## From Linear to Nonlinear: The Generalized-$\alpha$ Method:

We can express the Galerkin semi-discrete formulation in a compact form using the parlance of matrix algebra. Let:

$$\underline{R} = [R_P] = \text{residual vector}$$

$$\underline{d}(t) = [d_Q(t)] = \text{displacement vector}$$

$$\underline{d}_0 = [d_{0Q}] = \text{initial displacement vector}$$

$$\dot{\underline{d}}_0 = [\dot{d}_{0Q}] = \text{initial velocity vector}$$

where, with $P = ID(A,i)$ and $Q = ID(B,j)$, we have defined:

$$R_P = \eta(\hat{e}_A N_i ; \vec{u}^h, \vec{u}^h_{,tt}) - L(\hat{e}_A N_i)$$

$$d_Q(t) = (\vec{d}_j(t))_B$$

We can also accomodate Dirichlet boundary conditions with the residual vector. Defining BC in the same manner as in linear elastodynamics, we set:

$$R_P = (\vec{d}_i)_{A,tt} - (\vec{g}_i)_{A,tt} \qquad \text{if } BC(A,i) = 1$$

where $P = ID(A,i)$. Finally, we write $\underline{R} = \underline{R}(\underline{d}, \dot{\underline{d}}, \ddot{\underline{d}})$ to denote its explicit dependence on the displacement, velocity, and acceleration fields. With the above notation, we can rewrite (G) as follows:

$$\boxed{\begin{aligned} \underline{R}(\underline{d}(t), \dot{\underline{d}}(t), \ddot{\underline{d}}(t)) &= \underline{0}, \qquad t \in (0,T) \\ \underline{d}(0) &= \underline{d}_0 \\ \dot{\underline{d}}(0) &= \dot{\underline{d}}_0 \end{aligned}}$$

The above is a set of nonlinear ordinary differential equations for the coefficients $d_P(t)$. To solve this system, we employ the generalized-$\alpha$ method which is the generalization of HHT-$\alpha$ method allowing an exact control of the numerical damping.

The starting point of the generalized-$\alpha$ method is the satisfaction of the Newmark formulas:

$$\underline{d}_{n+1} = \underline{d}_n + \Delta t\, \underline{v}_n + \frac{(\Delta t)^2}{2}\left((1-2\beta)\,\underline{a}_n + 2\beta\, \underline{a}_{n+1}\right)$$

$$\underline{v}_{n+1} = \underline{v}_n + \Delta t\left((1-\gamma)\,\underline{a}_n + \gamma\, \underline{a}_{n+1}\right)$$

where $\underline{a}_{n+1} \approx \ddot{\underline{d}}(t_{n+1})$, $\underline{v}_{n+1} \approx \dot{\underline{d}}(t_{n+1})$, and $\underline{d}_{n+1} \approx \underline{d}(t_{n+1})$ and $\gamma$ and $\beta$ are tuneable parameters. To close our system, we must solve the equations of motion. In the generalized-$\alpha$ method, we solve:

$$\underline{R}(\underline{d}_{n+\alpha_f}, \underline{v}_{n+\alpha_f}, \underline{a}_{n+\alpha_m}) = \underline{0}$$

where:

$$\underline{d}_{n+\alpha_f} = \underline{d}_n + \alpha_f (\underline{d}_{n+1} - \underline{d}_n)$$

$$\underline{v}_{n+\alpha_f} = \underline{v}_n + \alpha_f (\underline{v}_{n+1} - \underline{v}_n)$$

$$\underline{a}_{n+\alpha_m} = \underline{a}_n + \alpha_m (\underline{a}_{n+1} - \underline{a}_n)$$

and $\alpha_f$ and $\alpha_m$ are again tuneable parameters. The full method is then as follows:

Generalized-$\alpha$ Method

Given $(\underline{d}_n, \underline{v}_n, \underline{a}_n)$, find $(\underline{d}_{n+1}, \underline{v}_{n+1}, \underline{a}_{n+1})$ s.t.:

$$\underline{R}(\underline{d}_{n+\alpha_f}, \underline{v}_{n+\alpha_f}, \underline{a}_{n+\alpha_m}) = \underline{0}$$

$$\underline{d}_{n+\alpha_f} = \underline{d}_n + \alpha_f (\underline{d}_{n+1} - \underline{d}_n)$$

$$\underline{v}_{n+\alpha_f} = \underline{v}_n + \alpha_f (\underline{v}_{n+1} - \underline{v}_n)$$

$$\underline{a}_{n+\alpha_m} = \underline{a}_n + \alpha_m (\underline{a}_{n+1} - \underline{a}_n)$$

$$\underline{v}_{n+1} = \underline{v}_n + \Delta t\,((1-\gamma)\,\underline{a}_n + \gamma\,\underline{a}_{n+1})$$

$$\underline{d}_{n+1} = \underline{d}_n + \Delta t\,\underline{v}_n + \frac{(\Delta t)^2}{2}((1-2\beta)\,\underline{a}_n + 2\beta\,\underline{a}_{n+1})$$

Note that we have four free parameters. We can reduce this number to two by requiring second-order accuracy and unconditional stability. Second-order accuracy is attained if:

$$\gamma = \frac{1}{2} - \alpha_f + \alpha_m$$

and:

$$\beta = \frac{1}{4}(1 - \alpha_f + \alpha_m)^2$$

while unconditional stability requires:

$$\alpha_m \geq \alpha_f \geq \frac{1}{2}$$

Note that Newmark's method is obtained if $\alpha_m = \alpha_f = 1$ and the HHT-$\alpha$ method is attained if $\alpha_m = 1$ and $\alpha_f = \alpha$. However, $\alpha_m$ and $\alpha_f$ can actually be optimally tuned to obtain desired numerical damping characteristics. We postpone this discussion until later.

Note that the generalized-$\alpha$ method gives a nonlinear algebraic system that must be solved at every time step. We utilize Newton's method to solve this system. To review,

Suppose that we are trying to find the root (or a root, if it is not unique) $x^*$ of a differentiable, nonlinear function $F(x)$ in one dimension. Beginning with $i=0$, we will construct a sequence of approximations $x^i$ such that $x^i \to x^*$ and $F(x^*)=0$. At any step $i$, we know that $x^* = x^i + \Delta x^i$ for some unknown $\Delta x^i$. Though unknown, we can use a Taylor expansion of $F$ to find an approximation to $\Delta x^i$. Namely:

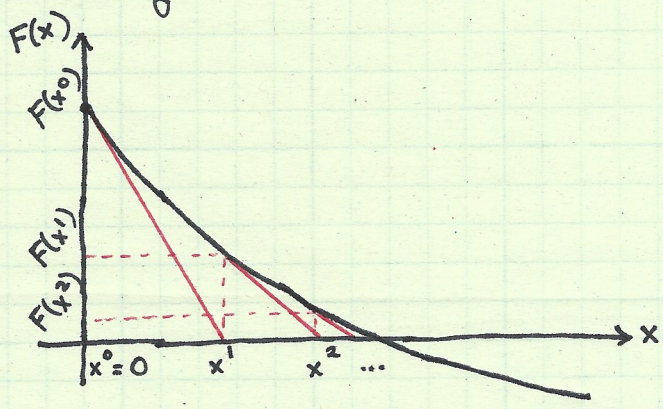
$$0 = F(x^*) = F(x^i + \Delta x^i)$$
$$\approx F(x^i) + \frac{dF}{dx}(x^i)\,\Delta x^i$$

$$\therefore \Delta x^i \approx -\left(\frac{dF}{dx}(x^i)\right)^{-1} F(x^i)$$

We then take:

$$x^{i+1} = x^i - \left(\frac{dF}{dx}(x^i)\right)^{-1} F(x^i)$$

and repeat the process. We continue to iterate until $F(x^i) \leq \varepsilon\,|F(x^0)|$ for some predetermined tolerance, $\varepsilon$. Visually:




In the context of solving the generalized-$\alpha$ equations, we view Newton's method as a two-phase predictor-multicorrector algorithm. There are various choices for the predictor. For nonlinear solid mechanics, the constant displacement predictor is preferred:

Predictor Phase: Set:

$$\overset{0}{\underline{d}}_{n+1} = \underline{d}_n$$
$$\overset{0}{\underline{a}}_{n+1} = -\frac{1}{\beta \Delta t}\,\underline{v}_n - \frac{(1-2\beta)}{2\beta}\,\underline{a}_n$$
$$\overset{0}{\underline{v}}_{n+1} = \underline{v}_n + \Delta t\left((1-\gamma)\,\underline{a}_n + \gamma\,\overset{0}{\underline{a}}_{n+1}\right)$$

Multicorrector Phase: Repeat the following steps for $i = 0, 1, 2, \ldots, i_{max}$ or until convergence is achieved.

1. Evaluate iterates at the intermediate time levels as:

$$\underline{d}^i_{n+\alpha_f} = \underline{d}_n + \alpha_f \left( \underline{d}^i_{n+1} - \underline{d}_n \right)$$

$$\underline{v}^i_{n+\alpha_f} = \underline{v}_n + \alpha_f \left( \underline{v}^i_{n+1} - \underline{v}_n \right)$$

$$\underline{a}^i_{n+\alpha_m} = \underline{a}_n + \alpha_m \left( \underline{a}^i_{n+1} - \underline{a}_n \right)$$

2. Use the intermediate solutions to assemble the residuals and the corresponding matrices in the linear system:

$$\frac{d\underline{R}^i}{d\underline{a}_{n+1}} \Delta \underline{a} = -\underline{R}^i_{n+1}$$

where:

$$\underline{R}^i_{n+1} = \underline{R}\left( \underline{d}^i_{n+\alpha_f}, \underline{v}^i_{n+\alpha_f}, \underline{a}^i_{n+\alpha_m} \right)$$

and:

$$\frac{d\underline{R}^i}{d\underline{a}_{n+1}} = \frac{d\underline{R}}{d\underline{a}_{n+1}} \left( \underline{d}^i_{n+\alpha_f}, \underline{v}^i_{n+\alpha_f}, \underline{a}^i_{n+\alpha_m} \right)$$

3. Having solved the linear system, update the iterates as:

$$\underline{a}^{i+1}_{n+1} = \underline{a}^i_{n+1} + \Delta \underline{a}$$

$$\underline{v}^{i+1}_{n+1} = \underline{v}^i_{n+1} + \gamma \Delta t \, \Delta \underline{a}$$

$$\underline{d}^{i+1}_{n+1} = \underline{d}^i_{n+1} + \beta (\Delta t)^2 \Delta \underline{a}$$

To obtain the tangent matrix $d\underline{R}^i / d\underline{a}_{n+1}$, we use repeated applications of the chain rule:

$$\frac{d\underline{R}^i}{d\underline{a}_{n+1}} = \frac{\partial \underline{R}}{\partial \underline{a}_{n+\alpha_m}} \overset{\alpha_m}{\frac{\partial \underline{a}_{n+\alpha_m}}{\partial \underline{a}_{n+1}}} + \frac{\partial \underline{R}}{\partial \underline{v}_{n+\alpha_f}} \overset{\alpha_f}{\frac{\partial \underline{v}_{n+\alpha_f}}{\partial \underline{v}_{n+1}}} \overset{\gamma \Delta t}{\frac{\partial \underline{v}_{n+1}}{\partial \underline{a}_{n+1}}} + \frac{\partial \underline{R}}{\partial \underline{d}_{n+\alpha_f}} \overset{\alpha_f}{\frac{\partial \underline{d}_{n+\alpha_f}}{\partial \underline{d}_{n+1}}} \overset{\beta (\Delta t)^2}{\frac{\partial \underline{d}_{n+1}}{\partial \underline{a}_{n+1}}}$$

$$= \alpha_m \underline{\underline{M}} + \alpha_f \gamma \Delta t \, \underline{\underline{C}} + \alpha_f \beta (\Delta t)^2 \underline{\underline{K}}$$

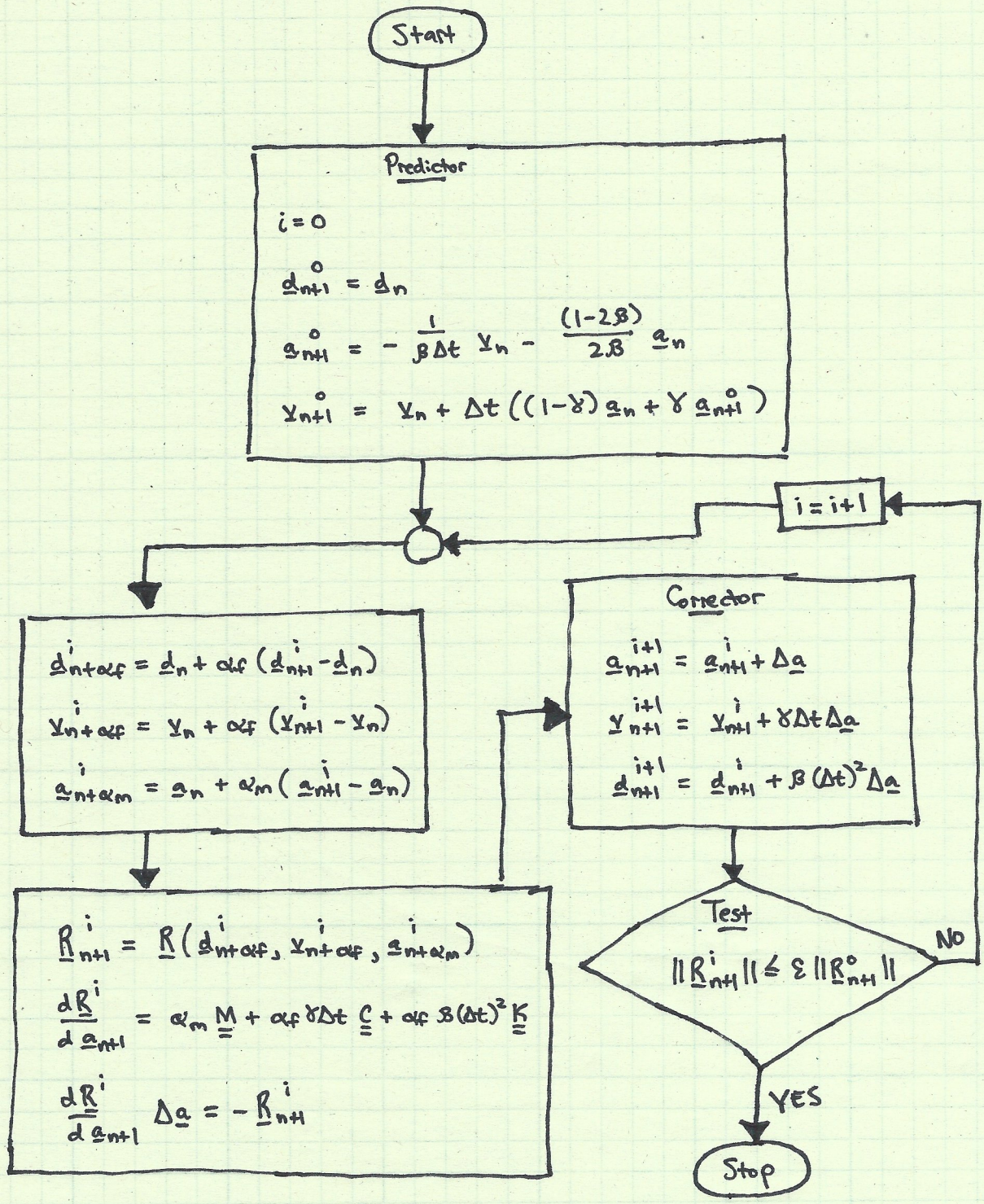
where:

$$\underline{\underline{M}} \equiv \frac{\partial \underline{R}}{\partial \underline{a}_{n+\alpha_m}} \left( \underline{d}^i_{n+\alpha_f}, \underline{v}^i_{n+\alpha_f}, \underline{a}^i_{n+\alpha_m} \right)$$

$$\underline{\underline{C}} \equiv \frac{\partial \underline{R}}{\partial \underline{v}_{n+\alpha_f}} \left( \underline{d}^i_{n+\alpha_f}, \underline{v}^i_{n+\alpha_f}, \underline{a}^i_{n+\alpha_m} \right)$$

$$\underline{\underline{K}} \equiv \frac{\partial \underline{R}}{\partial \underline{d}_{n+\alpha_f}} \left( \underline{d}^i_{n+\alpha_f}, \underline{v}^i_{n+\alpha_f}, \underline{a}^i_{n+\alpha_m} \right)$$

A flow chart for the predictor - multicorrector algorithm is given on the following page:




The most difficult part of the above algorithm is construction of $\underline{\underline{M}}$, $\underline{\underline{C}}$, and $\underline{\underline{K}}$. For the dynamics of a hyperelastic body, we have:

$$\underline{\underline{M}} = [M_{PQ}] \qquad M_{PQ} = (N_i \hat{e}_A, \rho_0 N_j \hat{e}_B)_{\Omega_0}$$

$$\underline{\underline{C}} = [C_{PQ}] \quad \text{with} \quad C_{PQ} = 0$$

$$\underline{\underline{K}} = [K_{PQ}] \qquad K_{PQ} \text{ depends on material}$$

where $P = ID(A,i)$ and $Q = ID(B,j)$ (note: Dirichlet BCs behave differently).

To compute the stiffness $\underline{\underline{K}}$, we must specify the constitutive equation for the stress. Suppose that:

$$\vec{\vec{S}} = 2\frac{\partial \Psi}{\partial \vec{\vec{C}}}$$

as in the case of a Neo-Hookean material. Then, we have:

$$K_{PQ} = \frac{\partial R_P}{\partial d_Q}\left(\underline{d}^{i}_{n+\alpha_f}, \underline{x}^{i}_{n+\alpha_f}, \underline{a}^{i}_{n+\alpha_m}\right)$$

$$= \frac{\partial}{\partial d_Q}\left(\int_{\Omega_0}\left(\left(\vec{\nabla}_X(N_i\hat{e}_A)\right):\vec{\vec{P}}\right)d\Omega_0\right)$$

$$= \frac{\partial}{\partial d_Q}\left(\int_{\Omega_0}\left(\left(\vec{\nabla}_X(N_i\hat{e}_A)\right):\left(\vec{\vec{F}}\vec{\vec{S}}\right)\right)d\Omega_0\right)$$

$$= \frac{\partial}{\partial d_Q}\left(\int_{\Omega_0}\left(\left(\vec{\nabla}_X(N_i\hat{e}_A)\right):\left(\vec{\vec{F}}\,2\frac{\partial \Psi}{\partial \vec{\vec{C}}}\right)\right)d\Omega_0\right)$$

where $P = ID(A,i)$, $Q = ID(B,j)$, $d_Q = (\vec{d}_j)_B$, $\vec{\vec{F}} = \vec{\vec{F}}(\underline{d}^{i}_{n+\alpha_f})$, and $\Psi = \Psi(\underline{d}^{i}_{n+\alpha_f})$.

Continuing, we have:

$$K_{PQ} = \int_{\Omega_0}\left(\vec{\nabla}_X(N_i\hat{e}_A)\right):\left(\frac{\partial}{\partial d_Q}\left(2\vec{\vec{F}}\frac{\partial \Psi}{\partial \vec{\vec{C}}}\right)\right)d\Omega_0$$

The product rule gives:

$$\frac{\partial}{\partial d_Q}\left(2\vec{\vec{F}}\frac{\partial \Psi}{\partial \vec{\vec{C}}}\right) = 2\frac{\partial \vec{\vec{F}}}{\partial d_Q}\frac{\partial \Psi}{\partial \vec{\vec{C}}} + 2\vec{\vec{F}}\frac{\partial}{\partial d_Q}\left(\frac{\partial \Psi}{\partial \vec{\vec{C}}}\right)$$

$$= \frac{\partial \vec{\vec{F}}}{\partial d_Q}\vec{\vec{S}} + 2\vec{\vec{F}}\frac{\partial}{\partial d_Q}\left(\frac{\partial \Psi}{\partial \vec{\vec{C}}}\right)$$

It is easily seen that:

$$\frac{\partial \vec{\vec{F}}}{\partial d_Q} = \vec{\nabla}_X(N_j\hat{e}_B) \quad \text{with } Q = ID(B,j)$$

By the chain rule:

$$\frac{\partial}{\partial d_Q}\left(\frac{\partial \Psi}{\partial C_{IJ}}\right) = \frac{\partial^2 \Psi}{\partial C_{IJ}\partial C_{KL}}\frac{\partial C_{KL}}{\partial F_{MN}}\frac{\partial F_{MN}}{\partial d_Q}$$

However:

$$\frac{\partial C_{KL}}{\partial F_{MN}} = \delta_{NK} F_{ML} + F_{MK} \delta_{NL}$$

Defining:

$$\mathbb{C}_{IJKL} = \frac{\partial^2 \psi}{\partial C_{IJ} \partial C_{KL}}$$

Remark: $\mathbb{C}_{IJKL}$ satisfies:

$$\mathbb{C}_{IJKL} = \mathbb{C}_{JIKL}$$
$$\mathbb{C}_{IJKL} = \mathbb{C}_{IJLK}$$
$$\mathbb{C}_{IJKL} = \mathbb{C}_{KLIJ}$$

we thus have:

$$\frac{\partial}{\partial d_Q}\left(\frac{\partial \psi}{\partial C_{IJ}}\right) = \mathbb{C}_{IJKL}\left(\delta_{NK} F_{ML} + F_{MK}\delta_{NL}\right)\frac{\partial F_{MN}}{\partial d_Q}$$

$$= \mathbb{C}_{IJKL}\left(F_{ML}\frac{\partial F_{MK}}{\partial d_Q} + F_{MK}\frac{\partial F_{ML}}{\partial d_Q}\right)$$

$$= 2\,\mathbb{C}_{IJKL} F_{ML}\frac{\partial F_{MK}}{\partial d_Q}$$

where we have exploited the minor symmetry $\mathbb{C}_{IJKL} = \mathbb{C}_{IJLK}$. It follows that:

$$\frac{\partial}{\partial d_Q}\left(\frac{\partial \psi}{\partial \vec{\vec{C}}}\right) = 2\,\vec{\vec{\vec{\vec{\mathbb{C}}}}}\,\vec{\vec{F}}^T \vec{\nabla}_X (N_j \hat{e}_B)$$

Therefore:

$$\frac{\partial}{\partial d_Q}\left(2\vec{\vec{F}}\frac{\partial \psi}{\partial \vec{\vec{C}}}\right) = \vec{\nabla}_X (N_j \hat{e}_B)\,\vec{\vec{S}} + 4\,\vec{\vec{F}}\,\vec{\vec{\vec{\vec{\mathbb{C}}}}}\,\vec{\vec{F}}^T \vec{\nabla}_X (N_j \hat{e}_B)$$

and thus:

$$\boxed{\begin{aligned} K_{PQ} = &\int_{\Omega_0} \left(\vec{\nabla}_X (N_i \hat{e}_A)\right) : \left(\vec{\nabla}_X (N_j \hat{e}_B)\,\vec{\vec{S}}\right) d\Omega_0 \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{Geometric Stiffness} \\ &+ \int_{\Omega_0} \left(\vec{\nabla}_X (N_i \hat{e}_A)\right) : \left(4\,\vec{\vec{F}}\,\vec{\vec{\vec{\vec{\mathbb{C}}}}}\,\vec{\vec{F}}^T \vec{\nabla}_X (N_j \hat{e}_B)\right) d\Omega_0 \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{Material Stiffness} \end{aligned}}$$

The stiffness matrix $K_{PQ}$ is symmetric by construction. The above definition is closed once $\psi$ is specified.

Now that we have finished our discussion of implementation, let us return to the choice of $\alpha_m$ and $\alpha_F$. The proceeding discussion appeared in the first paper on generalized-$\alpha$:

J. Chung, G. M. Hulbert, "A Time-Integration Algorithm for Structural Dynamics with Improved Numerical Dissipation: The Generalized-$\alpha$ Method," J. Appl. Mech., 60: 371-375. (1993)

In this paper, application of the generalized-$\alpha$ method to the model problem:

$$\ddot{u} + 2\xi\omega\dot{u} + \omega^2 u = f \qquad \text{for } t \in (0,T)$$
$$u(0) = u_0$$
$$\dot{u}(0) = v_0$$

was studied in detail where $\omega$ denotes the frequency of the 1-DOF system and $\xi$ denotes the damping ratio. To study stability, it suffices to consider $f = 0$. Then the generalized-$\alpha$ method applied to the above system takes the form:

$$\underline{X}_{n+1} = \underline{\underline{A}}\,\underline{X}_n$$

in which:

$$\underline{X}_n = \begin{bmatrix} d_n \\ \Delta t\, v_n \\ (\Delta t)^2 a_n \end{bmatrix}$$

$$\underline{\underline{A}} = \begin{bmatrix} \text{The So-Called} \\ \text{Amplification} \\ \text{Matrix} \end{bmatrix} = \underline{\underline{A}}(\Delta t, \gamma, \beta, \alpha_m, \alpha_f)$$

The stability of the generalized-$\alpha$ method is inherently related to the eigenvalues of $\underline{\underline{A}}$. Notably, the generalized-$\alpha$ method is unconditionally stable if the spectral radius:

$$\rho(\underline{\underline{A}}) = \max(|\lambda_1|, |\lambda_2|, |\lambda_3|)$$

satisfies $|\rho| \leq 1$ where $\lambda_1, \lambda_2$, and $\lambda_3$ are the eigenvalues of $\underline{\underline{A}}$. This condition is satisfied provided:

$$\alpha_m \geq \alpha_f \geq 1, \quad \beta \geq \tfrac{1}{4}(1 - \alpha_f + \alpha_m)^2$$

The spectral radius is also a measure of numerical dissipation. A smaller spectral radius corresponds to greater numerical dissipation.

Now, suppose $\gamma = \frac{1}{2} - \alpha_f + \alpha_m$ and $\beta = \frac{1}{4}(1 - \alpha_f + \alpha_m)^2$. Then there are two free parameters $\alpha_m$ and $\alpha_f$. In their seminal work, Chung and Hulbert showed that $\alpha_m$ and $\alpha_f$ can be chosen such that the spectral radius at infinite time step satisfies:

$$\rho_\infty = |\lambda_1| = |\lambda_2| = |\lambda_3|$$

That is, all three modes dissipate at the same rate. This results in a method with an optimal combination of high-frequency and low-frequency dissipation. For the second-order system in consideration, this optimal combination manifests when:

$$\alpha_m = \frac{2 - \rho_\infty}{1 + \rho_\infty}, \quad \alpha_f = \frac{1}{1 + \rho_\infty}$$

Thus we are left with a method with one free parameter $\rho_\infty$, the spectral radius at infinitely large time step. When $\rho_\infty = 1$, there is no damping, and when $\rho_\infty = 0$, there is maximal damping.

For first-order systems, such as those arising in fluid mechanics, the optimal $\alpha_m$ and $\alpha_f$ values are different. Instead, they are equal to:

$$\alpha_m = \frac{1}{2}\left(\frac{3-\rho_\infty}{1+\rho_\infty}\right), \quad \alpha_f = \frac{1}{1+\rho_\infty}$$

We will use these values when applying isogeometric analysis to incompressible fluid flow.