Theoretical Tools for Problem Setup

Shlomo Ta'asan
Carnegie Mellon University

1 Introduction

In this lecture we demonstrate the use of Fourier analysis to obtain a practical quantitative information about different optimization problems. Our main concern here is to study question regarding the formulation of the optimization problem, the choice of design variables and the choice of the cost functional. The complexity of the optimization problem will be shown to depend on these choices. We give examples to demonstrate this point and suggest a general approach for choosing design variables, and/or cost functional to achieve minimization problems that are good on one hand but also easy to solve on the other hand. For many engineering problems there is some freedom in the formulation of the problem and it is certainly an advantage to deal with the easier problems yet keeping the same engineering design tasks. We review basic facts from Fourier analysis and pseudo-differential operators and show its practical use for the analysis of Hessians. This analysis gives a very simple classification of problems based on the asymptotic behavior of the symbol of the Hessian at the high frequency range. It distinguishes between ill-posed (bad) problems, well-posed (good) problems, easy problems and difficult problems. This classification is of practical importance in the problem setup.

2 Review of Fourier Analysis for PDE

Fourier analysis has been a powerful tool for many years in analyzing different numerical procedures, and it goes back to Von Neumann. It can be used for a variety of tasks including a quantitative information about the behavior of solutions, their dependence on boundary values, etc. It can also be used to analyze numerical procedures that we use to solve partial differential equations or optimization problems.

We go here briefly on the main ideas we need for the analysis of optimization problems. Let us begin by a simple problem, the linearized small disturbance potential equation of fluid dynamics. This equation will demonstrate the different issues we are concerned with, and will set the foundation for the treatment of general problems, including systems of equations such as the Euler and the Navier-Stokes equations.

Example I: A Scalar Equation. Consider the equation

$$(1-M^2)\phi_{xx}+\phi_{yy}+\phi_{zz}=0, \tag{1}$$

in the domain $\Omega=\{(x,y,z)|z>0\}$. This equation is an approximation for a flow in the x direction, where perturbation velocities (around a mean velocity) are related to the potential ϕ as $u=\phi_x, v, \phi_y, w=\phi_z$, see Hirsch [12].

We want to analyze the solution in terms of its boundary values on $z=0$. We consider boundary data in terms of Fourier components

$$\phi(\mathbf{x},0)=\exp(i\mathbf{k}\cdot\mathbf{x}), \tag{2}$$

where $\mathbf{x}=(x,y)$ and $\mathbf{k}=(k_1,k_2)$. We look for an exponential behavior in the z direction

$$\phi(\mathbf{x},z)=\exp(i\mathbf{k}\cdot\mathbf{x})\exp(\sigma z), \tag{3}$$

and a substitution of this expression into equation (1) for ϕ leads to the relation

$$\sigma^2=(1-M^2)k_1^2+k_2^2. \tag{4}$$

Note that this equation for σ has two solutions. For $M<1$ one of them, $\sigma_1<0$, correspond to a decaying solution for ϕ as a function of z and is considered as the solution of interest. The other one, $\sigma_2>0$, correspond to an unbounded solution for ϕ and is discarded. Thus, for the subsonic case we have

$$\phi(\mathbf{x},z)=\exp(i\mathbf{k}\cdot\mathbf{x})\exp\left(-z\sqrt{(1-M^2)k_1^2+k_2^2}\right) \tag{5}$$

For the supersonic case, $M>1$, there are two bounded solutions for $(1-M^2)k_1^2+k_2^2<0$, and only one bounded solution for $(1-M^2)k_1^2+k_2^2\geq 0$. The two bounded solutions can be viewed as an incident wave and a reflected wave, whose sum satisfies the boundary condition,

$$\begin{aligned}\phi_i(\mathbf{x},z)&=A_i\exp(i\mathbf{k}\cdot\mathbf{x})\exp\left(iz\sqrt{-(1-M^2)k_1^2-k_2^2}\right) \quad (1-M^2)k_1^2+k_2^2<0\\ \phi_r(\mathbf{x},z)&=A_r\exp(i\mathbf{k}\cdot\mathbf{x})\exp\left(-iz\sqrt{-(1-M^2)k_1^2-k_2^2}\right) \quad (1-M^2)k_1^2+k_2^2<0\end{aligned} \tag{6}$$

A general solution for the problem can be written as a sum (integral) of these Fourier components.

Example II: A System of PDE. We next show the use of Fourier analysis for studying systems of partial differential equation. Consider the system

$$\begin{pmatrix}\beta^2\frac{\partial}{\partial x} & \frac{\partial}{\partial y}\\ \frac{\partial}{\partial y} & -\frac{\partial}{\partial x}\end{pmatrix}\begin{pmatrix}u\\ v\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix} \qquad \Omega \tag{7}$$

where $\Omega = \{(x,y)|y \geq 0\}$ and $\beta^2 = 1 - M^2$. Here one considers vector Fourier components, that is

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} A \\ B \end{pmatrix} \exp(ik_1 x)\exp(\sigma y). \tag{8}$$

The problem is to determine σ but also the relation of A and B. This is done by substituting this form into the equations leading to an algebraic equation for A, B,

$$\begin{pmatrix} i\beta^2 k_1 & \sigma \\ \sigma & -ik_1 \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{9}$$

This is a linear set of equations for the unknown A, B and it has a non zero solution if and only if the determinant of the system is zero. This gives an equation for σ in terms of k_1,

$$\beta^2 k_1^2 - \sigma^2 = 0. \tag{10}$$

For the subsonic case $\beta^2 > 0$ and we have only one solution $\sigma = -\beta|k_1|$, which correspond to a bounded solution in Ω. To find the relation of A and B we substitute this value of σ into the system, and solve for A, B yielding,

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} -\sigma \\ i\beta^2 k_1 \end{pmatrix} \qquad \sigma = -\beta|k_1|, \qquad \beta^2 > 0 \tag{11}$$

For the supersonic case, $\beta^2 < 0$, we have two solutions $\sigma_+ = i|\beta|k_1$ and $\sigma_- = -i|\beta|k_1$ corresponding to two bounded solution in Ω, and the coefficients A, B are given by

$$\begin{pmatrix} A_\pm \\ B_\pm \end{pmatrix} = \begin{pmatrix} -\sigma_\pm \\ i\beta^2 k_1 \end{pmatrix} \qquad \sigma_\pm = \pm i|\beta|k_1, \qquad \beta^2 < 0 \tag{12}$$

The solution for A, B given by (11) or (12) can be multiplied by an arbitrary constant. The actual value of that constant depends on the boundary condition which is required for the problem. This complete the solution of the problem in Ω.

Remark: In general, there may be several solutions for σ and to each of them there would correspond a vector solution such as (11) or (12) in the example above. The general solution is then a linear combination of these vectors with some weights to be determined from the boundary conditions. If the problem is well posed, then the number of boundary condition equals exactly to the number of bounded solutions that we are seeking and the analysis can be completed.

Some of the important question regarding general systems of equations can be answered to some extent using this tool. These include well-posedness of the problem, the choice of boundary conditions, their effect on the solutions and more.

3 On Pseudo-Differential Operators

Fourier analysis can be used to understand more complicated questions. For example, the relation of a function values to its normal derivative values on the boundary. Some relations between the quantities of interest may involve differential operators. Other relations may involve a more general class of operators, called pseudo-differential operators, which we briefly discuss here.

We consider elementary ideas from the theory of pseudo-differential operators which we demonstrate using some simple examples.

Example III Let ϕ be the solution of the equation

$$\begin{array}{ll} \Delta\phi = 0 & \Omega \\ \frac{\partial\phi}{\partial n} = \alpha & \partial\Omega, \end{array} \tag{13}$$

where the domain Ω is one of the two cases

$$\begin{array}{l} \Omega_2 = \{(x,z)|z > 0\} \\ \Omega_3 = \{(x,y,z)|z > 0\}, \end{array} \tag{14}$$

and $\frac{\partial}{\partial n}$ is the outward normal derivative at the boundary.

One of the questions we need to answer with regard to shape optimization or boundary control problems is the nature of the mappings between the control variable α, which in this case is a function defined on the boundary, and say the value of the solution on the boundary. That mapping is of course complicated and in general depends on the shape of the domain Ω. However, high frequencies in α have only local effect on the solution ϕ (this is a general property for elliptic equations which we exploit) and can be studied using Fourier techniques. We take

$$\alpha(x) = \exp(i\mathbf{k}\cdot\mathbf{x}) \tag{15}$$

where in case $\Omega = \Omega_3$ we take $\mathbf{k} = (k_1, k_2)$ and $\mathbf{x} = (x, y)$, and for $\Omega = \Omega_2$ we take $\mathbf{k} = k, \mathbf{x} = x$. Then ϕ is given by

$$\phi(\mathbf{x}, z) = A\exp(i\mathbf{k}\cdot\mathbf{x})\exp(\sigma z). \tag{16}$$

A substitution of this expression for ϕ into the interior equation in (13) implies the following equation for σ

$$\sigma^2 - |\mathbf{k}|^2 = 0, \tag{17}$$

and there are two solution

$$\phi_1(\mathbf{x}, z) = A\exp(i\mathbf{k}\cdot\mathbf{x})\exp(-|\mathbf{k}|z) \qquad \phi_2(\mathbf{x}, z) = A\exp(i\mathbf{k}\cdot\mathbf{x})\exp(|\mathbf{k}|z) \tag{18}$$

We look for a bounded solution in Ω and that is the one with $\exp(-|k|z)$. The above expression for ϕ satisfies the interior equations in (13) for any value of A, but only one value will also satisfy the boundary condition. Substituting (18) into that boundary condition, gives

$$A = \frac{1}{|\mathbf{k}|}. \tag{19}$$

The coefficient A that we have just found gives a very important information. It describes the mapping T between α and the values of ϕ on the boundary,

$$\phi|_{\partial\Omega} = T\alpha. \tag{20}$$

What we have just found is a description of the mapping T in the Fourier space, i.e., the symbol of T,

$$\hat{T}(\mathbf{k}) = \frac{1}{|\mathbf{k}|} = \begin{cases} \frac{1}{|k|} & \Omega = \Omega_2 \\ \frac{1}{\sqrt{k_1^2+k_2^2}} & \Omega = \Omega_3 \end{cases} \tag{21}$$

A Remark: For a general domain Ω we consider a small vicinity of a boundary point $x \in \partial\Omega$, and if the boundary is smooth at that point, we can flatten this piece of boundary by a proper transformation. We end up with a problem in half space, of the same form as above. The relation between α on the boundary, which is flat now, and ϕ inside the domain can be easily calculated using Fourier analysis.

3.1 The Symbol of an Operator

The operator T discussed above does not correspond to any differential operator. Only polynomials in $\mathbf{k}$ correspond to differential operators, and this is via a very simple relation. The differential operator $\partial/\partial x_j$ corresponds to the symbol (ik_j) since

$$\frac{\partial}{\partial x_j} \exp(i\mathbf{k}\cdot\mathbf{x}) = ik_j \exp(i\mathbf{k}\cdot\mathbf{x}). \tag{22}$$

From this we have,

$$\texttt{Symbol:} \quad \sum_{|\gamma|\le m} a_\gamma (i\mathbf{k})^\gamma \qquad \texttt{Differential Operator:} \quad \sum_{|\gamma|\le m} a_\gamma D^\gamma \tag{23}$$

where $i\mathbf{k} = (ik_1, \ldots, ik_n), \gamma = (\gamma_1, \ldots, \gamma_n), (i\mathbf{k})^\gamma = (ik_1)^{\gamma_1} \ldots (ik_n)^{\gamma_n}$ and $D^\gamma = (\partial^{\gamma_1}/\partial x_1^{\gamma_1}) \ldots (\partial^{\gamma_n}/\partial x_n^{\gamma_n})$.

For a general operator T, defined on functions in an infinite space, we define the symbol $\hat{T}(\mathbf{k})$ by

$$T \exp(i\mathbf{k}\cdot\mathbf{x}) = \hat{T}(\mathbf{k}) \exp(i\mathbf{k}\cdot\mathbf{x}). \tag{24}$$

This definition is for scalar PDE as well as for systems of PDE. In the second case the symbol will be a matrix whose elements are functions of $\mathbf{k}$. A definition in a bounded domain Ω can also be done, by considering a small vicinity of a point in Ω and a localization of the above.

Using the definition (24) we see that we can define a larger class of operators by considering symbols which are not polynomials. Under some assumptions which we omit here, one gets the class of pseudo-differential operators. They play a very important role in the study of boundary value problems for elliptic equations.

Example IV. We consider next the example

$$\begin{aligned} \Delta\phi = 0 \quad & \Omega \\ \phi = \alpha \quad & \partial\Omega \end{aligned} \tag{25}$$

where Ω is any of the two domains in the previous example. Following the same procedure as before,

$$\alpha(\mathbf{x}) = \exp(i\mathbf{k}\cdot\mathbf{x}) \tag{26}$$

$$\phi(\mathbf{x}, z) = A\exp(i\mathbf{k}\cdot\mathbf{x})\exp(\sigma z) \tag{27}$$

implies

$$\sigma^2 - |\mathbf{k}|^2 = 0 \tag{28}$$

and the relevant solution is

$$\phi(\mathbf{x}, z) = A\exp(i\mathbf{k}\cdot\mathbf{x})\exp(-|\mathbf{k}|z). \tag{29}$$

The value of A is determined from the boundary condition and clearly we have

$$A = 1. \tag{30}$$

The relation of $\frac{\partial\phi}{\partial n}_{|\partial\Omega}$ to the boundary values $\phi_{|\partial\Omega}$ is described by a mapping S

$$\frac{\partial\phi}{\partial n}_{|\partial\Omega} = S\alpha \tag{31}$$

whose symbol can be easily found by differentiating ϕ in the outward normal direction at the boundary, giving

$$\hat{S}(\mathbf{k}) = |\mathbf{k}|. \tag{32}$$

Notice that although we are dealing with differential problems, some relation between boundary values are not governed anymore by differential operators. The operators S, T from the last two examples are therefore not differential operators but pseudo-differential operators.

Shape design problems are related to boundary control problems, and therefore these type of operators play a very important role in shape optimization. They will help us to characterize the minimization problem in quantitative way that will allow the construction of very effective solvers.

4 Fourier Analysis For Optimization Problems

The above analysis using Fourier decomposition can also serve for the analysis of optimization problems. Of probably the main concern for us is to define optimization problems for our engineering tasks, that will be mathematically "good", or well-posed in mathematical terminology. We would like the problem to have a solution (existence), that the solution will be unique (uniqueness) and that the solution will depend in a continuous way on other parameters in the problem (continuous dependence). We will see that these properties of the problem, at least when considering the high frequency range, can be easily analyzed. The usual rigorous mathematical techniques for these question are very complex and may not be of a practical engineering use. Moreover, some important details which are of engineering importance are not present in the rigorous analysis, while they are present in the formal Fourier techniques.

The characterization of the minimizer for an optimization problem gives the equation

$$\nabla E(\alpha^*) = 0. \tag{33}$$

This is in general a nonlinear equation for the unknown α. Now lets say that we have an approximate solution α and we are seeking the correction $\tilde{\alpha}$ such that $\nabla E(\alpha+\tilde{\alpha}) = 0$. Using a Taylor expansion we see that $\tilde{\alpha}$ satisfies approximately the equation

$$\mathcal{H}\tilde{\alpha} = -\nabla E(\alpha) \tag{34}$$

where $\mathcal{H}$ is the Hessian of the functional.

If the design variable α can be decomposed in a Fourier series then important information can be obtain about the problem using Fourier analysis.

The symbol of the Hessian, $\hat{\mathcal{H}}(\mathbf{k})$ contains all of the necessary information for analyzing and designing optimization procedures. Notice that, at the vicinity of the minimum, the gradient of the functional, g, is linearly related to the error. In the Fourier space the relation is given by

$$\hat{\mathcal{H}}(\mathbf{k})\hat{\tilde{\alpha}}(\mathbf{k}) = -\hat{g}(\mathbf{k}) \tag{35}$$

where $\hat{\tilde{\alpha}}(\mathbf{k})$ and $\hat{g}(\mathbf{k})$ are the Fourier transforms of $\tilde{\alpha}, g$, respectively.

4.1 The Symbol of The Hessian

In order to get a quantitative description of the level curves of the cost functional and to be able to determine the structure of the functional near the minimum, a Fourier analysis of the Hessian of the functional is carried out. In many problems of engineering interest the design variable are associated with boundary quantities and the gradient of the cost function as well as the Hessian are quantities defined on part of the boundary as well.

We have seen in lecture 1 that the eigenvalue distribution of the Hessian plays an important role in convergence rates for the optimization problem. Moreover, the asymptotic behavior of the large eigenvalues of the Hessian is tightly related to the symbol of the Hessian. Its computation is therefore of practical importance.

In the next example we calculate the symbol of the Hessian for a control problem related to a shape design problem. We identify in this case the Hessian as a differential operator acting on functions defined on the boundary of the domain.

Example V Consider the following minimization problem,

$$\min_{\alpha} \frac{1}{2} \int_{\partial\Omega} (u - u^*)^2 dx \tag{36}$$

subject to

$$\begin{pmatrix} \beta^2 \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & -\frac{\partial}{\partial x} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{37}$$

with the boundary condition

$$v = \frac{\partial \alpha}{\partial x} \qquad \partial\Omega \tag{38}$$

where $\beta^2 = (1 - M^2)$ and $\Omega = \{(x, y) | y > 0\}$. We introduce adjoint variables (Lagrange multipliers) (λ, μ) which can be shown to satisfy

$$\begin{pmatrix} \beta^2 \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & -\frac{\partial}{\partial x} \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{39}$$

with the boundary condition

$$-\mu + (u - u^*) = 0 \qquad \partial\Omega. \tag{40}$$

At the minimum the following equation has to be satisfied

$$\frac{\partial \lambda}{\partial x} = 0 \qquad \partial\Omega. \tag{41}$$

The left hand side of this equation is the gradient of the functional subject to the PDE, and its behavior in the vicinity of the minimum needs to be analyzed. In order to do that we examine the perturbation in the solution as a result of a perturbation in the design function α. The linearity of the interior equation implies that the perturbation variables $\tilde{u}, \tilde{v}, \tilde{\lambda}, \tilde{\mu}$ satisfy the same equations as u, v, λ, μ in the interior of the domain, and the boundary conditions for them are

$$\begin{aligned} \tilde{v} - \tfrac{\partial \tilde{\alpha}}{\partial x} &= 0 \qquad \partial\Omega \\ \tilde{u} - \tilde{\mu} &= 0 \qquad \partial\Omega \end{aligned} \tag{42}$$

and the change in the gradient is given by

$$\frac{\partial \tilde{\lambda}}{\partial x}_{|\partial\Omega}. \tag{43}$$

The analysis continues by assuming $\tilde{\alpha}$ to be a Fourier component, that is,

$$\tilde{\alpha} = \exp(ikx). \tag{44}$$

From the boundary condition we get $\tilde{v}_{|\partial\Omega} = ik\exp(ikx)$ and using the interior equations (37) for $(\tilde{u}, \tilde{v})$ we conclude that

$$\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} = \begin{pmatrix} |k|/\beta \\ ik \end{pmatrix} e^{ikx} e^{-\beta|k|y}. \tag{45}$$

Using the boundary conditions for $\tilde{\mu}$ we get $\tilde{\mu}_{|\partial\Omega} = |k|/\beta\exp(ikx)$ and from the interior equations (39) for $(\tilde{\lambda}, \tilde{\mu})$ it is easy to see that

$$\begin{pmatrix} \tilde{\lambda} \\ \tilde{\mu} \end{pmatrix} = \begin{pmatrix} -ik/\beta^2 \\ |k|/\beta \end{pmatrix} e^{ikx} e^{-\beta|k|y}$$

Combining these results we obtain that the change in the gradient, corresponding to a change in the design variable by $\tilde{\alpha} = \exp(ikx)$ is

$$\frac{\partial \tilde{\lambda}}{\partial x}_{|\partial\Omega} = \frac{|k|^2}{\beta^2}\tilde{\alpha} \tag{46}$$

Thus, the symbol of the Hessian $\hat{\mathcal{H}}(k)$ is given by

$$\hat{\mathcal{H}}(k) = \frac{|k|^2}{\beta^2}. \tag{47}$$

This symbol correspond to the differential operator

$$\mathcal{H} = -\frac{1}{\beta^2}\frac{d^2}{dx^2}. \tag{48}$$

A General Remark: This analysis was performed on problems with constant coefficients, however, it is not limited to such cases. It can be applied to non-constant coefficients and nonlinear problems in general domains. In such cases one linearizes the problem (if it is nonlinear) and freezes coefficients at a point x_0 in the domain. A constant coefficient problem is obtained which describes the behavior of the problem in a small vicinity of that point. The validity of the resulting Fourier analysis for that problem is then restricted to a small vicinity. Usually, all expressions in the analysis will depend on the frozen coefficients and analyzing one point in the domain is enough to obtain the desired information about all points. The rigorous justification of this process is beyond the scope of our discussion here.

When considering the problem in a general domain one perform the analysis at a boundary point by transforming the vicinity of that point into 'half-space' and applying there the analysis presented above. Thus, smooth boundaries can be treated. This analysis is local and hence is relevant for high frequencies only. Low frequencies are affected by the shape of the boundary and cannot be analyzed using local techniques. However, in spite of this limitation, it is still a very useful tool for quantitative results regarding our problems.

5 Problems Classification

Using the symbol of the Hessian, we can classify problems according to the asymptotic behavior of $\hat{\mathcal{H}}(\mathbf{k})$ for large $|\mathbf{k}|$. We consider the equation

$$\mathcal{H}\alpha = f. \tag{49}$$

We want to classify problem into well-posed (good) problem, ill-posed (bad) problem, easy problems and difficult problems.

Well posed problems are characterized by having a unique solution that is stable to perturbation in the data of the problem. When we consider high frequencies, which is the range of frequencies where our analysis of the Hessian is accurate, the following property implies well posedness,

$$\hat{\mathcal{H}}(\mathbf{k}) = O(|\mathbf{k}|^{\gamma}) \qquad \gamma \geq 0. \tag{50}$$

Note that a small change in the design variable in the high frequency range causes large changes in the right hand side, or the gradient. Hence, small changes in the data results in small changes in the solution, and the solution has the desired stability properties. This is summarized by

$$|\hat{\alpha}(\mathbf{k})| \approx |\mathbf{k}|^{-\gamma}|\hat{g}(\mathbf{k})| \qquad\qquad \text{for large } |\mathbf{k}| \tag{51}$$

which follows from (35) and (50).

Ill-Posedness is referred to a case where the solution is not unique or that it is sensitive to data in the problem. Ill-posedness that result from the behavior of high frequencies can be characterized as

$$\hat{\mathcal{H}}(\mathbf{k}) = O(\frac{1}{|\mathbf{k}|^{\gamma}}) \qquad \gamma > 0. \tag{52}$$

To see that such a behavior causes ill-posedness, consider a solution α and a perturbation of it in the form $\alpha + c\exp(i\mathbf{k}\cdot\mathbf{x})$. The gradients evaluated for these two choices for the design variable are $\mathcal{H}\alpha$ and $\mathcal{H}\alpha + c\hat{\mathcal{H}}(\mathbf{k})\exp(i\mathbf{k}\cdot\mathbf{x})$. The latter can be approximated by

$$\mathcal{H}(\alpha + c\exp(i\mathbf{k}\cdot\mathbf{x})) = \mathcal{H}(\alpha) + c/|\mathbf{k}|^{\gamma}\exp(i\mathbf{k}\cdot\mathbf{x}) \approx \mathcal{H}\alpha \qquad \text{for large } |\mathbf{k}| \tag{53}$$

Since this is true for an arbitrary c and $|\mathbf{k}|$ sufficiently large, we get that if α is a solution then $\alpha + c\exp(i\mathbf{k}\cdot\mathbf{x})$ is an approximate solution for an arbitrary c and sufficiency large $|\mathbf{k}|$. This implies that small changes in the data of the problem will cause large changes in the solution. It is summarized in the relation

$$|\hat{\alpha}(\mathbf{k})| \approx |\mathbf{k}|^{\gamma}|\hat{g}(\mathbf{k})| \qquad \text{large } |\mathbf{k}| \tag{54}$$

which follows from (35) and (52), which shows that small changes in the gradient are amplified significantly in the design variables, for the high frequencies. Thus, high frequencies in the design variables are unstable.

The Discrete Problem. On a finite grid with mesh size h one consider $\mathbf{k}$ in the range $|\mathbf{k}| \leq \pi/h$. Thus, for well-posed problems satisfying (50), the eigenvalues of the Hessian corresponding to the highest frequencies behave as

$$\max_{|\mathbf{k}|\leq\pi/h} |\hat{\mathcal{H}}(\mathbf{k})| = O(\frac{1}{h^{\gamma}}). \tag{55}$$

Since the smallest eigenvalue of the discrete Hessian is given approximately by the corresponding eigenvalue of the differential Hessian, the condition number of the Hessian behaves as $O(\frac{1}{h^{\gamma}})$. This quantity is important in evaluating the performance of gradient based algorithms. As was mentioned in lecture 1, the convergence of gradient based methods is determined by $I - \delta\mathcal{H}$, and the high frequency components in the representation of the solution converge at a rate

$$\max_{\pi/(2h)\leq|\mathbf{k}|\leq\pi/h} |1 - \delta\hat{\mathcal{H}}(\mathbf{k})|. \tag{56}$$

The smallest eigenvalues of the Hessian are $O(1)$ being given approximately by their values from the continuous problem. From these observations we conclude that the expected rate of convergence for the full design problem is therefore

$$1 - O(h^{\gamma}), \tag{57}$$

and that the complexity of a given problem can be determined by the exponent γ.

- *Easy problems:* $0 \leq \gamma \ll 1$.
- *Difficult problems:* $\gamma \geq 1$

In summary, let the symbol of the Hessian satisfy

$$\hat{\mathcal{H}}(\mathbf{k}) = O(|\mathbf{k}|^{\gamma}) \tag{58}$$

then, we have the following

$$\begin{array}{lr} \texttt{Well-Posed ("good") optimization problems:} & \gamma \geq 0 \\ \texttt{Ill-Posed ("bad") Optimization problems:} & \gamma < 0 \\ \texttt{Easy optimization problems:} & 0 \leq \gamma \ll 1 \\ \texttt{Difficult optimization problems:} & \gamma \geq 1 \end{array} \tag{59}$$

6 Problem Reformulation

Unlike the constraint PDE which governs the optimization problem under study, the parameter space as well as the cost functional are not uniquely determined for a given engineering task. It is very likely that a proper choice for one or both can lead to well posed problems that are easier to solve.

We demonstrate this fact by an example. The problem given in section 4.1 is shown to have good stability properties for the high frequencies. However, $\gamma = 2$ for that problem, and the rate of convergence for gradient based methods is expected to be $1 - O(N^{-2})$ for N design variables since $h = O(\frac{1}{N})$. Different choices for design space and cost functional will be shown to have a very different behavior, although the engineering task remains roughly the same.

Example VI: Design Variables Reformulation Consider the minimization problem

$$\min_{\alpha} \frac{1}{2} \int (u - u^*)^2 dx \tag{60}$$

where (u, v) satisfy the equations

$$\begin{pmatrix} \beta^2 \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & -\frac{\partial}{\partial x} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \Omega \tag{61}$$

with the boundary condition

$$v = \alpha \qquad \partial\Omega \tag{62}$$

where the domain Ω and β are as before.

Following the same procedure as before, we introduce adjoint variables (Lagrange multipliers) (λ, μ) which can be shown to satisfy

$$-\begin{pmatrix} \beta^2 \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ \frac{\partial}{\partial y} & -\frac{\partial}{\partial x} \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{63}$$

with the boundary condition

$$-\mu + (u - u^*) = 0 \qquad \partial\Omega. \tag{64}$$

At the minimum the following equation has to be satisfied

$$-\lambda = 0 \qquad \partial\Omega, \tag{65}$$

and the gradient of the cost functional is given by

$$\nabla J = -\lambda_{|\partial\Omega}. \tag{66}$$

Following a similar derivation as before we find that the perturbations $\tilde{u}, \tilde{v}, \tilde{\lambda}, \tilde{\mu}$ corresponding to a perturbation $\tilde{\alpha} = \exp(ikx)$ are

$$\begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} = \begin{pmatrix} -ik/(\beta|k|) \\ 1 \end{pmatrix} e^{ikx} e^{-\beta|k|y} \tag{67}$$

and

$$\begin{pmatrix} \tilde{\lambda} \\ \tilde{\mu} \end{pmatrix} = \begin{pmatrix} -\frac{1}{\beta^2} \\ |k|/(i\beta k) \end{pmatrix} e^{ikx} e^{-\beta|k|y}$$

Combining these results we obtain that the change in the gradient, corresponding to a change in the design variable by $\tilde{\alpha} = \exp(ikx)$ is

$$\mathcal{H}\tilde{\alpha} = -\tilde{\lambda}_{|\partial\Omega} = \frac{1}{\beta^2}\exp(ikx) \tag{68}$$

Thus, the symbol of the Hessian is

$$\hat{\mathcal{H}}(k) = \frac{1}{\beta^2} \tag{69}$$

Remark. The boundary condition $v = \alpha_x$ correspond to a shape design problem where the shape is given by $\alpha(x)$, and the boundary condition there is $(u, v) \cdot \mathbf{n} = 0$. As a result of the above calculation we can derive the following conclusion. If the design variables are the slopes instead of the shape itself, a well-posed problem is still obtained. Moreover, this problem is much easier to solve that the one for the shape directly. Note that from the engineering point of view both problems can be used to perform the required design. In the second one, a reconstruction of the shape from the slopes has to be done and this itself is a stable problem.

References

[1] Arian E., Ta'asan S., Shape Optimization in One-Shot. Optimal Design and Control, Edited by J. Boggaard, J. Burkardt, M. Gunzburger, J. Peterson, Birkhauser Boston Inc. 1995

[2] Arian E., Ta'asan S., Multigrid One-Shot Methods for Optimal Control Problems: Infinite Dimensional Control. ICASE Report No. 94-52.

[3] E. Arian, S. Ta'asan, Analysis of the Hessian for Aerodynamics Optimization: Inviscid Flow. ICASE Report No. 96-28., submitted to Journal of Computational Physics

[4] Beux F., Dervieux A., A Hierarchical Approach for Shape Optimization, Inria Rapports de Recherche, N. 1868 (1993).

[5] Jameson A., Aerodynamics Design via Control Theory. J. Sci. Comp. Nov. 21, 1988

[6] Mandel B., Periaux J., Stoufflet B., Optimum Design Methods in Aerodynamics, AGARD-FDP-VKI Special Course, April 1994.

[7] Ta'asan S., One-Shot Methods for Optimal Control of Distributed Parameter Systems I: Finite Dimensional Control. ICASE Report No. 91-2, 1991

[8] S. Ta'asan, Trends in aerodynamics design and optimization: a mathematical view point. in Proceedings of the 12th AIAA Computational Fluid Dynamics Conference June 1 995, San Diego, CA. pp. 961-970 AIAA-95-1731-CP.

[9] Ta'asan S., Fast Solvers for MDO Problems. Multidisciplinary Design Optimization, state of the art. Edited by N.M. Alexandrov, M.Y. Hussaini, SIAM 1997.

[10] Ta'asan S., Salas M.D., Kuruvila G., Aerodynamics Design and Optimization in One-Shot. Proceedings of the AIAA 30th Aerospace Sciences and Exhibit, Jan 6-9 1992.

[11] Ta'asan S., Kuruvila G., Salas M.D., A New Approach to Aerodynamics Optimization. Proceedings of the First European Conference on Numerical Methods in Engineering. Sept 1992, Brussels Belguim.

[12] C. Hirsch, Numerical Computation of Internal and External Flows, Vol I, II, Wiley 1990.

[13] O. Pironneaux, Optimal Shape Design for Elliptic Systems, Springer Series in Computational Physics, 1983.

[14] A. Brandt, Multigrid Techniques: 1984 Guide with Applications to Fluid Dynamics. GMD-Studien.