To newton multigrid method,it's error propagation matrix T is defined as below:

$$x_0^{n+1} = x_0^n + e_0^n \tag{1}$$

$$r_0^n = b - A_0 x_0^n \tag{2}$$

$$A_1 = P^T A_0 P \tag{3}$$

$$r_1^n = P r_0^n \tag{4}$$

$$e_1^n = [A_1]^{-1} r_1^n \tag{5}$$

$$\begin{aligned} e_0^n &= P^T e_1^n \\ &= P^T [A_1]^{-1} r_1^n \\ &= P^T [P^T A_0 P]^{-1} P [b - A_0 x_0] \end{aligned} \tag{6}$$

$$T = I - [P(P^T A_0 P)^{-1} P^T] A \tag{7}$$

T's spectral radius smaller,speed of convergence quicker.
Newton multigrid is a method to treat linear system.As to nonlinear system,we use fas method.
To fas method,it's loss is defined as below:

$$b = \frac{\partial E}{\partial x}, A = \frac{\partial^2 E}{\partial x^2} \tag{8}$$

if given x,we can compute A and b in coarse mesh:

$$x_0^{n+1} = x_0^n + e_0^n \tag{9}$$

$$x_1^n = \widetilde{P} x_0^n \tag{10}$$

Notice that $\widetilde{P}$ is differ from P , $\widetilde{P}$ is used to interpolate the position of points to another layer, and P is used to interpolate the residuals to another layer.

$$A_1 = A(x_1^n) = A(\widetilde{P} x_0^n) \tag{11}$$

$$r_0^n = F(x_0^n) \tag{12}$$

$$r_1^n = P r_0^n = P F(x_0^n) \tag{13}$$

$$e_1^n = [A_1]^{-1} r_1^n \tag{14}$$

$$\begin{aligned} e_0^n &= P^T e_1^n \\ &= P^T [A_1]^{-1} r_1^n \\ &= P^T [A((\widetilde{P} x_0^n)]^{-1} P F(x_0) \end{aligned} \tag{15}$$

So we get x_0^{n+1},and we can use it to compute $F(x_0^{n+1})$. We hope the norm of $F(x_0^{n+1})$ to be as small as possible.

There are two matrices to modify, $\widetilde{P}$ and P.