensitivity of a response was generalized to allow consideration of system parameters that were functions rather than just scalars. By requiring the existence of partial Fréchet derivatives^{22,25} of the operators with respect to the state vector and the system parameters, the existence of an appropriate adjoint system was ensured. Although this work

generalized and extended the scope of the previously available sensitivity theory formulations, the existence of partial Fréchet derivatives is not actually essential for sensitivity analysis; as shown in Sec. III.A.2, the existence of the G differentials – for the Forward Method – or of the partial derivatives with respect to the state vectors only – for the Adjoint Method – are both necessary and sufficient.

Although the concept of an inner product has been essential to formulating the existing adjoint-function based approaches^{1-19,21} to sensitivity analysis, the implications associated with the particular use of this concept in these works have not been generally discussed. Clearly, the prerequisite for employing an inner product is that the problem under consideration must be formulated in an appropriate Hilbert (or at least pre-Hilbert) space. Furthermore, since a single definition for the inner product was used in each of these works when introducing adjoint operators, the underlying implication is that the problem being analyzed can only involve operators with ranges in the same Hilbert space to which the state vector belongs. By contrast, the Adjoint Method developed here makes use of two distinct inner products [cf, Eq. (14) *et seq.*]; this allows sensitivity analysis of problems involving operators whose ranges may be in a Hilbert space that differs from the Hilbert space to which either the state vector or the system parameters belong. Also, it is noted that no distinctions were made in previous works^{1-19,21} regarding the fundamental mathematical differences between the requirements underlying the "adjoint" formulations of sensitivity theory. The present work provides a basis for assessing the potentially important practical consequences of these differences.

The forgoing discussion has highlighted the major aspects regarding the specific uses of perturbation theory and variational approaches for applications to sensitivity analysis. For such applications, the common scope of these approaches is to obtain sensitivities. In reactor theory, for example, some authors^{3,34} regarded perturbation theory as an application of variational methods in the sense that a variational formulation "is employed to derive a generalized perturbation theory for estimating the change in the physical quantity of interest which would take place if the properties of the system were to be altered" (Ref. 3, p. 18). But the general uses of either perturbation theory or variational methods are not limited to deriving sensitivity functions. Similarities as well as distinctions between the perturbation theory and the variational approaches to sensitivity analysis, and the contributions that this work brings to sensitivity theory can be further clarified by briefly analyzing the relationships between perturbation theory, variational methods, and functional analysis from a broader perspective.

Perturbation theory and variational methods are not sharply defined disciplines; they are bodies of knowledge unified more by the respective method of approach than by clear-cut demarcation of their respective provinces. For example, the terminology "perturbation theory" is also encountered in celestial mechanics and in nonlinear oscillation theory. However, although these "perturbation theories" study systems deviating slightly from an ideal system for which the complete solution is known, the problems they treat and the tools they use are quite different from those used to derive sensitivities. In reactor theory, for example,³⁵ this latter use of perturbation theory has evolved from the work of Rayleigh on vibrating systems and of Schrödinger in quantum mechanics.

The works based on perturbation theory to derive sensitivities for problems involving linear operators tacitly assume that the eigenvalues and eigenvectors admit series expansions in a small parameter that measures the deviation of the "unperturbed operator" from the "unperturbed" one. Without a proof that the series actually converges, it is difficult to decide whether the first term of the series gives an adequate picture introduced by the perturbation, a fact well known in reactor theory,³⁵ for example. For applications to sensitivity analysis of problems involving linear operators, the underpinnings of the perturbation theory approach lie in linear functional analysis.

Although a systematic presentation of perturbation theory for linear operators is now available,³⁶ further work remains to be done to fully exploit these functional-analytic techniques for sensitivity analysis.

Variational methods, just like perturbation theory, are not developed specifically for sensitivity analysis, although the variational principles developed for this purpose are, of course, very useful. But variational principles, even those restricted to limited classes of variations, are difficult to formulate and for many nonlinear problems of interest (e.g., thermal hydraulics, heat and mass transfer) variational principles are not yet available.³⁷ Furthermore, a systematic and general treatment of variational principles for problems involving nonlinear operators must necessarily rely on the differential concepts of nonlinear functional analysis, i.e., Gateaux and Fréchet differentials and derivatives.

The present work attempts to provide a general framework for systematic sensitivity analysis of both linear and nonlinear systems. The scope of the theory formulated here is to derive sensitivities, to be used not only for predicting the behavior of the response when the system parameters are altered, but also for ranking the importance of these parameters, and for performing uncertainty analysis by combining the sensitivities with the appropriate parameter covariances.

The link between a rigorous perturbation theory (and/or variational) approach to sensitivity analysis and the sensitivity theory presented in this work is provided by functional analysis. In particular, the similar overall strategy and the use of adjoint operators stem from functional-analytic concepts. In this sense, the greater general validity and applicability of the present sensitivity theory also contributes to the development of perturbation theory for applications to nonlinear systems. Finally, it is noted that whenever the variational, differential, and perturbation theory approaches are rigorously applicable, the end results for the sensitivities are identical to those produced by the sensitivity theory presented in this work.

III.B. NONLINEAR SYSTEMS WITH RESPONSES DEFINED AT CRITICAL POINTS

III.B.1. Introduction

Concepts of nonlinear functional analysis have been employed in Sec. III.A to formulate a rigorous and comprehensive sensitivity theory for physical problems characterized by systems of coupled nonlinear equations. The formulation of this theory is centered on evaluating the Gâteaux (G) differential of the system's response (i.e., performance parameter) associated with the physical problem. This G-differential is a general and fundamental concept for defining the sensitivity of a response to variations in the system parameters.

As shown in Sec. III.A, a nonlinear functional can be used as a general representation for any response that is solely characterized by a numerical value. Note that only this numerical value changes when varying the system parameters. However, responses which cannot be characterized solely by a numerical value are often encountered in practice. In reactor safety and design, for example, responses of considerable interest are the maximum temperature in the cladding, the maximum power density, and the maximum normalized reactor power level (if point-kinetics equations are used in the transient reactor analysis code). Such responses are characterized both by the numerical value at the maximum and by the position in phase-space where the maximum occurs. In this case, varying the system parameters alters not only the value at this maximum but also alters the position of the maximum in phase-space. This is illustrated in the following section, where sensitivity theory is extended to allow treatment of responses that comprise, as particular cases, the representation of maxima, minima, and saddle points.

Although the responses treated in this section differ from those treated in Sec. III.A, the physical problem is the same as in that section. It is helpful to recall that, in the operator notation used in Sec. III.A, the problem is represented by the system of K coupled nonlinear equations

$$\mathbf{N}[u(x), \alpha(x)] = \mathbf{Q}[\alpha(x), x], \quad x \in \Omega \quad (1)$$

subject to boundary and initial conditions represented as

$$\{\mathbf{B}(e) - \mathbf{A}(\alpha)\}_{\partial\Omega} = \mathbf{O}, \quad (2)$$

where \mathbf{A} and \mathbf{B} are operators and $\partial\Omega$ is the boundary of Ω ; the operator $\mathbf{A}(\alpha)$ represents all inhomogeneous boundary terms.

III.B.2. System Response: A Functional Defined at a Critical Point of a Function of the System's State Vector and Parameters

Consider the system response R to be a functional of $e = (u, \alpha)$ defined at a critical point $y(\alpha)$ of a function $F(u, x, \alpha)$. Such a response can be represented as a functional of the form

$$R(e) = \int_{\Omega} F(u, x, \alpha) \prod_{i=1}^M \delta[x_i - y_i(\alpha)] \prod_{j=M+1}^J \delta(x_j - z_j) dx. \quad (3)$$

The quantities appearing in the integrand of Eq. (3) are defined as follows:

1. F is the nonlinear function under consideration.
2. $\delta(x)$ is the customary "delta" functional.
3. $\alpha \in R^I$, i.e., the components $\alpha_i, i = 1, \dots, I$, are restricted throughout this section to be real numbers.
4. $y(\alpha) = [y_1(\alpha), \dots, y_M(\alpha)]$, $M \leq J$, is a critical point of F . This critical point is defined here in one of the following two ways:
 - a. If the G-differential of F vanishes at $y(\alpha)$, then $y(\alpha)$ is a critical point defined implicitly as the solution of the system of equations.

$$\left\{ \frac{\partial F}{\partial x_i} \right\}_{y(\alpha)} = 0; \quad i = 1, \dots, J. \quad (4)$$

In this case, $y(\alpha)$ has J components, (i.e., $M = J$), and $\prod_{j=M+1}^J \delta(x_j - z_j) \equiv 1$ in the integrand of Eq. (3). Note that, in general, $y(\alpha)$ is a function of α .

- b. Occasionally, it may happen that $\partial F / \partial x_j$ takes on nonzero constant values (i.e., values that do not depend on x) for some of the variables $x_j, j = M + 1, \dots, J$. Then as a function of these variables x_j, F attains its extreme values at the points $x_j = z_j, z_j \in \partial\Omega$. Evaluating F at $z_j, j = M + 1, \dots, J$, yields a function G which depends on the remaining phase-space variables $x_i, i = 1, \dots, M$; G may then have a critical point at $y(\alpha) = [y_1(\alpha), \dots, y_M(\alpha)]$ defined implicitly as the solution of

$$\left\{ \frac{\partial G}{\partial x_i} \right\}_{y(\alpha)} = 0; \quad i = 1, \dots, M. \quad (5)$$

With the above specifications, the definition of $R(e)$ given in EQ. (3) is sufficiently general to include treatment of extrema (local, relative, or absolute), saddle, and inflexion points of the function F of interest. In practice, the base-case solution path, and therefore the specific nature and location of the critical point under consideration, are completely known prior to initiating the sensitivity studies.

It is thus apparent that in the formulation of a complete sensitivity theory, the components $y_i(\alpha), i = 1, \dots, M$, must be treated as responses in addition to $R(e)$. Hence, the objective of this sensitivity theory is twofold:

1. to determine the G-differential $VR(e^0; h)$ of $R(e)$ at the "base-case configuration point" $e^0 = (u^0, \alpha^0)$, which gives the sensitivity of $R(e)$ to changes $h = (h_u, h_\alpha)$ in the system's state functions and parameters, and
2. to determine the (column) vector $Vy(\alpha^0; h_\alpha) = (Vy_1, \dots, Vy_M)$ whose components $Vy_m(\alpha^0; h_\alpha)$ are the G-differentials of $y_m(\alpha)$ at α^0 , for $m = 1, \dots, M$. The vector $Vy(\alpha^0; h_\alpha)$ gives the sensitivity of the critical point $y(\alpha)$ to changes h_α .

To achieve the above objective, the "Forward Method" and the "Adjoint Method" will be developed along the same general lines as discussed in detail in Sec. III.A.

III.B.3. The Forward Method

Applying the definition of the G-differential to Eq. (3) shows that

$$\begin{aligned}
 VR(e^0;h) &= \int_{\Omega} [F'_u(e^0)h_u + F'_\alpha(e^0)h_\alpha] \prod_{i=1}^M \delta[x_i - y_i(\alpha^0)] \prod_{j=M+1}^J \delta(x_j - z_j) dx \\
 &+ \sum_{m=1}^M \left[-h_\alpha \left(\frac{dy_m}{d\alpha} \right)_{\alpha^0} \right] \int_{\Omega} F\delta'(x_m - y_m) \prod_{i=1, i \neq m}^M \delta(x_i - y_i) \prod_{j=M+1}^J \delta(x_j - z_j) dx.
 \end{aligned} \tag{6}$$

The last term on the right side of Eq. (6) vanishes, since

$$\begin{aligned}
 &\int_{\Omega} F\delta'(x_m - y_m) \prod_{i=1, i \neq m}^M \delta(x_i - y_i) \prod_{j=M+1}^J \delta(x_j - z_j) dx = \\
 &- \int_{\Omega} (\partial F / \partial x_m) \prod_{i=1}^M \delta(x_i - y_i) \prod_{j=M+1}^J \delta(x_j - z_j) dx = 0, \quad m = 1, \dots, M,
 \end{aligned} \tag{7}$$

in view of the well-known definition of the δ' functional and in view of either Eq. (4) if $M = J$, or of Eq. (5) if $M < J$. Therefore, the expression of $VR(e^0;h)$ simplifies to

$$VR(e^0;h) = \int_{\Omega} [F'_u(e^0)h_u + F'_\alpha(e^0)h_\alpha] \prod_{i=1}^M \delta[x_i - y_i(\alpha^0)] \prod_{j=M+1}^J \delta(x_j - z_j) dx. \tag{8}$$

Thus, the sensitivity $VR(e^0;h)$ of $R(e)$ to specified changes h_α can in principle be evaluated once the vector h_u is determined from the "forward sensitivity equations," i.e.,

$$VN(u^0, \alpha^0; h_u, h_\alpha) - VQ(\alpha^0; h_\alpha) = \mathbf{0} \tag{9}$$

$$\{VB(u^0, \alpha^0; h_u, h_\alpha) - VA(\alpha^0; h_\alpha)\}_{\partial\Omega} = \mathbf{0} .$$

As already mentioned, the sensitivity of the location in phase space of the critical point is given by the G-differential $Vy(\alpha^0; h_\alpha)$ of $y(\alpha)$ at α^0 . In view of either Eq. (4) or Eq. (5), each of the components $y_1(\alpha), \dots, y_M(\alpha)$ of $y(\alpha)$ is a real-valued function of the real variables $\alpha_1, \dots, \alpha_I$, and may be viewed as a functional defined on a subset of R^I . Therefore, each G-differential $Vy_m(\alpha^0; h_\alpha)$ of $y_m(\alpha)$ at α^0 is given by

$$Vy_m(\alpha^0; h_\alpha) = \left\{ \frac{dy_m}{d\alpha} \right\}_{\alpha^0} \cdot h_\alpha = \sum_{i=1}^I \left\{ \frac{\partial y_m}{\partial \alpha_i} \right\}_{\alpha^0} h_{\alpha_i}; \quad m = 1, \dots, M, \tag{10}$$

provided that $\partial y_m / \partial \alpha_i, i = 1, \dots, I$, exist at α^0 for all $m = 1, \dots, M$.

The explicit expression of $Vy(\alpha^0; h_\alpha)$ is obtained as follows. First, it is observed that both Eq. (4) and Eq (5) can be represented as

$$\int_{\Omega} (\partial F / \partial x_m) \prod_{i=1}^M \delta[x_i - y_i(\alpha)] \prod_{j=M+1}^J \delta(x_j - z_j) dx = 0; \quad m = 1, \dots, M. \quad (11)$$

Taking the G-differential of Eq. (11) at e^0 yields the following system of equations involving the components Vy_m :

$$\begin{aligned} & \int_{\Omega} \{ \partial(F'_u h_u + F'_\alpha h_\alpha) / \partial x_m \}_{e^0} \prod_{i=1}^M \delta[x_i - y_i(\alpha^0)] \prod_{j=M+1}^J \delta(x_j - z_j) dx \\ & - \sum_{s=1}^M Vy_s(\alpha^0; h_\alpha) \int_{\Omega} \{ \partial F / \partial x_m \}_{e^0} \delta'[x_s - y_s(\alpha^0)] \prod_{i=1, i \neq s}^M \delta[x_i - y_i(\alpha^0)] \\ & \quad \prod_{j=M+1}^J \delta(x_j - z_j) dx = 0; \quad m = 1, \dots, M. \end{aligned} \quad (12)$$

The above system is algebraic and linear in the components $Vy_s(\alpha^0; h_\alpha)$; therefore, it can be represented in matrix form as

$$\Phi Vy = \Gamma \quad (13)$$

by defining $\Phi = [\phi_{ms}]$ to be the $M \times M$ matrix with elements

$$\phi_{ms} \equiv \int_{\Omega} \{ \partial^2 F / \partial x_m \partial x_s \}_{e^0} \prod_{i=1}^M \delta[x_i - y_i(\alpha^0)] \prod_{j=M+1}^J \delta(x_j - z_j) dx, \quad \text{for } m, s = 1, \dots, M. \quad (14)$$

and by defining Γ to be the M -component (column) vector

$$\Gamma \equiv (f_1 + g_1, \dots, f_M + g_M)^T, \quad (15)$$

where

$$f_m \equiv - \int_{\Omega} \{ \partial(F'_\alpha h_\alpha) / \partial x_m \}_{e^0} \prod_{i=1}^M \delta[x_i - y_i(\alpha^0)] \prod_{j=M+1}^J \delta(x_j - z_j) dx; \quad m = 1, \dots, M, \quad (16)$$

and

$$g_m \equiv - \int_{\Omega} \{ \partial(F'_u h_u) / \partial x_m \}_{e^0} \prod_{i=1}^M \delta[x_i - y_i(\alpha^0)] \prod_{j=M+1}^J \delta(x_j - z_j) dx; \quad m = 1, \dots, M. \quad (17)$$

Notice that the definition of the δ' functional has been used to recast the second integral in Eq. (12) into the equivalent expression given in Eq. (14).

At this stage, the quantities ϕ_{ms} and f_m can be evaluated most efficiently by directly using Eqs. (14) and (16). It is of interest to observe here that if $M = J$, then Φ is the Hessian of F evaluated at the critical point $y(\alpha^0)$; alternatively, if $M < J$, then Φ is the Hessian of the function G [considered in Eq. (5)] evaluated at the respective critical point. The quantities g_m defined in Eq. (17) can also be evaluated, since h_u will have already been determined to compute the sensitivity $VR(e^0;h)$ given in Eq. (8). Upon completing the computation of the elements of Φ and Γ , Eq. (13) can be solved by employing methods of linear algebra to obtain

$$Vy(\alpha^0;h_\alpha) = \Phi^{-1}\Gamma. \quad (18)$$

As underscored by the derivations presented so far, the availability of the solution h_u of the "forward sensitivity equations" given in Eq. (9) is essential to evaluate both $VR(e^0;h)$ and $Vy(\alpha^0;h_\alpha)$. This is a distinctive characteristic of the Forward Method which, from an economical standpoint, makes this formalism ill-suited for sensitivity analysis of problems with large data bases (i.e., when α has many components).

III.B.4. The Adjoint Method

Since most of the problems encountered in practice are characterized by large data bases, the development of this formalism is motivated by the need for a tool to perform sensitivity analyses of such problems economically. To this end, the development of this formalism is centered on eliminating the explicit appearance of the unknown values of the vector h_u from Eqs. (8) and (18), and hence on circumventing the need to repeatedly solve Eq. (9). However, as detailed in Sec. III.A, h_u can be eliminated if and only if (iff) the following conditions are satisfied:

- (C.1) the partial G-derivatives at e^0 of $R(e)$ with respect to u and α exist,
- (C.2) the partial G-derivatives at e^0 of the operators \mathbf{N} and \mathbf{B} with respect to u and α exist,
- (C.3) the spaces E_u and E_Q are real Hilbert spaces, denoted by H_u and H_Q , respectively. For $u_1, u_2 \in H_u$, the inner product in H_u will be denoted by $[u_1, u_2]$ and is given by the integral $\int_{\Omega} u_1 \cdot u_2 dx$. The inner product in H_Q will be denoted by \langle, \rangle .

An examination of Eq. (8) shows that $VR(e^0;h)$ is linear in h . Hence, condition (C.1) is satisfied, and the H_u -dependent component of $VR(e^0;h)$, i.e., the "indirect effect term," can be written in inner product form as

$$\int_{\Omega} F'_u(e^0) h_u \prod_{i=1}^M \delta[x_i - y_i(\alpha^0)] \prod_{j=M+1}^J \delta(x_j - z_j) dx = [\nabla_u R(e^0), h_u], \quad (19)$$

where

$$\nabla_u R(e^0) = \prod_{i=1}^M \delta[x_i - y_i(\alpha^0)] \prod_{j=M+1}^J \delta(x_j - z_j) \left(\frac{\partial F(e^0)}{\partial u_1}, \dots, \frac{\partial F(e^0)}{\partial u_K} \right)^T. \quad (20)$$

The adjoint system is constructed by following the procedure set forth in Sec. III.A. (For brevity, details are omitted here.) Thus, condition (C.2) makes it possible to write the system of equations given in Eq. (9) as

$$\mathbf{N}'_u(e^0)h_u = V\mathbf{Q}(\alpha^0; h_\alpha) - \mathbf{N}'_\alpha(e^0)h_\alpha \quad (21)$$

and

$$\{\mathbf{B}'_u(e^0)h_u\}_{\partial\Omega} = \{V\mathbf{A}(\alpha^0; h_\alpha) - \mathbf{B}'_\alpha(e^0)h_\alpha\}_{\partial\Omega}. \quad (22)$$

Next, in view of Eq. (21) and condition (C.3), the following relationship holds for a vector $\nu \in H_Q$:

$$\langle \nu, \mathbf{N}'_u(e^0)h_u \rangle = [\mathbf{L}^*(e^0)\nu, h_u] + \{P[h_u; \nu]\}_{\partial\Omega}, \quad (23)$$

where $\mathbf{L}^*(e^0)$ is the operator formally adjoint to $\mathbf{N}'_u(e^0)$, and $\{P[h_u; \nu]\}_{\partial\Omega}$ is the associated bilinear form evaluated on $\partial\Omega$. The domain of $\mathbf{L}^*(e^0)$ is determined by selecting appropriate adjoint boundary conditions, represented here in operator form as

$$\{\mathbf{B}^*(\nu; e^0) - \mathbf{A}^*(e^0)\}_{\partial\Omega} = 0. \quad (24)$$

These boundary conditions are obtained by requiring that

1. they be independent of h_u, h_α , and G derivatives with respect to α , and
2. the substitution of Eqs. (22) and (24) into the expression of $\{P[h_u; \nu]\}_{\partial\Omega}$ must cause all terms containing unknown values of h_u to vanish.

This selection of the adjoint boundary conditions reduces $\{P[h_u; \nu]\}_{\partial\Omega}$ to a quantity designated here by $\hat{P}(h_\alpha, \nu; e^0)$, where \hat{P} contains boundary terms involving only known values of h_α, ν , and (possibly) e^0 . In general, P does not automatically vanish as a result of these manipulations, although it may do so in particular instances. Hence, Eq. (23) can be written as

$$[\mathbf{L}^*(e^0)\nu, h_u] = \langle \nu, V\mathbf{Q}(\alpha^0; h_\alpha) - \mathbf{N}'_\alpha(e^0)h_\alpha - \hat{P}(h_\alpha, \nu; e^0) \rangle, \quad (25)$$

where Eq. (21) was used to replace $\mathbf{N}'_u(e^0)h_u$. Comparing the left-hand side of Eq. (25) with the right-hand side of Eq. (19) shows that

$$\mathbf{L}^*(e^0)\nu = \nabla_u R(e^0). \quad (26)$$

Note that the uniqueness of the above relationship is ensured by the Riesz representation theorem. This completes the construction of the adjoint system. Furthermore, Eqs. (19), (25), and (26) can be used to express Eq. (8) as

$$\begin{aligned} VR(e^0; h) &= \int_\Omega F'_\alpha(e^0)h_\alpha \prod_{i=1}^M \delta[x_i - y_i(\alpha^0)] \prod_{j=M+1}^J \delta(x_j - z_j) dx \\ &+ \langle V\mathbf{Q}(\alpha^0; h_\alpha) - \mathbf{N}'_\alpha(e^0)h_\alpha, \nu \rangle - \hat{P}(h_\alpha, \nu; e^0). \end{aligned} \quad (27)$$

The desired elimination of the unknown values of h_u from the expression giving the sensitivity $VR(e^0;h)$ has thus been accomplished. Next, in view of Eq. (21) and condition (C.3), the following relationship holds for a vector $v \in H_Q$:

Unknown values of h_u can be eliminated from the expression of $Vy(\alpha^0;h_\alpha)$ given in Eq. (18), only if they can be eliminated from appearing in Eq. (17). Examination of Eq. (17) reveals that each quantity g_m is a functional that can be expressed in the equivalent form

$$g_m = \int_{\Omega} F'_u(e^0) h_u \delta'(x_m - y_m) \prod_{i=1, i \neq m}^M \delta(x_i - y_i) \prod_{j=M+1}^J \delta(x_j - z_j) dx. \quad (28)$$

by employing the definition of the δ' functional. In turn, the above expression can be written as the inner product

$$g_m = [\gamma_m(e^0), h_u], \quad (29)$$

where

$$\begin{aligned} \gamma_m(e^0) &\equiv \delta'[x_m - y_m(\alpha^0)] \prod_{i=1, i \neq m}^M \delta(x_i - y_i) \\ &\times \prod_{j=M+1}^J \delta(x_j - z_j) \left[\frac{\partial F(e^0)}{\partial u_1}, \dots, \frac{\partial F(e^0)}{\partial u_k} \right]^T. \end{aligned} \quad (30)$$

The desired elimination of the unknown values of h_u from Eq. (29) can now be accomplished by letting each of the functions $\gamma_m(e^0)$ play, in turn, the role previously played by $\nabla_u R(e^0)$ [cf. Eq. (20)], and by following the same procedure as that leading to Eq. (27). The end result is

$$g_m = \langle VQ(\alpha^0; h_\alpha) - N'_\alpha(e^0) h_\alpha, w_m \rangle - \hat{P}(h_\alpha, w_m; e^0), \quad (31)$$

where each function w_m is the solution of the adjoint system

$$\begin{aligned} \mathbf{L}^*(e^0) w_m &= \gamma_m(e^0) \\ \{\mathbf{B}^*(w_m; e^0) - \mathbf{A}^*(e^0)\}_{\partial\Omega} &= 0 \end{aligned} \quad (32)$$

for $m = 1, \dots, M$.

It is important to note that $\mathbf{L}^*(e^0)$, $\mathbf{B}^*(e^0)$, and $\mathbf{A}^*(e^0)$ appearing in Eq. (32) are the same operators as those appearing in Eqs. (26) and (24). Only the source term $\gamma_m(e^0)$ in Eq. (32) differs from the corresponding source term $\nabla_u R(e^0)$ in Eq. (26). Therefore, the computer code employed to solve the adjoint system given in Eqs. (26) and (24) can be used, with relatively trivial modifications, to compute the functions w_m from Eq. (32). Comparing now the right sides of Eqs. (25) and (31) reveals that the quantity $\hat{P}(h_\alpha, \nu; e^0)$ is formally identical to the quantity $\hat{P}(h_\alpha, w_m; e^0)$, and that the function $VQ(\alpha^0; h_\alpha) - N'_\alpha(e^0) h_\alpha$ appears in both the inner products denoted by \langle, \rangle . This indicates that the computer program employed to evaluate the second and third terms on the right side of Eq. (27) can also be used to evaluate the functionals $g_m, m = 1, \dots, M$, given in Eq. (31). Of course, the values of ν required to compute $VR(e^0;h)$ are to be replaced by the respective values of w_m when computing the g_m 's.

In most practical problems, the total number of parameters I greatly exceeds the number of phase-space variables J , and hence M , since $M \leq J$. Therefore, if the Adjoint Method can be developed as described in this section, then a large amount of computing costs can be saved by employing this formalism rather than the Forward Method. In this case, only $M + 2$ "large" computations (one for the "base-case," one for the adjoint function v , and M for the adjoint functions w_1, \dots, w_M) are needed to obtain the sensitivities $VR(e^0; h)$ and $Vy(\alpha^0; h_\alpha)$ to changes in *all* of the parameters. By contrast, $I + 1$ computations (one for the "base-case," and I to obtain the vector h_u) would be required if the Forward Method were employed.

III.B.5. Discussion

Note that, as shown in Eqs. (6)-(8), the contributions to $VR(e^0; h)$ arising from the α -dependence of $y(\alpha)$ vanish only because $y(\alpha)$ is a critical point of F . An important consequence of this fact can be demonstrated by *considering the point y not to be a function of α* . The response would then take on the form

$$R_1(e) = \int_{\Omega} F(u, x, \alpha) \prod_{i=1}^M \delta(x_i - y_i) \prod_{j=M+1}^J \delta(x_j - z_j) dx. \quad (33)$$

In the above equation, the subscript 1 indicates that the mathematical characteristics of $R_1(e)$ differ from those of $R(e)$, although both responses take on identical values at $e = e^0$, i.e.,

$$R_1(e^0) = R(e^0). \quad (34)$$

Calculating the G-differential $VR_1(e^0; h)$ of $R_1(e)$ at e^0 gives

$$VR_1(e^0; h) = \int_{\Omega} [F'_u(e)h_u + F'_\alpha(e^0)h_\alpha] \prod_{i=1}^M \delta(x_i - y_i) \prod_{j=M+1}^J \delta(x_j - z_j) dx. \quad (35)$$

Comparison of Eqs. (35) and (8) shows that

$$VR_1(e^0; h) = VR(e^0; h). \quad (36)$$

Consider now the total variations of $R(e)$ and $R_1(e)$ at $e = e^0$, i.e., $R(e^0 + h) - R(e^0) = VR(e^0; h) + \Delta(h)$, where

$$\lim_{t \rightarrow 0} [\Delta(th)/t] = 0, \quad (37)$$

and

$$R_1(e^0 + h) - R_1(e^0) = VR_1(e^0; h) + \Delta_1(h), \text{ where} \\ \lim_{t \rightarrow 0} [\Delta_1(th)/t] = 0. \quad (38)$$

Subtracting Eq. (38) from Eq. (37) and taking into account Eqs. (34) and (36), yields the relationship

$$R(e^0 + h) - R_1(e^0 + h) = \epsilon(h), \text{ where} \\ \lim_{t \rightarrow 0} [\epsilon(th)/t] = 0. \quad (39)$$

The result given in Eq. (39) can be readily strengthened if R is Fréchet differentiable [i.e., if $VR(e^0;h)$ is continuous and linear in h at e^0 , and is continuous in e at e^0]. In such a case, R_1 is also Fréchet differentiable; hence, $\lim_{t \rightarrow 0} [\Delta(th)/t] = 0$ in Eq. (37) and $\lim_{t \rightarrow 0} [\Delta_1(th)/t] = 0$ in Eq. (38) hold uniformly with respect to h on the set $\{h: \|h\| = 1\}$. Consequently, $\lim_{t \rightarrow 0} [\epsilon(th)/t] = 0$ in Eq. (39) also holds uniformly with respect to h on $\{h: \|h\| = 1\}$, and can be written in the equivalent form $\lim_{h \rightarrow 0} [|\epsilon(h)|/|h|] = 0$. Thus, the stronger result

$$\|R(e^0 + h) - R_1(e^0 + h)\| = O(\|h\|^2) \quad (40)$$

holds if R is Fréchet differentiable at $e = e^0$.

A simple illustration of the distinctions between $R(e)$ and $R_1(e)$ is shown in Fig. 1. There, the critical point $y_1(\alpha)$ of $F(u, x, \alpha)$ is a maximum occurring in the (one-dimensional) direction x_1 . Changes $h = (h_u, h_\alpha)$ would cause the new maximum of F to take on the value $R(e^0 + h)$ at $y_1(\alpha^0 + h_\alpha)$. The sensitivity $VR(e^0;h)$ of $R(e)$ at e^0 is given by Eq. (8) [or Eq. (27)], while the sensitivity $Vy_1(\alpha^0;h_\alpha)$ of $y_1(\alpha)$ at α^0 is given by Eq. (18). However, if y_1 is considered *not* to be a function of α , then $R_1(e^0 + h)$ would be the altered value of the functional $R_1(e)$. Nevertheless, the sensitivity $VR_1(e^0;h)$ of $R_1(e)$ at e^0 is the same as the sensitivity $VR(e^0;h)$ of $R(e)$ at e^0 , as shown in Eq. (36). This is only because $y_1(\alpha)$ is a critical point of $F(u, x, \alpha)$.

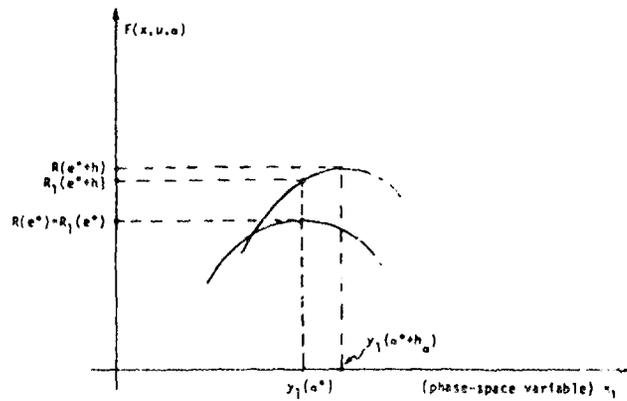


Fig. 1. Illustration of the distinction between $R(e)$ and $R_1(e)$.

III.B.6. Illustrative Application: Sensitivity Analysis of Extremum-Type Responses in Reactor Safety

In the following, we present the application of the sensitivity theory formulated so far in Section III.B to the single-phase modules of the MELT-IIIB fast reactor safety code. In particular, we present results regarding the sensitivity of the locations (in phase-space) of two important responses — the maximum fuel temperature and the maximum normalized reactor power level. The following derivations are based on the work by D. G. Cacuci, P. J. Maudlin, and C. V. Parks, "Adjoint Sensitivity Analysis of Extremum-Type Responses in Reactor Safety," *Nucl. Sci. Eng.* **83**, 112-135 (1983).

III.B.6.a. Problem Description

To determine the distribution of dependent variables, the MELT-IIIB code solves the following neutronic/thermal-hydraulic system of equations:

1. thermal-hydraulic equations, for each channel type j ($j=1,\dots,NC$), describing the average channel fuel pin and surrounding single-phase coolant
2. an equation describing the primary-loop hydraulics
3. neutron point-kinetics equations describing the reactor power.

This system of coupled nonlinear partial differential equations can be represented in operator form as

$$N[U(x),\alpha] = Q(\alpha) , \quad x \in \Omega_j , \quad (1)$$

where $U(x)$ satisfies boundary and initial conditions represented also in operator form as

$$B(U,\alpha) = A(\alpha) , \quad x \in \partial\Omega_j . \quad (2)$$

The quantities appearing in Eqs. (1) and (2) are defined as follows:

1. The quantity $x = (r,z,t)$ is the phase-space position vector whose components are the radial, axial, and time-independent variables, respectively.

2. The quantity $U = (T, T_c, T_s, P, u, n, C_1, \dots, C_{NG})$ is the state vector whose components are the dependent variables. Note that the vectors $T, T_c, T_s, P,$ and u (whose components are thermal-hydraulic dependent variables) are channel dependent; therefore, each of these vectors has NC components [e.g., $T = (T_1, \dots, T_j, \dots, T_{NC})$, where T_j refers to the temperature in the j 'th channel]. Thus, U is an M -component (column) vector, where $M = NG + 5 \times NC + 1$. The designation of each component of U is given in the NOMENCLATURE presented in Section III.B.6.h.

3. The quantity α is an I -component (column) vector whose components are the system parameters. (Here, I denotes the total number of these parameters.) Although the components of α may, in general, include functions of x and/or U , these components are restricted in this work to be real scalars. Such scalar system parameters include coefficients, scale factors, and initial conditions.

4. The quantities N and Q are M -component column vectors. The components of N and B are differential and algebraic operators acting on components of U and α . The components of Q and A represent inhomogeneous source and boundary terms, respectively. For convenience, their explicit expressions [and, consequently, the explicit form of Eqs. (1) and (2)] are given in Appendix A.

5. The domain Ω_j is the set $\Omega_j = \{r, z, t | r \in (0, R_f) \cup (R_g, R); z \in (0, L); t \in (0, t_f)\}$, and its boundary $\partial\Omega_j$ consists of the set of points $\partial\Omega_j = \{r = 0, R_f, R_g, R; z = 0, L; t = 0, t_f\}$. Thus, $x \in \Omega_j$ for the thermal-hydraulic equations, since these equations describe the physical behavior of the average channel fuel pin, surrounding coolant, and structure for each channel of type j . Similarly, $x \in \partial\Omega_j$ for the boundary and initial conditions associated with these thermal-hydraulic equations. Note that there are N_j pins in each channel j .

By contrast to the thermal-hydraulic equations, the point-kinetics equations and the primary-loop hydraulics equation apply to the total reactor domain; thus, these equations are time dependent but are channel independent. The total reactor domain, henceforth denoted by Ω , consists of the union of all the (pin) domains Ω_j , i.e.,

$$\Omega = \bigcup_{j=1}^{NC} \Omega_j \quad .$$

Thus, integrals over Ω are related to integrals over Ω_j through the relationship

$$\int_{\Omega} [] d\Omega = \sum_{j=1}^{NC} N_j \int_{\Omega_j} [] 2\pi r dr dz dt \quad . \quad (3)$$

Let $e = (U, \alpha)$ denote the concatenation of the state vector U and the vector α of system parameters. From the viewpoint of sensitivity theory, the maximum power response and the maximum fuel temperature response in any channel J are considered to be functionals of e , and are denoted by $R_n(e)$ and $R_T(e)$, respectively. Also, an examination of Eqs. (A.1), (A.6), (A.7), and (A.9) of Appendix A reveals that $n(t)$ and $T(x)$ are continuous; in particular, their first derivatives with respect to the independent variables exist at the locations where $n(t)$ and $T(x)$ attain their respective maxima.

The maximum power response can be represented as

$$R_n(e) = K \int_{\Omega} n(t) \delta[t - t_n(\alpha)] d\Omega, \quad (4)$$

where the constant

$$K \equiv 1 / \left[\sum_j N_j \pi L (R_j^2 + R^2 - R_g^2) \right]$$

serves as a normalization factor, and where $t_n(\alpha)$ represents the phase-space location of the maximum; $t_n(\alpha)$ is defined implicitly as the solution of

$$\{dn(t)/dt\} = 0 \quad , \quad \text{at } t = t_n(\alpha) \quad . \quad (5)$$

Note that $t_n(\alpha)$ is a function of α , so variations in the system parameters will induce variations in the phase-space location of the maximum power.

The maximum fuel temperature response for any channel J can be expressed as

$$R_T(e) = \int_{\Omega} (1/N_J) g_J \cdot T(x) \delta[x - x_T(\alpha)] d\Omega \quad . \quad (6)$$

In Eq. (6), g_J is the NC -component vector

$$g_J = (0, \dots, 0, 1, 0, \dots, 0) \quad , \quad (7)$$

whose only nonzero component is the J th component,

$$x_T(\alpha) = [r_T(\alpha), z_T(\alpha), t_T(\alpha)] \quad (8)$$

is the vector representing the location in phase space where the maximum fuel temperature in channel J occurs, and

$$\delta[x - x_T(\alpha)] = \delta[r - r_T(\alpha)] \delta[z - z_T(\alpha)] \delta[t - t_T(\alpha)] \quad (9)$$

is the customary three-dimensional delta functional. Note that $x_T(\alpha)$ is defined implicitly as the solution of the system of equations

$$(\partial T_J / \partial r)_{x_T(\alpha)} = 0 \quad , \quad (10a)$$

$$(\partial T_J / \partial z)_{x_T(\alpha)} = 0 \quad , \quad (10b)$$

$$(\partial T_J / \partial t)_{x_T(\alpha)} = 0 \quad , \quad (10c)$$

which express the conditions necessary for $T_J(x)$ to have a maximum at $x_T(\alpha)$. Due to the boundary condition $(\partial T_J / \partial r)_{r=0} = 0$ [see Appendix A, Eq. (A.9)], the fuel temperature will attain its maximum in the radial direction at the center of the fuel rod. Thus, the components $z_T(\alpha)$ and $t_T(\alpha)$ of $x_T(\alpha)$ are functions of (the system parameters) α , but

$$r_T = 0 \quad (11)$$

regardless of α .

III.B.6.b. Sensitivity Analysis: Theory

The sensitivity theory presented in Sections III.B.3 and III.B.4 is now applied to Eqs. (1), (2), (4), (5), (6), and (10), with the twofold objective:

1. to determine the sensitivities of R_n and R_T (i.e., of the numerical values of the maximum power and maximum fuel temperature responses, respectively) to changes in the system parameters α
2. to determine the sensitivities of the critical points t_n and x_T (i.e., of the phase-space locations where the respective maxima occur) to changes in the system parameters α .

Since all operators (including the responses) appearing in the mathematical formulation of the present problem can be generically represented by a nonlinear operator $S(e)$, the definition of the G-differential $VS(e^\circ; \mathbf{h})$ of $S(e)$ at e° becomes

$$VS(e^\circ; \mathbf{h}) \equiv \frac{d}{d\epsilon} \{[S(e^\circ + \epsilon \mathbf{h})]\}_{\epsilon=0} \quad , \quad (12)$$

where ϵ is a real scalar, and $\mathbf{h} = (\mathbf{h}_U, \mathbf{h}_\alpha)$ represents a fixed, but otherwise arbitrary, vector of "changes" around the base-case configuration $e^\circ = (U^\circ, \alpha^\circ)$. The vectors \mathbf{h}_U and \mathbf{h}_α have the same number of components as U and α , respectively; for example, $\mathbf{h}_U = (\mathbf{h}_T, \mathbf{h}_{T_c}, \mathbf{h}_{T_s}, \mathbf{h}_P, \mathbf{h}_u, \mathbf{h}_n, \mathbf{h}_{C_v}, \dots, \mathbf{h}_{C_{n_c}})$. Note that a G-differential $VS(e^\circ; \mathbf{h})$ that is linear in \mathbf{h} is customarily denoted by $DS(e^\circ; \mathbf{h})$. Necessary and sufficient conditions for VS to be linear in \mathbf{h} [i.e., for $VS(e^\circ; \mathbf{h}) = DS(e^\circ; \mathbf{h})$] are known, and their importance to sensitivity theory has been generally discussed in the previous sections of these lectures. As will soon become apparent (see also Appendix B), all operators acting on U satisfy these necessary and sufficient conditions, and therefore admit G-differentials that are linear in \mathbf{h} . Consequently, the notation $DS(e^\circ; \mathbf{h})$ is henceforth used to emphasize this important fact. Note also that the G-differential DS is related to the total variation $[S(e^\circ + \mathbf{h}) - S(e^\circ)]$ of $S(e^\circ)$ at e° through the relationship

$$\left. \begin{aligned} S(e^\circ + \mathbf{h}) - S(e^\circ) &= DS(e^\circ; \mathbf{h}) + \Delta(\mathbf{h}) \quad , \\ \lim_{\epsilon \rightarrow 0} [\Delta(\epsilon \mathbf{h})/\epsilon] &= 0 \end{aligned} \right\} \quad \text{where} \quad (13)$$

[Equation (13) actually holds in the most general case, i.e., with $VS(e^\circ; \mathbf{h})$ replacing $DS(e^\circ; \mathbf{h})$].

The G-differential of $R_n(e)$ at e° , which gives the sensitivity of R_n to changes \mathbf{h} , is obtained by applying the definition given in Eq. (12) to Eq. (4). This gives

$$\begin{aligned} DR_n(e^\circ; \mathbf{h}) &\equiv K \int_{\Omega} \left\{ \frac{d}{d\epsilon} \{ (n^\circ + \epsilon h_n) \delta[t - t_n(\alpha^\circ + \epsilon \mathbf{h}_\alpha)] \} \right\}_{\epsilon=0} d\Omega \\ &\equiv K \int_{\Omega} h_n \delta[t - t_n(\alpha^\circ)] d\Omega \\ &\quad - K \sum_{i=1}^I (\partial t_n / \partial \alpha_i)_{\alpha^\circ} h_{\alpha_i} \int_{\Omega} n^\circ \delta'[t - t_n(\alpha^\circ)] d\Omega \quad . \end{aligned} \quad (14)$$

Using the definition of the δ' functional, i.e.,

$$\int f(x) \delta'(x - x_0) dx = - \int (df/dx) \delta(x - x_0) dx \quad , \quad (15)$$

and recalling Eq. (5), the last term appearing on the right side of Eq. (14) can be shown to vanish, i.e.,

$$\int_{\Omega} n^{\circ}(t)\delta'(t-t_n)d\Omega = -[dn(t)/dt]_{t_n} = 0 \quad . \quad (16)$$

Thus, Eq. (14) simplifies to

$$DR_n = K \int_{\Omega} h_n(t)\delta[t-t_n(\alpha^{\circ})]d\Omega \quad . \quad (17)$$

The sensitivity $DR_T(e^{\circ};h)$ of $R_T(e)$ at e° is determined by using Eq. (6) and by following the same procedure as that leading to Eq. (17). The result is

$$DR_T = \int_{\Omega} (1/N_J)g_J h_T(x)\delta[x-x_T(\alpha^{\circ})]d\Omega \quad . \quad (18)$$

The sensitivities of the critical points $t_n(\alpha)$ and $x_T(\alpha)$ to changes h_{α} are given by the respective G-differentials of $t_n(\alpha)$ and $x_T(\alpha)$ at α° . In view of Eqs. (5), (8), and (10), each of the quantities t_n , z_T , and t_T is a real-valued function of the real variables $\alpha_1, \dots, \alpha_I$ and can therefore be regarded as a functional defined on a subset of \mathfrak{R}^I . Applying now the definition given in Eq. (12) to the functionals $t_n(\alpha)$, $z_T(\alpha)$, and $t_T(\alpha)$ yields

$$Dt_n(\alpha^{\circ};h_{\alpha}) = \sum_{i=1}^I h_{\alpha_i}(\partial t_n/\partial \alpha_i)_{\alpha^{\circ}} \quad , \quad (19)$$

$$Dz_T(\alpha^{\circ};h_{\alpha}) = \sum_{i=1}^I h_{\alpha_i}(\partial z_T/\partial \alpha_i)_{\alpha^{\circ}} \quad , \quad (20)$$

$$Dt_T(\alpha^{\circ};h_{\alpha}) = \sum_{i=1}^I h_{\alpha_i}(\partial t_T/\partial \alpha_i)_{\alpha^{\circ}} \quad . \quad (21)$$

In view of Eqs. (8), (11), (20), and (21), the sensitivity of the critical point $x_T(\alpha)$ to changes h_{α} around α° is given by the three-component column vector

$$Dx_T(\alpha^{\circ};h_{\alpha}) = (0, Dz_T, Dt_T) \quad . \quad (22)$$

The explicit expressions for Dt_n , Dz_T , and Dt_T are obtained by applying the general procedure outlined in Sections III.B.3 and III.B.4. Thus, to determine Dt_n , Eq. (5) is recast in the equivalent form

$$\int_{\Omega} (dn/dt)\delta[t-t_n(\alpha)]d\Omega = 0 \quad . \quad (23)$$

Taking the G-differential of Eq. (23) [by applying Eq. (12)] gives

$$\begin{aligned} & \int_{\Omega} (dh_n/dt)\delta[t-t_n(\alpha^{\circ})]d\Omega \\ & - Dt_n(\alpha^{\circ};h_{\alpha}) \int_{\Omega} (dn/dt)\delta[t-t_n(\alpha^{\circ})]d\Omega = 0 \quad . \end{aligned}$$

Using now Eq. (15) and solving the above equation for Dt_n gives

$$Dt_n = -(dh_n/dt)_{t_n(\alpha^\circ)} / (d^2n/dt^2)_{t_n(\alpha^\circ)} \quad (24)$$

The explicit expressions for Dz_T and Dt_T are obtained by following the same procedure as that leading to Eq. (24), i.e., by writing Eqs. (10b) and (10c) as

$$\int_{\Omega} [\partial(g_J T)/\partial z] \delta[x - x_T(\alpha)] d\Omega = 0$$

and

$$\int_{\Omega} [\partial(g_J T)/\partial t] \delta[x - x_T(\alpha)] d\Omega = 0$$

by taking G-differentials at α° of the above equations and by simultaneously solving the resulting equations for Dz_T and Dt_T . The result can be written in vector form as

$$\begin{pmatrix} Dz_T \\ Dt_T \end{pmatrix} = \mathbf{M}^{-1} \mathbf{F} \quad (25)$$

where

$$\mathbf{F} \equiv \begin{pmatrix} -(\partial h_T/\partial z)_{x_T(\alpha^\circ)} \\ -(\partial h_T/\partial t)_{x_T(\alpha^\circ)} \end{pmatrix} \quad (26)$$

and

$$\mathbf{M} \equiv \begin{pmatrix} (\partial^2 T_J/\partial z^2)_{x_T(\alpha^\circ)} & (\partial^2 T_J/\partial z \partial t)_{x_T(\alpha^\circ)} \\ (\partial^2 T_J/\partial t \partial z)_{x_T(\alpha^\circ)} & (\partial^2 T_J/\partial t^2)_{x_T(\alpha^\circ)} \end{pmatrix} \quad (27)$$

It is observed that for a given vector of changes \mathbf{h}_α , the sensitivities DR_n , DR_T , Dx_T , and Dt_n given, respectively, by Eqs. (17), (18), (22), and (24), can be evaluated only after determining the vector \mathbf{h}_U , since \mathbf{h}_α and \mathbf{h}_U are not independent. The (first-order) relationship between \mathbf{h}_α and \mathbf{h}_U is obtained by taking G-differentials of Eqs. (1) and (2) at e° . An examination of Eqs. (1) and (2) (see Appendix A) shows that each of the components of $\mathbf{N}(e)$, $\mathbf{Q}(\alpha)$, $\mathbf{B}(e)$, and $\mathbf{A}(\alpha)$ satisfies the necessary and sufficient conditions to admit *G-derivatives* at e° (this is illustrated in Appendix B). If a typical operator appearing in Eq. (1) or (2) is denoted by $\mathbf{S}(e)$, then the G-derivative of $\mathbf{S}(e)$ at e° is the operator $\mathbf{S}'_e(e^\circ)$ defined by the relationship

$$D\mathbf{S}(e^\circ; \mathbf{h}) = \mathbf{S}'_e(e^\circ) \mathbf{h} \quad (28)$$

where $D\mathbf{S}(e^\circ; \mathbf{h})$ is the linear G-differential. Furthermore, since $\mathbf{h} = (\mathbf{h}_U, \mathbf{h}_\alpha)$, the following relationship holds:

$$\mathbf{S}'_e(e^\circ) \mathbf{h} = \mathbf{S}'_U(e^\circ) \mathbf{h}_U + \mathbf{S}'_\alpha(e^\circ) \mathbf{h}_\alpha \quad (29)$$

where $\mathbf{S}'_U(e^\circ)$ is the partial G-derivative at e° of $\mathbf{S}(e)$ with respect to U , and $\mathbf{S}'_\alpha(e^\circ)$ is the partial G-derivative at e° of $\mathbf{S}(e)$ with respect to α . In view of Eqs. (28) and (29), the result of taking G-differentials at e° of Eqs. (1) and (2) is

$$\mathbf{N}'_U(\mathbf{e}^\circ)\mathbf{h}_U = [\mathbf{Q}'_\alpha(\alpha^\circ) - \mathbf{N}'_\alpha(\mathbf{e}^\circ)]\mathbf{h}_\alpha, \quad \mathbf{x} \in \Omega_j, \quad (30)$$

and

$$\mathbf{B}'_U(\mathbf{e}^\circ)\mathbf{h}_U = [\mathbf{A}'_\alpha(\alpha^\circ) - \mathbf{B}'_\alpha(\mathbf{e}^\circ)]\mathbf{h}_\alpha, \quad \mathbf{x} \in \partial\Omega_j. \quad (31)$$

The explicit representation of $\mathbf{N}'_U(\mathbf{e}^\circ)$ is the $M \times M$ matrix whose elements are the partial G-derivatives at \mathbf{e}° of the components of \mathbf{N} with respect to the components U_j of \mathbf{U} , i.e.,

$$\mathbf{N}'_U(\mathbf{e}^\circ) = [L_{ij}(\mathbf{e}^\circ)]; \quad L_{ij}(\mathbf{e}^\circ) = (N_i)'_{U_j}(\mathbf{e}^\circ); \quad i, j = 1, \dots, M. \quad (32)$$

The representation of $\mathbf{N}'_\alpha(\mathbf{e}^\circ)$ is the $M \times I$ matrix whose elements are the partial G-derivatives at \mathbf{e}° of the components of $\mathbf{N}(\mathbf{e})$ with respect to the components of α . The elements of the matrices representing $\mathbf{B}'_U(\mathbf{e}^\circ)$, $\mathbf{Q}'_\alpha(\alpha^\circ)$, and $\mathbf{A}'_\alpha(\alpha^\circ)$ are obtained in a similar manner. Note that $\mathbf{N}'_U(\mathbf{e}^\circ)\mathbf{h}_U$ and $\mathbf{B}'_U(\mathbf{e}^\circ)\mathbf{h}_U$ are linear in \mathbf{h}_U and are independent of \mathbf{h}_α ; on the other hand, $\mathbf{N}'_\alpha(\mathbf{e}^\circ)\mathbf{h}_\alpha$ and $\mathbf{B}'_\alpha(\mathbf{e}^\circ)\mathbf{h}_\alpha$ are linear in \mathbf{h}_α , and are independent of \mathbf{h}_U .

For a given vector of changes \mathbf{h}_α , Eqs. (30) and (31) could be solved to determine \mathbf{h}_U ; \mathbf{h}_U could then be used to evaluate the sensitivities DR_T , DR_n , Dx_T , Dt_n . However, due to the large number of system parameters, it would be prohibitively expensive to repeatedly solve Eqs. (30) and (31) for all vectors \mathbf{h}_α of possible interest to the sensitivity analysis of the problem at hand. An alternative procedure that avoids the need to repeatedly solve Eqs. (30) and (31) can be developed to evaluate the above-mentioned sensitivities by using adjoint operators.

Each of the functionals DR_n , DR_T , Dt_n , Dz_T , and Dt_T , [see Eqs. (17), (18), (24), and (25), respectively] is linear in \mathbf{h}_U . Considering now that $\mathbf{h}_U \in H$, where H is a Hilbert space equipped with the inner product

$$\langle V, W \rangle = \sum_{j=1}^M \int_{\Omega} v_j w_j d\Omega, \quad V \in H, \quad W \in H, \quad (33)$$

the Riesz representation theorem ensures that each of the functionals DR_n , DR_T , Dt_n , Dz_T , and Dt_T can be written as the inner product of \mathbf{h}_U with a uniquely defined vector in H . Thus, the functional DR_n given by Eq. (17) can be represented as

$$DR_n = \langle \mathbf{h}_U, \mathbf{S}_n^* \rangle, \quad (34)$$

where the M -component vector \mathbf{S}_n^* is defined as

$$\mathbf{S}_n^* \equiv K\delta[t - t_n(\alpha^\circ)] \quad (\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, 1, \mathbf{0}, \dots, \mathbf{0})^T. \quad (35)$$

Similarly, the functional DR_T given by Eq. (18) can be represented as

$$DR_T = \langle \mathbf{h}_U, \mathbf{S}_T^* \rangle, \quad (36)$$

where the M -component vector \mathbf{S}_T^* is defined as

$$\mathbf{S}_T^* \equiv (1/N_J)\delta[\mathbf{x} - \mathbf{x}_T(\alpha^\circ)] \quad (g_J, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \dots, \mathbf{0})^T. \quad (37)$$

The definition of the δ' functional given in Eq. (15) is used together with Eq. (33) to express Eq. (24) in the inner product form

$$Dt_n \equiv \langle \mathbf{h}_U, \mathbf{G}^* \rangle / (d^2n/dt^2)_{t_n(\alpha^\circ)} \quad , \quad (38)$$

where

$$\mathbf{G}^* \equiv K \delta'[t - t_n(\alpha^\circ)] \quad (\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, 1, 0, \dots, 0)^T \quad . \quad (39)$$

Similarly, Eq. (25) can be expressed as

$$\begin{pmatrix} Dz_T \\ Dt_T \end{pmatrix} = \mathbf{M}^{-1} \begin{pmatrix} \langle \mathbf{h}_U, \mathbf{F}_1^* \rangle \\ \langle \mathbf{h}_U, \mathbf{F}_2^* \rangle \end{pmatrix} \quad , \quad (40)$$

where

$$\mathbf{F}_1^* \equiv (1/N_J) \delta'[z - z_T(\alpha^\circ)] \delta[t - t_T(\alpha^\circ)] \delta(r) (g_J, \mathbf{0}, \dots, 0)^T \quad , \quad (41)$$

and

$$\mathbf{F}_2^* \equiv (1/N_J) \delta'[t - t_T(\alpha^\circ)] \delta[z - z_T(\alpha^\circ)] \delta(r) (g_J, \mathbf{0}, \dots, 0)^T \quad . \quad (42)$$

To proceed with the construction of the appropriate adjoint system, recall that both Eqs. (30) and (31) are linear in \mathbf{h}_U . Consequently, the following relationship holds for an arbitrary vector $V \in H$:

$$\langle V, \mathbf{N}'_U(e^\circ) \mathbf{h}_U \rangle = \langle \mathbf{h}_U, \mathbf{L}^*(e^\circ) V \rangle + [P(\mathbf{h}_U; V)]_{\partial\Omega} \quad . \quad (43)$$

In Eq. (43), $\mathbf{L}^*(e^\circ)$ is the operator formally adjoint to $\mathbf{N}'_U(e^\circ)$, and $[P(\mathbf{h}_U; V)]_{\partial\Omega}$ represents the associated bilinear form that consists of terms evaluated on the boundary $\partial\Omega$ of Ω . Note that the use of Eq. (33) in conjunction with Eq. (43) will require the introduction of appropriate normalization constants for those components of $\mathbf{N}'_U(e^\circ) \mathbf{h}_U$ that are functions of only some, rather than all, of the independent variables (r, z, t) .

The explicit form of $\mathbf{L}^*(e^\circ)$ is the $M \times M$ matrix

$$\mathbf{L}^*(e^\circ) = (l_{ij}) \quad ; \quad l_{ij} = L_{ji}^* \quad , \quad i, j = 1, \dots, M \quad , \quad (44)$$

obtained by transposing the formal adjoints of the operators $L_{ij}(e^\circ)$ given by Eq. (32). The explicit representation of each component of $\mathbf{L}^*(e^\circ) V$ is given in Appendix C. The domain of $\mathbf{L}^*(e^\circ)$ is determined by selecting appropriate adjoint boundary conditions, represented here in operator form as

$$\mathbf{B}^*(V; e^\circ) = \mathbf{A}^*(e^\circ) \quad , \quad x \in \partial\Omega_j \quad . \quad (45)$$

These boundary conditions are obtained by requiring that

1. they be independent of \mathbf{h}_U , \mathbf{h}_α , and G-derivatives with respect to α
2. substitution of Eqs. (31) and (45) into the expression of $[P(\mathbf{h}_U; V)]_{\partial\Omega}$ must cause all terms containing unknown values of \mathbf{h}_U to vanish.

The selection of the adjoint boundary conditions reduces $[P(\mathbf{h}_U; V)]_{\partial\Omega}$ to a quantity designated here by $\hat{P}(\mathbf{h}_\alpha; V; e^\circ)$, where \hat{P} contains boundary terms involving only known values of \mathbf{h}_α , V , and e° . [The explicit expressions for the adjoint boundary conditions represented by Eq. (45) can be found in Appendix C.] Hence, Eq. (43) can be written as

$$\langle \mathbf{h}_U, \mathbf{L}^*(e^\circ)V \rangle = \langle V, [Q'_\alpha(\alpha^\circ) - N'_\alpha(e^\circ)]\mathbf{h}_\alpha \rangle - \hat{P}(\mathbf{h}_\alpha; V; e^\circ) \quad , \quad (46)$$

where Eq. (30) is used to replace the quantity $N'_U(e^\circ)\mathbf{h}_U$.

Equations (45) and (46) hold for all (i.e., arbitrary) vectors $V \in H$. Five such vectors [i.e., each vector being an element of H and satisfying Eqs. (45) and (46)] will now be selected in a unique manner to successively eliminate the vector \mathbf{h}_U from Eqs. (34), (36), (38), and (40), so that alternative expressions for the sensitivities DR_n , DR_T , Dt_n , and Dx_T can be derived.

The alternative expression for DR_n is obtained by using Eq. (34) and by considering Eqs. (45) and (46) as written specifically for the vector $V_n^* \in H$, i.e.,

$$\langle \mathbf{h}_U, \mathbf{L}^*(e^\circ)V_n^* \rangle = \langle V_n^*, [Q'_\alpha(\alpha^\circ) - N'_\alpha(e^\circ)]\mathbf{h}_\alpha \rangle - \hat{P}(\mathbf{h}_\alpha; V_n^*; e^\circ) \quad , \quad (47)$$

and

$$\mathbf{B}^*(V_n^*; e^\circ) = \mathbf{A}^*(e^\circ) \quad , \quad x \in \partial\Omega_j \quad . \quad (48)$$

Comparing the left side of Eq. (47) with the right side of Eq. (34) shows that

$$\mathbf{L}^*(e^\circ)V_n^* = \mathbf{S}_n^* \quad , \quad x \in \Omega_j \quad . \quad (49)$$

Equations (34), (47), and (49) can now be used to express DR_n as

$$DR_n = \langle V_n^*, [Q'_\alpha(\alpha^\circ) - N'_\alpha(e^\circ)]\mathbf{h}_\alpha \rangle - \hat{P}(\mathbf{h}_\alpha; V_n^*; e^\circ) \quad . \quad (50)$$

With the derivation of Eq. (50), the unknown values \mathbf{h}_U that appeared in the original expression of DR_n [see Eq. (34)] have been eliminated. Now, once the *single* calculation to determine the adjoint vector V_n^* [by solving the adjoint system given by Eqs. (48) and (49)] is completed, Eq. (50) provides the most efficient means to obtain the sensitivity DR_n of $R_n(e)$ to changes \mathbf{h}_α around α° .

To derive an alternative expression for DR_T , the same procedure as outlined in the foregoing paragraph is applied to Eqs. (36), (45), and (46). This gives

$$DR_T = \langle V_T^*, [Q'_\alpha(\alpha^\circ) - N'_\alpha(e^\circ)]\mathbf{h}_\alpha \rangle - \hat{P}(\mathbf{h}_\alpha; V_T^*; e^\circ) \quad , \quad (51)$$

where the adjoint function V_T^* satisfies the adjoint system

$$\mathbf{L}^*(e^\circ)V_T^* = \mathbf{S}_T^* \quad , \quad x \in \Omega_j \quad , \quad (52)$$

subject to

$$B^*(V_T^*;e^\circ) = A^*(e^\circ) \quad , \quad x \in \partial\Omega_j \quad , \quad (53)$$

Repeating the above procedure, an alternative expression is obtained for Dt_n by using Eq. (38), and by considering Eqs. (45) and (46) as written specifically for the vector $Y^* \in H$. The ensuing result is

$$Dt_n = \{ \langle Y^*, [Q'_\alpha(\alpha^\circ) - N'_\alpha(e^\circ)] h_\alpha \rangle - \hat{P}(h_\alpha, Y^*; e^\circ) \} / (d^2n/dt^2)_{t,(\alpha^\circ)} \quad , \quad (54)$$

where Y^* satisfies the adjoint system

$$L^*(e^\circ)Y^* = G^* \quad , \quad x \in \Omega_j \quad , \quad (55)$$

subject to

$$B^*(Y^*;e^\circ) = A^*(e^\circ) \quad , \quad x \in \partial\Omega_j \quad . \quad (56)$$

The same procedure is repeated once again to derive an alternative expression for the left side of Eq. (40). The final result is

$$\begin{pmatrix} Dz_T \\ Dt_T \end{pmatrix} = M^{-1} \begin{pmatrix} \langle W_1^*, [Q'_\alpha(\alpha^\circ) - N'_\alpha(e^\circ)] h_\alpha \rangle \\ -\hat{P}(h_\alpha, W_1^*; e^\circ) \\ \langle W_2^*, [Q'_\alpha(\alpha^\circ) - N'_\alpha(e^\circ)] h_\alpha \rangle \\ -\hat{P}(h_\alpha, W_2^*; e^\circ) \end{pmatrix} \quad , \quad (57)$$

where $W_1^* \in H$ satisfies the adjoint system

$$L^*(e^\circ)W_1^* = F_1^* \quad , \quad x \in \Omega_j \quad , \quad (58)$$

subject to

$$B^*(W_1^*;e^\circ) = A^*(e^\circ) \quad , \quad x \in \partial\Omega_j \quad , \quad (59)$$

and where $W_2^* \in H$ satisfies the adjoint system

$$L^*(e^\circ)W_2^* = F_2^* \quad , \quad x \in \Omega_j \quad , \quad (60)$$

subject to

$$B^*(W_2^*;e^\circ) = A^*(e^\circ) \quad , \quad x \in \partial\Omega_j \quad , \quad (61)$$

Note from Eq. (57) that each sensitivity Dz_T and Dt_T depends on both W_1^* and W_2^* . Thus, two adjoint calculations are needed to evaluate the sensitivity Dx_T [see Eq. (22)] of the critical point x (where the fuel temperature attains a maximum) for *all* changes h_α around α° .

Note that the same operator, namely $L^*(e^\circ)$, appears on the left sides of Eqs. (49), (52), (55), (58), and (60); only the source terms appearing on the right sides of these equations differ from one another. Furthermore, as evidenced by a comparison of Eqs. (48), (53), (56), (59), and (61), the adjoint functions V_n^* , V_T^* , Y^* , W_1^* , and W_2^* satisfy formally identical boundary conditions. Therefore, apart from the relatively trivial modifications required to accommodate the distinct source terms S_n^* , S_T^* , G^* , F_1^* , and F_2^* , the same computer code can be used to solve all the respective adjoint systems to determine the functions V_n^* , V_T^* , Y^* , W_1^* , and W_2^* . An examination of the right sides of Eqs. (50), (51), (54), and (57) reveals that the function $[Q'_\alpha(\alpha^\circ) - N'_\alpha(e^\circ)]h_\alpha$ appears in all of the respective inner products denoted by \langle, \rangle ; furthermore, the quantities \hat{P} appearing in these equations are formally identical. Therefore, the computer code used to evaluate the sensitivity DR_n can also be used to evaluate the sensitivities DR_T , Dt_n , and Dx_T . Of course, the values of V_n^* required to compute DR_n are to be replaced by the respective values of V_T^* , Y^* , W_1^* , and W_2^* .

In reactor physics, the adjoint function has traditionally been interpreted as an "importance function." For the problem at hand, the components of the adjoint functions V_n^* , V_T^* , Y^* , W_1^* , and W_2^* can also be interpreted as importance functions. For this purpose, it is noted that each of these adjoint functions can be represented generically as the M -component vector

$$V = (H^*, H_c^*, H_s^*, p^*, m^*, n^*, C_1^*, \dots, C_{NG}^*) \quad (62)$$

Thus, when evaluating the response sensitivities DR_n and DR_T , V represents V_n^* and V_T^* , respectively. In this case, a dimensional analysis of Eqs. (50) and (51) shows that the dimensions $[V_j]$ of each component $V_j (j=1, \dots, M)$ of V are

$$[V_j] = \frac{\text{[response]}}{\left[\begin{array}{c} \text{terms in the} \\ j\text{th forward} \\ \text{equation} \end{array} \right] \left[\begin{array}{c} \text{region of} \\ \text{integration} \end{array} \right] \left[\begin{array}{c} \text{normalization} \\ \text{constant} \end{array} \right]} \quad (63a)$$

Similarly, V represents Y^* , W_1^* , and W_2^* when evaluating the critical point sensitivities Dt_n , Dz_T , and Dt_T , respectively. In this case, a dimensional analysis of Eqs. (54) and (57) shows that

$$[V_j] = \frac{\text{[response]}/\text{[component of critical point]}}{\left[\begin{array}{c} \text{terms in the} \\ j\text{th forward} \\ \text{equation} \end{array} \right] \left[\begin{array}{c} \text{region of} \\ \text{integration} \end{array} \right] \left[\begin{array}{c} \text{normalization} \\ \text{constant} \end{array} \right]} \quad (63b)$$

The considerations leading to Eqs. (63a) and (63b) hold generally for any maximum-type response. According to Eq. (63a), each component of the adjoint function used to compute response sensitivities can be viewed as a measure of the importance of the physical quantity described by the corresponding forward equation in contributing to the response. Furthermore, according to Eq. (63b), each component of the adjoint function used to compute critical point sensitivities can be viewed as a measure of the importance of the physical quantity described by the corresponding forward equation in contributing to the response movement in phase space.

Consider, for example, the coolant temperature equation for channel j [see Eq. (A.4)] and the corresponding component H_{cj}^* of V . The dimensions of the terms in this equation are $[\text{J}\cdot\text{cm}^{-1}\cdot\text{s}^{-1}]$. The respective region of integration is volume and time, with dimensions $[\text{cm}^3\cdot\text{s}]$. The dimension of the appropriate normalization constant, i.e., $1/\pi(R_f^2 - R^2 + R_g^2)$, is $[\text{cm}^{-2}]$. Thus, Eq. (63a) gives

$$[H_{cj}^*] = \frac{\frac{[\text{response}]}{[\text{J}\cdot\text{cm}^{-1}\cdot\text{s}^{-1}]}}{[\text{cm}^3\cdot\text{s}]} \frac{[\text{response}]}{[\text{cm}^2]} = \frac{[\text{response}]}{[\text{J}]}$$

For the responses R_n and R_T , the dimensions of H_{cj}^* are $[\text{MW}/\text{J}]$ and $[\text{K}/\text{J}]$, respectively. For the critical point t_n , Eq. (63b) indicates that H_{cj}^* has units of $[\text{MW}\cdot\text{s}^{-1}/\text{J}]$, while the units of H_{cj}^* corresponding to the components z_T and t_T of the critical point x_T are $[\text{K}\cdot\text{cm}^{-1}/\text{J}]$ and $[\text{K}\cdot\text{s}^{-1}/\text{J}]$, respectively. This dimensional analysis shows that H_{cj}^* is a measure of enthalpy importance in the coolant region of the j 'th channel. Similar analyses indicate that, for each channel, the components of H^* , H_s^* , m^* , and p^* are measures of enthalpy importance in the fuel pin region, enthalpy importance in the structure region, coolant mass importance, and momentum flux importance, respectively. Furthermore, n^* and C_1^*, \dots, C_{NG}^* are measures of power importance and precursor amplitude importances, respectively. Therefore, the adjoint variables H^* , m^* , etc. will henceforth be referred to as adjoint enthalpy, adjoint mass, etc.

The distinctions between the concepts underlying the derivations presented in this section and those underlying the derivations presented in work based on "perturbation theory" stem from distinctions between the concept of G-differentiation of an operator in a linear vector space and the concept of partial differentiation, in the elementary calculus sense, of a real-valued function of I variables. The concept of G-differentiation significantly generalizes the concept of differentiation customarily used in finite-dimensional calculus. For example, in the elementary I -dimensional calculus, the total differential of a real-valued function $f(x)$ defined on an open subset $Y \subset \mathbb{R}^I$ is expressed as

$$df = \sum_{i=1}^I (\partial f / \partial x_i) dx_i \quad .$$

On the other hand, the same function f is viewed in nonlinear functional analysis as the functional $f: Y \subset \mathbb{R}^I \rightarrow \mathbb{R}$. Elementary considerations show that if the G-differential $Vf(x; h)$ exists for all x in an open neighborhood of a point $x^\circ \in \text{int}(Y)$ and if, for all fixed $h \in \mathbb{R}^I$, $Vf(x; h)$ is continuous in x at x° , then

$$Vf(x; h) = \sum_{j=1}^I (\partial f / \partial x_j) h_j = \sum_{j=1}^I (\partial f / \partial x_j) dx_j = df \quad .$$

(Obviously, the components of h can be taken to be the differentials dx_i , of arbitrary magnitudes, of the independent variables x_i .)

In works that use perturbation theory, the response and the dependent variables are considered, at the outset, to be real-valued functions that depend (explicitly and implicitly) on a real *scalar* parameter α ; α stands, in turn, for each scalar system parameter (i.e., α represents any one of the components of the vector α used in this work). The sensitivity of the response R to a variation $\delta\alpha$ in any parameter α is then simply the customary derivative $dR/d\alpha$; the expression of $dR/d\alpha$ is determined by using partial differentiation (in the elementary calculus sense) of the response and of the equations describing the problem. Of course, this approach must *a priori* assume that the

above-mentioned differentiability and continuity conditions are satisfied by all the mathematical expressions describing the problem, including the response R . Consider now that all parameters α_i , $i=1,\dots,I$, are simultaneously varied by amounts $\delta\alpha_i$ around the base-case values α_i° . Then the total variation in the response, i.e., the total sensitivity, would be given (to first order) by the sum

$$\sum_{i=1}^I (dR/d\alpha_i)\delta\alpha_i \quad ,$$

where $dR/d\alpha_i$ includes both the "direct effects" (i.e., contributions of the type $\partial R/\partial\alpha_i$) and the "indirect effects" [i.e., contributions of the type $(\partial R/\partial V_j)(\partial V_j/\partial\alpha_i)$, where V_j represents a dependent variable]. But

$$\sum_{i=1}^I (dR/d\alpha_i)\delta\alpha_i$$

is precisely the quantity that gives the linear G-differential $DR(e^\circ; \mathbf{h})$ as used in this work (the variations $\delta\alpha_i$ are interpreted as the components of \mathbf{h}).

Since the G-derivative can be defined under conditions that are much weaker than those required for defining derivatives in the elementary calculus sense, the derivations presented in this work are considerably more general than those presented in works based on perturbation theory. Although the advantages offered by this generality have not been fully exploited in the course of the application presented in this work, the simplicity of using G-derivatives in a practical sensitivity analysis has nonetheless been highlighted. In Section III.A, we have discussed in detail the important practical advantages of using functional analysis concepts, such as the G-derivative, in sensitivity analysis. In particular, the use of G-derivatives opens the possibility of treating problems involving discontinuities and parameters that are functions (depending, for example, on space and/or time variables) rather than scalars.

III.B.6.c. Sensitivity Analysis: Results

For a numerical illustration of the theory presented in Sec. III.B.6.b, a subprompt-critical excursion in the Fast Flux Test Facility (FFTF) was selected for sensitivity analysis. This analysis uses the MELT-MELTADJ code system, which solves the forward (i.e., the original, nonlinear) and the appropriate adjoint systems of neutronic/thermal-hydraulic equations (given in Appendices A and C, respectively). The geometry of the FFTF is modeled with a two-channel representation of the reactor flow path. Channel two (for which $J = 2$) is designated as the hot channel and consists of 227 pins. Channel one represents the remainder of the FFTF core and consists of 15 624 pins. Only one flow loop is considered. The dimensions of the outer radii for the fuel, gap, and cladding are $R_f = 0.249$ cm, $R_g = 0.254$ cm, and $R = 0.292$ cm. The channel height is $L = 800$ cm; the bottom of the core is located axially at $z = 105.16$ cm, and the core length is 91.44 cm.

The subprompt-critical excursion is a protected transient involving a 0.23 dollar/s ramp reactivity insertion with scram that trips the control rods and primary pumps on high-power level. The power profile for this transient is given in Fig. 2. Although the high-power level for trip is attained at $t_0 = 0.518$ s, a time delay of $\Delta t = 0.19$ s postpones the actual control rod insertion and pump shutdown until 0.708 s after initiation of the transient conditions. Just after this time, the power $n(t)$ attains its maximum value of 467.7 MW. Note that the scram component $\bar{\rho}_{scram}(t)$ of the system's total reactivity $\bar{\rho}(t)$ is just a reactivity ramp ρ_0 that is switched on at $t_0 + \Delta t$, i.e.,

$$\bar{\rho}_{scram}(t) = \bar{\rho}_0(t - t_0 - \Delta t)1_+(t - t_0 - \Delta t), \quad (64)$$

where

$$1_+(t) = \begin{cases} 0, & t < 0 \\ 1, & t \geq 0 \end{cases}$$

is the customary unit-step functional. The large magnitude of ρ_0 and the discontinuous time derivatives of $\bar{\rho}_{scram}$ are the main cause of the highly nonlinear behavior of the power $n(t)$ and of the asymmetric shape of its maximum (see Fig. 2).

An examination of the temperature distribution $T(t, z, t)$ for this transient shows that the fuel temperature in the hot channel attains a maximum value of 2734.1 K. Spatially, this maximum is located at $r_T = 0$ and $z_T = 155$ cm, i.e., at the center of the fuel rod and just above the core midplane. The time variation of the temperature at this spatial location is plotted in Fig. 3, which shows that the maximum occurs in time at $t_T = 0.870$ s. Note that this maximum fuel temperature occurs later in time than the maximum power (see Figs. 2 and 3) due to the time delay in the power-to-thermal energy integration.

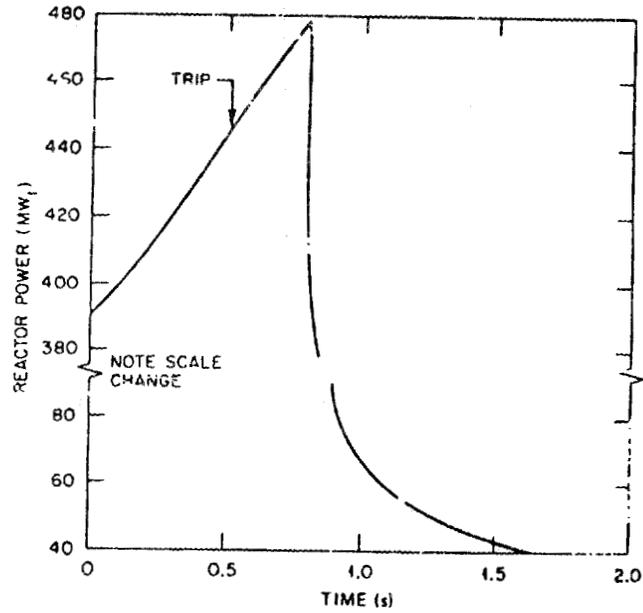


Fig. 2. Power profile for the kinetics/thermal-hydraulic transient with reactor scram on high-power level.

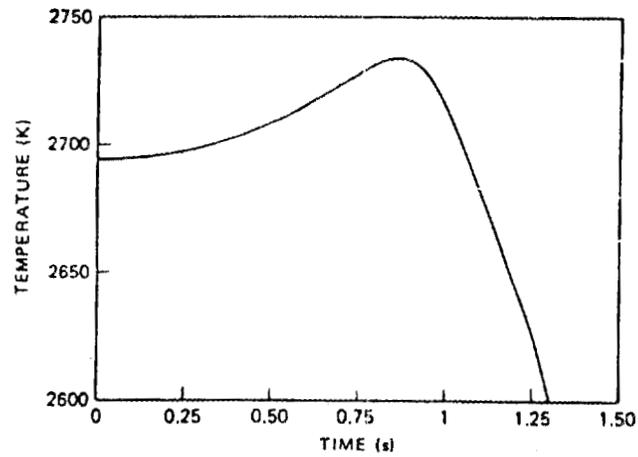


Fig. 3. Time variation of the temperature at the spatial location $r_T=0, z_T(\alpha^0) = 155$ cm.

The sensitivities of the numerical value of the *maximum fuel temperature response* $R_T(e)$ are calculated using Eq. (51), where the adjoint function V_T^* is the solution of the adjoint system represented by Eqs. (52) and (53). Numerical values for the quantities appearing in the source term S_T^* defined by Eq. (37) are $x_T(\alpha^0) = (0, 155, 0.870)$, $N_J = 227$, and $g_J = (0, 1)$.

Table I presents sensitivity results for those parameters that have the largest impact on the numerical value of the maximum fuel temperature response, R_T . Note that, for every value of the index i , these results correspond to a vector of changes h_α whose components are all zero except for h_{α_i} . The parameters are ranked in order of decreasing absolute magnitude of relative sensitivity. Based on these relative sensitivities, it can be concluded that the numerical value of the maximum fuel temperature is practically insensitive to variations in any of the system parameters except for variations in T , n_0 , and k_{fuel} .

The results presented in Table I also serve to illustrate the use of sensitivities for predicting the effects of parameter variations on the response. The basis for using sensitivities to predict the effects of parameter variations h_α on these responses is the general relationship expressed by Eq. (13), which takes on the particular form

$$R_T(e^0 + h) - R_T(e^0) = DR_T + O(\|h_\alpha\|^2). \quad (65)$$

TABLE I
Sensitivities for the Maximum Fuel Temperature Response $R_T(e)^*$

i	Parameter α_i	Relative Sensitivity ^a	Fractional Parameter Variation, h_{α_i}/α_i^0 (%)	Predicted Change ^b in Response Value (K)	Recalculated Change in Response Value (K) $R_T(e^0 + h) - R_T(e^0)$
1	T	0.746	0.5	10.20	10.0
2	n_0	0.155	0.3	1.27	1.3
3	k_{fuel}	-0.128	5.0	-17.5	-17.0
4	ρC_p	-0.015	1.0	-0.41	-0.4
5	T_{in}	0.013	1.0	0.35	0.2
6	h_g	-0.012	5.0	-1.58	-1.5
7	Δt	0.008	10.0	2.23	2.3
8	β	0.007	1.0	0.20	0.2
9	\tilde{n}	-0.007	0.4	-0.08	-0.1
10	$\bar{\rho}_{prog}$	-0.006	5.0	-0.82	-0.4
11	C_{pc}	-0.005	10.0	-1.29	-0.8

*The base-case value is $R_T(e^0) = 2734.1$ K, occurring at $x_T(\alpha^0) = (0, 155, 0.870)$.

^aRelative sensitivity $\equiv [DR_T(e^0; h)/R_T(e^0)](\alpha_i^0/h_{\alpha_i})$

^bPredicted change (in maximum value) $\equiv DR_T(e^0; h)$.

The $O(\|\mathbf{h}_\alpha\|^2)$ terms in Eq. (65) result from the facts that (a) DR_T is linear in \mathbf{h} , and (b) the vectors \mathbf{h}_U and \mathbf{h}_α are linearly related via Eqs. (30) and (31). As Eq. (65) indicates, the sensitivity DR_T predicts changes (i.e., deviations from the base-case value) that occur in the numerical value of the response R_T when the base-case parameter values α^0 are varied by \mathbf{h}_α . These predictions, though, do not take into account effects of second- and higher order terms in \mathbf{h}_α . For each specific fractional variation h_α/α_i^0 , the fifth column of Table I correspondingly lists the (sensitivity-based) predicted changes in the numerical value of the response R_T .

On the other hand, the results presented under the heading " $R_T(e^0 + \mathbf{h}) - R_T(e^0)$ " are the actual differences, obtained by direct recalculations, between the base-case numerical value of the response, i.e., $R_T(e^0)$, and the numerical value of the new maximum, i.e., $R_T(e^0 + \mathbf{h})$, that is attained at $x_T(\alpha^0 + \mathbf{h}_\alpha)$. Thus, for each specific h_α/α_i^0 , these results represent the corresponding numerical value taken on by the left side of Eq. (65). Note that for each α_i , the results presented in the fifth and sixth columns of Table I are in close agreement. This close agreement indicates that the nonlinear terms in \mathbf{h}_α , [see Eq. (65)] have relatively little practical impact on the numerical value of the maximum fuel temperature response, and highlights the usefulness of sensitivities for predicting the actual numerical value of the "perturbed" response $R_T(e^0 + \mathbf{h})$.

Sensitivity analysis results for *the critical point* $x_T(\alpha)$, at which the fuel temperature attains its maximum in phase space, are discussed next. As shown in Eq. (22), the sensitivity $Dx_T(\alpha^0; \mathbf{h}_\alpha)$ has two nonzero components, namely Dz_T and Dt_T . These components are evaluated using Eq. (57). The adjoint functions W_1^* and W_2^* , which must be determined prior to using Eq. (57), are obtained by solving the adjoint systems represented by Eq. (58) and (59) and by Eqs. (60) and (61), respectively. Just as in the case of Eq. (37), the source terms F_1^* and F_2^* [see Eqs. (41) and (42)] are calculated by using $N_J = 227$, $z_T(\alpha^0) = 155$ cm, $t_T(\alpha^0) = 0.870$ s, and $g_J = (0,1)$.

When the numerical calculations based on Eq. (57) were performed, it was found that all of the sensitivities Dz_T and most of the sensitivities Dt_T are negligibly small. Several of the sensitivities Dt_T , though, were found to be quite large. This implies that variations in the system parameters will affect almost exclusively the time component $t_T(\alpha)$ of $x_T(\alpha)$; they will have negligible effects on the axial component $z_T(\alpha)$, and, as discussed in Sec. III.B.6.b, they have no effects on the radial component r_T .

Table II presents sensitivity results for those parameters that have the largest impact on $t_T(\alpha)$. Just as in Table I, these results correspond to a vector of changes \mathbf{h}_α whose components are zero except for the *ith* component h_α . The parameters are ranked in order of decreasing absolute magnitude of relative sensitivity, a process equivalent to ranking the importance of their effects on $t_T(\alpha)$.

TABLE II

Sensitivities for the Time Component $t_T(\alpha)$ of $x_T(\alpha)$ *

i	Parameter α_i	Relative Sensitivity ^a	Fractional Parameter Variation, h_{α_i}/α_i (%)	Predicted Change ^b (s)	Recalculated Change ^c (s)
1	\tilde{T}	4.828	0.5	0.021	0.02
2	n_0	4.598	0.3	0.012	0.01
3	\tilde{n}	-3.448	0.4	-0.012	0.01
4	T_{in}	1.149	1.0	0.01	0.01
5	β	0.575	1.0	0.005	0.01
6	\bar{p}_{prog}	-0.575	5.0	-0.025	-0.03
7	k_{fuel}	-0.552	5.0	-0.024	-0.02
8	h_g	-0.322	5.0	-0.014	0.01
9	C_{pc}	-0.276	10.0	-0.024	-0.02
10	Δt	0.230	10.0	0.02	0.02
11	ρC_p	-0.115	1.0	-0.001	0.00

*The base-case value is $t_T(\alpha^0) = 0.870$ s.^aRelative sensitivity $\equiv [Dt_T(\alpha^0; h_\alpha)/t_T(\alpha^0)](\alpha_i^0/h_{\alpha_i})$.^bPredicted change $\equiv Dt_T(\alpha^0; h_\alpha)$.^cRecalculated change $\equiv t_T(\alpha^0 + h_\alpha) - t_T(\alpha^0)$.

The relative sensitivity results presented in Tables I and II indicate that, in both tables, the largest relative sensitivities involve the parameters \tilde{T} and n_0 . Thus, if varied, \tilde{T} and n_0 would have the largest impact on the numerical value of the maximum fuel temperature response, and would also cause the largest time shifts in the phase-space location of the resulting (i.e., new) maximum. Since all the relative sensitivities (in Tables I and II) to \tilde{T} and n_0 are positive, it follows that when a positive variation in \tilde{T} and/or n_0 is affected, the resulting maximum fuel temperature is both larger and occurs later in time than the original (i.e., the base-case) maximum fuel temperature.

Comparing the second and third columns in Table I to the respective columns in Table II, it becomes apparent that, except for \tilde{T} and n_0 , the parameter ranking in Table I differs from the ranking in Table II, although the same ranking procedure was used for both tables. The implications of this fact can be illustrated by considering the system parameters k_{fuel} and T_{in} . In Table I, k_{fuel} is ranked ahead of T_{in} , but this ranking is reversed in Table II. Consequently, a fractional variation in k_{fuel} causes a larger change in the numerical value of the maximum fuel temperature, but causes a smaller time shift of the maximum than does the same fractional variation in T_{in} .

Comparison of the relative sensitivities in Tables I and II also shows that, in general, the parameters affect the time location of the maximum fuel temperature significantly more than they affect the numerical value of this maximum. This conclusion is clearly illustrated by examining the two sensitivities to the initial value \tilde{n} of the neutronic power amplitude. It becomes readily apparent that a variation in \tilde{n} is of practically no importance to the numerical value of the maximum fuel temperature, but is of significant importance to $t_T(\alpha)$.

The sensitivities presented in Table II were used to predict the time shift [i.e., the difference between the time location $t(\alpha^0 + h_\alpha)$ at which the perturbed maximum $R_T(e^0 + h)$ occurs, and the time location $t(\alpha^0) = 0.870$ s at which $R_T(e^0)$ occurs] that is caused by each of the fractional parameter variations shown in the fourth column. The results for these predicted time shifts are presented in the fifth column. These predicted changes are in good agreement with the actual changes presented in the last column of Table II. These actual changes were obtained by direct recalculation of the fuel temperature, using the respective fractional parameter variations. It is informative to mention that, in all "forward" calculations, results are only printed at 0.01-s time intervals, although the actual time step used in such calculations is not fixed to 0.01 s, but varies as computed internally by the MELT-IIIB code.

III.B.6.d. Summary and Discussion of Results

The sensitivity theory for nonlinear systems with responses that are nonlinear functionals defined at critical points (e.g., maxima, minima, and saddle-points) has been developed in Sec. III.B by using concepts of nonlinear functional analysis. For the purposes of sensitivity analysis, the complete characterization of such responses requires consideration of both the numerical value of the response at the critical point and the phase-space location of the critical point.

This sensitivity theory has been successfully applied in Sec. III.B.6 to a problem of interest in reactor safety, namely a protected transient with scram on high-power level in the FFTF. To determine the base-case distribution of the dependent variables for this problem, the entire single-phase segment of the fast reactor safety code MELT-IIIB has been used, including

1. thermal-hydraulic equations, for each channel type, describing the behavior of the average channel fuel pin and surrounding coolant
2. an equation describing the primary loop hydraulics
3. neutron point-kinetics equations describing the reactor power level.

Two extremum-type responses have been considered — the maximum power response and the maximum fuel temperature response in the hot channel. Expressions for the sensitivities of the respective numerical values of these responses and expressions for the sensitivities of the phase-space locations at which the respective maxima occurred have been obtained in terms of adjoint functions. The adjoint systems satisfied, in turn, by each of these adjoint functions have been derived and solved. It has been shown that the complete sensitivity analysis of each (extremum-type) response requires the computation of as many adjoint functions as there are nonzero components of the respective critical point (e.g., maximum) in phase space, and of one additional adjoint function to evaluate the numerical value of the response. Once these adjoint functions have been computed, the sensitivities to *all* possible changes in the system parameters can be obtained by simple quadratures. For the problem at hand, a total of five adjoint calculations sufficed to perform the complete sensitivity analysis of the maximum fuel temperature and maximum power responses.

Note, though, that only the source terms in these five adjoint systems differ from one another; the form of the respective partial differential equations and corresponding boundary and final-time conditions is the same for all of these adjoint systems. Consequently, all five adjoint functions can be calculated by using the same code; only minor programming is required to implement the numerical calculation of the distinct source terms for the adjoint equations.

The results obtained for the sensitivities of the phase-space location (comprising time, axial, and radial components) of the maximum fuel temperature response show that variations in the system parameters affect mainly the time component of this maximum; such variations do not affect the radial component and produce negligible effects in the axial component. These sensitivities have been used to predict the phase-space location at which the new maximum fuel temperature occurs when the system parameters are varied. As has been shown, these predictions agreed well with direct recalculations using the perturbed parameter values. Similarly, the sensitivities obtained for the numerical value of the maximum fuel temperature response have been used to predict the numerical value of the perturbed maximum; these predictions also agreed well with direct recalculations.

The application presented has also highlighted the simplicity of using G-differentials and G-derivatives for sensitivity analysis of practical problems. The significant additional generality that stems from using such concepts in sensitivity analysis opens the possibility of treating problems which involve discontinuities and parameters that are functions rather than scalars.

III.B.6.e. Appendix A: Mathematical Representation of the MELT-IIIB Model

The thermal energy conservation equations for the average channel fuel pin, surrounding coolant, and structure are

$$\rho C_p \frac{\partial T}{\partial t} - \frac{1}{r} \frac{\partial}{\partial r} \left(r k \frac{\partial T}{\partial r} \right) - \psi n = \psi n_\gamma, \quad (A.1)$$

$$r \in (0, R_f) \cup (R_g, R); \quad z \in (0, L); \quad t \in (0, t_f),$$

$$A_c \rho_c C_{pc} \frac{DT_c}{Dt} - \frac{\partial}{\partial z} \left(A_c k_c \frac{\partial T_c}{\partial z} \right) - \psi n - 2\pi h \{ R [T(R, z, t) - T_c(z, t)] + R_s (T_s - T_c) \} = \psi n_\gamma, \quad z \in (0, L), \quad t \in (0, t_f), \quad (A.2)$$

$$A_s \rho_s C_{ps} \frac{\partial T_s}{\partial t} - 2\pi h R_s (T_c - T_s) = 0, \quad z \in (0, L), \quad t \in (0, t_f). \quad (A.3)$$

Mass and momentum conservation equations for each coolant channel are

$$A_c \frac{\partial \rho_c}{\partial t} + \frac{\partial (A_c \rho_c u)}{\partial z} = 0, \quad (A.4)$$

and

$$-\frac{\partial P}{\partial z} - \rho_c \frac{Du}{Dt} - \frac{f \rho_c u^2}{2D} - \rho_c g - \sum_m \delta(z - z_m) \frac{1}{2} \rho_c u^2 K_L = 0, \quad (A.5)$$

where $z \in (0, L)$ and $t \in (0, t_f)$.

Point kinetics equations are

$$\frac{dn}{dt} - \frac{\bar{\rho} - \beta}{\Lambda} n - \sum_{i=1}^{NG} \lambda_i C_i = 0, \quad (\text{A.6})$$

and

$$\frac{dC_i}{dt} + \lambda_i C_i - \frac{\beta_i n}{\Lambda} = 0, \quad (\text{A.7})$$

for $t \in (0, t_f)$, and $i = 1, \dots, NG$.

Initial and boundary conditions for Eqs. (A.1) through (A.7) are

$$T(t=0) = \bar{T}, \quad (\text{A.8})$$

$$\left[\frac{\partial T}{\partial r} \right]_{r=0} = 0, \quad (\text{A.9})$$

$$\left[rk \frac{\partial T}{\partial r} \right]_{R_f} - \left[rk \frac{\partial T}{\partial r} \right]_{R_t} = 0. \quad (\text{A.10})$$

$$\left[k \frac{\partial T}{\partial r} + h_g T \right]_{R_f} - (h_g T)_{R_t} = 0, \quad (\text{A.11})$$

$$\left[k \frac{\partial T}{\partial r} + hT \right]_R - hT_c = 0, \quad (\text{A.12})$$

$$T_c(t=0) = \bar{T}_c, \quad (\text{A.13})$$

$$T_c(z=0) = T_{in}, \quad (\text{A.14})$$

$$\left[A_c k_c \frac{\partial T_c}{\partial z} \right]_{z=L} = 0, \quad (\text{A.15})$$

$$T_s(t=0) = \bar{T}_s, \quad (\text{A.16})$$

$$u(t=0) = \bar{u}, \quad (\text{A.17})$$

$$P(z=L) = P_{exit}, \quad (\text{A.18})$$

$$P(z=0) = P_{inlet}, \quad (\text{A.19})$$

$$n(t=0) = \bar{n}, \quad (\text{A.20})$$

and

$$C_i(t=0) = \tilde{C}_i = \frac{\beta_i \tilde{n}}{\Lambda \lambda_i}. \quad (\text{A.21})$$

The loop-hydraulics equation provides a relation between Eqs. (A.18) and (A.19):

$$P_{inlet} - (P_{exit} - f_L W^2 + \Delta H + \Delta P) = 0, \quad (\text{A.22})$$

where

$$W = \sum_{j=1}^{NC} [N_j(\rho_c A_c u)_j]_{z=0}. \quad (\text{A.23})$$

As already mentioned, Eqs. (A.1) through (A.5) together with the corresponding initial and boundary conditions [i.e., Eqs. (A.8) through (A.19)] refer to the j th channel, but the channel subscript j , $j = 1, \dots, NC$, was suppressed for notational simplicity. Thus, there are a total of NC sets of equations of the form (A.1) through (A.5) and (A.8) through (A.19). Alternatively, Eqs. (A.1) through (A.5) together with Eqs. (A.8) through (A.19) may be viewed as vector equations that are satisfied by the NC -component vectors \mathbf{T} , \mathbf{T}_c , \mathbf{T}_s , \mathbf{P} , and \mathbf{u} . Note that coupling among the various channels occurs solely through Eq. (A.22) and is specifically due to the mass flow rate W defined in Eq. (A.23).

Thus, the left sides of Eqs. (A.1) through (A.7) constitute the components of the (column) vector \mathbf{N} , which appears on the left side of Eq. (1). Correspondingly, the right sides of Eqs. (A.1) through (A.7) form the components of the source $\mathbf{Q}(\alpha)$ of Eq. (1). Similarly, the left and right sides of Eqs. (A.8) through (A.22) are the components of $\mathbf{B}(\mathbf{U}, \alpha)$ and $\mathbf{A}(\alpha)$ [namely, Eq. (2)], respectively.

III.B.6.f. Appendix B: The G-Differential and the G-Derivative of a Nonlinear Operator in the MELT-IIIB Model

If a (nonlinear) operator $F(e)$ has a G-differential $DF(e^0; \mathbf{h})$ at e^0 that is linear in \mathbf{h} , then the operator $F'(e^0)$ that satisfies the relationship

$$DF(e^0; \mathbf{h}) = F'(e^0)\mathbf{h} \quad (\text{B.1})$$

is by definition the G-derivative of F at e^0 . Thus, an operation F admits a G-derivative at e^0 if and only if (iff) its G-differential is linear in \mathbf{h} ; on the other hand, $DF(e^0; \mathbf{h})$ is linear in \mathbf{h} iff²⁰

$$1. \quad F \text{ satisfies a weak Lipschitz condition at } e^0, \quad (\text{B.2})$$

$$2. \quad F(e^0 + \epsilon \mathbf{h}_1 + \epsilon \mathbf{h}_2) - F(e^0 + \epsilon \mathbf{h}_1) - F(e^0 + \epsilon \mathbf{h}_2) + F(e^0) = O(\epsilon), \quad (\text{B.3})$$

where ϵ is a real scalar, and \mathbf{h}_1 and \mathbf{h}_2 are vectors in the same space as \mathbf{h} , e , and e^0 . An operator $F(e)$ is said to satisfy a weak Lipschitz condition at e^0 if to each unit vector \mathbf{h} there corresponds a $\delta(\mathbf{h}) > 0$ such that if $|\epsilon| < \delta$, then

$$\|F(e^0 + \epsilon \mathbf{h}) - F(e^0)\| \leq C \|\epsilon \mathbf{h}\|, \quad (\text{B.4})$$

where C does not depend on \mathbf{h} .

Each of the components of $N(e)$, $B(e)$, $Q(\alpha)$, and $A(\alpha)$ (whose explicit expressions have been given in Appendix A) can be shown to satisfy Eqs. (B.3) and (B.4). Consequently, these operators admit G-differentials that are linear in \mathbf{h} and, hence, they admit G-derivatives. This assertion is illustrated in the following by considering the operator

$$\frac{1}{r} \frac{\partial}{\partial r} \left[rk(T) \frac{\partial T}{\partial r} \right],$$

which appears in one component of $N(e)$ [see Eqs. (A.1) and (1)]. Again, the channel subscript is omitted for notational simplicity.

Thus, consider that

$$F(e) \equiv \frac{1}{r} \frac{\partial}{\partial r} \left[rk(T) \frac{\partial T}{\partial r} \right], \quad (\text{B.5})$$

where $e = (U, \alpha)$ and $\mathbf{h} = (h_U, h_\alpha)$. Note that F operates only on the components h_T and h_k of h_U and h_α , respectively. Also note that $h_T = h_T(r, z, t)$ and $h_k = h_k(r, z, t)$, i.e., h_k can represent spatial- and/or time-dependent changes in $k(T)$.

The following demonstration proves that $F(e)$ satisfies Eq. (B.3):

$$\begin{aligned} & F(e^0 + \epsilon h_1 + \epsilon h_2) - F(e^0 + \epsilon h_1) - F(e^0 + \epsilon h_2) + F(e^0) \\ &= \frac{1}{r} \frac{\partial}{\partial r} \left\{ r \left[(k^0 + \epsilon h_{k_1} + \epsilon h_{k_2}) \frac{\partial}{\partial r} (T^0 + \epsilon h_{T_1} + \epsilon h_{T_2}) - (k^0 + \epsilon h_{k_1}) \frac{\partial}{\partial r} (T^0 + \epsilon h_{T_1}) \right. \right. \\ & \left. \left. - (k^0 + \epsilon h_{k_2}) \frac{\partial}{\partial r} (T^0 + \epsilon h_{T_2}) + k^0 \frac{\partial T^0}{\partial r} \right] \right\} = -\frac{1}{r} \frac{\partial}{\partial r} \left[r \epsilon^2 \left[h_{k_1} \frac{\partial h_{T_2}}{\partial r} + h_{k_2} \frac{\partial h_{T_1}}{\partial r} \right] \right]. \end{aligned}$$

Therefore

$$F(e^0 + \epsilon h_1 + \epsilon h_2) - F(e^0 + \epsilon h_1) - F(e^0 + \epsilon h_2) + F(e^0) = 0(\epsilon),$$

which demonstrates that Eq. (B.3) is satisfied.

The following demonstration proves that $F(e)$ satisfies Eq. (B.4):

$$\begin{aligned} \frac{1}{\epsilon} \|F(e^0 + \epsilon h) - F(e^0)\| &= \left\| \frac{1}{\epsilon} \left[\frac{1}{r} \frac{\partial}{\partial r} \left[r (k^0 + \epsilon h_k) \frac{\partial (T^0 + \epsilon h_T)}{\partial r} \right] - \frac{1}{r} \frac{\partial}{\partial r} \left[r k^0 \frac{\partial T^0}{\partial r} \right] \right] \right\| \\ &= \left\| \frac{1}{r} \frac{\partial}{\partial r} \left[r k^0 \frac{\partial h_T}{\partial r} \right] + \frac{1}{r} \frac{\partial}{\partial r} \left[r h_k \frac{\partial T^0}{\partial r} \right] + \beta \right\| = \|DF(e^0)h + \beta\|, \end{aligned} \quad (B.6)$$

where

$$\beta \equiv \frac{\epsilon}{r} \frac{\partial}{\partial r} \left[r h_k \frac{\partial h_T}{\partial r} \right],$$

and where $DF(e^0)$ is a vector of the same dimension as h . Note that the only nonzero components of $DF(e^0)$ are

$$\frac{1}{r} \frac{\partial}{\partial r} \left[r k^0 \frac{\partial [\cdot]}{\partial r} \right] \quad \text{and} \quad \frac{1}{r} \left[\frac{\partial}{\partial r} r [\cdot] \frac{\partial T^0}{\partial r} \right];$$

these nonzero components correspond, respectively, to the components h_T and h_k of h .

Since $\lim_{\epsilon \rightarrow 0} \|\beta\| = 0$, it follows from Eq. (B.6) that

$$\lim_{\epsilon \rightarrow 0} \|DF(e^0)h + \beta\| \leq \lim_{\epsilon \rightarrow 0} [\|DF(e^0)\| \|h\| + \|\beta\|] = \|DF(e^0)\| \|h\| < [\|DF(e^0)\| + 1] \|h\|.$$

Thus, there exists $\delta(h) > 0$ so that for $\epsilon < \delta h$ the inequality

$$\|F(e^0 + \epsilon h) - F(e^0)\| < \epsilon [\|DF(e^0)\| + 1] \|h\| = C \|\epsilon h\| \quad (B.7)$$

is satisfied. This proves that $F(e)$ satisfies Eq. (B.4).

By proving that $F(e)$ satisfies the necessary and sufficient conditions stated in Eqs. (B.2) and (B.3), the foregoing derivations constitute a nonconstructive proof that the operator $F(e)$ defined by Eq. (B.5) admits a G-derivative at e^0 . Alternatively, a constructive proof that $F(e)$ admits a G-derivative at e^0 can be devised by applying Eq. (12) to Eq. (B.5) to determine the G-differential of $F(e)$ at e^0 , and by subsequently showing that the resulting operator is linear in \mathbf{h} . Thus, applying Eq. (12) to Eq. (B.5) gives

$$\begin{aligned} VF(e^0; \mathbf{h}) &= \frac{d}{d\epsilon} \left\{ \frac{1}{r} \frac{\partial}{\partial r} \left[r(k^0 + \epsilon h_k) \frac{\partial(T^0 + \epsilon h_T)}{\partial r} \right] \right\}_{\epsilon=0} \\ &= \frac{1}{r} \frac{\partial}{\partial r} \left[r k^0 \frac{\partial h_T}{\partial r} \right] + \frac{1}{r} \frac{\partial}{\partial r} \left[r h_k \frac{\partial T^0}{\partial r} \right]. \end{aligned} \quad (\text{B.8})$$

Clearly, Eq. (B.8) is linear in \mathbf{h} ; hence,

$$VF(e^0; \mathbf{h}) = DF(e^0; \mathbf{h}) = F'(e^0) \mathbf{h}, \quad (\text{B.9})$$

where $F'(e^0)$ is the G-derivative of $F(e)$ at e^0 . Explicitly, $F'(e^0)$ is in this case the $(M + I)$ -dimensional column vector

$$F'(e^0) = \underbrace{(0, \dots, 0, f_T, 0, \dots, 0)}_{\text{Dimension of } \mathbf{h}_U = M} \underbrace{(0, \dots, 0, f_k, 0, \dots, 0)}_{\text{Dimension of } \mathbf{h}_\alpha = I}^T \quad (\text{B.10})$$

whose nonzero components

$$f_T \equiv \frac{1}{r} \frac{\partial}{\partial r} \left[r k^0 \frac{\partial [\]}{\partial r} \right], \quad (\text{B.11})$$

and

$$f_k \equiv \frac{1}{r} \frac{\partial}{\partial r} \left[r [\] \frac{\partial T^0}{\partial r} \right], \quad (\text{B.12})$$

occupy [in Eq. (B.10)] the same positions as occupied, respectively, by h_T and h_k among the components of \mathbf{h} . Furthermore,

$$F'(e^0) \mathbf{h} = F'_U(e^0) \mathbf{h}_U + F'_\alpha(e^0) \mathbf{h}_\alpha, \quad (\text{B.13})$$

where $F'_U(e^0)$ is the partial G-derivative at e^0 of $F(e)$ with respect to U , and is represented by the M -dimensional column vector

$$F'_U(e^0) = (0, \dots, 0, f_T, 0, \dots, 0)^T,$$

and where $F'_\alpha(e^0)$ is the partial G-derivative at e^0 of $F(e)$ with respect to α , and is represented by the I -dimensional column vector

$$F'_\alpha(e^0) = (0, \dots, 0, f_k, 0, \dots, 0)^T.$$

Note that $F'(e^0)$ [given in Eq. (B.13)] and the quantity $DF(e^0)$ that was used to obtain Eq. (B.7) are, in fact, identical. Of course, this relationship was not yet known at that stage.

III.B.6.g. Appendix C: Representation of the Adjoint System

The adjoint systems given in Eqs. (49), (52), (55), (58), and (60) are all of the form

$$L^*(e^\circ)V = S, \quad (C.1)$$

where

$$V = (H^*, H_c^*, H_s^*, p^*, m^*, n^*, C_1^*, \dots, C_{NG}^*)$$

is a generic representation of the vectors V_n^* , V_T^* , Y^* , W_1^* , and W_2^* , and

$$S = (S_1, S_2, S_3, S_4, S_5, S_6, \dots, S_{13})$$

is a generic representation of the vectors S_n^* , S_T^* , G^* , F_1^* , and F_2^* . The explicit form of Eq. (C.1) is

$$\begin{aligned} & -\rho(T^\circ)C_\rho(T^\circ) \frac{\partial H^*}{\partial t} - k(T^\circ) \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial H^*}{\partial r} \right) \\ & - \frac{n}{\lambda} \left(\frac{\partial \bar{p}}{\partial T} \right) n^* = S_1; \\ & r \in (0, R_f) \cup (R_g, R); \quad z \in (0, L); \quad t \in (0, t_f), \end{aligned} \quad (C.2)$$

$$\begin{aligned} & -A_c \rho_c C_{pc} \frac{\partial H_c^*}{\partial t} - \rho_c C_{pc} \frac{\partial}{\partial z} (A_c u H_c^*) - A_c k_c \frac{\partial^2 H_c^*}{\partial z^2} \\ & + 2\pi R h (H_c^* - H^*) + 2\pi R_s h (H_c^* - H_s^*) \\ & - \left(\frac{\partial P}{\partial z} \frac{1}{\rho_c} \frac{\partial \rho_c}{\partial T_c} - \frac{\rho_c u^2}{2D} \frac{\partial f}{\partial T_c} \right) p^* - A_c \frac{\partial \rho_c}{\partial T_c} \frac{Dm^*}{Dt} \\ & - \frac{n}{\lambda} \frac{\partial \bar{p}}{\partial T_c} n^* = S_2; \quad z \in (0, L), \quad t \in (0, t_f), \end{aligned} \quad (C.3)$$

$$\begin{aligned} & -A_c \rho_c C_{ps} \frac{\partial H_s^*}{\partial t} + 2\pi R_s h (H_s^* - H_c^*) = S_3; \\ & z \in (0, L), \quad t \in (0, t_f), \end{aligned} \quad (C.4)$$

$$\begin{aligned} & -\rho_c A_c \frac{\partial m^*}{\partial z} - \frac{\partial}{\partial t} (\rho_c p^*) - \frac{\partial}{\partial z} (u \rho_c p^*) \\ & + \rho_c \left[\frac{fu}{D} + \frac{u^2}{2D} \frac{\partial f}{\partial u} + \sum_m \delta(t - z_m) u K_L \right] p^* \\ & + A_c \rho_c C_{pc} \frac{\partial T_c}{\partial z} H_c^* = S_4; \\ & z \in (0, L), \quad t \in (0, t_f), \end{aligned} \quad (C.5)$$

$$-\frac{\partial p^*}{\partial z} = S_5, \quad z \in (0, L), \quad t \in (0, t_f),$$

$$\begin{aligned}
& -\frac{dn^*}{dt} - \frac{\bar{\rho} - \beta}{\Lambda} n^* - \sum_{i=1}^{NC} N_j \left\{ \int_r \int_z \psi H^* 2\pi r dr dz \right. \\
& \quad \left. + \int_z \psi H_c^* dz - \sum_{i=1}^{NC} \frac{\beta_i}{\Lambda} C_i^* - \delta(t - t_0) \left(\frac{1}{dn/dt} \right)_{t=t_0} \right. \\
& \quad \times \left[\frac{1}{\Lambda} \int_{t_0+\Delta t}^{t_f} n n^* \frac{\partial \bar{\rho}}{\partial \tau} dt + \int_{t_0+\Delta t_p}^{t_f} \frac{\partial(\Delta P)}{\partial \tau_p} \right. \\
& \quad \left. \left. \times \sum_{j=1}^{NC} N_j p_j^* dt \right] \right\} = S_6, \quad t \in (0, t_f), \quad (C.7)
\end{aligned}$$

$$\begin{aligned}
& -\frac{dC_i^*}{dt} + \lambda_i(C_i^* - n^*) = S_{6+i}, \\
& \quad i = 1, \dots, NG, \quad t \in (0, t_f). \quad (C.8)
\end{aligned}$$

The superscript $^\circ$ is used in Eq. (C.2) to explicitly denote that the components of $\mathbf{L}^*(e^\circ)$ depend on the base-case value e° of $e = (U, \alpha)$. Although the explicit display of this dependence was subsequently omitted for notational simplicity, it should be understood throughout this Appendix that all e -dependent quantities are evaluated at e° ; thus, these quantities are just coefficients whose values are known from the base-case solution. Also, note that the term involving $\delta(t - t_0)$ in Eq. (C.7) is due to the particular problem (i.e., a protected transient with scram on high power) analyzed in this work. Other types of scram initiators or switches are discussed in Ref. 16.

The adjoint boundary conditions and final-time conditions for Eqs. (C.2) through (C.8) are

$$V = 0 \text{ at } t = t_f, \quad (C.9)$$

$$H_c^* = 0 \text{ at } z = 0, \quad (C.10)$$

$$u \rho_c C_{pc} H_c^* + k_c \frac{\partial H_c^*}{\partial z} + u \frac{\partial \rho_c}{\partial T_c} m^* = 0 \text{ at } z = L, \quad (C.11)$$

$$A_c m^* + p^* u = 0 \text{ at } z = L, \quad (C.12)$$

$$\left(k \frac{\partial H^*}{\partial r} + h_g H^* \right)_{r=R_f} - (h_g H^*)_{r=R_g} = 0, \quad (C.13)$$

$$\left(r k \frac{\partial H^*}{\partial r} + h H^* \right)_{r=R} - h H_c^* = 0, \quad (C.14)$$

$$\left(\frac{\partial H^*}{\partial r} \right)_{r=0} = 0, \quad (C.15)$$

and, for the j 'th channel,

$$(up^* + A_c m^*)_j - 2f_{t,1} \sqrt{A_c} \sum_{i=1}^{NC} N_i p_i^* = 0 \text{ at } z = 0 . \quad (\text{C.16})$$

Of course, there is a one-to-one correspondence between Eqs. (C.2) through (C.8) and Eqs. (A.1) through (A.7). Thus, just like in the case of Eqs. (A.1) through (A.5), Eqs. (C.2) through (C.6) are actually vector equations (in that they refer to NC distinct channels) although, as written here, this fact was not explicitly indicated. Note that coupling between channels in the adjoint system occurs solely via Eq. (C.16); this is explicitly indicated in Eq. (C.16) by the use of the index j .

III.B.6.h. Nomenclature

A_c/A_s	=	cross-sectional area of coolant/structure, m^2
C_i	=	precursor amplitude for group i , MW
C_i^*	=	adjoint precursor amplitude for group i , (response dimension)/MW
$C_p/C_{pc}/C_{ps}$	=	heat capacity of fuel pin/coolant/structure, J/kg K
D	=	equivalent diameter for coolant channel, m
D/Dt	=	one-dimensional substantial (i.e., total) derivative operator
f	=	channel friction factor
f_L	=	loop friction multiplier, $(Pa/kg/s)^2$
g	=	gravitational constant = $9.807 m/s^2$
ΔH	=	cold leg pressure head, Pa
$H^*/H_c^*/H_s^*$	=	adjoint enthalpy of fuel pin/coolant/structure, (response dimension)/J
h_g/h	=	heat transfer coefficient for gap/coolant, $W/m^2 K$
K_L	=	pressure head loss coefficient for an abrupt area change
k/k_c	=	thermal conductivity in fuel pin/coolant, $W/m^2 K$
L	=	length of coolant channel, m
m^*	=	adjoint mass, (response dimension)/kg
N_j	=	number of pins in channel j
NC	=	number of coolant channels
NG	=	number of precursor groups
n	=	neutronic power amplitude, MW
n^*	=	adjoint power, (response dimension)/MW
n_γ	=	gamma-ray heating power amplitude, MW
n_o	=	trip power level for reaction scram, MW
ΔP	=	pressure drop across pump, Pa
P	=	coolant channel pressure, Pa
P_{inlet}	=	inlet plenum pressure, Pa
P_{exit}	=	exit plenum pressure, Pa
P^*	=	adjoint momentum flux, (response dimension)/[$kg m/s/m^2$]
$R_f/R_g/R$	=	outer radius of fuel/gas/cladding, m

R_s	=	inner radius of structure, m
r	=	fuel pin radius variable, m
$T/T_c/T_s$	=	temperature in fuel pin/coolant/structure, K
T_{in}	=	inlet coolant temperature, K
t	=	time variable, s
t_f	=	final time value (also used to initiate adjoint calculation), s
t_o	=	trip time of reactor scram, s
Δt	=	time delay between scram trip and scram reactivity insertion, s
Δt_p	=	time delay between scram trip and pump coastdown, s
u	=	channel coolant velocity, m/s
W	=	reactor mass flow rate, kg/s
z	=	axial direction variable, m

Greek Symbols

β_i	=	delayed neutron fraction for precursor group i
β	=	$\sum_{i=1}^{NG} \beta_i$
Λ	=	prompt neutron generation time, s
λ_i	=	precursor decay constant for group i , s^{-1}
$\rho/\rho_c/\rho_s$	=	density in fuel pin/coolant/structure, kg/m^3
$\bar{\rho}$	=	total reactivity
$\bar{\rho}_{prog}$	=	programmed input reactivity
$\bar{\rho}_{scram}$	=	programmed scram reactivity
$\bar{\rho}_{Dopp}$	=	Doppler and other feedback reactivity
ψ	=	normalized power shaping function, which includes the coolant regions, $W/(m^3 MW)$

Subscripts

m	=	coolant channel abrupt area change m
-----	---	--

Superscripts

\sim	=	steady-state quantity
*	=	adjoint quantity
\circ	=	base-case value

III.C. NONLINEAR SYSTEMS WITH FEEDBACK

The material presented in this section is largely based on the following article: D. G. Cacuci and M. C. G. Hall, "Efficient Estimation of Feedback Effects with Application to Climate Models," *J. Atm. Sci.*, **41** (13), 1984.

III.C.1. Introduction

A simplification that frequently occurs in large computer models is the use of experimentally observed values for what should be prognostically determined variables (i.e. dependent variables). Such a simplification is useful because the interactive modeling of a particularly complicated or ill-understood process can be postponed, while it is still possible to make physically meaningful comparisons between experimental observations and quantities that remain prognostically determined. When a computer model is used predictively, the experimentally prescribed quantities are in reality subject to change due to forcing influences. When the effect of such a change is taken into account, this is usually referred to as including the effect of feedback because quantities that are normally input in the model (e.g., parameters or data) are allowed to depend on the model output (i.e., the dependent variables).

When there is more than one plausible way of incorporating feedback in a model, it is useful to experiment with various forms of the feedback. But for more complex models, recalculation for each of several different forms of the feedback can be prohibitively expensive. However, the Adjoint Method cannot be applied directly to estimate the effect of feedback because the variations in the parameters are not prescribed, but depend on the output of the model. Consequently, in Section III.C.2. the use of the term *sensitivity to feedback* is defined and justified. It is then shown how this sensitivity can be estimated using the Adjoint Method.

III.C.2. Definition and Estimation of Sensitivity to Feedback

We consider, as before, that the physical process is modeled mathematically by:

$$\left. \begin{aligned} \mathbf{N}[\mathbf{u}(\mathbf{x}), \alpha(\mathbf{x})] &= 0 \\ \mathbf{B}[\mathbf{u}(\mathbf{x}), \alpha(\mathbf{x})] &= 0 \end{aligned} \right\} \quad (1)$$

where the meaning of the various variables is the same as in Section III.A.

For clarity, a simple one-dimensional climate model illustrates the following general development. This illustrative model is described by the equations

$$\left. \begin{aligned} du/dt + \alpha_1 u^4 + \alpha_2 &= 0 \\ u(a) - u_a &= 0 \end{aligned} \right\} \quad (2)$$

For this model, the only independent variable is time t which varies from a to b , the only dependent variable is the temperature $u(t)$, and the two parameters $\alpha = (\alpha_1, \alpha_2)$ are constants that depend on the physical properties of the system such as heat capacity, incident radiation, albedo and emissivity. The initial value of u , is u_a .

A scalar result R of a model described by (1) can in general be expressed as a functional of \mathbf{u} and α . For example, the average longwave radiation in the illustrative model described by (2) is proportional to the functional

$$R(\mathbf{u}, \alpha) = \int_a^b dt (\alpha_1 u^4). \quad (3)$$

In the following, α^0 denotes the parameters' nominal values, and \mathbf{u}^0 denotes the nominal solution of (1). Thus the nominal solution satisfies

$$\left. \begin{aligned} \mathbf{N}(\mathbf{u}^0, \alpha^0) &= 0 \\ \mathbf{B}(\mathbf{u}^0, \alpha^0) &= 0 \end{aligned} \right\} \quad (4)$$

and the nominal result is $R(\mathbf{u}^0, \alpha^0)$.

Feedback can be introduced into the model by allowing some of the parameters α to depend on the components of \mathbf{u} . Without loss of generality, this feedback can be specified by adding an operator $\mathbf{A}(\mathbf{u})$ to the parameters' nominal values α^0 . Thus, in the presence of feedback, the parameters' values become $\alpha^0 + \mathbf{A}(\mathbf{u})$, and the solution \mathbf{u}^f with feedback satisfies

$$\left. \begin{aligned} \mathbf{N}[\mathbf{u}^f, \alpha^0 + \mathbf{A}(\mathbf{u}^f)] &= 0 \\ \mathbf{B}[\mathbf{u}^f, \alpha^0 + \mathbf{A}(\mathbf{u}^f)] &= 0 \end{aligned} \right\} \quad (5)$$

The result with feedback is $R[\mathbf{u}^f, \alpha^0 + \mathbf{A}(\mathbf{u}^f)]$. In the illustrative model, feedback can be introduced by allowing the emissivity to depend on temperature. For example, when the value of α_1 is allowed to be $\alpha_1^0 + \lambda(u - u_a)$, where λ is a constant specifying the strength of the feedback, then the feedback operator $\mathbf{A}(u)$ is the vector

$$\mathbf{A}(u) = [\lambda(u - u_a), 0]. \quad (6)$$

The solution u^f with feedback now satisfies

$$\left. \begin{aligned} du^f/dt + [\alpha_1^0 + \lambda(u^f - u_a)](u^f)^4 + \alpha_2^0 &= 0 \\ u^f(a) - u_a &= 0 \end{aligned} \right\}$$

and the result (3) with feedback is

$$\int_a^b dt [\alpha_1^0 + \lambda(u^f - u_a)](u^f)^4.$$

The difference

$$R[\mathbf{u}^f, \alpha^0 + \mathbf{A}(\mathbf{u}^f)] - R(\mathbf{u}^0, \alpha^0) \quad (7)$$

gives the actual effect of the feedback \mathbf{A} on the result $R(\mathbf{u}^0, \alpha^0)$. In practice, this difference can be evaluated exactly only by introducing feedback into the model explicitly, calculating \mathbf{u}^f , and re-evaluating the result. For more complex climate models such as atmospheric general circulation models, rerunning the model more than once can be prohibitively expensive. This provides the motivation to develop a more efficient method of evaluating (7).

Note that the Adjoint Method developed in Sec. III.A cannot be applied directly to the evaluation of (7) because, with feedback, the variation $\mathbf{A}(\mathbf{u}^f)$ in α^0 is not prescribed but depends on \mathbf{u}^f . The purpose of the following development is to provide an approximate expression for (7) that can be evaluated efficiently using the adjoint method.

The functional $VR_{\mathbf{A}}(\mathbf{h})$ is defined by

$$VR_{\mathbf{A}}(\mathbf{h}) = \lim_{\epsilon \rightarrow 0} \left\{ R[\mathbf{u}^0 + \epsilon \mathbf{h}, \alpha^0 + \epsilon \mathbf{A}(\mathbf{u}^0 + \epsilon \mathbf{h})] - R(\mathbf{u}^0, \alpha^0) \right\} / \epsilon, \quad (8)$$

where \mathbf{h} is an arbitrary set of increments to the dependent variables \mathbf{u}^0 , and ϵ is a real number. The functional $\Delta R_{\mathbf{A}}(\mathbf{h})$ is defined by

$$\Delta R_{\mathbf{A}}(\mathbf{h}) = \left\{ R[\mathbf{u}^0 + \mathbf{h}, \alpha^0 + \mathbf{A}(\mathbf{u}^0 + \mathbf{h})] - R(\mathbf{u}^0, \alpha^0) \right\} - VR_{\mathbf{A}}(\mathbf{h}). \quad (9)$$

Note that if the arbitrary value of \mathbf{h} is chosen to be $\mathbf{u}^f - \mathbf{u}^0$, then (9) can be written

$$R[\mathbf{u}^f, \alpha^0 + \mathbf{A}(\mathbf{u}^f)] - R(\mathbf{u}^0, \alpha^0) = VR_{\mathbf{A}}(\mathbf{h}) + \Delta R_{\mathbf{A}}(\mathbf{h}). \quad (10)$$

Thus for $\mathbf{h} = \mathbf{u}^f - \mathbf{u}^0$, $\Delta R_{\mathbf{A}}(\mathbf{h})$ is the discrepancy between the actual effect of feedback (7) and the functional (8).

The properties of the functional (9) can be determined as follows. Since both \mathbf{A} and \mathbf{h} are arbitrary in (9) they can be replaced respectively by $\epsilon \mathbf{A}$ and $\epsilon \mathbf{h}$ to give

$$\Delta R_{\epsilon \mathbf{A}}(\epsilon \mathbf{h}) = \left\{ R[\mathbf{u}^0 + \epsilon \mathbf{h}, \alpha^0 + \epsilon \mathbf{A}(\mathbf{u}^0 + \epsilon \mathbf{h})] - R(\mathbf{u}^0, \alpha^0) \right\} - VR_{\mathbf{A}}(\epsilon \mathbf{h}). \quad (11)$$

The definition in (8) shows that $VR_{\epsilon \mathbf{A}}(\epsilon \mathbf{h}) = \epsilon VR_{\mathbf{A}}(\mathbf{h})$, and so dividing (11) by ϵ and letting $\epsilon \rightarrow 0$ gives

$$\lim_{\epsilon \rightarrow 0} \Delta R_{\epsilon A}(\epsilon \mathbf{h})/\epsilon = 0. \quad (12)$$

This equation shows that $\Delta R_{\epsilon A}(\mathbf{h})$ contains no first-order terms in \mathbf{A} or \mathbf{h} . Thus, for $\mathbf{h} = \mathbf{u}^f - \mathbf{u}^0$, (10) and (12) show that the functional $VR_{\mathbf{A}}(\mathbf{h})$ given by (8) is an estimate of the actual effect of feedback (7) correct to first order in \mathbf{A} and \mathbf{h} . Consequently, $VR_{\mathbf{A}}(\mathbf{h})$ can be called the *sensitivity of R to feedback \mathbf{A}* .

In practice, it is more convenient to use the following definition of $VR_{\mathbf{A}}$ which is equivalent to (8):

$$VR_{\mathbf{A}} = \left\{ (d/d\epsilon) R[\mathbf{u}^0 + \epsilon \mathbf{h}, \alpha^0 + \epsilon \mathbf{A}(\mathbf{u}^0 + \epsilon \mathbf{h})] \right\}_{\epsilon=0}. \quad (13)$$

For nearly all physical models, performing the differentiation in (13) gives

$$VR_{\mathbf{A}} = R'_1 \mathbf{h} + R'_2 \mathbf{A}(\mathbf{u}^0), \quad (14)$$

where R'_1 and R'_2 denote, respectively, the partial Gâteaux derivatives at (\mathbf{u}^0, α^0) of $R(\mathbf{u}, \alpha)$ with respect to its first and second arguments. For example, with the result R defined by (3) and the feedback \mathbf{A} defined by (6), $VR_{\mathbf{A}}$ is obtained as follows:

$$\begin{aligned} VR_{\mathbf{A}} &= \left\{ (d/d\epsilon) \int_a^b dt [\alpha_1^0 + \epsilon \lambda(u^0 + \epsilon h - u_a)] X (u^0 + \epsilon h)^4 \right\}_{\epsilon=0} \\ &= \int_a^b dt [\alpha_1^0 4(u^0)^3 h] + \int_a^b dt [(u^0)^4 \lambda(u^0 - u_a)]. \end{aligned}$$

Note that for this model R'_1 is the operator

$$R'_1(\quad) = \int_a^b dt [\alpha_1^0 4(u^0)^3 (\quad)].$$

and R'_2 is the operator

$$R'_2(\quad) = \left[\int_a^b dt [(u^0)^4 \lambda(u^0 - u_a)], 0 \right] \cdot (\quad).$$

To evaluate the sensitivity to feedback given by (14), Eqs. (4) and (5) are needed to determine $\mathbf{h} = \mathbf{u}^f - \mathbf{u}^0$. Subtracting (4) from (5) gives

$$\left. \begin{aligned} \mathbf{N}[\mathbf{u}^f, \alpha^0 + \mathbf{A}(\mathbf{u}^f)] - \mathbf{N}(\mathbf{u}^0, \alpha^0) &= 0 \\ \mathbf{B}[\mathbf{u}^f, \alpha^0 + \mathbf{A}(\mathbf{u}^f)] - \mathbf{B}(\mathbf{u}^0, \alpha^0) &= 0 \end{aligned} \right\} \quad (15)$$

Relationships equivalent to (10) hold for the operators \mathbf{N} and \mathbf{B} ; applying these relationships to (15) gives

$$\left. \begin{aligned} \mathbf{V}\mathbf{N}_A(\mathbf{h}) + \Delta\mathbf{N}_A(\mathbf{h}) &= 0 \\ \mathbf{V}\mathbf{B}_A(\mathbf{h}) + \Delta\mathbf{B}_A(\mathbf{h}) &= 0 \end{aligned} \right\} \quad (16)$$

where

$$\lim_{\epsilon \rightarrow 0} \Delta\mathbf{N}_{\epsilon A}(\epsilon\mathbf{h})/\epsilon = \lim_{\epsilon \rightarrow 0} \Delta\mathbf{B}_{\epsilon A}(\epsilon\mathbf{h})/\epsilon = 0. \quad (17)$$

Also, relationships equivalent to (14) hold for $\mathbf{V}\mathbf{N}_A$ and $\mathbf{V}\mathbf{B}_A$; applying these relationships to (16) gives

$$\left. \begin{aligned} \mathbf{N}'_1\mathbf{h} + \mathbf{N}'_2\mathbf{A}(\mathbf{u}^0) + \Delta\mathbf{N}_A(\mathbf{h}) &= 0 \\ \mathbf{B}'_1\mathbf{h} + \mathbf{B}'_2\mathbf{A}(\mathbf{u}^0) + \Delta\mathbf{B}_A(\mathbf{h}) &= 0 \end{aligned} \right\} \quad (18)$$

For the illustrative model, with N and B defined by (2) and \mathbf{A} defined by (6), $\mathbf{V}\mathbf{N}_A(h)$ and $\mathbf{V}\mathbf{B}_A(h)$ can be obtained as follows:

$$\begin{aligned} \mathbf{V}\mathbf{N}_A(h) &= ((d/d\epsilon)\{(d/dt)(u^0 + \epsilon h) + [\alpha_1^0 + \epsilon\lambda(u^0 + \epsilon h - u_a)](u^0 + \epsilon h)^4 + \alpha_2^0\})_{\epsilon=0} \\ &= [d/dt + \alpha_1^0 4(u^0)^3]h + (u^0)^4\lambda(u^0 - u_a), \\ \mathbf{V}\mathbf{B}_A(h) &= (d/d\epsilon)[u(a) + \epsilon h(a) - u_a]_{\epsilon=0} \\ &= h(a). \end{aligned}$$

Note that for this model, \mathbf{N}'_1 is the operator

$$\mathbf{N}'_1(\quad) = [d/dt + \alpha_1^0 f(u^0)^3](\quad),$$

\mathbf{N}'_2 is the operator

$$\mathbf{N}'_2(\quad) = [(u^0)^4, 1] \cdot (\quad),$$

and the boundary conditions in (16) and (18) become

$$h(a) = -\Delta\mathbf{B}_A(h). \quad (19)$$

The problem of efficiently evaluating the sensitivity (14) where $\mathbf{A}(\mathbf{u}^0)$ is known and \mathbf{h} is determined by (18) is precisely the problem addressed in Sec. III.A in the development of the adjoint method. The purpose of this method is to evaluate the sensitivity (14) without explicitly evaluating \mathbf{h} , thereby avoiding the need to solve (18) anew for every different feedback $\mathbf{A}(\mathbf{u}^0)$. The adjoint method starts by defining an operator \mathbf{L}^* adjoint to \mathbf{N}'_1 as follows:

$$\langle \mathbf{q} | \mathbf{N}'_1 \mathbf{r} \rangle = \langle \mathbf{r} | \mathbf{L}^* \mathbf{q} \rangle + P(\mathbf{q}, \mathbf{r}), \quad (20)$$

where \mathbf{q} and \mathbf{r} are arbitrary functions of \mathbf{x} , $\langle \mathbf{q} | \mathbf{r} \rangle$ denotes the scalar product of \mathbf{q} and \mathbf{r} in the region of physical interest Ω , and $P(\mathbf{q}, \mathbf{r})$ is a term evaluated on the boundary of this region. For the illustrative model, the scalar product is

$$\langle q | r \rangle = \int_a^b dt [q(t)r(t)].$$

For this model, (20) can be written

$$\begin{aligned} & \int_a^b dt \left\{ q \left[\frac{d}{dt} + \alpha_1^0 4(u^0)^3 \right] r \right\} \\ &= \int_a^b dt \left\{ r \left[-\frac{d}{dt} + \alpha_1^0 4(u^0)^3 \right] q \right\} + [qr]_a^b. \end{aligned}$$

Thus L^* is the operator

$$L^* = \left[-\frac{d}{dt} + \alpha_1^0 4(u^0)^3 \right],$$

and $P(q, r)$ is the term

$$P(q, r) = [qr]_a^b.$$

The adjoint solution $\mathbf{v}(\mathbf{x})$ is the solution to the system of equations

$$\left. \begin{aligned} \mathbf{L}^* \mathbf{v} &= \mathbf{s} \\ \mathbf{B}^* \mathbf{v} &= 0 \end{aligned} \right\} \quad (21)$$

where \mathbf{s} is a source term defined by

$$\langle \mathbf{s} | \mathbf{q} \rangle = R'_1 \mathbf{q}, \quad (22)$$

and \mathbf{B}^* is an operator representing the adjoint boundary conditions that will be defined later. For the illustrative model, (22) becomes

$$\int_a^b dt [sq] = \int_a^b dt \alpha_1^0 4(u^0)^3 q.$$

Thus s is the term $\alpha_1^0 4(u^0)^3$, and the first of Eqs. (21) becomes

$$[-d/dt + \alpha_1^0 4(u^0)^3]v = \alpha_1^0 4(u^0)^3. \quad (23)$$

The adjoint method concludes by expressing the sensitivity VR_A in terms of an adjoint solution as follows:

$$VR_A = R_2' A(u^0) - \langle v | N_2' A(u^0) \rangle - P(h, v) + \Delta_A(h), \quad (24)$$

where

$$\lim_{\epsilon \rightarrow 0} \Delta_{\epsilon A}(\epsilon h) / \epsilon = 0.$$

The adjoint boundary conditions are chosen to eliminate the unknown values of h from $P(h, v)$ in (24). For example, with the illustrative model Eq. (24) becomes

$$\begin{aligned} VR_A &= \int_a^b dt [(u^0)^4 \lambda(u^0 - u_a)] \\ &- \int_a^b dt [v(u^0)^4 \lambda(u^0 - u_a)] - [hv]_a^b + \Delta_A(h). \end{aligned} \quad (25)$$

The value of $h(a)$ in this equation is known from the initial conditions (19). Thus the only unknown value of h , i.e., $h(b)$, can be eliminated from (25) by choosing the adjoint boundary condition

$$v(b) = 0. \quad (26)$$

The advantage of the adjoint method is that the adjoint solution is independent of the feedback being considered, and all values of h for VR_A are known without having to solve (18). Thus, once the adjoint solution v has been calculated, it is possible to estimate the effect of many different feedbacks without solving any additional differential equations. This advantage of the adjoint method can be seen for the illustrative model. The adjoint equations (23) and (26) do not contain any terms arising from the feedback operator. Moreover, the equivalent of (25) can be derived from (24) for a general feedback operator $[A_1(u), A_2(u)]$:

$$\begin{aligned} VR_A &= \left[\int_a^b dt (u^0)^4, 0 \right] \cdot [A_1(u^0), A_2(u^0)] \\ &- \int_a^b dt \{ v [(u^0)^4, 1] \cdot [A_1(u^0), A_2(u^0)] \} + \Delta B_A(h) v(a) + \Delta_A(h). \end{aligned}$$

Thus, once (23) and (26) have been solved, the above equation can be used to estimate the effect of any feedback where the Δ terms are neglected.

IV. SUMMARY AND CONCLUSIONS

The methods and concepts of abstract analysis have been employed to formulate a sensitivity theory for physical problems described by systems of coupled nonlinear equations, and nonlinear responses. Greater generality has been achieved by considering the problem and the response as mappings defined on the product space $E = E_u \times E_\alpha$. Consequently, it has been possible to circumvent the need to assume a specific form for the response $\mathbf{R}(e)$. The scope and versatility of the present formulation of sensitivity theory have also been extended by defining the sensitivity of the response to variations in the system parameters (α) as the G-differential $\mathcal{V}\mathbf{R}(e^0;h)$ of $\mathbf{R}(e)$ at e^0 .

Two alternative formalisms have been developed to evaluate the sensitivity $\mathcal{V}\mathbf{R}(e^0;h)$ of $\mathbf{R}(e)$: the Forward Method and the Adjoint Method. As has been shown, there are clear distinctions between the necessary and sufficient conditions required for the validity of each formalism. On the one hand, it has been demonstrated that the Forward Method can be rigorously formulated in normed linear spaces, and that the existence of the G-differentials of all operators appearing in the original nonlinear equations are the necessary and sufficient conditions underlying the validity of this formalism. It has also been emphasized that these G-differentials are not linear operators.

On the other hand, it has been shown that the necessary and sufficient conditions underlying the validity of the Adjoint Method are more restrictive. Most prominent among these conditions is the requirement that all operators acting on the state vector u must admit densely defined partial G-derivatives at $e^0 = (u^0, \alpha^0)$ with respect to u . Furthermore, the underlying normed linear spaces have to be complete in order that the Adjoint Method be unique and generally valid. By setting the development of this formalism in Hilbert spaces, the Riesz representation theorem was shown to play a fundamental role. Although this theorem does not hold in general in a pre-Hilbert space [e.g., $\nabla_u R(e^0)$ in Eq. (20) of Section III.A. may not exist], in many practical applications it may do so. Thus, the Adjoint Method may still be applicable to certain problems which fit naturally in a pre-Hilbert space that may not be convenient to complete in practice. (Theoretically, of course, pre-Hilbert spaces can always be completed.)

Note that the need to introduce *any* derivatives of operators acting *solely* on the system parameters α , or derivatives of the state vector with respect to α , has been completely eliminated. As has been shown, the existence of the G-differentials $\mathcal{V}\mathbf{Q}(\alpha^0;h_\alpha)$ and $\mathcal{V}\mathbf{A}(\alpha^0;h_\alpha)$ is both necessary and sufficient. Furthermore, the use of distinct inner products makes it possible to treat problems involving operators whose range is not in the same Hilbert space as the state vector. Finally, the results obtained by employing the previous approaches^{1-19,21} to sensitivity theory can be recovered as particular forms of the results obtained here. Altogether, these factors contribute to the greater generality and applicability of the Adjoint Method presented here.

It is of practical interest to mention that, in particular applications, additional conditions may need to be imposed on the operators $\mathbf{N}, \mathbf{B}, \mathbf{Q}$, and \mathbf{A} , in order to solve Eqs. (1) and (2) of Section III.A. by some particular numerical procedure. For example, several of the most widely used numerical methods³¹ for solving nonlinear operator equations require the existence of Fréchet derivatives of \mathbf{N} and \mathbf{B} at e^0 ; in such cases, the conditions underlying the validity of the Adjoint Method would automatically be satisfied.

The sensitivity theory presented in Section III.A.2.b.(i). has been restricted to responses that are functionals in order to highlight the intimate connection between the construction of the adjoint system and the mathematical nature of the response. This connection is underscored by recalling the essential role played by the Riesz representation theorem when identifying the sensitivity of the response with an inner product [cf., Eqs. (20 and (22) of Section III.A.]

Sensitivity theory has also been formulated in Section III.A.2.b.(ii). for nonlinear systems with general operators as responses. It has been shown that there are essentially no conceptual and computational differences between the treatment of operators and the treatment of functionals as responses within the Forward Method. However, there is a considerable difference between the treatment of these two types of responses within the Adjoint Method. Thus, the Adjoint Method can be developed only if the h_u -dependence of the G-differential giving the sensitivity of the operator-type response is expressible as a linear combination of linear functionals of h_u . For this purpose, it has been necessary and sufficient to consider the response $\mathbf{R}(e)$ to be an element of the Hilbert space $H_R(\Omega_R)$, to introduce an orthonormal basis for $H_R(\Omega_R)$, and to require the existence of the G-derivative of $\mathbf{R}(e)$ at e^0 . The indirect effect term has then been expressed as a linear combination of basis elements, each of these elements being multiplied by a linear functional of h_u which contained the entire h_u -dependence of the response sensitivity. This h_u -dependence has in turn been eliminated from the expression of each of these functionals by using adjoint functions satisfying appropriately constructed adjoint systems.

When derived via the Adjoint Method, the exact expression of the sensitivity of an operator-type response contains as many adjoint functions as there are non-zero terms in the linear combination of basis elements. This linear combination may, in principle, contain infinitely many terms. To minimize the computation of adjoint functions, it becomes important to select a basis and a notion of convergence to represent the indirect effect term as accurately as possible with the smallest number of basis elements. It is also desirable to derive, if possible, bounds and/or asymptotic expressions for the remainder after truncating the linear combination expressing the indirect effect term.

It has already been established (in many works on sensitivity analysis) that the Adjoint Method is the most economical to use, whenever possible, if the physical problem involves a large data base (or many alterations in the data) and comparatively few functional-type responses. For operator-type responses, however, the specific needs of sensitivity analysis, the number of system parameters and responses and the characteristics of each response must be examined to determine whether computational costs warrant the use of the Adjoint Method.

Section III.B. presented the formulation of sensitivity theory for nonlinear systems with responses that are functionals defined at a critical point of a function $F(u, x, \alpha)$ of the system's state vector and parameters. In practice, this critical point may represent any extremum, saddle, or inflexion point of $F(u, x, \alpha)$. It has been shown that changes in the system parameters affect both the numerical value of the response and the critical point itself. Expressions for the sensitivity of the numerical value of the response and for the sensitivity of the critical point have been obtained within the context of the Forward Method by directly applying the definition of the G-differential. However, since it is expensive to use this method to answer all sensitivity questions that might arise in practice, the Adjoint Method has been developed to yield alternative expressions for the desired sensitivities. This method requires the computation of as many adjoint functions as there are components of the critical point in phase-space, and of one additional adjoint function to evaluate the sensitivity of the numerical value of the response. Once these adjoint functions have been computed, the sensitivities to *all* possible changes in the system parameters can be obtained by simple quadratures. This makes the Adjoint Method the most cost-efficient method to use whenever possible, although, as has been discussed, the necessary and sufficient conditions underlying its validity are more restrictive than those underlying the validity of the Forward Method.

The sensitivity theory presented in Secs. III.B.1. through III.B.4. has been applied in Sec. III.B.6. to a problem of interest in reactor safety, namely a protected transient with scram on high-power level in the Fast Flux Test Facility (FFTF). To determine the base-case distribution of the dependent variables for this problem, the entire single-phase segment of the fast reactor safety code MELT-IIIB has been used, including (i) thermal-hydraulic equations, for each channel type, describing the behavior of the average channel fuel pin and surrounding coolant, (ii) an equation describing the primary loop hydraulics, and (iii) neutron point-kinetics equations describing the reactor power level.

Two extremum-type responses have been considered - the maximum power response and the maximum fuel temperature response in the hot channel. Expressions for the sensitivities of the respective numerical values of these responses and expressions for the sensitivities of the phase-space locations at which the respective maxima occurred have been obtained in terms of adjoint functions. The adjoint systems satisfied, in turn, by each of these adjoint functions have been derived and solved. It has been shown that the complete sensitivity analysis of each (extremum-type) response requires the computation of as many adjoint functions as there are non-zero components of the respective critical point (e.g., maximum) in phase space, and of one additional adjoint function to evaluate the numerical value of the response. Once these adjoint functions have been computed, the sensitivities to *all* possible changes in the system parameters can be obtained by simple quadratures. For this illustrative reactor safety application, a total of five adjoint calculations sufficed to perform the complete sensitivity analysis of the maximum fuel temperature and maximum power responses. Note, though, that only the source terms in these five adjoint calculations differ from one another; the form of the respective partial differential equations and corresponding boundary and final-time conditions is the same for all of these adjoint systems. Consequently, all five adjoint functions can be calculated by using the same code; only minor programming is required to implement the numerical calculation of the distinct source terms for the adjoint equations.

Section III.C. has presented the theoretical formulation of an efficient sensitivity analysis method for estimating the effect of feedback in a mathematical model. A feedback operator acting on the model's dependent variables defines a feedback mechanism by modifying the values of parameters or data in the model. Although the effect of prescribed variations in the parameters can be evaluated efficiently using the Adjoint Method, this method cannot be applied directly to estimate the effect of feedback; this is because the parameter variations are not prescribed but depend on the output of this model. Therefore, we have defined a quantity called *sensitivity to feedback* that can be estimated using the Adjoint Method. It has been shown that the sensitivity to feedback is an estimate of the actual effect of feedback correct to first order in the strength of feedback, and it has also been shown how the sensitivity can be estimated using the Adjoint Method. The principal advantage of this application of the Adjoint Method is that, once the adjoint solution has been calculated, the effect of a variety of different feedbacks can be estimated with minimal additional computing time. The comprehensive sensitivity theory presented in Sec. III.C. for estimating the effect of feedback is likely to be useful for models where extensive recalculation with a variety of feedbacks is impractical. An approximate yet quantitative indication of the effects of a wide range of potentially important feedbacks will help identify sources of uncertainty in model predictions, and will indicate for incorporating feedbacks rigorously.

The theoretical advances which this work contributes to sensitivity theory were made possible by the use of concepts of nonlinear functional analysis. Nonetheless, the potential of using such concepts to extend further the scope of sensitivity theory warrants more research. Present research is divided between developing sensitivity theory and applying existing theory to new areas.

An important but presently open question regarding sensitivity theory is the effect of the higher-order G -differentials of the response. The possibility of using concepts of nonlinear functional analysis to estimate this effect is currently being researched. The incorporation of this effect into an uncertainty analysis formalism would result in a reliable and efficient tool for comprehensive sensitivity and uncertainty analyses of complex physical problems.

V. REFERENCES

1. S. Glasstone and M. C. Edlund, *The Elements of Nuclear Reactor Theory* (Van Nostrand, Princeton, New Jersey, 1952).
2. E. P. Wigner, "Effects of Small Perturbations on Pile Period," Chicago Report CP-G-3048 (1945).
3. W. M. Stacey, Jr., *Variational Methods in Nuclear Reactor Physics* (Academic, New York, 1974).
4. H. Levine and J. Schwinger, *Phys. Rev.* **75**, 1423 (1949).
5. R. Roussopolos, *C. R. Acad. Sci.*, **236**, 1858 (1953).
6. L. N. Usachev, *J. Nucl. Energy* **18**, 571 (1964).
7. G. C. Pomraning, *J. Math. Phys.* **8**, 149 (1967).
8. A. Gandini, *J. Nucl. Energy* **21**, 755 (1967).
9. J. Lewins, *Nucl. Sci. Eng.* **31**, 160 (1968).
10. W. M. Stacey, Jr., *J. Math. Phys.* **13**, 1119 (1972).
11. S. A. Gerstl and W. M. Stacey, Jr., *Nucl. Sci. Eng.* **51**, 339 (1973).
12. E. Greenspan, "Developments in Perturbation Theory," in *Advances in Nuclear Science and Technology*, Vol. 9 (Academic, New York, 1976).
13. M. Becker, *Nucl. Sci. Eng.* **62**, 296 (1977).
14. M. L. Williams, *Nucl. Sci. Eng.* **70**, 20 (1979).
15. E. M. Oblow, *Nucl. Sci. Eng.* **68**, 332 (1978).
16. E. Greenspan, *Nucl. Sci. Eng.* **74**, 185 (1980).
17. C. V. Parks and P. J. Maudlin, *Nucl. Technol.* **54**, 38 (1981).
18. A. Gandini, *Nucl. Sci. Eng.* **67**, 91 (1978); also "Corrigendum," *Nucl. Sci. Eng.* **70**, 112 (1979). See also: E. Greenspan, D. Gilai, and E. M. Oblow, *Nucl. Sci. Eng.* **68**, 1 (1978).
19. J. Lewins and M. Becker, Eds., *Advances in Nuclear Sciences and Technology*, Vol. 14 (Plenum, New York, 1981). See also, J. H. Marable, C. R. Weisbin, and G. De Saussure, *Nucl. Sci. Eng.* **75**, 30 (1980).
20. B. R. Sehgal, private communication (EPRI, 1978).
21. D. G. Cacuci, C. F. Weber, E. M. Oblow, and J. H. Marable, *Nucl. Sci. Eng.* **75**, 88 (1980).
22. M. M. Vainberg, *Variational Methods for the Study of Nonlinear Operators* (Holden-Day, San Francisco, 1964); see also, L. A. Liusternik and V. J. Sobolev, *Elements of Functional Analysis* (Ungar, New York, 1961).
23. E. K. Blum, "The Calculus of Variations, Functional Analysis, and Optimal Control Problems," in *Topics in Optimization*, Vol. 31, Mathematics in Science and Engineering Series, edited by G. Leitmann (Academic, New York, 1967).
24. T. L. Saaty, *Modern Nonlinear Equations* (McGraw-Hill, New York, 1967).
25. L. B. Rall, editor, *Nonlinear Functional Analysis and Applications* (Academic, New York, 1971); see also, M. S. Berger, *Nonlinearity and Functional Analysis* (Academic, New York, 1977).

26. V. Hutson and J. S. Pym, *Applications of Functional Analysis and Operator Theory* (Academic, New York, 1980).
27. R. Tomovic and M. Vukobratovic, *General Sensitivity Theory* (American Elsevier, New York, 1972), p. 3.
28. D. G. Cacuci, "Sensitivity Theory for Nonlinear Systems. I. Nonlinear Functional Analysis Approach," *J. Math. Phys.*, **22**, 2794 (1981), and II. Extensions to Additional Classes of Responses," *J. Math. Phys.* **22**, 2803 (1981).
29. The terminology for $V\mathbf{R}(e^0;h)$ as defined by Eq. (4) is not used uniformly in the literature. Some authors refer to $V\mathbf{R}$ as the Gâteaux variation of \mathbf{R} at e^0 , reserving the term "differential" for those instances when $V\mathbf{R}$ is linear in its second argument h . Others refer to $V\mathbf{R}$ as the weak differential of \mathbf{R} at e^0 , to distinguish it from the strong (or Fréchet) differential of \mathbf{R} at e^0 . However, the terminology employed in this study appears to occur most frequently in the latest works on nonlinear operators.
30. H. Sagan, *Introduction to the Calculus of Variations* (McGraw-Hill, New York, 1969).
31. J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables* (Academic, New York, 1970).
32. Although the terminology employed here seems to occur most frequently in the literature, attention is drawn to the fact that some authors use the term "Gâteaux derivative" to designate a G-differential that exists in a neighborhood of e^0 .
33. In principle, \hat{P} , could have been forced to vanish if appropriate extensions (in the operator sense) of $N'_u(e^0)$ had been considered. Of course, this would have necessitated the reformulation of Eqs. (12) and (13). In practice, however, it is more convenient to avoid such complications, since \hat{P} appears (ultimately) only as a readily computable quantity in the final expression the "indirect effect" term [cf., Eq. (22)].
34. G. I. Bell and S. Glasstone, *Nuclear Reactor Theory* (Van Nostrand Reinhold, New York, 1970), p. 291.
35. A. M. Weinberg and E. P. Wigner, *The Physical Theory of Neutron Chain Reactors* (University of Chicago Press, Chicago, 1958).
36. T. Kato, *Perturbation Theory for Linear Operators* (Springer, New York, 1966).
37. B. A. Finlayson, *The Method of Weighted Residuals and Variational Principles* (Academic, New York, 1972).

INTERNAL DISTRIBUTION

- | | |
|-------------------------|--------------------------------------|
| 1. R. G. Alsmiller, Jr. | 35. E. M. Oblow |
| 2. S. M. Babcock | 36. P. J. Otaduy |
| 3. J. Barhen | 37. R. W. Peelle |
| 4. C. Bottcher | 38. F. G. Perey |
| 5. B. L. Broadhead | 39. R. B. Perez |
| 10-14. D. G. Cacuci | 40. F. G. Pin |
| 11. B. A. Carreras | 41. H. Postma |
| 12. G. de Saussure | 42. V. Protopopescu |
| 13. F. C. Difilippo | 43. M. W. Rosenthal |
| 14. G. F. Flanagan | 44. R. T. Santoro |
| 15. R. E. Funderlic | 45. A. D. Solomon |
| 16. R. H. Gardner | 46. V. R. Uppuluri |
| 17. G. A. Geist | 47. R. C. Ward |
| 18. D. G. Gosslee | 48. C. R. Weisbin |
| 19. L. J. Gray | 49. A. Zucker |
| 20. P. M. Haas | 50. P. W. Dickson, Jr. (Consultant) |
| 21. W. R. Hamel | 51. G. H. Golub (Consultant) |
| 22. M. T. Heath | 52. R. M. Haralick (Consultant) |
| 23. H. R. Hicks | 53. D. Steiner (Consultant) |
| 24. D. T. Ingersoll | 54. EPMD Reports Office |
| 25-29. J. K. Ingersoll | 55. Central Research Library |
| 30. F. C. Maienschein | 56. K-25 Plant Library |
| 31. J. March-Leuba | 57. ORNL Patent Office |
| 32. B. F. Maskewitz | 58. Y-12 Technical Library |
| 33. T. J. Mitchell | Document Reference Section |
| 34. F. R. Mynatt | 59. Laboratory Records-RC |
| | 60-64. Laboratory Records Department |

EXTERNAL DISTRIBUTION

65. David L. Auton, Defense Nuclear Agency, 6801 Telegraph Rd., Alexandria, VA 22310-3398
66. Paul Baybutt, Battelle Memorial Institute, 505 King Ave., Columbus, OH 43201
67. M. Becker, School of Engineering, Office of the Dean, Rensselaer Polytechnic, Tibbets Ave., Troy, NY 12181
68. W. J. Conover, Department of Mathematics, Texas Technical University, Lubbock, TX 79409
69. J. C. Helton, Department of Mathematics, Arizona State University, Tempe, AZ 85287

70. R. L. Iman, Division 1223, Sandia Laboratories, Albuquerque, NM 87185
71. M. D. McKay, Statistics Group, Los Alamos National Laboratory, Los Alamos, NM 87545
72. D. Okrent, 5532 Boelter Hall, UCLA, Los Angeles, CA 90024
73. R. Pierls, Applied Mathematics Department, Brookhaven National Laboratory, Upton, NY 11973.
74. T. H. Pigford, Department of Nuclear Engineering, University of California, Berkeley, CA 94720
75. J. Vaurio, Bldg. 207, Argonne National Laboratory, 9700 South Cass Ave., Argonne, IL 60439
76. L. Zadeh, Division of Computer Sciences, University of California, Berkeley, CA 94720
- 77-78. Technical Information Center (TIC)
79. Office of Assistant Manager for Energy Research and Development, DOE-ORO, Oak Ridge, Tennessee 37831