

NUMERICAL COMPUTATION OF SENSITIVITIES AND THE ADJOINT APPROACH *

ROBERT MICHAEL LEWIS †

Abstract. We discuss the numerical computation of sensitivities via the adjoint approach in optimization problems governed by differential equations. We focus on the adjoint problem in its weak form. We show how one can avoid some of the problems with the adjoint approach, such as deriving suitable boundary conditions for the adjoint equation. We discuss the convergence of numerical approximations of the costate computed via the weak form of the adjoint problem and show the significance for the discrete adjoint problem.

1. Introduction. In this paper we discuss the numerical computation of sensitivities via the adjoint approach in optimization problems governed by differential equations. We focus on the weak form of the adjoint problem. The weak form of the adjoint problem allows one to finesse the issue of identifying the adjoint problem as a conventional boundary-value, initial-value, or initial-boundary-value problem. We also discuss how one can use the weak form of the adjoint problem to compute numerical approximations of the costate and relate this approach to the discrete adjoint problem.

The context for this discussion is the nonlinear programming problem

$$(1.1) \quad \begin{aligned} &\text{minimize} && F(a) = f(a, u(a)) \\ &\text{subject to} && C(a) = c(a, u(a)) \geq 0, \end{aligned}$$

where $u(a)$ is the solution of some set of differential equations,

$$(1.2) \quad S(a, u(a)) = 0.$$

We will refer to a as the design variable and u as the state variable. For simplicity, we restrict our attention to the question of computing derivatives associated with the objective $F(a)$ and ignore the constraints $C(a)$.

The adjoint approach [3, 11, 14] allows one to compute the derivative $F'(a)$ of $F(a)$ with respect to a in a very efficient manner. The primary cost in computing $F'(a)$ via the adjoint approach is the calculation of an intermediate quantity λ , called the costate or adjoint state, as the solution of the adjoint problem, which is a linear problem associated with the governing equation (1.2).

However, the adjoint approach is not without attendant difficulties. These difficulties are particularly associated with the calculation of the intermediate quantity λ . One problem that arises is that of determining the appropriate adjoint problem. It is not always the case that one can easily identify the adjoint problem as some manner of conventional initial-value, boundary-value, or initial-boundary-value problem. Objectives in optimization problems for which one cannot identify the appropriate adjoint problem in a straightforward way are sometimes called “inadmissible.”

*Revision: 1.6 Date: 1998/01/13 22:01:07

†ICASE, Mail Stop 403, NASA Langley Research Center, Hampton, Virginia 23681-0001, buckaroo@icase.edu. This research was supported by the National Aeronautics and Space Administration under NASA Contract No. NAS1-19480 while the author was in residence at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA 23681-0001.

The problem of inadmissibility is reported, for instance, in [2, 6], where the authors conclude that only certain objectives are admissible in aerodynamic optimization problems. The difficulty with the “inadmissible” cost functionals encountered by these authors stems from the absence of suitable boundary terms to cancel terms that appear due to the choice of cost functional. Another question associated with the adjoint approach is that of the convergence of numerical approximations of the costate λ . Finally, from the point of view of numerical optimization, there is the paramount question of the approximation of the derivative $F'(a)$ and the convergence of such approximations under refinement of the discretization of (1.1)–(1.2).

We discuss how studying the adjoint problem in its weak form gives us one way to address these issues. The weak form of the adjoint problem is based on the identity of the costate λ as a linear functional.

The weak form of the adjoint problem always exists, which allows us to resolve the problem of “inadmissibility.” Furthermore, we can solve the weak form of the adjoint problem numerically and can derive convergence estimates for λ that depend on the convergence estimates for the solution of the forward problem (1.2). While the approximations of λ may converge in a very weak sense, the ensuing convergence estimates for λ nevertheless enable us to derive useful convergence estimates for approximations of the derivative $F'(a)$. We also relate the numerical approximation of the weak form of the adjoint problem to the adjoint of the discretized problem.

2. Issues in the numerical calculation of sensitivities. We begin with a discussion of some of the issues in the numerical calculation of sensitivities for optimization problems governed by differential equations.

2.1. Consistency of sensitivity calculations. We assume that the discretization of u (and possibly a) is parameterized by h , where the discretization is refined as $h \rightarrow 0$. There are two senses of consistency for numerical approximations of sensitivities for the problem (1.1)–(1.2).

Consistency with the infinite-dimensional problem. As the discretization is refined, the approximation λ^h of the costate and the approximation $F'_h(a)$ of the derivative should converge in a suitable sense to the correct quantities. One would also hope for estimates of the rate of convergence.

Consistency with the finite-dimensional computational problem. At any level of discretization h , the approximation of the derivative $F'(a)$ should correspond to the derivative of the finite-dimensional optimization problem obtained by discretizing (1.1)–(1.2) at level h .

The necessity of the first sense of consistency is clear; if the numerical scheme is to make sense then it should converge under refinement of the discretization. The necessity of the second sense of consistency is subject to debate. One could, for instance, derive the adjoint problem and discretize it in a manner independent of the discretization of (1.1)–(1.2). Asymptotically the sensitivities derived by these two approaches will be consistent (provided both are consistent with the infinite-dimensional problem), which might lead one to conclude that consistency of derivatives with the discretized optimization problem is unnecessary. However, inconsistencies at the discrete level may cause an optimization algorithm applied to the discretized version of the optimization problem to terminate prematurely.

This is not so much an issue in the case of unconstrained optimization. For instance, consider the situation in Fig. 2.1, in which $a \in \mathbb{R}^2$ is the current value of the design variables, and $-g(a)$ is an approximation to the negative gradient $-\nabla F(a)$ at a . In the case of unconstrained minimization, $-g(a)$ is well within the cone of descent

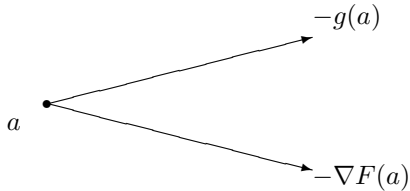


FIG. 2.1. *Acceptability of errors in the derivative in the unconstrained case.*

directions for F and is a reasonable approximation of $-\nabla F(a)$. The analysis in [5] of the use of inexact gradient information in trust region methods for unconstrained minimization indicates that optimization algorithms for unconstrained minimization are surprisingly insensitive to such errors.

However, the situation changes if we add even the simplest of constraints. Consider the previous situation with the addition of two bound constraints that are binding at a , as depicted in Fig. 2.2. In this picture the feasible region is denoted by Ω .

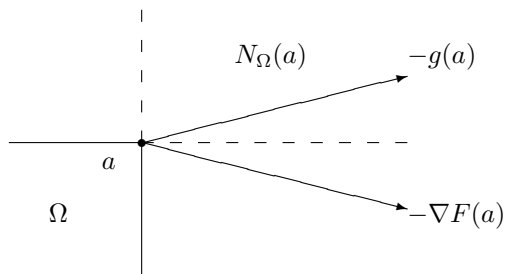


FIG. 2.2. *Unacceptability of errors in the derivative in the constrained case.*

Now the approximation $-g(a)$ lies in the normal cone of Ω at a , $N_{\Omega}(a)$. This would lead us to the erroneous conclusion that a is a Karush-Kuhn-Tucker point (i.e., a constrained stationary point), since it appears that at a we cannot improve F locally without violating the constraints. In the constrained problem, the cone of feasible descent directions is not determined just by $-\nabla F(a)$ but also by the constraint geometry, and in general this cone is much smaller than in the case of unconstrained minimization. As a consequence, in constrained optimization even small errors in the derivative can lead to problems.

Thus small errors in the gradient can lead to serious problems in the application of optimization algorithms. Consistency of the numerical approximations of sensitivities with the discretized optimization problem is one step to reducing the presence of such errors. From a practical point of view, moreover, consistency of derivatives with the discretized problem makes it easier to check the correctness of their implementation.

In the case the finite-dimensional computational problem, the desired sense of consistency is clear: sensitivities should be numerically consistent (i.e., consistent to the extent allowed by the limitations of machine precision). Consistency with the infinite-dimensional problem, that is, the sense in which λ^h and $F'_h(a)$ should converge to λ and $F'(a)$, is more subtle.

Both λ and $F'(a)$ are linear functionals; let $\langle \lambda, v \rangle$ denote the value obtained by

applying the linear functional λ to the vector v . One typically posits that the action of λ can be expressed as

$$(2.1) \quad \langle \lambda, v \rangle = \int \lambda(x)v(x)$$

where $\lambda(x)$ is a function sufficiently smooth to permit integration by parts against the linearization of the state equation (1.2). One then attempts to identify λ as the solution of an adjoint differential equation.

If this can be done, one can then apply the convergence analysis of the schemes for the solution of this differential equation to obtain estimates of the convergence of an approximation λ^h to λ . However, there is a problem with this approach. An arbitrary numerical scheme for discretizing the adjoint equation will not necessarily be numerically consistent with the discretized problem. We will return to this point in §4.1.

2.2. Deriving the adjoint problem. Moreover, it is not always the case that we can write λ in the form (2.1). A related problem is that of “inadmissible” objectives, mentioned in the introduction, where one cannot identify λ as the solution of a differential equation. The following example illustrates both points. For $a \in \mathbb{R}$, let

$$\begin{aligned} u'(x) &= a & x \in (0, 1) \\ u(0) &= 0 \end{aligned}$$

and define $F(a) = u'(1)$; note that $F(a) = a$. If we try to apply the adjoint method in the usual way to compute $F'(a)$, beginning with (2.1) and integrating by parts, we obtain

$$\int_0^1 \lambda(x)u'(x) = \lambda(1)u(1) - \int_0^1 \lambda'(x)u(x) = u'(1).$$

for all $u \in C^1[0, 1]$, say. Now we are stuck; if we require λ to satisfy the adjoint ODE $-\lambda'(x) = 0$ for $x \in (0, 1)$ we obtain the requirement

$$\lambda(1)u(1) = u'(1),$$

for all $u \in C^1[0, 1]$ and no choice for the terminal condition $\lambda(1)$ will insure this. The problem is that $\lambda = \delta'(1)$ and this is supported at $x = 1$, where we also need to impose a terminal condition for λ , which makes λ over-determined at $x = 1$. Also note that λ is not a function, so the assumption (2.1) is not, strictly speaking, justified. On the other hand, the objective $F(a) = a$ is unequivocally a differentiable function of a single real variable, so any obstruction must be illusory.

2.3. Representation of linear functionals and directions of steepest descent. Related to the question of convergence of numerical approximations to λ and $F'(a)$ is the question of just what such approximations are converging to. In a very real sense, linear functionals such as $F'(a)$ and λ have no intrinsic identity other than their values when applied to their arguments.

Consider, for instance, the familiar δ -functional defined on the Sobolev space $H^1[-1, 1]$ by $\langle \delta, u \rangle = u(0)$. This functional is bounded on $H^1[-1, 1]$. We can represent the action of δ in many different ways. For instance, we can write δ as the limit of the sequence

$$(2.2) \quad \delta_k(x) = k\chi_k(x),$$

where χ_k is the characteristic function of the interval $[-1/2k, 1/2k]$:

$$\langle \delta, u \rangle = u(0) = \lim_{k \rightarrow \infty} \int_{-1}^1 \delta_k(x) u(x) dx.$$

We can also represent δ as integration against a measure μ defined by

$$\mu(\Omega) = \begin{cases} 0 & x \notin \Omega \\ 1 & x \in \Omega \end{cases}.$$

Then

$$\langle \delta, u \rangle = \int_{-1}^1 u(x) d\mu(x).$$

Furthermore, since δ is a bounded linear functional on $H^1[-1, 1]$, the Riesz representation theorem [9] guarantees the existence of $w \in H^1[-1, 1]$ such that in the H^1 inner product $(\cdot, \cdot)_{H^1}$,

$$\langle \delta, u \rangle = (w, u)_{H^1[-1, 1]} = \int_{-1}^1 w(x)u(x) + w'(x)u'(x).$$

In this case one can check that

$$(2.3) \quad w(x) = \begin{cases} -\frac{e^{-x} - e^2 e^x}{2(1 + e^2)} & -1 \leq x \leq 0 \\ -\frac{e^x - e^2 e^{-x}}{2(1 + e^2)} & 0 \leq x \leq 1 \end{cases}.$$

Thus we see that δ has no intrinsic identity save $\langle \delta, u \rangle = u(0)$. While δ can be represented by the reasonable function $w(x)$ in (2.3), δ can also be viewed as the limit of the sequence in (2.2), which is not convergent in L^2 and pointwise is converging to zero almost everywhere.

Given that it is not clear what it means to “look at” λ and $F'(a)$ as linear functionals, how then are these quantities to be interpreted? The notion of a direction of steepest provides one answer to this question.

The direction of steepest descent for F with respect to the norm $\|\cdot\|_X$ is a solution of

$$(2.4) \quad \begin{aligned} & \text{minimize} && \langle F'(a), p \rangle \\ & \text{subject to} && \|p\|_X \leq 1, \end{aligned}$$

provided a solution exists. Note that this direction depends on the choice of norm; the direction of steepest descent tells us the direction that infinitesimally yields the greatest decrease in F *per unit distance*, and distance depends on the choice of norm. Since the direction of steepest descent is a direction in X , it gives us something meaningful to examine (e.g., actually plot). We would replace $F'(a)$ by λ in (2.4) if λ were the functional of interest.

Since directions of steepest descent are directions in the design space, they can be used in optimization algorithms. For instance, in the usual adjoint approach, one hopes to obtain a representation of the form

$$\langle F'(a), p \rangle = \int gp$$

for $g \in L^2$. The Cauchy-Schwarz inequality then says that g defines the direction of steepest descent with respect to the L^2 norm.

On the other hand, there is no L^2 direction of steepest descent for the δ -functional because δ is not bounded on L^2 . At the same time, the representer w of δ in the H^1 inner product given in (2.3) defines (again by the Cauchy-Schwarz inequality) the direction of steepest descent for δ in the H^1 norm. The computation of directions of steepest descent in the H^k norm is discussed in [10].

2.4. Further remarks on the representation of the costate. Actually, the assumption that λ can be written in the form (2.1) is, in general, almost true in a sense we can make precise. This observation depends on the density of functionals of the form (2.1) in the negative norm Sobolev spaces [1, 13] and other spaces. We will use this observation in §4.1 in connection with numerical schemes for computing λ .

The Riesz representation theorem allows us to identify H^k with its dual (i.e., the space of bounded linear functionals on H^k): given any linear functional ℓ on H^k we can find $w \in H^k$ for which $\langle \ell, u \rangle = (w, u)_{H^k}$ for all $u \in H^k$. The negative norm space H^{-k} is an alternative representation of the dual of H^k . For $w \in L^2$, the negative Sobolev norm is defined to be

$$\|w\|_{H^{-k}} = \sup_{\|u\|_{H^k}=1} \left| \int w(x)u(x) \right|.$$

The negative norm Sobolev space H^{-k} is defined to be the completion of L^2 with respect to this norm.

This means that functionals of the form (2.1) are dense in H^{-k} , so if $\lambda \in H^{-k}$ then λ can be approximated arbitrarily well (in H^{-k} norm) by functionals of the form (2.1). Functionals of the form (2.1) are dense in many other dual spaces, as well, such as $(C^k)'$. Thus we should not be surprised that the assumption (2.1) turns out to be correct as often as it does. On the other hand, the examples of the previous sections show that λ need not have such a representation.

3. The weak form of the adjoint problem. We now turn from these general questions concerning the calculation of sensitivities to the adjoint problem in its weak form. We give two derivations of the weak form of the adjoint problem. The first is based on implicit differentiation, while the second is based on the variation of the Lagrangian.

We assume that $a \in X$, $u \in U$, and $S : (a, u) \in X \times U \mapsto S(a, u) \in V$, where X, U, V are Banach spaces. We also assume that f is differentiable on $X \times U$, and that given (a, u) for which $S(a, u) = 0$ (i.e., u solves (1.2)), the partial Fréchet derivative of S with respect to u , $S_u(a, u)$, is boundedly invertible as a map from U to V .

We denote by X', U' , and V' the dual spaces of X, U, V (i.e., the spaces of bounded linear functionals on X, U, V). We denote the adjoint [9] of a bounded linear operator A by A^\times ; if $A : Y \rightarrow Z$, then $A^\times : Z' \rightarrow Y'$ is defined by $\langle A^\times z', y \rangle_{Y'} = \langle z', Ay \rangle_Z$ for all $y \in Y$.

3.1. Derivation via implicit differentiation. Under the preceding hypotheses, the classical implicit function theorem [8] assures us that $S(a, u(a)) = 0$ defines $u(a)$ as a smooth function of a and that the derivative of $u(a)$ with respect to a is

$$(3.1) \quad \frac{du}{da}(a) = -S_u^{-1}S_a \Big|_{(a, u(a))}.$$

This formula for the Jacobian of u with respect to a is formally just the result of applying implicit differentiation to $S(a, u(a)) = 0$. The derivative of F is then

$$(3.2) \quad F'(a) = f_a + f_u \frac{du}{da} \Big|_{(a, u(a))} = f_a - f_u S_u^{-1} S_a \Big|_{(a, u(a))}.$$

At this point the costate $\lambda \in V'$ appears. Define

$$(3.3) \quad \lambda = -f_u S_u^{-1};$$

then $F'(a) = f_a + \lambda S_a$, where the right-hand side is evaluated at $(a, u(a))$.

Next we derive the weak form of the adjoint problem and make clear the role adjointness plays. From (3.3), $\lambda \in V'$ satisfies

$$(3.4) \quad \langle \lambda S_u, \nu \rangle_U = -\langle f_u, \nu \rangle_U,$$

for all $\nu \in U$. However, by definition

$$(3.5) \quad \langle \lambda S_u, \nu \rangle_U = \langle \lambda, S_u \nu \rangle_V = \langle S_u^\times \lambda, \nu \rangle_U,$$

so

$$(3.6) \quad \lambda = -(S_u)^{-\times} f_u.$$

Note that $S_u^{-\times} : f_u \in U' \rightarrow (S_u)^{-\times} f_u \in V'$.

From (3.4) and (3.5) we obtain

$$(3.7) \quad \langle \lambda, S_u \nu \rangle_V = -\langle f_u, \nu \rangle_U$$

for all $\nu \in U$. This relation we call the *weak form* of the adjoint problem since it defines λ only through the action of λ as a linear functional on vectors in its domain.

The weak form of the adjoint problem always exists and is a distributional differential equation. Any boundary conditions are implicit. Such equations are discussed in [12], for instance. There remains the question of relating the weak form of the adjoint problem to a conventional differential equation. As we have noted before, usually one can identify an adjoint differential equation (and boundary conditions) that is equivalent to (3.7); however, we have also seen that this not always the case. The interpretation of distributional differential equations in terms of conventional differential equations arises more generally in the solution of differential equations with distributional data and is not always possible [12]. On the other hand, the weak form of the adjoint problem always exists, avoiding the problem of ‘‘inadmissibility’’ mentioned previously.

We can relate these formulae to $F'(a)$, beginning with (3.2). Given $\eta \in X$,

$$\begin{aligned} \langle F'(a), \eta \rangle_X &= \left\langle f_a + f_u \frac{du}{da}, \eta \right\rangle_X \\ &= \langle 1, f_a \eta \rangle_{\mathbb{R}} + \left\langle f_u, \frac{du}{da} \eta \right\rangle_U = \langle f_a^\times 1, \eta \rangle_X + \left\langle \frac{du}{da}^\times f_u, \eta \right\rangle_X. \end{aligned}$$

or just

$$(3.8) \quad F'(a) = f_a^\times 1 + \frac{du}{da}^\times f_u.$$

From (3.1) we obtain

$$(3.9) \quad \frac{du^\times}{da} f_u = -S_a^\times (S_u)^{-\times} f_u.$$

Then (3.6), (3.8), and (3.9) yield

$$(3.10) \quad F'(a) = f_a^\times 1 - S_a^\times (S_u)^{-\times} f_u = f_a^\times 1 + S_a^\times \lambda.$$

In the finite-dimensional case the first identity in (3.10) corresponds to transposing (3.2) to obtain $\nabla F(a) = \nabla_a f - S_a^T S_u^{-T} \nabla_u f$.

3.2. Derivation via variation of the Lagrangian. We can also derive the weak form of the adjoint problem by examining the variation of the Lagrangian due to variations in a . This approach, another form of the calculations of the preceding section, begins with the Lagrangian $L(a, u; \lambda) = f(a, u) + \langle \lambda, S(a, u) \rangle$. One typically assumes that the action of the linear functional λ can be represented as an integral of the form (2.1), but as discussed in §2 this may not be justified. For that reason we adhere to the abstract representation $\langle \lambda, S(a, u) \rangle$ for the action of λ on the equation residual $S(a, u)$.

The effect δL of a variation δa on the Lagrangian L is given by

$$\delta L = \delta f + \langle \lambda, \delta S \rangle = f_a \delta a + f_u \delta u + \langle \lambda, S_a \delta a + S_u \delta u \rangle,$$

where to first-order the perturbation δu in u due to the perturbation δa satisfies $S_a \delta a + S_u \delta u = 0$. Rearranging terms we have

$$\delta L = f_a \delta a + \langle \lambda, S_a \delta a \rangle + f_u \delta u + \langle \lambda, S_u \delta u \rangle.$$

If we require λ to satisfy

$$(3.11) \quad f_u \delta u + \langle \lambda, S_u \delta u \rangle = 0$$

for all perturbations δu , then $F'(a) \delta a = f_a \delta a + \langle \lambda, S_a \delta a \rangle$ for all δa , which determines $F'(a)$.

Typically one insures that (3.11) holds for all δu by starting with the representation (2.1) for the action of λ and applying integration by parts to derive an adjoint problem independent of δu , which, if solved by λ , implies that (3.11) holds for all δu . However, as noted previously the assumption that the action of λ can be written in the form (2.1) for some suitable function $\lambda(x)$ is not always justified. This explains some of the difficulties with “inadmissible” objectives that have been reported—the costate λ in those problems does not have the hypothesized form (2.1), and the adjoint method reaches an impasse.

On the other hand, if we take as the definition of the adjoint problem the weak formulation (3.11) (which is equivalent to (3.7)), we can avoid these problems since we do not presume *a priori* any particular representation for the action of λ . Moreover, this definition via duality sheds light on the convergence of numerical schemes for the approximation of λ and $F'(a)$, as we discuss in the next section.

4. Convergence of numerical schemes for the solution of the weak form of the adjoint equation. First we discuss how suitable numerical solutions λ^h of the weak form of the adjoint problem converge to λ , establishing consistency with the infinite-dimensional problem. We then give a concrete example of such a scheme that is also consistent with the finite-dimensional optimization problem.

Let $L = S_u$. We assume that for any v we can solve the linearized problem $Lw = v$ and that we approximate solutions w by $w^h \in U^h$, where h is a discretization parameter and U^h is an approximating space.

In the discretization of the adjoint problem, we require that λ^h satisfy

$$\langle \lambda^h, Lw^h \rangle = \langle f_u^h, w^h \rangle$$

for all $w^h \in U^h$. This defines the action of λ^h on the range V^h of U^h under L . There remains the definition of λ^h on the complement of V^h , which we do as follows. For $v \in V$, let w^h be the approximate solution of $Lw = v$ at the discretization level h . Then we require

$$\langle \lambda^h, v \rangle = \langle f_u^h, w^h \rangle.$$

This agrees with the preceding definition of λ^h on V^h . Moreover,

$$\langle \lambda^h, Lw \rangle = \langle f_u^h, w^h \rangle = \langle \lambda^h, Lw^h \rangle,$$

so we have the adjoint Galerkin condition:

$$(4.1) \quad \langle \lambda^h, Lw - Lw^h \rangle = 0$$

for all w, w^h .

Given $v \in V$, we can find $w \in U$ for which $Lw = v$. Then

$$\begin{aligned} \langle \lambda - \lambda^h, v \rangle &= \langle \lambda - \lambda^h, Lw \rangle \\ &= \langle \lambda, Lw - Lw^h \rangle + \langle \lambda - \lambda^h, Lw^h \rangle + \langle \lambda^h, Lw^h - Lw \rangle \\ &= \langle f_u, w - w^h \rangle + \langle f_u - f_u^h, w^h \rangle + \langle \lambda^h, Lw^h - Lw \rangle. \end{aligned}$$

Applying (4.1) reduces this to

$$(4.2) \quad \langle \lambda - \lambda^h, v \rangle = \langle f_u, w - w^h \rangle + \langle f_u - f_u^h, w^h \rangle.$$

This is the basic identity for approximating the error $\lambda - \lambda^h$; we will give examples of its use in §5.

From this discussion we see that the error in the approximation of λ is governed by how well we approximate f_u and by the rate of convergence of w^h to w for solutions of $Lw = v$. Approximation of f_u is not trivial: unless f is linear in u , f_u will involve the solution of the forward problem (1.2). We will also see in §5 that the sense in which w^h must approximate w need not be as strong as the norm on U .

These estimates guarantee convergence in what might be only a very weak norm, such as a negative Sobolev norm; this sense of convergence may be odd. For instance, $\cos kx \rightarrow 0$ in $H^{-1}[0, 2\pi]$ as $k \rightarrow \infty$. The possibility of convergence of numerical approximations of λ in a weak sense is not a purely theoretical possibility, either; an example of this phenomenon in optimal control is discussed in [4]. Nevertheless, we can still use the weak form of the adjoint problem to obtain meaningful approximations of $F'(a)$.

4.1. The convergence of the discrete adjoint approach. The discrete adjoint problem is one practical instance of the abstract approach outlined in the preceding section. Furthermore, the discrete adjoint problem is consistent with the discretized forward problem.

For this discussion we assume that we are approximating the forward problem (and its linearization) by a Galerkin finite element method. Extension to the more general Petrov-Galerkin methods is straightforward.

Let $\{\phi_1^h, \dots, \phi_N^h\}$ be the basis for the finite element space U^h , where $N = N(h)$. We assume that the linearized operator $L = S_u$ is a map $L : H^m \rightarrow H^k$, $k \geq 0$, and that f_u is bounded on H^m . That is, $U = H^k$, $U' = H^{-k}$, $V = H^m$, and $V' = H^{-m}$. The case $k < 0$ can be handled along the lines we discuss here, but requires a slightly different argument. For brevity we omit this discussion.

From the discussion in §2.4, we know that we can approximate λ (in H^{-k} norm) by functionals of the form

$$\langle \lambda^h, v \rangle = \int \lambda^h(x)v(x),$$

where $\lambda^h(x) \in L^2$. In particular, we can approximate λ by functionals λ^h of this form for which $\lambda^h(x) \in U^h$:

$$\lambda^h(x) = \sum_{i=1}^N \lambda_i^h \phi_i^h(x).$$

If we do so, then (3.7) becomes

$$(4.3) \quad \sum_{i=1}^N \sum_{j=1}^N \lambda_i^h w_i^h \int \phi_i^h(x) L \phi_j^h(x) = \sum_{j=1}^N w_j^h \langle f_u, \phi_j^h \rangle.$$

The standard matrix associated with the finite element method has appeared. Let

$$a_{ij} = \int \phi_i^h(x) L \phi_j^h(x).$$

Then the matrix $A = (a_{ij})$ is the stiffness matrix from the finite element method. If we define

$$\begin{aligned} \tilde{\lambda}_h &= (\lambda_1^h, \dots, \lambda_N^h)^T \\ \tilde{w}_h &= (w_1^h, \dots, w_N^h)^T \\ \tilde{f}_u &= (\langle f_u^h, \phi_1 \rangle, \dots, \langle f_u^h, \phi_N \rangle)^T, \end{aligned}$$

we can rewrite (4.3) as

$$(4.4) \quad \tilde{\lambda}_h^T A \tilde{w}_h = \tilde{f}_u^T \tilde{w}_h.$$

If (4.4) is to hold for all \tilde{w}_h , we must have

$$(4.5) \quad A^T \tilde{\lambda}_h = \tilde{f}_u.$$

This we recognize as the discrete adjoint, where we take the transpose (in the usual sense of linear algebra) of the linearized forward equation. This is the system that would result if we were to treat the finite element basis coefficients as the state variables in the optimization problem at level h .

The discrete adjoint is thus consistent with both the finite-dimensional discretized problem and the infinite-dimensional problem. However, we stress that this latter consistency is in the sense of convergence as linear functionals. If λ can be represented as

the solution of a differential equation, there is no guarantee that the finite-dimensional transpose of the scheme for the solution of the linearized forward problem will be consistent (in the usual sense of finite-difference and finite element methods) with the adjoint differential equation.

This stronger sense of consistency is known not to hold in general; [16] gives a counterexample. There the authors examine upwind differencing and show how transposition of the discretization of the linearized forward problem yields a numerical scheme for the adjoint problem that does not necessarily inherit either the consistency or order of accuracy (in the usual sense of finite-difference schemes) of the finite-difference scheme for the linearized forward problem. Given the correspondence of upwind differencing and the lowest-order discontinuous Galerkin methods [7] one may be able to adapt this counterexample to the finite element setting as well.

The example in [16] is not at odds with our discussion. We are interested in consistency (i.e., convergence) of λ^h and λ in the sense of linear functionals. As we have discussed, this may be weaker than the convergence of λ^h to λ as functions (e.g., pointwise convergence or convergence in L^2). It is this latter sense of consistency that is discussed in [16].

4.2. Consequences for the calculation of sensitivities. In this section we discuss the consequences of the preceding analysis for the calculation of $F'(a)$. From (3.10) we have

$$F'(a) - F'_h(a) = (f_a - f_a^h)^\times 1 + S_a^\times (\lambda - \lambda^h) + (S_a - S_a^h)^\times \lambda^h.$$

Since $\|T^\times\| = \|T\|$ for any bounded linear operator T ,

$$\|F'(a) - F'_h(a)\| \leq \|f_a - f_a^h\| + \|S_a\| \|\lambda - \lambda^h\| + \|S_a - S_a^h\| \|\lambda^h\|.$$

Estimates on $\|\lambda - \lambda^h\|$ and $\|\lambda^h\|$ then apply as part of the bound on the error in approximating $F'(a)$. However, the approximation of $F'(a)$ also depends on how well we approximate f_a and S_a , so the approximation of λ may be only part of the story of approximating $F'(a)$.

Also note that the error $\lambda - \lambda^h$ appears in the term $S_a^\times (\lambda - \lambda^h)$ in $F'(a) - F'_h(a)$, and

$$\langle S_a^\times (\lambda - \lambda^h), \eta \rangle = \langle \lambda - \lambda^h, S_a \eta \rangle = \langle f_u - f_u^h, L^{-1} S_a \eta \rangle.$$

Only the error in λ restricted to the range $R = S_a X \subset V$ plays a role in the approximation of $F'(a)$. A bound on $\|\lambda - \lambda^h\|$ would consider the error over all of V . It might be possible to use this observation suggests it to obtain sharper estimates on the error $F'(a) - F'_h(a)$ in certain situations (say, if $L^{-1}R$ were to consist of very smooth functions).

5. Illustration. We close our discussion with the following illustrative example. Let Ω be the square $\{(x_1, x_2) \mid -1 \leq x_1, x_2 \leq 1\}$ and consider the problem

$$(5.1) \quad \begin{aligned} -\Delta u &= a && \text{in } \Omega \\ u &= 0 && \text{on } \partial\Omega. \end{aligned}$$

The design variable is $a = a(x)$. The objectives we consider are convex quadratic functions of a , so the optimization problems are trivial. However, these objective serve to illustrate the points we wish to make about computing the costate and derivative of the objective.

Since (5.1) is linear in u , the linearization $L = S_u$ has the same form: $Lw = v$ is given by

$$(5.2) \quad \begin{aligned} -\Delta w &= v && \text{in } \Omega \\ w &= 0 && \text{on } \partial\Omega. \end{aligned}$$

We also have $S_a \eta = -\eta$ or $S_a = -I$.

We will approximate λ using the discrete adjoint approach, as discussed in 4.1. Applying piecewise linear elements to (5.1) and (5.2) and assuming a reasonable mesh, we have the estimate

$$\|u - u^h\|_{L^2} \leq Ch^2 \|u\|_{H^2}$$

and a similar estimate for w, w^h .

The first objective we consider is

$$F(a) = \frac{1}{2} \int_{\Omega} (u - \psi)^2,$$

where ψ is a datum we wish to match. We will assume that $a \in L^2$; then $u \in H^2$ and $S, L : H^2 \rightarrow L^2$. Consequently, we expect $\lambda \in L^2$.

We have

$$\langle f_u, w - w^h \rangle = \int_{\Omega} (u - \psi) (w - w^h)$$

from which we obtain

$$|\langle f_u, w - w^h \rangle| \leq \|u - \psi\|_{L^2} \|w - w^h\|_{L^2} \leq Ch^2 \|v\|_{L^2}.$$

Observe that convergence of w^h to w in L^2 suffices and we do not need convergence in H^1 or H^2 .

Meanwhile, assuming $\psi \in U^h$ for simplicity, we have

$$\langle f_u - f_u^h, w^h \rangle = \int_{\Omega} (u - u^h) w^h,$$

so

$$|\langle f_u - f_u^h, w^h \rangle| \leq \|u - u^h\|_{L^2} \|w^h\|_{L^2} \leq Ch^2 \|v\|_{L^2}.$$

From these estimates and (4.2) we arrive at

$$\|\lambda - \lambda^h\|_{L^2} \leq Ch^2.$$

We obtained this result via a general argument using the weak form of the adjoint problem. We could also have arrived at this estimate by deriving the adjoint problem explicitly and examining its solutions. On the other hand, the objective

$$F(a) = \frac{1}{2} |\nabla u(0)|^2$$

is more easily handled by the weak form of the adjoint problem. Objectives of this sort arise when trying to match data measured only at points.

In this case we assume $a \in C^2(\overline{\Omega})$. Then $u \in H^4$ and $F(a)$ is defined. Continuing with piecewise linear finite elements, for any $\varepsilon \in (0, 1)$ we have [15] for some C_ε depending on ε ,

$$\| \nabla u - \nabla u^h \|_{L^\infty} \leq C_\varepsilon h^\varepsilon \| a \|_{L^\infty}$$

$$\| \nabla w - \nabla w^h \|_{L^\infty} \leq C_\varepsilon h^\varepsilon \| v \|_{L^\infty} \leq C_\varepsilon h^\varepsilon \| v \|_{H^2}.$$

Taking care to keep $x = 0$ interior to an element, we have

$$\begin{aligned} | \langle f_u - f_u^h, w \rangle | &= | (\nabla u(0) - \nabla u^h(0))^T \nabla w(0) | \\ &\leq C \| \nabla u - \nabla u^h \|_{L^\infty} \| w \|_{H^2} \leq C_\varepsilon h^\varepsilon \| a \|_{L^\infty} \| v \|_{L^2} \end{aligned}$$

and

$$\begin{aligned} | \langle f_u, w - w^h \rangle | &= | \nabla u(0)^T (\nabla w(0) - \nabla w^h(0)) | \\ &\leq \| \nabla u \|_{L^\infty} \| \nabla w - \nabla w^h \|_{L^\infty} \leq C_{\varepsilon, u} h^\varepsilon \| v \|_{H^2}. \end{aligned}$$

From these bounds and (4.2) we arrive at

$$\| \lambda - \lambda^h \|_{H^{-2}} \leq C_\varepsilon h^\varepsilon.$$

This holds despite the fact that λ is not even in L^2 : formally,

$$\begin{aligned} -\Delta \lambda &= (\partial_{x_1} u(0) \partial_{x_1} + \partial_{x_2} u(0) \partial_{x_2}) \delta && \text{in } \Omega \\ \lambda &= 0 && \text{on } \partial\Omega. \end{aligned}$$

We can write λ in terms of the fundamental solution for the Laplacian:

$$(5.3) \quad \lambda(x) = \frac{1}{2\pi} (\partial_{x_1} u(0) \partial_{x_1} + \partial_{x_2} u(0) \partial_{x_2}) \log |x| + r(x),$$

where r is smooth. The singular leading term is not in L^2 ; one can check that $\lambda \in H^{-1} \setminus L^2$.

Finally, we can estimate the error in $F'(a)$. Note that $f_a = 0$ and $S_a = S_a^h = -I$, so

$$\| F'(a) - F'_h(a) \|_{(C^2)'} = \| \lambda - \lambda^h \|_{H^{-2}} \leq C_\varepsilon h^\varepsilon.$$

Since $F'(a) = \lambda \notin C^2$ we cannot take (5.3) as defining a gradient or descent direction for the optimization problem. One would need to compute a direction of steepest descent with respect to a suitable norm as discussed in §2.3; see [10] for a further discussion of this point.

6. Conclusion. We have discussed some of the important analytical and computational issues in numerical approaches to the adjoint problem with an emphasis on the systematic approach to the adjoint problem via its weak form and its numerical solution. In the process we have shown how solutions of the discrete adjoint problem are suitably consistent with both the finite-dimensional and infinite-dimensional optimization problems. From numerical approximations of $F'(a)$ we can compute the directions of steepest descent and quasi-Newton directions needed for nonlinear programming algorithms.

One topic that deserves closer attention is the case where the governing equations are solved using finite-difference and finite-volume methods. These cases are of particular interest since they are most effective when the state u is smooth (except possibly for well-defined loci of discontinuity). The smoother the class of possible u the larger the class of possible cost functionals, and thus the broader the possibilities for λ and $F'(a)$.

A related question is the effect of numerical errors due to quadrature and other approximations. Depending on the locality of the interesting features of λ , these errors (and other “variational crimes” discussed in [17]) may be significant. The interpretation of some finite-difference and finite-volume methods in terms of Galerkin and Petrov-Galerkin finite element schemes with particular quadrature rules may allow one to relate results for finite elements to finite-difference and finite-volume methods (and vice versa).

7. Acknowledgments. The author wishes to thank David Keyes for answering many questions concerning finite element methods.

REFERENCES

- [1] R. A. ADAMS, *Sobolev Spaces*, Academic Press, 1975.
- [2] W. K. ANDERSON AND V. VENKATKRISHNAN, *Aerodynamic design optimization on unstructured grids with a continuous adjoint formulation*. AIAA-97-0643, 1997. Presented at the 35th Aerospace Sciences Meeting and Exhibit, Reno, Nevada, January 6–10, 1997.
- [3] H. T. BANKS AND K. KUNISCH, *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser, 1989.
- [4] J. BURNS, K. ITO, AND G. PROPST, *On nonconvergence of adjoint semigroups for control systems with delays*, SIAM Journal on Control and Optimization, 26 (1988), pp. 1442–1454.
- [5] R. G. CARTER, *On the global convergence of trust region algorithms using inexact gradient information*, SIAM Journal on Numerical Analysis, 28 (1991), pp. 251–265.
- [6] A. JAMESON, N. A. PIERCE, AND L. MARTINELLI, *Optimum aerodynamic design using the Navier-Stokes equations*. AIAA Paper 97-0101, 1997. Presented at the 35th Aerospace Sciences Meeting and Exhibit, Reno, Nevada, January 6–10, 1997.
- [7] C. JOHNSON, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Cambridge University Press, 1987.
- [8] L. V. KANTOROVICH AND G. P. AKILOV, *Functional Analysis in Normed Spaces*, MacMillan, 1964.
- [9] E. KREYSZIG, *Introductory Functional Analysis with Applications*, John Wiley & Sons, 1978.
- [10] R. M. LEWIS, *A nonlinear programming perspective on sensitivity calculations for systems governed by state equations*, Tech. Report 97-12, ICASE, NASA Langley Research Center, Hampton, Virginia, 23681-0001, February 1997.
- [11] J. L. LIONS, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, 1971.
- [12] J. L. LIONS AND E. MAGENES, *Non-homogeneous Boundary Value Problems and Applications*, Springer-Verlag, 1972.
- [13] J. T. ODEN AND J. N. REDDY, *An Introduction to the Mathematical Theory of Finite Elements*, John Wiley & Sons, 1976.
- [14] O. PIRONNEAU, *Optimal shape design for elliptic systems*, Springer-Verlag, 1984.
- [15] R. RANNACHER, *Zur L^∞ -Konvergenz linearer finiter Elemente beim Dirichlet-Problem*, Mathematische Zeitschrift, 149 (1976), pp. 69–77.
- [16] A. SEI AND W. W. SYMES, *A note on consistency and adjointness for numerical schemes*, Tech. Report TR95-04, Department of Computational and Applied Mathematics, Rice University, January 1995. Submitted to Mathematical and Computer Modelling.
- [17] G. STRANG AND G. J. FIX, *An Analysis of the Finite Element Method*, Prentice Hall, 1973.