Chapter 8

FINITE DIFFERENCE ESTIMATES OF SENSITIVITIES

8.1 Introduction

A common method of approximating derivatives is to use finite differences. Since the sensitivities are derivatives, or approximations of them, we would like to compute finite difference estimates of the quantities whose sensitivites we are studying. Such an independent computation will be useful as a check of both the accuracy and efficiency of the discretized sensitivity approach.

We compare the definitions, computations, accuracy, and usefulness of discretized sensitivities and finite difference sensitivities. We show that if shape parameters are involved, it is more difficult to make comparisons, particularly if we are interested in comparing individual finite element coefficients, that is, values associated with nodes. This is mainly because the nodes move, while a sensitivity must refer to changes at a fixed location.

We consider an adjustment to finite difference calculations involving finite element coefficients. This adjustment allows a comparison with the discretized sensitivities. We show data that suggests that, for our finite element formulation, the approximation error, as measured by the difference between the discretized sensitivities and the adjusted finite difference quantities, decreases linearly with h.

This is empirical evidence that the discretized sensitivities, which are cheaper to compute than discrete sensitivities, have good enough approximation abilities for most uses.

8.2 Finite Coefficient Differences and Finite Physical Differences

In earlier discussions of sensitivities with respect to a parameter, we have mentioned the idea of comparing a discretized sensitivity and a finite difference estimate. We did not stop to explain this concept; it is a common technique, and, for explicit parameters, does not require any special thought or consideration.

However, for a geometric parameter, particularly in the discrete case, we will see that the calculation of a finite difference sensitivity, and comparison to a discretized sensitivity requires some extra care. Let us suppose that, for some geometric parameter value α, we have a solution to the discrete problem, comprising a set of finite element coefficients. To be more particular, let us suppose that one of those items is u_i, a velocity coefficient associated with node i, which is itself located at node (x_i, y_i).

If we change the value of α, the quantities that change may include not just u_i, but also the location of the associated node. At this point there are two paths we could follow. If we simply compute the finite difference quotient

$$\frac{u_i(\alpha + \delta\alpha) - u_i(\alpha)}{\delta\alpha} \tag{8.1}$$

then we are approximating the behavior of the finite element coefficient itself, which we can term a *finite coefficient difference*.

To approximate the behavior of the the physical quantity u at (x_i, y_i), we would need to determine where this point lies in the finite element mesh associated with $\alpha + \delta\alpha$, evaluate the finite element approximation, and then compute the quotient:

$$\frac{u^h(x_i, y_i, \alpha + \delta\alpha) - u^h(x_i, y_i, \alpha)}{\delta\alpha}. \tag{8.2}$$

This is a proper approximation of the physical quantity, and hence we might term this a *finite physical difference.*

It may not be clear, but it is much easier to compute the finite coefficient differences than the finite physical differences. As a geometric parameter varies, the nodes and elements move, complicating the evaluation of the perturbed solution. But the structure of the finite element coefficient vector itself will generally not change, and it is a trivial matter to compute differences of comparable vector elements.

To get an idea of the distinction between these quantities, we will look at plots of the discretized sensitivities (which approximate the same quantities as the finite physical differences) versus the finite coefficient differences. Figure 8.1 displays the discretized α-sensitivities for velocity, while Figure 8.2 shows a corresponding finite coefficient difference field, computed by comparing coefficient values.

Significant information can be found by comparing the two plots. In particular, we note that the discretized sensitivity field is nonzero on the bump, whereas the corresponding finite difference sensitivities vanish; the discretized sensitivities seem to satisfy the continuity equation (the discretized sensitivity equation forces this!) while this is not true for the finite difference sensitivities; the discretized sensitivities and finite difference sensitivities seem to agree in the portions of the region before and after the bump, where the geometry is never distorted.

The plot should make it clear that discretized sensitivities and finite coefficient differences can

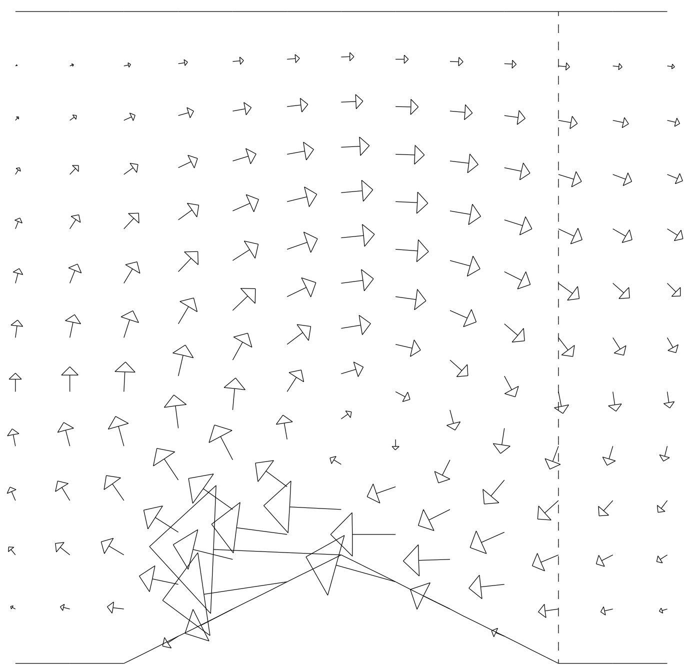

Figure 8.1: Discretized α-sensitivity nodal values.
Notice, in particular, the nonzero boundary condition on the bump.
The sensitivities estimate change at a fixed location.

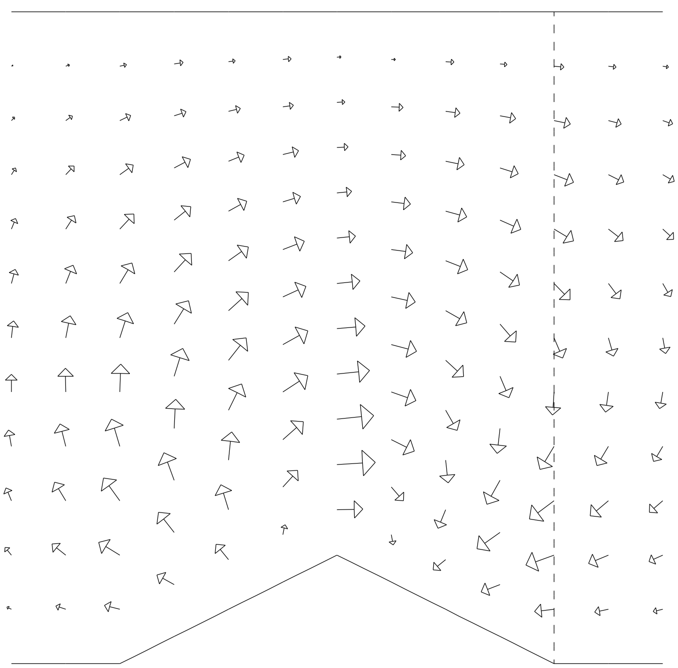

Figure 8.2: The corresponding α-finite difference nodal values.
The boundary values are exactly zero, and the "vortex" is quite different.
These values apply at a fixed node, but a moving location.

be quite different quantities. We propose to explain what this difference is, and how it can be quantified. We are interested in how well our discretized sensitivities approximate the true continuous sensitivities and the discrete sensitivities. Understanding the finite coefficient differences will help us in this task.

8.3 Differencing with Respect to Geometric Parameters

Throughout this chapter, we will assume that we are varying a single shape parameter which we shall denote by α. We will occasionally want to write a formula that references state variable values such as $u(x, y)$. When we are varying a geometric parameter, there may be times when the point (x,y) does not lie within the current flow region $\Omega(\alpha)$, so that we don't have a well-defined value for u. In such cases, let us agree that u will simply be set to 0. While better choices exist to extend the solution to the neighboring area, right now we are only interested in assuring ourselves that an expression like $u(x,y)$ will always have a defined value.

In order to distinguish the information produced by a finite coefficient difference sensitivity from that in a discretized sensitivity, we must consider how a finite difference quotient is computed. Again, we assume that we have a particular value of α of interest to us, with a corresponding solution (u,v,p), and as usual, we will focus our attention on u. Typically, once the current solution u is calculated at α, we slightly perturb the value of the parameter to $\alpha+\Delta\alpha$ and compute a new solution, which we may write $u(\alpha+\Delta\alpha)$. We then form a difference quotient whose numerator is the change in the solution at a fixed location (x,y), and whose denominator is the change in the parameter. Since the (total) derivative of a quantity depending on α is defined as the limit of such quotients as $\Delta\alpha \rightarrow 0$, we can expect, for reasonably smooth problems and small $\Delta\alpha$, that we can compute a good approximation to this quantity.

If we focus our attention on a point (x,y) which lies in $\Omega(\alpha)$, and consider how the value of $u(x,y)$ will vary with α, then we can approximate the desired partial derivative $\frac{\partial u}{\partial \alpha}(x,y,\alpha)$ by the finite difference quotient:

$$\frac{\Delta u(x,y,\alpha)}{\Delta\alpha} \equiv \frac{u(x,y,\alpha\pm\Delta\alpha)-u(x,y,\alpha)}{\pm\Delta\alpha} \tag{8.3}$$

Perhaps we should explain the cumbersome form of the quotient on the right hand side. In this formula, we assume that the perturbation $\Delta\alpha$ is small enough, and that we have chosen the proper sign for it, so that the point (x,y) is not only in $\Omega(\alpha)$ but also in $\Omega(\alpha+\Delta\alpha)$. Simple control of the size of the perturbation $\Delta\alpha$ is enough to guarantee that a point (x,y) in the interior of $\Omega(\alpha)$ will also lie in the interior of $\Omega(\alpha+\Delta\alpha)$. But for points that actually lie on the bump, we need to control the sign of the perturbation as well, and even then, we can't guarantee that the point will remain in the perturbed region unless we can show that:

$$\frac{\partial Bump}{\partial \alpha}(x, \alpha) \neq 0. \tag{8.4}$$

For a bump defined by a piecewise linear function, for instance, we can guarantee that Condition (8.4) is true everywhere except at the fixed endpoints, where we don't care.

So now if u is the solution of a continuous problem, Equation (8.3) may be used to define a finite difference quotient $\frac{\Delta u}{\Delta \alpha}(x, y, \alpha)$ which has values throughout the flow region, and which estimates the influence of α. The sensitivity of the solution of the continuous problem at (x, y) is the partial derivative $\frac{\partial u(x, y, \alpha)}{\partial \alpha}$, but this derivative is precisely the limit of the finite difference quotients defined by Equation (8.3), so we may assert immediately that, for the continuous problem,

$$\lim_{\Delta\alpha \to 0} \frac{\Delta u}{\Delta \alpha}(x, y, \alpha) = u_\alpha(x, y, \alpha), \tag{8.5}$$

and thus, finite difference quotients can be used as a check on the accuracy of the sensitivity calculation.

Now, by contrast, let us suppose that we are interested in the value of u at a *moving* point. This might seem a needless complication, until one realizes that points on the bump surface move, that we have nodes on the bump, and in our formulation, all the interior nodes above the bump will also move if the bump moves. The interior nodes determine the shape of the elements, the form of the basis functions, and the physical "anchor" for the coefficient values, so a lot changes when a node moves.

Both on the bump, and at the interior nodes, we expect to find nonzero sensitivities. We will keep our problem simple by assuming that only the y coordinate of such a point varies with α, so that we may write its coordinates as $(x, y(\alpha))$. Further, we assume that this point remains in $\Omega(\alpha)$ for all values of α of interest to us. At such a moving point, we now must consider u to be effectively a function of only x and α. To study the influence of the parameter α, we would need to find the total derivative of u with respect to α, which we write $\frac{Du}{D\alpha}(x, y(\alpha), \alpha)$, and we may approximate this total derivative by the finite difference quotient:

$$\frac{\Delta u(x, y(\alpha), \alpha)}{\Delta \alpha} = \frac{u(x, y(\alpha + \Delta\alpha), \alpha + \Delta\alpha) - u(x, y(\alpha), \alpha)}{\Delta \alpha}. \tag{8.6}$$

Moreover, we note that for this situation, we have a simple relationship between the total derivative and the sensitivity:

$$\frac{Du}{D\alpha}(x, y(\alpha), \alpha) = \frac{\partial u}{\partial y}(x, y(\alpha), \alpha)\frac{dy(\alpha)}{d\alpha} + \frac{\partial u}{\partial \alpha}(x, y(\alpha), \alpha), \tag{8.7}$$

which breaks down our total derivative into a change due to spatial movement, and a change due to variations in α.

From what we have discussed so far, we can now explain a few things about Figure 8.1 and Figure 8.2. First of all, in the regions before and after the bump, none of the nodes move with changes in α. Therefore, in this region the discretized sensitivities and the finite coefficient differences should agree closely.

We can also explain the "boundary value discrepancy", that is, why on the surface of the bump the discretized sensitivities are nonzero while the finite coefficient differences are exactly zero. The discretized sensitivity approximates the true sensitivity, which is computed as the limit of finite difference quotients that compare the value of u to values of u for nearby values of α, but at the same location. Clearly, if α decreases, the velocity at the point (which is now zero) will increase, that is, fluid will pass through the point, moving to the right. Therefore, the sensitivity, which is the limit of ratios of this positive change in u to this negative change in α is represented as a *negative* horizontal velocity increment.

By contrast, the finite coefficient difference compares the current value of u (which is zero) to values of u at nearby values of α, at locations that move (up or down) with the bump. But a point on the bump always has zero horizontal velocity u, so the finite coefficient difference associated with a node on the bump will always be zero.

Thus, we arrive at quite different results, based on whether we allow the point under consideration to move with α or stay fixed.

8.4 Finite Differences for a Discretized Problem

Now let us suppose that, corresponding to a parameter value of α, we have the solution of a discrete problem, which may be represented a set of coefficient data $u_i^h(\alpha)$, which in turn determine a continuously differentiable function of x, y, and α:

$$u^h(x, y, \alpha) \equiv \sum_i u_i^h(\alpha)\ w_i(x, y, \alpha). \tag{8.8}$$

where, as usual, if the geometric point of evaluation happens to be the i-th node, (x_i, y_i), then

$$u^h(x, y_i(\alpha), \alpha) = u_i^h(\alpha). \tag{8.9}$$

As in the continuous case, we may define the following finite difference approximation to the partial derivative $\frac{\partial u^h(x,y,\alpha)}{\partial \alpha}$:

$$\frac{\Delta u^h(x, y, \alpha)}{\Delta \alpha} \equiv \frac{u^h(x, y, \alpha \pm \Delta\alpha) - u^h(x, y, \alpha)}{\pm\Delta\alpha}. \tag{8.10}$$

where we choose the magnitude of $\Delta\alpha$ and, if necessary, the sign, so that the point (x, y) is in $\Omega(\alpha + \Delta\alpha)$ as well as in $\Omega(\alpha)$.

From this definition, and assuming u^h is twice continuously differentiable in α, we may use a Taylor expansion to see that

$$\frac{\Delta u^h(x, y, \alpha)}{\Delta \alpha} = \frac{\partial u^h(x, y, \alpha)}{\partial \alpha} + O(\Delta\alpha), \tag{8.11}$$

and hence, for small $\Delta\alpha$, the finite difference quotient is a good approximation to the sensitivity.

Now, if we again assume we are interested in the changes in u^h at a point whose y-coordinate changes with α, and if we assume that the point is always in the flow region, then we consider the difference quotient:

$$\frac{\Delta u^h(x,y(\alpha),\alpha)}{\Delta\alpha} = \frac{u^h(x,y(\alpha+\Delta\alpha),\alpha+\Delta\alpha) - u^h(x,y(\alpha),\alpha)}{\Delta\alpha}, \tag{8.12}$$

and we can use a Taylor expansion to show that

$$\begin{aligned} \frac{\Delta u^h(x,y(\alpha),\alpha)}{\Delta\alpha} &= \frac{Du^h(x,y(\alpha),\alpha)}{D\alpha} + O(\Delta\alpha) && (8.13)\\ &= \frac{\partial u^h}{\partial y}\frac{dy}{d\alpha} + \frac{\partial u^h}{\partial\alpha} + O(\Delta\alpha). && (8.14) \end{aligned}$$

In particular, let us take the moving point to be velocity node i, so that the y-coordinate of the node may be written $y_i(\alpha)$. In this case, we may write

$$\begin{aligned} \frac{\Delta u^h(x,y_i(\alpha),\alpha)}{\Delta\alpha} &= \frac{u^h(x,y_i(\alpha+\Delta\alpha),\alpha+\Delta\alpha) - u^h(x,y_i(\alpha),\alpha)}{\Delta\alpha} && (8.15)\\ &= \frac{u_i^h(\alpha+\Delta\alpha) - u_i^h(\alpha)}{\Delta\alpha}, && (8.16) \end{aligned}$$

where we have written $u_i^h(\alpha)$ to indicate the dependence of the i-th horizontal velocity finite element coefficient on the parameter α. In other words, the total derivative of u^h along the path of the moving node is simply the difference quotient of the associated coefficient $u_i^h(\alpha)$.

Combining Equations (8.14) and (8.16), we see that:

$$\frac{u_i^h(\alpha+\Delta\alpha) - u_i^h(\alpha)}{\Delta\alpha} = \frac{\partial u^h}{\partial y}(x,y_i(\alpha),\alpha)\frac{dy_i(\alpha)}{d\alpha} + \frac{\partial u^h}{\partial\alpha}(x,y_i(\alpha),\alpha) + O(\Delta\alpha), \tag{8.17}$$

which means that we may estimate the true sensitivity at node i by

$$\begin{aligned} u_\alpha^h(x,y_i(\alpha),\alpha) &\equiv \frac{\partial u^h(x,y_i(\alpha),\alpha)}{\partial\alpha} && (8.18)\\ &\approx \frac{u_i^h(\alpha+\Delta\alpha) - u_i^h(\alpha)}{\Delta\alpha} - \frac{\partial u^h}{\partial y}(x,y_i(\alpha),\alpha)\frac{dy_i(\alpha)}{d\alpha}. && (8.19) \end{aligned}$$

Now, if we call the quantity on the right hand side of Equation (8.19) the "adjusted finite coefficient difference", we can now see how to relate three quantities of interest to us: the discretized sensitivities, the finite coefficient differences, and the discrete sensitivities. We will make our comparison by considering individual finite element coefficients. For a particular value of α, consider a flow solution (u^h, v^h, p^h), and some finite element node i set at $(x, y_i(\alpha))$. The discrete sensitivity of the solution is $u^h_\alpha(x, y_i(\alpha))$. We can now see that this quantity is approximated to order $O(\Delta\alpha)$ by the adjusted finite coefficient difference. Moreover, the discrete sensitivity at $(x, y_i(\alpha))$ is also approximated, to an order depending in some way on the mesh parameter h (which we expect to be roughly $O(h)$ for our Taylor Hood approach), by the discretized sensitivity $(u_\alpha)^h(x, y_i(\alpha))$, which is simply the discretized sensitivity coefficient at node i. Thus, for suitably small values of $\Delta\alpha$ and h, the discretized sensitivity coefficient and the adjusted finite coefficient difference should be comparable.

Therefore, for a discrete problem, we can compare discretized sensitivities and finite difference sensitivities at a node by comparing the corresponding finite element coefficients, but we must make an adjustment to the finite difference sensitivity coefficients in cases where the associated node actually moves with changes in α.

To demonstrate this fact, we adjust the finite coefficient difference velocity field that was shown Figure 8.2 (and which matched so poorly with our discretized sensitivity field) and display the result in Figure 8.3. There are no arrows displayed along the boundary, because we don't have unknown coefficients associated with those nodes. In the interior, however, the agreement with the discretized velocity sensitivity field of Figure 8.1 is now striking.

Equation (8.18) tells us how to adjust the finite coefficient differences associated with a moving node, so that they approximate the sensitivity of a solution quantity at a fixed point. Conversely, the equation also tells us how to transform the discretized sensitivity coefficient associated with a node so that we can estimate the change in the solution coefficient as the

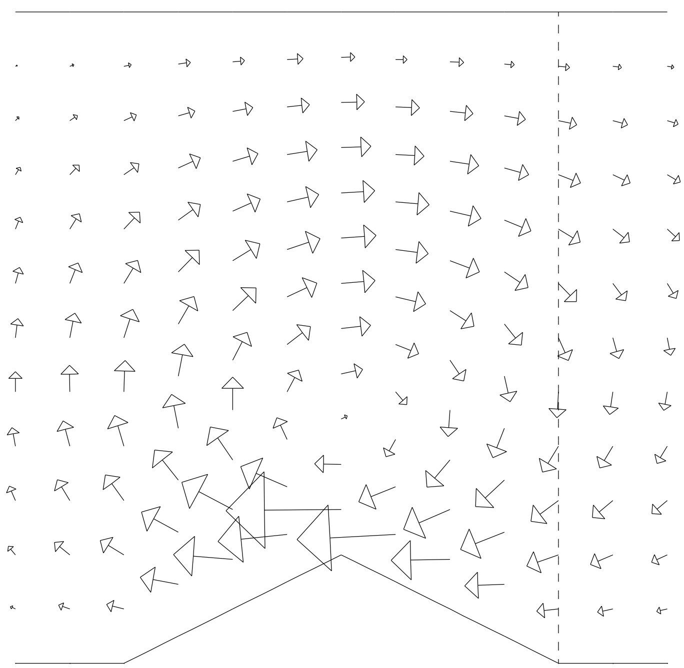

Figure 8.3: Adjusted finite difference α-sensitivity nodal values.

parameter changes and the node moves. The corresponding formula is simply:

$$\frac{u_i^h(\alpha+\Delta\alpha)-u_i^h(\alpha)}{\Delta\alpha} \approx (u_\alpha)^h(x,y_i(\alpha),\alpha)+\frac{\partial u^h}{\partial y}(x,y_i(\alpha),\alpha)\frac{dy_i(\alpha)}{d\alpha}. \quad (8.20)$$

Whereas Equation (8.18) is useful for estimating the closeness of our approximation to the true sensitivities, Equation (8.20) is useful if we wish to use discretized sensitivities to estimate the change that will occur in our finite element coefficients associated with a moving node, when a shape parameter is varied. This is exactly the sort of calculation we must make when trying to produce a suitable starting point for the Picard or Newton iteration.

8.5 Sensitivity Approximation with Decreasing h

We now wish to make a comparison, for a shape parameter, between the discretized sensitivities and the adjusted and unadjusted finite coefficient differences. We already saw a fairly poor agreement in the previous chapter, where we compared discretized sensitivities and unadjusted finite coefficient differences.

Table 8.1 is based on solutions of a problem involving three parameters, (λ, α, Re), and a flow solution computed at $(0.5, 0.5, 10.0)$, on grids with $h = 0.25$, $h = 0.125$, and $h = 0.0625$. For each mesh parameter, we computed the discretized sensitivities, the finite coefficient differences, and the adjusted finite coefficient differences, that is, the results of Equation (8.18). In Table 8.1 we print out the largest of each of these quantities, broken down into the components associated with u, v and p. The last column gives the maximum difference between the finite element coefficients for the discretized sensitivities and the adjusted finite coefficient differences.

From the table, it is clear that we are not getting the superb agreement between discretized sensitivities and finite coefficient differences that we have seen for explicit parameters; even

Table 8.1: Finite coefficient differences versus discretized α-sensitivities. Table entries represent the maximum absolute coefficient value.

h	Finite Coef Difference (FCD)	Adjusted FCD (AFCD)	Discretized Sensitivity (DS)	AFCD - DS
		U		
0.25	0.309	0.802	0.654	0.148
0.125	0.326	1.07	1.02	0.0834
0.0625	0.327	1.33	1.30	0.0327
		V		
0.25	0.242	0.260	0.271	0.108
0.125	0.260	0.273	0.272	0.0584
0.0625	0.258	0.276	0.272	0.0232
		P		
0.25	0.438	0.438	0.471	0.133
0.125	0.400	0.493	0.507	0.0684
0.0625	0.346	0.462	0.476	0.0338

the adjusted finite coefficient differences have a significant disagreement with the discretized sensitivities. However, we expected that fact. The first interesting thing to note is how much better the agreement is between the discretized sensitivities and the adjusted finite coefficient differences. Evidently, the adjustment for moving nodes is quite significant. The more important question is, how does the disagreement behave with decreasing mesh parameter h?

We ask this question because we really want to know how the discretized sensitivities approximate the true sensitivities. That information is not available to us, since we have not tried to compute the true sensitivities. However, we assume that the adjusted finite coefficient differences are a good estimate of the true sensitivities. If the discretized sensitivities approximate the adjusted finite coefficient differences well, then we expect that they are almost as good an approximation of the true sensitivities.

We must keep in mind, though, that the adjusted finite coefficient differences are by no

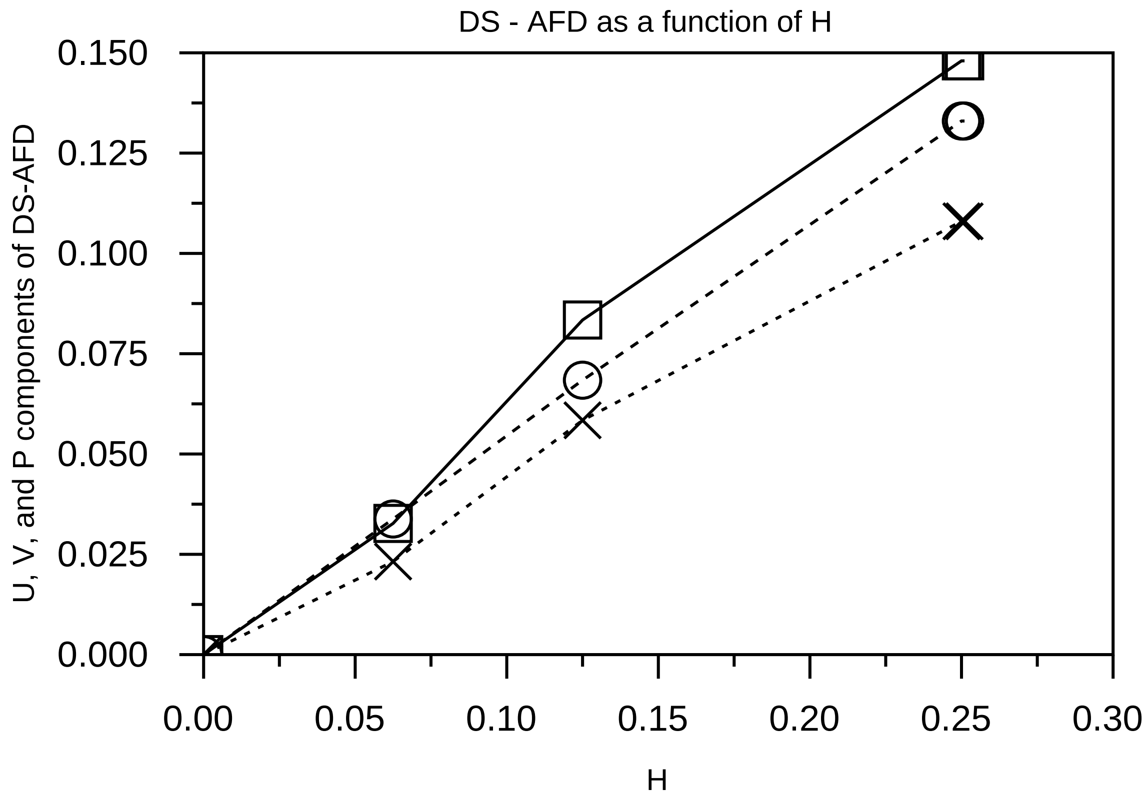

Figure 8.4: Discretized sensitivity/adjusted finite coefficient difference discrepancies. The box marks u data, the cross v, the circle p.

means perfect estimates of the true sensitivities. The errors incurred by the adjusted finite coefficient differences include the usual finite difference error, but also the added inaccuracy due to the use of the Taylor approximation involving $\frac{\partial u^h}{\partial y}$ for the adjustment term.

The discrepancies between the discretized sensitivities and the adjusted finite coefficient differences are plotted in Figure 8.4. There is a clear trend in the data. For each solution component, the discrepancy between the discretized sensitivity and the adjusted finite coefficient differences drops roughly in step with h; as h is halved, so is the discrepancy. This bears out our claim that, even for geometric parameters, the discretized sensitivities converge to the true sensitivities as h goes to zero. The actual order of convergence depends, of course, on the particular discretization applied to the problem, and seems, for our problem, to be behaving linearly in h, as we have already estimated.

The table contains more information, if we look closely. As h decreases, the infinity norms of the horizontal velocity and the pressures increase significantly. This should be surprising. It suggests that the grid is still not fine enough to capture important features of the flow. But this hasn't happened on other problems, so what is different here? The difference is surely the fact that there is a strong boundary condition being applied on a curved boundary. Very important effects are occurring just on the boundary, and we are missing some of these details if the mesh is not fine enough.