Infinite Dimensional Preconditioners for Optimal Design Problems

Shlomo Ta'asan
Carnegie Mellon University

1 Introduction

In this lecture we discuss the use of elementary theory of pseudo-differential operators to analyze Hessians for optimization problems governed by partial differential equations. The role of this analysis is to come up with Hessian approximation that will accelerate gradient based methods. Quasi-Newton methods such as BFGS cannot cope efficiently with a very large dimension of the design space. Their efficiency is very good for a small dimensional design space, and it deteriorate as the dimension of the design space increases. The approach presented here is a complementing approach for the classical quasi-Newton methods. It uses the asymptotic behavior of the symbol of the Hessian to construct an accurate Hessian approximation for the high frequency range, in the representation of the design space. The resulting approximate Hessians are differential or pseudo-differential operators.

2 The Main Idea

Gradient based methods can be viewed as relaxation methods for the equation

$$\mathcal{H}\alpha = g \tag{1}$$

where g is the gradient and $\mathcal{H}$ is the Hessian of the functional considered. For example, a Jacobi relaxation for equation (1) has the form

$$\alpha \leftarrow \alpha + \delta(g - \mathcal{H}\alpha) \tag{2}$$

which is essentially the steepest descent method for minimizing the cost functional. The observation that the convergence rate for gradient descent methods is governed by $I - \delta\mathcal{H}$ suggests that effective Preconditioners can be constructed using the behavior of the symbol of the Hessian $\hat{\mathcal{H}}(\mathbf{k})$, for large $\mathbf{k}$. The idea is simple. Assume that

$$\hat{\mathcal{H}}(\mathbf{k}) = O(|\mathbf{k}|^{\gamma}) \qquad \text{for large } |\mathbf{k}| \tag{3}$$

and let $\mathcal{R}$ be an operator whose symbol satisfies

$$\hat{\mathcal{R}}(\mathbf{k}) = O(\frac{1}{|\mathbf{k}|^\gamma}) \qquad \text{for large } |\mathbf{k}|. \tag{4}$$

The behavior of the preconditioned method

$$\alpha \leftarrow \alpha - \delta \mathcal{R} g \tag{5}$$

is determined by

$$I - \delta \mathcal{R} \mathcal{H} \tag{6}$$

whose symbol

$$1 - \delta \hat{\mathcal{R}}(\mathbf{k}) \hat{\mathcal{H}}(\mathbf{k}) \tag{7}$$

approaches a constant for large $|\mathbf{k}|$. A proper choice of δ leads to a convergence rate which is independent of the dimensionality of the design space. This is not the case if the symbol of the iteration operator has some dependence on $\mathbf{k}$.

It is desired not to change the behavior of the low frequencies by the use of the preconditioner, since the analysis we do for the Hessian does not hold in the limit $|\mathbf{k}| \to 0$. That is, we would like the symbol of the preconditioner to satisfy also,

$$\hat{\mathcal{R}}(\mathbf{k}) \to 1 \qquad \text{for } |\mathbf{k}| \to 0. \tag{8}$$

2.1 Constructing The Preconditioner from Its Symbol

We come now to the question of constructing the preconditioner from its symbol. Since we are only interested in acceleration of certain numerical procedure, it is enough to use approximations for the true Hessians. Let us begin with the simplest examples. We have seen the correspondence between differential operators and symbols,

$$\texttt{Symbol:} \quad ik_j \qquad \qquad \texttt{Operator:} \quad \frac{\partial}{\partial x_j} \qquad j = 1, \ldots, 3 \tag{9}$$

and therefore

$$\texttt{Symbol:} \quad (ik_j)^m \qquad \qquad \texttt{Operator:} \quad \frac{\partial^m}{\partial x_j^m} \qquad j = 1, \ldots, 3 \tag{10}$$

Polynomials in ik_j, even in several dimensions, correspond to differential operators which are easily found as was shown in a previous lecture.

Example I. Consider the problem given in example V in lecture no. 2.

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

subject to

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

with the boundary condition

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

where $\beta^2 = (1 - M^2)$ and $\Omega = \{(x, y)|y > 0\}$. It was shown there that $\hat{\mathcal{H}}(k) = \frac{|k|^2}{\beta^2}$. This implies that

$$\mathcal{H}_{approx} = -\frac{1}{\beta^2}\frac{d^2}{dx^2} \tag{11}$$

An effective preconditioner $\mathcal{R}$ must satisfy $\hat{\mathcal{R}}(k) = \frac{\beta^2}{|k|^2}$ for large $|k|$ and this is obtained for

$$\mathcal{R}^{-1} = \mu I - \frac{1}{\beta^2}\frac{d^2}{dx^2} \tag{12}$$

The addition of the operator μI was to ensure that the preconditioner does not affect the low frequency range. A choice $\mu = 1$ can be taken although some approximation of the first eigenvalue can give a better choice. Thus, the implementation of a preconditioned iteration for that problem consist of repeated application of the two steps

$$\begin{array}{l} \mu\psi - \frac{1}{\beta^2}\frac{d^2}{dx^2}\psi = -\lambda_x \\ \alpha \leftarrow \alpha - \delta\psi \end{array} \tag{13}$$

where δ is found using a line search on coarse grids and, $\delta = 1$ on fine grids. Note that the construction of the preconditioner was done on the differential level but the numerical implementation is using some approximation of it, e.g., finite difference approximation.

A good discretization (h-elliptic) of the state equation uses staggered grid. We demonstrate it on a rectangular domain with a uniform grid of spacing h. Let the grid points be labeled $\{(i, j)|0 \le i \le N_x; 0 \le j \le N_y\}$. The discrete variables approximating u will be located at the middle of the vertical cell edges, i.e., will be parameterized as $u_{i,j+1/2}$. The discrete approximations to v will be located at the middle of the horizontal edges, i.e., parameterized by $v_{i+1/2,j}$. Discretization of the first equation is done at the cell centers and the second equations at the vertices, both using central differences. Design variables are located at the boundary nodes and the boundary condition is given by

$$v_{i+1/2,0} = \frac{1}{h}(\alpha_{i+1} - \alpha_i). \tag{14}$$

A calculation of the cost functional for the discrete problem requires the values of u on the boundary $j=0$. This is done by introducing ghost variables $u_{i+1/2,-1/2}$. An extra equation for these ghost values is introduced at the boundary nodes, approximating the second interior equation. We introduce adjoint variables (Lagrange multipliers) (λ, μ) discretized as $\lambda_{i+1/2,j+1/2}$ and $\mu_{i,j}$ with ghost points for λ. The adjoint variables satisfy the same equation as (u, v) but at points shifted by $(1/2, 1/2)$. A straightforward calculation shows that the gradient is given by

$$\frac{1}{2h}(\lambda_{i+1/2,1/2} - \lambda_{i-1/2,1/2} + \lambda_{i+1/2,-1/2} - \lambda_{i-1/2,-1/2}). \tag{15}$$

The discrete preconditioner is done as follows. Let $\psi_j, j = 1, \ldots, N$ be the solution of the discrete problem

$$\begin{aligned}\beta^2 h^2 \mu \psi_j - \psi_{j-1} + 2\psi_j - \psi_{j+1} = \; & -\tfrac{\beta^2 h}{2}(\lambda_{i+1/2,1/2} - \lambda_{i-1/2,1/2} \\ & +\lambda_{i+1/2,-1/2} - \lambda_{i-1/2,-1/2})\end{aligned} \tag{16}$$

for $j = 2, \ldots, N-1$, where h is the mesh size used for the discretization and $\psi_0 = \psi_N = 0$. The design variables are updated by

$$\alpha_j = \alpha_j - \delta\psi_j \qquad j = 1, \ldots, N \tag{17}$$

Note that applying the preconditioner requires the solution of a differential equation on the boundary where the control is given. This is a typical case. The equation defining the preconditioner is in one dimension less than the state and the costate equations.

Example II. We now move to a more challenging case which is the construction of an approximation to a Hessian with a symbol $\hat{\mathcal{H}}(\mathbf{k}) = |\mathbf{k}|$, and the problem is on the boundary of a domain in three space dimensions. Recall that in our lecture no 2 in this volume we have discussed the mapping

$$\phi_{|\Gamma} = T\frac{\partial \phi}{\partial n}_{|\Gamma}, \tag{18}$$

where ϕ is the solution of a Laplace equation in the domain Ω

$$\Delta\phi = 0, \tag{19}$$

and we have found that its symbol is

$$\hat{T}(\mathbf{k}) = |\mathbf{k}|. \tag{20}$$

The construction of an operator T from functions defined on the boundary of a domain, to functions defined on the same boundary, whose symbol is $|\mathbf{k}|$ is done as follows. Let g be a function defined on the boundary of Ω, we define Tg by

$$Tg = \frac{\partial \phi}{\partial n}_{|\partial\Omega}, \tag{21}$$

where

$$\begin{array}{ll} \Delta\phi = 0 & \Omega \\ \phi = g & \partial\Omega. \end{array} \tag{22}$$

Another case we consider is an operator S whose symbol is

$$\hat{S}(\mathbf{k}) = \frac{1}{|\mathbf{k}|}. \tag{23}$$

It can be approximated as

$$Sg = \phi_{|\partial\Omega} \tag{24}$$

where ϕ is the solution of

$$\begin{array}{ll} \Delta\phi = 0 & \Omega \\ \frac{\partial\phi}{\partial n} = g & \partial\Omega \end{array} \tag{25}$$

This follows from certain relations that we obtained in a previous lecture.

Example III. Here we construct an operator whose symbol is $i(ak_1 + bk_2)/|\mathbf{k}|$. We have a product of symbols, and each of them is something that we already know. A product of symbols correspond to applying the corresponding operators one after the other (with the proper order for systems of differential equations).

The symbol $i(ak_1 + bk_2)$ correspond to the operator $a\frac{\partial}{\partial t_1} + b\frac{\partial}{\partial t_2}$ where t_1, t_2 are the tangential coordinate corresponding to the wave directions k_1, k_2 respectively. Let ϕ be the solution of (25) then,

$$Tg = (a\frac{\partial}{\partial t_1} + b\frac{\partial}{\partial t_2})\phi_{|\partial\Omega} \tag{26}$$

has the desired symbol.

Remark: The operators that we have constructed in the last example are nonlocal, and one may construct also integral operators for them, with singular kernels. We prefer this approach since in the context of the optimal design problems one already has a (fast) solver for the equations needed for these pseudo-differential operators.

2.2 Preconditioners for Finite Dimensional Design Space

The previous section discussed the construction of preconditioners from their symbol in case the design space was a space of functions defined, for example, on the boundary of a domain. This is the infinite dimensional design space. In many applications one uses a fixed finite dimensional representation of the design space, using a set of prescribed shape functions. When the number of these functions is small one can use acceleration techniques such as BFGS. When that number grows and the number of BFGS steps required to solve the problem increases significantly, one may combine a preconditioner which is based on the infinite dimensional analysis.

We consider two design spaces. The first, $\mathcal{A}$, is a space of functions which is infinite dimensional and the second one, $\mathcal{A}_q$, is a subspace of the first and is represented in terms of q functions, $f_j, j = 1, \ldots, q$. We also assume that the set $f_j, j = 1, \ldots, q$ is orthonormal with respect to the usual L_2 inner product on the boundary,

$$\int_{\partial\Omega} f_j(s) f_k(s) ds = \delta_{j,k}, \tag{27}$$

where $\delta_{j,k}$ is the Kronecker delta. Functions in $\mathcal{A}_q$ are linear combination of the form $\sum_{j=1}^{d-1} \alpha_j f_j$, and the space $\mathcal{A}_q$ can be identified $I\!R^q$. We construct a mapping P from $\mathcal{A}$ to $I\!R^q$ by,

$$(P\alpha)_j = \int_{\partial\Omega} \alpha(s) f_j(s) ds \qquad j = 1, \ldots, q. \tag{28}$$

Note that the transpose of the operator P acts from the finite dimensional space $I\!R^q$ to $\mathcal{A}$ and is given by

$$(P^T \vec{\alpha})(s) = \sum_{j=1}^{d-1} \alpha_j f_j(s) \tag{29}$$

where $\vec{\alpha} = (\alpha_1, \ldots, \alpha_q)$.

Now we come to the point of relating gradients calculated with respect to the design space $\mathcal{A}$ to those calculated with respect to $I\!R^q$. Let

$$\delta J = \epsilon \int_{\partial\Omega} \tilde{\alpha}(s) g(s) ds + \frac{1}{2}\epsilon^2 \int_{\partial\Omega} (\mathcal{H}\tilde{\alpha})(s) \tilde{\alpha}(s) ds + O(\epsilon^3) \tag{30}$$

be the variation of the functional corresponding to a change in the design variables by $\epsilon\tilde{\alpha}$. The gradient with respect to $\mathcal{A}$ is certainly $\nabla J = g$, and the Hessian is $\mathcal{H}$. Now if the change in the design variables are done in the subspace $\mathcal{A}_q$, we consider $\tilde{\alpha} = \sum_{j=1}^{q} \tilde{\alpha}_j f_j$ and then a substitution into the above expression for δJ gives

$$\delta J = \epsilon \sum_{j=1}^{q} \tilde{\alpha}_j \int_{\partial\Omega} f_j(s) g(s) ds + \frac{1}{2}\epsilon^2 \sum_{j=1}^{q} \sum_{k=1}^{q} \tilde{\alpha}_k \tilde{\alpha}_j \int_{\partial\Omega} (\mathcal{H} f_j)(s) f_k(s) ds + O(\epsilon^3) \tag{31}$$

and in that case

$$\nabla J = \begin{pmatrix} g_1 \\ \vdots \\ g_q \end{pmatrix} = \vec{g} \tag{32}$$

where $\vec{g} = (g_1, \ldots, g_q)$, $g_j = \int_{\partial\Omega} f_j(s) g(s) ds$, and the Hessian for the subspace, $\mathcal{H}_q$, is related to the full Hessian $\mathcal{H}$ as

$$(\mathcal{H}_q)_{j,k} = \int_{\partial\Omega} (\mathcal{H} f_j)(s) f_k(s) ds \qquad j,k = 1, \ldots, q. \tag{33}$$

Notice that for the finite dimensional design space we can write,

$$\begin{aligned} &\tilde{\alpha}(s) = (P^T \vec{\alpha})(s) \\ &\nabla J = Pg \\ &\mathcal{H}_q = P \mathcal{H} P^T. \end{aligned} \tag{34}$$

These are the abstract formulas for the discrete quantities for the subspace as a function of the same quantities on the infinite dimensional space.

A preconditioner for the finite dimensional design can be obtained by constructing first the infinite dimensional preconditioner and then using the above formula to get $\mathcal{R}_q = \mathcal{H}_q^{-1}$. The preconditioned iteration is

$$\vec{\alpha} \leftarrow \vec{\alpha} - \delta \mathcal{R}_q \vec{g}. \tag{35}$$

Example Consider the case $\mathcal{H} = -\frac{1}{\beta^2} \frac{d^2}{dx^2}$ which appeared in one of the previous lectures. The finite dimensional preconditioner is constructed from the inverse of the finite dimensional Hessian

$$(\mathcal{H}_q)_{j,k} = -\frac{1}{\beta^2} \int_{\partial\Omega} f_j(s) \frac{d^2 f_k}{dx^2}(s) ds. \tag{36}$$

3 Application to Shape Design: Fluid Dynamics

Application of the ideas discussed in the previous sections to examples of shape design problems for fluid dynamics models will be demonstrated, For simplicity, we use the small disturbance potential equations and the Linearized Euler equations. Discussion starting from a truly shape design problems is given in Arian and Ta'asan [3].

3.1 Small Disturbance Potential Equation

Let $\Omega = \{(x, y, z) | z > 0\}$ and let ϕ satisfy

$$\begin{array}{ll} (1 - M^2)\phi_{xx} + \phi_{yy} + \phi_{zz} = 0 & \Omega \\ -\frac{\partial \phi}{\partial z} = u_0 \alpha_x & \partial\Omega \end{array} \tag{37}$$

and consider the minimization problem

$$\min_{\alpha} \frac{1}{2} \int_{\partial\Omega} (\rho_0 u_0 \phi_x - f^*)^2. \tag{38}$$

This problem is related to a shape design problem governed by the Full Potential equation in a general domain, using the cost functional $\frac{1}{2}\int_{\partial\Omega}(p-p^*)^2 ds$.

It can be shown using standard computation, as explained in a previous lecture, that if λ satisfies the equation

$$\begin{array}{ll} (1-M^2)\lambda_{xx}+\lambda_{yy}+\lambda_{zz}=0 & \Omega \\ \frac{\partial\lambda}{\partial z}-\rho_0 u_0(\rho_0 u_0\phi_x-f^*)_x=0 & \partial\Omega \end{array} \tag{39}$$

then the gradient of the functional is given by

$$\nabla_\alpha J=-u_0\frac{\partial\lambda}{\partial x}. \tag{40}$$

We would like to compute the Hessian for this problem and to construct an infinite dimensional preconditioner for it. We assume a perturbation in α of the form

$$\tilde{\alpha}(\mathbf{x})=\hat{\alpha}(\mathbf{k})\exp{(i\mathbf{k}\cdot\mathbf{x})} \tag{41}$$

where $\mathbf{x}=(x,y)$ and $\mathbf{k}=(k_1,k_2)$, and then the corresponding change in ϕ is

$$\tilde{\phi}(\mathbf{x},z)=\hat{\phi}(\mathbf{k})\exp{(i\mathbf{k}\cdot\mathbf{x})}\exp{(\sigma z)}, \tag{42}$$

and the change in the adjoint variable is

$$\tilde{\lambda}(\mathbf{x},z)=\hat{\lambda}(\mathbf{k})\exp{(i\mathbf{k}\cdot\mathbf{x})}\exp{(\bar{\sigma}z)} \tag{43}$$

where $\sigma,\bar{\sigma}$ are solutions for the following algebraic equation,

$$\begin{array}{l} -(1-M^2)k_1^2-k_2^2+\sigma^2=0 \\ -(1-M^2)k_1^2-k_2^2+\bar{\sigma}^2=0. \end{array} \tag{44}$$

In these expressions k_1,k_2 are given, as well as $\hat{\alpha}(\mathbf{k})$ which amount to perturbing the shape by one frequency with a given amplitude.

The Choice of $\sigma,\bar{\sigma}$. There is a nontrivial point with respect to the choice of the roots that needs some explanation. In the subsonic case, $M<1$, we have two real roots for σ. One is negative and correspond to a bounded solution in Ω, the other is positive and correspond to an unbounded solution for ϕ,λ and it is discarded in our analysis. In the supersonic case, $M>1$, and the expression $(1-M^2)k_1^2+k_2^2$ may be either positive or negative. If it is positive we take for σ the root mentioned above. If it is negative then σ has two imaginary roots. One of these correspond to an incident wave and the other to a reflected wave. The perturbation in the incident wave is zero, since this wave comes from infinity and there was no change there. The reflected wave arise from the change in shape. In the supersonic regime, at the outflow there are no boundary conditions for ϕ, while the adjoint variable λ has two boundary conditions. Therefore, for the perturbation in the adjoint variable no waves are going toward the outflow. This means that the sign of $\bar{\sigma}$ is opposite to that of σ in the supersonic case when wave solutions exist. Therefore, the roots are

$$\begin{array}{ll} \sigma = \bar{\sigma} = -\sqrt{(1-M^2)k_1^2 + k_2^2} & (1-M^2)k_1^2 + k_2^2 > 0 \\ \sigma = -\bar{\sigma} = i\sqrt{-(1-M^2)k_1^2 - k_2^2} & (1-M^2)k_1^2 + k_2^2 \leq 0. \end{array} \tag{45}$$

From the boundary condition for ϕ and λ we see that the following relations hold,

$$\begin{array}{c} -\sigma\hat{\phi}(\mathbf{k}) = ik_1 u_0 \hat{\alpha}(\mathbf{k}) \\ \bar{\sigma}\hat{\lambda}(\mathbf{k}) = \rho_0^2 u_0^2 k_1^2 \hat{\phi}(\mathbf{k}) \end{array} \tag{46}$$

and from these the change in the gradient as a result of a change in α is given by $-\tilde{\lambda}_x$ and therefore,

$$\hat{\mathcal{H}}(\mathbf{k}) = \rho_0^2 u_0^4 \frac{k_1^4}{\bar{\sigma}\sigma}. \tag{47}$$

Now notice that

$$\begin{array}{lll} \bar{\sigma}\sigma & = (1-M^2)k_1^2 + k_2^2 & (1-M^2)k_1^2 + k_2^2 > 0 \\ \bar{\sigma}\sigma & = -(1-M^2)k_1^2 - k_2^2 & (1-M^2)k_1^2 + k_2^2 \leq 0 \end{array} \tag{48}$$

and the two can be combined as

$$\bar{\sigma}\sigma = |(1-M^2)k_1^2 + k_2^2|. \tag{49}$$

In summary, the symbol of the Hessian is

$$\hat{\mathcal{H}}(\mathbf{k}) = \rho_0^2 u_0^4 \frac{k_1^4}{|(1-M^2)k_1^2 + k_2^2|}. \tag{50}$$

3.1.1 Two Dimensional Case

In this case k_2 does not exists and we have a symbol

$$\hat{\mathcal{H}}(k_1) = \frac{k_1^2}{|1-M^2|}. \tag{51}$$

This is the symbol of the differential operator

$$-\frac{1}{|1-M^2|}\frac{\partial^2}{\partial x^2}. \tag{52}$$

A preconditioned gradient descent method have the form

$$\begin{array}{c} \mu\psi - \frac{1}{|1-M^2|}\frac{d^2}{dx^2}\psi = -u_0\frac{\partial\lambda}{\partial x} \\ \alpha \leftarrow \alpha - \delta\psi \end{array} \tag{53}$$

That is, we have to solve an ODE on the boundary with the gradient as a source term, before using it as a direction of change for the design variable. According to our analysis the new method converges at a rate with is independent of the number of design variables, since the symbol for the modified iteration does not depend on **k**. Note that the construction of the preconditioner was done on the differential level but the numerical implementation is using some approximation of it, e.g., finite difference approximation.

3.1.2 Three Dimensional Case

We distinguish here two cases, the purely subsonic case and the general case which may include transonic regimes.

Purely Subsonic Case. In this case $1 - M^2 > 0$ and hence $|(1-M^2)k_1^2 + k_2^2| = (1-M^2)k_1^2 + k_2^2$. The preconditioner symbol is

$$\hat{\mathcal{R}}(\mathbf{k}) = \left[(1-M^2)k_1^2 + k_2^2\right]/(\rho_0^2 u_0^4 k_1^4). \tag{54}$$

The preconditioned iteration will have in the Fourier space the direction

$$\hat{\tilde{\alpha}}(\mathbf{k}) = \hat{\mathcal{R}}(\mathbf{k})\hat{g}(\mathbf{k}) \tag{55}$$

which after some rearrangements reads as

$$\rho_0^2 u_0^4 k_1^4 \hat{\tilde{\alpha}}(\mathbf{k}) = ((1-M^2)k_1^2 + k_2^2)\hat{g}(\mathbf{k}). \tag{56}$$

This equation in the Fourier space can be translated into the following differential equation for the change in the design variable,

$$\rho_0^2 u_0^4 \frac{\partial^4 \tilde{\alpha}}{\partial x^4} = -(1-M^2)\frac{\partial^2 g}{\partial x^2} - \frac{\partial^2 g}{\partial y^2}. \tag{57}$$

Actually, since our analysis was accurate only for the high frequency changes, we may not want to use that preconditioning for the very smooth components in the solution. This suggest a combination of the standard gradient descent method and this preconditioning, which for example, can be employed as

$$\mu\tilde{\alpha} + \rho_0^2 u_0^4 \frac{\partial^4 \tilde{\alpha}}{\partial x^4} = -(1-M^2)\frac{\partial^2 g}{\partial x^2} - \frac{\partial^2 g}{\partial y^2} + \mu g \tag{58}$$

The addition of the μ term is so that the low frequency range will not be affected by this preconditioner and would just use the gradient direction. High frequency on the other hand, are accurately analyzed by our method and should use the above preconditioner. It is also possible to use BFGS method in conjunction with the infinite dimensional preconditioner developed here.

Supersonic and Transonic Cases. In this case the term $|(1-M^2)k_1^2+k_2^2|$ cannot be simplified and we have to treat a certain pseudo differential operator. To approximate $|(1-M^2)k_1^2+k_2^2|$ we use the relation

$$\bar{\sigma}\sigma = |(1-M^2)k_1^2+k_2^2| \tag{59}$$

which was derived before, using the interior equations. We want to derive an implementation in real space of the equation whose form in the Fourier space is

$$\mu\hat{\alpha}(\mathbf{k}) + \rho_0^2 u_0^4 k_1^4 \alpha(\mathbf{k}) = |(1-M^2)k_1^2+k_2^2|\hat{g}(\mathbf{k}) + \mu\hat{g}(\mathbf{k}).$$

The symbol $\bar{\sigma}\sigma$ represent two normal derivatives to solution of the small disturbance equation we started with, one with the ϕ equation and the other with the λ equation. The difference between the two is at the far field boundary condition which is responsible for the proper choice in $\sigma, \bar{\sigma}$. The operator whose symbol is $|(1-M^2)k_1^2+k_2^2|$ is therefore constructed in two step.

$$\begin{aligned} (1-M^2)\psi_{xx} + \psi_{yy} + \psi_{zz} = 0 \quad & \Omega \\ \psi = g \quad & \partial\Omega \end{aligned} \tag{60}$$

$$\begin{aligned} (1-M^2)\bar{\psi}_{xx} + \bar{\psi}_{yy} + \bar{\psi}_{zz} = 0 \quad & \Omega \\ \bar{\psi} = \tfrac{\partial\psi}{\partial z} \quad & \partial\Omega \end{aligned} \tag{61}$$

and the full preconditioned direction for $\tilde{\alpha}$ is therefore

$$\mu\tilde{\alpha} + \rho_0^2 u_0^4 \frac{\partial^4\tilde{\alpha}}{\partial x^4} = \frac{\partial\bar{\psi}}{\partial z} + \mu g \tag{62}$$

where $\bar{\psi}, \psi$ satisfy the above equations.

3.2 The Linearized Euler Equations

Consider the linearized Euler equation around a mean flow $(\rho_0, \rho_0 u_0, 0, 0, p_0)$,

$$\begin{pmatrix} Q & \rho_0\partial_x & \rho_0\partial_y & \rho_0\partial_z & 0 \\ 0 & Q & 0 & 0 & \frac{1}{\rho_0}\partial_x \\ 0 & 0 & Q & 0 & \frac{1}{\rho_0}\partial_y \\ 0 & 0 & 0 & Q & \frac{1}{\rho_0}\partial_z \\ 0 & \rho_0 c_0^2\partial_x & \rho_0 c_0^2\partial_y & \rho_0 c_0^2\partial_z & Q \end{pmatrix} \begin{pmatrix} \rho \\ u \\ v \\ w \\ p \end{pmatrix} = 0 \tag{63}$$

in a domain $\Omega = \{(x,y,z)|z \geq 0\}$, where $Q = \vec{U_0}\cdot\vec{\nabla}$ ($\vec{U_0} \equiv (u_0, 0, 0)$ denotes the velocity vector), with the solid wall boundary condition

$$w = u_0\alpha_x \qquad \partial\Omega. \tag{64}$$

It is assumed that the problem was obtained by a linearization in a vicinity of a boundary point, and that the far field boundary conditions were given in terms of characteristic variables, which are not used explicitly in the derivation of the approximate Hessian. The minimization problem is

$$\min_{\alpha} \frac{1}{2} \int_{\partial\Omega} (p - p^*)^2 dx. \tag{65}$$

If a change $\tilde{\alpha}$ produces a change $\tilde{p}$ in the pressure then, the variation in this functional can be written as

$$\delta J = \int_{\partial\Omega} (p - p^*)\tilde{p} dx + \frac{1}{2} \int_{\partial\Omega} \tilde{p}^2 dx + O(\|\tilde{p}\|^3) \tag{66}$$

We calculate the Hessian in a slightly different way than before to illustrate another approach. If one can express the quadratic term in $\tilde{p}$ in terms of $\tilde{\alpha}$ one can identify the Hessian. That is,

$$\int \tilde{p}^2 dx = \int (\mathcal{H}\tilde{\alpha})\tilde{\alpha} dx. \tag{67}$$

This means that we can calculate the Hessian without going through the adjoint variable. We need to express $\tilde{p}$ in terms of $\tilde{\alpha}$, and we do it in the Fourier space. From the boundary condition at the wall

$$\hat{\tilde{w}}(\mathbf{k}) = ik_1 u_0 \hat{\tilde{\alpha}}(\mathbf{k}). \tag{68}$$

The calculation of $\hat{\tilde{p}}(\mathbf{k})$ in terms of $\hat{\tilde{\alpha}}(\mathbf{k})$ is done by solving the system of the linearized Euler equation with the above boundary condition for $\tilde{w}$. We look for solution of the form

$$\tilde{U} = \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \end{pmatrix} \exp(i\mathbf{k}\cdot\mathbf{x}) \exp(ik_3 z). \tag{69}$$

The term ik_3 here is the analog of our σ in the previous example. It is more convenient here due to the form of the symbol of the full equation. The following relation follows by substituting the above expression for $\tilde{U}$ into the Linearized Euler equations (63),

$$\hat{L}(\mathbf{k}) \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \end{pmatrix} = \begin{pmatrix} u_0 k_1 & \rho_0 k_1 & \rho_0 k_2 & \rho_0 k_3 & 0 \\ 0 & \rho_0 u_0 k_1 & 0 & 0 & k_1 \\ 0 & 0 & \rho_0 u_0 k_1 & 0 & k_2 \\ 0 & 0 & 0 & \rho_0 u_0 k_1 & k_3 \\ 0 & \rho_0 c_0^2 k_1 & \rho_0 c_0^2 k_2 & \rho_0 c_0^2 k_3 & u_0 k_1 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \\ A_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{70}$$

This is a linear system for $(A_1, \ldots, A_5)$ and it has a nontrivial solution when the determinant is zero,

$$det\hat{L}(\mathbf{k}) = k_1^3(u_0^2k_1^2 - c_0^2(k_1^2 + k_2^2 + k_3^2)) = 0 \tag{71}$$

Note that there are five solutions for this equations. Each of them has a corresponding solution for the vector $(A_1, \ldots, A_5)$,

$$\begin{array}{lr} V_1 = (\rho_0 u_0 c_0^2 k_1, -k_1, -k_2, -k_3, \rho_0 u_0 k_1) & k_3 = \sigma_1 \\ V_2 = (\rho_0 u_0 c_0^2 k_1, -k_1, -k_2, k_3, \rho_0 u_0 k_1) & k_3 = \sigma_2 \\ V_3 = (1, 0, 0, 0, 0,) & k_1 = 0 \\ V_4 = (0, 1, 0, 0, 0,) & k_1 = 0 \\ V_5 = (0, 0, -k_3, k_2, 0) & k_1 = 0 \end{array} \tag{72}$$

where

$$\begin{array}{ll} \sigma_1 = -\sigma_2 = \sqrt{(1-M^2)k_1^2 + k_2^2} & (1-M^2)k_1^2 + k_2^2 > 0 \\ \sigma_1 = -\sigma_2 = -i\sqrt{-(1-M^2)k_1^2 - k_2^2} & (1-M^2)k_1^2 + k_2^2 \leq 0 \end{array} \tag{73}$$

Note that for the subsonic case σ_1 correspond to the bounded solution, while σ_2 to the unbounded one. When $(1-M^2)k_1^2 + k_2^2 < 0$ we have two bounded solution. In that case σ_2 correspond to the incident wave and therefore its amplitude is zero for the perturbation variables. Thus, we are left with σ_1 for both subsonic and supersonic cases. The three solution corresponding to $k_1 = 0$ are not important for our analysis since they do not affect the changes in pressure (see the corresponding eigenvectors).

To summarize, only V_1 contributes to the pressure changes as a result of changes to the design variables by $\tilde{\alpha}$. The solution for $\tilde{U}$ is given by AV_1 for some scalar A. The w component in this solution is $-Ak_3$ and this must equal to $iu_0k_1\hat{\tilde{\alpha}}(\mathbf{k})$ form the boundary condition which in the Fourier space is given by (68). From that we find $A = -iu_0\frac{k_1}{k_3}\hat{\tilde{\alpha}}(\mathbf{k})$. Thus the solution is,

$$\tilde{U} = -iu_0\frac{k_1}{k_3}\hat{\tilde{\alpha}}(\mathbf{k}) \begin{pmatrix} \rho_0 u_0 c_0^2 k_1 \\ -k_1 \\ -k_2 \\ -k_3 \\ \rho_0 u_0 k_1 \end{pmatrix} \tag{74}$$

The last component in this vector gives us the change in the pressure

$$\hat{\tilde{p}}(\mathbf{k}) = -i\rho_0 u_0^2 \frac{k_1^2}{k_3}\hat{\tilde{\alpha}}(\mathbf{k}) \tag{75}$$

and from this we get

$$|\hat{\tilde{p}}(\mathbf{k})|^2 = \rho_0^2 u_0^4 \frac{k_1^4}{k_3\bar{k}_3}|\hat{\tilde{\alpha}}(\mathbf{k})|^2. \tag{76}$$

Notice that we have taken the complex conjugate of $\hat{\bar{p}}(\mathbf{k})$, and since k_3 is a complex number its conjugate was taken as well. Since $k_3\bar{k}_3 = |(1-M^2)k_1^2 + k_2^2|$ we obtain the symbol of the Hessian in the form,

$$\hat{\mathcal{H}}(\mathbf{k}) = \rho_0^2 u_0^4 \frac{k_1^4}{|(1-M^2)k_1^2 + k_2^2|}. \tag{77}$$

A preconditioner for this problem is done exactly as in the small disturbance equations using (60)-(62). It is also possible to construct the preconditioner based on solution of the linearized Euler equations, but is more complicated and unnecessary. The gradient g appearing in (60)-(62) has to be changed to the gradient for this problem, using the adjoint formulation.

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.