nzsee
NEW ZEALAND SOCIETY FOR EARTHQUAKE ENGINEERING

Source of numerical non-convergence in the analysis of bilinear SDOF systems

N. Morris, R. Chandramohan & C.R. McGann
University of Canterbury, Christchurch, New Zealand.

ABSTRACT

Numerical non-convergence is an important factor impeding the research into and widespread adoption of the nonlinear dynamic analysis. Upon encountering non-convergence at an analysis time step when using an implicit time integration scheme, heuristic methods are typically employed to enforce convergence in order to proceed with the analysis. These heuristic methods are, however, not guaranteed to always work. This study investigates the convergence behaviour of bilinear SDOF systems analysed using implicit Newmark time integration schemes. The source of numerical non-convergence is observed to be related to the rounding and truncation of floating-point operations. Finally, an algorithm is developed to compute the lower limit of the convergence tolerance to be used, in order to completely eliminate the likelihood of encountering numerical non-convergence in the dynamic analysis of a bilinear SDOF system.

1 INTRODUCTION

A notable roadblock when performing nonlinear dynamic analysis using an implicit time integration scheme is numerical non-convergence. Numerical non-convergence causes the analysis to stop prematurely, rather than calculating the full response of the system to the applied load, typically an earthquake ground motion. Heuristic methods exist which can be used when non-convergence is encountered, for instance, using a different solution algorithm, decreasing the time step, or increasing the tolerance. These methods, however, do not always prevent non-convergence. In the SDOF systems of this study it was found that increasing the tolerance does work, provided the tolerance is increased above a specific value, rather than arbitrarily.

This study investigated SDOF systems, with simple piecewise hysteretic behaviour. Mitigation methods were formulated to eliminate the likelihood of encountering numerical non-convergence when performing a nonlinear time history analysis on a bilinear inelastic SDOF system. By understanding the causes of non-convergence in SDOF systems, the potential causes of non-convergence in more complex SDOF and MDOF systems can be further investigated.

A set of over one million inelastic SDOF systems was created and analysed to investigate the causes of non-convergence. The systems were designed to have different combinations of natural periods (ranging from 0.1 s to 5s), strength reduction factors (ranging from 1 to 8), and damping ratios (ranging from 0% to 15%). Three different piecewise linear models were also used to represent the hysteretic behaviour of the inelastic spring: an elastic-perfectly-plastic model, a bilinear strain hardening model, and a trilinear strain softening model, which had 0 stiffness once the spring force reached 0 kN on the softening curve. The El Centro ground motion from the 1940 Imperial Valley earthquake in California was used to analyse all the systems. The average acceleration and linear acceleration time integration schemes, from the Newmark family of schemes (Newmark 1959), were used to conduct the analyses.

2 DYNAMIC ANALYSIS OF AN SDOF SYSTEM USING A NEWMARK SCHEME

For a nonlinear SDOF system, the equation of motion is

$$m\ddot{u}(t) + c\dot{u}(t) + f_s(t) = p(t) \tag{1}$$

where $u(t)$, $\dot{u}(t)$, and $\ddot{u}(t)$ are the time-varying displacement, velocity, and acceleration, respectively, m is the mass, c is the damping coefficient, $f_s(t)$ is the time-varying spring force, and $p(t)$ is the time-varying external force. Numerical time integration schemes are commonly used to solve this equation of motion. Newmark's constant average acceleration and linear acceleration schemes, considered in this study, are two of the most widely used schemes (Chopra 2020 §5.4.2). The governing equations of the Newmark family of schemes help write the equation of motion at time step i+1 as

$$\alpha_1 u_{i+1} + (f_s)_{i+1} = \hat{p}_{i+1} \tag{2}$$

where

$$\hat{p}_{i+1} = p_{i+1} + \alpha_1 u_i + \alpha_2 \dot{u}_i + \alpha_3 \ddot{u}_i \tag{3}$$

$$\alpha_1 = \frac{1}{\beta(\Delta t)^2} m + \frac{\gamma}{\beta \Delta t} c \tag{4}$$

$$\alpha_2 = \frac{1}{\beta \Delta t} m + \left(\frac{\gamma}{\beta} - 1\right) c \tag{5}$$

$$\alpha_3 = \left(\frac{1}{2\beta} - 1\right) m + \Delta t \left(\frac{\gamma}{2\beta} - 1\right) c \tag{6}$$

In which Δt is the analysis time step, and β and γ are the two constants that parameterise the Newmark family of time integration schemes. Conducting a dynamic analysis entails solving Equation 2 at each time step to calculate the time-varying response of the structure. Since these are implicit schemes and the system in nonlinear in nature, Equation 2 can only be solved iteratively. When the left and right hand sides of the equation of motion are balanced at a time step, a state of *dynamic equilibrium* is said to be achieved. When the Newton-Raphson method is employed to solve Equation 2, the residual force at the end of each iteration is computed as

$$r_{i+1}^{j} = \hat{p}_{i+1} - \alpha_1 u_{i+1}^{j} - (f_s)_{i+1}^{j} \tag{7}$$

where r_{i+1}^{j} is the residual force, and j represents the iteration number. To avoid introducing unnecessary rounding errors into the analysis, according to the Sterbenz lemma (Sterbenz 1973), the order of $\alpha_1 u_{i+1}^{j}$ and $(\mathrm{f_s})_{\mathrm{i+1}}^{\mathrm{j}}$ should be retained in the order seen in Equation 7, since the $\alpha_1 u_{i+1}$ term will always be larger in magnitude than the spring force (Paultre 2013). The displacement is the only quantity that is iteratively refined within a time step until dynamic equilibrium is achieved; the velocity and the acceleration are updated only at the end of the time step. Within each iteration, the displacement increment, $\delta\mathrm{u}_{\mathrm{i+1}}^{\mathrm{j}}$, is calculated as

$$\delta u_{i+1}^{j} = \frac{r_{i+1}^{j}}{\hat{k}_{i+1}^{j}} \tag{8}$$

where

$$\hat{k}_{i+1}^{j} = \alpha_1 + k_{i+1}^{j} \tag{9}$$

k_{i+1}^{j} is the stiffness of the nonlinear spring at iteration *j* and $\hat{k}_{i+1}^{j}$ is the effective stiffness of the system at iteration *j*.

A scalar error index is typically computed after each iteration within a time step. When this index is found to be smaller than a pre-defined convergence tolerance, ϵ (typically chosen to be an integral exponent of 10), the system is assumed to have reached a state of dynamic equilibrium. The analysis is now said to have converged to a solution at that time step and can move on to the next. Alternatively, if the number of iterations conducted within a time step exceeds a pre-defined upper bound, the analysis is considered to have not converged and is aborted. Common error indices used include the residual force, displacement increment, and energy increment error indices (Chopra 2020 §5.7.1). The discussion in this paper will focus on the residual force error index, however, the other error indices experience non-convergence for the same reason, with different minimum allowable tolerance values.

3 IDENTIFIED SOURCE OF NUMERICAL NON-CONVERGENCE

Each inelastic SDOF system was analysed repeatedly using progressively lower convergence tolerances until numerical non-convergence was observed. The final iterations, at the time step at which non-convergence was encountered, were closely examined to identify the source of non-convergence. The root cause of non-convergence, across all the analysed systems, was identified to be the rounding and truncation involved in floating-point operations.

A fixed number of bits are allocated to store any floating-point variable. A double precision floating-point variable, also known as a 64-bit floating-point variable, is stored using 64 bits. The IEEE 754 standard (IEEE 2019) dictates that 1 bit is used to store the sign of the value, *s*, 11 bits are used to store the exponent, *e*, and 52 bits are used to store the mantissa, *m*, also referred to as the fraction. When written in scientific form, any non-zero binary number is written as

$$(-1)^{s} * (1.m)_2 * 2^{e} \tag{10}$$

where the mantissa stores the digits after the binary point. The leading 1 exists for every number and therefore is implied rather than stored. The value of a unit change in the last position of the mantissa is called the quantum and represents the smallest amount by which the floating-point variable can be changed. Since a floating-point variable is stored in a finite number of bits, it can only store a finite number of values.

Within the context of structural analysis, the solution to Equation 2 must be chosen from among a finite set of displacement values that can be stored as a floating-point variable in the computer's memory. Non-convergence is typically encountered when the residual forces corresponding to the displacement values computed at successive iterations are all larger than the convergence tolerance. Three distinct modes of non-convergence were identified based on the relationship between the displacement values and the displacement quantum when non-convergence is encountered: "no-increment non-convergence", "single increment non-convergence", and "double increment non-convergence". The initial discussion will focus on single increment non-convergence, while the other two modes will be discussed later.

3.1 Single increment non-convergence

If the solution algorithm reaches a stage where the displacement alternates indefinitely between two adjacent representable values (whose difference is equal to the quantum of the displacement) at successive iterations, and the residual forces corresponding to these two displacement values are positive and negative respectively with magnitudes larger than the convergence tolerance, a single increment mode of non-convergence is said to have occurred. An example of iterations leading to the single increment mode of non-convergence is illustrated in Figure 1 and described in Table 1.

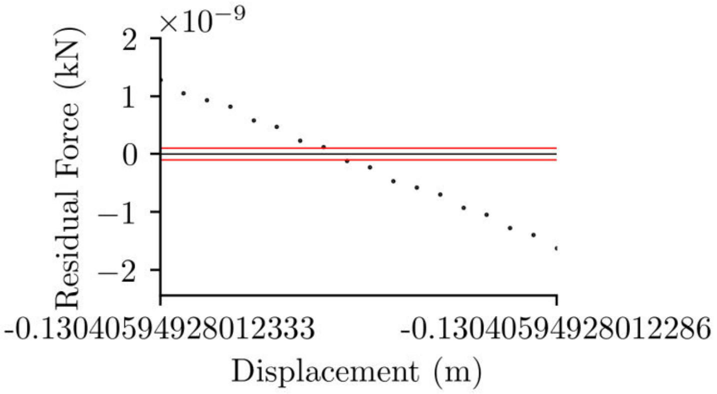

Figure 1: Residual force values of the representable displacement values for a time step in which single increment non-convergence is encountered. Model and analysis parameters: m = 150 tonnes, T_n = 3.6 s, R = 4.0, ξ = 2%, ϵ = 1 x 10^{-10} kN, $\Delta t = 0.01$ s, EPP hysteretic spring model, Newmark's average acceleration scheme.

Figure 1 shows the finite set of floating-point values the displacement and the associated residual force can assume, for an instance of single increment non-convergence. The red lines show the tolerance bounds of ±10^{-10} kN. This example relates to a time step encountering single increment non-convergence for an SDOF system with a mass of 150 tonnes, natural period of 3.6 s, strength reduction factor of 4, and damping ratio of 2%, modelled using an elastic-perfectly-plastic spring, and analysed using the average acceleration scheme with a time step of 0.01 s. It is observed that none of the representable displacement values correspond to residual forces that lie within the tolerance bounds. The slope of the imaginary line passing approximately through the dots is equal to $-\hat{k}$, however, due to the rounding of the term $\alpha_1 u_{i+1}$ after α_1 and u_{i+1} are multiplied together, however, the slope of each individual line segment is roughly equal to the value of $\hat{k}$ rounded either up or down to the nearest power of 2.

The same instance of non-convergence is shown in Table 1. Representing the displacements in hexadecimal form, where each hexadecimal digit represents four binary digits, provides better insight into how the binary representation changes than the decimal representation, while being more concise than the full binary representation. The quantum of the binary representation is equal to that of the hexadecimal representation, and any mention of the quantum in this paper will refer to this value. It can be seen that the displacement increments by the quantum in every iteration after iteration 1, and the residual forces associated with the two displacement values are larger than the convergence tolerance.

3.2 Factors controlling the single increment mode of non-convergence

The factors influencing the occurrence of non-convergence via the single increment mode can be inferred from Figure 1. A larger convergence tolerance, a smaller effective stiffness $\hat{k}$, and a smaller minimum displacement increment would all aid convergence, since they would increase the likelihood of one of the dots lying within the tolerance bounds, which the employed solution algorithm would find.

The size of the displacement quantum at a time step directly depends on the displacement at that time step, the largest quantum expected in the analysis would be encountered at the time step where the peak inelastic displacement is reached. Hence, the peak inelastic displacement can be used to predict the convergence behaviour of the analysis. Table 2 compares different ranges of displacements to the corresponding minimum changes in residual force, assuming an effective stiffness of 6,000,000 kNm^{-1}.

Table 1: Values of the displacement and residual force for iterations within a time step in which single increment non-convergence was encountered.

Iteration	Displacement (m) in Decimal form	Displacement (m) in Hexadecimal* form	Residual Force Error Index (kN)
0	-0.1272004670453954	-0x1.0481adaff799ep-3	-19240
1	-0.1304059492801231	-0x1.0b12463ae54d2p-3	-1.137 x 10^{-10}
2	-0.13040594928012314	-0x1.0b12463ae54d3p-3	1.192 x 10^{-10}
3	-0.1304059492801231	-0x1.0b12463ae54d2p-3	-1.137 x 10^{-10}
4	-0.13040594928012314	-0x1.0b12463ae54d3p-3	1.192 x 10^{-10}
⋮	⋮	⋮	⋮

*Hexadecimal numeral to be read as '*sign* 0x 1.*mantissa* p *exponent'*

The displacement ranges are split into powers of 2 so that the exponent of the floating-point variable is constant within each range, and therefore, the quantum takes the same value for all displacements within each range. Two changes in residual force are shown, computed by multiplying the displacement quantum by the effective stiffness, rounded down and up to the nearest powers of 2, denoted by floor and ceiling respectively.

The ability of the maximum inelastic displacement to predict convergence behaviour can also be seen in Figure 2, where contours of 0.125 m and 0.25 m maximum inelastic displacements are superimposed onto a plot showing instances of non-convergence for SDOF systems with different natural period and strength reduction factors. The analyses shown in this plot used a mass of 150 tonnes, a damping ratio of 2%, a time step of 0.01 s, and a residual force tolerance of 10^{-10} kN. The springs were modelled as elastic-perfectly-plastic and the average acceleration integration scheme was used, resulting in $\hat{k}$ values of roughly 6,000,000 kNm^{-1} across all systems. There is a clear trend in which systems with maximum inelastic displacements below 0.125 m do not encounter non-convergence, whereas most of the systems with maximum inelastic displacements above 0.125 m do encounter non-convergence. This is because the change in residual force is only larger than 2 x 10^{-10} kN, the spacing between the tolerance bounds for the analyses, when the displacement is above 0.125 m, as seen in Table 2. The relationship between the displacement,

Table 2: Comparison between the displacement ranges and the minimum changes to the residual force, assuming an effective stiffness of 6,000,000 kNm^{-1}

Displacement **Range** **(m)**	**Displacement** **Hexadecimal Form** (Bolded displacement) **(m)**	**Displacement** **Quantum** **Decimal Form** **(m)**	**Change in** **Residual Force** **(Floor)** **(kN)**	**Change in** **Residual Force** **(Ceiling)** **(kN)**
0.0625-0.1249	0x1.0000000000000p-4	~1.388 x 10^{-17}	5.82 x 10^{-11}	1.16 x 10^{-10}
0.125-0.249	0x1.0000000000000p-3	~2.776 x 10^{-17}	1.16 x 10^{-10}	2.33 x 10^{-10}
0.25-0.49	0x1.0000000000000p-2	~5.551 x 10^{-17}	2.33 x 10^{-10}	4.66 x 10^{-10}
0.5-0.99	0x1.0000000000000p-1	~1.110 x 10^{-17}	4.66 x 10^{-10}	9.31 x 10^{-10}

effective stiffness, and tolerance is only a predictor of whether non-convergence is possible and does not predict if non-convergence will occur. In Figure 2 there are many systems with natural periods around 2 to 2.5s and strength reduction factors between 6.5 and 8 which did no encounter non-convergence despite maximum inelastic displacements larger than 0.125 m.

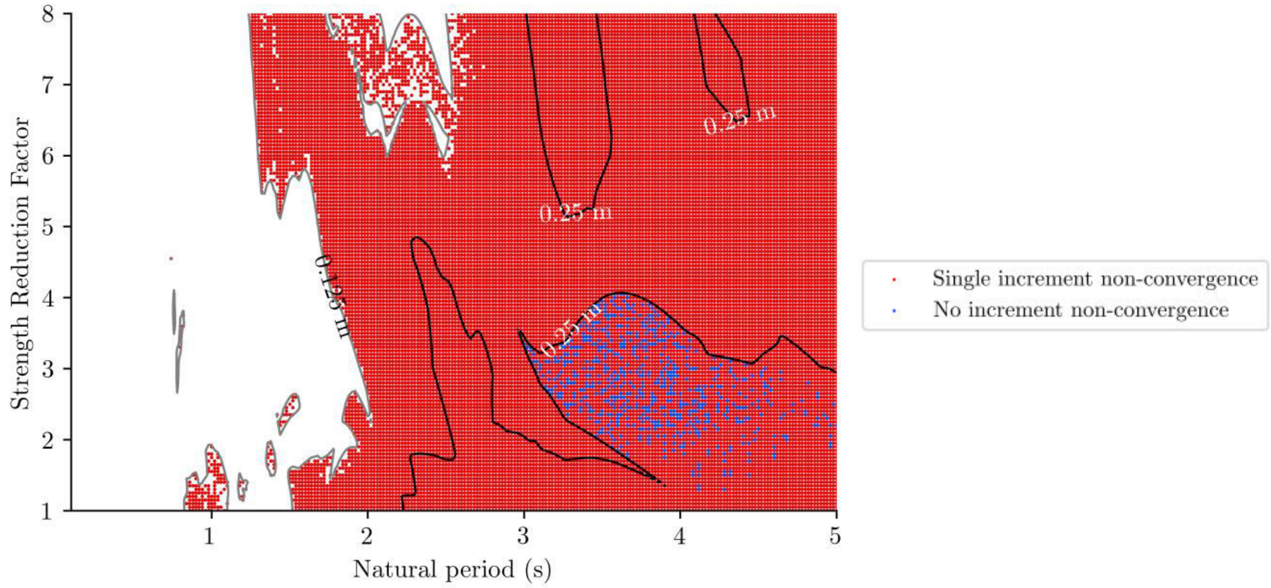

Figure 2: Instances of non-convergence for systems with different natural periods and strength reduction factors. m = 150 tonnes, ξ = 2%, Δt = 0.01 s, ϵ_R = 1 x 10^{-10} kN. Contours of 0.125 m (grey) and 0.25 m (black) max inelastic displacement are included.

3.3 No-increment and double increment non-convergence

No-increment non-convergence is said to have occurred when the displacement does not change between iterations. If the change in displacement calculated in Equation 8 is less than half of the quantum the displacement will be rounded to the same value it began the iteration at, and the same behaviour repeats. Since the state of the system does not change between iterations, the same behaviour will repeat until non-convergence is encountered.

For the systems in Figure 2, the effective stiffness in each of the systems is roughly 6,000,000 kNm^{-1}. Knowing that iterations only occur when the residual force is outside of the tolerance bounds, the smallest change in displacement calculated in Equation 8 is around 1.67 x 10^{-17} m, resulting from a residual force equal to the tolerance. Comparing the smallest calculated change in displacement to the displacement quanta, from Table 2, 1.67 x 10^{-17} m will apply a change when the displacement is below 0.25 m, and round back to the same displacement when the displacement is above 0.25. No-increment non-convergence can, therefore, only occur in systems with maximum inelastic displacements above 0.25 m, which is observed in Figure 2.

For both the no-increment and single increment modes of non-convergence, no representable points exist within the tolerance bounds. A point does exist, however, when encountering double increment non-convergence. An example of the representable displacements and residual forces for a time step which encounters double increment non-convergence is shown in Figure 3. This example relates to an SDOF system with a mass of 125 tonnes, natural period of 3.6 s, strength reduction factor of 3.95, and damping ratio of 2%, modelled using an elastic-perfectly-plastic spring, and analysed using the average acceleration scheme with a time step of 0.01 s. Although a representable displacement is in dynamic equilibrium, the algorithm oscillates between the two points indicated by the arrows until the analysis aborts. Iterations do not

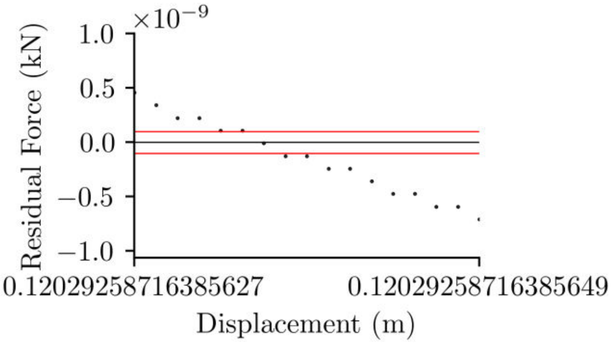

Figure 3: Residual forces for the representable displacement values at a time step encountering double increment non-convergence. m =125 tonnes, T_n = 3.6 s, R = 3.95, $\xi = 2\%$, ϵ = 1 x 10^{-10} kN, Δt = 0.01 s, EPP hysteretic spring model, Newmark's average acceleration scheme

reach the point within the tolerance bounds because the calculated change in displacement is larger than 1.5 times the quantum, and therefore calculated displacement change rounds to two quanta rather than one.

Double increment non-convergence is dependent on the effective stiffness and was found to only be encountered when the effective stiffness is in the range of

$$2^x < \hat{k} \leq \frac{4}{3} 2^x$$

where 2^x is $\hat{k}$ rounded down to the nearest power of 2. When the effective stiffness is outside of this range, the iterations will reach a representable point within the tolerance bounds, if one exists. The $\hat{k}$ values for the systems in Figure 2 are outside of this range, and therefore the systems will not encounter double increment non-convergence. Higher modes, such as triple increment non-convergence, do not exist.

4 SELECTING A TOLERANCE TO ELIMINATE THE LIKELIHOOD OF ENCOUNTERING NON-CONVERGENCE

The relationship between the tolerance, effective stiffness, and the maximum inelastic displacement can predict of non-convergence is possible. The algorithm below calculates a tolerance for the residual force error index above which non-convergence will not be encountered in the dynamic analysis of a bilinear SDOF system:

1. Estimate the maximum inelastic displacement, *u*

   The maximum inelastic displacement is required since it relates to the largest displacement quantum expected throughout the analysis. Any one of the many pre-existing techniques, such as Chopra and Chintanapakdee (2004), Hatzigeorgiou and Beskos (2009), Miranda (2001), and Miranda and Ruiz-García (2002), can be used to estimate the maximum inelastic displacement. The exponent of the displacement value, in binary form, is the key takeaway from the estimation.

   When using a strain softening model, the largest allowable deformation should be used instead, above which the system is deemed to be unstable.

2. Calculate the values of α_1 and *k*

   α_1 depends on the analysis parameters and Equation 4, and *k* is a characteristic of the system.

3. Reduce the displacement to the nearest power of 2, and increase $\alpha_1$and *k* to the nearest powers of 2

   The variables, rounded to the appropriate powers of 2, will be indicated as $\tilde{u}$, $\tilde{\alpha}_1$, and $\tilde{k}$.

4. Select a tolerance such that a representable value exists within the bounds

   The tolerance should be chosen such that is satisfies:

   $$2\epsilon_r > 2^{-52} \times \tilde{u}\tilde{\alpha}_1 + 2^{-52} \times \tilde{u}\tilde{k} \quad (12)$$

   $$2\epsilon_d > 2^{-52} \times \tilde{u}\tilde{\alpha}_1 \quad (13)$$

$$2\epsilon_e > \frac{1}{2}(2^{-52} \times \tilde{u})(2^{-52} \times \tilde{u}\tilde{\alpha}_1 + 2^{-52} \times \tilde{u}\tilde{k}) \tag{14}$$

where ϵ_r, ϵ_d, and ϵ_e are the tolerances for the residual force, displacement increment, and incremental energy error indices respectively, multiplied by 2 to represent the tolerance bounds. Since the three variables have been rounded to their respective powers of 2, $2^{-52} \times \tilde{u}$ is used to obtain the value of the displacement quantum.

5. Check for the possibility of double increment non-convergence

   If any expected values of $\hat{k}$ satisfy Equation 11, the righthand side of Equation 12 and 13 should be doubled, and the righthand side of Equation 14 should be quadrupled.

5 CONCLUSION

In this study, the cause of numerical non-convergence in SDOF systems was investigated. A sample set of over 1 million SDOF systems was analysed, using three piecewise linear models to describe the hysteretic behaviour, and subjected to the 1940 El Centro ground motion. The Newmark average acceleration and linear acceleration numerical time integration schemes were used, alongside the residual force error index. From the analysis it was observed that non-convergence in the analysis of SDOF systems is caused by rounding within floating-point operations. Three modes of non-convergence were identified, no-increment, single increment, and double increment non-convergence. The relationship between the tolerance, effective stiffness, and maximum inelastic displacement was a predictor for all three types of non-convergence. When the minimum change to the residual force exceeded the tolerance bounds, non-convergence could occur. To mitigate the likelihood of encountering non-convergence, an algorithm was developed to calculate a minimum allowable tolerance for the residual force error index. Using a tolerance equal to or above the minimum allowable tolerance ensures that non-convergence will not be encountered in an analysis.

REFERENCES

Chopra AK and Chintanapakdee C (2004) Inelastic deformation ratios for design and evaluation of structures: Single-degree-of-freedom bilinear systems". *Journal of Structural Engineering*, **130**(9): 1309-1319. https://doi.org/10.1061/(ASCE)0733-9445(2004)130:9(1309)

Chopra AK (2020). "*Dynamics of structures: Theory and Applications to Earthquake Engineering. in SI units"* 5th Edition. ISBN 9781292249186, Pearson Education Limited, London, United Kingdom

Hazigeorgiou GC and Beskos DE (2009). "Inelastic displacement ratios for SDOF structures subjected to repeated earthquakes". *Engineering Structures*, **31**(11): 2744-2755. https://doi.org/10.1016/j.engstruct.2009.07.002

IEEE (2019). "*IEEE Standard for Floating-Point Arithmetic"* IEEE Std 753-2019. Institute of Electrical and Electronics Engineers, Piscataway, United States, 84pp. https://ieeexplore.ieee.org/document/8766229

Miranda E (2001). "Estimation of inelastic deformation demands of SDOF systems", *Journal of Structural Engineering*. **127**(9): 1005-1012. https://doi.org/10.1061/(ASCE)0733-9445(2001)127:9(1005)

Miranda E and Ruiz-García J (2002). "Evaluation of approximate methods to estimate inelastic displacement demands". *Earthquake Engineering & Structural Dynamics*, **31**(3): 539-560.

Newmark NM (1959). "A Method of Computation for Structural Dynamics". *Journal of the Engineering Mechanics Division*, **85**(3): 67-94. http://doi.org/10.1061/JMCEA3.0000098

Paultre P (2013). "*Dynamics of Structures"*. ISBN 9781848210639, John Wiley and Sons, Hoboken, United States, 816 pp.

Sterbenz PH (1973). "*Floating-Point Computation"*. ISBN 978-0133224955, Prentice Hall, Englewood Cliffs, United States https://archive.org/embed/SterbenzFloatingPointComputation