SPICE Workshop
September 7, 2005

Smolenice, Slovakia

Tutorial on Boundary Error

Robert J. Geller

Department of Earth and Planetary Science, Univ. of Tokyo

bob@eps.s.u-tokyo.ac.jp

DSM Theory I

The Direct Solution Method (DSM, Geller and Ohminato, 1994) is a Galerkin weak form method for solving the elastic equation of motion in the frequency domain

$$\left(\omega^2 \mathbf{T} - \mathbf{H}\right) \mathbf{c} = -\mathbf{g},$$

where ω is the frequency, $\mathbf{T}$ is the mass (kinetic energy) matrix, $\mathbf{H}$ is the stiffness (potential energy) matrix, $\mathbf{c}$ is the vector of expansion coefficients for the trial functions, and $\mathbf{g}$ is the force vector.

The matrix and vector elements are as follows:

$$T_{mn} = \int_V (\phi_i^{(m)})^* \rho \, \phi_i^{(n)} \, dV$$

$$H_{mn} = \int_V (\phi_{i,j}^{(m)})^* C_{ijkl} \, \phi_{k,l}^{(n)} \, dV$$

$$g_m = \int_V (\phi_i^{(m)})^* \, f_i \, dV,$$

where $\phi_i{}^{(m)}$ is the i-component of the m-th trial function, "$,j$" denotes spatial differentiation with respect to the j-coordinate, ρ is the density, C_{ijkl} is the elastic modulus, (We assume anelastic attenuation is included in the elastic moduli.) f_i is the external body force, and $*$ denotes complex conjugation.

The summation convention applies to subscripts corresponding to physical (x, y or z)-coordinates, but not to indices corresponding to abstract vector spaces, such as those denoting trial functions. The displacement is represented as a linear combination of the trial functions:

$$u_i = \sum_n c_n \phi_i^{(n)}.$$

The operators defined above will not in general be exact. We formally denote the exact operators by $\mathbf{T}^{(\mathbf{0})}$ and $\mathbf{H}^{(\mathbf{0})}$, the exact solution by $\mathbf{c}^{(\mathbf{0})}$, the error of the numerical operators by $\delta\mathbf{T}$ and $\delta\mathbf{H}$, and the error of the numerical solution by $\delta\mathbf{c}$,

$$\mathbf{T} = \mathbf{T}^{(0)} + \delta\mathbf{T}$$

$$\mathbf{H} = \mathbf{H}^{(0)} + \delta\mathbf{H}$$

$$\mathbf{c} = \mathbf{c}^{(0)} + \delta\mathbf{c}.$$

In order to make a formal error estimate we will now represent the exact solution, numerical solution, and error of the numerical solution by eigenfunction expansions. It is not necessary to know the actual numerical values of the eigenvalues and eigenvectors; also, the following results are independent of the basis.

The normal modes satisfy

$$\left(\omega_m^2 \mathbf{T}^{(0)} - \mathbf{H}^{(0)}\right) \mathbf{c_m} = 0,$$

where ω_m is the eigenfrequency of the m-th mode, and $\mathbf{c_m}$ is the eigenvector. We assume that the modes are orthonormalized so that

$$\mathbf{c_m}^* \mathbf{H}^{(0)} \mathbf{c_n} = \omega_m^2 \mathbf{c_m}^* \mathbf{T}^{(0)} \mathbf{c_n} = \omega_m^2 \delta_{mn},$$

where δ_{mn} is a Kronecker-delta.

Error Estimates I

The exact equation of motion can be formally written as follows:

$$\left(\omega^2 \mathbf{T}^{(0)} - \mathbf{H}^{(0)}\right) \mathbf{c}^{(0)} = -\mathbf{g}.$$

The error of the solutions, $\delta\mathbf{c}$, can be estimated using the first order Born approximation:

$$\left(\omega^2 \mathbf{T}^{(0)} - \mathbf{H}^{(0)}\right) \delta\mathbf{c} = -\left(\omega^2 \delta\mathbf{T} - \delta\mathbf{H}\right) \mathbf{c}^{(0)}.$$

Error Estimates II

We represent the exact solution in terms of an eigenfunction expansion:

$$\mathbf{c}^{(\mathbf{0})} = \sum_m d_m^{(0)} \mathbf{c_m}.$$

The expansion coefficient of the m-th mode is given by

$$d_m^{(0)} = -g_m/(\omega^2 - \omega_m^2),$$

where

$$g_m = \mathbf{c_m}^* \mathbf{g}.$$

The denominator of the r.h.s. will be small, and thus $d_m^{(0)}$ will be large, when ω is close to ω_m but can be neglected elsewhere.

Error Estimates III

We also represent the error of the numerical solution in terms of an eigenfunction expansion:

$$\delta \mathbf{c} = \sum_m \delta d_m \mathbf{c_m}.$$

The expansion coefficient of the m-th mode is given by

$$\delta d_m = -\frac{\sum_n (\omega^2 \delta T_{mn} - \delta H_{mn}) d_n^{(0)}}{(\omega^2 - \omega_m^2)}.$$

Error Estimates IV

By the same argument used above, the m-component of the error (i.e., δd_m) will be large only when ω is close to ω_m. But in the vicinity of $\omega = \omega_m$ only $d_m^{(0)}$ will be large; the expansion coefficients of all the other modes will be negligible. Therefore in the vicinity of $\omega = \omega_m$ the $n \neq m$ terms in the summation can be neglected. The expansion coefficient of the m-th mode for the numerical error is therefore approximately given by

$$\delta d_m = -\frac{(\omega^2 \delta T_{mm} - \delta H_{mm}) d_m^{(0)}}{(\omega^2 - \omega_m^2)}.$$

Error Estimates V

The relative error of the numerical solution in the vicinity of $\omega = \omega_m$ can thus be approximated by

$$\frac{\delta d_m}{d_m^{(0)}} = -\frac{(\omega^2 \delta T_{mm} - \delta H_{mm})}{(\omega^2 - \omega_m^2)} = -\frac{\delta T_{mm}(\omega^2 - \delta H_{mm}/\delta T_{mm})}{(\omega^2 - \omega_m^2)}.$$

Note that in the cases of interest ω will be real, but ω_m will include a small imaginary part due to anelastic attenuation. Therefore, the denominators in the above will become small as $\omega \rightarrow \omega_m$, but will never exactly equal zero.

The Criterion

The above equation shows that in general the relative error will greatly increase as ω approaches ω_m. However, if the numerator is also proportional to $(\omega^2 - \omega_m^2)$ the relative error will not worsen appreciably as $\omega \rightarrow \omega_m$. Such proportionality can be achieved if and only if the errors of the numerical operators approximately satisfy

$$\omega_m^2 \delta T_{mm} - \delta H_{mm} = 0$$

for every mode.

Errors for OPT operators

The above equation shows that the relative error of the numerical solution obtained using modified operators that approximately satisfy the criterion is approximately given by:

$$\left| \frac{\delta d_m}{d_m^{(0)}} \right| = |\delta T_{mm}| \,,$$

even as $\omega \rightarrow \omega_m$. On the other hand, if the operators do not approximately satisfy the criterion the relative error will worsen drastically as $\omega \rightarrow \omega_m$.

Boundary errors I

As shown above, the spatial pattern of the error of the numerical solution in the vicinity of $\omega = \omega_m$ is given by the eigenfunction of the m-th mode. A large numerical error of the operators near a free surface or internal discontinuity does not mean that the error of the numerical solution will be unusually large there; it simply means that the scalar quantity

$$\frac{\omega^2 \delta T_{mm} - \delta H_{mm}}{\omega^2 - \omega_m^2},$$

which, as shown by the criterion, is the factor controlling the relative error of the expansion coefficient of the m-th mode in the vicinity of $\omega = \omega_m$, will be somewhat larger.

Boundary errors II

This means that the relative error of the numerical solution will be uniformly larger everywhere in the medium, rather than that the error will be especially large at some particular points. Thus we do not have to be especially concerned with reducing the error of the operators at particular points, even points where we particularly desire accurate solutions.

Intuitive Explanation

Near any frequency ω where the solution is large, there will be a dominant mode, whose frequency is ω_m and whose eigenfunction is $\mathbf{c}_m$.

To first order, the exact solution, numerical solution, and **the error of the numerical solution** will **all** be proportional to $\mathbf{c}_m$.

Therefore an **operator error** at the boundary will cause a contribution to the **solution error** which is (to lowest order) proportional to $\mathbf{c}_m$.

The above discussion is conducted in the frequency domain, but applies equally to the time domain.