Chapter 5

STATE SENSITIVITIES WITH RESPECT TO A PARAMETER

5.1 Introduction

As long as we confine ourselves to the feasible set, each set of values for the parameters defined in Chapter 2 defines a corresponding fluid flow problem. From Chapter 3, we may expect that there is a corresponding solution of the problem, and from Chapter 4 we know that, for the same set of parameters, there is also a discretized flow problem whose solution may be computed, and which approximates the solution of the continuous problem. These facts are summarized by stating that, for each set of feasible parameter values, β, there is (at least) one corresponding continuous flow solution (u, v, p), and discretized flow solution (u^h, v^h, p^h).

We would like, if possible, to regard these flow solutions as actual *functions* of the parameters. Despite the fact that we know that solutions to the continuous flow problem are not unique, we will find that, away from certain critical Reynolds numbers, the flow solutions form continuous branches of solutions, which are "locally unique".

Once we regard the flow solutions as *functions* of the parameters, it is natural to investigate the influence of each problem parameter on the components of the solution, a concept called the *sensitivity* of the solution with respect to the parameter.

We show how the state equations that define the state variables may be differentiated to yield equations for the sensitivity of the state variables with respect to a given problem parameter. This operation may be carried out on continuous or discrete state equations.

Since discrete state equations are generally derived from a continuous state equation, we might consider applying the same discretization process to the sensitivities of the continuous equation, and consider how the resulting *discretized sensitivities* are related to the sensitivities of the discrete state equation.

We work through a simple linear differential equation with a parameter, computing the sensitivities and discretized sensitivities, and showing that in this case discretization and differentiation may be interchanged; that is, they *commute*. These operations are not actually guaranteed to commute in all cases, and we briefly mention the case of parameters that affect the problem geometry. The consequences of non-commutativity are discussed in later chapters.

5.2 Questions About the Solutions of Parameterized Problems

In many fluid problems, there are a number of physical quantities which influence the problem, and hence the solution. These quantities might include the viscosity, the strength of the inflow, the temperature, the roughness of the walls, and so on. Especially when the item can be described by a few numerical values, we call such a varying quantity a flow parameter. Some flow parameters might be under the explicit control of the experimenter, while some might represent physical conditions which can't be controlled, but may be measured beforehand.

A special class of flow parameters is called *geometric* or *shape parameters.* Such parameters affect the shape of the flow region, or the location or shape of obstacles in the flow, or the location at which boundary conditions are imposed. The bump parameters that we discussed in Chapter 2 are an example of geometric parameters. As we proceed to the optimization problem we are ultimately interested in, we will see that geometric parameters introduce special complications into the problems we study. We will handle these difficulties at the appropriate time.

We will suppose, however, that if the values of the relevant parameters are known, then the particular fluid flow problem is completely defined, and can be solved to produce at least one flow solution. Let us consider some generic set of one or more parameters which we will call β. We will let u momentarily represent all three flow solution quantities for the continuous problem. If we fix all other aspects of the fluid flow problem, then we would like to regard the flow solution as a function of the parameters that are varying, a relationship we would represent by writing $u(\beta)$. Although we will frequently use such expressions, they are not quite proper. We have already discussed the fact that for higher Reynolds numbers, there may be multiple solutions of a given flow problem, so that the notation $u(\beta)$, suggesting a unique solution, is not strictly correct.

However, there is a sort of "local uniqueness" result for solutions to the flow problem. That is, there are certain generally isolated values of the Reynolds number, where the nature of the solution changes, and the graph of the solution may bifurcate. If we remove the data at those critical Reynolds numbers, the remaining solution graph breaks apart into a collection of open sets, each of which is a smoothly varying function of the Reynolds number, called a *smooth solution branch.* In general, this means that for a given set of parameters β_0, if we compute a flow solution u_0, we may reasonably write $u(\beta)$ to represent the local family of flow solutions that lie along the open branch that passes through the given solution $u(\beta_0) \equiv u_0$.

Once we confine our attention to a smooth branch of solutions, we may pose certain natural questions. We phrase these questions in terms of the continuous problem, though we will be interested in the discretized problem as well. Questions we might ask about the relationship between u and β include:

- Suppose that, for some set of parameter values β_0, we have computed a particular solution u_0. Can we estimate the value of flow solutions for nearby values of β?
- Can we, in fact, guarantee that there *is* a solution at a nearby value of β?
- Can we guarantee that the solution is *unique*?
- If we have the value $J(u_0)$ of some functional J at u_0, can we estimate the value of J at flow solutions for nearby values of β?
- At the given solution u_0, can we estimate the value of $\frac{\partial J}{\partial \beta}$?

We will show that the answer to these questions is generally positive, and that the primary tools we need are the linearizations (or, for the discrete case, the partial derivatives) of the state equations with respect to u and β, and the partial derivatives of the state variable u with respect to the parameters β, called the solution *sensitivity*.

5.3 The Sensitivity of a Continuous Solution

Let us assume that we have a solution space of pairs $((u,v,p),\beta)$ of state variable functions (u,v,p) and parameter sets $\beta \in R^m$. We assume that each state function is a suitably differentiable function defined over the open domain $\Omega \subset R^p$ and taking values in R^q. We assume that Ω has a piecewise smooth boundary Γ, and we allow the possibility that the domain Ω depends on β.

We assume that there is a state function, written $F((u,v,p),\beta) = 0$, with $u, v \in C^2(\Omega)$, $p \in C^1(\Omega)$, and $\beta \in R^n$. The function $F((u,v,p),\beta)$ comprises internal constraints, which are applied at all points $x \in \Omega$, and boundary constraints that hold just at points $x \in \Gamma$. A solution of the parameterized Navier Stokes equations may then be considered as a pair $((u,v,p),\beta)$ for which the equation $F((u,v,p),\beta) = 0$ holds.

Suppose we have already found a solution to the state equation, that is, a pair $((u_0, v_0, p_0)\beta_0)$ from the solution space, which satisfy the state equations. Having found such a solution, we are interested in understanding how the two components (u,v,p) and β are related, and whether, for small changes in β_0, we can make correspondingly small changes in (u_0, v_0, p_0) and come up with a new pair that again solves the state equation.

We would like to assert the existence of an implicitly defined function $h(\beta)$ which allows us to view the state solution locally as an explicit function of β. The function $h(\beta)$, if it exists, should have the property that it solves the flow problem for the parameters β, that is, for β "near" to β_0:

$$F(h(\beta),\beta) = 0 \ . \tag{5.1}$$

Moreover, we would like $h(\beta)$ to be continuous in β, and in fact continuously differentiable. In that case, assuming that the state system F has a continuous Frechet derivative, we may differentiate Equation (5.1) with respect to any component parameter β_i to arrive at:

$$F_{uvp}(h(\beta),\beta)\ h_{\beta_i}(\beta) + F_{\beta_i}(h(\beta),\beta) = 0 \ . \tag{5.2}$$

If the function $h(\beta)$ exists, then we may abuse notation and write $(u(\beta), v(\beta), p(\beta))$ instead, to emphasize the idea that, locally at least, the flow solution may be regarded as a function of the parameters. In that case, instead of $h_{\beta_i}(\beta)$ we write $(u_{\beta_i}(\beta), v_{\beta_i}(\beta), p_{\beta_i}(\beta))$.

We call Equation (5.2) the *continuous sensitivity equation*. We call the quantity $u_{\beta_i}(\beta)$ the *continuous sensitivity* of the solution component $u(\beta)$ with respect to the parameter β_i, although this is, of course, simply another name for the partial derivative of $u(\beta)$ with respect to β_i.

5.4 The Sensitivity of a Discrete Solution

When we talk about the solution of a discrete state equation, we are usually referring to the solution of a *discretized* continuous state equation. That is, it is understood that the discrete state equation was derived, via discretization, from some continuous state equation. This will be the case for all the problems we consider.

Whether or not there is a corresponding continuous state equation, we will assume that the discrete state equations do *not* involve any differentiation or integration with respect to the parameters, but rather are purely algebraic in u^h, v^h, p^h and β.

In fact, the finite element equations we have been considering are fairly simple functions of the unknown coefficients u_i^h, v_i^h and p_i^h. It might seem that the discretized finite element equations, Equations (4.10)-(4.12) plus the boundary conditions, do not have an algebraic form, since they involve integrals and derivatives of the state variables with respect to the spatial variables. But since the basis functions and approximating functions used in the finite element method are piecewise polynomials, the integrals such as

$$\int_\Omega Re\ u^h \frac{\partial u^h}{\partial x} w_i \, d\Omega \ , \tag{5.3}$$

are integrals of polynomials. This term, in particular, is equivalent to the expression:

$$\sum_{j=1}^{N_w} \sum_{k=1}^{N_w} \left(\int_\Omega Re\ w_j \frac{\partial w_k}{\partial x} w_i \, d\Omega \right) u_j^h u_k^h \ , \tag{5.4}$$

where the quantity in parentheses is simply a numerical quantity to be determined.

Because of complications that may occur with isoparametric elements, or nonpolynomial source terms, or boundary conditions, these integrals are usually carried out using numerical quadrature. Nonetheless, it should be clear now that the discretized finite element equations are merely a set of *quadratic* equations in the finite element coefficients.

We begin consideration of the discrete case by quoting a version of the *implicit function theorem for* R^n. This theorem gives us conditions under which a differentiable *relationship* involving state variables and parameters guarantees that there is actually an implicitly defined local differentiable *function* for the state variables in terms of the parameters.

Theorem 5.1 (The Implicit Function Theorem for R^n**)** *Suppose* $f : \mathcal{D} \rightarrow R^n$, *where* $\mathcal{D} \subset R^n \times R^m$ *is an open set, and that* $f \in C^1(\mathcal{D})$, *that* $f(x_0, y_0) = 0$ *for some* $(x_0, y_0) \in \mathcal{D}$, *and that* $rank[f_x](x_0, y_0) = n$.

Then there exists an open set $U \subset R^n \times R^m$, *containing* (x_0, y_0), *an open set* $V \subset R^m$ *containing* y_0, *and a function* $h : V \rightarrow R^n$ *such that* $h(y_0) = x_0$ *and* $f(h(y), y) = 0$ *for all* $y \in V$.

Furthermore, h *is uniquely determined by* $(h(y), y) \in U$ *for all* $y \in V$, *and* $h \in C^1(V)$ *and* $h'(y)$ *satisfies:*

$$f_x(h(y), y)h'(y) + f_y(h(y), y) = 0. \tag{5.5}$$

For our purposes, we want to slightly modify the hypotheses of this theorem, and change the terminology to fit our own more closely. In particular, we replace the rank condition by a stronger condition on the solvability of a certain linear system. This is in part because in order to get the "base solution" $((u_0^h, v_0^h, p_0^h), \beta_0)$, we will have to solve exactly such a linear system, and in part because such a formulation is closer to that used for the continuous problem. In the following theorem, we will commit a slight abuse of notation, by writing

$u^h(\beta)$ for the implicit function that was denoted by $h(y)$ in the previous theorem, and $u^h_\beta(\beta)$ for the derivatives.

Theorem 5.2 (Local Parameterization of Solutions in R^n) *Suppose that we have a set of n algebraic state equations involving a discrete state variable $u^h \in R^n$ and a set of parameters $\beta \in R^m$, which we write*

$$F^h(u^h, \beta) = 0; \tag{5.6}$$

that these state equations are well-defined and continuously differentiable in both u^h and β, throughout some open domain $\mathcal{D} \subset R^n \times R^m$; that we have a pair of values u^h_0 and β_0, with $(u^h_0, \beta_0) \in \mathcal{D}$, for which $F^h(u^h_0, \beta_0) = 0$; and that the linear system of equations

$$F^h_u(u^h_0, \beta_0)\psi = \phi \tag{5.7}$$

is uniquely solvable for any value ϕ.

Then there exists an open set $U \subset R^n \times R^m$, containing (u^h_0, β_0), an open set $V \subset R^m$ containing β_0, and a function $u^h : V \to R^n$ such that $u^h(\beta_0) = u_0$ and $F^h(u^h(\beta), \beta) = 0$ for all $\beta \in V$.

Furthermore, the function $u^h(\beta)$ is uniquely determined by $(u^h(\beta), \beta) \in U$ for all $\beta \in V$, and $u^h(\beta) \in C^1(V)$, and the derivative function $u^h_\beta(\beta)$ satisfies

$$F^h_u(u^h(\beta), \beta)u^h_\beta(\beta) + F^h_\beta(u^h(\beta), \beta) = 0. \tag{5.8}$$

Proof: Condition (5.7) implies that

$$rank(F^h_u(u^h_0, \beta_0)) = n, \tag{5.9}$$

and since all the other hypotheses of Theorem 5.2 are assumed, we may now assert the desired result. □

Thus, the implicit function theorem says that if we solve the discrete problem for a particular set of parameters β_0, getting a solution u_0^h, then for values of β "near" to β_0, there is a uniquely determined "nearby" state solution value, which we may write as $u^h(\beta)$. The pair $(u^h(\beta), \beta)$ satisfy the state equation, and $u^h(\beta)$ is continuously differentiable.

Since the discrete state equations are algebraic, F_u^h is simply the Jacobian matrix, and Condition (5.7) requires that this matrix be nonsingular. Given that, the derivative function $u_\beta^h(\beta)$ is defined by a solvable linear system.

Thus if we can assume the conditions of Theorem 5.2, then when we have a particular solution to the parameterized state equations, we know that there is actually a smoothly parameterized curve of solutions, which we may write as $u^h(\beta)$, of "nearby" solutions to "nearby" problems. We are interested in studying exactly how these nearby solutions vary with changes in the parameter. We make the following definitions:

Definition 5.1 (Discrete Sensitivity Equation) *Equation (5.8) is called the* **discrete sensitivity equation** *corresponding to the discrete state equation (5.6).*

Definition 5.2 (Discrete Sensitivity) *Under the assumptions of Theorem 5.2, the function $u_\beta^h(\beta)$ which solves the discrete sensitivity equation (5.8) is defined to be the* **discrete sensitivity** u_0^h *with respect to the parameters β.*

The expression $u_\beta^h(\beta)$ stands for the sensitivity of the discrete solution with respect to all of the parameters. For our problem, u^h actually represents three discretized functions, which are in turn represented by $2N_w + N_q$ real numbers. Supposing that we have $NPar$ parameters, then $u_\beta^h(\beta)$ could be expressed as a rectangular matrix of $2N_w + N_q$ rows and $NPar$ columns. A particular column would represent the sensitivity of all of the solution coefficients with respect to a particular parameter β_i, and we could denote this by $u_{\beta_i}^h(\beta)$, which is also equal to the partial derivative of $u^h(\beta)$ with respect to β_i.

Under the assumption that the state equations $F(u^h, \beta) = 0$ contain no derivatives of the state variables with respect to β, the sensitivity equation (5.8) is *linear* in the unknown function u^h_β.

Perhaps we should emphasize the meaning of Equation (5.8). It tells us that to compute the sensitivity of a particular solution to the state equations, we must differentiate the state equations with respect to the parameter of interest, and evaluate the resulting quantities at the particular solution, to arrive at the sensitivity equations.

5.5 The Discretized Sensitivity of a Solution

If we start with a continuous state equation for a state variable u, discretize it to a system for a discrete state variable u^h, we know from the previous section that we may also derive a sensitivity equation from this discretized state equation, for a quantity u^h_β, and that the sensitivity u^h_β is exactly the partial derivative of the discrete state variable u^h with respect to β.

Suppose, instead, that, beginning with the same continuous state equation, we first derive the continuous sensitivity equation, and *then* discretize that. We have interchanged the two operations that gave us the sensitivity u^h_β. What can we say about the results of this operation?

When we discretize the state equation first, we derive the sensitivity equation by differentiation of the discretized system. This can require the differentiation of numerous terms that do not appear directly in the state equation, but are simply part of the discretization process that may depend on the parameters, such as the location of quadrature points, the area of finite elements, the form of basis functions, and so on.

By contrast, it is generally easy to differentiate the continuous state equation with respect to a parameter. Once this is done, the same process that discretized the state equation can be applied to the sensitivity equation. This means that there is far less work required to design the algorithm, and we will see later that there will also be far less work to carry out the algorithm as well.

Before we proceed, we give a name to the solution of the system that is produced by discretizing the continuous sensitivity equation.

Definition 5.3 (Discretized Sensitivity of a Solution) *Suppose that we have a continuous state equation $F(u, \beta) = 0$ with solution u, and the corresponding continuous sensitivity equation with sensitivity u_β. Suppose that a discretization can be applied, yielding a discretized state equation $F^h(u^h, \beta) = 0$. Suppose that this same discretization can be applied to the continuous sensitivity equation (5.2), yielding what will be called the* **discretized sensitivity equation***:*

$$(F_u)^h(u^h(\beta), \beta)\ (u_\beta)^h + (F_\beta)^h(u^h(\beta), \beta) = 0\ . \tag{5.10}$$

The solution $(u_\beta)^h$ of this equation, if any, is called the **discretized sensitivity** *of the continuous solution u.*

Because a discretization operation has been applied, the discretized sensitivity equation is not guaranteed to be a sensitivity equation, that is, it need not represent the defining equation for a partial derivative.

Moreover, when we solve the discretized sensitivity equations, the discretized sensitivities are *not* guaranteed to actually be the partial derivatives of any quantity. In particular, they are not guaranteed to be the partial derivatives of the discrete state variables with respect to the parameters.

It is easy to miss this fact, because in many common cases, the operations of differentiation and discretization may commute. That is, we *may* have the case that:

$$(u_\beta)^h = u_\beta^h, \tag{5.11}$$

but such a situation must be verified.

In our setting, because of the convergence properties of the finite element method, we can assert that, as $h \to 0$ the discretized sensitivity $(u_\beta)^h$ approaches u_β, the sensitivity of the continuous solution. But even this statement doesn't guarantee commutativity, since we haven't shown that u_β^h also approaches u_β.

5.6 Example: An Explicit Parameter in a Linear ODE

We will now consider a simple case where the sensitivities are easy to compute. Consider the following linear ordinary differential equation, with parameter β, which will play the role of our continuous state equation:

$$\frac{du(t,\beta)}{dt} = \beta\ u(t,\beta), \tag{5.12}$$

$$u(0,\beta) = u_0. \tag{5.13}$$

For any particular value of β, the exact solution, that is, our continuous state solution, is:

$$u(t,\beta) = u_0\ e^{\beta t}, \tag{5.14}$$

from which we may immediately compute the continuous sensitivity:

$$u_\beta(t,\beta) = \frac{\partial u}{\partial \beta} = u_0\ t\ e^{\beta t}\ . \tag{5.15}$$

We now show that we do not need to have an explicit formula for the dependence of u on β in order to compute u_β. To show this, we compute the continuous sensitivity equations:

$$\frac{du_\beta(t,\beta)}{dt} = \beta\ u_\beta(t,\beta) + u(t,\beta), \tag{5.16}$$

$$u_\beta(0,\beta) = 0. \tag{5.17}$$

For a fixed value of β, we can solve this system directly for $u_\beta(t,\beta)$. We would not need a formula for u valid for all β, but only the form of $u(t,*)$ for the current values of the parameters.

Now we suppose that we wish to define and solve a discretized version of this problem. We discretize by setting down a mesh of equally spaced points t_i, with a spacing of h, and with the initial point set to $t_0 = 0$, and then using the forward Euler approximation for the derivative with respect to t. Then we may immediately write the resulting discrete state equations as:

$$\frac{u^h(t_{i+1},\beta) - u^h(t_i,\beta)}{h} = \beta u^h(t_i,\beta), \tag{5.18}$$

$$u^h(t_0,\beta) = u_0. \tag{5.19}$$

Now suppose that we have solved the discrete state equation for the state variables $u^h(t_i,\beta)$. We may then differentiate the discrete state equations with respect to β to get the sensitivity equations for the discretized problem:

$$\frac{u^h_\beta(t_{i+1},\beta) - u^h_\beta(t_i,\beta)}{h} = \beta u^h_\beta(t_i,\beta) + u^h(t_i,\beta), \tag{5.20}$$

$$u^h_\beta(t_0,\beta) = 0\ . \tag{5.21}$$

If we instead discretize the continuous sensitivity equation, we arrive at the discretized sensitivity equations:

$$\frac{(u_\beta)^h(t_{i+1},\beta) - (u_\beta)^h(t_i,\beta)}{h} = \beta (u_\beta)^h(t_i,\beta) + u^h(t_i,\beta), \tag{5.22}$$

$$(u_\beta)^h(t_0,\beta) = 0\ . \tag{5.23}$$

The equations defining u^h_β and $(u_\beta)^h$ are the same, and so their solutions must be the same. For this problem and this discretization, then, differentiation and discretization may be applied in either order without affecting the result.

5.7 Sensitivities for Geometric Parameters

We have seen that the derivation of the sensitivity equations can be quite straightforward when the parameter in question appears explicitly in the state equations. Simply differentiating the state equations produces the desired system.

But geometric parameters do not appear in such a simple way. For instance, a state equation involving a geometric parameter might have the form $u(\beta) = 0$, where β is to be interpreted here as a spatial location. Our simple technique of differentiation will not produce a proper result here.

To produce a proper formulation of the sensitivity with respect to a geometric parameter, it is best, if possible, to return to the definition of a sensitivity as a difference quotient. In other words, given a solution $u(x, \beta_0)$, we can compute its sensitivity to the parameter β by imagining that we compute a solution at $\beta_0 + \Delta\beta$, and considering the limit

$$u_\beta(x, \beta_0) = \lim_{\Delta\beta \to 0} \frac{u(x, \beta_0 + \Delta\beta) - u(x, \beta_0)}{\Delta\beta} . \qquad (5.24)$$

The difference quotient makes a comparison between values of u at the same spatial coordinates. Because a geometric parameter can actually alter the shape of the region, we may find that even for small perturbations of the parameter, a particular point x no longer lies within the region, and has no associated value of u. This problem is particularly acute when the point x lies on the boundary of the region. However, there are generally ways to handle such difficulties by extending or extrapolating the solution as necessary. We will see, when we consider the sensitivities with respect to bump parameters, that for our formulation, we can always choose the sign of the perturbation of the bump parameter so that the point of interest remains in the flow region, and hence has defined solution values that may be used to make the necessary comparison.

It should be clear that geometric parameters are more difficult to handle when computing sensitivities. In the next chapter, we will see that geometric parameters can also create situations in which the discretization and differentiation operations do not commute.