Chapter 8

Nonlocal Damage Models

8.1 Basic Types of Nonlocal Damage Formulations

8.1.1 Formulations Motivated by Isotropic Damage

A number of nonlocal concepts giving local response in the linear elastic range have been proposed in the literature. We will illustrate some of them using the simple isotropic damage model from Section 5.2. Certain models use a formulation in which the role of the equivalent strain is played by the damage energy release rate[1]

$$Y = \frac{1}{2}\boldsymbol{\varepsilon} : \boldsymbol{D}_e : \boldsymbol{\varepsilon} \tag{8.1}$$

The internal variable κ can then be identified with the maximum value of Y ever reached in the history of the material,

$$Y_{max}(t) = \max_{\tau \le t} Y(\tau) \tag{8.2}$$

Recall that the basic stress-strain equation reads

$$\boldsymbol{\sigma} = (1-\omega)\boldsymbol{D}_e : \boldsymbol{\varepsilon} \tag{8.3}$$

where

$$\omega = g(Y_{max}) \tag{8.4}$$

is the damage parameter. Function g is usually designed such that $\omega = 0$ for Y_{max} below a certain threshold value, Y_0.

Now, several nonlocal versions of the model can be constructed:

1. The model originally proposed by Pijaudier-Cabot and Bažant [242] averages the damage energy release rate Y computed from (8.1) and evaluates the damage parameter corresponding to the maximum previously reached *nonlocal* value $\overline{Y}_{max}$. As long as $Y \le Y_0$ at every point, $\overline{Y}_{max}$ is also below the threshold and the response is linear elastic.

2. Bažant and Pijaudier-Cabot [30] suggested that alternatively one could average the damage parameter ω computed from (8.4) and substitute its nonlocal value into (8.3). As long as the material remains (everywhere) elastic, ω is equal to zero and so its nonlocal average is also zero. Equation (8.3) then reduces to the law of linear elasticity.

3. The smeared crack model of Bažant and Lin [25] dealt with nonlocal strain. Of course, we cannot substitute the averaged strain into (8.3) because then the model would be nonlocal already in the elastic range. However, if we use the nonlocal strain only in (8.1) when computing Y and keep the strain in (8.3) as local then the local character of the initial linear elastic response is preserved.

[1] Of course, Y is not a rate in the sense of a derivative with respect to time. It equals minus the derivative of the free energy $\psi = (1-\omega)\boldsymbol{\varepsilon} : \boldsymbol{D}_e : \boldsymbol{\varepsilon}/2$ with respect to the damage parameter, and so it represents the "rate" at which energy is released as the damage parameter increases (at constant strain and temperature).

formulation	isotropic damage model	general model
$\omega(\overline{\varepsilon})$	$\boldsymbol{\sigma} = [1-\omega(\overline{\boldsymbol{\varepsilon}})]\boldsymbol{D}_e : \boldsymbol{\varepsilon}$	$\boldsymbol{\sigma} = \boldsymbol{D}_s(\overline{\boldsymbol{\varepsilon}}) : \boldsymbol{\varepsilon}$
$\overline{Y}$	$\boldsymbol{\sigma} = [1-\omega(\overline{Y(\boldsymbol{\varepsilon})})]\boldsymbol{D}_e : \boldsymbol{\varepsilon}$	$\boldsymbol{\sigma} = \boldsymbol{D}_s(\boldsymbol{\Omega}(\overline{\boldsymbol{Y}(\boldsymbol{\varepsilon})})) : \boldsymbol{\varepsilon}$
$\overline{\omega}$	$\boldsymbol{\sigma} = [1-\overline{\omega(\boldsymbol{\varepsilon})}]\boldsymbol{D}_e : \boldsymbol{\varepsilon}$	$\boldsymbol{\sigma} = \overline{\boldsymbol{D}_s(\boldsymbol{\varepsilon})} : \boldsymbol{\varepsilon}$
$\overline{\gamma}$	$\boldsymbol{\sigma} = [1+\overline{\gamma(\boldsymbol{\varepsilon})}]^{-1}\boldsymbol{D}_e : \boldsymbol{\varepsilon}$	$\boldsymbol{\sigma} = [\boldsymbol{C}_e + \overline{\boldsymbol{C}_i(\boldsymbol{\varepsilon})}]^{-1} : \boldsymbol{\varepsilon}$
$\overline{s}$	$\boldsymbol{\sigma} = \boldsymbol{D}_e : \boldsymbol{\varepsilon} - \overline{\omega(\boldsymbol{\varepsilon})\boldsymbol{D}_e : \boldsymbol{\varepsilon}}$	$\boldsymbol{\sigma} = \boldsymbol{D}_e : \boldsymbol{\varepsilon} - \overline{\boldsymbol{s}(\boldsymbol{\varepsilon})}$
$\overline{\Delta s}$	$\dot{\boldsymbol{\sigma}} = (1-\omega)\boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} - \overline{\dot{\omega}\boldsymbol{D}_e : \boldsymbol{\varepsilon}}$	$\dot{\boldsymbol{\sigma}} = \boldsymbol{D}_u : \dot{\boldsymbol{\varepsilon}} - \overline{\dot{\boldsymbol{s}}(\boldsymbol{\varepsilon}, \dot{\boldsymbol{\varepsilon}})}$
$s(\overline{\varepsilon})$	$\boldsymbol{\sigma} = \boldsymbol{D}_e : \boldsymbol{\varepsilon} - \omega(\overline{\boldsymbol{\varepsilon}})\boldsymbol{D}_e : \overline{\boldsymbol{\varepsilon}}$	$\boldsymbol{\sigma} = \boldsymbol{D}_e : \boldsymbol{\varepsilon} - \boldsymbol{s}(\overline{\boldsymbol{\varepsilon}})$

Table 8.1: Overview of nonlocal formulations

4. Pijaudier-Cabot and Bažant [242] also mentioned that a nonlocal model could be obtained by averaging of the specific fracturing strain. Applying this idea to the isotropic damage law we rewrite (8.3) as

$$\boldsymbol{\varepsilon} = (1+\gamma)\boldsymbol{C}_e : \boldsymbol{\sigma} \tag{8.5}$$

where $\gamma = \omega/(1-\omega)$ is the specific fracturing strain. Replacing γ by its weighted average, $\overline{\gamma}$, we construct a nonlocal version of the isotropic damage model.

The above nonlocal formulations are summarized in the upper section of Table 8.1. For easy reference, we will denote them by symbols $\overline{Y}$, $\overline{\omega}$, etc.; see the first column of Table 8.1. These formulations were motivated by the isotropic damage model but they can be extended to the class of constitutive laws that express the stress as the product of a secant (damaged) stiffness and the strain, $\boldsymbol{\sigma} = \boldsymbol{D}_s : \boldsymbol{\varepsilon}$. The generalized forms are shown in the last column of Table 8.1. For example, for formulation $\omega(\overline{\varepsilon})$ we use the nonlocal strain as input for the evaluation of the secant stiffness while for formulation $\overline{\omega}$ we first evaluate the secant stiffness locally and then compute its nonlocal average. A natural extension of formulation $\overline{\gamma}$ to the anisotropic case is a model that averages the inelastic compliance, $\boldsymbol{C}_i$.

Anisotropic damage models usually work with a certain damage tensor $\boldsymbol{\Omega}$. A natural extension of formulation $\overline{Y}$ is a model applying nonlocal averaging to the tensor $\boldsymbol{Y}$ that is work-conjugate with $\boldsymbol{\Omega}$.

8.1.2 General Formulations

In addition to nonlocal formulations motivated by the isotropic damage model it is possible to develop nonlocal models written directly in a general format.

5. The elastic response remains local if we average a quantity that is in the elastic state equal to zero, e.g., the inelastic strain. This concept applies to any type of constitutive law formally written as

$$\boldsymbol{\sigma} = \boldsymbol{D}_e : (\boldsymbol{\varepsilon} - \boldsymbol{e}) \tag{8.6}$$

where $\boldsymbol{e}$ is the inelastic strain (fracturing strain, plastic strain, etc.). A nonlocal version of the law is obtained when we replace the inelastic strain by its nonlocal counterpart. If the elastic moduli are uniform throughout the body, this is fully equivalent to a model averaging the inelastic stress

$$\boldsymbol{s} = \boldsymbol{D}_e : \boldsymbol{e} = \boldsymbol{D}_e : \boldsymbol{\varepsilon} - \boldsymbol{\sigma} \tag{8.7}$$

The nonlocal law then reads

$$\boldsymbol{\sigma} = \boldsymbol{D}_e : \boldsymbol{\varepsilon} - \overline{\boldsymbol{s}} \tag{8.8}$$

This is the standard version of the generalized nonlocal concept due to Bažant [20]. Note that Bažant worked with the inelastic stress rate

$$\dot{\boldsymbol{s}} = \boldsymbol{D}_e : \dot{\boldsymbol{\varepsilon}} - \dot{\boldsymbol{\sigma}} \tag{8.9}$$

which can be integrated to yield (8.7).

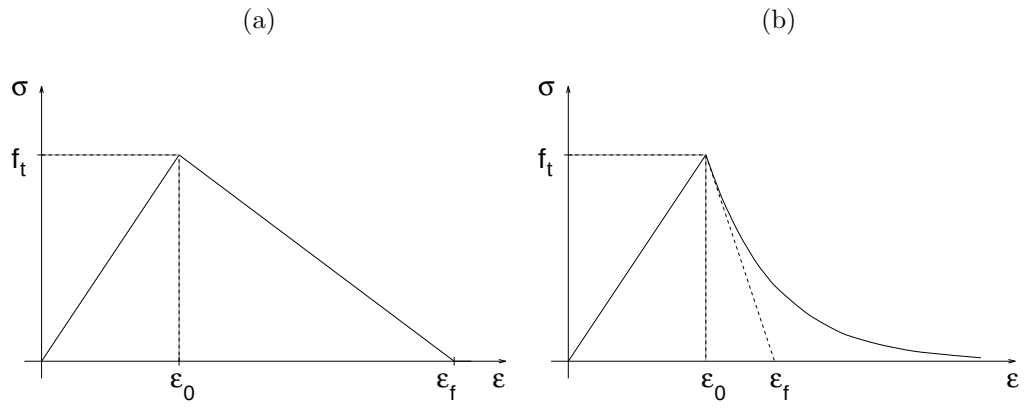


Figure 8.1: Local stress-strain law with a) linear softening, b) exponential softening.

6. Alternatively, we could define the inelastic stress rate as

$$\dot{\boldsymbol{s}} = \boldsymbol{D}_u : \dot{\boldsymbol{\varepsilon}} - \dot{\boldsymbol{\sigma}} \tag{8.10}$$

This approach has been taken by Jirásek and Bažant [144]. An important difference compared to (8.9) is that the elastic stress rate is now $\boldsymbol{D}_u : \dot{\boldsymbol{\varepsilon}}$ where $\boldsymbol{D}_u$ is the stiffness matrix valid for unloading. For models with degradation of the elastic moduli, $\boldsymbol{D}_u$ varies during the loading process, and $\dot{\boldsymbol{s}}$ defined by (8.10) is no longer the time derivative of the quantity $\boldsymbol{s}$ defined by (8.7). Integration of the nonlocal constitutive law

$$\dot{\boldsymbol{\sigma}} = \boldsymbol{D}_u : \dot{\boldsymbol{\varepsilon}} - \dot{\overline{\boldsymbol{s}}} \tag{8.11}$$

then yields a result different from (8.8).

7. Finally, Bažant et al. [33] postulated a general nonlocal constitutive law in the form

$$\boldsymbol{\sigma} = \boldsymbol{D}_e : \boldsymbol{\varepsilon} - \boldsymbol{s}(\overline{\boldsymbol{\varepsilon}}) \tag{8.12}$$

where $\boldsymbol{s}(\overline{\boldsymbol{\varepsilon}})$ is the inelastic stress calculated from the nonlocal strain, $\overline{\boldsymbol{\varepsilon}}$. This means that the actual stress is obtained as the sum of an elastic part evaluated from local strain and an inelastic part evaluated from nonlocal strain.

Of course, general nonlocal formulations 5–7 can be specialized to the isotropic damage model by substituting $\boldsymbol{s} = \omega \boldsymbol{D}_e : \boldsymbol{\varepsilon}$ for the inelastic stress. The resulting stress-strain equations are listed in the lower section of Table 8.1.

8.2 Evaluation of Nonlocal Formulations in One Dimension

8.2.1 Analytical and Numerical Solutions

Let us now test the behavior of individual nonlocal formulations in an elementary localization problem—tensile failure of a straight uniform bar of length L; see Fig. 7.4. The bar is supposed to be divided into a finite number of elements with a linear displacement interpolation inside each element and with one Gauss integration point per element.

We will use simple local stress-strain relations with linear elasticity up to the peak stress and either linear or exponential softening; see Fig. 8.1. The nonlocal formulations will pass the test if, for sufficiently large bar elongations, the residual resistance vanishes and the strain profile keeps its localized character.

Before starting numerical simulations we will study a simple problem solvable by hand. Let us consider a bar divided into three equally sized elements. To render the hand solution feasible we use the local stress-strain relation with linear softening. Of course, such a crude model will not lead to realistic shapes of the load-displacement diagram but it will help us to identify the nature of the problems occuring for some of the formulations.

To facilitate the calculations we fix the parameters of the local constitutive law (Fig. 8.1a) to $f_t = 1$, $\varepsilon_0 = 1$, and $\varepsilon_f = 3$, and we consider a bar of cross-sectional area $A = 1$ and length $L = 3$. For linear softening with the chosen parameters, the dependence of the damage parameter on the maximum previously reached strain, ε_{max}, is given by

$$\omega = \begin{cases} 0 & \text{if } \varepsilon_{max} \leq 1 \\ 1.5\left(1 - \dfrac{1}{\varepsilon_{max}}\right) & \text{if } 1 \leq \varepsilon_{max} \leq 3 \\ 1 & \text{if } 3 \leq \varepsilon_{max} \end{cases} \tag{8.13}$$

Furthermore, we assume that the interaction radius R from (7.60) is only slightly larger than the element size and that the discretized nonlocal averaging formulae are

$$\overline{f}_1 = 0.9f_1 + 0.1f_2 \tag{8.14}$$

$$\overline{f}_2 = 0.1f_1 + 0.8f_2 + 0.1f_3 \tag{8.15}$$

$$\overline{f}_3 = 0.1f_2 + 0.9f_3 \tag{8.16}$$

where f_i are local values and $\overline{f}_i$ are nonlocal values of an arbitrary variable f at the center of element number i. Such an assumption corresponds to slightly different values of R for individual elements ($R_1 = R_3 = 1.2247$, $R_2 = 1.2438$). With this choice, the model response is qualitatively the same as for a uniform interaction radius and the coefficients in the averaging formulae are easy to handle.

- Let us start with formulation $\omega(\overline{\varepsilon})$. At peak stress, $\sigma = f_t = 1$, the load-displacement diagram has a multiple bifurcation point. Besides the uniform solution there exist several solutions with one or two elastically unloading elements. It is possible to show that the steepest descent of the load-displacement diagram is obtained if damage localizes into one of the elements at the boundary while the other two elements unload. Under displacement control, this solution corresponds to the stable branch of the diagram (see [21], Section 10.2).

  Provided that damage localizes in element number 1, the constitutive law $\sigma = [1 - \omega(\overline{\varepsilon})]\varepsilon$ applied at element centers yields

  $$\sigma_1 = \left(\frac{1.5}{0.9\varepsilon_1 + 0.1\varepsilon_2} - 0.5\right)\varepsilon_1 \tag{8.17}$$

  $$\sigma_2 = \varepsilon_2 \tag{8.18}$$

  $$\sigma_3 = \varepsilon_3 \tag{8.19}$$

  From equilibrium conditions $\sigma_1 = \sigma_2 = \sigma_3$ we obtain the solution

  $$\sigma = \varepsilon_2 = \varepsilon_3 = \sqrt{18.0625\varepsilon_1^2 + 15\varepsilon_1} - 4.75\varepsilon_1 \tag{8.20}$$

  expressed in terms of strain ε_1, which plays the role of a parameter controling the loading process. The solution remains valid as long as $\overline{\varepsilon}_1 \leq 3$. At $\varepsilon_1 = 3.333$ we have $\sigma = \varepsilon_2 = \varepsilon_3 = 0$ and $\overline{\varepsilon}_1 = 3$. The load is fully relaxed and additional increments of applied displacement do not have to oppose any residual resistance. The load-displacement diagram is represented by the dashed curve in Fig. 8.2a. The solid curve in the same graph corresponds to a numerical solution with 30 elements while the straight dotted line is the uniform (not localized) solution, i.e., a rescaled local stress-strain curve. The solution with a large number of elements exhibits snapback and the final elongation at complete failure is smaller but the essential feature investigated in the present section is the same as for the solution with 3 elements—the load is fully relaxed.

- Formulation $\overline{Y}$ leads for the present problem to a quartic equation, which cannot be easily solved by hand. However, the numerically obtained solutions with 3 and 30 elements are very similar to the preceding formulation; see Fig. 8.2b.

- The response is substantially different for formulation $\overline{\omega}$. Application of the nonlocal law $\sigma = (1 - \overline{\omega})\varepsilon$ at element centers leads to

  $$\sigma_1 = 1.35 - 0.35\varepsilon_1 \tag{8.21}$$

  $$\sigma_2 = \left(0.85 + \frac{0.15}{\varepsilon_1}\right)\varepsilon_2 \tag{8.22}$$

  $$\sigma_3 = \varepsilon_3 \tag{8.23}$$

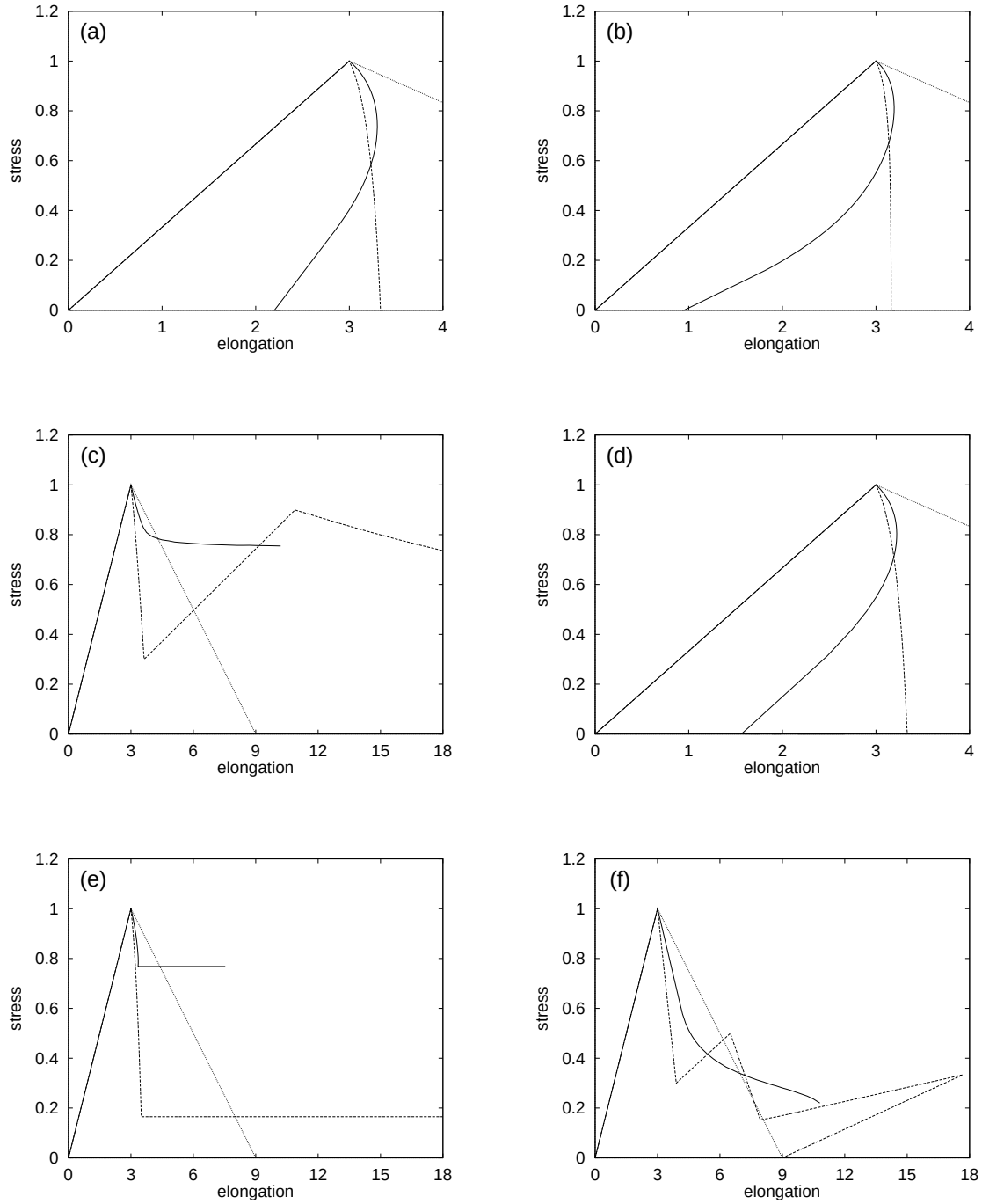


Figure 8.2: Load-displacement diagrams for formulation a) $\omega(\overline{\varepsilon})$, b) $\overline{Y}$, c) $\overline{\omega}$, d) $\overline{\gamma}$, e) $\overline{\Delta s}$, f) $\overline{s}$ and $s(\overline{\varepsilon})$. Dashed curves have been obtained with 3 elements, solid curves with 30 elements.

and from equilibrium we get the solution

$$\varepsilon_2 = \frac{(1.35 - 0.35\varepsilon_1)\varepsilon_1}{0.85\varepsilon_1 + 0.15} \tag{8.24}$$

$$\sigma = \varepsilon_3 = 1.35 - 0.35\varepsilon_1 \tag{8.25}$$

These expressions remain valid until $\varepsilon_1 = 3$. At this state, element number 1 is fully *locally* damaged ($\omega_1 = 1$) but the nonlocal damage $\overline{\omega}_1 = 0.9\omega_1 + 0.1\omega_2 = 0.9 < 1$, and so the element can still transfer stress. Moreover, during the subsequent stage of loading no further damage is produced because ω_1 cannot grow anymore and strains in elements 2 and 3 are below the elasticity limit (these elements have been unloaded to strains $\varepsilon_2 = 0.333$ and $\varepsilon_3 = 0.3$). This means that the model responds elastically (with reduced stiffness of elements 1 and 2) until ω_2 starts growing at $\varepsilon_2 = 1$. The load-displacement curve is again rising up to a stress comparable to the tensile strength; see the dashed line in Fig. 8.2c. The final stage, during which local damage in element 2 is growing from 0 to 1, can be described by

$$\varepsilon_1 = \frac{(1.2 - 0.3\varepsilon_2)\varepsilon_2}{0.15 - 0.05\varepsilon_2} \tag{8.26}$$

$$\varepsilon_3 = \frac{(1.2 - 0.3\varepsilon_2)\varepsilon_2}{0.15 + 0.85\varepsilon_2} \tag{8.27}$$

$$\sigma = 1.2 - 0.3\varepsilon_2 \tag{8.28}$$

Surprisingly, as ε_2 approaches 3, ε_1 tends to infinity while σ tends to 0.3. This means that the load-displacement curve asymptotically approaches a horizontal line well above the line of zero stress (this would become obvious if Fig. 8.2c was plotted for a larger range of elongation values). It might be argued that such a paradoxical result is caused by the poor spatial resolution of the model and that the behavior improves after mesh refinement. A simulation with 30 elements (which is certainly enough to capture all essential features of the solution) gives a monotonically decreasing post-peak curve but again ceases to provide full load relaxation; see the solid curve in Fig. 8.2c. Thus it must be concluded that formulation $\overline{\omega}$ does not meet the fundamental requirement postulated at the beginning of this section. It exhibits a special type of stress locking.

- Formulation $\overline{\gamma}$ leads to full load relaxation and the load-displacement diagram is similar to those produced by formulations $\omega(\overline{\varepsilon})$ and $\overline{Y}$; see Fig. 8.2d. The post-peak solution is described by

$$\varepsilon_2 = \frac{(2.7 - 0.7\varepsilon_1)\varepsilon_1}{0.3 + 1.3\varepsilon_1} \tag{8.29}$$

$$\sigma = \varepsilon_3 = \frac{(3 - \varepsilon_1)\varepsilon_1}{0.3 + 1.3\varepsilon_1} \tag{8.30}$$

An interesting difference compared to formulations $\omega(\overline{\varepsilon})$ and $\overline{Y}$ is that, at complete failure, the strain does not localize into a single element (at $\varepsilon_1 = 3$ we have $\sigma = \varepsilon_3 = 0$ but $\varepsilon_2 = 0.333 \neq 0$).

- The initial post-peak response of the model with nonlocal inelastic stress rate is described by

$$\varepsilon_2 = 1.5 - 0.5\varepsilon_1 + 0.3\ln\varepsilon_1 \tag{8.31}$$

$$\sigma = \varepsilon_3 = 1.5 - 0.5\varepsilon_1 + 0.15\ln\varepsilon_1 \tag{8.32}$$

This corresponds to a reasonable descending branch in the load-displacement diagram; see Fig. 8.2e. However, at $\varepsilon_1 = 3$ the stress ceases to decrease and the diagram continues by a horizontal line. The reason is that local inelastic stress increments in all elements are now zero (element 1 is fully locally damaged and elements 2 and 3 are locally in the elastic range). As the unloading stiffness of element 1 is also zero, no stress change is possible in that element. The same type of behavior, only with a larger value of the residual stress, is exhibited by the model with 30 elements.

- Finally, for formulation $\overline{s}$ we get a diagram with alternating ascending and descending straight segments; see Fig. 8.2f. Each descending segment corresponds to softening in one of the elements while the other elements are locally either in the elastic range or fully damaged. This alternating effect is indeed due to the poor spatial resolution and is not present in the simulation with 30

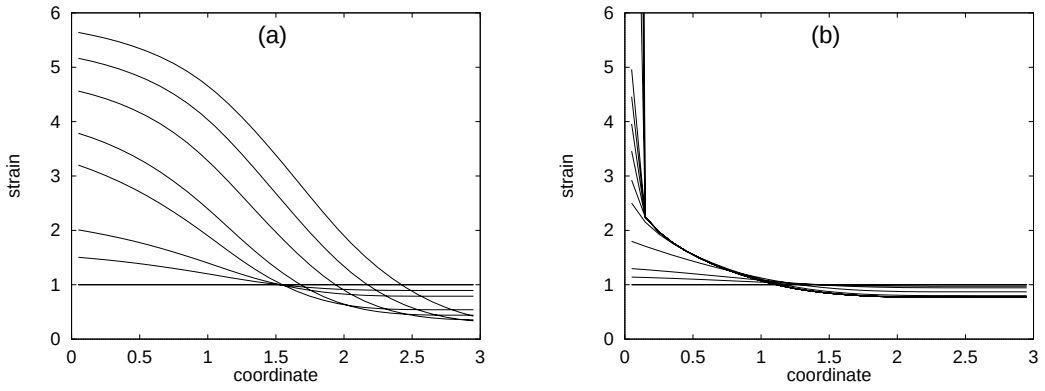


Figure 8.3: Evolution of strain profile for formulation a) $\overline{s}$, b) $\overline{\Delta s}$.

elements. However, the important point is that independently of the discretization the stress drops down to zero only after all elements have been fully damaged! Consequently, the final strain profile is not localized but uniform.

Even though formulations $\overline{s}$ and $s(\overline{\varepsilon})$ are in general different, the corresponding load-displacement diagrams (not the strain profiles) happen to be the same (for the present simple uniaxial problem).

8.2.2 Theoretical Analysis of Locking Mechanisms

The fact that formulation $\overline{s}$ must give a uniform strain profile at complete failure can be proven theoretically without resorting to finite element discretization. At complete failure, the stress at every point is zero, and so the nonlocal constitutive relation (8.8) combined with (8.7) and reduced to one dimension gives

$$E\left[\varepsilon(x) - \int_0^L \alpha(x,\xi) e(\xi)\, d\xi\right] = 0 \tag{8.33}$$

Young's modulus E is positive, and so the expression in brackets must vanish. This condition can be rewritten as

$$\int_0^L \alpha(x,\xi)\,[\varepsilon(x) - e(\xi)]\; d\xi = 0 \tag{8.34}$$

because the weight function is normalized, $\int_0^L \alpha(x,\xi)\, d\xi = 1$. Note that (8.34) must hold for any x. Let us denote by x_0 the point with the largest strain. The inelastic strain can nowhere exceed $\varepsilon(x_0)$, i.e., $\varepsilon(x_0) - e(\xi) \geq 0$ for any ξ. Moreover, $\alpha(x_0,\xi)$ is nonnegative for any ξ and is strictly positive for $\xi \in (x_0 - R, x_0 + R)$ where R is the interaction radius. Consequently, (8.34) can hold at $x = x_0$ only if $\varepsilon(x_0) = e(\xi)$ for any $\xi \in (x_0 - R, x_0 + R)$. For a weight function with unbounded support this means that $\varepsilon(\xi) = \varepsilon(x_0) = \text{const}$ everywhere. But even if the weight function has a bounded support, we can recursively apply the same argument at x_0 shifted by $\pm nR/2$, $n = 1, 2, \ldots$, and arrive at the conclusion that the strain is constant along the entire bar. This explains why the formulation with nonlocal inelastic stress cannot properly represent localized deformation at complete failure. The progressive expansion of the process zone is documented in Fig. 8.3a, which shows the evolution of the strain profile obtained numerically for the test problem analyzed in the preceding subsection.

Similarly, we can explain the stress-locking behavior of formulation $s(\overline{\varepsilon})$. Analyzing the situation at complete failure when $\sigma = E\varepsilon - s(\overline{\varepsilon}) = 0$ and using the fact that the inelastic stress cannot exceed the elastic stress computed for the same strain, $s(\overline{\varepsilon}) \leq E\overline{\varepsilon}$, we can derive an inequality

$$\int_0^L \alpha(x,\xi)\,[\varepsilon(x) - \varepsilon(\xi)]\; d\xi \leq 0 \tag{8.35}$$

that must be satisfied for every $x \in \langle 0, L\rangle$. Again, it can be concluded that the strain profile must be uniform.

Let us now look at the behavior of formulation $\overline{\omega}$. The constitutive equation in one dimension reads

$$\sigma = (1 - \overline{\omega}) E \varepsilon \tag{8.36}$$

where

$$\overline{\omega}(x) = \int_0^L \alpha(x, \xi) \omega(\xi) \, d\xi \tag{8.37}$$

At complete fracture we have $\sigma(x) = 0$ and so there must exist a point x_0 at which

$$\overline{\omega}(x_0) = 1 \tag{8.38}$$

otherwise the strain would have to vanish identically and the total extension of the bar would be zero. However, as $\omega(\xi) \leq 1$ for any ξ, (8.38) can hold only if $\omega(\xi) = 1$ whenever $\alpha(x_0, \xi) > 0$. For a weight function with unbounded support this means that every point of the bar must be completely damaged. The rigorous proof of a similar statement for a weight function with bounded support would be more tricky but even in this case the model is incapable of capturing localized damage at complete failure.

We can also explain the mechanism of stress locking for formulation $\overline{\Delta s}$. For a plasticity-type model, in which unloading takes place with the initial stiffness, the formulation is identical with the approach using nonlocal inelastic stress, and the criticism of formulation $\overline{s}$ applies. For a damage-type model with degradation of elastic stiffness, the problem appears as soon as the point x_0 with maximum strain reaches the state of complete local damage. The current unloading modulus E_u at x_0 is now zero and arbitrary strain increments at x_0 do not affect the stress state. Therefore, strain increments fully localize into this single point while the stress remains constant (and different from zero). This behavior is documented in Fig. 8.3b, which shows the evolution of the strain profile obtained numerically for the test problem analyzed in the preceding subsection.

8.3 Localization Zone

8.3.1 Structure of Localization Zone

We have shown that certain nonlocal formulations are inherently incapable of reproducing the entire material degradation process up to complete failure. Unless we are interested only in the response at the onset of localization, models that exhibit the special type of stress locking described in the previous section should be avoided. Theoretical analysis of the locking mechanisms revealed that the pathological behavior must appear independently of the particular value of the internal length or interaction radius.

From now on, we will restrict our attention to the formulations that do have the potential of properly describing localized damage up to the formation of a stress-free crack. Formulations $\overline{Y}$ and $\omega(\overline{\varepsilon})$ deal respectively with nonlocal damage energy release rate and nonlocal strain, and so they are quite similar because the damage energy release rate can be interpreted as the square of a generalized strain norm. In one dimension we simply have $Y = E\varepsilon^2/2$, which means that formulation $\overline{Y}$ averages the square of strain while formulation $\omega(\overline{\varepsilon})$ averages the strain itself. Numerical experience with simulations of tensile failure indicates that in general there are only minor differences between the results obtained with the two approaches. From computational point of view it is less expensive to average the damage energy release rate because it is a scalar quantity. Also, averaging of energy seems to be somewhat more logical from the physical point of view.

On the other hand, the formulation with nonlocal strain is more general because it can be extended to the class of constitutive laws written in the form $\boldsymbol{\sigma} = \boldsymbol{D}_s(\boldsymbol{\varepsilon}) : \boldsymbol{\varepsilon}$; see Table 8.1. In the nonlocal version, we evaluate the unloading (secant) stiffness matrix from the nonlocal strain and then multiply it by the local strain to obtain the actual stress. This concept can be applied for example to the microplane model [29] and to the fixed or rotating crack model [25, 150, 154].

Alternatively, formulation $\overline{\gamma}$ could be used for the same purpose. Nonlocal averaging would be applied to the inelastic material compliance matrix $\boldsymbol{C}_i$, which requires an even larger amount of computational work than averaging of the strain.

It is well known that a properly formulated nonlocal version of a strain-softening constitutive model provides solutions with damage and strain localized into bands whose width is controled by the internal length. When describing these bands we should pay attention to the choice of the quantity that serves as an indicator of the intensity of localization. This quantity could be local or nonlocal strain (largest

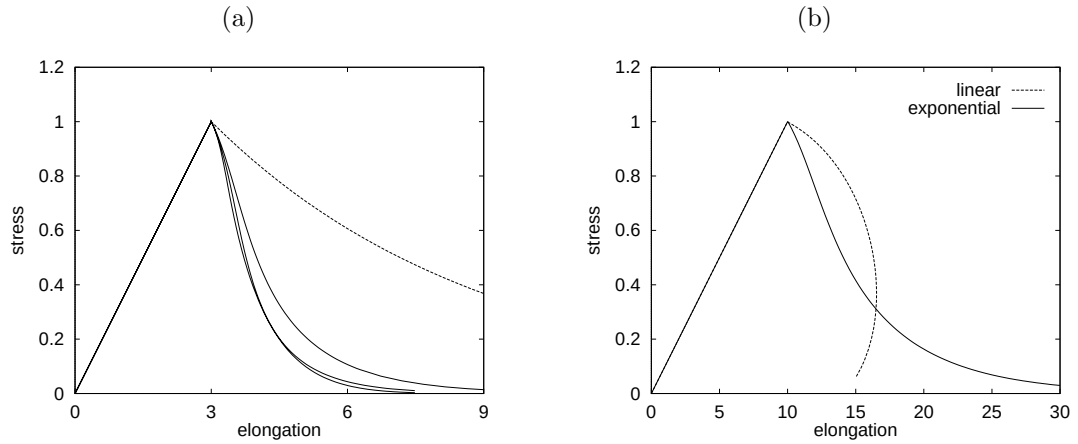


Figure 8.4: Load-displacement diagrams for nonlocal formulations: a) with exponential local law, dashed curve corresponds to the uniform solution, solid curves are stable paths for formulations $\omega(\overline{\varepsilon})$, $\overline{\gamma}$, and $\overline{Y}$ (from top to bottom); b) with linear and exponential local law

principal value), the damage parameter, the damage energy release rate, etc. For simplicity we will now consider the formulation with nonlocal strain but similar arguments would apply to the formulation with nonlocal damage energy release rate. We focus on a loading process in which the bar elongation grows beyond any bounds. From the basic expression for stress, $\sigma = (1-\omega)E\varepsilon$, it is obvious that as the bar approaches the state of complete failure, $\sigma = 0$, either the strain must tend to zero, or the damage parameter must tend to one, or both. For the present purpose, it is convenient to work with the discretized model rather than with the original one-dimensional continuum.

Let us consider a softening law with nonzero stress for any positive strain, e.g., the exponential softening law. For such a law, the value of the damage parameter at a given element tends to one if and only if the nonlocal strain at the integration point of that element tends to infinity. Thus we can define localization zone Z_ω as the set of elements for which $\omega \to 1$ or, equivalently, as the set of elements for which $\overline{\varepsilon} \to \infty$. Another observation that we make is that there must exist at least one element at which the local strain tends to infinity because otherwise the total extension would have to remain finite; let us denote the set of such elements by Z_ε.

It is easy to see that if the local strain in an element grows beyond any bounds then the nonlocal strain is unbounded in all elements whose distance from the current one (measured from integration point to integration point) is smaller than the interaction radius, R. And vice versa, the nonlocal strain in an element is unbounded only if there exists another element with unbounded local strain within the interaction radius. Consequently, Z_ω consists of elements from Z_ε plus all their neighbors located closer than R. In the limit for infinitely small elements we can expect Z_ε to be an interval in the center of the zone and Z_ω to be Z_ε extended on each side by an interval of length R.

8.3.2 Evolution of Localization Zone

Theoretical predictions from the preceding subsection can be confirmed by numerical simulations. As is clear from Fig. 8.2a,b,d, the shape of diagrams obtained with a nonlocal model using a linear local softening law are not at all realistic. More reasonable response is produced by an exponential local softening law; see Fig. 8.4a.

Fig. 8.5 illustrates a typical evolution of the localization zone. It shows the profiles of local strain, nonlocal strain, and damage parameter at several stages of the loading process. The simulation has been done for formulation $\omega(\overline{\varepsilon})$ with the bell-shaped weight function and the exponential softening law from Fig. 8.1b. Note that the zone of increasing local strain gets thinner as the loading process continues while the zone of increasing nonlocal strain keeps an approximately constant width. In the damage profile we observe a very fast transition between the elastically unloading regions with zero damage and the process zone with almost complete damage. This is of course related to the fact that the points where the nonlocal strain remains at the limit elastic value do not shift.

Fig. 8.6 shows the profiles of nonlocal strain and damage parameter for a simulation with the Gaussian

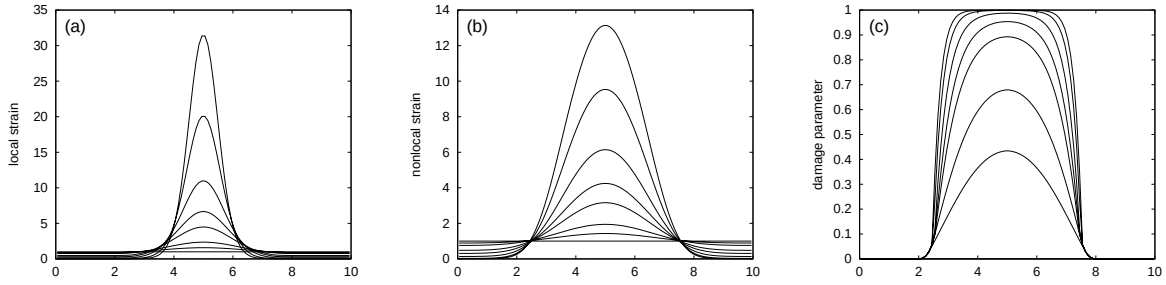


Figure 8.5: Profiles obtained with bell function: a) local strain, b) nonlocal strain, c) damage parameter

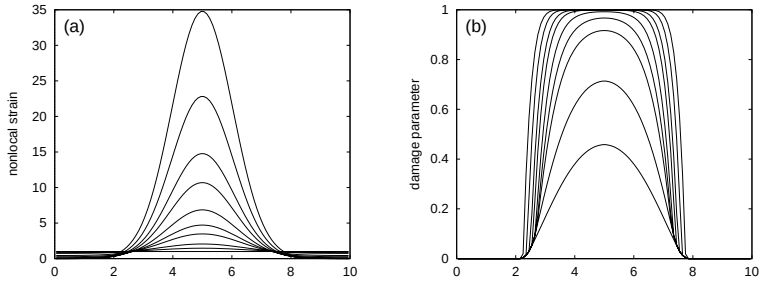


Figure 8.6: Profiles obtained with Gauss function: a) nonlocal strain, b) damage parameter

weight function. The initial development is very similar to the previous case but at later stages the damaged zone starts expanding. This is consistent with our theoretical predictions because for a weight function with unbounded support the zone Z_ω in which damage tends to one extends over the entire specimen. Theoretically, for very large elongations, all elements of the model are almost completely damaged. This seems to be merely an academic problem because there is no need for simulating the tensile failure of a bar up to extremely large elongations. However, in two- or three-dimensional fracture simulations we often obtain a zone of localized strain that corresponds to a macroscopic crack propagating across the specimen. The stress-free part of the crack sometimes opens very wide while the structure still keeps nonnegligible resistance. For example, in analyses of a crack propagating in a gravity dam the structural resistance increases due to the stabilizing effect of the dead load and the crack opening is quite large already before the peak load. In such situations, it may happen that the nonlocal transfer of damage reduces the stiffness of material in an unrealistically large band around the crack. This can lead to erroneous results if, e.g., a precracked body is unloaded and subsequently reloaded by a different type of external forces. It is therefore recommended to use weight functions with bounded support, for which the maximum width of the damaged zone Z_ω is finite.

Let us now investigate the core of the process zone, Z_ε, characterized by local strains growing beyond any bounds. It can be expected that this zone is relatively small because the band of growing local strains is shrinking during the loading process. First, consider the linear softening law (or any other law with full stress relaxation at finite strain). In the moment when the element with the largest nonlocal strain reaches the state of complete damage, the stress in this element must be zero. From equilibrium we conclude that the stress in all other elements must be zero as well. However, the other elements are not yet fully damaged (because their nonlocal strain is smaller than in the critical element), and so zero stress can be produced only by zero local strain. In other words, local strain is fully localized in one critical element, independently of the number of elements or type of nonlocal weight function. The only assumption that we made was that the local softening law gives exactly zero stress at some finite value of strain.

It might seem from the above considerations that the state at complete failure depends on the finite element discretization. This is true only to a limited extent. The final strain is localized in one element of size $h = L/n$ where n is the total number of elements. Outside this element, local strain is zero. Local strain in the failing element is determined by the condition that the nonlocal strain in that element should be equal to a material parameter ε_f, which specifies the point in the local stress-strain diagram where

the stress first vanishes. The nonlocal strain evaluated at the integration point x_0 (element center) is

$$\overline{\varepsilon} = h\alpha(x_0, x_0)\varepsilon \tag{8.39}$$

and the local strain can be expressed as

$$\varepsilon = \Delta L/h \tag{8.40}$$

where ΔL is the total bar elongation. Substituting into the critical condition $\overline{\varepsilon} = \varepsilon_f$ we get

$$\Delta L = \frac{\varepsilon_f}{\alpha(x_0, x_0)} \tag{8.41}$$

which means that the bar elongation at complete failure depends only on material parameters and not on the finite element discretization. It is obvious that in the limit for $n \to \infty$ the local strain distribution tends to ΔL times the Dirac delta function. The zone Z_ε therefore contains only one point, $Z_\varepsilon = \{x_0\}$, and the zone of complete damage is an interval of length $2R$ centered at x_0,

$$Z_\omega = (x_0 - R, x_0 + R) \tag{8.42}$$

The same statement can be justified even for the exponential softening law. The proof is relatively lengthy and is therefore omitted.

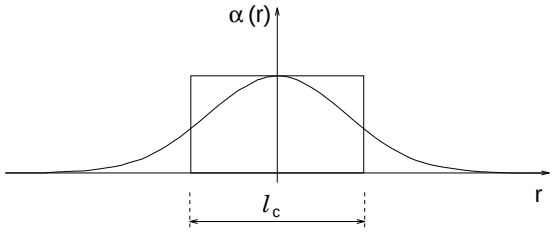


Figure 8.7: Definition of characteristic length

Formula (8.41) can be given an interesting physical meaning. If the point x_0 is sufficiently far from the boundary we have

$$\alpha(x_0, x_0) = \frac{\alpha_0(0)}{\displaystyle\int_{-\infty}^{\infty} \alpha_0(r)\,\mathrm{d}r} \tag{8.43}$$

We can define the characteristic length ℓ_c such that the area of the rectangle of width ℓ_c and height $\alpha_0(0)$ is the same as the area under the graph of the weight function; see Fig. 8.7. By definition,

$$\ell_c = \frac{\displaystyle\int_{-\infty}^{\infty} \alpha_0(r)\,\mathrm{d}r}{\alpha_0(0)} = \frac{1}{\alpha(x_0, x_0)} \tag{8.44}$$

and so (8.41) can be written as $\Delta L = \varepsilon_f \ell_c$. The total elongation at failure is equal to the local strain at failure times the characteristic length.

Note that the internal length ℓ defined by (7.61) and the characteristic length ℓ_c defined by (8.44) are two different quantities. For a given type of weight function their ratio is fixed but different from 1.

8.4 Energy Dissipation in Nonlocal Continuum

8.4.1 Shape of Load-Displacement Curve

As we have seen, formulations $\omega(\overline{\varepsilon})$ and $\overline{Y}$ give full stress relaxation but the shape of the load-displacement diagrams in Fig. 8.2a,b,d is not realistic. In the previous section we explained that the value of total bar elongation at complete failure is determined by the local strain at zero stress and the characteristic length of the nonlocal weight function. The diagrams therefore become more ductile if we increase the

parameter ε_f and/or the characteristic length (which is proportional to the internal length). Ductility of the local stress-strain law is determined by the ratio $\varepsilon_f/\varepsilon_0$ where $\varepsilon_0 = f_t/E$ is the strain at peak stress. Ductility of the structural response is also affected by the ratio of the internal length to the total length of the bar, ℓ/L.

The dashed line in Fig. 8.4b represents the load-displacement diagram calculated with parameters $\varepsilon_f/\varepsilon_0 = 5$ and $\ell/L = 0.1$. But even if we increase $\varepsilon_f/\varepsilon_0$, the calculated response still exhibits snapback. This happens because the local softening law is linear and the zone of increasing local strain is shrinking as damage progresses. Better results are obtained for the exponential softening law; see the solid curve in Fig. 8.4b. In this case, the shrinking of the zone of localized strain is less dramatic and the long tail of the local stress-strain curve forces the global curve to develop a long tail as well. This should be kept in mind when selecting an appropriate local law for a nonlocal simulation. Linear softening is sometimes used as a crude approximation of the actual softening curve. In the context of the crack-band approach this usually leads to results that do not perfectly match experimental data but are within a reasonable tolerance. However, in nonlocal simulations linear softening laws have to be avoided because, except for early nonlinear response, the results are totally unrealistic.

8.4.2 Dissipated Energy

Selection of suitable material parameters for a nonlocal simulation is further complicated by the fact that the response is affected not only by the local law but also by the internal length. An important material property that has to be properly reflected by models used in fracture simulations is the fracture energy. Let us leave aside the problems with an objective definition of fracture energy as a true material property that should not depend on the size of the specimen, type of loading, etc. For the present purpose we simply look at the total energy that our model dissipates in a uniaxial tensile test. Suppose that the desired value of this energy is prescribed (known from experiments) and we have to select the model parameters accordingly. For the standard crack-band approach this task is very easy because strain localizes in one element and the total energy dissipation is equal to the area under the stress-strain curve (dissipation per unit volume) multiplied by the volume of the element. However, for a nonlocal model the size of the zone in which dissipation takes place varies as damage progresses, and it is not clear how to determine the dissipation volume. At first we might think that one should use the volume of the zone Z_ω in which damage converges to one. But for a nonlocal model the dissipation density in this zone is not constant. Note that for a local model the density of energy dissipated at a material point is a unique function of the final value of the damage parameter at that point and for $\omega = 1$ corresponds to the area under the stress-strain curve. On the other hand, for the nonlocal model the graph of stress versus local strain does not follow a unique curve because the damage parameter depends on the values of strain in the neighborhood of that point. This is illustrated in Fig. 8.8 showing the local stress-strain law and the actual stress-strain evolution for integration points of five different elements. Points that are located close to the boundary of the damaged zone start unloading (in the sense of decreasing local strain) sooner than points around the center of the zone. If we replot the graph in the logarithmic scale (Fig. 8.8b) it becomes clear that only the rightmost curve tends to infinity while all the other curves return to the origin. Even though all of the points are inside the zone where damage tends to one, the density of dissipation (area under the stress-strain curve) varies from very large values around the center of the zone to small values at its boundary.

The total energy dissipated by a nonlocal model can be determined by numerical integration of the area under the load-displacement curve. Dividing the total dissipated energy by the dissipation density that corresponds to the local model (and by the cross-sectional area, which has a unit value in our calculations) we obtain a certain length that we suggest to call the dissipation length, ℓ_d. It represents the width of a cracking band in an equivalent local continuum that would dissipate the same energy as the actual band formed in the nonlocal continuum. Unlike the width of the damaged zone Z_ω, the dissipation length is not proportional to the internal length imposed by the nonlocal weight function. Of course, it increases with increasing ℓ but it is also affected by other factors.

Let us characterize the local constitutive law by the ductility parameter η, defined as the ratio of the local dissipation density to the elastic energy density at peak. For the linear softening law, $\eta = \varepsilon_f/\varepsilon_0 = E\varepsilon_f/f_t$ is simply the ratio of the strain at complete failure to the limit elastic strain, while for the exponential softening law we have $\eta = 2\varepsilon_f/\varepsilon_0 - 1$. Recall that, for the exponential law, ε_f is the strain at the intersection of the horizontal axis with the tangent to the softening curve at peak stress. Fig. 8.9 shows the relative dissipation length ℓ_d/ℓ as a function of the ductility parameter for different

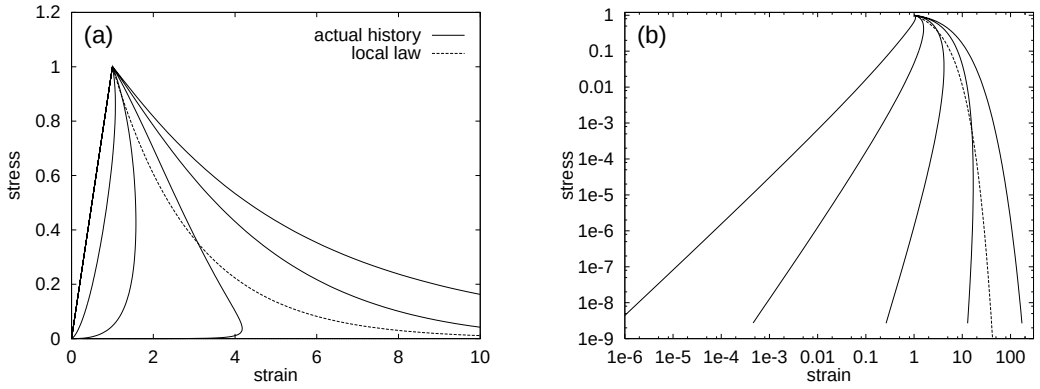


Figure 8.8: Actual stress-strain histories compared to the local law: a) natural scale, b) logarithmic scale

types of nonlocal weight functions and local softening laws. Each curve in the plot is labeled by three letters. The first letter denotes the type of softening law (L=linear, E=exponential), the second letter depends on the type of nonlocal weight function (B=bell function, G=Gauss function), and the last letter indicates the type of nonlocal formulation (Y=formulation $\overline{Y}$, E=formulation $\omega(\overline{\varepsilon})$). For highly ductile local laws the dissipation length varies only slightly and is almost independent of the type of softening law and of the type of nonlocal formulation. Its value is somewhat smaller for the Gauss weight function than for the bell-shaped function.

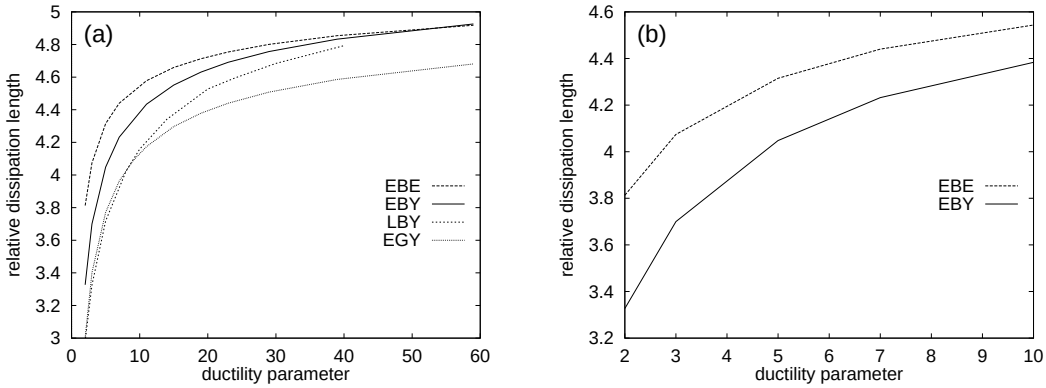


Figure 8.9: Relative dissipation length as function of the ductility parameter: a) global picture, b) close-up

8.4.3 Parameter Identification

Substantial variation of the dissipation length is observed for ductility parameters smaller than 10. For two important cases (exponential softening law, bell-shaped weight function, nonlocal formulations $\overline{Y}$ and $\omega(\overline{\varepsilon})$) the curves are replotted in the range $2 \leq \eta \leq 10$ in Fig. 8.9b. The graphs can be exploited for the selection of model parameters. Suppose that we want to set up a model with Young's modulus $E = 25$ GPa, uniaxial tensile strength $f_t = 3$ MPa, and fracture energy (meaning here the energy per unit area dissipated in the uniaxial tensile test) $G_f = 80$ N/m. Furthermore, suppose that based on some relationship between the material microstructure and the internal length we decide to set $\ell = 20$ mm. The nonlocal model should be based on formulation $\overline{Y}$, and our goal is to design an appropriate local constitutive law. In view of the recommendations resulting from the theoretical analysis in previous sections of this report, we decide to use exponential softening and the bell-shaped weight function. The

interaction radius, $R = \ell\sqrt{7} = 53$ mm, is easy to determine. We also know the parameters E and f_t of the local constitutive law. The remaining parameter to be determined, ε_f, controls the fracture energy. It can be evaluated in an iterative manner. We start with an initial estimate for the relative dissipation length, e.g., we take the high-ductility limit, $\ell_d/\ell = 5$. For the corresponding dissipation length, $\ell_d = 5\ell = 100$ mm, the local dissipation density would have to be $g_f = G_f/\ell_d = 80/0.1$ N/m^2 $= 800$ N/m^2. As the elastic energy density at peak stress is $g_p = f_t^2/2E = 180$ N/m^2 we would have to select a law with ductility parameter $\eta = g_f/g_p = 800/180 = 4.44$. For this ductility parameter we can find the corresponding relative dissipation length $\ell_d/\ell = 3.95$ from Fig. 8.9b. Now we repeat the whole procedure and we obtain improved estimates $\ell_d = 79$ mm, $g_f = 1010$ N/m^2, and $\eta = 5.6$. For this value the relative dissipation length from Fig. 8.9b is $\ell_d/\ell = 4.1$, and we get $\ell_d = 82$ mm, $g_f = 975$ N/m^2, and $\eta = 5.4$. The corresponding relative dissipation length would differ only a little from the previous iteration, and so we can accept the current estimate. For the exponential law, the ductility parameter is $\eta = 2\varepsilon_f/\varepsilon_0 - 1$, and therefore $\varepsilon_f = \varepsilon_0(\eta+1)/2 = f_t(\eta+1)/2E = 3 \times 6.4/2 \times 25000 = 0.384 \times 10^{-3}$.

8.5 Computational Issues

8.5.1 Mesh-Induced Directional Bias

Even if sensitivity to the element size is suppressed by the mesh-dependent softening modulus technique or by a localization limiter, the results may still exhibit excessive dependence on the shape and orientation of finite elements. This is especially true for problems with strain localization because the direction of a simulated process zone may be biased by the structure of the mesh. In calculations using a standard continuum (even with an adjustment of the stress-strain diagram), fracture tends to propagate along certain preferred directions, which lead along element sides or across element diagonals (for quadrilateral elements). This effect is documented in Fig. 8.10, which shows the fracture patterns obtained for the three-point bending specimen from Fig. 6.4a on a skewed quadrilateral mesh. The crack should run along the vertical axis of symmetry. The standard rotating crack model exhibits strong directional bias, and the simulated crack band propagates along a preferred mesh direction; see Fig. 8.10a. This sensitivity is partly due to stress locking that would be produced if the band was not aligned with the element sides.

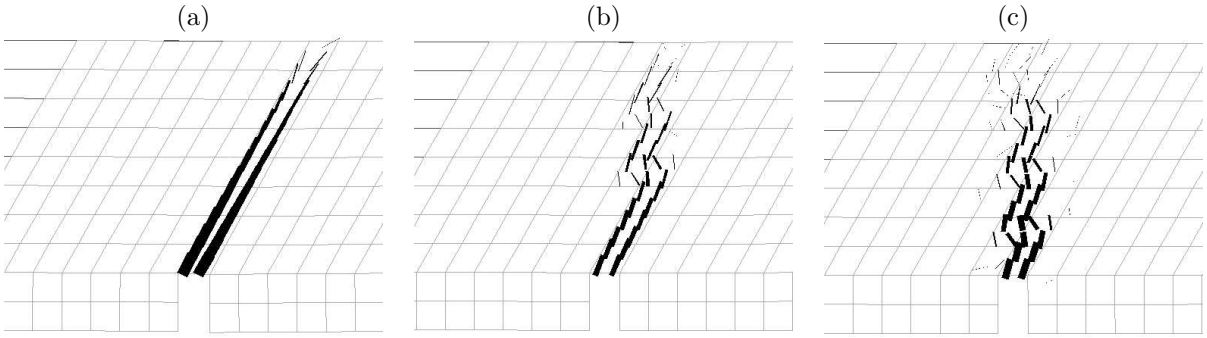


Figure 8.10: Three-point bending specimen with a notch: crack patterns simulated by a) standard RC model, b) local RC-SD model, c) nonlocal RC-SD model

The RC-SD model [154] described in Section 6.3 combines the rotating crack with a scalar damage concept. Transition to a damage formulation alleviates stress locking and consequently leads to an improved crack trajectory; see Fig. 8.10b. However, the effect of mesh bias is still considerable. A nonlocal formulation leads to a fracture pattern that approximates the actual one as well as possible on this relatively coarse mesh; see Fig. 8.10c. The directions of "local cracks" plotted as dark rectangles at individual Gauss points oscillate and are not aligned with the overall crack trajectory. However, this is quite natural because these local cracks, defined just for the purpose of visualization of the results, are computed from the local inelastic strains. At very late stages of the stiffness degradation process, local strains localize in a single layer of elements, even though the model is nonlocal. This cannot be judged as a deficiency of the model. What matters is that the load-displacement diagram and the overall direction of the cracking band are reproduced correctly, independently of the mesh size and orientation.

8.5.2 Efficient Implementation of Nonlocal Averaging

Numerical implementation of the nonlocal damage model based on averaging of equivalent strain is relatively straightforward. The evaluation of stresses from given strains remains fully explicit, and no internal iteration loop is needed. Of course, equilibrium iteration on the structural level cannot be avoided, same as for any other nonlinear model. An existing computer code with a certain local damage model can be extended to a nonlocal formulation without excessive effort. All that one needs is to implement the algorithm of weighted spatial averaging and, before damage is evaluated, replace the local equivalent strain by its nonlocal counterpart.

The values of nonlocal equivalent strain must be traced at individual Gauss integration points of the finite element model, because these are the points at which stresses need to be evaluated. Let us denote the coordinates of Gauss points as $\boldsymbol{x}_k$, $k = 1, 2, \ldots N_{GP}$, where N_{GP} is the total number of Gauss points in the finite element model. In the numerical algorithm, the averaging integral (7.63) is replaced by

$$\bar{\varepsilon}_k = \sum_{l=1}^{N_{GP}} w_l J_l \alpha_{kl} \tilde{\varepsilon}_l \tag{8.45}$$

where w_l is the integration weight of Gauss point number l, J_l is the Jacobian of the isoparametric transformation evaluated at this point, and α_{kl} is the weight of nonlocal interaction between points k and l, determined as

$$\alpha_{kl} = \frac{\alpha_0(\|\boldsymbol{x}_k - \boldsymbol{x}_l\|)}{\sum_{m=1}^{N_{GP}} w_m J_m \alpha_0(\|\boldsymbol{x}_k - \boldsymbol{x}_m\|)} \tag{8.46}$$

It is important to note that α_{kl} vanishes if the distance $\|\boldsymbol{x}_k - \boldsymbol{x}_l\|$ between points k and l is larger than the nonlocal interaction radius R. So the sums in (8.45) and in the denominator of (8.46) do not need to be taken over all Gauss points l but only over those that are located inside the sphere or circle of radius R centered at point k. Moreover, since the factors $a_{kl} \equiv w_l J_l \alpha_{kl}$ are needed at every iteration of every incremental step and do not vary during the simulation, they should be evaluated only once and then stored, rather than recomputed each time they are needed.

An efficient numerical scheme for nonlocal averaging has the following structure:
Loop over all Gauss points $\boldsymbol{x}_k$, $k = 1, 2, \ldots N_{GP}$.

1. Find all Gauss points $\boldsymbol{x}_l$ whose distance from $\boldsymbol{x}_k$ is smaller than R, and for each of them evaluate $a_{kl} = w_l J_l \alpha_0(\|\boldsymbol{x}_k - \boldsymbol{x}_l\|)$.
2. Compute the sum $a_k = \sum_l a_{kl}$.
3. Divide each a_{kl} by a_k and store it in a table along with a reference to point l. The table is associated with point k.

This procedure is activated as a part of the initialization tasks performed before the actual incremental-iterative solution of the nonlinear problem starts. Each Gauss point is associated with a nonlocal interaction table that gives access to its "neighbors" (i.e., to Gauss points at distance smaller than R) and to the corresponding interaction coefficients a_{kl}.

The search for nonlocal neighbors can be done by checking the distance between all pairs of Gauss points, but for larger problems this step can become a bottleneck of the computation. A much more efficient implementation is based on the quadtree (in two dimensions) or octree (in three dimensions) technique. The entire body is placed in a rectangle or cube that is hierarchically divided into smaller rectangles or cubes down to the required level, and this structure is stored in a tree. Each Gauss point is assigned to one of the leafs of the tree, depending on its position in space. The search for neighbors can then be restricted to a limited number of leafs that are easily identified and accessed.

The stress evaluation procedure repeatedly called during the incremental-iterative solution makes use of the nonlocal interaction tables when the nonlocal equivalent strain (or any other nonlocal variable used by the specific nonlocal model) is computed. To obtain $\bar{\varepsilon}_k$, it is sufficient to get the local equivalent strains $\tilde{\varepsilon}_l$ at all points l that are included in the table associated with point k, multiply each of them by the coefficient a_{kl} and take a sum of these products.

Nonlocal averaging requires a certain extra effort as compared to the corresponding local model, but the convergence of the equilibrium iteration on the global level is usually more regular for the nonlocal model than for the local one. The reason is that the spatial distribution of strain is relatively smooth

even after the onset of localization, and the transition from a diffuse damage pattern to a localized one is not so abrupt. Consequently, the increased numerical effort per iteration is partially compensated by the reduced number of iterations needed to restore equilibrium. Since the nonlocal model completely removes the pathological sensitivity to the mesh size and also at least partially alleviates the mesh-induced directional bias, this extra effort is indeed worthwhile.

8.5.3 Nonlocal Tangent Stiffness

In general terms, the structure of the tangent operator for nonlocal damage models was first discussed in [243]. A consistent derivation and algorithmic implementation of the tangent stiffness matrix was presented in [147]. The final formula can be written in the engineering notation as

$$\boldsymbol{K} = \boldsymbol{K}_u - \sum_{k,l} g'_k w_k J_k a_{kl} \boldsymbol{B}_k^T \bar{\boldsymbol{\sigma}}_k \boldsymbol{\eta}_l^T \boldsymbol{B}_l \tag{8.47}$$

where

$$\boldsymbol{K}_u = \sum_k w_k J_k (1-\omega_k) \boldsymbol{B}_k^T \boldsymbol{D}_e \boldsymbol{B}_k \tag{8.48}$$

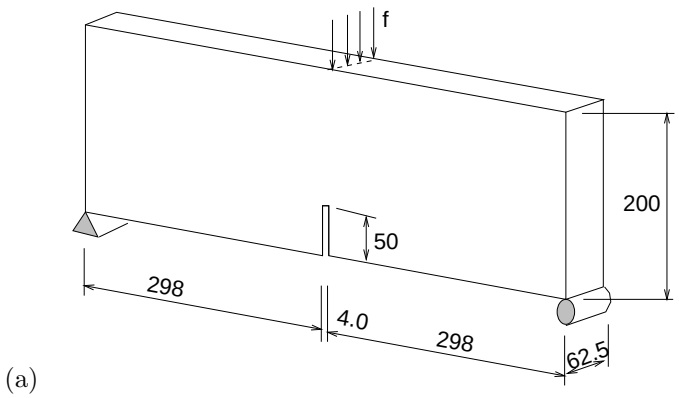
In the above, subscripts k and l refer to the integration points of the finite element model, w_k are the corresponding integration weights, J_k are Jacobians reflecting the size of the contributing volume around each integration point, $\boldsymbol{B}$ is the usual B-matrix containing the derivatives of the shape functions, g' denotes the derivative of the damage function g with respect to its argument κ and is set to zero if unloading takes place at the corresponding integration point, coefficients a_{kl} that express the strength of nonlocal interaction between integration points number k and l were defined in the preceding subsection, $\boldsymbol{D}_e$ is the elastic material stiffness matrix, $\bar{\boldsymbol{\sigma}} = \boldsymbol{D}_e \boldsymbol{\varepsilon}$ is the column matrix with the effective stress components, and $\boldsymbol{\eta} = \partial \tilde{\varepsilon} / \partial \boldsymbol{\varepsilon}$ is the column matrix with derivatives of the equivalent strain with respect to the engineering strain components.

The first term in (8.47), $\boldsymbol{K}_u$, represents the secant stiffness matrix valid if all the material points are in the unloading regime (when $g'_k = 0$ for all k). The double sum in (8.47) provides a correction of the secant stiffness due to additional damage growth. The individual terms in the sum represent the contributions of nonlocal interaction between integration points $\boldsymbol{x}_k$ and $\boldsymbol{x}_l$ to the overall stiffness. Of course, the sum needs to be taken only over those pairs of integration points whose distance is smaller than the interaction radius R, because for all the other pairs a_{kl} vanishes. Each pair of interacting Gauss points contributes to only a small block of the global stiffness matrix, with rows corresponding to internal forces at the element that contains point $\boldsymbol{x}_k$, and with columns corresponding to nodal displacements at the element that contains point $\boldsymbol{x}_l$. The size of this block is therefore equal to the size of the usual element stiffness matrix. This means that the complete stiffness matrix can be assembled from much smaller matrices, similar to the usual assembly procedure. The difference is that the code numbers associated with the rows are in general different from the code numbers associated with the columns. This can be handled by a minor modification of the usual assembly routine. Compared to the local case, the bandwidth increases due to the nonlocal interaction, and the global stiffness matrix is in general not symmetric.

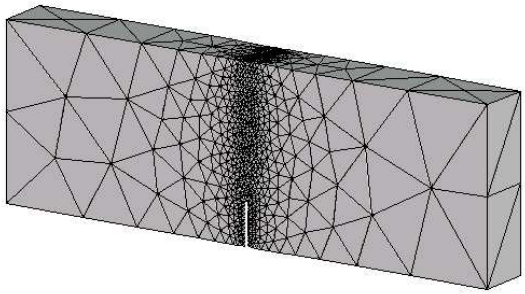
As follows from (8.47), the nonlocal tangent stiffness matrix has several particular properties that have to be taken into account when selecting an economical storage scheme and an efficient solver.

First of all, the stiffness matrix is nonsymmetric, but this is often the case for local models as well. A symmetric tangent stiffness of a local damage model is obtained only if the loading function is expressed in terms of the thermodynamic forces conjugate to the internal variables and the evolution laws are postulated as normality rules. Only the model relating damage to the damage energy release rate meets these criteria. Any other definition of equivalent strain gives linearly independent $\boldsymbol{b}_\sigma \equiv \boldsymbol{B}^T \bar{\boldsymbol{\sigma}}$ and $\boldsymbol{b}_\eta \equiv \boldsymbol{B}^T \boldsymbol{\eta}$, and the symmetry is lost already in the local case. Associated models with nice symmetry properties are appealing from the theoretical point of view, they limit the number of constitutive parameters and facilitate the numerical implementation, but they do not always capture all the essential features of the real material behavior. Therefore, the lack of symmetry cannot be considered as a substantial drawback specific to nonlocal formulations.

A more important complication caused by nonlocality is the evolutive character of the profile of the stiffness matrix. For the class of material models considered here, the elastic response remains local, and so the initial distribution of nonzero entries in the stiffness matrix is the same as in the standard local case. When the damage threshold is exceeded and the damage zone starts evolving, new nonzero entries



(a)



(b)

Figure 8.11: Three-point bending test simulated in three dimensions: (a) geometry and loading, (b) finite element mesh

appear due to the activated interaction between the Gauss points belonging to different elements, and the profile of the stiffness matrix must be dynamically adapted. Of course, one can also allocate the storage space for all possible nonzero entries from the very beginning, but this would be a big waste of resources because the process zone usually extends over only a small part of the structure and most of the allocated entries would remain zero throughout the entire simulation. If the final extent of the process zone is known in advance, one could allocate only those entries that will later become nonzero. However, such *a priori* information is available only in some academic examples but not in general applications.

An efficient and versatile solution scheme must be based on dynamic memory allocation. During the evolution of the process zone, the bandwidth of the stiffness matrix increases, but the matrix still remains sparse. This must be taken into account when selecting the most appropriate solver.

The consistent tangential stiffness matrix can be exploited in the global equilibrium iteration procedure. The resulting acceleration of the convergence rate is illustrated by the example of a three-point bending test simulated in three dimensions using the nonlocal isotropic damage model with a Rankine-type definition of equivalent strain (5.26) and with exponential softening; see [147] for details. The geometry and loading are shown in Fig. 8.11a. The finite element mesh contains 6461 nodes and 33158 constant-strain tetrahedral elements; see Fig. 8.11b.

The nonlinear response is analyzed in 15 increments of applied displacements at the loaded edge. Solution strategies based on the secant stiffness matrix (SSM) and tangent stiffness matrix (TSM) with either direct or iterative solvers are exploited. The direct solver is based on LU decomposition using the generalized "skyline" storage scheme and the profile optimization algorithm proposed in [293]. The iterative solver is based on the Generalized Minimum Residual Method (GMRES) using the compressed row (CompRow) storage scheme and preconditioning by an incomplete LU factorization (ILU) with fill-up [272].

Fig. 8.12a shows the evolution of error (norm of the out-of-balance forces) during equilibrium iteration with the TSM. The curves correspond to three typical steps, denoted as A, B and C. The convergence rate is quadratic, which confirms that the tangent stiffness is indeed consistent, and convergence is achieved within 3 to 7 iterations.

For a high required accuracy (relative tolerance of the out-of-balance forces set to 10^{-6}), the complete analysis with a secant stiffness and a direct solver takes 27 hours and 37 minutes (of the total user time),

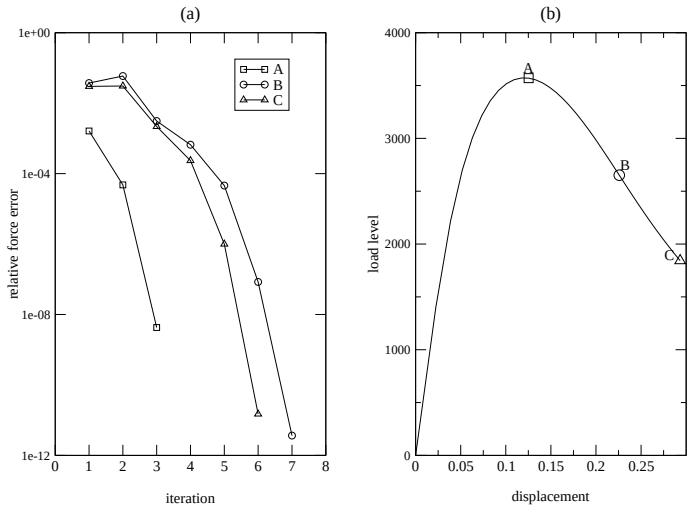


Figure 8.12: Three-point bending test simulated in three dimensions: (a) convergence characteristics for tangent stiffness, (b) load-displacement diagram

the analysis with a tangent stiffness and a direct solver takes 136 hours and 41 minutes, but the analysis with a tangent stiffness and an iterative solver takes only 15 hours and 23 minutes. The simulations have been run on a Pentium III workstation with 1 GB of memory, running at 933 MHz.

8.6 Mesh-Adaptive Techniques

8.6.1 General Structure of the Adaptive Procedure

Nonlocal models lead to smooth solutions with a continuous variation of strain. However, to resolve narrow bands of highly localized strains, it is necessary to use sufficiently fine computational grids. Fortunately, the mesh must be fine only in the process zone, while the remaining part of the structure can be reasonably represented by a coarser mesh. The localization pattern is in general not known in advance, and it is extremely tedious to construct suitably refined meshes "by hand". Efficiency of the analysis can be greatly increased by an adaptive technique, which automates the whole process.

The basic components of the adaptive procedure include

- an error estimator or indicator,
- a remeshing criterion,
- transfer algorithms for primary unknowns and internal variables, and
- a mesh generator interface.

The general algorithm of nonlinear adaptive analysis can be described as follows: After reaching the equilibrium state (corresponding to a given load increment) and updating the solution state, an *a posteriori* error estimation is performed, in order to evaluate the error distribution. Then, a remeshing criterion uses the information about the error distribution and determines the subsequent activity. If the obtained error level is still acceptable, the analysis continues with the next load increment on the current mesh. If the evaluated error exceeds a limit defined by the user, the required mesh density is determined from the error distribution and a new spatial discretization is generated.

In a truly adaptive approach, after generating a new discretization, the data structures corresponding to the newly generated mesh are created, and the transfer of displacements and internal variables from the old mesh to the new one is performed. After the mapping, the internal variables are used together with the strain computed from the mapped displacements to update the internal state of each integration point on the new mesh (to achieve local consistency). Once the transfer has been finished, the old discretization is deleted and the mapped configuration is brought into global equilibrium by iteration at constant value

of the loading parameter (e.g., of the applied displacement). Afterwards, the solution continues with the next load increment.

Another possibility is to restart the analysis from the initial state after the new discretization has been generated [150, 261]. This approach does not require the transfer of the current state from the old discretization to the new one, but from the computational point of view is less efficient compared to the truly adaptive approach, especially if the remeshing is done frequently.

8.6.2 Error Estimators and Indicators

For linear problems, error analysis of the finite element solution can be developed in a mathematically rigorous way [11, 337]. In the nonlinear range, however, rigorous error estimates can be constructed only for a restricted class of problems. A general theory is not available since there are various sources and forms of nonlinearities. In the linear elastic case the error arises essentially from the discretization of the domain (so-called spatial error). In the nonlinear case, the error depends on the time discretization for history-dependent solids, and a part of the error is always induced by the incremental-iterative technique. The path dependency renders the problem more complex and, consequently, a reliable error estimation becomes more difficult, especially for nonconventional theories of enriched continua. Nevertheless, considerable progress has been made in recent years. For instance, in [261] the authors proposed a sophisticated error estimator for nonlocal damage models. Ladevèze and coworkers [175, 105] developed *a posteriori* estimators based on the error in the constitutive relation, and Comi and Perego [60] adapted this technique to their nonlocal damage theory [59]. However, the implementation of these complicated estimators requires a considerable effort.

A simple and convenient alternative to rigorous error estimators is provided by heuristic error indicators. They are often based on physical intuition and insight into the problem at hand. Error indicators are typically based on variables that characterize the nonlinear dissipative process, e.g., on the inelastic strain. In the present work we use directly the damage variable ω for isotropic damage models, and the maximum principal value of the damage tensor for anisotropic damage models. Therefore, we will use the expression "damage indicator" instead of "error indicator".

8.6.3 Remeshing Criterion

A high value of the damage indicator allows to identify the process zone and to determine the required mesh size. The maximum allowable mesh size inside the process zone is related to the nonlocal interaction radius R. It should be a fraction of R, to make sure that the nonlocal interactions between individual integration points are properly activated. The minimum requirement needed for nonlocal averaging to work reasonably well is to have at least four remote integration points influencing the given point in an equivalent one-dimensional situation. Extending this simple one-dimensional rule, the corresponding maximum element size can be estimated. For example, for quadratic elements with 3×3 integration points the maximum size allowed in the process zone is $h_1 = 3 \times 2R/4 = 1.5R$. The optimal recommended value that allows to accurately represent the localized character of the solution and is still feasible from the computational point of view is about 10 integration points per width of the process zone. The proposed remeshing criterion, based on the damage indicator ω, uses element size h_0 for $\omega \leq \omega_0$ and element size h_1 for $\omega \geq \omega_1$, with linear interpolation in between. Here, ω_0 is typically selected as damage at peak stress in a uniaxial test (i.e., zero for models that remain linear elastic up to the peak), ω_1 is taken, e.g., as the damage level corresponding to softening to 50 % of the peak stress, h_0 is a typical element size in a coarse mesh, and h_1 is the element size that enables good nonlocal interaction.

As an additional improvement, the remeshing criterion based on the damage indicator is combined with an error estimator for elastic problems, e.g., with the Zienkiewicz-Zhu estimator [337]. This allows to optimize the mesh in the elastic or slightly damage part of the domain. The transition of the mesh size between the linear elastic part and the nonlinear damaged part can be handled by the mesh generator or by a suitable smoothing technique.

8.6.4 Transfer Operators

The truly adaptive approach requires the mapping of displacements and internal history variables (in the present case of the damage variable or damage tensor), allowing to continue the analysis from the current state, instead of restarting the analysis from scratch after each mesh refinement. Mapping of the primary

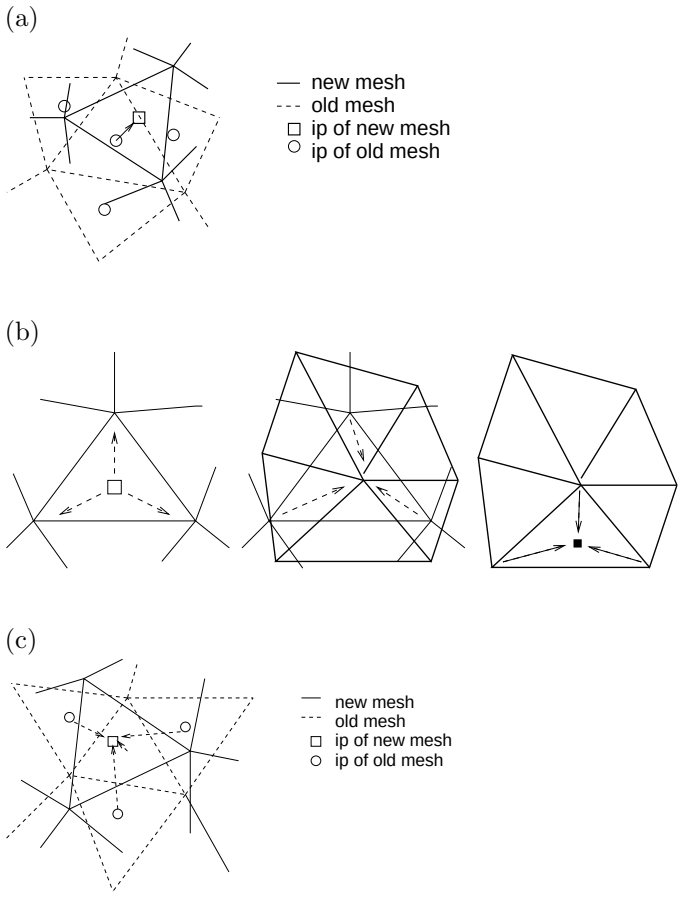


Figure 8.13: Mapping techniques: (a) closest point transfer, (b) shape function projection, (c) least square projection

unknowns (nodal displacements) is usually done using the shape function projections. Mapping of the damage variables is more intricate. It can be done using one of the following techniques:

- CPT – closest point transfer (Fig. 8.13a)
- SFT – shape function projection transfer (Fig. 8.13b)
- LST – least-square projection transfer (Fig. 8.13c)

A comparative study [148] revealed that SFT leads to an artificial damage diffusion, i.e., to an expansion of the damaged zone after the mapping, while CPT and LST preserve the size of the damaged zone. CPT results into lower stress oscillations after mapping, but when these oscillations are removed by equilibrium iterations (at constant loading parameter), the original shape of the damage profile is best reproduced by LST. It can be concluded that SFT is not a suitable mapping technique for the damage variable, and that LST is probably the most accurate technique but computationally more expensive than CPT, which provides acceptable results at low cost.

Spurious damage diffusion produced by SFT is confirmed in a two-dimensional example of a four-point shear test on a single-edge-notched specimen; see the damage patterns in Fig. 8.14.

8.6.5 Examples

The proposed adaptive methodology has been successfully implemented by the first author into his object-oriented finite element system OOFEM [236]. To illustrate its performance, we present two examples of failure simulations.

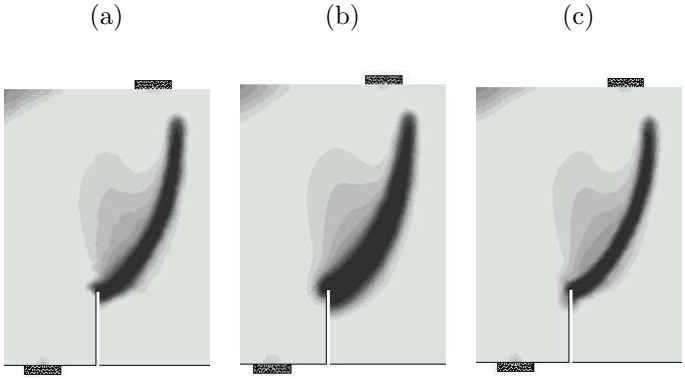


Figure 8.14: Four-point shear test — comparison of damage patterns for different transfer operators: (a) CPT, (b) SFT, (c) LST

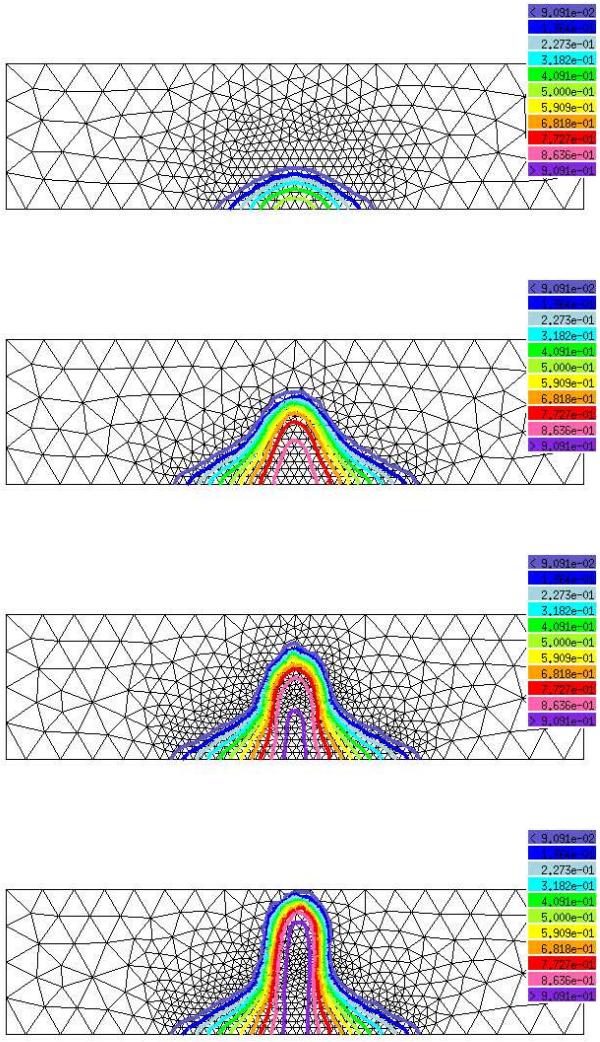


Figure 8.15: Adaptive analysis of damage evolution in a three-point bending test

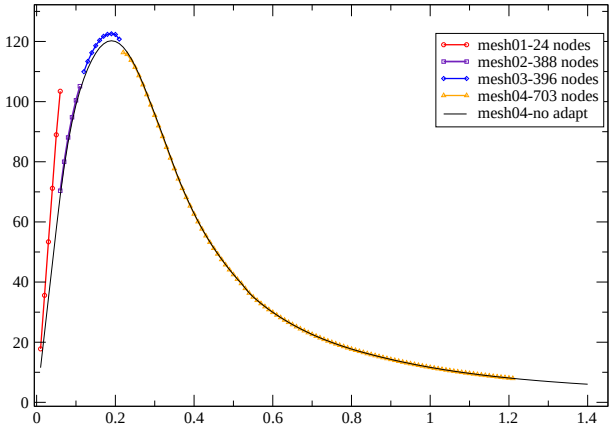


Figure 8.16: Load-displacement curve for three-point bending test

Three-Point Bending Beam

The first example illustrates the analysis of a three-point bending test. The nonlocal formulation of the isotropic damage model with the Rankine-like equivalent strain measure (5.26) and the exponential damage law (5.29) has been used. The constitutive properties have been set to: Young's modulus of elasticity $E = 30$ GPa, Poisson's ratio $\nu = 0.3$, limit elastic strain $\varepsilon_0 = 0.0001$, equivalent strain corresponding to the fully damaged state $\varepsilon_f = 0.012$. The bell-shaped nonlocal average function (7.60) has been used with nonlocal averaging radius $R = 10$ mm. The damage-based error indicator has been used together with the remeshing criterion based on a linear interpolation scheme. The mesh size has been set to $h_0 = 40$ mm in the elastic regions and to $h_1 = 10$ mm in the damaged regions.

The initial coarse mesh contained only 24 nodes and 28 constant strain elements. Only four remeshing steps were necessary to obtain the final mesh, containing 703 nodes and 1328 elements. The initial damage pattern is rather diffuse because there is no initial notch, but later it gets localized, see Fig. 8.15. The loading has been controlled by a prescribed displacement under the applied force. The resulting load-displacement diagram is shown in Fig. 8.16. In this plot, the parts of diagram are distinguished as corresponding to different adaptive steps. The load-displacement diagram obtained by a standard (non-adaptive) analysis on the final mesh is also shown. A very good agreement has been obtained, demonstrating the quality of the transfer operators. The differences in the initial part of the load-displacement diagram are mainly due to the fact that the initial meshes are too coarse and thus naturally lead to a much stiffer response than the final fine mesh.

DEN Specimen with Curved Cracks

As another example of damage propagation along a curved trajectory, we present the adaptive simulation of the double-edge-notched (DEN) specimen that was tested by Nooru-Mohamed [217] using the experimental setup in Fig. 8.17a. The specimen can be subjected to a combination of shear and tension (or compression). Nooru-Mohamed performed his experiments for a number of loading paths, some of them even nonproportional. Among the most interesting loading scenarios were paths 4a, 4b and 4c. During the first stage, the specimen was loaded by an increasing "shear" force, P_s, while keeping the "normal" force, P, at zero. After reaching a certain load level, the type of loading was changed. During the second stage, the force P_s was kept constant and the test was controled by increasing the "normal" displacement δ. For path 4a the change of loading occured at $P_s = 5$ kN, for path 4b at $P_s = 10$ kN, and for path 4c at the maximum shear force that the specimen could sustain, $P_s = P_s^{max} = 27.5$ kN. In all the cases, the failure pattern consisted of two macroscopic cracks propagating from the notches in an inclined direction. For path 4a, these cracks were almost horizontal and close to each other (Fig. 8.17b top), while for path 4c they were highly curved and farther apart (Fig. 8.17b bottom).

Failure of the DEN specimen under loading paths 4a and 4c has been simulated using the nonlocal version of the anisotropic damage model proposed in [141]. The material parameters have been deduced from the data provided by [217]: compressive strength measured on cubes $f_c = 46.24$ MPa for path 4a and $f_c = 46.19$ MPa for path 4c, and splitting tensile strength $f_s = 3.67$ MPa for path 4a and $f_s = 3.78$

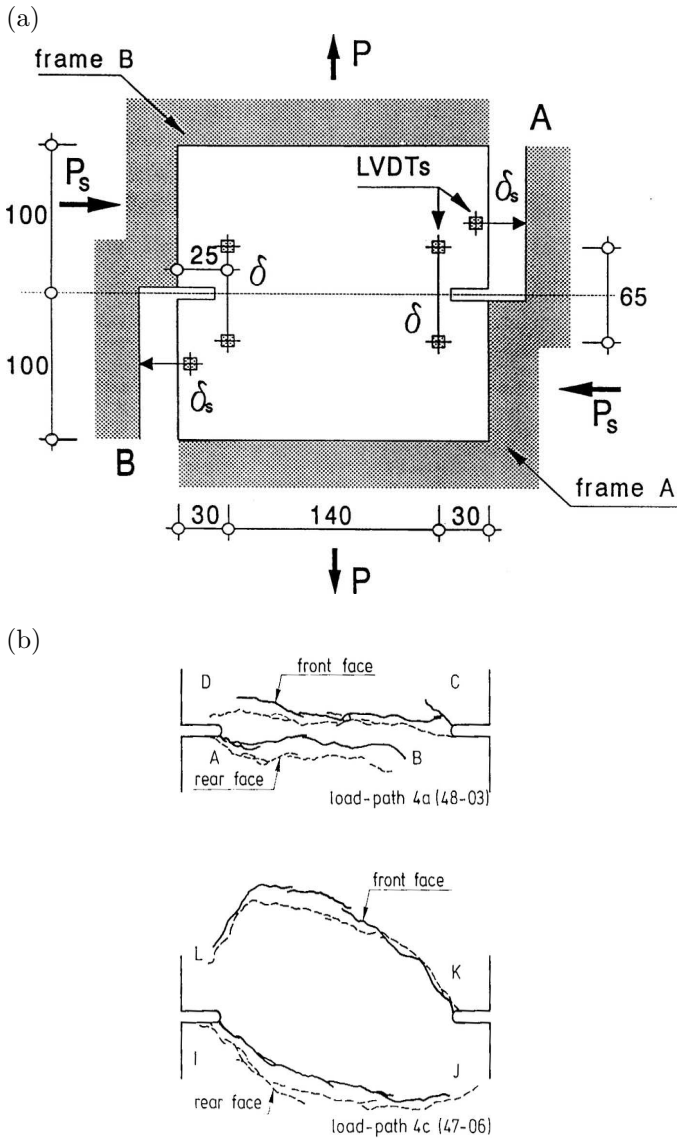


Figure 8.17: Curved crack propagation in a DEN specimen: (a) experimental setup and (b) observed crack patterns (reproduced from Nooru-Mohamed 1992)

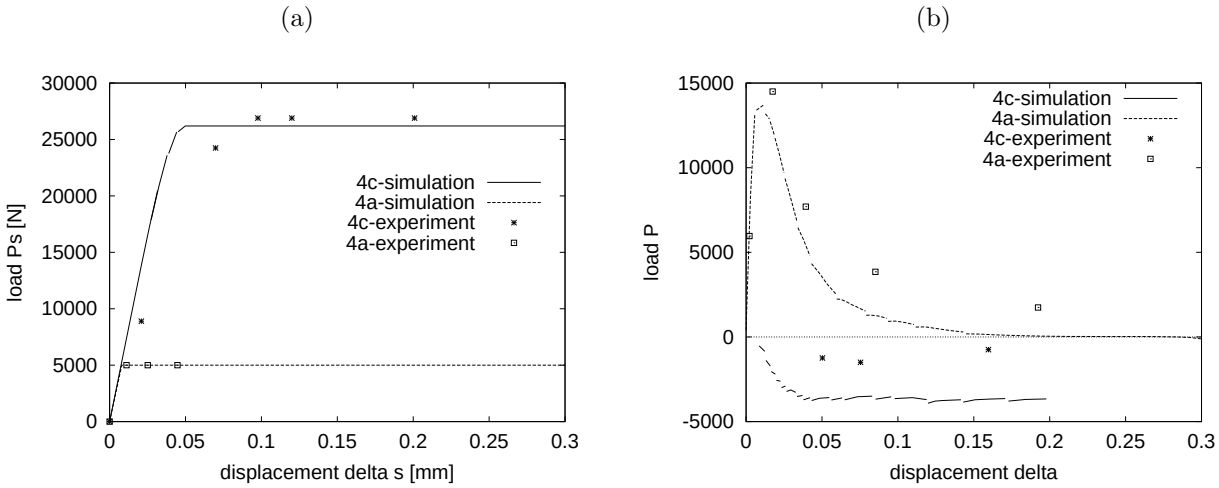


Figure 8.18: Measured and simulated load-displacement curves in terms of (a) shear force P_s and displacement δ_s, (b) normal force P and displacement δ

MPa for path 4c. The compressive strength is slightly above the value that corresponds to concrete C-30 according to the CEB-FIP Model Code (1991). Interpolation between the values of tensile strength and Young's modulus corresponding to concretes C-30 and C-40 gives $f_t = 3$ MPa and $E = 29$ GPa. The value of tensile strength is in agreement with the empirical formula $f_t = 0.8 f_s$. The fracture energy is considered by the same value as in [80], i.e., $G_f = 110$ J/m^2, and the interaction radius is set to $R = 5$ mm.

The adaptive analysis with parameters $\omega_0 = 0.6$, $h_0 = 8$ mm, $\omega_1 = 0.8$ and $h_1 = 2.8$ mm required 13 remeshings for path 4a and 23 remeshings for path 4c. The evolving meshes with the isolines of damage and strain are shown in Figs. 8.19 and 8.20. It is clear that for both loading paths the numerical prediction is in an excellent agreement with the experimental results. Even the highly curved cracks generated by path 4c are reproduced very accurately.

The load-displacement curves, plotted in Fig. 8.18, reveal certain discrepancies. For low and moderate levels of the shear force P_s, the response is linear, but the measured displacements δ_s exceed the computed displacements by as much as 50%. To get a perfect agreement in this range, the elastic modulus used in the simulation would need to be reduced to about 20 GPa. It is not very likely that the actual elastic modulus was really that low. The measured displacement was probably not only due to the elastic deformation of the specimen. The fact that the displacements are underpredicted is a general problem observed in other simulations of this test published in the literature; see e.g. [80].

Leaving aside the difference between the experimental and numerical compliances, one can say that the agreement between the test and the simulation is very satisfactory. For loading path 4a, the response to the shear loading is linear and the normal loading generates a tensile reaction force increasing up to 15 kN in the test and 14 kN in the simulation. For loading path 4c, the first clearly visible deviation from linearity during the shear loading appears at $P_s = 20$ kN in both test and simulation, and the peak shear force is 27.5 kN in the test and 26.2 kN in the simulation. The simulation correctly predicts that the reaction force generated by increasing the normal displacement at constant shear force is compressive. The maximum magnitude of the normal force is 1.5 kN in the test and below 4 kN in the simulation, which should not be considered as a huge relative error because these forces are very small compared to the previous loading case. The compressive force generated by an extension of the specimen can be explained by the fact that the resistance of the specimen was already exhausted by the shear force, and adding a tensile force under load control would lead to an unstable crack propagation.

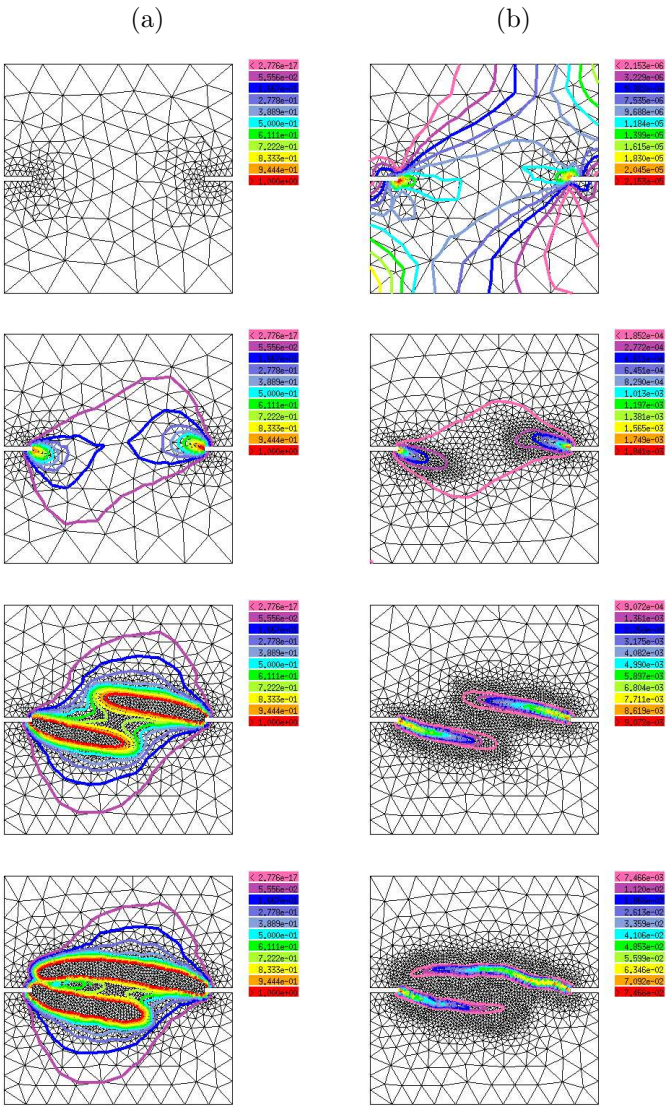


Figure 8.19: Loading path 4a: evolving meshes with isolines of (a) damage and (b) strain

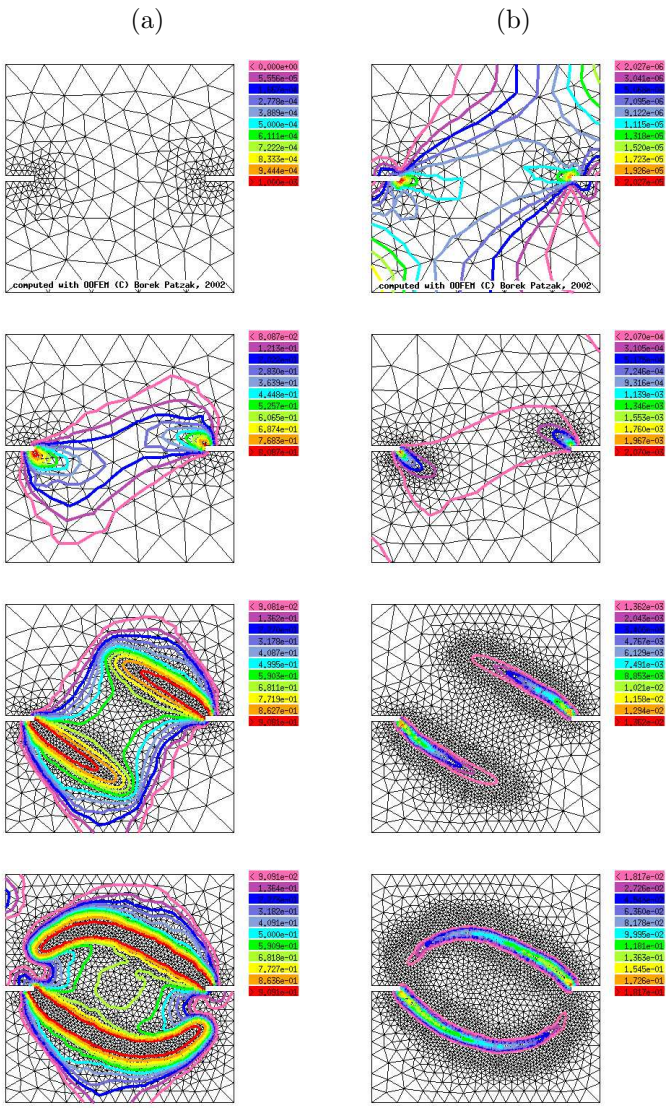


Figure 8.20: Loading path 4c: evolving meshes with isolines of (a) damage and (b) strain