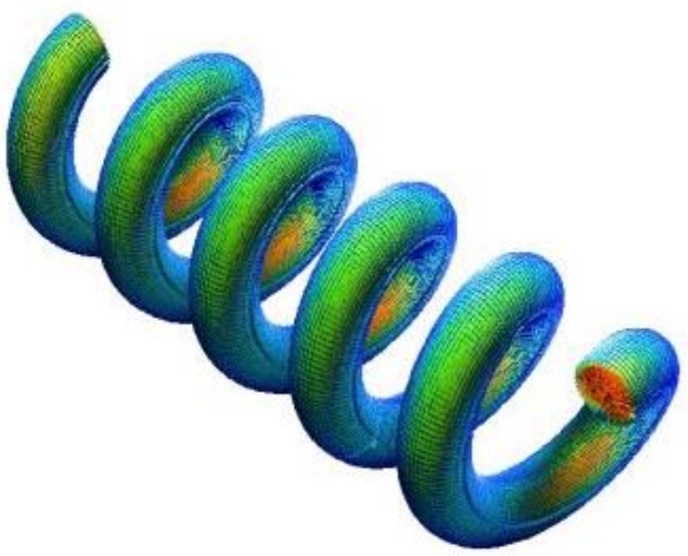
BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS

FACULTY OF CIVIL ENGINEERING

NONLINEAR MECHANICS

LECTURE NOTES



Contents

1 Introduction to Nonlinear Phenomena

Let's get to the point very quickly.

Nature is nonlinear.

Nonlinearity is all around us. Nonlinear phenomena are phenomena, which, in contrast to a linear system, cannot be explained by a mathematical relationship of proportionality (that is, a linear relationship between two variables). Simply speaking, twice as much is not necessarily twice as good. The spread of an infectious disease is most often exponential, rather than linear, with time. The same way, the materials that surround us all exhibit either large deformations or inelastic material behavior. This document is of course entitled to the questions of solid mechanics, but this only means, that the terminology to be used is fashioned to fit a specific subdomain of natural phenomena and the most important concepts introduced throughout the document are applicable in a much broader context.

In solid mechanics, there are three major, and most common sources of nonlinear structural behavior:

(a) **Material nonlinearity** This type of nonlinearity is probably the one that you are most familiar with and refers to the ability for a material to exhibit a nonlinear stress-strain (constitutive) response. Elasto-plastic, crushing, and cracking are good examples, but this can also include time-dependent effects such as visco-elasticity or visco-plasticity (creep). Material nonlinearity is often characterized by a gradual weakening of the structural response as an increasing force is applied, due to some form of internal decomposition. Most metals have a fairly linear stress/strain relationship at low strain values; but at higher strains the material yields, at which point the response becomes nonlinear and irreversible. Also, material nonlinearity may be related to factors other than strain. Strain-rate-dependent material data and material failure are both forms of material nonlinearity, furthermore material properties can also be a function of temperature and other predefined fields.

(b) **Boundary nonlinearity** Boundary nonlinearity occurs if the boundary conditions change during the analysis. In highly flexible components and assemblies with multiple components, progressive displacement can cause self-contact or component-to-component contact. In such circumstances, the stiffness of the structure or assembly may change when two or more parts either contact or separate from initial contact – sometimes only a little, but sometimes quite a bit. A good example is the blowing a sheet of material into a mold. The sheet expands relatively easily under the applied pressure until it begins to contact the mold. From then on the pressure must be increased to continue forming the sheet because of the change in boundary conditions.

(c) **Geometric nonlinearity** The third source of nonlinearity is related to changes in the geometry of the structure during the analysis which must be considered in formulating the constitutive and equilibrium equations. Structures whose stiffness depends on the displacement they may undergo are geometrically nonlinear. This accounts for phenomena such as the stiffening of a loaded clamped plate, and buckling or 'snap-through' behavior in slender structures or components. Geometric nonlinearity is usually more difficult to characterize than purely material considerations. Geometric effects may be both sudden and unexpected, but without taking them into account any computational simulation may completely fail to predict the real structural behavior.

Many structures exhibit combinations of these three main sources of nonlinearity, and the algorithms which solve nonlinear equations are generally set up to handle nonlinear effects from a variety of sources.

To find evidence of possible nonlinear behavior, look for characteristics such as permanent deformations, and any gross changes in geometry. Cracks, necking, thinning, distortions in open section beams, buckling, stress values which exceed the elastic limits of the materials, evidence of local yielding, shear bands, and temperatures above 30 % of the melting temperature are all indications that nonlinear effects may play a significant role in understanding the structural behavior.

2 One-Dimensional Strain Measures

To understand the different ways in wich strains can be measured, we need to understand how motion can be described in the first place. To introduce the basic ideas, we consider first a one-dimensional, prismatic truss element, the geometry of which can be characterized by its length and cross-sectional area. Let's generalize the following notation for this section: a quantity referring to the undeformed state is denoted by capital, the same quantity in the deformed state is denoted by small letters. Accordingly, prior to the deformations the truss has length L and cross sectional area A, which are then changed to l and a (Figure 1).

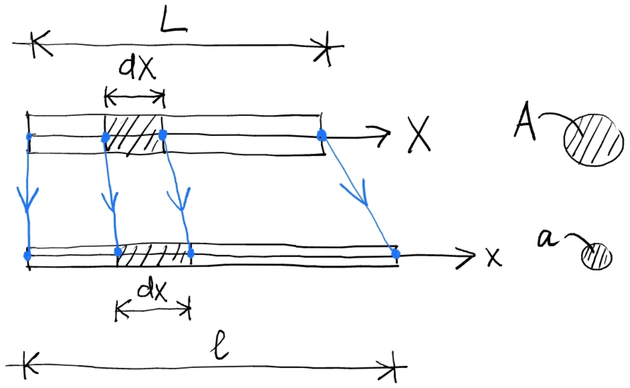
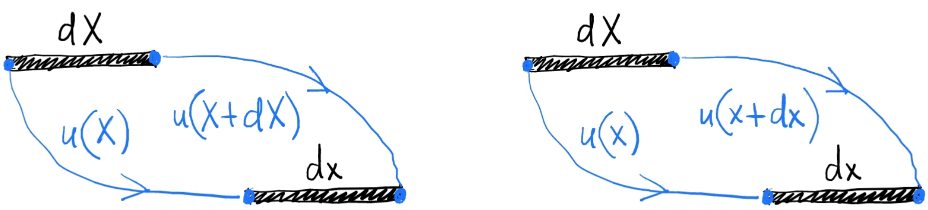


Figure 1. Deformation of a truss member.

The question to be answered is "How L and A transform to l and a", or more generally, "What is the connection between geometrical quantities referring to the initial and the deformed configurations?". It is intuitive to think that if we are able to describe the deformation of an infinitesimally small portion of the truss, we should be able to extend that knowledge to the whole domain.



(a) Lagrangian description. (b) Eulerian description.

Figure 2. Lagrangian and Eulerian specification of the same motion.

On the left and right of Figure 2 we can see two equally valid approaches to the specification of the longitudinal displacement of a typical infinitesimal section, originally occupying the portion of the axis of the truss between X and $X + dX$, which is then changed to x and $x + dx$. The two approaches simply mean that we pose different questions to observe the same phenomenon:

How much has the point moved being at X prior to deformation?
or
How much has the point moved being at x after deformation?

The first approach is to formulate the unknown function as $u(X)$, as a function of the initial or material coordinates and posing the question '$u(X) =?$', while in the second case the unknown function is formulated as a function of the current or spatial coordinates and we pose the question '$u(x) =?$'. Material and spatial coordinates are also called *Lagrangian* and *Eulerian* coordinates, hence the terminology Lagrangian or Eulerian description is often used in the literature. The two approaches both have their benefits and are both suitable to the description of certain kinds of motions. In the field of solid mechanics we most often use the Lagrangian description, while for the description of motion of fluids an Eulerian specification of the flow field is much more suitable. To further illustrate the differences, think about the following questions and try to figure out which description fits the case:

'How much does a truss deform due to a given load?' versus 'What amount of load deforms the truss into a given length?'
or
'How far can I get with 100 $ in my pocket?' versus 'How much money do I need to get to a given destination?'

We don't want to suggest the false image that every problem can only be solved by one or the other approach, but clearly, there are situations, where one of the two is more adviseable to use. Description of the change of cross-sectional area requires exactly the same considerations, and the motion of a linear section is illustrated on Figure 3 using a Lagrangian specification of the radial displacement function v.

2.1 Lagrangian Strain Measures

If we decide to use a Lagrangian description, we formulate the unknowns as functions of material coordinates, therefore:

$$x(X) = X + u(X), \qquad x(X) + dx = X + dX + u(X + dX), \tag{1}$$

$$r(R) = R + v(R), \qquad r(R) + dr = R + dR + v(R + dR), \tag{2}$$

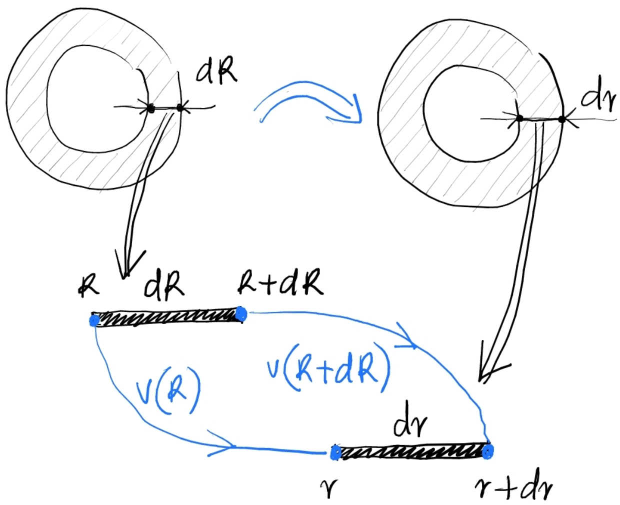


Figure 3. Radial deformation of a truss using Lagrangian specification.

where $u(X)$ and $v(R)$ are unknown functions of material coordinates and can be approximated using a linear model:

$$u(X + dX) \cong u(X) + \left.\frac{\partial u(X)}{\partial X}\right|_X dX, \tag{3}$$

$$v(R + dR) \cong v(R) + \left.\frac{\partial v(R)}{\partial R}\right|_R dR. \tag{4}$$

Combining Equations (1) with Equation (3) and Equations (2) with Equation (4) gives

$$\begin{aligned} dx &= (x(X) + dx) - x(X) && (5)\\ &= dX + \left.\frac{\partial u(X)}{\partial X}\right|_X dX && (6)\\ &= dX\left(1 + \left.\frac{\partial u(X)}{\partial X}\right|_X\right) && (7)\\ &= \lambda(X)\, dX, && (8) \end{aligned}$$

and

$$\begin{aligned} dr &= (r(R) + dr) - r(R) && (9)\\ &= dR + \left.\frac{\partial v(R)}{\partial R}\right|_R dR && (10)\\ &= dR\left(1 + \left.\frac{\partial v(R)}{\partial R}\right|_R\right) && (11)\\ &= \mu(R)\, dR, && (12) \end{aligned}$$

where $\lambda(X)$ and $\mu(R)$ are called *stretch* and *contraction*, and are both fundamental quantities, extremely useful for the classification of different possible strain measures. Let's assume now a homogeneous (or affine) deformation process, that is $\lambda(X) = \lambda$ and $\mu(R) = \mu$ and integrate both sides of Equations (8) and (12) we get

$$\int_l dx = \lambda \int_L dX, \tag{13}$$

$$l = \lambda\, L = (1 + \varepsilon_E)\, L, \tag{14}$$

and

$$\int_r dr = \mu \int_R dR, \tag{15}$$

$$r = \mu\, R = (1 + \eta_E)\, R, \tag{16}$$

where ε_E and η_E are the engineering definitions of axial and radial strain, which, after sorting Equations (14) and (16) are:

$$\varepsilon_E = \frac{l - L}{L} \quad \text{and} \quad \eta_E = \frac{r - R}{R}. \tag{17}$$

The strain definitions of Equations (17) can be expressed with the homogeneous form of stretch and contraction by sorting equations (14) and (16):

$$\varepsilon_E = \lambda - 1 \quad \text{and} \quad \eta_E = \mu - 1, \tag{18}$$

where

$$\lambda = \frac{l}{L} \quad \text{and} \quad \mu = \frac{r}{R}. \tag{19}$$

We note that in many textbooks, engineering strain is introduced as a definition, whereas it is clear from the above derivations, that these quantities can be deduced with proper mathematical reasoning. In our humble opinion, it is always preferable to introduce an idea as something to be understood, rather than a concept to be memorized.

Now the reader might expect, that nonlinear strains are gonna be introduced by involving higher order terms in the approximations of $u(X)$ and $v(R)$ by Equations (3) and (4). Of course, the appereance of the small strain definition is a direct consequence of the linear approximation in Equation (**??**). Now let's see what happens when using a quadratic model as:

$$u(X + dX) \cong u(X) + \left.\frac{\partial u(X)}{\partial X}\right|_X dX + \frac{1}{2}\left.\frac{\partial^2 u(X)}{\partial X^2}\right|_X dX^2. \tag{20}$$

3 One-Dimensional Stress and Strain Measures

When external forces are applied to objects made of elastic materials, they produce chenges in shape and size of the object. The distinction between changes of shape and size is no coincidence, it is important to note, that an object can change size without changing shape and vice versa. The questions of this chapter to be answered are:

1. How to quantify stress and strain?
2. Given a definition of stress and strain, how to find the associate material parameters to accurately predict mechanical work?

Strain is a measure of deformation expressing the relative change in shape or size of an object due to externally-applied forces.

Stress is the measure of internal force (per unit area, etc.) associated with a particular strain definition.

As an introduction to the different ways in which large strains can be measured, we consider first a one-dimensional truss element of undeformed length L and cross-sectional area A. After applying the external loads, the length and the cross-sectional area changes to l and a. Similarly in this chapter, denote the quantities referring to the initial configuration with capital letters, and use small letters to describe quantities referring to the actual state of the body.

Possibly the simplest, naturally emerging quantity that we can use to measure strain in the bar is the so-called *engineering strain* ε_E defined as

$$\varepsilon_E = \frac{l - L}{L}. \tag{21}$$

Although the expression is simple enough to seem like a result of pure intuition, it can be easily derived by studying an infinitesimal section of the bar of length dX, shown on Figure .

Although there are a wide variety of strain definitions to be used, they must obey some requirements. The above definition of strain directly implies that it is dimensionless and has no units, and it is just logical to expect that in a small elongation scenario any of the applied strain definitions must yield the same value, being the value of the linear strains.

Stretch

The fraction of a deformed and the corresponding undeformed geometrical quantity is called stretch and is denoted by:

$$\lambda = \frac{l}{L} \tag{22}$$

for the longitudinal deformation, and

$$\mu = \sqrt{\frac{a}{A}} \tag{23}$$

for the deformation of the cross section. If we have a rod with a circular cross section of diameters D and d in the undeformed and deformed configurations, the latter expression equals to:

$$\mu = \frac{d}{D} \tag{24}$$

It is informative to explore the limits of a deformation measure, if it has any. One possible limiting case could be the one, in which the cross-sectional area vanishes, since that comes with an infinitely large normal stress value. The longitudinal and cross sectional deformations are linked by the Poisson ratio, thus:

$$-\nu\lambda = \mu \tag{25}$$

$$-\nu\frac{l}{L} = \frac{d}{D} \tag{26}$$

$$d = -l\nu\frac{D}{L} \tag{27}$$

From the last expression it follows, that the diamater of the deformed rod is zero if either the Poisson ratio, or the rod's deformed length are zero. These two conditions are limiting cases of the deformation measure at hand.

Now lets find the energy pair of the deformation measure. The concept is that the stored elastic energy withing the domain of the structure should equal the work done by the external force F due to an infinitesimal displacement of the point of application. In the present case the longitudinal displacement u can be expressed as:

$$u = l - L, \tag{28}$$

therefore

$$du = dl, \tag{29}$$

whilst a differential increment of the deformation measure is:

$$d\lambda = \frac{dl}{L}. \tag{30}$$

From these expressions, the work of the external force can be expressed as:

$$\begin{aligned} F du &= F dl && (31)\\ &= F\, d\lambda\, L && (32) \end{aligned}$$

To obtain an expression in the form of $F\ du = volume * stress * deformation$, the following can be done:

$$Fdu = F\ d\lambda\ L = \frac{A}{A} F\ d\lambda\ L = \frac{F}{A}\ V\ d\lambda = V\ \sigma_n\ d\lambda \tag{33}$$

It turns out, that the resulting energy pair of stretch is the so-called nominal stress $\sigma_n = \frac{F}{A}$.

Engineering strain

The definition of engineering strain components for the problem at hand and their relationship are:

$$\varepsilon_l = \frac{l - L}{L} = \lambda - 1 \tag{34}$$

$$\varepsilon_d = \frac{d - D}{D} = \mu - 1 \tag{35}$$

$$\varepsilon_d = -\nu\, \varepsilon_l \tag{36}$$

Looking for the possible conditions of a vanishing cross-section:

$$-\nu(\lambda - 1) = \mu - 1 \tag{37}$$

$$-\nu(\lambda - 1) = \frac{d}{D} - 1 \tag{38}$$

$$d = D[1 - \nu(\lambda - 1)] \tag{39}$$

From the condition $d > 0$:

$$\lambda < \frac{1 + \nu}{\nu} \qquad \text{that is} \qquad l < L\frac{1 + \nu}{\nu}, \tag{40}$$

which is a limiting case for the linear strain.
Looking for the energy pair ($d\varepsilon_l = d\lambda$):

$$F du = F\, d\varepsilon_l\, L = \frac{A}{A} F\, d\varepsilon_l\, L = \frac{F}{A}\, V\, d\lambda = V\, \sigma_n\, d\varepsilon_l \tag{41}$$

or

$$F du = F\, d\varepsilon_l\, L = \frac{a}{a} F\, d\varepsilon_l\, L = \frac{F}{a} a\, l\, \frac{L}{l}\, d\varepsilon_l = \sigma\, \lambda^{-1}\, v\, d\varepsilon_l \tag{42}$$

$$= \sigma\, \lambda^{-1}\, J\, V\, d\varepsilon_l = \sigma_{p1}\, V\, d\varepsilon_l, \tag{43}$$

where

σ_{p1}	is the 1st Piola-Kirchhoff stress;
$J = \frac{v}{V}$	is the Jacobian of deformation;
$\sigma = \frac{F}{a}$	is the true stress;

The energy pair of the linear strain is the 1st Piola-Kirchhoff stress, which in the present case equals the nominal stress:

$$\sigma_{p1} = \sigma\, \lambda^{-1}\, J = \sigma_n\, \mu^{-2}\, \lambda^{-1}\, J = \sigma_n \tag{44}$$

Green - Lagrange strain

The definition of strain components for the problem at hand and their relationship are:

$$\varepsilon_l = \frac{l^2 - L^2}{2L^2} = \frac{1}{2}(\lambda^2 - 1) \tag{45}$$

$$\varepsilon_d = \frac{1}{2}(\mu^2 - 1) \tag{46}$$

$$\varepsilon_d = -\nu\,\varepsilon_l \tag{47}$$

Looking for the possible conditions of a vanishing cross-section:

$$-\nu\frac{1}{2}(\lambda^2 - 1) = \frac{1}{2}(\mu^2 - 1) \tag{48}$$

$$d = D\sqrt{1 - \nu(\lambda^2 - 1)} \tag{49}$$

To be sure that the argument of the square root is greater than zero, the following must be true:

$$1 - \nu(\lambda^2 - 1) > 0 \qquad \text{that is} \qquad l < L\,\sqrt{\frac{1+\nu}{\nu}} \tag{50}$$

Looking for the energy pair ($d\varepsilon_l = \lambda\, d\lambda = \lambda\frac{dl}{L}$):

$$F du = F\, d\varepsilon_l\,\frac{L}{\lambda} = \frac{A}{A}F\, d\varepsilon_l\,\frac{L}{\lambda} = \sigma_n\,\lambda^{-1}\,V d\varepsilon_l = \sigma_{p2}\,V\,d\varepsilon_l \tag{51}$$

The energy pair of the Green - Lagrange strain is the 2nd Piola-Kirchhoff stress $\sigma_{p2} = \sigma_n\,\lambda^{-1}$.

Logarithmic strain

The definition of strain components for the problem at hand and their relationship are:

$$\varepsilon_l = ln(\lambda) \tag{52}$$

$$\varepsilon_d = ln(\mu) \tag{53}$$

$$\varepsilon_d = -\nu\,\varepsilon_l \tag{54}$$

Looking for the possible conditions of a vanishing cross-section:

$$-\nu ln(\lambda) = ln(\mu) \tag{55}$$

$$d = D\lambda^{-\nu} \tag{56}$$

The condition $d > 0$ holds for any choice of the geometrical quantities, there is no limiting case for this deformation measure.
Looking for the energy pair ($d\varepsilon_l = \frac{d\lambda}{\lambda} = \frac{dl}{\lambda L}$):

$$F du = F\ d\varepsilon_l\ \lambda\ L = \frac{A}{A} F\ d\varepsilon_l\ \lambda\ L = \sigma_n\ \lambda\ V d\varepsilon_l = \sigma_K\ V\ d\varepsilon_l \tag{57}$$

The energy pair of the Green - Lagrange strain is the Kirchhoff stress $\sigma_K = \sigma_n\ \lambda$.

4 Use of Variational Methods

In the classical sense, a variational principle has to do with the finding the extremum or stationary values of of a functional with respect to the variables of the problem. In solid and structural mechanics the functional represents the total energy of the system, and in other problems, it is simply an integral representation of the governing equations. Many problems of mechanics are posed in terms of finding the extremum and thus, by their nature, can be formulated in terms of variational statements. However, there are problems that can be formulated by other means, such as conservation laws, but these can also be formulated by means of variational principles. For example, to answer this question: "What is the shape of a chain suspended at both ends?" we can use the variational principle that the shape must minimize the gravitational potential energy.

The classical use of the phrase "variational formulations" refers to the construction of a functional or a variational principle that is equivalent to the governing equations of a problem. The modern use of the phrase refers to the formulation in which the governing equations are translated into weighted-integral statements that are not necessary equivalent to a variational principle. Even those problems that do not admit variational principles in the classical sense can now be formulated using weighted-integral statements. The importance of variational formulations of physical laws, goes far beyond its use as simply an alternate to other formulations. Variational formulations form a powerful basis for obtaining approximate solutions to practical problems, many of which are intractable otherwise. Variational formulations can also serve to unify diverse fields, suggest new theories, and provide a powerful means to study the existence and uniqueness of solutions to problems. While all sufficiently smooth fields lead to meaningful variational forms, the converse is not true: There exist physical phenomena which can be adequately modeled mathematically only in a variational setting; the are nonsensical when viewed locally.

5 Approximate Solution of Boundary Value Problems with Variational Methods

Let say that we have the following differential equation

$$\frac{d}{dx}(a(x)\frac{du}{dx}) + f(x) = 0 \tag{58}$$

defined over a one-dimensional domain Ω, and boundary conditions in the form of

$$Bu = \hat{u} \tag{59}$$

where

$a(x), f(x)$	are known functions;
$u(x)$	is the unknown function;
B	is an operator;
$\hat{u}$	is the prescribed value of the unknown function on the boundary of Ω.

We seek the approximate solution over the entire domain of the problem in the following form:

$$u(x) \approx U_N(x, c_j) = \sum_{j=1}^{N} c_j\varphi_j(x) + \varphi_0(x) \tag{60}$$

where the c_j are coefficients to be determined and $\varphi_j(x)$ and $\varphi_0(x)$ are approximation functions chosen such that the specified boundary conditions of the problem are satisfied by the N-parameter approximate solution $U_N(x)$. In eq. (60) the part containing the unknowns ($\sum c_j\varphi_j$) is termed the homogeneous part and the other is the nonhomogeneous part (φ_0) that has the sole purpose of satisfying the specifiend boundary conditions of the problem. Since φ_0 satisfies the boundary conditions, the sum $\sum c_j\varphi_j$ must satisfy, for arbitrary c_j, the homogeneous form of the boundary conditions. In addition it is required, that the approximation funcitons be such that U_N is continously differentiable as many times as called for in the original differential equation.

Since U_N is only an approximation of u, the equality in eq. (58) will not generally hold, and the difference is called the *residual*, denoted by R:

$$R(x, c_j) = \frac{d}{dx}(a(x)\frac{dU_N(x, c_j)}{dx}) + f(x) \neq 0 \tag{61}$$

When looking for an approximate solution, our goal is to obtain N linearly independent equations among c_j, such that they make the residual to vanish, in a global sense. For that we apply the fundamental lemma of variational calculus and multiply the residual with an N number of linearly independent weight functions $w_i(x)$, then integrate the resulting experssions over the problem domain. The ith equation takes the folowing form:

$$\int_\Omega w_i(x)R(x, c_j)dx = 0 \tag{62}$$

This is called the *weighted-residual statement* of the problem. The difference between eqs. (58) and (62) is that they both require the residual to vanish, but (58) is a local, or point-to-point, while (62) is a global, or integral requirement. Methods that use the weighted-residual form of the problem differ in the selection of the weight functions w_i. We have the following variational methods as special cases:

Petrov-Galerkin method : $w_i = \psi_i \neq \varphi_i$

Galerkin's method : $w_i = \varphi_i$

Least squares method : $w_i = \frac{d}{dx}(a(x)\frac{d\varphi_i}{dx})$

Collocation method : $w_i = \delta(x - x_i)$

Here x_i is the ith collocation point of the domain of the problem and $\delta(\cdot)$ is the Dirac delta function defined such that its value is zero for all nonzero values of its arguments:

$$\delta(x - x_0) = 0 \quad \text{when} \quad x \neq x_0, \qquad \int_{-\infty}^{\infty} f(x)\delta(x - x_0)dx = f(x_0) \tag{63}$$

If we plan to use the approximation functions φ_i for w_i, it makes sense to shift half of the derivatives from u to w so that both are differentiated equally, and finally we end up having weaker continuity requirements on φ_j. This can be done by partial integration of the weighted residual statement:

$$\begin{aligned}\int_\Omega w_i(x)R(x, c_j)dx = &\int_\Omega w_i(x)\frac{d}{dx}(a(x)\frac{dU_N(x, c_j)}{dx})dx + \int_\Omega w_i(x)f(x)dx \\ =&[w_i(x)a(x)\frac{dU_N(x, c_j)}{dx}]_\Gamma - \int_\Omega \frac{dw_i(x)}{dx}a(x)\frac{dU_N(x, c_j)}{dx}dx \\ &+ \int_\Omega w_i(x)f(x)dx = 0 \qquad \text{for} \quad i = 1 \dots N\end{aligned}$$

This is called the variational, or *weak-form* of the problem, and the variational method is referred to as the *Ritz-method.* An important characteristic of the weak form is, that it includes the natural boundary conditions of the problem, therefore the approximate solution U_N is required to satisfy only the essential boundary conditions. The weak-form equivalent of eq. (62) for the Ritz-method is:

$$\int_\Omega \frac{dw_i(x)}{dx} a(x) \frac{dU_N(x, c_j)}{dx} dx - [w_i(x) a(x) \frac{dU_N(x, c_j)}{dx}]_\Gamma = \int_\Omega w_i(x) f(x) dx \quad (64)$$

Either we use a weighted-residual method and eq. (62) or the Ritz-method with eq. (64) to construct the necessary number of independent equations among the unknown parameters c_j, after carrying out the integrations, a system of linear equations can be formed:

$$\mathbf{Kc} = \mathbf{f} \quad (65)$$

where

c	is the vector of unknowns;
K	is the matrix containing the coefficients of the unknowns after integration;
f	is the vector that contains the terms that are not multiplied by any of the unknown coefficients.

Due to the different choices of construction, the system of algebraic equations will have different characteristics. For linear differential equations of any order, only the least-squares method yields a system of matrix equations whose coefficient matrix is symmetric. One other method that has the symmetry property is the Ritz-method, which uses the weak form of self-adjoint differential equations. We note here, that since any physical law which can be expressed as a variational principle describes a self-adjoint operator, the Ritz-method has special significance in the field of solid mechanics.

6 Examples

We illustrate the steps of the previous section on a simple example of a rod of length L, constant cross-sectional modulus EA, subjected to concentrated and distributed loads F and $p(x)$ (see Figure 4). The differential equation and the

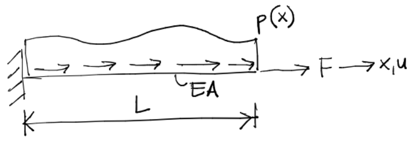


Figure 4. Rod subject to a distributed and a contentrated load.

boundary conditions of the problem:

$$EA\frac{d^2u(x)}{dx^2} + p(x) = 0 \quad \text{on} \quad \Omega = (0, L) \tag{66}$$

$$u(0) = 0, \quad N(L) = EA\frac{du(x)}{dx}\bigg|_{x=L} = F \tag{67}$$

In the following examples we wish to approximate the unknown function $u(x)$ by

$$u(x) \approx U_2(x) = \varphi_0(x) + c_1\varphi_1(x) + c_2\varphi_2(x) \tag{68}$$

therefore we need two linearly independent equations that relate c_1 and c_2.

Weighted-residual methods The approximate solution U_N has to be twice continuously differentiable and satisfy both the essential and natural boundary conditions. The approximation function φ_0 has to satisfy the non-homogeneous boundary conditions, that is

$$EA\frac{d\varphi_0(x)}{dx}\bigg|_{x=L} = F \tag{69}$$

This can be guaranteed by

$$\varphi_0(x) = \frac{F}{EA}x \tag{70}$$

The functions φ_1 and φ_2 must satisfy the homogeneous boundary conditions, that is

$$\varphi_1(0) = \varphi_2(0) = \frac{d\varphi_1(x)}{dx}\bigg|_{x=L} = \frac{d\varphi_2(x)}{dx}\bigg|_{x=L} = 0 \tag{71}$$

The requirements are satisfied by the following functions:

$$\varphi_1(x) = x^2 - 2Lx, \quad \varphi_2(x) = x^3 - 3L^2x \tag{72}$$

The final approximation with its first and second derivatives:

$$U_2(x, c_1, c_2) = \frac{F}{EA}x + c_1(x^2 - 2Lx) + c_2(x^3 - 3L^2x) \tag{73}$$

$$\frac{dU_2(x, c_1, c_2)}{dx} = \frac{F}{EA} + c_1(2x - 2L) + c_2(3x^2 - 3L^2) \tag{74}$$

$$\frac{d^2U_2(x, c_1, c_2)}{dx^2} = 2c_1 + 6xc_2 \tag{75}$$

The residual function:

$$R(x, c_1, c_2) = EA\frac{d^2U_2(x, c_1, c_2)}{dx^2} + p(x) = EA(2c_1 + 6xc_2) + p(x) \tag{76}$$

The weighted-residual statements of eq. (62) for the weight functions w_1 and w_2 relating the unknown parameters c_1 and c_2:

$$EA\int_0^L w_1(x)(2c_1 + 6xc_2)dx = -\int_0^L w_1(x)p(x)dx \tag{77}$$

$$EA\int_0^L w_2(x)(2c_1 + 6xc_2)dx = -\int_0^L w_2(x)p(x)dx \tag{78}$$

Ritz-method The Ritz-method uses the weak-form of eq. (64). Utilizing the weaker continuity requirements, when solving the problem at hand with the Ritz-method, we use the following approximation functions:

$$\varphi_0(x) = 0, \quad \varphi_1(x) = x, \quad \varphi_2(x) = x^2. \tag{79}$$

The final approximation and its first derivative:

$$U_2(x, c_1, c_2) = c_1x + c_2x^2, \quad \frac{dU_2(x, c_1, c_2)}{dx} = c_1 + 2c_2x \tag{80}$$

The approximation functions φ_i vanish at $x = 0$ as they satisfy the specified essential boundary conditions, and for the problem at hand expression (64) simplifies to the following:

$$EA\int_0^L \frac{dw_i(x)}{dx}\frac{dU_2(x, c_j)}{dx}dx = \int_0^L w_i(x)p(x)dx + w_i(L)F \tag{81}$$

The two independent equations relating c_1 and c_2 for the Ritz-method:

$$EA\int_0^L \frac{dw_1(x)}{dx}(c_1 + 2c_2x)dx = \int_0^L w_1(x)p(x)dx + w_1(L)F \tag{82}$$

$$EA\int_0^L \frac{dw_2(x)}{dx}(c_1 + 2c_2x)dx = \int_0^L w_2(x)p(x)dx + w_2(L)F \tag{83}$$

With the two-parameter approximate solution U_2 of (68), the structure of the algebraic equation system (65) is the following:

$$\begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix} \tag{84}$$

6.1 Example 1

Consider a bar of length $L = 10.0$, cross section modulus $EA = 5.0$, distributed force $p(x) = 0.1$ and concentrated force $F = 1.0$, illustrated on Figure 5. The problem is solved with the methods mentioned above, and with the Finite Element Method for verification. The components of the equation system (84) and the displacement at $x = L$ are listed in Table 1 with four significant digits.

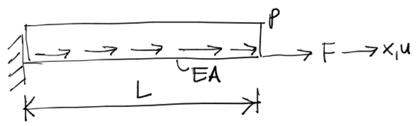


Figure 5. Figure of Example 1.

6.2 Example 2

Consider a bar of length $L = 10.0$, cross section modulus $EA = 5.0$, a linearly distributed force with maximum intensity $pmax = 0.2$ and concentrated force $F = 1.0$, illustrated on Figure 6. The problem is solved with the methods mentioned above, and with the Finite Element Method for verification. The components of the equation system (84) and the displacement at $x = L$ are listed in Table 2 with four significant digits.

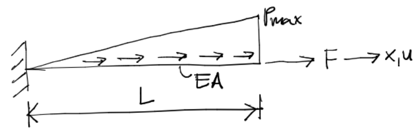


Figure 6. Figure of Example 2.

	w_1	w_2	K_{11}	K_{12}	K_{21}	K_{22}	f_1	f_2	c_1	c_2	$u(L)$
Petrov-Galerkin	1	x	100.0	1500	500.0	10e3	-1.000	-5.000	-0.01000	0	3.000
Galerkin	φ_1	φ_2	-6667	-125e3	-125e3	-2.4e6	66.67	1250	-0.01000	0	3.000
Least-squares	$EA\frac{d^2\varphi_1}{dx^2}$	$EA\frac{d^2\varphi_2}{dx^2}$	1000	15e3	15e3	30e4	-10.00	-150.0	-0.01000	0	3.000
Collocation	$\delta(x-0)$	$\delta(x-L)$	10.00	0	10.00	300.0	-0.1000	-0.1000	-0.01000	0	3.000
Ritz	φ_1	φ_2	50.00	500.0	500.0	6667	15.00	133.3	0.4000	-0.01000	3.000
FEM	-										3.000

Table 1. Results of Example 1

	w_1	w_2	K_{11}	K_{12}	K_{21}	K_{22}	f_1	f_2	c_1	c_2	$u(L)$
Petrov-Galerkin	1	x	100.0	1500	500.0	10e3	-1.000	-6.667	0	-0.0006667	3.333
Galerkin	φ_1	φ_2	-6667	-125e3	-125e3	-2.4e6	83.33	1600	0	-0.0006667	3.333
Least-squares	$EA\dfrac{d^2\varphi_1}{dx^2}$	$EA\dfrac{d^2\varphi_2}{dx^2}$	1000	15e3	15e3	30e4	-10.00	-200.0	0	-0.0006667	3.333
Collocation	$\delta(x-0)$	$\delta(x-L)$	10.00	0	10.00	300.0	0	-0.2000	0	-0.0006667	3.333
Ritz	φ_1	φ_2	50.00	500.0	500.0	6667	16.67	150.0	0.4333	-0.01000	3.333
FEM	-										3.333

Table 2. Results of Example 2

7 Solution of Nonlinear Equations

From all the possible cases that may lead to having a nonlinear equation, the two most likely ones are

- The obvious case of being face-to-face with a nonlinear equation,
- The problem of minimizing or maximizing a nonlinear function or functional (more about functionals later).

Of course this means, that -with the necessary mathematical manipulations,- the problem of finding stationary points of functions leads to the solution of nonlinear algebraic equations. * In mechanical problems, we can either directly formulate the nonlinear equations that govern the behaviour of the structure, or find a function or functional and a corresponding variational pronciple, that represents the energy balance of it. Needless to say, **both formulations must lead to the same unique solution**.

Let's illustrate these ideas with the example of a simple spring system depicted on Figure 7, and the assumption that the force S necessary to deform a spring from a state of having the initial length L, to the state of having the current length l is proportional to the change of length, with the factor of proportionality k being called the stiffness of the spring, thus:

$$S(l) = (l - L)k. \tag{85}$$

Our goal is to determine the unknown positions x_1 and x_2 of the mass points, given

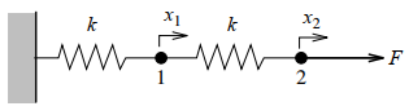


Figure 7. Two-degrees-of-freedom linear spring structure

the value of the external load F. For the discrete case we have, the principle of minimum potential energy (which is a proper variational principle) states that the energy is at a stationary position, if any infinitesimal change of position involves no change in energy. The potential energy function (TPE) for this case using equation (85) is

$$\Pi(\boldsymbol{x}) = \frac{1}{2}kx_1^2 + \frac{1}{2}k(x_2 - x_1)^2 - Fx_2, \tag{86}$$

*For the sake of completeness we note that analogously, finding a stationary path of a functional leads to having to solve partial differential equations in the general case.

where $\boldsymbol{x} = (x_1, x_2)^T$ and x_1 and x_2 are displacements of the joints 1 and 2. Believe it or not†, the previos TPE formulation is equal to:

$$\Pi(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{K}\boldsymbol{x} - \boldsymbol{x}^T\boldsymbol{f}, \tag{87}$$

with

$$\boldsymbol{K} = \begin{bmatrix} 2k & -k \\ -k & k \end{bmatrix}, \quad \boldsymbol{f} = \begin{pmatrix} 0 \\ F \end{pmatrix}. \tag{88}$$

Now consider the TPE due to a change in displacements given by the increment vector $\boldsymbol{u} = (u_1, u_2)^T$ as

$$\Pi(\boldsymbol{x}+\boldsymbol{u}) = (\boldsymbol{x}+\boldsymbol{u})^T\boldsymbol{K}(\boldsymbol{x}+\boldsymbol{u}) - (\boldsymbol{x}+\boldsymbol{u})^T\boldsymbol{f}, \tag{89}$$

$$= \frac{1}{2}k(x_1+u_1)^2 + \frac{1}{2}k(x_2+u_2-x_1-u_1)^2 - F(x_2+u_2). \tag{90}$$

According to our variational principle, the function $\Pi(\boldsymbol{x})$ is stationary at $\boldsymbol{x}^*$, if and only if ‡

$$\Pi(\boldsymbol{x}^*) = \Pi(\boldsymbol{x}^*+\boldsymbol{u}) \qquad \text{for any } \boldsymbol{u} \text{ possible.} \tag{91}$$

What is a possible choice for $\boldsymbol{u}$? Anything that is geometrically feasible, which is governed by the so-called geometric equations. Since a vector is known to have a direction and a magnitude §, it can be scaled to change magnitude, and rotated to change direction. Let's reformulate our previous statement like this:

The function $\Pi(\boldsymbol{x})$ is stationary at $\boldsymbol{x}^*$, if and only if for any infinitesimal scalar α and arbitrary direction $\boldsymbol{u}$

$$\Pi(\boldsymbol{x}^*) = \Pi(\boldsymbol{x}^*+\alpha\boldsymbol{u}) \qquad \textit{for any } \alpha \in \mathbb{R}, \boldsymbol{u} \in \mathbb{R}^2. \tag{92}$$

This indeed implies that $\boldsymbol{x}^*$ must be such, that it renders $\Pi(\boldsymbol{x}^*)$ to be stationary, at least in the very small neighbourhood of $\boldsymbol{x}^*$. From the *principle of minimum energy* , which is essentially a restatement of the second law of thermodynamics (not to be confused with minimum total potential energy principle), we also know that for a closed system, with constant external parameters and entropy, the internal energy will decrease and approach a minimum value at equilibrium. In other words we know in advance that our stationary point needs to be a minimum point and our problem can be stated as

$$\text{to find } \boldsymbol{x}^* \text{ which minimises } \Pi(\boldsymbol{x}). \tag{93}$$

†If not, well please take a pen, a paper, and go for it. This applies for the rest of the document.

‡The asterisk is used to distinguish between the solution $\boldsymbol{x}^*$ the possible candidate $\boldsymbol{x}$.

§Don't take this as a precise definition of a vector.

Note here for later reference, that the domain from where the candidate $\boldsymbol{x}$ can be selected is almost always bounded by constraint functions (representing boundary conditions, etc.) and we say to have a *constrained optimization problem*, we only transformed it into an *unconstrained optimization problem* by hard-coding the selection of applied forces and boundary conditions.

The obvius way of approaching the nonlinear governing equations directly is to write the equailibrium equations of the two mass points:

$$R_1(\boldsymbol{x}) = k(x_2 - x_1) - kx_1 = 0 \tag{94}$$

$$R_2(\boldsymbol{x}) = F - k(x_2 - x_1) = 0 \tag{95}$$

whic is in a more compact form

$$\boldsymbol{R}(\boldsymbol{x}) = \boldsymbol{0}, \quad \boldsymbol{R}(\boldsymbol{x}) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad \boldsymbol{x} \in \mathbb{R}^2. \tag{96}$$

$\boldsymbol{R}(\boldsymbol{x})$ is the so-called *residual* or *out-of-balance* vector, and a solution for $\boldsymbol{x}$ is achieved when $\boldsymbol{R}(\boldsymbol{x}) = \boldsymbol{0}$. For now it may be apparent for the reader, that the expressions $R_1(\boldsymbol{x})$ and $R_2(\boldsymbol{x})$ are actually linear in x_1 and x_2. This is simply because the spring force in equation (85) is a linear function of its variable, but it doesn't violate the generality of any the statements of the section. Before we move to solving our problem, we need to clarify a few basic mathematical ideas and notations.

7.1 Basic Concepts of Unconstrained Optimization

The idea of a *line* is important, and is the set of points

$$\boldsymbol{x}(= \boldsymbol{x}(\alpha)) = \boldsymbol{x}' + \alpha \boldsymbol{s} \tag{97}$$

for all α (sometimes for all $\alpha \geq 0$; this is strictly a half-line), in which $\boldsymbol{x}'$ is a fixed point along the line (corresponding to $\alpha = 0$), and $\boldsymbol{s}$ is the *direction* of the line.

The calculus of any *function* of many variables, $f(\boldsymbol{x})$ say, is clearly important too. In general it is assumed that the function $f(\boldsymbol{x})$ is smooth enough, so that the forgoing definitions hold. For a function $f(\boldsymbol{x})$ at any point $\boldsymbol{x}$ there is a *vector of first partial derivatives*, or *gradient vector*

$$\begin{pmatrix} \partial f / \partial x_1 \\ \partial f / \partial x_2 \\ \vdots \\ \partial f / \partial x_n \end{pmatrix}_{\boldsymbol{x}} = \nabla f(\boldsymbol{x}) \tag{98}$$

These expressions can be used to determine derivatives of $f(\boldsymbol{x})$ along any line $\boldsymbol{x}(\alpha)$. By the chain rule we have

$$\frac{df}{d\alpha} = \boldsymbol{s}^T \nabla f = \nabla f^T \boldsymbol{s}. \tag{99}$$

The quantity $df/d\alpha$ is called the *directional derivative* of f along direction $\boldsymbol{s}$. To put more emphasis on the meaning, we will adopt the general notation:

$$Df(\boldsymbol{x})[\boldsymbol{s}] = \left.\frac{df(\boldsymbol{x}(\alpha))}{d\alpha}\right|_{\alpha=0}, \tag{100}$$

which indicates directional derivative of f at $\boldsymbol{x}$ in the direction of an increment $\boldsymbol{s}$. Generally speaking, the directional derivative is the generalization of a derivative in that it provides the **change in an item due to a small change in something upon which the item depends on**.

The concepts introduced here are far more general than this chapter implies. For example, we can find the directional derivative of the determinant of matrix $\boldsymbol{A}$ in the direction of the change $\boldsymbol{U}$. Here the function $det(\boldsymbol{X})$ takes the role of the function $f(\boldsymbol{x})$, where $\boldsymbol{X}$ is an $N \times N$ square matrix. In this case, the line can be defined as the set of $N \times N$ matrices

$$\boldsymbol{X}(= \boldsymbol{X}(\alpha)) = \boldsymbol{A} + \alpha \boldsymbol{U} \tag{101}$$

for all α (sometimes for all $\alpha \geq 0$; this is strictly a half-line), in which $\boldsymbol{A}$ is a fixed point along the line (corresponding to $\alpha = 0$), and $\boldsymbol{U}$ is the *direction* of the line. Let's have $N = 2$, then the determinant of the matrix $\boldsymbol{A}$ and $\boldsymbol{A} + \alpha\boldsymbol{U}$ are

$$det(\boldsymbol{A}) = A_{11}A_{22} - A_{12}A_{21} \tag{102}$$

$$det(\boldsymbol{A} + \alpha\boldsymbol{U}) = (A_{11} + \alpha U_{11})(A_{22} + \alpha U_{22}) \tag{103}$$

$$- (A_{12} + \alpha U_{12})(A_{21} + \alpha U_{21}), \tag{104}$$

therefore

$$Ddet(\boldsymbol{A})[\boldsymbol{U}] = \frac{d}{d\alpha} det(\boldsymbol{X}(\alpha)) = \tag{105}$$

$$A_{22}U_{11} + A_{11}U_{22} - A_{21}U_{12} - A_{12}U_{21}. \tag{106}$$

Now being familiar with the meaning and notation of the directional derivative we can say, that the statement

The function $f(\boldsymbol{x})$ is stationary at $\boldsymbol{x}^*$, if and only if for any infinitesimal scalar α and arbitrary direction $\boldsymbol{s}$

$$f(\boldsymbol{x}^*) = f(\boldsymbol{x}^* + \alpha\boldsymbol{s}) \qquad \textit{for any } \alpha \in \mathbb{R}, \boldsymbol{s} \in \mathbb{R}^N. \tag{107}$$

is equivalent to

$$Df(\boldsymbol{x})[\boldsymbol{s}] = 0. \tag{108}$$

Now consider the solution of the set of nonlinear algebraic equations

$$\boldsymbol{R}(\boldsymbol{x}) = \boldsymbol{0}, \tag{109}$$

where, for a simple case with two equations and two unknowns

$$\boldsymbol{R}(\boldsymbol{x}) = \begin{pmatrix} R_1(x_1, x_2) \\ R_2(x_1, x_2) \end{pmatrix}; \quad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{110}$$

Typically, nonlinear equations of this type are solved using a Newton-Raphson iterative process whereby given a solution estimate $\boldsymbol{x}_k$ at iteration k, a new value $\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{\Delta}$ is obtained from an increment $\boldsymbol{\Delta}$ by establishing the linear approximation

$$\boldsymbol{R}(\boldsymbol{x} + \boldsymbol{\Delta}) \approx \boldsymbol{R}(\boldsymbol{x}) + D\boldsymbol{R}(\boldsymbol{x})[\boldsymbol{\Delta}] = \boldsymbol{0}. \tag{111}$$

The directional derivative is evaluated according to the chain rule to give

$$D\boldsymbol{R}(\boldsymbol{x})[\boldsymbol{\Delta}] = \boldsymbol{J}_R(\boldsymbol{x})\boldsymbol{\Delta}, \tag{112}$$

where $\boldsymbol{J}_R$ is an $N \times N$ *Jacobian matrix* over the function $\boldsymbol{R}(\boldsymbol{x})$. Due to the physical meaning, in solid mechanics the Jacobian matrix is denoted by $\boldsymbol{K}(\boldsymbol{x})$ and is called the *Stiffness matrix* of the structure with

$$\boldsymbol{K}(\boldsymbol{x}_k) = [K_{ij}(\boldsymbol{x}_k)]; \qquad K_{ij}(\boldsymbol{x}_k) = \left.\frac{\partial R_i}{\partial x_j}\right|_{(\boldsymbol{x}_k)}. \tag{113}$$

End of section 7.1.

Now we can finish our example. Evaluating the directional derivative of the TPE function and applying the variational proncipule gives

$$\begin{aligned} D\Pi(\boldsymbol{x})[\mathbf{u}] &= \left.\frac{d}{d\alpha}\right|_{\alpha=0} \Pi(\boldsymbol{x} + \alpha\boldsymbol{u}) && (114) \\ &= kx_1u_1 + k(x_2 - x_1)(u_2 - u_1) - Fu_2 && (115) \\ &= \boldsymbol{u}^T(\boldsymbol{K}\boldsymbol{x} - \boldsymbol{f}) = 0 && (116) \end{aligned}$$

where

$$\boldsymbol{K} = \begin{bmatrix} 2k & -k \\ -k & k \end{bmatrix}, \quad \boldsymbol{f} = \begin{pmatrix} 0 \\ F \end{pmatrix}. \tag{117}$$

The equality must hold for any $\boldsymbol{u}$, which is only guaranteed if

$$\boldsymbol{Kx} = \boldsymbol{f} \tag{118}$$

The other way around to solve the problem is to solve equations (94) and (95) by the application of the Newton-Raphson method, namely equation (111). It is apparent that the displacement vector $\boldsymbol{u} = (u_1, u_2)^T$ can play the role of the correction $\boldsymbol{\Delta}$ in equation (111), thus:

$$\boldsymbol{Ku} = -\boldsymbol{R} \tag{119}$$

Let's have another example of a truss member shown in Figure 8 with initial and loaded lengths, cross-sectional areas and volumes : L, A, V and l, a, v respectively. The residual is expressed by the equation for vertical equilibrium at the sliding

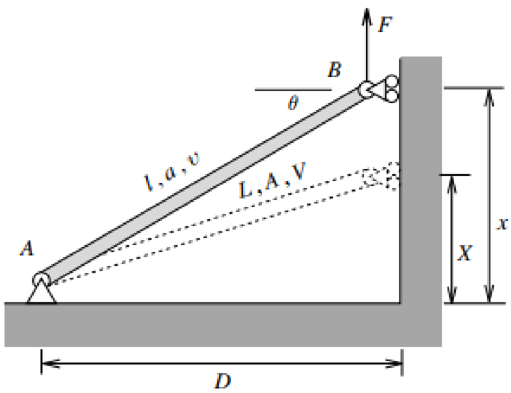


Figure 8. Single truss member.

joint B:

$$R(x) = T(x) - F = 0; \quad T = \sigma a sin\theta; \quad sin\theta = \frac{x}{l}; \tag{120}$$

where $T(x)$ is the vertical component, at B, of the internal force in the truss member and x gives the truss position. According to Equation (113), the stiffness is $K = dR/dx$. Since the load F is constant, the stiffness is the change in the vertical component T :

$$K = \frac{dR}{dx} = \frac{dT}{dx} = \frac{d}{dx}\left(\sigma a \frac{x}{l}\right) = \frac{d\sigma}{dx} a \frac{x}{l} + \sigma \frac{d}{dx}\left(a \frac{x}{l}\right) \tag{121}$$

To evaluate the last term in the above equation, here we assume that the material is incompressible and hence $V = v$ or $a = AL/l$, thus

$$\frac{dT}{dx} = \frac{d\sigma}{dx} AL \frac{x}{l^2} + \sigma AL \frac{d}{dx}\left(\frac{x}{l^2}\right). \tag{122}$$

Using the parameters of Figure 8, the length of the bar is $l = \left(D^2 + x^2\right)^{1/2}$, and consequently $dl/dx = x/l$ as you should verify. Using this formula and the basic rules of differentiation (chain rule, product rule, quotient rule, etc.):

$$\frac{dT}{dx} = \frac{d\sigma}{dl} AL \frac{x^2}{l^3} + \sigma AL \frac{l^2 - 2x^2}{l^4}. \tag{123}$$

To evaluate the term $d\sigma \backslash dl$, two constitutive equations are chosen based, without explanation at the moment, on Green's and a logarithmic definition of strain, hence the Cauchy, or true stress σ is either:

$$\sigma = E\frac{l^2 - L^2}{2L^2} \quad \text{or} \quad \sigma = Eln\frac{l}{L}; \tag{124}$$

where E is a (Young's modulus like) constitutive constant that, in ignorance, can be chosen to the same irrespective of the strain measure being used. With the above strain definitions:

$$\left(\frac{d\sigma}{dl}\right)_G = \frac{El}{L^2} \quad \text{and} \quad \left(\frac{d\sigma}{dl}\right)_L = \frac{E}{l} \tag{125}$$

and the stiffnesses are

$$K_G(x) = E\frac{A}{L}\left(\frac{x}{l}\right)^2 + \sigma AL\frac{l^2 - 2x^2}{l^4}; \tag{126}$$

$$K_L(x) = EAL\left(\frac{x}{l^2}\right)^2 + \sigma AL\frac{l^2 - 2x^2}{l^4}. \tag{127}$$

The subscripts G and L refer to the boundness of the obtained expressions to the choice of strain definition. Finally, the k-th iteration of the resulting Newton-Raphson iteration goes like:

$$\begin{aligned}
&R(x_k + u) = R(x_k) + K(x_k)u = 0,\\
&u = -R(x_k)/K(x_k),\\
&x_{k+1} = x_k + u.
\end{aligned}$$

A simple python code for solving the one-degree-of-freedom truss example using the logarithmic definition of strain is given below.

```python
import math

#input data
D = 2500.
X = 2500
A = 100.
E = 5e5
F = 1.5e7
```

```python

#calculated data
L = math.sqrt(X**2 + D**2)
V = A*L

def residual(x):
    T = s*a*x/l
    return T-F

def stiffness(x):
    K = E*A*L*(x/l/l)**2 + s*A*L*(l**2-2*x**2)/l**4
    return K

#initialize variables
s = 0.
l = L
a = A
x = X

cnorm = 1e-12 #numerical zero for the stop condition
miter = 200 #max. number of iterations
rnorm = cnorm*2 #initial value for the residual norm
niter = 0 #initial value for the number of iterations
while (rnorm > cnorm) and (niter<miter):
    niter = niter + 1

    #stiffness and residual
    K = stiffness(x)
    if abs(K) < 1e-20:
        print('Near zero stiffness --> STOP')
    R = residual(x)

    #geometry increment
    u = -R/K
    x = x + u
    l = math.sqrt(x**2 + D**2)
    a = V/l

    #stress
    s = E*math.log(l/L)

    #residual norm
    rnorm = abs(R)
    print('residual norm after iteration #{} : {}'.format(niter,
   rnorm))

print('solution : x = {}'.format(x))
```