INTRODUCTION TO MICROMECHANICS AND NANOMECHANICS
Lecture Notes (CE236/C214)

SHAOFAN LI
Department of Civil and Environmental Engineering,
University of California, Berkeley, CA94720, USA

University of California
Berkeley, California

Contents

Chapter 1

INTRODUCTION

What is micromechanics ? Generally speaking, micromechanics is a scientific discipline that studies: (1) mechanical, electrical, and, in general, thermodynamical behaviors of a material with microstruture, or (2) materials' behaviors at micro (nano) or mesoscale.

In recent years, micromechanics has become an indispensible part of theoretical foundation for many engineering fields and emergying technologies such as nanotechnology and biomedical technology.

The term "micromechanics" has become a truly interdiscipline jargon. It has been used with different meanings in different contexts. Traditionally, in the area of applied mechanics, micromechanics is referred to as a hierarchical mechanics paradigm that deals the effective material properties that are statistical averagies of a nested two level structure: microscopic and macroscopic structures. A material point at a macrolevel can be viewed as an ensemble microscope material space. The physical laws at macrolevel or the material behaviors at macro-level are derived from the ensemble average of massive micro-objects governed by the physical laws at microlevel. For instance, the effective material properties at macrolevel are the average of material properties of microstructures at fine scale. In general, the two-level paradigm is a special mathematical abstraction that is not associated with any fixed length scale. When studying material properties of a metal, $1mm$ may be viewed as macroscale, and the length scale at microlevel may range from $\mathring{A}$ to nm; whereas studying the deformation of a dam, the macroscale could be up to 10^3 m, and the length scale at microlevel may be around 10^{-2} m. In this sense, traditional micromechanics is essentially a particular (in some sense classical) averaging theory that takes into account the overall effects of microstructures. In practice, it deals with subjects of a broad spectrum: material properties of composite/synthetic materials, e.g. composite structures, cementitious materials, geotechnical materials, and phase transformations; material properties of bio-materials, e.g. constitutive modeling of bone, muscle, blood flow; environmental problems e.g. air pollutions, ground water transport and diffusion, oil spill in the ocean, etc.

In condensed matter physics and today in applied mechanics as well, the term micromechanics is used to describe a three-level physics realm: micromechanics at molecular or atomic level ($\overset{\circ}{\mathrm{A}}$), meso-mechanics at nm length scale, and macroscopic phenomenological theory at mm level or up.

The main task of contemporary micromechanics, or nano-mechanics, is to seek unknown physical laws or mechanics regulations at the nano-scale. Different from traditional micromechanics, a salient feature of nanomechanics is its multiscale and multi-physics character. It includes some features that are present in quantum mechanics, or quantum statistical mechanics, a manifestation of the effects at atomic or sub-atomic level; on the other hand, it also shares with many features from the description of continuum mechanics, because of the size statistical ensemble.

The impetus for contemporary micromechanics or nano-mechanics is primarily due to the emergence of nanoscience and bio-medical technology. It appears that physics along is not sufficient to deal with the many problems that are appearing from today's nano-technologies and nano-engineering. There is a call for a nano-mechanics and nano-computational mechanics to serve as the infra-structure of these emerging engineering fields. For instatnces, much attention has been focused on material properties of thin film, manufacturing devices and components of a microelectromechanical system (MEMS), e.g. sub-micro size sensors, motors, the mechanics of nanotube and nanowire, computer-aided material design, and micro-biophysics/biochemistry systems, e.g. protein/DNA interaction in biomolecular simulation (e.g. Schlik et al [1999ab]), etc.

From the perspective of higher learning and intellectual advancement, micromechanics has developed into a rigorous mathematical theory, philosophical methodology, and beautiful computational realization. Forty years ago, micro-elasticity started with simple definitions of eigenstrain and inclusion, came along with Eshelby's elegant equivalent homogenization theory (Eshelby [1957],[1959],[1961]) and Hashin & Shtrikman's variational principle (Hashin and Shtrikman [1962ab],[1964]), it is now the foundation of an entire composite material industry.

Less than ten years ago, Lattice Boltzmann method first debuted as a numerical emulation of continuous Boltzmann equation in statistical physics. Today, Lattice Boltzmann method has become a bona fide computational mesomechanics paradigm, and it has been used to solve problems such as turbulence flow (Qian et al [1992][1993]), combustion, and flow pass through porous media and even cooling of packed flowers (Van der Sman [1997][2000]); In later 1980s, Clementi and his co-workers [1988] initiated the idea of multiscale modeling, or multiscale simulation, i.e. using super-computers to conduct large scale computations that combine ab initio modeling, classical molecular dynamic modeling, and phenomenological modeling in a single simulation. The unified macroscopic, atomistic, ab initio dynamics (MAAD) description brings all three descriptions together into a seamless union, embracing all the size scales, from the very small to the very big (e.g. Abraham et al [1996],[1997ab],[2000]).

The simplest and earlist multi-scale modeling notion is the so-called Cauchy-Born rule. By combining this concept with the finite element methods, the so-called quasicontinuum method was developed by Tadmor, Ortiz, Phillips and their co-workers (Tadmor et al 1996). The Cauchy-Born rule is ensentially a simplistic "homogenization postulation" in lattice kinematics, and it serves as passage to link between the molecular dynamics and continuum mechanics. The Born rule assumes that the continuum energy density W can be computed using an atomic potential, with the link to the continuum being the deformation gradient $\mathbf{F}$. To briefly review continuum mechanics, the deformation gradient $\mathbf{F}$ maps an undeformed line segment $d\mathbf{X}$ in the reference configuration onto a deformed line segment $d\mathbf{x}$ in the current configuration,

$$d\mathbf{x} = \mathbf{F} d\mathbf{X} \tag{1.1}$$

In general, $\mathbf{F}$ can be written as

$$\mathbf{F} = \mathbf{I} + \frac{d\mathbf{u}}{d\mathbf{X}} \tag{1.2}$$

where $\mathbf{u}$ is the displacement vector. If there is no displacement in the continuum, the deformation gradient is equal to unity.

The major restriction and implication of the Cauchy-Born rule is that the continuum deformation must be homogeneous. This results from the fact that the underlying atomic system is forced to deform according to the continuum deformation gradient $\mathbf{F}$. By using the Born rule, one may be able to derive a continuum stress tensor and tangent stiffness directly from the interatomic potential, which allowed the usage of the standard nonlinear finite element method. This procedure is now called as the so-called quasi-continuum method.

Apparently, the contemporary mico-mechanics or nano-mechanics is only at its infancy. There are many unknown approaches to be explored and many new phenonmena to be studied. In this lecture notes, we are attempting to synthesize the most recent research results in the forefront of nano-mechanics while presenting traditional micro-mechanics in a coherent fashion. By doing so, we hope that it may serve as a stepping stone for us to reach a new height in the quest for a multiscale nano-mechanics of our time.

Chapter 2

PRELIMINARY

2.1 Vectors and Tensors

2.1.1 Vectors

Consider a Cartisian coordinate in a three dimensional space with unit vector basis, $\{\mathbf{e}_i\}, i = 1, 2, 3.$ An arbitrary position vector, $\mathbf{x}$, may be expressed as

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 = x_i\mathbf{e}_i = (\mathbf{x}\cdot\mathbf{e}_i)\mathbf{e}_i \tag{2.1}$$

where Einstein convention is used that the repeated indices indicates summation from 1 to 3.

Consider two vectors, $\mathbf{V} = V_i\mathbf{e}_i$ and $\mathbf{W} = W_j\mathbf{e}_j$. The scalar (dot) product of two vectore, $\mathbf{V}$ and $\mathbf{W}$, is defined as

$$\mathbf{V}\cdot\mathbf{W} = \Big(V_i\mathbf{e}_i\Big)\cdot\Big(W_j\mathbf{e}_j\Big) = V_iW_j\Big(\mathbf{e}_i\cdot\mathbf{e}_j\Big) = V_iW_j\delta_{ij} = V_iW_i \tag{2.2}$$

where

$$\mathbf{e}_i\cdot\mathbf{e}_j = \left\{\begin{array}{ll} 1, & i = j \\ 0, & i \neq j \end{array}\right\} =: \delta_{ij} \tag{2.3}$$

is called Keronecker delta.

A cross product of two vectors, $\mathbf{A} = A_i\mathbf{e}_i, \mathbf{B} = B_j\mathbf{e}_j$, is defined as

$$\mathbf{A}\times\mathbf{B} = (A_i\mathbf{e}_i)\times(B_j\mathbf{e}_j) = A_iB_j\mathbf{e}_i\times\mathbf{e}_j = e_{kij}A_iB_j\mathbf{e}_k \tag{2.4}$$

where $\mathbf{e}_i\times\mathbf{e}_j = e_{kij}\mathbf{e}_k$, and e_{kij} is called the permutation symbol,

$$e_{ijk} = \left\{\begin{array}{ll} 1, & \text{for an even permutation of } 1,2,3 \\ -1, & \text{for an odd permutation of } 1,2,3 \\ 0, & \text{repeated indices} \end{array}\right. \tag{2.5}$$

This definition can be explained as a permutation rule that change of any two adjcent indces of the symbol, there is a negative sign (-1) occurs.

For example, since $e_{123} = 1$, then

$$e_{132} = (-1)e_{123} = (-1)(1) = -1$$

and

$$e_{312} = (-1)e_{132} = (-1)(-1)e_{123} = (-1)(-1)1 = 1$$

The cross product of two vectors can also written as

$$\begin{aligned}
\mathbf{A} \times \mathbf{B} &= e_{kij}A_iB_j\mathbf{e}_k = e_{1ij}A_iB_j\mathbf{e}_1 + e_{2ij}A_iB_j\mathbf{e}_2 + e_{3ij}A_iB_j\mathbf{e}_3 \\
&= (A_2B_3 - A_3B_2)\mathbf{e}_1 + (A_3B_1 - A_1B_3)\mathbf{e}_2 + (A_1B_2 - A_2B_1)\mathbf{e}_3 \\
&= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \qquad (2.6)
\end{aligned}$$

Therefore

$$\mathbf{e}_i \times \mathbf{e}_j = e_{kij}\mathbf{e}_k, \quad \Rightarrow \quad e_{kij} = (\mathbf{e}_i \times \mathbf{e}_j) \cdot \mathbf{e}_k \qquad (2.7)$$

Since

$$\mathbf{e}_i \times \mathbf{e}_j = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \delta_{1i} & \delta_{2i} & \delta_{3i} \\ \delta_{1j} & \delta_{2j} & \delta_{3j} \end{vmatrix} \qquad (2.8)$$

then

$$e_{kij} = e_{ijk} = (\mathbf{e}_i \times \mathbf{e}_j) \cdot \mathbf{e}_k = \begin{vmatrix} \delta_{1k} & \delta_{2k} & \delta_{3k} \\ \delta_{1i} & \delta_{2i} & \delta_{3i} \\ \delta_{1j} & \delta_{2j} & \delta_{3j} \end{vmatrix} = \begin{vmatrix} \delta_{1i} & \delta_{2i} & \delta_{3i} \\ \delta_{1j} & \delta_{2j} & \delta_{3j} \\ \delta_{1k} & \delta_{2k} & \delta_{3k} \end{vmatrix} \qquad (2.9)$$

This provides a link between permutation symbol and Keronecker delta. Consider the product of two permutation symbols,

$$\begin{aligned}
e_{ijk}e_{rst} &= \begin{vmatrix} \delta_{1i} & \delta_{2i} & \delta_{3i} \\ \delta_{1j} & \delta_{2j} & \delta_{3j} \\ \delta_{1k} & \delta_{2k} & \delta_{3k} \end{vmatrix} \begin{vmatrix} \delta_{1r} & \delta_{2r} & \delta_{3r} \\ \delta_{1s} & \delta_{2s} & \delta_{3s} \\ \delta_{1t} & \delta_{2t} & \delta_{3t} \end{vmatrix} \\
&= \begin{vmatrix} \delta_{1i} & \delta_{2i} & \delta_{3i} \\ \delta_{1j} & \delta_{2j} & \delta_{3j} \\ \delta_{1k} & \delta_{2k} & \delta_{3k} \end{vmatrix} \begin{vmatrix} \delta_{1r} & \delta_{2s} & \delta_{3t} \\ \delta_{1r} & \delta_{2s} & \delta_{3t} \\ \delta_{1r} & \delta_{2s} & \delta_{3t} \end{vmatrix} \\
&= \begin{vmatrix} \delta_{ir} & \delta_{is} & \delta_{it} \\ \delta_{jr} & \delta_{js} & \delta_{jt} \\ \delta_{kr} & \delta_{ks} & \delta_{kt} \end{vmatrix} \qquad (2.10)
\end{aligned}$$

One may show that for any second order tensor $\mathbf{A}$, 1 When $i = r$, $e_{ijk}e_{ist} = \delta_{js}\delta_{kt} - \delta_{jt}\delta_{ks}$;

2 When $i = r$ and $j = s$, $e_{ijk}e_{ij\ell} = 2\delta_{k\ell}$;

3 When $i = r$, $j = s$, and $k = t$, $e_{ijk}e_{ijk} = 3! = 6$.

which are call $e - \delta$ identities.

2.1.2 Tensor Algebra

Consider two vectors, $\mathbf{A} = A_i\mathbf{e}_i$ and $\mathbf{B} = B_j\mathbf{e}_j$. One can form a second order tensor, $\mathbf{C}$ by using the tensor product

$$\mathbf{C} = \mathbf{A} \otimes \mathbf{B} = \Big(A_i\mathbf{e}_i\Big) \otimes \Big(B_j\mathbf{e}_j\Big) = A_iB_j\mathbf{e}_i \otimes \mathbf{e}_j \tag{2.11}$$

The dyad is called the second order tensor [1], and its basis, $\mathbf{e}_i \otimes \mathbf{e}_j$, is called dyadic basis. In this case, the components of the second order tensor are $C_{ij} = A_iB_j$.

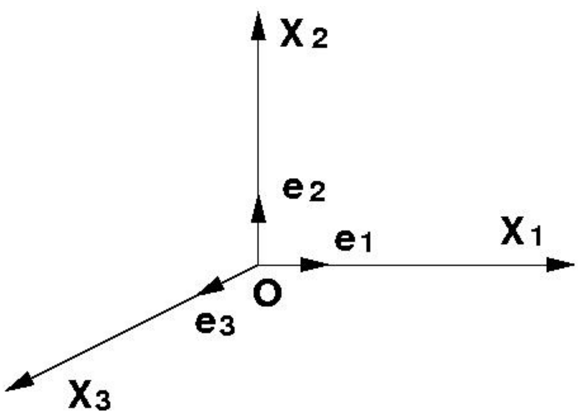


Figure 2.1. Cartesian Coordinate

In fact, every second order tensor can be expressed in a dyadic basis, such as

$$\boldsymbol{\sigma} = \sigma_{ij}\mathbf{e}_i \otimes \mathbf{e}_j \tag{2.12}$$
$$\boldsymbol{\epsilon} = \epsilon_{ij}\mathbf{e}_i \otimes \mathbf{e}_j \tag{2.13}$$

[1]One may call the vector as the first order tensor.

A conjugate of a dyad (second order tensor) is defined as

$$\Big(\boldsymbol{\epsilon}\Big)^T := \epsilon_{ji}\mathbf{e}_i \otimes \mathbf{e}_j \tag{2.14}$$

This is why in linear elasticity we may define the infinitesimal strain tensor as

$$\boldsymbol{\epsilon} = \frac{1}{2}\Big(\nabla \otimes \mathbf{u} + (\nabla \otimes \mathbf{u})^T\Big) = \frac{1}{2}\Big(u_{j,i} + u_{i,j}\Big)\mathbf{e}_i \otimes \mathbf{e}_j \tag{2.15}$$

or in component form $\epsilon_{ij} = \frac{1}{2}\Big(u_{j,i} + u_{i,j}\Big)$.

In general, a n-th order tensor is a polyads, or has a polyadic representation, e.g.

$$\mathbf{C} = C_{ijk\ell}\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_\ell \tag{2.16}$$

is a forth order tensor.

Analogous to the scalar product of vectors, the *double contraction* of two tensors are defined as two dot products among of Cartesian tensor bases, i.e. if $\mathbf{A} = A_{ij}\mathbf{e}_i \otimes \mathbf{e}_j$ and $\mathbf{B} = B_{k\ell}\mathbf{e}_k \otimes \mathbf{e}_\ell$, then

$$\begin{aligned} \mathbf{A} : \mathbf{B} &= (A_{ij}\mathbf{e}_i \otimes \mathbf{e}_j) : (B_{k\ell}\mathbf{e}_k \otimes \mathbf{e}_\ell) = A_{ij}B_{k\ell}(\mathbf{e}_i \cdot \mathbf{e}_k)(\mathbf{e}_j \cdot \mathbf{e}_\ell) \\ &= A_{ij}B_{k\ell}\delta_{ik}\delta_{j\ell} = A_{ij}B_{ij} \end{aligned} \tag{2.17}$$

The trace of a second order tensor is defined as

$$tr\mathbf{A} := \mathbf{A} : \mathbf{1}^{(2)} = A_{ii} = A_{11} + A_{22} + A_{33} \tag{2.18}$$

In each contraction, there are two bases annihilated. Consider a forth order tensor $\mathbf{C} = C_{ijk\ell}\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_\ell$ and a second order tensor $\boldsymbol{\epsilon} = \epsilon_{ij}\mathbf{e}_i \otimes \mathbf{e}_j$. There are total six basis vectors. A double contraction between the two will annihilate four basis vectors and produce a second order tensor, i.e.

$$\begin{aligned} \boldsymbol{\sigma} &= \mathbf{C} : \boldsymbol{\epsilon} = \Big(C_{ijk\ell}\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_{ell}\Big) : \Big(\epsilon_{st}\mathbf{e}_s \otimes \mathbf{e}_t\Big) \\ &= C_{ijk\ell}\epsilon_{st}\mathbf{e}_i \otimes \mathbf{e}_j\delta_{ks}\delta_{\ell t} = C_{ijk\ell}\epsilon_{k\ell}\mathbf{e}_i \otimes \mathbf{e}_j \end{aligned} \tag{2.19}$$

In component form, $\sigma_{ij} = C_{ijk\ell}\epsilon_{k\ell}$.

We say that a second order tensor is symmetric, if

$$\mathbf{A} = \Big(\mathbf{A}\Big)^T, \ \text{ or in component form } \ A_{ij} = A_{ji} \tag{2.20}$$

A second order tensor is skew symmetric, if

$$\mathbf{A} = -\Big(\mathbf{A}\Big)^T, \ \text{ or in component form } \ A_{ij} = -A_{ji} \tag{2.21}$$

In general, an arbitrary second order tensor can be expressed as

$$A_{ij} = \frac{1}{2}\Big(A_{ij} + A_{ji}\Big) + \Big(A_{ij} - A_{ji}\Big) = A_{(ij)} + A_{[ij]} \tag{2.22}$$

Denote an arbitrary second order Cartesian basis as

$$\mathbf{e}_{ij} = \mathbf{e}_i \otimes \mathbf{e}_j \ . \tag{2.23}$$

The second order unit tensor and the forth order unit tensor are constructed based on the following rules:

$$\mathbf{1}^{(2)} := \Big(\mathbf{e}_i \cdot \mathbf{e}_j\Big)\mathbf{e}_i \otimes \mathbf{e}_j = \delta_{ij}\mathbf{e}_i \otimes \mathbf{e}_j = \delta_{ij}\mathbf{e}_{ij} \tag{2.24}$$

$$\begin{aligned} \mathbf{1}^{(4)} &:= \Big(\mathbf{e}_i \otimes \mathbf{e}_j\Big) : \Big(\mathbf{e}_k \otimes \mathbf{e}_\ell\Big)\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_\ell \\ &= (\mathbf{e}_{ij} : \mathbf{e}_{k\ell})\mathbf{e}_{ij} \otimes \mathbf{e}_{k\ell} = \delta_{ik}\delta_{j\ell}\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_\ell \end{aligned} \tag{2.25}$$

The superscript indicates the order. It is interesting to note that the fourth order unit tensor defined in (2.25) is not symmetric with all indices.

To represent symmetric tensors, it may be expedient to first define symmectric tensor basis. The second order symmetric basis is defined as

$$\mathbf{e}^S_{ij} = \frac{1}{2}\Big(\mathbf{e}_{ij} + \mathbf{e}^T_{ij}\Big) = \frac{1}{2}\Big(\mathbf{e}_i \otimes \mathbf{e}_j + \mathbf{e}_j \otimes \mathbf{e}_i\Big) \tag{2.26}$$

Any second order symmetric tensor can then be expressed as $\mathbf{S} = S_{ij}\mathbf{e}^S_{ij}$. One may denote the space of all second order symmetric tensors as

$$T^{(2s)} = \{\mathbf{S} \;\Big|\; \mathbf{S} = S_{ij}\mathbf{e}^S_{ij}\} \tag{2.27}$$

The corresponding second order symmetric unit tensor is then defined as

$$\begin{aligned} \mathbf{1}^{(2s)} &= \frac{1}{2}\Big(\mathbf{e}_i \cdot \mathbf{e}_j + \mathbf{e}_j \cdot \mathbf{e}_i\Big)\mathbf{e}_i \otimes \mathbf{e}_j \\ &= \delta_{ij}\mathbf{e}_i \otimes \mathbf{e}_j = \mathbf{1}^{(2)} \end{aligned} \tag{2.28}$$

One may also define the second order anti-symmetric tensor as $\mathbf{e}^A_{ij} = \frac{1}{2}\Big(\mathbf{e}_i \otimes \mathbf{e}_j - \mathbf{e}_j \otimes \mathbf{e}_i\Big)$.

The fourth-order symmetric tensor bases is built upon the second order symmetric tensor bases, i.e.

$$\mathbf{e}^S_{ijk\ell} = \mathbf{e}^S_{ij} \otimes \mathbf{e}^S_{k\ell} \tag{2.29}$$

and the fourth-order symmectric tensor space is defined as

$$T^{(4s)} = \{\mathbf{S} \;\Big|\; \mathbf{S} = S_{ijk\ell}\mathbf{e}^S_{ijk\ell}\} \tag{2.30}$$

The corresponding fourth-order unit tensor is defined as

$$\mathbf{1}^{(4s)} := \mathbf{e}^S_{ij} : \mathbf{e}^S_{k\ell}\mathbf{e}^S_{ijk\ell} == \frac{1}{2}\Big(\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}\Big)\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_\ell \qquad (2.31)$$

It may be noted that the fourth-order unit tensor can be decomposed to symmetric part and antisymmetric part in terms of the first and second indices, or the of the third and forth indices,

$$\begin{aligned} 1^{(4)}_{ijk\ell} &:= \delta_{ik}\delta_{j\ell} = \frac{1}{2}(\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}) + \frac{1}{2}(\delta_{ik}\delta_{j\ell} - \delta_{i\ell}\delta_{jk}) \\ &= 1^{(4s)}_{ijk\ell} + 1^{(4a)}_{ijk\ell} \end{aligned} \qquad (2.32)$$

One may show that for given second-order tensor, $\mathbf{A}$,

$$\mathbf{1}^{(4)} : \mathbf{A} \;\rightarrow\; \mathbf{A} \qquad (2.33)$$

$$\mathbf{1}^{(4s)} : \mathbf{A} \;\rightarrow\; \frac{1}{2}\Big(\mathbf{A} + \mathbf{A}^T\Big) \qquad (2.34)$$

$$\mathbf{1}^{(4a)} : \mathbf{A} \;\rightarrow\; \frac{1}{2}\Big(\mathbf{A} - \mathbf{A}^T\Big) \qquad (2.35)$$

Note that $\mathbf{1}^{(4)} \neq \mathbf{1}^{(2)} \otimes \mathbf{1}^{(2)}$.

2.1.3 Inversion formula for fourth-order isotropic tensor

Consider general form of fourth order isotropic tensor,

$$\mathbf{Q} = m\mathbf{1}^{(2)} \otimes \mathbf{1}^{(2)} + 2w\mathbf{1}^{(4s)} \qquad (2.36)$$

Let $\mathbf{Q}^{-1}$ be its inverse tensor. According to the well-known Sherman-Morrision formula (e.g. Dahlquist and Bjorck [1974]),

$$\mathbf{Q}^{-1} = -\frac{m}{2w(3m+2w)}\mathbf{1}^{(2)} \otimes \mathbf{1}^{(2)} + \frac{1}{2w}\mathbf{1}^{(4s)} \;. \qquad (2.37)$$

In component form,

$$Q_{ijk\ell} = m\delta_{ij}\delta_{k\ell} + w(\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}) \qquad (2.38)$$

$$Q^{-1}_{ijk\ell} = -\frac{m}{2w(3m+2w)}\delta_{ij}\delta_{k\ell} + \frac{1}{4w}(\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}) \qquad (2.39)$$

A more straightforward approach to invert an isotropic tensor is to adopt the following E-basis orthogonal decomposition. Let

$$\mathbf{E}^{(1)} := \frac{1}{3}\mathbf{1}^{(2)} \otimes \mathbf{1}^{(2)} \;,\; E^{(1)}_{ijk\ell} = \frac{1}{3}\delta_{ij}\delta_{k\ell} \qquad (2.40)$$

$$\begin{aligned} \mathbf{E}^{(2)} &:= -\frac{1}{3}\mathbf{1}^{(2)} \otimes \mathbf{1}^{(2)} + \mathbf{1}^{(4s)} \\ \Rightarrow E^{(2)}_{ijk\ell} &= -\frac{1}{3}\delta_{ij}\delta_{k\ell} + \frac{1}{2}(\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}) \end{aligned} \qquad (2.41)$$

The E-bases have the following special properties,

$$\begin{aligned}
\mathbf{E}^{(1)} + \mathbf{E}^{(2)} &= \mathbf{1}^{(4s)} \\
\mathbf{E}^{(1)} : \mathbf{E}^{(1)} &= \mathbf{E}^{(1)}, \text{ and } \mathbf{E}^{(2)} : \mathbf{E}^{(2)} = \mathbf{E}^{(2)} \\
\mathbf{E}^{(1)} : \mathbf{E}^{(2)} &= \mathbf{E}^{(2)} : \mathbf{E}^{(1)} = \mathbf{0}\,.
\end{aligned}$$

We now use E-basis approach to verify Sherman-Morrison formula. Let,

$$\mathbf{Q} = (3m + 2w)\mathbf{E}^{(1)} + 2w\mathbf{E}^{(2)} \tag{2.42}$$

and

$$\mathbf{Q}^{-1} = h\mathbf{E}^{(1)} + v\mathbf{E}^{(2)} \tag{2.43}$$

By definition,

$$\begin{aligned}
\mathbf{Q} : \mathbf{Q}^{-1} &= \mathbf{1}^{(4s)} = \mathbf{E}^{(1)} + \mathbf{E}^{(2)} \\
(3m + 2w)h\mathbf{E}^{(1)} + 2wv\mathbf{E}^{(2)} &= \mathbf{E}^{(1)} + \mathbf{E}^{(2)}
\end{aligned}$$

which then leads to

$$h = \frac{1}{3m + 2w} \tag{2.44}$$

$$v = \frac{1}{2w} \tag{2.45}$$

Consequently, we can write that

$$\begin{aligned}
\mathbf{Q}^{-1} &= (h - v)\mathbf{E}^{(1)} + v(\mathbf{E}^{(1)} + \mathbf{E}^{(2)}) \\
&= -\frac{3m}{2w(3m + 2w)}\mathbf{E}^{(1)} + \frac{1}{2w}\mathbf{1}^{(4s)} \\
&= -\frac{m}{2w(3m + 2w)}\mathbf{1}^{(2)} \otimes \mathbf{1}^{(2)} + \frac{1}{2w}\mathbf{1}^{(4s)}
\end{aligned}$$

Let's practice more examples.

EXAMPLE 2.1 *Consider an isotropic elastic tensor,*

$$\begin{aligned}
\mathbf{C} &= \lambda\mathbf{1}^{(2)} \otimes \mathbf{1}^{(2)} + 2\mu\mathbf{1}^{(4s)} \\
&= 3K\mathbf{E}^{(1)} + 2\mu\mathbf{E}^{(2)}
\end{aligned}$$

Since by definition, $\mathbf{C} : \mathbf{D} = \mathbf{1}^{(4s)}$*, it can be readily shown that*

$$\begin{aligned}
\mathbf{D} &= \frac{1}{3K}\mathbf{E}^{(1)} + \frac{1}{2\mu}\mathbf{E}^{(2)} \\
&= -\frac{\lambda}{2\mu(3\lambda + 2\mu)}\mathbf{1}^{(2)} \otimes \mathbf{1}^{(2)} + \frac{1}{2\mu}\mathbf{1}^{(4s)}
\end{aligned}$$

EXAMPLE 2.2 *For spherical inclusion, the Eshelby tensor is*

$$\begin{aligned}
\mathbf{S}^{\Omega} &= \frac{5\nu-1}{15(1-\nu)}\mathbf{1}^{(2)}\otimes\mathbf{1}^{(2)} + \frac{2(4-5\nu)}{15(1-\nu)}\mathbf{1}^{(4s)} \\
&= \frac{(1+\nu)}{3(1-\nu)}\mathbf{E}^{(1)} + \frac{2(4-5\nu)}{15(1-\nu)}\mathbf{E}^{(2)} \\
&= s_1\mathbf{E}^{(1)} + s_2\mathbf{E}^{(2)}
\end{aligned}$$

where $s_1 = \dfrac{1+\nu}{3(1-\nu)}$ *and* $s_2 = \dfrac{2(4-5\nu)}{15(1-\nu)}$.

Then

$$\begin{aligned}
(\mathbf{S}^{\Omega})^{-1} &= \frac{3(1-\nu)}{1+\nu}\mathbf{E}^{(1)} + \frac{15(1-\nu)}{2(4-5\nu)}\mathbf{E}^{(2)} \\
&= \frac{(1-\nu)(3-5\nu)}{2(1+\nu)(4-5\nu)}\mathbf{1}^{(2)}\otimes\mathbf{1}^{(2)} + \frac{15(1-\nu)}{2(4-5\nu)}\mathbf{1}^{(4s)}
\end{aligned}$$

Moreover,

$$\begin{aligned}
\mathbf{T}^{\Omega} &= \mathbf{1}^{(4s)} - \mathbf{C} : \mathbf{S}^{\Omega} : \mathbf{D} \\
&= (\mathbf{E}^{(1)} + \mathbf{E}^{(2)}) - (3K\mathbf{E}^{(1)} + 2\mu\mathbf{E}^{(2)}) : (s_1\mathbf{E}^{(1)} + s_2\mathbf{E}^{(2)}) \\
&\quad : \Big(\frac{1}{3K}\mathbf{E}^{(1)} + \frac{1}{2\mu}\mathbf{E}^{(2)}\Big) \\
&= (1-s_1)\mathbf{E}^{(1)} + (1-s_2)\mathbf{E}^{(2)}
\end{aligned}$$

2.1.4 Tensor analysis

Define gradient operator as

$$\nabla = \frac{\partial}{\partial x_i}\mathbf{e}_i \tag{2.46}$$

It is a vector operation.

Applying gradient operator to a scalar function, $f \in C^0(\Omega)$, $\Omega \subset \mathbb{R}^d$, will result a vector. In other words, the gradient of a scalar function (zero-th order tensor) is a first order tensor, i.e.

$$grad\, f := \nabla f = \Big(\frac{\partial}{\partial x_i}\mathbf{e}_i\Big) f = \frac{\partial f}{\partial x_i}\mathbf{e}_i \tag{2.47}$$

For a vector function, $\mathbf{A}(\mathbf{x}) = A_i(\mathbf{x})\mathbf{e}_i$, its gradient is a tensor product between the gradient operator and the vector field,

$$grad\, \mathbf{A} := \nabla \otimes \mathbf{A} = \Big(\frac{\partial}{\partial x_i}\mathbf{e}_i\Big) \otimes A_j\mathbf{e}_j = \frac{\partial A_j}{\partial x_i}\mathbf{e}_i \otimes \mathbf{e}_j \tag{2.48}$$

The gradient of a vector field, a first order tensor field, is a second order tensor. In general, the gradient operation increases the ordero f a tensorial field up to one order higher.

On the other hand, the scalar product or contraction between a gradient operator and a tensorial field is called *divergence* operation, which will result a new tensorial field with reduced order. Consider a vector field, $\mathbf{A} = A_i \mathbf{e}_i$. Its divergence is being defined as

$$div \mathbf{A} := \nabla \cdot \mathbf{A} = \Big(\frac{\partial}{\partial x_i} \mathbf{e}_i \Big) \cdot \Big(A_j \mathbf{e}_j \Big) = \frac{\partial A_j}{\partial x_i} (\mathbf{e}_i \cdot \mathbf{e}_j) = \frac{\partial A_i}{\partial x_i} \tag{2.49}$$

The cross product between the gradient operator and a tensorial field. $\mathbf{A} = A_i \mathbf{e}_i$, is called the *Curls* or *rot* of the tensorial field.

$$Curl \mathbf{A} := \nabla \times \mathbf{A} = \frac{\partial A_j}{\partial x_i} \Big(\mathbf{e}_i \times \mathbf{e}_j \Big) = e_{ijk} \partial_i A_j e_k = e_{ijk} \partial_j A_k e_k \tag{2.50}$$

In what follows, a few integral transformations are listed.

Suppose that there is a continuous function, $f(x) \in C^1(\Omega)$, defined in a domain $\Omega \in \mathbb{R}^d$ with smooth boundary $\partial\Omega$. A well-known integral theorem is

$$\int_\Omega \nabla f d\Omega = \int_{\partial\Omega} f \mathbf{n} dS \tag{2.51}$$

or in component form

$$\int_\Omega \frac{\partial f}{\partial x_i} d\Omega = \int_{\partial\Omega} f n_i dS \tag{2.52}$$

In general for a smooth tensorial field, $\mathbf{A}$, we have the following statement,

$$\int_\Omega \nabla \otimes \mathbf{A} d\Omega = \int_{\partial\Omega} \mathbf{n} \otimes \mathbf{A} dS \tag{2.53}$$

Consider a continuous m-order tensorial field, $A(x) \in [C^1(\Omega)]^m \times d$, the well known divergence theorem can be expressed in a Cartesian coordinate as

$$\int_\Omega \nabla \cdot \mathbf{A} d\Omega = \int_{\partial\Omega} \mathbf{n} \cdot \mathbf{A} dS \tag{2.54}$$

If $\mathbf{A}$ is a vector field, i.e. $\mathbf{A} = A_i \mathbf{e}_i$, the divergence theorem can be expressed in a component form as

$$\int_\Omega \frac{\partial A_i}{\partial x_i} d\Omega = \int_{\partial\Omega} n_i A_i dS \tag{2.55}$$

If we consider the volume integration of a cross product between gradient operator and the tensorial field, we can have the following integral transformation,

$$\int_{\Omega} \nabla \times \mathbf{A} d\Omega = \int_{\partial\Omega} \mathbf{n} \times \mathbf{A} dS \tag{2.56}$$

Again, if $\mathbf{A}$ is a vector field, we may write its Cartesian component form,

$$\int_{\Omega} e_{ijk} \frac{\partial A_k}{\partial x_j} d\Omega = \int_{\partial\Omega} e_{ijk} n_j A_k dS \tag{2.57}$$

2.2 Review of Linear Elasticity Theory

To set the stage, we first review the basic formulations of infinitesimal, linear elasticity theory.

• Equations of motion

Denote $\boldsymbol{\sigma} = \sigma_{ij} \mathbf{e}_i \otimes \mathbf{e}_j$ as Cauchy stress tensor, and $\mathbf{u} = u_i \mathbf{e}_i$ as the infinitesimal displacement field, ρ as the density of the continuum, and $\mathbf{b} = b_i \mathbf{e}_i$ as the body force per unity volume. The equation of motion of a material particle can be expressed in a Cartesian coorinate as $\forall x \in \Omega$,

$$\nabla \cdot \boldsymbol{\sigma} + \rho \mathbf{b} = \rho \frac{\partial^2 \mathbf{u}}{\partial t^2} \tag{2.58}$$

For convenience, we often write the component form

$$\sigma_{ji,j} + \rho b_i = \frac{\partial^2 u_i}{\partial t^2} \tag{2.59}$$

where $u_{ji,j} = \dfrac{\partial u_{ji}}{\partial x_j}$.

• Geometric relation

The infinitesimal strain field $\boldsymbol{\epsilon} = \epsilon_{ij} \mathbf{e}_i \otimes \mathbf{e}_j$ is defined as

$$\boldsymbol{\epsilon} = \frac{1}{2}\Big(\nabla \otimes \mathbf{u} + (\nabla \otimes \mathbf{u})^T\Big) \tag{2.60}$$

Note that $\nabla \otimes \mathbf{u} = u_{j,i} \mathbf{e}_i \otimes \mathbf{e}_j$. Hence $(\nabla \otimes \mathbf{u})^T = u_{i,j} \mathbf{e}_i \otimes \mathbf{e}_j$.

Therefore in component form,

$$\epsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i}) \tag{2.61}$$

• Constitutive equations

For linear elastic solids, the constitutive equations have the following form,

$$\boldsymbol{\sigma} = \mathbf{C} : \boldsymbol{\epsilon} \;\; \Rightarrow \;\; \sigma_{ij} = C_{ijkl} \epsilon_{kl} \tag{2.62}$$

where $\mathbf{C} = C_{ijkl}\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \otimes \mathbf{e}_l$ is the elasticity tensor.

For isotropic elastic media, it has the form,

$$\mathbf{C} = \lambda \mathbf{I} \otimes \mathbf{I} + 2\mu \mathbf{1}^{(4s)} \tag{2.63}$$

where λ, μ are Lame constants. In component form, it reads

$$C_{ijkl} = \lambda \delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \tag{2.64}$$

Inversely, one may write that

$$\boldsymbol{\epsilon} = \mathbf{C}^{-1} : \boldsymbol{\sigma} = \mathbf{D} : \boldsymbol{\sigma} \;\; \epsilon_{ij} = D_{ijkl}\sigma_{kl} \tag{2.65}$$

where the fourth order tensor, $\mathbf{D}$, is called compliance tensor. For isotropic materials, it has the form

$$D_{ijkl} = -\frac{\lambda}{2\mu(3\lambda + 2\mu)}\delta_{ij}\delta_{kl} + \frac{1}{4\mu}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \tag{2.66}$$

• Compatibility condition

Compatibility conditions for infinitesimal deformation field may be expressed as (Melvan [1969]),

$$\nabla \times \boldsymbol{\epsilon} \times \nabla = \mathbf{0} \tag{2.67}$$

In indicial natation, it reads,

$$e_{pki}e_{qlj}\epsilon_{ij,kl} = 0 \tag{2.68}$$

or alternatively

$$\epsilon_{ij,kl} + \epsilon_{kl,ij} - \epsilon_{ik,jl} - \epsilon_{il,jk} = 0 \tag{2.69}$$

• Elastic potential energy

The strain energy density is defined as

$$U(\boldsymbol{\epsilon}) = \int_0^{\boldsymbol{\epsilon}} \boldsymbol{\sigma}(\boldsymbol{\epsilon}^{'}) : d\boldsymbol{\epsilon}^{'} \tag{2.70}$$

Based on foundamental theorem of calculus, one may find its inverse relationship as

$$\frac{\partial U}{\partial \boldsymbol{\epsilon}} = \boldsymbol{\sigma}, \;\; \frac{\partial U}{\partial \epsilon_{ij}} = \sigma_{ij} \tag{2.71}$$

The complementary strain energy density can be obtained via Legendre transform,

$$U^*(\boldsymbol{\sigma}) = \boldsymbol{\sigma} : \epsilon - U(\boldsymbol{\epsilon}) \tag{2.72}$$

Or one may define

$$U^*(\boldsymbol{\sigma}) = \int_0^{\boldsymbol{\sigma}} \boldsymbol{\epsilon}(\boldsymbol{\sigma}^{'})d\boldsymbol{\sigma}^{'} \tag{2.73}$$

One may derive that

$$\boldsymbol{\epsilon} = \frac{\partial U^*}{\partial \boldsymbol{\sigma}}, \ \ or \ \ \epsilon_{ij} = \frac{\partial U^*}{\partial \sigma_{ij}} \tag{2.74}$$

For linear elastic materials,

$$C_{ijkl}\epsilon_{kl} = \frac{\partial U}{\partial \epsilon_{ij}} \ \ \Rightarrow \ \ C_{ijkl} = \frac{\partial^2 U}{\partial \epsilon_{ij}\partial \epsilon_{kl}} \tag{2.75}$$

In general, for hyperelastic media, the elastic stiffness tensor can be calculated based on the formula

$$C_{ijkl} = \frac{\partial^2 U}{\partial \epsilon_{ij}\partial \epsilon_{kl}} \tag{2.76}$$

Similarly, one may find elastic compliance tensor by calculation

$$D_{ijkl} = \frac{\partial^2 U^*}{\partial \sigma_{ij}\partial \sigma_{kl}} \tag{2.77}$$

Change the order of differentiation in Eq.(2.66),

$$\frac{\partial^2 U}{\partial \epsilon_{ij}\partial \epsilon_{kl}} = \frac{\partial^2 U}{\partial \epsilon_{kl}\partial \epsilon_{ij}} \tag{2.78}$$

One may derive that $C_{ijkl} = C_{klij}$.

Furthermore since $\epsilon_{ij} = \epsilon_{ji}$ and $\epsilon_{kl} = \epsilon_{lk}$, $C_{ijkl} = C_{jikl} = C_{ijlk} = C_{jlik}$. These are called minor symmetry.

Similar conclusions can be drawn from elastic compliance tensors as well.

Both elastic tensor $\mathbf{C}$ and compliance tensor $\mathbf{D}$ are positive definite, because both strain energy density and complementary strain energy density must be positive, i.e.

$$U(\boldsymbol{\epsilon}) = \frac{1}{2}\boldsymbol{\epsilon} : \mathbf{C} : \boldsymbol{\epsilon} = \frac{1}{2}C_{ijkl}\epsilon_{ij}\epsilon_{kl} > 0$$

$$U^*(\boldsymbol{\sigma}) = \frac{1}{2}\boldsymbol{\sigma} : \mathbf{D} : \boldsymbol{\sigma} = \frac{1}{2}D_{ijkl}\sigma_{ij}\sigma_{kl} > 0$$

By definition that a fourth-order tensor, C_{ijkl}, is positive-definite, when

$$\frac{1}{2}C_{ijkl}\epsilon_{ij}\epsilon_{kl} > 0, \ \ \forall \epsilon_{ij} \tag{2.79}$$

where equality holds only if $\epsilon_{ij} = 0$.

2.2.1 Betti's reciprocal theorem and Somigliana Identity

Consider two sets of different self-equilibrating states: $\left\{\mathbf{u}^{(\alpha)}, \boldsymbol{\epsilon}^{(\alpha)}, \boldsymbol{\sigma}^{(\alpha)}, \mathbf{f}^{(\alpha)}\right\}$, $\alpha = 1, 2$,

$$\nabla \cdot \boldsymbol{\sigma}^{(\alpha)} + \mathbf{f}^{(\alpha)} = 0 \tag{2.80}$$

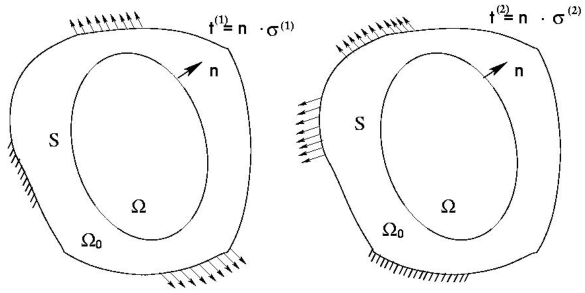


Figure 2.2. Two sets of different self-equilibrating states

with boundary conditions,

$$\mathbf{n}\cdot\boldsymbol{\sigma}(\alpha) = \mathbf{t}^{(\alpha)0},\ \forall \mathbf{x}\in\Gamma_t^0 \tag{2.81}$$
$$\mathbf{u}^{(\alpha)} = \mathbf{u}^{(\alpha)0},\ \forall \mathbf{x}\in\Gamma_u^0,\ \ \alpha=1,2 \tag{2.82}$$

acting in a same object Ω_0.

The Betti's reciprocal theorem[2] states that: the work done by the first set of self-equilibrating surface traction, $\mathbf{t}^{(1)}$, and body force $\mathbf{f}^{(1)}$ in any interior region $\Omega\subset\Omega_0$, going through the displacement field, $\mathbf{u}^{(2)}$, of the second self-equilibrating system, equals the work done by the second set of tractions, $\mathbf{t}^{(2)}$, and the body force, $\mathbf{f}^{(2)}$, in the same interior region going through the displacement field, $\mathbf{u}^{(1)}$, of the first self-equilibrating system, i.e.

$$\int_\Omega f_i^{(1)}u_i^{(2)}d\Omega+\int_{\partial\Omega}t_i^{(1)}u_i^{(2)}dS=\int_\Omega f_i^{(2)}u_i^{(1)}d\Omega+\int_{\partial\Omega}t_i^{(2)}u_i^{(1)}dS \tag{2.83}$$

Proof:

Consider both states being equilibrium states. It has

$$\begin{aligned}\int_\Omega f_i^{(1)}u_i^{(2)}d\Omega &= -\int_\Omega \sigma_{ji,j}^{(1)}u_i^{(2)}d\Omega\\ &= -\int_{\partial\Omega}\sigma_{ji}^{(1)}n_ju_i^{(2)}dS+\int_\Omega\sigma_{ji}^{(1)}u_{i,j}^{(2)}d\Omega\\ &= -\int_{\partial\Omega}t_i^{(1)}u_i^{(2)}dS+\int_\Omega\sigma_{ji}^{(1)}\epsilon_{ji}^{(2)}d\Omega \end{aligned}\tag{2.84}$$

Moving the first term of the right-hand side of (2.74) to the left-hand side yields

$$\int_\Omega f_i^{(1)}u_i^{(2)}d\Omega+\int_{\partial\Omega}t_i^{(1)}u_i^{(2)}dS=\int_\Omega\sigma_{ij}^{(1)}\epsilon_{ij}^{(2)}d\Omega \tag{2.85}$$

[2]Precisely speaking, it is the Betti's second reciprocal theorem.

Similarly, one may show that

$$\int_\Omega f_i^{(2)} u_i^{(1)} d\Omega + \int_{\partial\Omega} t_i^{(2)} u_i^{(1)} dS = \int_\Omega \sigma_{ij}^{(2)} \epsilon_{ij}^{(1)} d\Omega \tag{2.86}$$

Consider the fact that the two systems exist in the same material

$$\int_\Omega \sigma_{ij}^{(1)} \epsilon_{ij}^{(2)} d\Omega = \int_\Omega C_{ijkl} \epsilon_{kl}^{(1)} \epsilon_{ij}^{(2)} d\Omega = \int_\Omega C_{klij} \epsilon_{kl}^{(1)} \epsilon_{ij}^{(2)} d\Omega = \int_\Omega \epsilon_{kl}^{(1)} \sigma_{kl}^{(2)} d\Omega$$

Compare the both sides of (2.75) and (2.76), the theorem holds.

In addition, the equality

$$\int_\Omega \sigma_{ij}^{(1)} \epsilon_{ij}^{(2)} d\Omega = \int_\Omega \sigma_{ij}^{(2)} \epsilon_{ij}^{(1)} d\Omega \tag{2.87}$$

is called Betti's first reciprocal theorem.

To derive Somigliana identity, we first consider Dirac's delta function, which is the limit of the following function, $\delta(x) = lim_{\epsilon \to 0} \delta_\epsilon(x)$,

$$\delta_\epsilon(x) = \lim_{\epsilon \to 0} \begin{cases} 0; & x < -\epsilon/2 \\ 1/\epsilon; & -\epsilon/2 < x < \epsilon/2 \\ 0; & x > \epsilon/2 \end{cases} \tag{2.88}$$

A graph of Dirac's delta function is shown in Fig. 2.3.

Dirac delta function has following properties

$$(1) \qquad \int_{-\infty}^{\infty} \delta(x) dx = 1 \tag{2.89}$$

$$(2) \qquad \int_{-\infty}^{\infty} \delta(x-y) f(y) dy = f(x) \tag{2.90}$$

The first property (2.79) can be easily shown by definition that

$$\int_{-\infty}^{\infty} \delta(x) dx = \int_{-\epsilon/2}^{\epsilon/2} \frac{1}{\epsilon} dx = 1 \tag{2.91}$$

To show the second property, we let $x - y = z$ and $dy = -dz$. Thus

$$\begin{aligned} \int_{-\infty}^{\infty} \delta(x-y) f(y) dy &= -\int_{\infty}^{-\infty} \delta(z) f(x-z) dz = \int_{-\infty}^{\infty} \delta(z) f(x-z) dz \\ &= \frac{1}{\epsilon} \int_{-\epsilon/2}^{\epsilon/2} f(x-z) dz = \frac{1}{\epsilon} f(x - \zeta \frac{\epsilon}{2}) \int_{-\epsilon/2}^{\epsilon/2} dz \\ &= f(x), \;\; as \;\; \epsilon \to 0 \end{aligned} \tag{2.92}$$

where $-1 < \zeta < 1$.

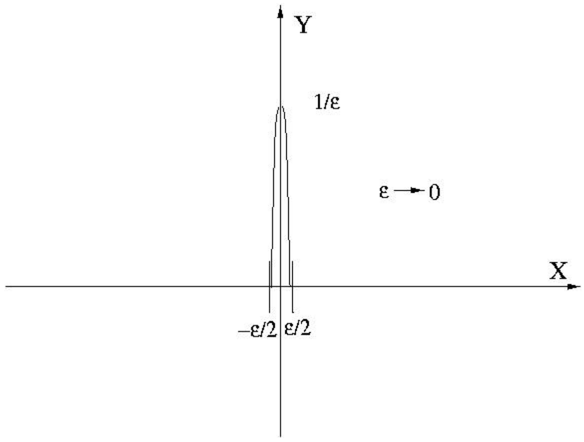


Figure 2.3. Dirac's delta function

Consider an infinitely space filled with homogeneous elastic medium. The body force is form of concentrated load at a fixed point y,

$$\mathbf{f} = \delta(\mathbf{x} - \mathbf{y})\delta_{mk}\mathbf{e}_k \tag{2.93}$$

The subscript index m is in the direction of m.

The equilibrium equations then have the form,

$$\nabla \cdot \boldsymbol{\sigma}_m + \delta(\mathbf{x} - \mathbf{y})\delta_{mk}\mathbf{e}_k = 0, \;\; \forall \mathbf{x} \in \mathbb{R}^3 \tag{2.94}$$

The displacement solution of this problem is called foundamental solution of Navier equation, or the Green's function for an infinitely extended homogeneous elastic domain. Denote the displacement solution as

$$\mathbf{u}_m = \mathbf{G}^{\infty}_m(\mathbf{x}, \mathbf{y}) = G^{\infty}_{mi}(\mathbf{x}, \mathbf{y})\mathbf{e}_i \tag{2.95}$$

The corresponding strain and stress fields are:

$$\epsilon^{G^\infty_m}_{ij} = \frac{1}{2}\Big(G^{\infty}_{mi,j} + G^{\infty}_{mj,i}\Big), \;\; \sigma^{G^\infty_m}_{ij} = C_{ijkl}\epsilon^{G^\infty_m}_{ij} \tag{2.96}$$

Next, we consider a singly connected finite region $\Omega \subset \mathbb{R}^3$. The finite region Ω is in a self-equilibrating state, i.e., there is a body force distribution: $\nabla \cdot \boldsymbol{\sigma} + \mathbf{f} = \mathbf{0}, \;\; \forall \mathbf{x} \in \Omega$, and a traction force distribution: $\mathbf{t} = \mathbf{n} \cdot \boldsymbol{\sigma}, \;\; \forall \mathbf{x} \in \partial\Omega$.

Let

$$\begin{aligned}
\mathbf{f}^{(1)}(\mathbf{x}) &= \delta(\mathbf{x} - \mathbf{y})\delta_{mk}\mathbf{e}_k, \;\; \mathbf{u}^{(1)}(\mathbf{x}) = G^{\infty}_{mi}(\mathbf{x}, \mathbf{y})\mathbf{e}_i && (2.97)\\
\mathbf{t}^{(1)}(\mathbf{x}) &= \sigma^{G^\infty_m}_{ij}(\mathbf{x}) n_j \mathbf{e}_i && (2.98)\\
\mathbf{f}^{(2)}(\mathbf{x}) &= f_i(\mathbf{x})\mathbf{e}_i, \;\; \mathbf{u}^{(2)}(\mathbf{x}) = u_i(\mathbf{x})\mathbf{e}_i && (2.99)\\
\mathbf{t}^{(2)}(\mathbf{x}) &= \sigma_{ij}(\mathbf{x}) n_j \mathbf{e}_i && (2.100)
\end{aligned}$$

Apply Betti's reciprocal theorem,

$$\begin{aligned}
&\int_{\Omega} \delta(\mathbf{x}-\mathbf{y})\delta_{mi} u_i(x) d\Omega_x + \int_{\partial\Omega} n_j \sigma_{ji}^{G_m^\infty} u_i(\mathbf{x}) dS_x \\
= &\int_{\Omega} f_i(\mathbf{x}) G_{mi}^\infty(\mathbf{x},\mathbf{y}) d\Omega_x + \int_{\partial\Omega} n_j \sigma_{ji} G_{mi}^\infty(\mathbf{x},\mathbf{y}) dS_x \qquad (2.101)
\end{aligned}$$

Considering the property of Dirac delta function, one can obtain:

$$\begin{aligned}
u_m(\mathbf{y}) = &\int_{\Omega} f_i(\mathbf{x}) G_{mi}^\infty(\mathbf{x},\mathbf{y}) d\Omega_x + \int_{\partial\Omega} t_i(\mathbf{x}) G_{mi}^\infty(\mathbf{x},\mathbf{y}) dS_x \qquad (2.102) \\
&- \int_{\partial\Omega} t_i^{G_m^\infty}(\mathbf{x},\mathbf{y}) u_i(\mathbf{x}) dS_x, \quad m = 1,2,3
\end{aligned}$$

Equation (2.92) is the well-known Somigliana identity.

2.3 Exercises

PROBELM 2.1 *Let $\delta\mathbf{u}$ be a virtual displacement field and $\boldsymbol{\sigma}$ be a self-equilibrium stress field. Show*

$$\Big(\nabla \cdot \boldsymbol{\sigma}\Big) \cdot \delta\mathbf{u} = \nabla \cdot \Big(\boldsymbol{\sigma} \cdot \delta\mathbf{u}\Big) - \boldsymbol{\sigma} : (\nabla \otimes \delta\mathbf{u}) \qquad (2.103)$$

PROBELM 2.2 *Assume body force $\mathbf{f} = 0$. The elastostatic equilibrium equation takes the form:*

$$\sigma_{ji,j} = 0, \ \ or \ \ \nabla \cdot \boldsymbol{\sigma} = \mathbf{0} \qquad (2.104)$$

Show

$$\int_{\Omega} \sigma : \boldsymbol{\epsilon} d\Omega = \int_{\partial\Omega} \mathbf{t} \cdot \mathbf{u} dS \qquad (2.105)$$

where $\mathbf{t} = \mathbf{n} \cdot \boldsymbol{\sigma}$.

(Hint: use Gauss theorem, the divergence theorem.)

PROBELM 2.3 *Suppose that there are two different solutions of equilibrium equation,*

$$\nabla \cdot \boldsymbol{\sigma}_1 = \mathbf{0}, \ \ \nabla \cdot \boldsymbol{\sigma}_2 = \mathbf{0} \qquad (2.106)$$

which satisfy the same boundary conditions,

$$\begin{cases} \mathbf{u}_1 = \mathbf{u}^0, \\ \mathbf{u}_2 = \mathbf{u}^0; \end{cases} \quad \forall \mathbf{x} \in \Gamma_u \qquad (2.107)$$

$$\begin{cases} \mathbf{n} \cdot \boldsymbol{\sigma}_1 = \mathbf{t}^0, \\ \mathbf{n} \cdot \boldsymbol{\sigma}_2 = \mathbf{t}^0; \end{cases} \quad \forall \mathbf{x} \in \Gamma_t \qquad (2.108)$$

where $\Gamma_u \bigcup \Gamma_t = \partial\Omega$.

By using the positive-definiteness of elastic tensor and compliance tensor, show:

$$\Delta\boldsymbol{\sigma} = \boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2 = \mathbf{0} \tag{2.109}$$

$$\Delta\boldsymbol{\epsilon} = \epsilon_1 - \epsilon_2 = \mathbf{0} \tag{2.110}$$

PROBELM 2.4 *Show that for a given second-order tensor,* $\mathbf{A}$*,*

$$\mathbf{1}^{(4)} : \mathbf{A} \rightarrow \mathbf{A} \tag{2.111}$$

$$\mathbf{1}^{(4s)} : \mathbf{A} \rightarrow \frac{1}{2}\Big(\mathbf{A} + \mathbf{A}^T\Big) \tag{2.112}$$

$$\mathbf{1}^{(4a)} : \mathbf{A} \rightarrow \frac{1}{2}\Big(\mathbf{A} - \mathbf{A}^T\Big) \tag{2.113}$$

Chapter 3

HOMOGENIZATION I — CLASSICAL AVERAGING METHOD

"Curiouser and curiouser!" cried Alice,"Now I'm opening out like the largest telecope that everwas!"

— Lewis Carroll, *Alice in Wonderland*

3.1 Representative volume element

One of the foundamental concept in classical micromechanics is the so-called *Representative volume element*, or RVE.

The classical micromechanics paradigm is a two-level hierarchical mechanical structure: Macro-level and Micro-level, or it consists of two elements: macr o-element and micro-element. At macro-level, a continuum is made of many material points, and each material point is related with a micro-space. A macro material point is also called a macro-element, or volume element. Its associated micro-space contains many micro-elements. In fact, it is a microscopic continuum. If a material is statistically homogeneous at macro-level, to study material behaviors, we only need to examine material properties at an arbitrary (typical) macro-point, and the micro-space associated with that macro-point is called the representative volume element.

An RVE for a material point of a continuum mass is a statistical ensemble of microscale objects surrounding or constituting the macro material point. This means that an RVE should contain a very large number of micro-elements such that it can be a statistically representative of the local continuum properties, or it is statistically stable.

In essence, the concept of representative volume element in classical micromechanics is a mathematical paradigm. It has no fixed length scale associated with each level.

The length scales associated macro-level and microlevel are relative. If you study effective material properties of a heterogeous metal, the lengthscale of microlevel maybe from a few *nm* to μm, and the lengthscale of macrolevel may be from a few mm to centimeter. If you study the stiffness of a dam, the lengthscale of microlevel could be from centimeters, whereas the lengthscale of macro-level could be meters.

In classical mechanics, at macro-level, the material properties are always assumed to be homogeneous but unknown, whereas at micro-level, i.e., inside the RVE, the material properties are heterogeneous but known.

At microlevel, the heterogeneous micro-structure is known and physical laws is known. The task of micromechanics is based on information of microstructure to find homogeneous material properties at macro-level, which is often called *overall material properties* or *effective material properties*.

The methodology to find effective material properties is called *homogenization.* Homogenization is another word that has been widely used in many different contexts. In this book, the term "homogenization" is used to mean statistical averaging. There are mainly two sets of homogenization methods, mathematical homogenization and mechanical homogenization.

The objectives of micromechanics is to find both material properties at macro-level, or overall (effective) material properties and physical laws at macro-level.

The first subject of continuum micromechanics if micro-elasticity. The basic premises of microelasticity is to assume that inside an RVE, the micro-constitutive relation of a material is elastic, and in more cases, they are assumed to be linear elastic. In micromechanics, the concept of the RVE is used to derive material properties due to microstructures. In most cases, the micro-structures are often independent with gravity or other types of body forces. Therefore, in micro-continuum mechanics, the body force effect is often negleted. The equilibrium equations inside an RVE is often written as

$$\nabla \cdot \boldsymbol{\sigma} = 0 \quad \Rightarrow \quad \sigma_{ij,j} = 0. \tag{3.1}$$

3.2 Average stress in an RVE

Definition of average operator $< \cdot >$. Suppose that **T(x,X)** is a general tensor field defined in an RVE. Note that here **x** is the spatial coordinate inside an RVE for a fixed material point, whereas **X** is the spatial coordinate of the material point with respect to a macro-coordinate. If at macro-level, material is homogeneous, i.e. material properties at macro-level do no change from place to place, **X** is often dropped out. We simply write **T = T(x)**, which means that one RVE is sufficient to represent all the material points in the object that is under investigation.

To associate a micro-level tensor field with a tensorial quantity at macro-level is called homogenization. To do so, we first define the so-called average operator. The average value of the tensor field **T(x)** at a material point is defined as

$$< \mathbf{T} >_{\mathbf{X}} := \frac{1}{V} \int_V \mathbf{T}(\mathbf{x}, \mathbf{X}) dV_x \tag{3.2}$$

If the material is homogeneous at macro-level, we have

$$< \mathbf{T} > := \frac{1}{V} \int_V \mathbf{T}(\mathbf{x}) dV_x \tag{3.3}$$

For instance, if $\mathbf{T} = \boldsymbol{\sigma}(\mathbf{x})$ is a micro-stress field, the macro-stress at a material point will be $\boldsymbol{\Sigma} =< \boldsymbol{\sigma} >$. Similarly, if $\mathbf{T} = \boldsymbol{\epsilon}(\mathbf{x})$ is a micro-strain field, the macro-strain at a material point is $\mathcal{E} =< \boldsymbol{\epsilon} >$.

A very useful average theorem about micro-Cauchy stress tensor may be stated as follows:

THEOREM 3.1 *Suppose an RVE is subjected to natural boundary condition, and the traction on remote boundary of an RVE (∂V) is generated by a constant stress tensor, $\boldsymbol{\sigma}^0$. Then the average stress at this material point, or the macro stress at the material point,*

$$\boldsymbol{\Sigma} =< \boldsymbol{\sigma} >= \boldsymbol{\sigma}^0 \tag{3.4}$$

Note that the point here is that one only knows the traction distribution on the remote boundary of the RVE, but one does not know the exact stress distribution inside the RVE.

Proof

Consider,

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij} \ \text{ and } \ \sigma_{ji,j} = 0 \tag{3.5}$$

One then can express Cauchy stess inside an RVE as

$$\begin{aligned} \sigma_{ij} &= \sigma_{ik}\delta_{kj} = \sigma_{ik}\delta_{jk} = \Big(\sigma_{ik}\frac{\partial x_j}{\partial x_k}\Big) \\ &= (\sigma_{ik}x_j)_{,k} - \sigma_{ik,k}x_j = (\sigma_{ik}x_j)_{,k} \end{aligned} \tag{3.6}$$

Therefore,

$$\begin{aligned} < \sigma_{ij} > &= \frac{1}{V}\int_V \sigma_{ij} dV = \frac{1}{V}\int_V \Big(\sigma_{ik}x_j\Big)_{,k} dV \\ &= \frac{1}{V}\oint_{\partial V} \sigma_{ik}x_j n_k dS = \frac{1}{V}\oint_{\partial V} \sigma^0_{ik} x_j n_k dS \\ &= \frac{\sigma^0_{ik}}{V}\oint_{\partial V} x_j n_k dS = \frac{\sigma^0_{ik}}{V}\int_V \frac{\partial x_j}{\partial x_k} dV \\ &= \frac{\sigma^0_{ik}}{V}\int_V \delta_{jk} dV = \frac{\sigma^0_{ik}}{V}\delta_{jk} V = \sigma^0_{ij} \end{aligned} \tag{3.7}$$

3.3 Average strain and strain rate

Consider a displacement field, $\mathbf{u} = u_i \mathbf{e}_i$, inside an RVE. Suppose that on the remote boundary of the RVE, the displacement filed is prescribed,

$$u_i(\mathbf{x}) = u_i^0(\mathbf{x}), \quad \forall \mathbf{x} \in \partial V \tag{3.8}$$

One can find the average displacement gradient field in terms of boundary data, i.e.,

$$< u_{i,j} >= \frac{1}{V} \int_V u_{i,j} dV = \frac{1}{V} \int_{\partial V} n_j u_i^0 dS \tag{3.9}$$

Note that you don't know exact distribution of the displacement field inside the RVE.

Moreover, one may find the average strain and rotation fields in terms of boundary displacement data,

$$< \epsilon_{ij} >= \frac{1}{2}\Big(< u_{i,j} > + < u_{j,i} >\Big) = \frac{1}{2V} \oint_{\partial V} (n_j u_i^0 + n_i u_j^0) dS$$

$$< \omega_{ij} >= \frac{1}{2}\Big(< u_{i,j} > - < u_{j,i} >\Big) = \frac{1}{2V} \oint_{\partial V} (n_j u_i^0 - n_i u_j^0) dS$$

REMARK 3.3.1 *in general, the average displacement fields of an RVE can not be expressed in terms of remote surface data. To see this, one may evaluate the average displacement field. Using the trick,*

$$u_i = u_k \delta_{ki} = u_k \delta_{ik} = u_k \frac{\partial x_i}{\partial x_k} = (u_k x_i)_{,k} - u_{k,k} x_i$$

Hence

$$\begin{aligned} < u_i > &= \frac{1}{V} \int_V u_i dV = \frac{1}{V} \int \Big((u_k x_i)_{,k} - u_{k,k} x_i \Big) dV \\ &= \frac{1}{V} \Big(\oint_{\partial V} u_k^0 x_i n_k dS - \int_V u_{k,k} x_i dV \Big) \end{aligned} \tag{3.10}$$

It is clear that $< u_i >$ *can not be expressed in terms of boundary data, unless* $u_{k,k} = 0$.

However, for incompressible materials, such as rubber or plastic zone of ductile materials, it is often true that $u_{k,k} = 0$. *Therefore,*

$$< u_i >= \frac{1}{V} \int_V u_i dV = \frac{1}{V} \oint_{\partial V} u_k^0 x_i n_k dS \tag{3.11}$$

An average theorem for infinitesimal strain can be stated as follows.

THEOREM 3.2 *Suppose that an RVE is only subjected to essential boundary condition. On the remote surface of the RVE, its displacement fields are prescribed as*

$$\mathbf{u}^0 = \boldsymbol{\epsilon}^0 \cdot \mathbf{x}, \quad \Rightarrow \quad u_i^0 = \epsilon_{ij}^0 x_j \tag{3.12}$$

where ϵ_{ij}^0 *is a constant strain tensor. Then, the average strain field of the RVE equals the constant strain tensor, i.e.*

$$< \boldsymbol{\epsilon} > = \boldsymbol{\epsilon}^0, \quad \Rightarrow \quad < \epsilon_{ij} > = \epsilon_{ij}^0 \tag{3.13}$$

Proof:

First of all, the prescribed essential boundary condition does not necessarily generate a constant strain field inside the RVE, i.e.

$$\epsilon_{ij}(\mathbf{x}) \neq \epsilon_{ij}^0$$

In fact

$$\epsilon_{ij}(\mathbf{x}) = \epsilon_{ij}^0 + \tilde{\epsilon}_{ij}(\mathbf{x}), \ \ \forall \mathbf{x} \in V$$

and the perturbation strain field satisfying $\tilde{\epsilon}_{ij}(\mathbf{x}) = 0, \ \ \forall \mathbf{x} \in \partial V$.

By definition,

$$\begin{aligned} < \epsilon_{ij} > &= \frac{1}{V}\int_V \epsilon_{ij} dV = \frac{1}{2V}\int_V \Big(u_{i,j} + u_{j,i}\Big) dV \\ &= \frac{1}{2V}\oint_{\partial V} (u_i^0 n_j + u_j^0 n_i) dS \\ &= \frac{1}{2V}\oint_{\partial V} (x_k \epsilon_{ki}^0 n_j + x_k \epsilon_{kj}^0 n_i) dS \\ &= \frac{1}{2V}\oint_{\partial V} (\epsilon_{ki}^0 \delta_{kj} V + \epsilon_{kj}^0 \delta_{ki} V) = \epsilon_{ij}^0 \end{aligned} \tag{3.14}$$

One may also show the following identities about average virtual work and average strain energy density.

$$< \boldsymbol{\sigma} : \delta\boldsymbol{\epsilon} > = \frac{1}{V}\oint_{\partial V} \mathbf{t} \cdot \delta \mathbf{u} dS \tag{3.15}$$

$$\begin{aligned} & < \boldsymbol{\sigma} : \boldsymbol{\epsilon} > - < \boldsymbol{\sigma} > : < \boldsymbol{\epsilon} > \\ = \ & \frac{1}{V}\oint_{\partial V} \Big(\mathbf{u} - \mathbf{x}\cdot < \nabla \otimes \mathbf{u} > \Big) \cdot \Big(\mathbf{n} \cdot (\boldsymbol{\sigma} - < \boldsymbol{\sigma} >)\Big) dS \end{aligned} \tag{3.16}$$

Since $\sigma_{ij}\delta\epsilon_{ij} = \frac{1}{2}\sigma_{ij}(\delta u_{i,j} + \delta u_{j,i}) = \sigma_{ij}\delta u_{i,j}$,

$$\begin{aligned}
\frac{1}{V}\int_V \sigma_{ij}\delta\epsilon_{ij}dV &= \frac{1}{V}\int_V \sigma_{ij}\delta u_{i,j}dV \\
&= \frac{1}{V}\int_V \Big(\sigma_{ij}\delta u_i\Big)_{,j} dV = \frac{1}{V}\oint_{\partial V} \sigma_{ij}\delta u_i n_j dS \\
&= \frac{1}{V}\oint_{\partial V} t_i \delta u_i dS \qquad (3.17)
\end{aligned}$$

where $t_i := n_j\sigma_{ji}$. Hence,(3.15)holds.

To show (3.16), one may write

$$\begin{aligned}
& \frac{1}{V}\int_{\partial V}\Big(u_i - x_j < u_{i,j} >\Big)\Big(n_k(\sigma_{ki} - < \sigma_{ki} >\Big)dS \\
= & \frac{1}{V}\int_{\partial V}\Big(u_i n_k \sigma_{ki} - u_i n_k < \sigma_{ki} > - x_j < u_{i,j} > n_k \sigma_{ki} \\
& + x_j < u_{i,j} > n_k < \sigma_{ki} >\Big)dS \\
= & \frac{1}{V}\int_V \sigma_{ki} u_{i,k} dV - \Big(\frac{1}{V}\int_V u_{i,k} dV\Big) < \sigma_{ki} > \\
& - \delta_{jk} < u_{i,j} > \frac{1}{V}\int_V \sigma_{ki} dV + < \epsilon_{ij} >< \sigma_{ij} > \\
= & < \sigma_{ij}\epsilon_{ij} > - < \sigma_{ij} >< \epsilon_{ij} > \qquad (3.18)
\end{aligned}$$

3.4 Definition of eigenstrain, eigenstress, and inclusion

’Eigenstrain’ is a generic name to describe a transformation strain field that can equivalently represent induced strain due to misfit of inhomogeneities, thermal expansion, plastic strain, residual strain , phase transformation, etc., all of which, when homogeneously applied produce a compatible deformation field without generating stresses. The German word "*eigen*" means characteristic. It is believed that any strain field generated by an inhomogeneity distribution may have a one-to-one correspondence to a fictitious eigenstrain field, which is characteristically equivalent (in the sense of mechanical variables, such as stress, strain, and displacements) to the induced strain field generated by the inhomogeneity distribution.

’Eigenstress’ is a generic name given to self-equilibrated transformation stress (internal) field that can generate equivalent perturbed stress and strain distributions caused by one or several of there eigenstrains in bodies which are free from any other external forces and surface constraints. The eigenstress field is created by the incompatibility of the eigenstrains.

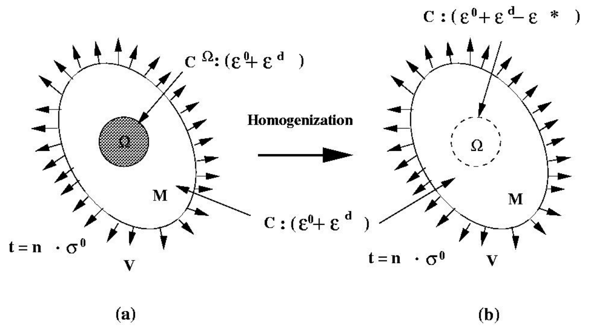


Figure 3.1. Illustration of Eshelby's equivalent eigenstrain principle. (a)Initial heterogeneous body, (b) equivalent homogeneous body ($V = \Omega \cup M$).

The term *inclusion* denotes a subdomain in the matrix subjected to transformation strains (eigenstrains), while the inhomogeneity is a subdomain with properties distinct from those from the matrix.

3.5 Eshelby's equivalent eigenstrain method I: Traction boundary condition

Eshelby's equivalent eigenstrain principle is a homogenization method. It establishes the equivalency between an eigenstrain (eitenstress) field and an inhomogeneity distribution, such that distribution of inhomogeneities may be replaced by the eigenstrain field with the equivalent mechanical effect. This equivalency mapping process translates the heterogeneity of material into an added non-uniform strain distribution, while making the material properties become homogeneous again.

Let's consider an Elastic solid, V, with elasticity tensor, $\mathbf{C}$, and compliance tensor, $\mathbf{D}$. Inside the elastic solid, there is an inhomogeneity, a subdomain, Ω, with different elastic constants, $\mathbf{C}^{\Omega}$ and $\mathbf{D}^{\Omega}$ (see Fig. 3.1).

The so-called Eshelby's equivalent eigenstrain principle, or Mura's equivalent eigenstrain principle, is to replace the inhomogeneity with a homogenized inclusion, within which an eigenstrain field is prescribed, such that the homogenized field is mechanical equivalent to the original inhomogeneous field.

Consider that the original inhomogeneous solid is subjected to a traction boundary condition, $\mathbf{t} = \mathbf{n} \cdot \boldsymbol{\sigma}^0$. The presence of inhomogeneity will produce stress perturbation and hence the strain field perturbation,

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d, \quad \boldsymbol{\epsilon} = \boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d \,.$$

The stress and strain distributions inside the inhomogeneous solid are

$$\begin{aligned} \boldsymbol{\sigma} &= \begin{cases} \mathbf{C} : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d) & \mathbf{x} \in M \\ \mathbf{C}^\Omega : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d) & \mathbf{x} \in \Omega \end{cases} \\ \boldsymbol{\epsilon} &= \begin{cases} \mathbf{D} : (\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d) & \mathbf{x} \in M \\ \mathbf{D}^\Omega : (\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d) & \mathbf{x} \in \Omega \end{cases} \end{aligned} \tag{3.19}$$

The Eshelby's equivalent eigenstrain homogenization method is to choose a suitable strain field,

$$\boldsymbol{\epsilon} = \begin{cases} 0, & \forall \mathbf{x} \in M \\ \boldsymbol{\epsilon}^*, & \forall \mathbf{x} \in \Omega \end{cases} \tag{3.20}$$

to superpose with the actual strain field,$\boldsymbol{\epsilon} = \boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d$, such that the total strain field of homogenized solid is equivalent to the total strain field of inhomogeneous solid, i.e.

$$\begin{aligned} \boldsymbol{\sigma}(\mathbf{x}) &= \mathbf{C} : (\boldsymbol{\epsilon}(\mathbf{x}) - \boldsymbol{\epsilon}^*(\mathbf{x})) \\ &= \begin{cases} \mathbf{C} : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d) \\ \mathbf{C} : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d - \boldsymbol{\epsilon}^*) \end{cases} = \begin{cases} \mathbf{C} : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d), & \mathbf{x} \in M \\ \mathbf{C}^\Omega : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d), & \mathbf{x} \in \Omega \end{cases} \end{aligned} \tag{3.21}$$

Consider $\boldsymbol{\epsilon}^0 = \mathbf{D} : \boldsymbol{\sigma}^0$. Under the chosen traction boundary condition, $< \boldsymbol{\sigma} >= \boldsymbol{\sigma}^0$, but $\boldsymbol{\epsilon}^0 \neq < \boldsymbol{\epsilon} >$.

From (3.21), one may derive that

$$\boldsymbol{\sigma}^d(\mathbf{x}) = \mathbf{C} : (\boldsymbol{\epsilon}^d(\mathbf{x}) - \boldsymbol{\epsilon}^*(\mathbf{x})), \ \forall \mathbf{x} \in V \tag{3.22}$$

$$\mathbf{C}^\Omega(\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d) = \mathbf{C} : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d - \boldsymbol{\epsilon}^*), \ \forall \mathbf{x} \in \Omega \tag{3.23}$$

where Eq.(3.23) is called "*stress consistency condition*". It is the criterion for choosing suitable eigenstrain field. Note that $\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d - \boldsymbol{\epsilon}^*$ is the total *elastic* strain.

Alternatively, Eqs (3.21) and (3.22) can be recast into following forms,

$$\boldsymbol{\sigma} = \mathbf{C} : (\boldsymbol{\epsilon} - \boldsymbol{\epsilon}^*) \ \Rightarrow \ \epsilon = \mathbf{D} : \boldsymbol{\sigma} + \epsilon^* \tag{3.24}$$

$$\boldsymbol{\sigma}^d = \mathbf{C} : (\epsilon^d - \boldsymbol{\epsilon}^*) \ \Rightarrow \ \epsilon^d = \mathbf{D} : \boldsymbol{\sigma}^d + \boldsymbol{\epsilon}^* \tag{3.25}$$

3.6 Eshelby's equivalent eigenstress method II: Displacement boundary condition

Consider the same inhomogeneous solid and following displacement boundary condition

$$\mathbf{u}(\mathbf{x}) = \boldsymbol{\epsilon}^0 \cdot \mathbf{x}, \ \forall \mathbf{x} \in \partial V \tag{3.26}$$

The inhomogeneity inside the solid will generate a disturbance stress field, $\boldsymbol{\sigma}$. The total stress field is

$$\epsilon(\mathbf{x}) = \begin{cases} \mathbf{D} : (\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d) \\ \mathbf{D}^\Omega : (\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d) \end{cases} \tag{3.27}$$

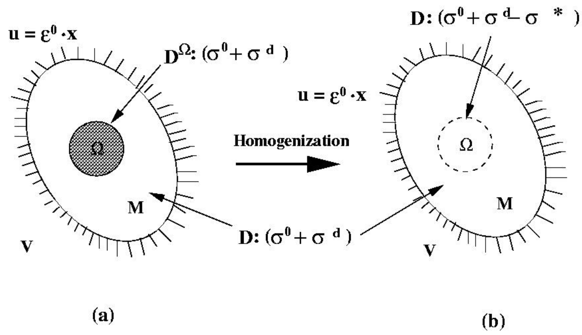


Figure 3.2. Illustration of Eshelby's equivalent eigenstress principle. (a) Initial heterogeneous body, (b)equivalent homogeneous body ($V = \Omega \cup M$).

As proved in previous section, under prescribed boundary condition, the average strain, $< \boldsymbol{\epsilon} >= \boldsymbol{\epsilon}^0$. On the other hand,$< \boldsymbol{\sigma} > \neq \boldsymbol{\sigma}^0$.

To homogenize the heterogeneous medium, we introduce the following eigenstress distribution,

$$\boldsymbol{\sigma}^*(\mathbf{x}) = \begin{cases} 0, & \forall \mathbf{x} \in M \\ \boldsymbol{\sigma}^*, & \forall \mathbf{x} \in \Omega \end{cases} \tag{3.28}$$

such that

$$\boldsymbol{\epsilon}(\mathbf{x}) = \begin{cases} \mathbf{D} : (\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d) \\ \mathbf{D} : (\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d - \boldsymbol{\sigma}^*) \end{cases} = \begin{cases} \mathbf{D} : (\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d), & \mathbf{x} \in M \\ \mathbf{D}^\Omega : (\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d), & \mathbf{x} \in \Omega \end{cases} \tag{3.29}$$

From Eq.(3.29), we can derive that

$$\boldsymbol{\epsilon}^d(\mathbf{x}) = \mathbf{D} : (\boldsymbol{\sigma}^d(\mathbf{x}) - \boldsymbol{\sigma}^*), \ \forall \mathbf{x} \in V \tag{3.30}$$

$$\mathbf{D}^\Omega(\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d) = \mathbf{D} : (\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d - \boldsymbol{\sigma}^*), \ \forall \mathbf{x} \in \Omega \tag{3.31}$$

where Eq.(3.31) is called "*strain consistency condition.*"

Alternatively,

$$\boldsymbol{\epsilon}^d(\mathbf{x}) = \mathbf{D} : (\boldsymbol{\sigma}^d(\mathbf{x}) - \boldsymbol{\sigma}^*) \ \Rightarrow \ \boldsymbol{\sigma}^d = \mathbf{C} : \boldsymbol{\epsilon}^d + \boldsymbol{\sigma}^* \tag{3.32}$$

Comparing Eq.(3.32) with (3.25) yield the following identities,

$$\boldsymbol{\epsilon}^* + \mathbf{D} : \boldsymbol{\sigma}^* = 0, \ \text{ or } \ \boldsymbol{\sigma}^* + \mathbf{C} : \boldsymbol{\epsilon}^* = 0 \tag{3.33}$$

3.7 Effective material properties via eigenstrain method

In this section, we illustrate how to use equivalent eigenstrain method to find overal material properties.

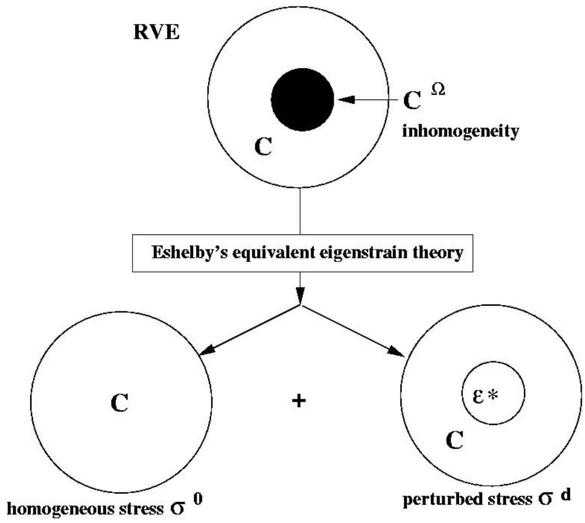


Figure 3.3. Illustration of Eshelby's equivalent eigenstrain principle

We still consider the previous problem: an RVE with only on inhomogeneity. Denote the total volume of RVE as V, the volume of the matrix as M, and the volume of the inhomogeneity as Ω. Assume that the RVE is a heterogeneous linear elastic medium and the micro-constitutive relations are:

$$\boldsymbol{\epsilon} = \mathbf{D} : \boldsymbol{\sigma}, \ \ \mathbf{x} \in M \tag{3.34}$$

$$\boldsymbol{\epsilon} = \mathbf{D}^{\Omega} : \boldsymbol{\sigma}, \ \ \mathbf{x} \in \Omega \tag{3.35}$$

Our objective is to find the constitutive relation at macro-level,i.e.

$$\boldsymbol{\Sigma} = \bar{\mathbf{C}} : \mathcal{E} \ \ \Rightarrow \ \ <\boldsymbol{\sigma}>= \bar{\mathbf{C}} :<\boldsymbol{\epsilon}> \tag{3.36}$$

Note that here we have already assumed that the constitutive relation at macro-level is also linear elastic. The only unknown is the effective compliance tensor, or effective elastic tensor. This shows the primitive feature of classical micro-elasticity. In contemporary micromechanics, one does not know whether the material behaviors at macro-level is linear elastic or some other forms. One determines macro behaviors of the material as an outcome of homogenization.

Apply the traction boundary condition on the remote surface of the RVE,

$$\mathbf{t} = \mathbf{n} \cdot \boldsymbol{\sigma}^0$$

As mentioned before, under such boundary condition, $<\boldsymbol{\sigma}>= \boldsymbol{\sigma}^0$, nevertheless, $<\boldsymbol{\epsilon}>\neq \boldsymbol{\epsilon}^0$, i.e. $<\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d>\neq \boldsymbol{\epsilon}^0$. Therefore, our goal is to find the effective elastic compliance tensor such that $<\boldsymbol{\epsilon}>= \bar{\mathbf{D}} : \boldsymbol{\sigma}^0$

Denote the average strain and stress in the matrix and in the inhomogeneity as

$$\bar{\boldsymbol{\epsilon}}^M := \frac{1}{M}\int_M \boldsymbol{\epsilon}(\mathbf{x})dV \ , \ \bar{\boldsymbol{\sigma}}^M := \frac{1}{M}\int_M \boldsymbol{\sigma}(\mathbf{x})dV; \tag{3.37}$$

$$\bar{\boldsymbol{\epsilon}}^\Omega := \frac{1}{\Omega}\int_\Omega \boldsymbol{\epsilon}(\mathbf{x})dV \ , \ \bar{\boldsymbol{\sigma}}^\Omega := \frac{1}{\Omega}\int_\Omega \boldsymbol{\sigma}(\mathbf{x})dV; \tag{3.38}$$

Therefore, $\bar{\boldsymbol{\epsilon}}^M = \mathbf{D} : \bar{\boldsymbol{\sigma}}$, and$\bar{\boldsymbol{\epsilon}}^\Omega = \mathbf{D}^\Omega : \bar{\boldsymbol{\sigma}}^\Omega$.

Consider $V = M \cup \Omega$ and let $f := \left|\frac{\Omega}{V}\right|$.Then

$$\begin{aligned}
\bar{\boldsymbol{\epsilon}} &= \frac{1}{V}\int_V \boldsymbol{\epsilon} dV = \frac{1}{V}\int_{M\cup\Omega} \boldsymbol{\epsilon} dV \\
&= \frac{1}{V}\Big(\frac{M}{M}\int_M \boldsymbol{\epsilon} dV + \frac{\Omega}{\Omega}\int_\Omega \boldsymbol{\epsilon} dV\Big) = \frac{M}{V}\bar{\boldsymbol{\epsilon}}^M + \frac{\Omega}{V}\bar{\boldsymbol{\epsilon}}^\Omega
\end{aligned} \tag{3.39}$$

Hence,

$$\begin{aligned}
\frac{M}{V}\bar{\boldsymbol{\epsilon}}^M &= < \boldsymbol{\epsilon} > - f\bar{\boldsymbol{\epsilon}}^\Omega \\
&= \bar{\mathbf{D}} : \boldsymbol{\sigma}^0 - f\mathbf{D}^\Omega : \bar{\boldsymbol{\sigma}}^\Omega
\end{aligned} \tag{3.40}$$

On the other hand,

$$\begin{aligned}
\frac{M}{V}\bar{\boldsymbol{\epsilon}}^M &= \frac{M}{V}\mathbf{D} : \bar{\boldsymbol{\sigma}}^M = \mathbf{D} : \Big(\frac{M}{V}\frac{1}{M}\int_M \boldsymbol{\sigma}(\mathbf{x})dV\Big) \\
&= \mathbf{D} : \Big(\frac{1}{V}\int_{V-\Omega} \boldsymbol{\sigma}(\mathbf{x})dV\Big) \\
&= \mathbf{D} : \Big(\frac{1}{V}\int_V \boldsymbol{\sigma}(\mathbf{x})dV - \frac{1}{V}\int_\Omega \boldsymbol{\sigma}(\mathbf{x})dV\Big) \\
&= \mathbf{D} : \Big(\boldsymbol{\sigma}^0 - f\bar{\boldsymbol{\sigma}}^\Omega\Big)
\end{aligned} \tag{3.41}$$

Compare Eqs. (3.40) and (3.41),

$$\bar{\mathbf{D}} : \boldsymbol{\sigma}^0 - f\mathbf{D}^\Omega : \bar{\boldsymbol{\sigma}}^\Omega = \mathbf{D} : \boldsymbol{\sigma}^0 - f\mathbf{D} : \bar{\boldsymbol{\sigma}}^\Omega \tag{3.42}$$

Therefore,

$$\Big(\mathbf{D} - \bar{\mathbf{D}}\Big) : \boldsymbol{\sigma}^0 = f\Big(\mathbf{D} - \mathbf{D}^\Omega\Big) : \bar{\boldsymbol{\sigma}}^\Omega = f\Big(\mathbf{D} - \mathbf{D}^\Omega\Big) :< \boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d >_\Omega \tag{3.43}$$

The equaqtion is often referred to as **The Basic Equation for Average Stress**.

By definition,

$$\bar{\boldsymbol{\sigma}}^\Omega = \mathbf{C}^\Omega :< \boldsymbol{\epsilon}^0 + \epsilon^d >_\Omega \tag{3.44}$$

From the stress consistency condition, one may obtain

$$\boldsymbol{\epsilon}^* = \mathbf{C}^{-1} : (\mathbf{C} - \mathbf{C}^\Omega) : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d) = \Big(\mathbf{A}^\Omega\Big)^{-1} : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d) \tag{3.45}$$

where $\mathbf{A}^\Omega = (\mathbf{C} - \mathbf{C}^\Omega)^{-1} : \mathbf{C}$.

If one can relate the perturbed strain with the eigenstrain, i.e.

$$\boldsymbol{\epsilon}^d = \mathbf{S}^\Omega : \boldsymbol{\epsilon}^* \tag{3.46}$$

Eq (3.45) may be rewritten as

$$\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d = \mathbf{A}^\Omega : \boldsymbol{\epsilon}^* \ \ \Rightarrow \ \ \boldsymbol{\epsilon}^* = \Big(\mathbf{A}^\Omega - \mathbf{S}^\Omega\Big)^{-1} : \boldsymbol{\epsilon}^0 \tag{3.47}$$

Subsequently,

$$\begin{aligned} \boldsymbol{\epsilon}(\mathbf{x}) &= \boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d = \mathbf{A}^\Omega : \boldsymbol{\epsilon}^* = \mathbf{A}^\Omega : (\mathbf{A}^\Omega - \mathbf{S}^\Omega)^{-1} : \boldsymbol{\epsilon}^0 \\ &= \mathbf{A}^\Omega : (\mathbf{A}^\Omega - \mathbf{S}^\Omega)^{-1} : \mathbf{D} : \boldsymbol{\sigma}^0, \ \ \forall \mathbf{x} \in \Omega \end{aligned} \tag{3.48}$$

In the literature, we denote $\mathcal{A}^\Omega = \mathbf{A}^\Omega : (\mathbf{A}^\Omega - \mathbf{S}^\Omega)^{-1}$ as the so-called "concentration tensor", because it represents the relation ship between the background strain field and the actual strain field in the inhomogeneity, i.e. how are the strains concentrated. Suppose both the Eshelby tensor $\mathbf{S}^\Omega$ and tensor $\mathbf{A}^\Omega$ are constant tensors, then $\mathcal{A}^\Omega = const.$, and

$$\boldsymbol{\epsilon}(\mathbf{x}) = \mathcal{A}^\Omega : \boldsymbol{\epsilon}^0, \ \ \forall \mathbf{x} \in \Omega \quad \Rightarrow \quad \bar{\boldsymbol{\epsilon}}^\Omega = \mathcal{A}^\Omega : \boldsymbol{\epsilon}^0 \tag{3.49}$$

Therefore,

$$\bar{\boldsymbol{\sigma}}^\Omega = \mathbf{C}^\Omega : \mathbf{A}^\Omega : (\mathbf{A}^\Omega - \mathbf{S}^\Omega)^{-1} : \mathbf{D} : \boldsymbol{\sigma}^0, \ \ \forall \mathbf{x} \in \Omega \tag{3.50}$$

Substituting the expression (3.50) into (3.43) yields

$$\Big(\bar{\mathbf{D}} - \mathbf{D}\Big) : \boldsymbol{\sigma}^0 = f(\mathbf{D}^\Omega - \mathbf{D}) : \mathbf{C}^\Omega : \mathbf{A}^\Omega : (\mathbf{A}^\Omega - \mathbf{S}^\Omega)^{-1} : \mathbf{D} : \boldsymbol{\sigma}^0 \tag{3.51}$$

Consider

$$\Big(\mathbf{D}^\Omega - \mathbf{D}\Big) : \mathbf{C}^\Omega = \mathbf{1}^{(4s)} - \mathbf{D} : \mathbf{C}^\Omega$$

and

$$\begin{aligned} \Big(\mathbf{A}^\Omega\Big)^{-1} &= \Big((\mathbf{C} - \mathbf{C}^\Omega)^{-1} : \mathbf{C}\Big)^{-1} = \mathbf{C}^{-1} : (\mathbf{C} - \mathbf{C}^\Omega) \\ &= \mathbf{1}^{(4s)} - \mathbf{C}^{-1} : \mathbf{C}^\Omega \\ &= \mathbf{1}^{(4s)} - \mathbf{D} : \mathbf{C}^\Omega = \Big(\mathbf{D}^\Omega - \mathbf{D}\Big) : \mathbf{C}^\Omega \\ &\Rightarrow \ \Big(\mathbf{D}^\Omega - \mathbf{D}\Big) : \mathbf{C}^\Omega = \Big(\mathbf{A}^\Omega\Big)^{-1} \end{aligned} \tag{3.52}$$

Therefore

$$\left(\bar{\mathbf{D}} - \mathbf{D}\right) : \boldsymbol{\sigma}^0 = f\left(\mathbf{A}^\Omega - \mathbf{S}^\Omega\right)^{-1} : \mathbf{D} : \boldsymbol{\sigma}^0 \tag{3.53}$$

It is the straightforward to derive

$$\boxed{\bar{\mathbf{D}} = \left(\mathbf{1}^{(4s)} + f(\mathbf{A}^\Omega - \mathbf{S}^\Omega)^{-1}\right) : \mathbf{D}} \tag{3.54}$$

Note that the crucial step of this derivation is the assumption that disturbance strain field can be related to eigenstrain distribution, i.e. $\boldsymbol{\epsilon}^d = \mathbf{S}^\Omega : \boldsymbol{\epsilon}^*$, where the tensor $\mathbf{S}^\Omega$ is called the Eshelby tensor. Chapter 6 will be devoted to derive Eshelby tensor for specific shapes of inhomogeneities or inclusions.

3.8 Jock Eshelby (I)

John Douglas Eshelby was born in Puddington, Cheshire, On December 21, 1916, the eldest son of Alan Douglas Eshelby. Because of ill health he missed his formal schooling from the age 13 and ilved at the family home in north Somerset, where he learned instead from tutors. So, as he used to say, he had to work many things our for himself, and perhaps this helped to make him such an original and creative thinker. Ovservant of people and things, he had a deep physical insight into the workings of nature around him. As a child, watching his father's diesel generator, he noticed how a moving belt ratains its shape when struck; and recently he was to be seen studying the spider's web pattern of cracks in broken windows, while he pondered on the limitations of the present theory of elastic plates.

Through a contact with Professor Mott (now Sir Nevill) he went early to the University of Bristol and obtained a first in physics there in 1937. During the second World War he served first at the Admiralty, degaussing ships, and then in the technical branch of the Royal Air Force, where he reached the rank of squadron leader. He flew sometimes in Sunderlands out of Pembroke Dock, and there is in the Science Museum some radar equipment that he helped to design.

He returned to Bristol in 1946, at an exciting time for solid state physics when rapid advances were made in the theory of the deformation of crystals. The opportunity arose for him to take up theoretical research, and here he made his initial mark in dislocation theory, revealing quite suddently to those around him a mastery of some of the most difficult problems of the time. he obtained his Ph.D. in 1950 and two years later spent a year at the University of Illinois.

There followed some ten years at the University of Birmingham, a period in 1963 as visiting prefessor at the Technische Hochschule, Sturgar, and then two years at Cambridge, where he became a Fellow and College Lecture at Churchhill. In 1966 he went to the University of Sheffield, holding a readership and, from 1971, a personal chair in the theory of materials.

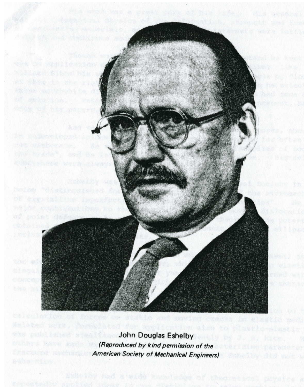


Figure 3.4. Illustration of Eshelby's equivalent eigenstrain principle

His work was a great part of his life. His general field was the theoretical physics of the deformation, strength and fracture of engineering materials, and his principal interests were lattice defects and continuum mechanics.

Though motivated by the desire to understand he kept a firm eye on application and had no time for useless erudition, like willard Gibbs his object was to make things appear simple by "looking at them in the right way". With a keen discrimination he selected those worthwhile difficult problems whcich nevertheless had some chance of solution. Entirely unconcerned with personal advancement, he hoped only of his paper that each would be a "little gem".

And so it is. Many indeed are treasure houses, abounding in undeveloped asides on which others may later build, for often he did not elaborate. He regarded himself as a modest "supplier of tools for the trade", and he felt to others their day to day use. His colleagues everywhere were always consulting him.

Eshelby was elected a Fellow of the Toyal Society in 1974, being "distinguished for his theoretical studies of the micromechanics of crystalline imperfections and material inhomogeneities". he made major contributions to the theory of static and moving dispocations and of point defects. By an elegant use of the theory of the potential he obtained some remarkable results on the elastic fields of ellipsoidal inclusions and inhomogeneities.

In 1951 he introduced, in analogy with the Maxwell tensor, the elastic energy momentum tensor, which yields forces on elastic singularities. During his later years he was much concerned with this concept and its developments, which can provide parameters characterizing the singular fields.

In 1968 he published accounts of its application to the calculation of forces on static and moving cracks inelastic media. Related work, formulated for application also to plastic-elastic media, was published simultaneously and independently by J.R.Rice. Many others have made widespread use of these characterizing parameters in fracture mechanics, sometimes in a way to which Eshelby did not wholly subscribe.

Eshelby had a wide knowledge of theoretical physics and repeatedly applied ideas in one discipline to solve problems in another. He drew much inspiration from masters of the past and liked to regard some of his most important works as amusing applications of the theorem of Gauss.

But his scholarly interests went far beyond science. He read French, German and Russian and could find his way about a Chinese dictionary; indeed, he knew a great deal about languages and the ancient world and enjoyed holding his own in discussions with professionals in these fields. His dry jokes and sayings will long be remembered:

"It's obvious", he would say,"I forget exactly why". One of his great pleasures was to find good secondhand books.

Just before his death he was in correspondence with former colleagues about some implications of recent calculations he had made of forces on defects in liquid crystals; and also about cracks in metal fatigue. He was also preparing lectures to be given in California in the new year.

3.9 Exercises

PROBELM 3.1 *Let*

$$w(\mathbf{x}) = \frac{1}{\pi R^3} \exp(-\frac{\mathbf{x} \cdot \mathbf{x}}{R^2}), \tag{3.55}$$

representing a Gaussian distribution .

For any smooth vector field, $\mathbf{A} \in \mathbb{R}^{\mathbf{3}}$*, define weighted average operation,*

$$< \mathbf{A} > (\mathbf{x}) := \int_{\mathbb{R}^3} w(\mathbf{x} - \mathbf{x}')\mathbf{A}(\mathbf{x}')d\Omega_{\mathbf{x}'} \tag{3.56}$$

where $d\Omega_{\mathbf{x}'} := dx_1' dx_2' dx_3'$*.*

Show that

$$\nabla \cdot < \mathbf{A} > = < \nabla \cdot \mathbf{A} > \tag{3.57}$$

(Hint: Use Gauss theorem (divergence theorem), and the fact that $w(\mathbf{x}) \to 0$ *as* $|\mathbf{x}| \to \infty$*.)*

PROBELM 3.2 *Use identidy*

$$e_{ijk}e_{rst} = \begin{vmatrix} \delta_{ir} & \delta_{is} & \delta_{it} \\ \delta_{jr} & \delta_{js} & \delta_{jt} \\ \delta_{kr} & \delta_{ks} & \delta_{kt} \end{vmatrix} \tag{3.58}$$

show:

$$e_{ijk}e_{ijk} = 3! = 6 \tag{3.59}$$
$$e_{ijk}e_{ij\ell} = 2\delta_{k\ell} \tag{3.60}$$
$$e_{ijk}e_{i\ell m} = \delta_{j\ell}\delta_{km} - \delta_{jm}\delta_{k\ell} \tag{3.61}$$

PROBELM 3.3 *Prove*

$$\mathbf{S}^{\Omega} + \mathbf{D} : \mathbf{T}^{\Omega} : \mathbf{C} = \mathbf{1}^{(4)} \tag{3.62}$$
$$\mathbf{T}^{\Omega} + \mathbf{C} : \mathbf{S}^{\Omega} : \mathbf{D} = \mathbf{1}^{(4)} \tag{3.63}$$

where $\mathbf{S}^{\Omega}$ *and* $\mathbf{T}^{\Omega}$ *are the Eshelby tensor and the conjugate Eshelby tensor respectively.*

Hint: First show that

$$\boldsymbol{\sigma}^d = \mathbf{C} : (\boldsymbol{\epsilon}^d - \boldsymbol{\sigma}^*)\ , \ \text{ and } \ \boldsymbol{\sigma}^* + \mathbf{C} : \boldsymbol{\epsilon}^* = 0\ . \tag{3.64}$$

PROBELM 3.4 *Consider eigenstress homogenization problem illustrated in Fig. (3.2). Suppose that the disturbance stress field,* $\boldsymbol{\sigma}^d$*, can be related to the eigenstress field,* $\boldsymbol{\sigma}^*$*, i.e.*

$$\boldsymbol{\sigma}^d = \mathbf{T}^{\Omega} : \boldsymbol{\sigma}^*, \ \ \forall \mathbf{x} \in \Omega \tag{3.65}$$

where $\mathbf{T}^{\Omega}$ *is the so-called conjugate Eshelby tensor. Show that the effective elastic tensor is equal to*

$$\bar{\mathbf{C}} = \Big[\mathbf{1}^{(4s)} + f(\mathbf{B}^{\Omega} - \mathbf{T}^{\Omega})^{-1}\Big] : \mathbf{C} \tag{3.66}$$

where the tensor, $\mathbf{B}^{\Omega} := (\mathbf{D} - \mathbf{D}^{\Omega})^{-1} : \mathbf{D}$.

PROBELM 3.5 *Suppose that an RVE (V) is subjected the following pure traction boundary condition,*

$$\mathbf{n} \cdot \boldsymbol{\sigma} = \bar{\mathbf{t}} = \mathbf{n} \cdot \boldsymbol{\sigma}^0, \ \ \forall \mathbf{x} \in \partial V \tag{3.67}$$

Show that

$$< \boldsymbol{\sigma} : \delta\boldsymbol{\epsilon} >= \boldsymbol{\sigma}^0 :< \delta\boldsymbol{\epsilon} > \ \ . \tag{3.68}$$

Chapter 4

GREEN'S FUNCTION AND FOURIER TRANSFORM

To this end, the key problem of micro-elasticity is to find the relationship between disturbance strain and eigenstrain (transformation strain). In specific,

$$\text{Find } \mathbf{S}^\Omega \text{ such that } \boldsymbol{\epsilon}^d = \mathbf{S}^\Omega : \boldsymbol{\epsilon}^* \tag{4.1}$$

or to find the conjugate Eshelby tensor,

$$\text{Find } \mathbf{T}^\Omega \text{ such that } \boldsymbol{\sigma}^d = \mathbf{T}^\Omega : \boldsymbol{\sigma}^* \tag{4.2}$$

A systematic and elegant procedure to derive $\mathbf{S}^\Omega$ and $\mathbf{T}^\Omega$ was established by Jock Eshelby, which is one of the most important contribution in classical elasticity in the twentieth century.

To understand Eshelby's inclusion/eigenstrain theory, we first review basic theory of Green's function and Fourier transform.

4.1 Green's Function

Suppose $\mathbf{L}$ is a general differential operator, i.e.

$$\begin{aligned} L[u] &= f(\mathbf{x}), \ \ \forall \mathbf{x} \in \Omega && (4.3)\\ B[u] &= h(\mathbf{x}), \ \ \forall \mathbf{x} \in \partial\Omega && (4.4) \end{aligned}$$

Suppose the above boundary value problem (BVP) is well posed. Choose $f(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{y})$ (Dirac's delta function). Then, the solution of BVP (4.3)-(4.4) is called Green's function, and it is denoted as $G(\mathbf{x}, \mathbf{y})$, i.e.

$$\begin{aligned} L[G(\mathbf{x}, \mathbf{y})] &= \delta(\mathbf{x} - \mathbf{y}), \ \ \forall \mathbf{x} \in \Omega && (4.5)\\ B[G(\mathbf{x}, \mathbf{y})] &= h(\mathbf{x}), \ \ \forall \mathbf{x} \in \partial\Omega && (4.6) \end{aligned}$$

Why are we interested in Green's function, why are we so fond of Green's function? What makes it so special?

To answer this question, we first consider a differential operator, L. Suppose that there exists an inverse operator to L, and it is denoted as L^{-1}, such that,

$$LL^{-1} = L^{-1}L = \mathbf{I} \tag{4.7}$$

The simplest differential operator is,

$$L = \frac{d}{dx}\Big(\cdot\Big) \;\; \Leftrightarrow \;\; L^{-1} = \int(\cdot)dx \tag{4.8}$$

For general differential operator L, its inverse operator may be written as

$$L^{-1}(\cdot) = \int \mathcal{K}(\mathbf{x}-\mathbf{y})(\cdot)d\mathbf{y}$$

where $\mathcal{K}$ is the so-called kernel function. Once the kernel function is determined, the inverse operator L^{-1} is determined.

Suppose that we have already known the inverse operator of L in Eqs.(4.3) and (4.4). We then can solve the differential equation by applying the inverse operation,

$$\begin{aligned} L^{-1}L[u] &= L^{-1}(f(\mathbf{x})) \\ u(\mathbf{x}) &= L^{-1}(f(\mathbf{x})) = \int \mathcal{K}(\mathbf{x}-\mathbf{y})f(\mathbf{y})d\mathbf{y} \end{aligned} \tag{4.9}$$

Equation (4.9) is usually termed as "the superposition principle".

Next question: what is the kernel function? Or how to find the kernel function for a differential operator L?

Since

$$\begin{aligned} u(\mathbf{x}) &= \mathbf{I}u(\mathbf{x}) = LL^{-1}(u(\mathbf{x})) = L\int \mathcal{K}(\mathbf{x}-\mathbf{y})u(\mathbf{y})d\mathbf{y} \\ &= \int L\mathcal{K}(\mathbf{x}-\mathbf{y})u(\mathbf{y})d\mathbf{y} \end{aligned} \tag{4.10}$$

Comparing (4.10) with

$$u(\mathbf{x}) = \int \delta(\mathbf{x}-\mathbf{y})u(\mathbf{y})d\mathbf{y}$$

one may find that $L\mathcal{K}(\mathbf{x}-\mathbf{y}) = \delta(\mathbf{x}-\mathbf{y})$. Therefore, one can deduce that the kernel function of a differential operator L is its Green's function:

$$\mathcal{K}(\mathbf{x}-\mathbf{y}) = G(\mathbf{x}-\mathbf{y}) \tag{4.11}$$

In principle, if the Green's function of a BVP has been found, the BVP is considered to be solved. This is becaruse one can obtain the general solution of the differential equation $L[u] = f(\mathbf{x})$ via superposition through certain reciprocal formula.

EXAMPLE 4.1 *We consider Euler-Bernoulli beam equation with clamped boundary conditions*

$$\begin{aligned} L[u] &= \frac{d^2}{dx^2}\Big(EI\frac{d^2u}{dx^2}\Big) = f(x),\ \forall x \in (0,l) &(4.12)\\ u(0) &= u(l) = 0,\ \text{ and } u^{'}(0) = u^{'}(l) = 0 &(4.13)\end{aligned}$$

Suppose that we have found the Green's function related to this problem, i.e.

$$\begin{aligned} L[G] &= \frac{d^2}{dx^2}\Big(EI\frac{d^2G(x,y)}{dx^2}\Big) = \delta(x-y),\ \forall x,y \in (0,l) &(4.14)\\ G(0,y) &= G(l,y) = 0,\ \text{ and } G^{'}(0,y) = G^{'}(l,y) = 0 &(4.15)\end{aligned}$$

Via integration by parts, one can show that

$$\begin{aligned} \int_0^l u\Big(\frac{d^2}{dx^2}EI\frac{d^2v}{dx^2}\Big)dx &= \Big[u\Big(\frac{d}{dx}EI\frac{d^2v}{dx^2}\Big)\Big]_0^l - \Big[\Big(\frac{du}{dx}\Big)\Big(EI\frac{d^2v}{dx^2}\Big)\Big]_0^l\\ &+ \int_0^l \Big(\frac{d^2u}{dx^2}\Big)EI\Big(\frac{d^2v}{dx^2}\Big)dx &(4.16)\end{aligned}$$

Let $v = G(x,y)$. We will have the following reciprocal formula

$$\begin{aligned} &\int_0^l uL[G]dx - \int_0^l GL[u]dx\\ =\ &\Big[u\Big(\frac{d}{dx}EI\frac{d^2G}{dx^2}\Big)\Big]_0^l - \Big[\Big(\frac{du}{dx}\Big)\Big(EI\frac{d^2G}{dx^2}\Big)\Big]_0^l\\ &-\Big[G\Big(\frac{d}{dx}EI\frac{d^2u}{dx^2}\Big)\Big]_0^l + \Big[\Big(\frac{dG}{dx}\Big)\Big(EI\frac{d^2u}{dx^2}\Big)\Big]_0^l &(4.17)\end{aligned}$$

Consider the fact that both $u(x)$ and $G(x,y)$ satisfy the same homogeneous essential boundary conditions. A simple reciprocal holds

$$\int_0^l uL(G)dx = \int_0^l GL(u)dx \tag{4.18}$$

which leads to

$$\int_0^l u(y)\delta(x-y)dy = \int_0^l G(x-y)f(y)dy \tag{4.19}$$

and consequently,

$$u(x) = \int_0^l G(x-y)f(y)dy \tag{4.20}$$

In structural engineering, the Green's function solution represents the concentrated load solution, and the Green's function is called the influence funtion. Eq.(4.20) is obtained as an argument of superposition.

EXAMPLE 4.2 *In the second example, we consider Poisson's equation,*

$$\nabla^2 u = f_1(\mathbf{x}), \text{ and } \nabla^2 v = f_2(\mathbf{x}), \ \forall \mathbf{x} \in \Omega \tag{4.21}$$

One can derive the following identity via integration by parts,

$$\begin{aligned}\int_\Omega u\nabla\cdot(\nabla v)d\Omega &= \int_\Omega \{\nabla\cdot(u\nabla v) - (\nabla u)\cdot(\nabla v)\}d\Omega \\ &= \int_{\partial\Omega}\Big(\frac{\partial v}{\partial n}\Big)udS - \int_\Omega (\nabla u)\cdot(\nabla v)d\Omega \end{aligned}\tag{4.22}$$

Interchange the position of u and v,

$$\int_\Omega v\nabla\cdot(\nabla u)d\Omega = \int_{\partial\Omega}\Big(\frac{\partial u}{\partial n}vdS - \int_\Omega (\nabla v)\cdot(\nabla u)d\Omega \tag{4.23}$$

Subsraction of (4.22) from (4.23) yields the so-called Green's reciprocal theorem:

$$\int_\Omega \Big(u\nabla^2 v - v\nabla^2 u\Big)d\Omega = \int_{\partial\Omega}\Big\{\Big(u\frac{\partial v}{\partial n}\Big) - \Big(v\frac{\partial u}{\partial n}\Big)\Big\}dS \tag{4.24}$$

Let $v(\mathbf{x}) = G(\mathbf{x},\mathbf{y})$, $f_1(\mathbf{x}) = f(\mathbf{x})$, and $f_2(\mathbf{x}) = \delta(\mathbf{x}-\mathbf{y})$. We can then show that

$$u(\mathbf{x}) = \int_{\partial\Omega}\Big\{\Big(\frac{\partial G}{\partial n}\Big)u - G\Big(\frac{\partial u}{\partial n}\Big)\Big\}dS_{\mathbf{y}} + \int_\Omega G(\mathbf{x},\mathbf{y})f(\mathbf{y})d\Omega_{\mathbf{y}} \tag{4.25}$$

Note that in 4.25, the Green's function solution does not necessarily have the same boundary data as unknown function, $u(\mathbf{x})$, as in the previous example. Often times, the Green's function in the infinite domain is chosen in a reciprocal representation.

4.2 Fourier transform

Consider a function, $f(x) \in L^1(\mathbb{R})$, or $\displaystyle\int_{-\infty}^{\infty}|f(x)|dx < \infty$. We define the Fourier transform as

$$\bar{f}(\xi) = \mathcal{F}[f] = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)exp(-i\xi x)dx \tag{4.26}$$

$$f(x) = \mathcal{F}^{-1}[\bar{f}] = \int_{-\infty}^{\infty}\bar{f}(\xi)exp(i\xi x)d\xi \tag{4.27}$$

In generalized Fourier transform, ξ is a complex number. Assume that function $f(x)$ has the property such that $exp(C_1 x)|f(x)| \to 0$ as $x \to \infty$ and $exp(-C_2 x)|f(x)| \to 0$ as $x \to -\infty$. The inversion foumula may be expressed as the following contour integral

$$f(x) = \int_{-\infty - i\gamma}^{\infty - i\gamma} \bar{f}(\xi) exp(i\xi x) d\xi \tag{4.28}$$

where $C_1 > \gamma > C_2$. The integration contour is usually referred as the Bromwich contour(Thomas John l'Anson Bromwich (1875-1929)).

LEMMA 4.3 (JORDAN) *Suppose that on the circular arc C_R shown in Fig.(4.2) we have $f(\xi) \to 0$ uniformly as $R \to \infty$. Then*

$$\lim_{R\to\infty} exp(ix\xi) f(\xi) d\xi = 0, \quad (x > 0)$$

We note that if $x < 0$ similar result holds for the contour in lower half space.

THEOREM 4.4 (CAUCHY-GOUSAT) *if $f(z)$ is an analytical function at each point within and on a closed contour C, then*

$$\oint_C f(z) dz = 0 \tag{4.29}$$

THEOREM 4.5 (CAUCHY'S RESIDUE THEOREM) *if $f(z)$ is analytical inside a closed contour C (taken in the positive sense) except at points, $z_1, z_2, \cdots, z_n$, where $f(z)$ has singularities, then*

$$\oint_C f(z) dz = 2\pi i \sum_{j=1}^{n} \text{Residue of } f(z) \; at \; z_j \tag{4.30}$$

Now, the question becomes what is a residue and how to calculate it. The answer involves with the singularity of $f(z)$. For a function of complex varible, $f(z)$, one may express $f(z)$ in a local region by its Laurent expansion – an extension of Taylor expansion of real variable. For instance around a fixed point z_j, we may write

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_j)^n + \sum_{n=1}^{\infty} a_{-n} (z - z_j)^{-n}, \quad 0 < |z - z_j| < a \tag{4.31}$$

The residue is defined as the coefficient a_{-1}.

There are three types of singularities:(1) essential singularity, (2) removable singularity, and (3)pole of order n.

• The essential singularity refers to a singularity, or pole of infinity order. For instance, for the pole $z = 0$,

$$\cos\left(\frac{1}{z}\right) = 1 - \frac{1}{2!z^2} + \frac{1}{4!z^4} - \frac{1}{6!z^6} + \cdots$$

$z = 0$ is an essential singularity.

• The removable singularity is an unsubstantial singularity, i.e. the alleged singularity disappears in Laurent expansion. For instance, at $z = 0$,

$$f(z) = \frac{\sin z}{z} = 1 - \frac{z^2}{3!} + \frac{z^4}{5!} - \frac{z^6}{7!} + \cdots$$

• Pole of order n: Consider the function,

$$f(z) = \frac{1}{z+1} + \frac{1}{(z-1)^3}$$

This function has two singularities at $z = -1$ and $z = 1$. For singularity at $z = -1$, its order is one, and it is called a pole of order one. For singularity at $z = 1$, its order is three, and it is called a pole of order 3.

The formula to calculate the residue for a pole, z_j, of order n is

$$\text{Residue at } (z = z_j) = \frac{1}{(n-1)!} \lim_{z \to z_j} \frac{d^{n-1}}{dz^{n-1}} \Big[(z - z_j)^n f(z) \Big] \tag{4.32}$$

We call the pole of order one as *simple pole*. For simple pole,

$$\text{Residue of a simple pole at } (z = z_j) = \lim_{z \to z_j} (z - z_j) f(z) \tag{4.33}$$

If $f(z) = p(z)/q(z)$, one may also write

$$\text{Residue of a simple pole at } (z = z_j) = \frac{p(z_j)}{q'(z_j)} \tag{4.34}$$

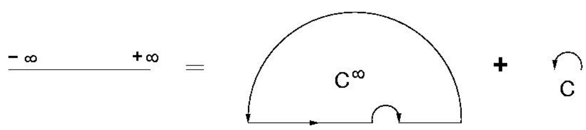


Figure 4.1. Contour integral and the count of residue

EXAMPLE 4.6 *In this example, we apply Cauchy's residue theorem to evaluate the following line integral.*

$$\int_{-\infty}^{\infty} \frac{\exp(ikt)}{(t-x)(t-ia)} dt$$

where $k > 0$ and $a > 0$.

Since $k > 0$, based on Jordan's lemma, we can use the following contour integral to replace the line integral,

$$\int_{-\infty}^{\infty} \frac{\exp(ikt)}{(t-x)(t-ia)} dt = \int_{C^\infty} \frac{\exp(ikt)}{(t-x)(t-ia)} dt + \int_C \frac{\exp(ikt)}{(t-x)(t-ia)} dt$$

where the contour integral is a half circle. Thus,

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{\exp(ikt)}{(t-x)(t-ia)} dt &= 2\pi i \; Residue(f(ia)) + \pi i \; Residue(f(x)) \\ &= -2\pi i \frac{\exp(-ka)(x+ia)}{x^2+a^2} + \pi i \frac{\exp(ikx)(x+ia)}{x^2+a^2} \end{aligned} \tag{4.35}$$

The simple pole at x is only counted for half of the residue is because that it has only half circle.

THEOREM 4.7 (CAUCHY'S INTEGRAL FORMULA) *Let $f(z)$ be analytical interior to and on a simple closed contour C. Then at any interior point z*

$$f(z) = \frac{1}{2\pi i} \oint_C \frac{f(\zeta)}{\zeta - z} d\zeta \tag{4.36}$$

THEOREM 4.8 (CONVOLUTION) *If $f(x), g(x) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, the following identity holds*

$$\int_{-\infty}^{\infty} \bar{f}(\xi)\bar{g}(\xi) \exp(i\xi x) d\xi = \frac{1}{2\pi} \int_{-\infty}^{\infty} g(x-y) f(y) dy \tag{4.37}$$

Proof:
by definition,

$$\begin{aligned} \int_{-\infty}^{\infty} \bar{f}(\xi)\bar{g}(\xi) \exp(i\xi x) d\xi &= \int_{-\infty}^{\infty} \Big[\bar{g}(\xi)\Big(\frac{1}{2\pi}\int_{-\infty}^{\infty} f(y)\exp(-i\xi y)dy\Big)\Big] \exp(i\xi x) d\xi \\ &= \int_{-\infty}^{\infty} f(y)\Big[\frac{1}{2\pi}\int_{-\infty}^{\infty} \bar{g}(\xi)\exp(i\xi(x-y))d\xi\Big] dy \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} g(x-y) f(y) dy \end{aligned} \tag{4.38}$$

In 3D, we have

$$\int_{-\infty}^{\infty} \bar{f}(\boldsymbol{\xi})\bar{g}(\boldsymbol{\xi}) \exp(i\boldsymbol{\xi}\cdot\mathbf{x}) = \frac{1}{(2\pi^3)} \int_{-\infty}^{\infty} g(\mathbf{x}-\mathbf{y}) f(\mathbf{y}) d\mathbf{y} \tag{4.39}$$

EXAMPLE 4.9 *Consider Heaviside function,*

$$H(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases} \tag{4.40}$$

Note that at x=0 Heaviside function is not defined.

To find the Fourier transform of the Heaviside function,

$$\begin{aligned} \bar{H}(\xi) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} H(x)\exp(-i\xi x)dx \\ &= \frac{1}{2\xi}\int_{0}^{\infty}\exp(-i\xi x)dx = \frac{1}{2\pi}\frac{(-1)}{i\xi}\exp(-i\xi x)\Big|_0^{\infty} \\ &= \frac{1}{2\pi i\xi} \end{aligned} \tag{4.41}$$

The result implies that $\exp(-i\xi\infty) \to 0$*, which requires that* $Im(\xi) < 0$*. Lighthill showed that in the sense of generalized function,*

$$\bar{H}(\xi) = \exp\Big(-\frac{\pi i}{2}\mathrm{sgn}(\xi)\Big)\frac{1}{2\pi|\xi|}, \ \text{ where } \mathrm{sgn}\xi := \begin{cases} 1 & \xi > 0 \\ -1 & \xi < 0 \end{cases} \tag{4.42}$$

Note that $H(x) \notin L^1(\mathbb{R})$*. Therefore, Fourier transform of Heaviside function does not really exit for* $f \in L^1$*.* $\displaystyle\int_{-\infty}^{\infty} |f(x)| < \infty$ *is a very stringent condition. It is why many functions that has Laplace transform do not possess Fourier transform, which is the reason why sometimes we use Laplace transform instead of Fourier transform. By the way, if* ξ *is taken as a complex number, Fourier transform is equivalent to bilateral Laplace transform.*

EXAMPLE 4.10 *To find the Fourier transform of the Dirac's delta function,*

$$\bar{\delta} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\delta(x)\exp(-ix\xi)dx = \frac{1}{2\pi} \tag{4.43}$$

Inversely,

$$\delta(x) = \int_{-\infty}^{\infty}\bar{\delta}(\xi)\exp(i\xi x)d\xi = \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(i\xi x)d\xi \tag{4.44}$$

EXAMPLE 4.11 *On the other hand, consider the inversion formula,*

$$\int_{-\infty}^{\infty}\delta(\xi)\exp(i\xi x)d\xi = \exp(i0x) = 1, \quad \Rightarrow \quad \bar{1}(\xi) = \delta(\xi) \tag{4.45}$$

Hence

$$\bar{1}(\xi) = \delta(\xi) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\exp(-i\xi x)dx \tag{4.46}$$

In three-dimensional space, we have the identity,

$$\delta(\boldsymbol{\xi}) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \exp(-i\boldsymbol{\xi} \cdot \mathbf{x}) d\mathbf{x} \tag{4.47}$$

Combining (4.44) and (4.47), one may draw conclusion that

$$\delta(\boldsymbol{\xi}) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} \cos(\boldsymbol{\xi} \cdot \mathbf{x}) d\mathbf{x} \tag{4.48}$$

EXAMPLE 4.12 *The Fourier transform of $f(x)$ is*

$$\bar{f}(\xi) = \frac{1}{2\pi} \frac{ia}{\xi(\xi^2 - ia\xi - a)} \tag{4.49}$$

Find $f(x)$?

$\bar{f}(\xi)$ *has three poles in the complex plane:*

$$\xi_1 = 0, \; and \; \xi_{2,3} = \frac{ia}{2} \pm \sqrt{a - \frac{a^2}{4}} = \frac{ia}{2} \pm \beta, \; \beta := \sqrt{a - a^2/4} \tag{4.50}$$

Therefore,

$$\begin{aligned}
f(x) &= \int_{-\infty - i\gamma}^{\infty - i\gamma} \bar{f}(\xi) \exp(i\xi x) d\xi \\
&= \oint_C \frac{1}{2\pi} \frac{ia \exp(i\xi x)}{(\xi - 0)(\xi - \xi_2)(\xi - \xi_3)} d\xi \\
&= \pi i \text{Residue of } \xi \; at \; \xi_1 + 2\pi i \sum_{j=2}^{3} \text{Residue of } \xi \; at \; \xi_j \\
&= ia \Big\{ \frac{\exp(i\xi_1 x)}{\xi_2 \xi_3} + \frac{\exp(i\xi_2 x)}{\xi_2(\xi_2 - \xi_3)} + \frac{\exp(i\xi_3 x)}{\xi_3(\xi_3 - \xi_2)} \Big\} \\
&= (-a) \Big\{ \frac{1}{-a} + \frac{\exp[ix(\frac{ia}{2} + \beta)]}{\Big(\frac{ia}{2} + \beta\Big) 2\sqrt{a - \frac{a^2}{4}}} - \frac{\exp[ix(\frac{ia}{2} - \beta)]}{\Big(\frac{ia}{2} - \beta\Big) 2\sqrt{a - \frac{a^2}{4}}} \\
&= 1 - \frac{\exp\Big(-\frac{xa}{2}\Big)}{2\sqrt{a - \frac{a^2}{4}}} \Big\{ \Big(\beta - \frac{ia}{2}\Big) \exp(i\beta x) + \Big(\beta + \frac{ia}{2}\Big) \exp(-i\beta x) \Big\} \\
&= 1 - \frac{\exp\Big(-\frac{xa}{2}\Big)}{2\sqrt{a - \frac{a^2}{4}}} \Big(2\beta \cos x + a \sin \beta x\Big)
\end{aligned} \tag{4.51}$$

4.3 Examples of Green's Function

EXAMPLE 4.13 *Find the Green's function of two-dimensional Poission's equation in infinite domain,*

$$\nabla^2 G(\mathbf{x},\mathbf{y}) + \delta(|\mathbf{x}-\mathbf{y}|) = 0, \ \ \forall \mathbf{x} \in \mathbb{R}^2 \tag{4.52}$$

Use the polar coordinate $\nabla^2 = \frac{1}{r}\frac{d}{dr}\Big(r\frac{d}{dr}\Big)$ *and denote* $\mathbf{x}^{'} = \mathbf{x}-\mathbf{y}$. *We have*

$$\frac{1}{r}\frac{d}{dr}\Big(r\frac{d}{dr}G\Big) = -\delta(x_1^{'})\delta(x_2^{'}) \tag{4.53}$$

and

$$\int_0^{2\pi}\int_0^{r} \frac{1}{r^{'}}\frac{d}{dr^{'}}\Big(r^{'}\frac{d}{dr^{'}}G\Big)r^{'}dr^{'}d\theta = -\int_\Omega \delta(dx_1^{'})\delta(dx_2^{'})dx_1^{'}dx_2^{'} \tag{4.54}$$

where $r^{'}$ *is the dummy variable and* $r = |\mathbf{x}-\mathbf{y}| = \sqrt{(x_1-y_1)^2+(x_2-y_2)^2}$. *The integration domain is a circular region centered at* $\mathbf{x}=\mathbf{y}$ *and with the radius r.*

Therefore,

$$2\pi\Big(r\frac{d}{dr}G\Big) = -1, \quad \Rightarrow \frac{d}{dr}G = -\frac{1}{2\pi}\frac{1}{r} \tag{4.55}$$

Finally, we find that

$$G(\mathbf{x}-\mathbf{y}) = -\frac{1}{2\pi}lnr \tag{4.56}$$

EXAMPLE 4.14 *Consider one dimensional Helmhotz equation,*

$$\frac{d^2u}{dx^2} + k^2u = \delta(|x-y|) \tag{4.57}$$

Apply Fourier transform,

$$\begin{aligned}
\bar{u}(\xi) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} u(x)\exp(-i\xi x)dx \\
\mathcal{F}\Big(\frac{d^2u}{dx^2}\Big) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{d^2u}{dx^2}\exp(-i\xi x)dx = -\xi^2\bar{u}(\xi) \\
\bar{\delta}(|x-y|) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\delta(x-y)\exp(i\xi x)d\xi = \frac{1}{2\pi}\exp(-i\xi y)
\end{aligned} \tag{4.58}$$

and

$$\bar{u}(\xi) = \frac{1}{2\pi}\frac{1}{k^2-\xi^2}\exp(-i\xi y) \tag{4.59}$$

Therefore,

$$
\begin{aligned}
u(x) &= \int_{-\infty}^{\infty} \bar{u}(\xi)\exp(i\xi x)d\xi \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1}{(k+\xi)(k-\xi)}\exp(i\xi(x-y))d\xi \\
&= \frac{1}{2\pi}\oint_R \frac{\exp(i\xi(x-y))}{(k+\xi_)k-\xi)}d\xi = \frac{i}{2}\sum_{i=1}^{2}\text{Residues of } \xi\ at\ \xi_i \\
&= -\Big(\frac{i}{2}\Big)\Big\{\frac{1}{2k}\exp(ik(x-y)) - \frac{1}{2k}\exp(-ik(x-y))\Big\} \\
&= -\frac{i}{4k}\Big\{\Big(\cos k(x-y) + i\sin k(x-y)\Big) - \Big(\cos k(x-y) - i\sin k(x-y)\Big)\Big\} \\
&= -\frac{i}{4k}\Big(2i\sin k(x-y)\Big) = \frac{1}{k}\sin k(x-y) \qquad (4.60)
\end{aligned}
$$

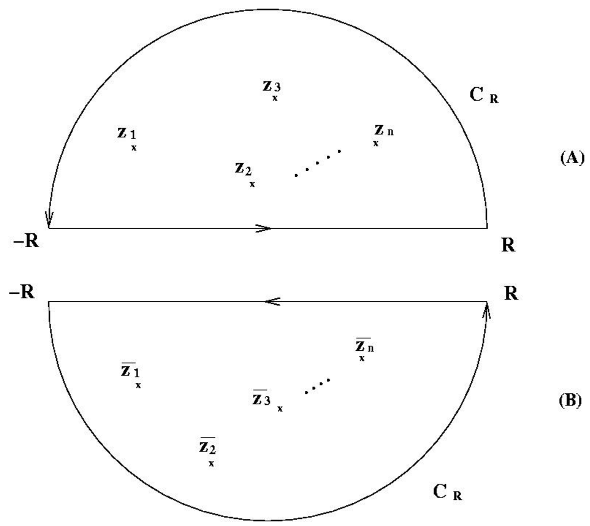


Figure 4.2. Inversion paths of Fourier transform

EXAMPLE 4.15 *Find Green's function for three-dimensional Poisson's equation,*

$$\nabla^2 G + \delta(\mathbf{x}-\mathbf{x}') = 0 \;\Rightarrow\; G_{,ii} + \delta(\mathbf{x}-\mathbf{x}^{'}) = 0 \qquad (4.61)$$

where $\nabla^2 = \dfrac{\partial^2}{\partial x_i \partial x_i}, i = 1,2,3$ *and* $\delta(\mathbf{x}-\mathbf{x}') = \delta(x_1 - x_1')\delta(x_2 - x_2')\delta(x_3 - x_3')$

Consider the fact that

$$\bar{\delta}(\mathbf{x}-\mathbf{x}') = \frac{1}{\pi^3}\exp\Big(-i\boldsymbol{\xi}\cdot\mathbf{x}'\Big) \Rightarrow \delta(\mathbf{x}-\mathbf{x}') = \frac{1}{2\pi^3}\exp\Big(i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{x}')\Big)d\boldsymbol{\xi}$$

Therefore, based on definition,

$$G(\mathbf{x}-\mathbf{x}') = \int_{-\infty}^{\infty}\bar{G}(\boldsymbol{\xi})\exp\Big(i\boldsymbol{\xi}\cdot\mathbf{x}\Big)d\boldsymbol{\xi}$$

one may derive that

$$G_{,ii}(\mathbf{x}-\mathbf{x}') = -\int_{-\infty}^{\infty}\bar{G}(\boldsymbol{\xi})\xi_i\xi_i\exp(i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{x}'))d\boldsymbol{\xi} \tag{4.62}$$

and

$$\bar{G}(\boldsymbol{\xi})\xi_i\xi_i = \frac{1}{(2\pi)^3} \quad \Rightarrow \quad \bar{G}(\boldsymbol{\xi}) = \frac{1}{(2\pi)^3}\Big(\frac{1}{\xi_i\xi_i}\Big) \tag{4.63}$$

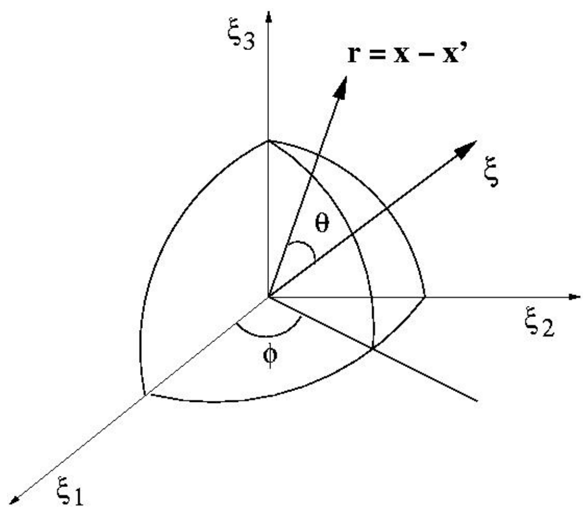


Figure 4.3. Inversion of three-dimensional Fourier transform

Let $\xi^2 := \xi_i\xi_i, \mathbf{r} = \mathbf{x} - \mathbf{x}^{'}, r := |\mathbf{x} - \mathbf{x}^{'}|$*, and* $\boldsymbol{\xi} \cdot (\mathbf{x} - \mathbf{x}^{'}) = \xi r \cos\theta$*. Then,*

$$\begin{aligned}
G(\mathbf{x} - \mathbf{x}^{'}) &= \frac{1}{(2\pi^3)} \int_{-\infty}^{\infty} \frac{1}{\xi^2} \exp(i\boldsymbol{\xi} \cdot (\mathbf{x} - \mathbf{x}^{'}))d\boldsymbol{\xi} \\
&= \frac{1}{(2\pi^3)} \int_0^{\infty} \int_0^{\pi} \int_0^{2\pi} \frac{1}{\xi^2} \exp(i\boldsymbol{\xi} \cdot (\mathbf{x} - \mathbf{x}^{'}))\xi^2 d\xi \sin\theta d\theta d\phi \\
&= \frac{1}{(2\pi^3)} \int_0^{\infty} \int_1^{-1} \int_0^{2\pi} \exp(i\xi r \cos\theta) d\xi(-d\cos\theta)d\phi \\
&= \frac{1}{(2\pi^2)} \int_0^{\infty} \int_{-1}^{1} \exp(i\xi r t) d\xi dt \\
&= \frac{1}{(2\pi^2)} \int_0^{\infty} d\xi \int_{-1}^{1} \Big[\cos(\xi r t) + i\sin(\xi r t)\Big] dt \qquad (4.64)
\end{aligned}$$

Consider the fact that

$$\int_{-1}^{1} \cos(\xi r t)dt = \frac{1}{\xi r} \sin(\xi r t)\Big|_{-1}^{1} = \frac{2\sin\xi r}{\xi r} \qquad (4.65)$$

$$\int_{-1}^{1} \sin(\xi r t)dt = 0 \qquad (4.66)$$

Hence

$$\begin{aligned}
G(\mathbf{x} - \mathbf{x}^{'}) &= \frac{1}{2\pi^2} \int_0^{\infty} \frac{\sin\xi r}{\xi r} d\xi \\
&= \frac{1}{2\pi^2 r} \int_0^{\infty} \frac{\sin\xi r}{\xi r} d(\xi r) = \frac{1}{2\pi^2 r} Si(\infty) \qquad (4.67)
\end{aligned}$$

where $Si(x) := \displaystyle\int_0^x \frac{\sin t}{t} dt$ *and* $Si(\infty) = \dfrac{\pi}{2}$*. Finally, we have*

$$G(\mathbf{x} - \mathbf{x}^{'}) = \frac{1}{4\pi} \frac{1}{|\mathbf{x} - \mathbf{x}^{'}|} \qquad (4.68)$$

4.4 Static Green's function for 3D linear elasticity

The Green's function for static, linear, isotropic elasticity was derived by Lord Kelvin (1882). The derivation shown below employs the Fourier integral transform, which is a systematic and elegant procedure to find Green's function for partial differential equations. Consider the Navier equation,

$$\sigma_{ji,j} + f_i = 0 \qquad (4.69)$$

Denote Green's function vector of the displacement field as

$$u_i^m(\mathbf{x}, \mathbf{y}) = G_{mi}^{\infty}(\mathbf{x}, \mathbf{y}) \qquad (4.70)$$

We let

$$\sigma_{ij}^{G_m^\infty} = C_{ijkl}\epsilon_{kl}^{G^\infty} \tag{4.71}$$

$$f_i^m = \delta(\mathbf{x}-\mathbf{y})\delta_{mi} \tag{4.72}$$

where $\delta(\mathbf{x}-\mathbf{y}) := \delta(x_1-y_1)\delta(x_2-y_2)\delta(x_3-y_3)$, and the integer m is a free index, which indicates the direction of the concentrated load.

Then,

$$\sigma_{ij}^{G_m^\infty} = C_{ijkl}\epsilon_{kl}^{G^\infty} = C_{ijkl}G_{mk,l}^\infty \;\rightarrow\; \sigma_{ij,j}^{G_m^\infty} = C_{ijkl}G_{mk,lj}^\infty$$

Then Green's function for an infinite linear elastic medium is the solution of the following equatin,

$$C_{ijkl}G_{mk,lj}^\infty + \delta(\mathbf{x}-\mathbf{y})\delta_{mi} = 0 \tag{4.73}$$

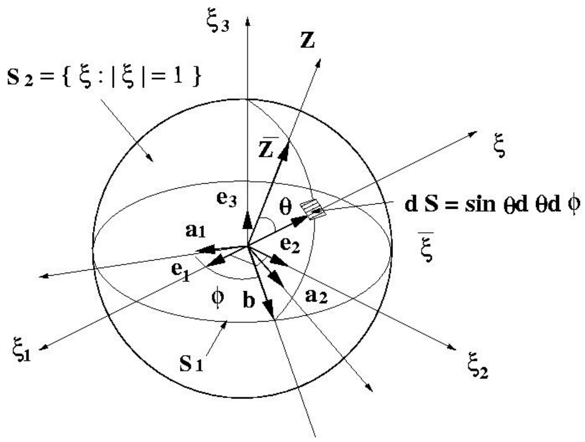


Figure 4.4. The unit sphere S^2 in the ξ-space. Green's function at point $\mathbf{z}$ is expressed by a line integral along S^1 which lies on the plane perpendicular to $\mathbf{z}$

Apply Fourier integral transform,

$$G_{mk}^\infty(\mathbf{x}-\mathbf{y}) = \int_{-\infty}^{\infty} \bar{\mathbf{G}}_{mk}^\infty(\boldsymbol{\xi})\exp(i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{y}))d\boldsymbol{\xi} \tag{4.74}$$

where $\displaystyle\int_{-\infty}^{\infty} = \int\int\int_{-\infty}^{\infty}$, and $d\boldsymbol{\xi} = d\xi_1 d\xi_2 d\xi_3$.

Consider

$$G_{mk,lj}^\infty(\mathbf{x}-\mathbf{y}) = -\int_{-\infty}^{\infty} \bar{\mathbf{G}}_{mk}^\infty(\boldsymbol{\xi})\xi_l\xi_j\exp(i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{y}))d\boldsymbol{\xi} \tag{4.75}$$

$$\delta(\mathbf{x}-\mathbf{y}) = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\exp(i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{y}))d\boldsymbol{\xi} \tag{4.76}$$

We obtain the following algebraic equations in Fourier space,

$$C_{ijkl}\bar{\mathbf{G}}^{\infty}_{mk}(\boldsymbol{\xi})\xi_l\xi_j = \frac{1}{(2\pi)^3}\delta_{im} \tag{4.77}$$

Let

$$K_{ik} = C_{ijkl}\xi_j\xi_l \;\;\Rightarrow\;\; K_{ik}\bar{\mathbf{G}}^{\infty}_{mk} = \frac{1}{(2\pi)^3}\delta_{im} \tag{4.78}$$

Consider Laplace expansion,

$$N_{ji}(\boldsymbol{\xi})K_{ik}(\boldsymbol{\xi}) = D(\boldsymbol{\xi})\delta_{jk} \tag{4.79}$$

where N_{ji} is the cofactor of K_{ji} and $D(\boldsymbol{\xi}) = det\{K_{ij}(\boldsymbol{\xi})\}$.
Multiplying (4.78) with N_{ji} yields

$$N_{ji}(\boldsymbol{\xi})K_{ik}(\boldsymbol{\xi})\bar{\mathbf{G}}^{\infty}_{mk}(\boldsymbol{\xi}) \;=\; \frac{1}{(2\pi)^3}N_{ji}(\boldsymbol{\xi})\delta_{im} \tag{4.80}$$

$$D(\boldsymbol{\xi})\delta_{jk}\bar{\mathbf{G}}^{\infty}_{mk}(\boldsymbol{\xi}) \;=\; \frac{1}{(2\pi)^3}N_{jm}(\boldsymbol{\xi}) \tag{4.81}$$

which leads to

$$\bar{\mathbf{G}}^{\infty}_{jm}(\boldsymbol{\xi}) = \frac{1}{(2\pi)^3}\frac{N_{jm}(\boldsymbol{\xi})}{D(\boldsymbol{\xi})} \tag{4.82}$$

Change indices $j \leftrightarrow i$ and $m \leftrightarrow j$. Via inverse Fourier transform, one may find that

$$G^{\infty}_{ij}(\mathbf{x}-\mathbf{y}) = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\frac{N_{ij}(\boldsymbol{\xi})}{D(\boldsymbol{\xi})}\exp(i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{y}))d\boldsymbol{\xi} \tag{4.83}$$

For linear isotropic material, one may find that

$$N_{ij}(\boldsymbol{\xi}) \;=\; \mu\xi^2\Big((\lambda+2\mu)\delta_{ij}\xi^2 - (\lambda+\mu)\xi_i\xi_j\Big) \tag{4.84}$$

$$D(\boldsymbol{\xi}) \;=\; \mu^2(\lambda+2\mu)\xi^6 \tag{4.85}$$

Let $\mathbf{z} = \mathbf{x} - \mathbf{y}$. We have

$$G^{\infty}_{ij}(\mathbf{z}) = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\frac{1}{\mu(\lambda+2\mu)\xi^4}\Big((\lambda{+}2\mu)\delta_{ij}\xi^2{-}(\lambda{+}\mu)\xi_i\xi_j\Big)\exp(i\boldsymbol{\xi}{\cdot}\mathbf{z})d\boldsymbol{\xi} \tag{4.86}$$

To integrate (4.86), we donote S^2 as a unit sphere where $|\boldsymbol{\xi}| = 1$, and denote S^1 as a unit circle on the surface of S^2, where S^2 is intersected by a plane perpendicular to vector **z**.

Apply Radon decompositon,

$$d\boldsymbol{\xi} = dV_\xi = d\xi_1 d\xi_2 d\xi_3 \;\;\Rightarrow\;\; dV_\xi = \xi^2 d\xi dS \tag{4.87}$$

where $\xi^2 = \xi_1^2 + \xi_2^2 + \xi_3^2$ and dS is the surface element on the unit sphere S^2 in ξ-space. Imagine that the ξ-space is a expanded spherical balloon.

Denote $\bar{\boldsymbol{\xi}} = \bar{\xi}_i \mathbf{e}_{\xi i}$ as a unit vector pointing from the origin to the surface of S^2 along $\boldsymbol{\xi}$ direction and denote $\bar{\mathbf{z}} = \bar{z}_i \mathbf{e}_{zi}$ as another unit vector point from the origint to the surface of S^2 along $\mathbf{z}$ direction. Therefore, $\boldsymbol{\xi} = \xi \bar{\boldsymbol{\xi}}$ and $\mathbf{z} = z\bar{\mathbf{z}}$ where $\xi = \sqrt{\xi_1^2 + \xi_2^2 + \xi_3^2}$ and $z = \sqrt{z_1^2 + z_2^2 + z_3^2}$. Obviously, $\bar{\xi}_i = \xi_i/\xi$ and $\bar{z}_i = z_i/z$.

Then Eq.(4.86) can be written as

$$\begin{aligned} G_{ij}(\mathbf{z}) &= \frac{1}{(2\pi)^3}\int_0^\infty d\xi \int_{S^2} \frac{1}{\mu(\lambda+2\mu)}\Big((\lambda+2\mu)\delta_{ij} - (\lambda+\mu)\bar{\xi}_i\bar{\xi}_j\Big) \\ &\quad \cdot \exp\{i\xi z \bar{\boldsymbol{\xi}}\cdot\bar{\mathbf{z}}\} dS(\bar{\boldsymbol{\xi}}) \end{aligned} \tag{4.88}$$

Consider the symmetry property (change $\boldsymbol{\xi} \to -\boldsymbol{\xi}$ of Eq.(4.86)). We may also have

$$\begin{aligned} G_{ij}(\mathbf{z}) &= \frac{1}{(2\pi)^3}\int_0^\infty d\xi \int_{S^2} \frac{1}{\mu(\lambda+2\mu)}\Big((\lambda+2\mu)\delta_{ij} - (\lambda+\mu)\bar{\xi}_i\bar{\xi}_j\Big) \\ &\quad \cdot \exp\{-i\xi z \bar{\boldsymbol{\xi}}\cdot\bar{\mathbf{z}}\} dS(\bar{\boldsymbol{\xi}}) \end{aligned} \tag{4.89}$$

Change the scalar $\xi \to -\xi$. Eq.(4.89) yields

$$\begin{aligned} G_{ij}(\mathbf{z}) &= \frac{1}{(2\pi)^3}\int_{-\infty}^0 d\xi \int_{S^2} \frac{1}{\mu(\lambda+2\mu)}\Big((\lambda+2\mu)\delta_{ij} - (\lambda+\mu)\bar{\xi}_i\bar{\xi}_j\Big) \\ &\quad \cdot \exp\{i\xi z \bar{\boldsymbol{\xi}}\cdot\bar{\mathbf{z}}\} dS(\bar{\boldsymbol{\xi}}) \end{aligned} \tag{4.90}$$

Combining (4.88) with (4.90) yields

$$\begin{aligned} G_{ij}(\mathbf{z}) &= \frac{1}{2(2\pi)^3}\int_{-\infty}^\infty d\xi \int_{S^2} \frac{1}{\mu(\lambda+2\mu)}\Big((\lambda+2\mu)\delta_{ij} - (\lambda+\mu)\bar{\xi}_i\bar{\xi}_j\Big) \\ &\quad \cdot \exp\{i\xi z \bar{\boldsymbol{\xi}}\cdot\bar{\mathbf{z}}\} dS(\bar{\boldsymbol{\xi}}) \end{aligned} \tag{4.91}$$

since

$$\int_{-\infty}^\infty \exp(i\xi z\bar{\boldsymbol{\xi}}\cdot\bar{\mathbf{z}})d\xi = 2\pi\delta(z\bar{\boldsymbol{\xi}}\cdot\bar{\mathbf{z}}) \tag{4.92}$$

one has

$$G_{ij}(\mathbf{z}) = \frac{1}{2(2\pi)^2}\oint_{S^2} \delta(z\bar{\boldsymbol{\xi}}\cdot\bar{\mathbf{z}}) \frac{[(\lambda+2\mu)\delta_{ij} - (\lambda+\mu)\bar{\xi}_i\bar{\xi}_j]}{\mu(\lambda+2\mu)} dS(\bar{\boldsymbol{\xi}}) \tag{4.93}$$

To integrate (4.93), one has to evaluate the following two integrals:

$$\int_{S^2} \delta(z\bar{\boldsymbol{\xi}}\cdot\bar{\mathbf{z}})dS? \ \ and \ \ \int_{S^2} \bar{\xi}_i\bar{\xi}_j \delta(z\bar{\boldsymbol{\xi}}\cdot\bar{\mathbf{z}})dS?$$

Consider $\bar{\boldsymbol{\xi}} \cdot \bar{\mathbf{z}} = \cos\theta$, $d\cos\theta = -\sin\theta d\theta$. One may decompose the surface element on S^2 into: $dS(\bar{\boldsymbol{\xi}}) = \sin\theta d\theta d\phi = -d(\bar{\boldsymbol{\xi}} \cdot \bar{\mathbf{z}})d\phi$, where $\theta \to [0, \pi]\Big(\cos\theta \to [1, -1]\Big)$ and $\phi \to [0, 2\pi]$. If we let $t = \bar{\xi} \cdot \bar{\mathbf{z}}$,

$$\int_{S^2} \delta(z\bar{\boldsymbol{\xi}} \cdot \bar{\mathbf{z}})dS = \int_{-1}^{1} \delta(zt)dt \int_0^{2\pi} d\phi = \frac{2\pi}{z} \tag{4.94}$$

On the other hand,

$$\int_{S^2} \bar{\xi}_i\bar{\xi}_j\delta(z\bar{\boldsymbol{\xi}} \cdot \bar{\mathbf{z}})dS = \int_{-1}^{1}\int_0^{2\pi} \delta(zt)\bar{\xi}_i\bar{\xi}_j dt d\phi \tag{4.95}$$

Consider the projection of vector $\bar{\boldsymbol{\xi}}$,

$$Proj_{\bar{\mathbf{z}}\|}\bar{\boldsymbol{\xi}} = \cos\theta\bar{\mathbf{z}} = \cos\theta\bar{z}_i\mathbf{e}_i \tag{4.96}$$

$$Proj_{\bar{\mathbf{z}}\perp}\bar{\boldsymbol{\xi}} = \sin\theta\mathbf{b} = \sin\theta(\cos\phi\mathbf{a}_1 + \sin\phi\mathbf{a}_2) \tag{4.97}$$

Considering,

$$\mathbf{a}_1 = (\mathbf{a}_1 \cdot \mathbf{e}_i)\mathbf{e}_i; \;\; \mathbf{a}_2 = (\mathbf{a}_2 \cdot \mathbf{e}_i)\mathbf{e}_i$$

one has

$$\begin{aligned} \bar{\boldsymbol{\xi}} &= x_i\mathbf{e}_i = \cos\theta\bar{\mathbf{z}} + \sin\theta\mathbf{b} \\ &= \cos\theta\bar{z}_i\mathbf{e}_i + \sin\theta\Big(\cos\phi a_{1i} + \sin\phi a_{2i}\Big)\mathbf{e}_i \end{aligned} \tag{4.98}$$

Thereby,

$$\begin{aligned} \bar{\xi}_i &= \cos\theta\bar{z}_i + \sin\theta(\cos\phi a_{1i} + \sin\phi a_{2i}) \\ \Rightarrow \bar{\xi}_i\bar{\xi}_j &= \Big(\cos\theta\bar{z}_i + \sin\theta(\cos\phi a_{1i} + \sin\phi a_{2i})\Big) \\ &\quad \cdot\Big(\cos\theta\bar{z}_j + \sin\theta(\cos\phi a_{1j} + \sin\phi a_{2j})\Big) \\ &= \cos^2\theta\bar{z}_i\bar{z}_j + \sin\theta\cos\theta\Big[\bar{z}_i(\cos\phi a_{1j} + \sin\phi a_{2j}) \\ &\quad +\bar{z}_j(\cos\phi a_{1i} + \sin\phi a_{2i})\Big] \\ &\quad + \sin^2\theta(\cos\phi a_{1i} + \sin\phi a_{2i})(\cos\phi a_{1j} + \sin\phi a_{2j}) \\ &= t^2\bar{z}_i\bar{z}_j + t\sqrt{1-t^2}\Big[\bar{z}_i(\cos\phi a_{1j} + \sin\phi a_{2j}) \\ &\quad +\bar{z}_j(\cos\phi a_{1i} + \sin\phi a_{2i})\Big] \\ &\quad +(1-t^2)(\cos\phi a_{1i} + \sin\phi a_{2i})(\cos\phi a_{1j} + \sin\phi a_{2j}) \end{aligned} \tag{4.99}$$

where $t = \cos\theta$.

Consider the fact that

$$\begin{aligned}\int_{-1}^{1} t^2\delta(zt)dt &= 0 \\ \int_{-1}^{1} t\sqrt{1-t^2}\delta(zt)dt &= 0 \end{aligned} \tag{4.100}$$

We have

$$\begin{aligned}\oint_{S^2} \delta(zt)\bar{\xi}_i\bar{\xi}_j dS &= \int_{-1}^{1}\delta(zt)\int_0^{2\pi}\{\cos^2\phi a_{1j}a_{1i} \\ &\quad + \cos\phi\sin\phi(a_{1j}a_{2i}+a_{1i}a_{2j}) + \sin^2\phi a_{2j}a_{2i}\}dtd\phi \\ &= \frac{\pi}{z}(a_{1i}a_{1j}+a_{2i}a_{2j}) = \frac{\pi}{z}\Big(\delta_{ij}-\bar{z}_i\bar{z}_j\Big) \end{aligned} \tag{4.101}$$

because $a_{1i}a_{1j}+a_{2i}a_{2j}+\bar{z}_i\bar{z}_j=\delta_{ij}$. Note that $\mathbf{a}_1$, $\mathbf{a}_2$, and $\bar{\mathbf{z}}$ form a triads. Let $Q_{1i}=a_{1i}$, $Q_{2i}=a_{2i}$ and $Q_{3i}=\bar{z}_i$. From $Q_{ik}Q^T_{kj}=Q_{ik}Q_{jk}=\delta_{ij}$, one derives that $a_{1i}a_{1j}+a_{2i}a_{2j}+\bar{z}_i\bar{z}_j=\delta_{ij}$.

Consequently,

$$\begin{aligned} G^\infty_{ij}(\mathbf{z}) &= \frac{1}{(2\pi)^2}\frac{1}{2z}\Big[\frac{2\pi(\lambda+2\mu)\delta_{ij}-\pi(\lambda+\mu)(\delta_{ij}-\bar{z}_i\bar{z}_j)}{\mu(\lambda+2\mu)}\Big] \\ &= \frac{1}{8\pi}\frac{1}{z}\frac{1}{\mu}\frac{(\lambda+\mu)}{(\lambda+2\mu)}\Big\{\frac{\lambda+3\mu}{\lambda+\mu}\delta_{ij}+\bar{z}_i\bar{z}_j\Big\} \\ &= \frac{1}{16\pi\mu(1-\nu)|\mathbf{x}-\mathbf{y}|}\Big\{(3-4\nu)\delta_{ij}+\frac{(x_i-y_i)(x_j-y_j)}{|\mathbf{x}-\mathbf{y}|^2}\Big\} \end{aligned} \tag{4.102}$$

4.5 Variation in a Theme: Radon Transform

Let $\mathbf{x}=(x_1,x_2,x_3)$ be the positoin vector of a spatial point in $\mathbb{R}^3$ and consider a regular function $f(\mathbf{x})$ (image density) defined in $\mathbb{R}^3$. The Radon transform of $f(\mathbf{x})$ is defined as

$$\hat{f}(s,\mathbf{n}) = \mathbf{R}\{f(\mathbf{x})\} = \int_{-\infty}^{\infty} f(\mathbf{x})\delta(s-\mathbf{n}\cdot\mathbf{x})d\mathbf{x} \tag{4.103}$$

$\hat{f}$ is the projectin of $f(\mathbf{x})$ on the plane $\mathbf{n}\cdot\mathbf{x}=s$, where $\mathbf{n}$ is a unit vector, and s is the distance from the plane to the origin of the coordinate (see Fig. (4.5)). The integral is the integration of image density, $f(\mathbf{x})$, along the plane. The collection of all $\hat{f}(s,\mathbf{n})$ for all unit vector $\mathbf{n}$ is called the Radon transform.

The inverse Radon transform is carried out by two steps:

1. $$\tilde{f}(s,\mathbf{n}) = \partial_s^2\hat{f}(s,\mathbf{n}) \tag{4.104}$$

2. $$f(\mathbf{x}) = \mathbf{R}^{-1}(\tilde{f}) = -\frac{1}{8\pi^2}\int_{|\mathbf{n}|=1}\tilde{f}(\mathbf{n}\cdot\mathbf{x},\mathbf{n})dS(\mathbf{n}) \tag{4.105}$$

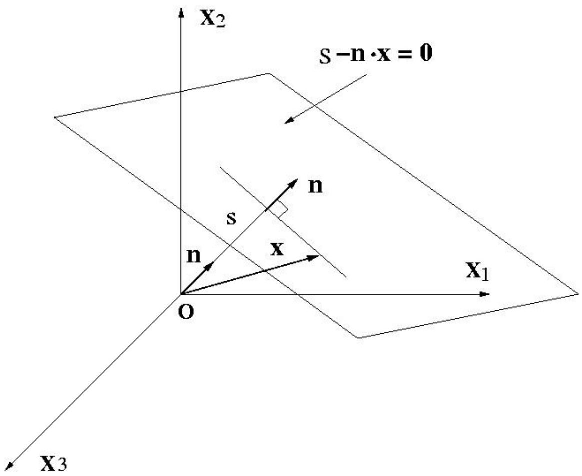


Figure 4.5. Projection plane of three-dimensional Radon transform

The Radon transform has the following properties:

1 $\hat{f}(s, \mathbf{n})$ is an even and homogeneous, of order -1, function, i.e. $\hat{f}(\alpha s, \alpha \mathbf{n}) = |\alpha|^{-}1\hat{f}(s, \mathbf{n})$;

2 linearity: $\mathbf{R}(c_1 f + c_2 g) = c_1 \hat{f} + c_2 \hat{g}$;

3 transform of derivatives:

$$\begin{aligned} \mathbf{R}(\partial_i f(\mathbf{x})) &= n_i \partial_s \hat{f}(s, \mathbf{n}) \\ \mathbf{R}(\partial_i \partial_j f(\mathbf{x})) &= n_i n_j \partial_s^2 \hat{f}(s, \mathbf{n}) \end{aligned}$$

EXAMPLE 4.16 *Consider an image density function, $g(x, y)$. The two-dimensional Radon transform may be defined as*

$$\hat{g}(\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) \delta(\rho - x \cos\theta - y \sin\theta) dx dy \tag{4.106}$$

which is identical to the following line integral

$$\hat{g}(\rho, \theta) = \int_{-\infty}^{\infty} g(\rho \cos\theta + t \sin\theta, \rho \sin\theta - t \cos\theta) dt \tag{4.107}$$

where parameter, t, is the length of straight line $\cos\theta x + \sin\theta y = \rho$. It is shown in Fig (4.6) that

$$x = \rho \cos\theta + t \sin\theta, \text{ and } y = \rho \sin\theta - t \cos\theta \tag{4.108}$$

In Fig. (4.6), it can be seen that two very bright spots are found in the Radon transform, and the postion shown the parameters of the lines in the real physical image.

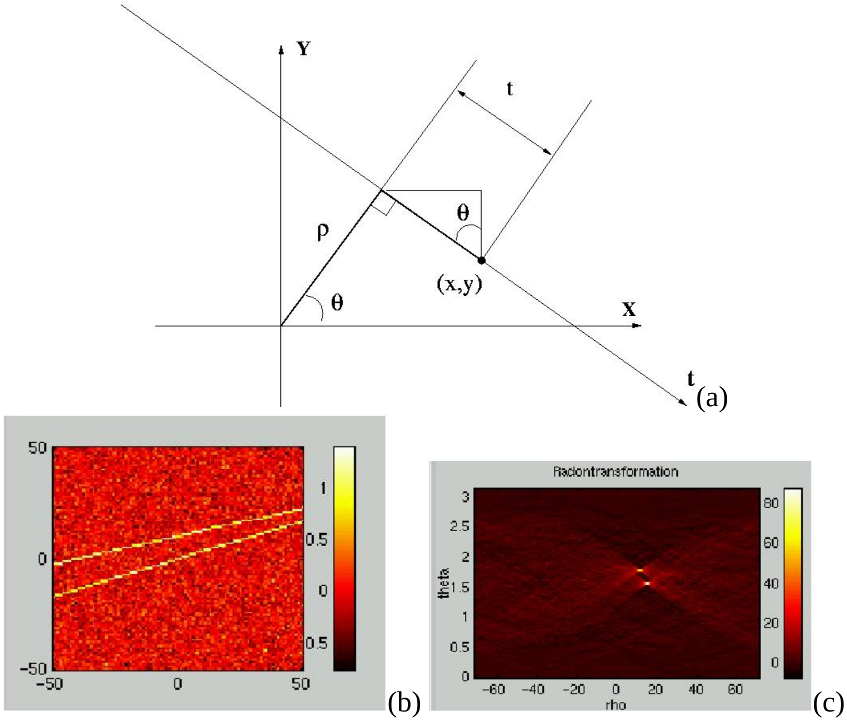


Figure 4.6. Two-dimensional Radon transform: (a) prjectinline, (b) image in the physical space, and (c) image in the Radon transform space

EXAMPLE 4.17 *Let $f(\mathbf{x}) = \delta(\mathbf{x})$. The Radon transform of Dirac's delta function is*

$$\hat{\delta}(s, \mathbf{n}) = \mathbf{R}(\delta) = \int_{-\infty}^{\infty} \delta(\mathbf{x})\delta(s - \mathbf{n} \cdot \mathbf{x})d\mathbf{x} = \delta(s) \tag{4.109}$$

where $s = n_i x_i$.

Subsequently,

$$\tilde{\delta}(s, \mathbf{n}) = \delta''(s) \tag{4.110}$$

and the inverse Radon transform is

$$\delta(\mathbf{x}) = -\frac{1}{8\pi^2} \int_{S^2} \delta''(n_k x_k) dS \tag{4.111}$$

One can verify this by considering the identity (4.94), i.e.

$$\int_{S^2} \delta(n_k x_k) dS = \frac{2\pi}{|\mathbf{x}|} \tag{4.112}$$

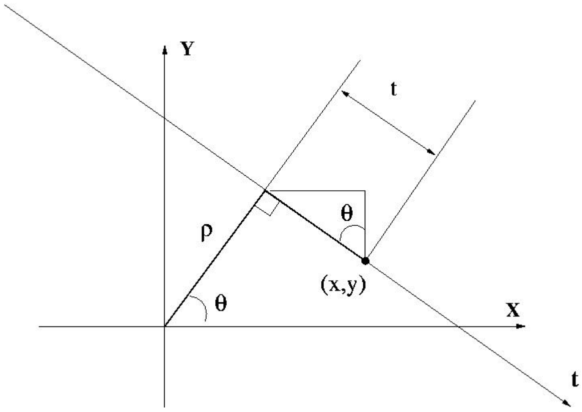


Figure 4.7. Two-dimensional Radon transform: (a) projection line, (b) image in the physical space, and (c) image in the Radon transform space

Applying the harmonic operator $\nabla^2 = \dfrac{\partial^2}{\partial x_i \partial x_i}$ *to the above identity and considering Example (4.15) (Eq.(4.68)) yields*

$$\int_{S^2} \delta''(n_k x_k) n_i n_i dS = 2\pi \nabla^2 \Big(\frac{1}{|\mathbf{x}|}\Big) = -8\pi^2 \delta(\mathbf{x}) \tag{4.113}$$

Now we use the Radon transform to derive 3D static Green's function of a linear elastic medium. Consider the concentrated load is acting at the origin of the coordinat $(\mathbf{y} = \mathbf{0})$.

$$C_{ijkl} G_{km,lj} + \delta(\mathbf{x}) \delta_{im} = 0 \tag{4.114}$$

Assume that the Green's function can be written as a form of inverse Radon transform,

$$G^{\infty}_{km}(\mathbf{x}) = -\frac{1}{8\pi^2} \int_{S^2} \tilde{G^{\infty}_{km}}(\bar{\xi}_n x_n) dS \tag{4.115}$$

Then

$$G^{\infty}_{km,lj}(\mathbf{x}) = -\frac{1}{8\pi^2} \int_{S^2} \tilde{G^{\infty}_{km}}''(\bar{\xi}_n x_n) \bar{\xi}_l \bar{\xi}_j dS \tag{4.116}$$

On the other hand,

$$\delta(\mathbf{x}) = \mathbf{R}^{-1}\Big(\tilde{\delta}(s)\Big) = -\frac{1}{8\pi^2} \int_{S^2} \delta''(\bar{\xi}_n x_n) dS \tag{4.117}$$

We then obtain

$$C_{ijkl} \bar{\xi}_j \bar{\xi}_l \tilde{G^{\infty}_{km}}''(\bar{\xi}_n x_n) = -\delta_{im} \delta''(\bar{\xi}_n x_n) \tag{4.118}$$

which leads to

$$\tilde{G^{\infty}_{ij}}(\bar{\xi}) = -K^{-1}_{ij}(\xi)\delta(\bar{\xi}_n x_n) + C_1\bar{\xi}_n x_n + C_0 \tag{4.119}$$

where

$$K^{-1}_{ri} C_{ijkl}\bar{\xi}_l\bar{\xi}_j = \delta_{rk}, \ \ or \ \ K_{ij} = \frac{N_{ij}(\bar{\xi})}{D(\xi)} \tag{4.120}$$

Note that $C_1 = C_0 = 0$ because it is required that $G^{\infty}_{ij}(\mathbf{x}) \to 0$, as $\mathbf{x} \to \infty$.

For isotropic materials,

$$K^{-1}_{ij}(\bar{\boldsymbol{\xi}}) = \frac{1}{\mu}\Big[\delta_{ij} - \frac{(\lambda+\mu)\bar{\xi}_i\bar{\xi}_k}{(\lambda+2\mu)}\Big] \tag{4.121}$$

and, correspondingly,

$$G^{\infty}_{ij}(\mathbf{x}) = \frac{1}{8\pi^2}\int_{S^2} K^{-1}_{ij}(\bar{\boldsymbol{\xi}})\delta(\bar{\xi}_n x_n) dS \tag{4.122}$$

and subsequently,

$$G^{\infty}_{ij}(\mathbf{x}) = \frac{1}{4\pi\mu}\Big[\frac{\delta_{ij}}{|\mathbf{x}|} - \frac{(\lambda+\mu)}{2(\lambda+2\mu)}|\mathbf{x}|_{,ij}\Big] \tag{4.123}$$

4.6 Joseph Fourier(I)

Joseph Fourier was born in 1768 in Auxerre, the ninth child of a master tailor. Although the death of his father left him an orphan at the age of ten, his intelligence gained him a free place at the local Benedictine school. At the end of a brilligent school career he applied to enter the artillery only to be informed that such a profession was only open to those of noble blood and was closed to him 'even if he were a second Newton'.

Fourier began to prepare to enter the Benedictine teaching order but, whatever his plans may have been, the course of his life was violently altered by the outbreak of the French Revolution, ... The situation of the new Republic called for ruthless measures which the government, conscious of its own revolutionary virtue, was well prepared to take. Treachery was fought by a political terror in which opponents both to the left and right were executed and, as the definition of treachery was extended, it became clear that no one was safe. Fourier himself was arrested, released and then rearrested. A deputation from Auxerre which, with considerable courage, went to Paris to plea his case, was told-'Yes, he speaks well, but we nolonger have any need of musical patriots.' Only the fall of Robespierre saved Fourier's head.

However Fourier's release did not mark the end of his troubles. As coup d'etate follows coup d'eta, and the revolution swung erractically to the right he

Figure 4.8. Joseph Fourier

would remain a marked man. No one had been executed in Auxerre but Fourier had been an agent of the terror there. His arrest was on a charge of H'ebertism and the H'ebertists were to the left of Robespierre. The word 'terrorist' then, like 'Trotskyist' now, denoted a defeated yet feared opponent.

Luckily an opportunity to leave Auxerre now presented itself. A new college (the Echole Normale) was being set up in Paris to help train teachers and Fourier could now study under men like Lagrange, Monge and Laplace and excape his terrorist past. Fourier's talents were soon noted, but the college was not successful and its closure was followed by further problems for Fourier.

'We shudder when we think that the pupils of the Ecole Normale were chosen under the reign of Robespierre and his proteges. It is only too true that Balme and Fourier, pupils of the department of Yonne have long prefessed the atrocious principles and infernal maxims of the tyrants. Nevertheless they prepare to become teachers of our children. Is it not to vomit their poison in the bosim of innocence (From an address to the National Convention, quoate by Herivel)'

Fourier was again arrested, released, rearrested and finally, following yet another political swing, released to become a teacher at the new Ecole Polytechnique.

Here Fourier remained for three years. That his talent was recognized is shown by the fact that he succeeded Lagrange in the Chair of Analysis and Mechanics. The quiet interlude was ended by a gonernment order to join the invasion of Egypt. Ostensible intended to liberate Egypt from the Turks and to threaten the British position in India, the expedition may have been seen by the government as a way of keeping a troublesome general as far away as possible and by the general (Napoleon) as the first step toward becoming Emperor of the East. Fourier wa one of a ghroup of scientists and intellectualls intended to form part of the immense cultural benefits that France was to bestow on Egypt.

Both before and after Napolean's departure, Fourier occupied several important administrative and political posts in Egypt. When the French expedition finally surrended in 1801 and Fourier was repatriated, Napoleon offered him the post of Prefect of the Department of the Isere centred round Grenoble (France had been divided into 83 Departments and each Prefect governed his Department of behalf of the central government.)

Although he could have continued a Professor at the Polytechnique, Fourier accepted the offer. Herivel suggests that Egypt had given him a taste for administration and that he hoped to rise higher. Herivel also accounts that Fourier's close association with Kleber after Napoleon's departure account for the fact that these hopes were not fulfilled.

Fourier seems to have been popular and efficient Prefect. His greatest achievement during his 14 years of office was by reconciling the conflicting interests of some forty communities to enable the swamps of Bourgion to be drained. The draining of twenty thousand acres of swamps resulted in major economic and health benefits and was achieved during a period morenoted for grandiose paper plans than for concrete achievements. Fourier's other administrative memorial was a new road across the Alps (now Route 91).

Apart from his perfectorial duties Fourier helped organize the Description of Egypt. This work written by the intellectuals attached to the Egyptian expedetion did much to inspire European interest in Egypt and was thus one of the two permanent results of the expediton. (The other was the discovery of the Rosetta Stone, atrilingual inscription which was to provide the key to the deciphering of hieriglyphics.)

Fourier's main contribution was the general introduction – a survey of Egyptian history up to modern times. An Egyptologist with whom I discussed this described the introduction as a masterpiece and a turning point in the subject, was surprised to hear that Fourier also had a reputation as a mathematician!

–T.W.Korner From *Fourier Analysis*

4.7 Exercises

PROBELM 4.1 *Find the Green's function for a both end clamped Euler-Bernoulli beam, i.e.*

$$\frac{d^2}{dx^2} EI \frac{d^2 G(x,y)}{dx^2} = \delta(x-y), \quad \forall x, y \in (0,\ell) \tag{4.124}$$

and

$$G(0,y) = G(\ell,y) = 0, \quad G'(0,y) = G'(\ell,y) = 0\ . \tag{4.125}$$

PROBELM 4.2 *For isotropic materials, elasticity tensor has the form*

$$C_{ijk\ell} = \lambda \delta_{ij} \delta_{k\ell} + \mu(\delta_{i\ell}\delta_{jk} + \delta_{ik}\delta_{j\ell}) \tag{4.126}$$

Show

1.

$$K_{ik}(\boldsymbol{\xi}) = C_{ijk\ell}\xi_j\xi_\ell = (\lambda+\mu)\xi_i\xi_k + \mu\delta_{ik}\xi_j\xi_j \tag{4.127}$$

2. *(Hint : use $e_{ijk}e_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}$.)*

$$\begin{aligned} N_{ij}(\boldsymbol{\xi}) &= \frac{1}{2} e_{ik\ell} e_{jmn} K_{km} K_{\ell n} \\ &= \mu\xi^2((\lambda+2\mu)\delta_{ij}\xi^2 - (\lambda+\mu)\xi_i\xi_j) \end{aligned} \tag{4.128}$$

3.

$$D(\boldsymbol{\xi}) = \mu^2(\lambda+2\mu)\xi^6 \tag{4.129}$$

PROBELM 4.3 *The Green's function, $G^\infty(\mathbf{x},\mathbf{x}')$, satisfies the 2D Laplace equation,*

$$\nabla^2 G^\infty(\mathbf{x},\mathbf{x}') + \delta(\mathbf{x}-\mathbf{x}') = \mathbf{0}, \ \ \forall \mathbf{x} \in \mathbb{R}^{\mathbf{2}} \tag{4.130}$$

where $\nabla^2 = \dfrac{\partial^2}{\partial x_1^2} + \dfrac{\partial^2}{\partial x_2^2} = \dfrac{\partial^2}{\partial x_\alpha \partial x_\alpha}$, $\alpha = 1,2$. *And* $\delta(\mathbf{x}-\mathbf{x}') = \delta(x_1 - x_1')\delta(x_2 - x_2')$. *Use Fourier transform method to derive*

$$G^\infty(\mathbf{x}-\mathbf{x}') = -\frac{1}{2\pi}\ln|\mathbf{x}-\mathbf{x}'|\ . \tag{4.131}$$

Hints

$$\delta(\mathbf{x}-\mathbf{x}') = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\Big(i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{x}')\Big) d\boldsymbol{\xi} \tag{4.132}$$

and

$$\begin{aligned} &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\exp(i(\xi_1 x_1 + \xi_2 x_2))}{\xi_1^2 + \xi_2^2} d\xi_1 \xi_2 \\ = &-2\pi lnR \end{aligned} \tag{4.133}$$

where $R = \sqrt{(x_1 - x_1')^2 + (x_2 - x_2')^2}$.

PROBELM 4.4 *In isotropic materials, the static Green's function of linear elasticity is*

$$G^{\infty}_{ij}(\mathbf{x}, \mathbf{x}') = \frac{1}{4\pi\mu} \frac{\delta_{ij}}{|\mathbf{x} - \mathbf{x}'|} - \frac{1}{16\pi\mu(1-\nu)} \frac{\partial^2}{\partial x_i \partial x_j} |\mathbf{x} - \mathbf{x}'| \tag{4.134}$$

Let $\bar{\mathbf{x}} = \mathbf{x} - \mathbf{x}'$ *and* $\bar{x} = |\bar{\mathbf{x}}| = |\mathbf{x} - \mathbf{x}'|$. *Show that for isotropic materials,*

$$C_{j\ell mn} G_{ij,\ell} = \frac{-1}{8\pi(1-\nu)} \left\{ (1-2\nu) \frac{\delta_{mi}\bar{x}_n + \delta_{ni}\bar{x}_m - \delta_{mn}\bar{x}_i}{\bar{x}^3} + 3 \frac{\bar{x}_m \bar{x}_n \bar{x}_i}{\bar{x}^5} \right\} \tag{4.135}$$

where ν *is the Poisson ratio, and* μ, λ *are the Lamé constants with*

$$\lambda = \frac{2\mu\nu}{1-2\nu}, \quad \mu = \frac{\lambda(1-2\nu)}{2\nu}, \quad \nu = \frac{\lambda}{2(\lambda+\mu)} \tag{4.136}$$

Hint: ($C_{j\ell mn} = \lambda \delta_{j\ell}\delta_{mn} + \mu(\delta_{jm}\delta_{\ell n} + \delta_{jn}\delta_{\ell m})$*).*

Chapter 5

EIGENSTRAIN THEORY

There are mainly two homogenization meghods used in engineering applications today. The first one is Eshelby's, or Mura's eigenstrain theory. The central part of the theory is Eshelby's eigenstrain solution for ellipsoidal inclusion. The theory has been further refined, detailed and articulated by Professor Mura and his co-workers. Today, it is called eigenstrain theory, and it has widespread applications.

5.1 Fundamental equations of micro-elasticity

Consider equilibrium equation in an RVE

$$\sigma_{ji,j} = 0 \tag{5.1}$$

After homogenization, inhomogeneities are replaced by a eigenstrain distribution ${\epsilon^*}_{ij}(\mathbf{x})$. Assuming that material is linear elastic, and the total strain is the sum of elastic strain and eigenstrain,

$$\epsilon_{ij} = e_{ij} + {\epsilon^*}_{ij} \tag{5.2}$$

The total strain is defined as $\epsilon_{ij} = \dfrac{1}{2}(u_{i,j} + u_{j,i})$. And elastic strain is related with Cauchy stress by Hooke's law

$$\sigma_{ij} = C_{ijk\ell}(\epsilon_{k\ell} - {\epsilon^*}_{k\ell}) = C_{ijk\ell}(u_{k,\ell} - {\epsilon^*}_{k\ell}) \tag{5.3}$$

The equilibrium equation then takes a form

$$C_{ijk\ell}u_{i,\ell j} - C_{ijk\ell}{\epsilon^*}_{k\ell,j} = 0 \tag{5.4}$$

Note that one interprets the effect of eigenstrain distribution as a type of body force, $f_i = -C_{ijk\ell}{\epsilon^*}_{k\ell,j}$, and the original equilibrium equation has the form $\sigma_{ji,j} + f_i = 0$.

Let,

$$
\begin{aligned}
u_k(\mathbf{x}) &= \int_{-\infty}^{\infty} \bar{u}_k(\boldsymbol{\xi}) \exp(i\boldsymbol{\xi} \cdot \mathbf{x}) d\mathbf{x} \\
&= \int_{-\infty}^{\infty} \bar{u}_k(\boldsymbol{\xi}) \exp(i\xi_m x_m) d\mathbf{x} && (5.5) \\
\epsilon^*{}_{k\ell}(\mathbf{x}) &= \int_{-\infty}^{\infty} \bar{\epsilon}^*_{k\ell}(\boldsymbol{\xi}) \exp(i\xi_m x_m) d\mathbf{x} && (5.6)
\end{aligned}
$$

Hence

$$
\begin{aligned}
u_{k,\ell j}(\mathbf{x}) &= -\int_{-\infty}^{\infty} \bar{u}_k \xi_\ell \xi_j(\boldsymbol{\xi}) \exp(i\xi_m x_m) d\mathbf{x} && (5.7) \\
\epsilon^*{}_{k\ell,j}(\mathbf{x}) &= i\int_{-\infty}^{\infty} \bar{\epsilon}^*_{k\ell}(\boldsymbol{\xi}) \xi_j \exp(i\xi_m x_m) d\mathbf{x} && (5.8)
\end{aligned}
$$

Substituting (5.7) and (5.8) into (5.4) yields

$$
\int_{-\infty}^{\infty} (C_{ijk\ell}\bar{u}_k \xi_\ell \xi_j + iC_{ijk\ell}\bar{\epsilon}^*_{k\ell}(\boldsymbol{\xi})\xi_j) \exp(i\xi_m x_m) d\mathbf{x} = 0 \qquad (5.9)
$$

which leads to

$$
C_{ijk\ell}\xi_j \xi_\ell \bar{u}_k = -iC_{ijk\ell}\bar{\epsilon}^*_{k\ell}(\boldsymbol{\xi})\xi_j \qquad (5.10)
$$

Denote

$$
\begin{aligned}
K_{ik}(\boldsymbol{\xi}) &= C_{ijk\ell}\xi_j \xi_\ell && (5.11) \\
\bar{f}_i &= -iC_{ijk\ell}\bar{\epsilon}^*_{k\ell}\xi_j && (5.12)
\end{aligned}
$$

They are related by

$$
\begin{bmatrix} K_{11} & K_{12} & K_{13} \\ K_{21} & K_{22} & K_{23} \\ K_{31} & K_{32} & K_{33} \end{bmatrix} \begin{pmatrix} \bar{u}_1 \\ \bar{u}_2 \\ \bar{u}_3 \end{pmatrix} = \begin{pmatrix} \bar{f}_1 \\ \bar{f}_2 \\ \bar{f}_3 \end{pmatrix} \qquad (5.13)
$$

We find that

$$
\bar{u}_i(\boldsymbol{\xi}) = \frac{N_{ij}(\boldsymbol{\xi})}{D(\boldsymbol{\xi})}\bar{f}_j = K^{-1}_{ij}\bar{f}_j \qquad (5.14)
$$

where

$$
\begin{aligned}
N_{ij}(\boldsymbol{\xi}) &= \frac{1}{2} e_{ik\ell} e_{jmn} K_{km} K_{\ell n} && (5.15) \\
D(\boldsymbol{\xi}) &= e_{mn\ell} K_{m1} K_{n2} K_{\ell 3} && (5.16)
\end{aligned}
$$

For isotropic materials,

$$\begin{aligned}
\mathbf{K}(\boldsymbol{\xi}) &= \boldsymbol{\xi}\cdot\mathbf{C}\cdot\boldsymbol{\xi} = \boldsymbol{\xi}\cdot\Big\{\lambda\mathbf{1}^{(2)}\otimes\mathbf{1}^{(2)} + 2\mu\mathbf{1}^{(4s)}\Big\}\cdot\boldsymbol{\xi} \\
&= \lambda\boldsymbol{\xi}\otimes\boldsymbol{\xi} + \mu\Big(\boldsymbol{\xi}\otimes\boldsymbol{\xi} + |\boldsymbol{\xi}|^2\mathbf{1}^{(2)}\Big) \\
&= (\lambda+\mu)\boldsymbol{\xi}\otimes\boldsymbol{\xi} + \mu|\boldsymbol{\xi}|^2\mathbf{1}^{(2)} \qquad (5.17)
\end{aligned}$$

Denote

$$\mathbf{Q}(\boldsymbol{\xi}) = \mathbf{K}^{-1}(\boldsymbol{\xi}) \qquad (5.18)$$

$\mathbf{Q}$ must be an isotropic second order tensor in Fourier space as well. Assume that

$$\mathbf{Q}(\boldsymbol{\xi}) = \{\boldsymbol{\xi}\cdot\mathbf{C}\cdot\boldsymbol{\xi}\}^{-1} = A\boldsymbol{\xi}\otimes\boldsymbol{\xi} + B\mathbf{1}^{(2)} \qquad (5.19)$$

then

$$\Big[(\lambda+\mu)\boldsymbol{\xi}\otimes\boldsymbol{\xi} + \mu|\boldsymbol{\xi}|^2\mathbf{1}^{(2)}\Big]\cdot\Big[A\boldsymbol{\xi}\otimes\boldsymbol{\xi} + B\mathbf{1}^{(2)}\Big] = \mathbf{1}^{(2)} \qquad (5.20)$$

subsequently,

$$\Big[A(\lambda+2\mu)|\boldsymbol{\xi}|^2 + B(\lambda+\mu)\Big]\boldsymbol{\xi}\otimes\boldsymbol{\xi} + B\mu|\boldsymbol{\xi}|^2\mathbf{1}^{(2)} = \mathbf{1}^{(2)} \qquad (5.21)$$

One can then determine the constant A and B,

$$\begin{aligned}
A &= -\frac{(\lambda+\mu)}{\mu(\lambda+2\mu)|\boldsymbol{\xi}|^4} \qquad (5.22)\\
B &= \frac{1}{\mu|\boldsymbol{\xi}|^2} \qquad (5.23)
\end{aligned}$$

Hence,

$$\mathbf{Q}(\boldsymbol{\xi}) = \Big(\boldsymbol{\xi}\cdot\mathbf{C}\cdot\boldsymbol{\xi}\Big)^{-1} = \frac{|\boldsymbol{\xi}|^{-2}}{\mu}\Big\{-\frac{(\lambda+\mu)}{\mu(\lambda+2\mu)|\boldsymbol{\xi}|^2}\boldsymbol{\xi}\otimes\boldsymbol{\xi} + \mathbf{1}^{(2)}\Big\} \qquad (5.24)$$

or in component form,

$$Q_{ij} = K_{ij}^{-1} = \frac{|\boldsymbol{\xi}|^{-2}}{\mu}\Big\{-\frac{(\lambda+\mu)}{\mu(\lambda+2\mu)|\boldsymbol{\xi}|^2}\xi_i\xi_j + \delta_{ij}\Big\} \qquad (5.25)$$

Consider

$$\bar{u}_i(\boldsymbol{\xi}) = Q_{ij}(\boldsymbol{\xi})\bar{f}_j = -iC_{j\ell mn}\bar{\epsilon}^*_{mn}\xi_\ell\frac{N_{ij}(\boldsymbol{\xi})}{D(\boldsymbol{\xi})} \qquad (5.26)$$

Applying Fourier inverse transform,

$$\begin{aligned}
u_i(\mathbf{x}) &= -i\int_{-\infty}^{\infty} C_{j\ell mn}\bar{\epsilon}^*_{mn}(\boldsymbol{\xi})\xi_\ell\frac{N_{ij}(\boldsymbol{\xi})}{D(\boldsymbol{\xi})}\exp(i\boldsymbol{\xi}\cdot\mathbf{x})d\boldsymbol{\xi} \qquad (5.27)\\
&= -\int_{-\infty}^{\infty}\bar{f}_j(\boldsymbol{\xi})\frac{N_{ij}(\boldsymbol{\xi})}{D(\boldsymbol{\xi})}\exp(i\boldsymbol{\xi}\cdot\mathbf{x})d\boldsymbol{\xi} \qquad (5.28)
\end{aligned}$$

Consequences of (5.27) are

$$\begin{aligned} \epsilon_{ij}(\mathbf{x}) &= \frac{1}{2}\int_{-\infty}^{\infty} C_{k\ell mn}\bar{\epsilon}^{*}_{mn}(\boldsymbol{\xi})\xi_\ell\Big(N_{ik}(\boldsymbol{\xi})\xi_j + N_{jk}(\boldsymbol{\xi})\xi_i\Big)D^{-1}(\boldsymbol{\xi}) \\ &\quad \cdot \exp(i\boldsymbol{\xi}\cdot\mathbf{x})d\boldsymbol{\xi} \qquad (5.29) \\ \sigma_{ij}(\mathbf{x}) &= C_{ijk\ell}\Big\{\int_{-\infty}^{\infty}\Big(C_{pqmn}\bar{\epsilon}^{*}_{mn}(\boldsymbol{\xi})\xi_q\xi_\ell N_{kp}(\boldsymbol{\xi})D^{-1}(\boldsymbol{\xi}) \\ &\quad \cdot \exp(i\boldsymbol{\xi}\cdot\mathbf{x})\Big)d\boldsymbol{\xi} - {\epsilon^{*}}_{k\ell}(\mathbf{x})\Big\} \qquad (5.30) \end{aligned}$$

5.2 Method of Green's Functions

Consider

$$G_{ij}(\mathbf{x}-\mathbf{y}) = \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty} N_{ij}(\boldsymbol{\xi})D^{-1}(\boldsymbol{\xi})\exp(i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{y}))d\boldsymbol{\xi} \qquad (5.31)$$

Based on convolution theorem and according to (5.28) and (5.29), one can derive that

$$\begin{aligned} u_i(\mathbf{x}) &= -\int_{-\infty}^{\infty} C_{j\ell mn}{\epsilon^{*}}_{mn}(\mathbf{y})G_{ij,\ell}(\mathbf{x}-\mathbf{y})d\mathbf{y} \qquad (5.32) \\ u_i(\mathbf{x}) &= \int_{-\infty}^{\infty} G_{ij}(\mathbf{x}-\mathbf{y})f_j(\mathbf{y})d\mathbf{y} \qquad (5.33) \end{aligned}$$

The corresponding expressions for stress and strain are

$$\begin{aligned} \epsilon_{ij}(\mathbf{x}) &= -\frac{1}{2}\int_{-\infty}^{\infty} C_{k\ell mn}{\epsilon^{*}}_{mn}(\mathbf{y})\{G_{ik,\ell j}(\mathbf{x}-\mathbf{y}) \qquad (5.34) \\ &\quad +G_{jk,\ell i}(\mathbf{x}-\mathbf{y})\}d\mathbf{y} \qquad (5.35) \\ \sigma_{ij}(\mathbf{x}) &= -C_{ijk\ell}\Big\{\int_{-\infty}^{\infty} C_{pqmn}{\epsilon^{*}}_{mn}(\mathbf{y})G_{kp,q\ell}(\mathbf{x}-\mathbf{y})d\mathbf{y} \\ &\quad +{\epsilon^{*}}_{k\ell}(\mathbf{x})\Big\} \qquad (5.36) \end{aligned}$$

Eq.(5.37) is rewritten by Mura(1963) as the following form

$$\sigma_{ij}(\mathbf{x}) = C_{ijk\ell}\int_{-\infty}^{\infty} e_{sth}e_{\ell nh}C_{pqmn}G_{kp,qt}(\mathbf{x}-\mathbf{y}){\epsilon^{*}}_{sm}d\mathbf{y} \qquad (5.37)$$

To prove the equivalenct between (5.37) and (5.38), we use the identity $e_{sth}e_{\ell nh} = \delta_{s\ell}\delta_{tn} - \delta_{sn}\delta_{t\ell}$ to expand (5.38),

$$\begin{aligned} \sigma_{ij}(\mathbf{x}) &= C_{ijk\ell}\int_{-\infty}^{\infty} C_{pqmn}\Big(\delta_{s\ell}\delta_{tn} - \delta_{sn}\delta_{tl}\Big)G_{kp,qt}(\mathbf{x}-\mathbf{y}){\epsilon^{*}}_{sm}d\mathbf{y} \\ &= C_{ijk\ell}\int_{-\infty}^{\infty} C_{pqmn}\Big(G_{kp,qn}(\mathbf{x}-\mathbf{y}){\epsilon^{*}}_{\ell m} - G_{kp,q\ell}(\mathbf{x}-\mathbf{y}){\epsilon^{*}}_{nm}\Big)d\mathbf{y} \end{aligned} \qquad (5.38)$$

The first term of the integrand is

$$C_{pqmn}G_{kp,qn}(\mathbf{x}-\mathbf{y}) = G_{mnpq}G_{kp,qn}(\mathbf{x}-\mathbf{y}) = -\delta_{mk}\delta(\mathbf{x}-\mathbf{y}) \tag{5.39}$$

Therefore,

$$\begin{aligned} & C_{ijk\ell}\int_{-\infty}^{\infty} C_{pqmn}G_{kp,qn}(\mathbf{x}-\mathbf{y}){\epsilon^*}_{\ell m}d\mathbf{y} \\ = & -C_{ijk\ell}\int_{-\infty}^{\infty}\delta(\mathbf{x}-\mathbf{y}){\epsilon^*}_{k\ell}d\mathbf{y} = -C_{ijk\ell}{\epsilon^*}_{k\ell} \end{aligned} \tag{5.40}$$

We then recover (5.37).

Recall,

$$G_{ij}^{\infty}(\mathbf{x}-\mathbf{y}) = \frac{1}{8\pi^2}\int_{S^2}\delta\Big((\mathbf{x}-\mathbf{y})\cdot\boldsymbol{\xi}\Big)Q_{ij}(\boldsymbol{\xi})dS \tag{5.41}$$

where $Q_{ij}(\boldsymbol{\xi}) = N_{ij}(\boldsymbol{\xi})/D(\boldsymbol{\xi})$.

Substitute (5.42) into (5.34),

$$\begin{aligned} u_i(\mathbf{x}) &= -\frac{1}{8\pi^2}\int_{-\infty}^{\infty}\Big(\int_{S^2}\delta\Big((\mathbf{x}-\mathbf{y})\cdot\boldsymbol{\xi}\Big)Q_{ij}(\boldsymbol{\xi})dS\Big)f_j(\mathbf{y})d\mathbf{y} \\ &= -\frac{1}{8\pi^2}\int_{S^2}Q_{ij}(\boldsymbol{\xi})\Big[\int_{-\infty}^{\infty}f_j(\mathbf{y})\delta(s-y_m\xi_m)d\mathbf{y}\Big]dS \\ &= -\frac{1}{8\pi^2}\int_{S^2}Q_{ij}(\boldsymbol{\xi})\hat{f}(s,\boldsymbol{\xi})dS \end{aligned} \tag{5.42}$$

where $s = x_m\xi_m$ and

$$\hat{f}_j(s,\boldsymbol{\xi}) = \int_{-\infty}^{\infty}f_j(\mathbf{y})\delta(s-y_m\xi_m)d\mathbf{y}$$

is the Radon transform of $f_j(\mathbf{y})$.

EXAMPLE 5.1 *Assume that a linearly distributed eigenstrain is prescribed in a spherical ball* $(|\mathbf{x}| \le a)$.

$${\epsilon^*}_{k\ell} = \begin{cases} \frac{1}{2}(c_k x_\ell + c_\ell x_k) & |\mathbf{x}| \le a \\ 0 & |\mathbf{x}| > a \end{cases} \tag{5.43}$$

Hence

$${\epsilon^*}_{k\ell,j} = \frac{1}{2}\Big(c_k\delta_{\ell j} + c_\ell\delta_{kj}\Big) \tag{5.44}$$

and for isotropic materials

$$f_i = -C_{ijk\ell}{\epsilon^*}_{k\ell,j} = \frac{1}{2}\Big(C_{ijkj}c_k + C_{ijj\ell}c_\ell\Big) = -(\lambda+4\mu)c_i$$

The area of intersection of the plane $\xi_m x_m = s$ with the sphere of radus a is $\pi(a^2 - s^2)$, if $|s| \leq a$ and zero otherwise. Thus

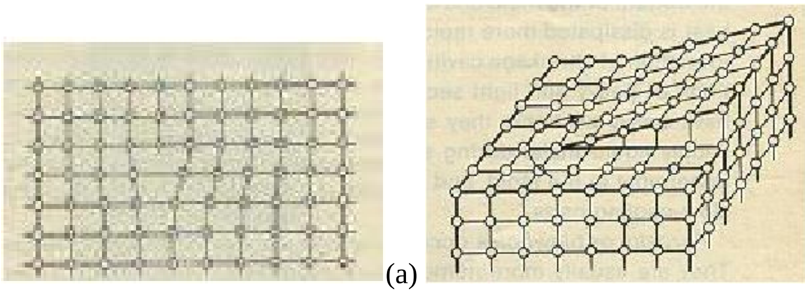
$$\begin{aligned}
\hat{f}_j(s, \boldsymbol{\xi}) &= -\int_{-\infty}^{\infty} (\lambda + 4\mu) c_i \delta(s - x_m \xi_m) d\mathbf{y} \\
&= -\int_{S^a \cap \{\xi_m x_m = s\}} (\lambda + 4\mu) c_i dS = -(\lambda + 4\mu) c_i \pi (a^2 - s^2) \\
&= -(\lambda + 4\mu) c_i \pi (a^2 - (\xi_m x_m)^2) \qquad (5.45)
\end{aligned}$$

Therefore, the induced displacement field inside the sphere is

$$u_i(\mathbf{x}) = \frac{(\lambda + 4\mu)}{8\pi^2} \int_{S^2} Q_{ij}(\boldsymbol{\xi}) c_j (a^2 - (\xi_m x_m)^2) H(a^2 - x_m x_m) dS \quad (5.46)$$

where $H(\cdot)$ is the Heaviside function, $\xi_m \xi_m = 1$, and

$$Q_{ij}(\boldsymbol{\xi}) = \frac{1}{\mu}\Big[\delta_{ij} - \frac{(\lambda + \mu)\xi_i \xi_j}{(\lambda + 2\mu)}\Big]$$



(a) (b)

Figure 5.1. Illustraions of dislocations: (a)edge dislocation, and (b) screw dislocation

5.3 Application I: Dislocation problems

A dislocation is a distorted region among substantially perpect crystal lattice environment. In other words, a dislocation is a linear defect around which some of the atoms are misaligned or crystal lattice being distorted. There are two types of dislocations: (1) edge dislocation, and (2) screw dislocation (see Fig. 5.1). Use of eigenstrain theory to describe the effect of dislocations and their induced disturbance mechanical fields is a success. Eigenstrain theory has been an important approach in the development of dislocation theory. Here we only introduce some simple examples.

Consider a straight screw dislocation on a half space. There is a jump or discontinuity in displacement at $x_2 = 0$ and $-\infty < x_1 < 0$, with the magnitude of b(burgers vector). A ficticious eigenstrain field is prescribed on the slip plane to mimic the mechanical effect of dislocation,

$$\epsilon^*{}_{23} = \begin{cases} \frac{1}{2} b\delta(x_2)H(-x_1), & \mathbf{x} \in \Omega \\ 0. & \mathbf{x} \in \mathbb{R}^3/\Omega \end{cases} \tag{5.47}$$

where the slip surface may be described as

$$\Omega = \Big\{(x_1, 0, x_3)\Big| x_1 < 0, -\infty < x_3 < \infty\Big\}$$

and $H(\cdot)$ is the heaviside function.

The eigenstrain field may be considered as the consequence of the displacement field,

$$u_3^*(\mathbf{x}) = bH(x_2)H(-x_1) \tag{5.48}$$

since

$$\epsilon^*{}_{23} = \frac{1}{2}\Big(\frac{\partial u_3^*}{\partial x_2} + \frac{\partial u_2^*}{\partial x_3}\Big) = \frac{b}{2}\delta_{x_2} H(-x_1)$$

(Question: what about $\epsilon^*{}_{31}$?)

Apply Fourier transform

$$\begin{aligned} \bar{\epsilon}_{23}^*(\boldsymbol{\xi}) &= \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty} \epsilon^*{}_{23}(\mathbf{x}) \exp(-i\boldsymbol{\xi}\cdot\mathbf{x}))d\mathbf{x} \\ &= \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty} \frac{b}{2}\delta(x_2)H(-x_1)\exp(-i\boldsymbol{\xi}\cdot\mathbf{x})d\mathbf{x} \end{aligned} \tag{5.49}$$

Consider

$$\int_{-\infty}^{\infty} \delta(x_2)\exp(-i\xi_2 x_2)dx_2 = 1 \tag{5.50}$$

$$\int_{-\infty}^{\infty} H(-x_1)\exp(-i\xi_1 x_1)dx_1 = \int_{-\infty}^{0} \exp(-i\xi_1 x_1)dx_1$$

$$= \frac{i}{\xi_1} \quad Im(\xi_1) < 0$$

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-i\xi_3 x_3)dx_3 = \delta(\xi_3) \tag{5.51}$$

Therefore,

$$\bar{\epsilon}_{23}^* = \frac{1}{(2\pi)^2}\frac{b}{2}\Big(\frac{i}{\xi_1}\Big)\delta(\xi_3) \tag{5.52}$$

Substituing (5.53) into the general formula of micro-elasticity,

$$\begin{aligned} u_i(\mathbf{x}) &= -i\int_{-\infty}^{\infty} C_{j\ell mn}\bar{\epsilon}_{mn}^*\xi_\ell Q_{ij}(\boldsymbol{\xi})\exp(i\boldsymbol{\xi}\cdot\mathbf{x})d\boldsymbol{\xi} \\ &= -2i\int_{-\infty}^{\infty} C_{j\ell 23}\bar{\epsilon}_{23}^*\xi_\ell Q_{ij}(\boldsymbol{\xi})\exp(i\boldsymbol{\xi}\cdot\mathbf{x})d\boldsymbol{\xi} \\ &= \Big(\frac{2b}{2(2\pi^2)}\Big)\int_{-\infty}^{\infty}\Big(\frac{\delta(\xi_3)}{\xi_1}\Big)C_{j\ell 23}Q_{ij}(\boldsymbol{\xi})\exp(i\boldsymbol{\xi}\cdot\mathbf{x})d\boldsymbol{\xi} \end{aligned} \tag{5.53}$$

where the factor 2 is due to the presence of $\epsilon^*{}_{32}$, if the minor synmmetry is being considered. For isotropic materials,

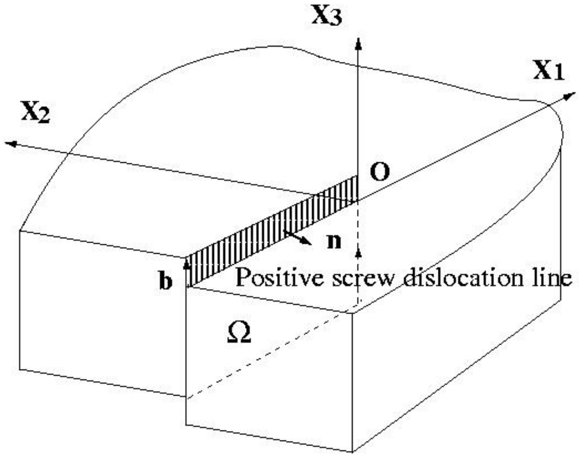


Figure 5.2. A screw dislocation

$$\begin{aligned} C_{j\ell 23} &= \lambda\delta_{j\ell}\delta_{23} + \mu(\delta_{j2}\delta_{\ell 3} + \delta_{j3}\delta_{\ell 2}) \\ &= \mu(\delta_{j2}\delta_{\ell 3} + \delta_{j3}\delta_{\ell 2}) \end{aligned}$$

The only non-zero components are $C_{2323} = \mu$ and $C_{3223} = \mu$. Therefore,

$$\begin{aligned} u_1(\mathbf{x}) &= \Big(\frac{b}{(2\pi)^2}\Big)\int_{-\infty}^{\infty}\Big(\frac{\delta(\xi_3)}{\xi_1}\Big)\Big(C_{2323}Q_{12}(\boldsymbol{\xi})\xi_3 + C_{3223}Q_{13}(\boldsymbol{\xi})\xi_2\Big) \\ &\quad \exp(i\xi\cdot\mathbf{x})d\boldsymbol{\xi} \\ u_2(\mathbf{x}) &= \Big(\frac{b}{(2\pi)^2}\Big)\int_{-\infty}^{\infty}\Big(\frac{\delta(\xi_3)}{\xi_1}\Big)\Big(C_{2323}Q_{22}(\boldsymbol{\xi})\xi_3 + C_{3223}Q_{23}(\boldsymbol{\xi})\xi_2\Big) \\ &\quad \exp(i\xi\cdot\mathbf{x})d\boldsymbol{\xi} \\ u_3(\mathbf{x}) &= \Big(\frac{b}{(2\pi)^2}\Big)\int_{-\infty}^{\infty}\Big(\frac{\delta(\xi_3)}{\xi_1}\Big)\Big(C_{2323}Q_{32}(\boldsymbol{\xi})\xi_3 + C_{3223}Q_{33}(\boldsymbol{\xi})\xi_2\Big) \\ &\quad \exp(i\xi\cdot\mathbf{x})d\boldsymbol{\xi} \end{aligned}$$

in which,

$$\begin{aligned} Q_{12}(\boldsymbol{\xi}) &= -\frac{(\lambda+\mu)}{\mu(\lambda+2\mu)}\frac{\xi_1\xi_2}{\xi^4} \\ Q_{22}(\boldsymbol{\xi}) &= \frac{[(\lambda+2\mu)\xi^2 - (\lambda+\mu)\xi_2^2]}{\mu(\lambda+2\mu)\xi^4} \\ Q_{13}(\boldsymbol{\xi}) &= -\frac{(\lambda+\mu)}{\mu(\lambda+2\mu)}\frac{\xi_1\xi_3}{\xi^4} \end{aligned}$$

$$\begin{aligned}
Q_{23}(\boldsymbol{\xi}) &= -\frac{(\lambda+\mu)}{\mu(\lambda+2\mu)}\frac{\xi_2\xi_3}{\xi^4}\\
Q_{32}(\boldsymbol{\xi}) &= Q_{23}(\boldsymbol{\xi})\\
Q_{22}(\boldsymbol{\xi}) &= \frac{[(\lambda+2\mu)\xi^2-(\lambda+\mu)\xi_3^2]}{\mu(\lambda+2\mu)\xi^4}
\end{aligned} \tag{5.54}$$

Obviously,

$$\begin{aligned}
\int_{-\infty}^{\infty}\delta(\xi_3)Q_{12}(\boldsymbol{\xi})\xi_3 d\xi_3 &= 0\\
\int_{-\infty}^{\infty}\delta(\xi_3)Q_{13}(\boldsymbol{\xi})\xi_2 d\xi_3 &= 0\\
\int_{-\infty}^{\infty}\delta(\xi_3)Q_{22}(\boldsymbol{\xi})\xi_3 d\xi_3 &= 0\\
\int_{-\infty}^{\infty}\delta(\xi_3)Q_{23}(\boldsymbol{\xi})\xi_2 d\xi_3 &= 0\\
\int_{-\infty}^{\infty}\delta(\xi_3)Q_{32}(\boldsymbol{\xi})\xi_3 d\xi_3 &= 0\\
\int_{-\infty}^{\infty}\delta(\xi_3)Q_{33}(\boldsymbol{\xi})\xi_2 d\xi_3 &= \frac{1}{\mu}\frac{\xi_2}{(\xi_1^2+\xi_2^2)}
\end{aligned}$$

Thereby, $u_1(\mathbf{x}) = u_2(\mathbf{x}) = 0$, and

$$\begin{aligned}
u_3(\mathbf{x}) &= \frac{b}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\xi_2}{\xi_1(\xi_1^2+\xi_2^2)}\exp\Big(i(\xi_1x_1+\xi_2x_2)\Big)d\xi_1 d\xi_2\\
&= \frac{b}{\pi}\tan^{-1}\Big(\frac{x_2}{x_1}\Big)
\end{aligned} \tag{5.55}$$

according to the inverse Fourier transform (Mura's book page 17),

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\xi_2}{\xi_1(\xi_1^2+\xi_2^2)}\exp\Big(i(\xi_1x_1+\xi_2x_2)\Big)d\xi_1 d\xi_2 = 2\pi\tan^{-1}\Big(\frac{x_2}{x_1}\Big)$$

5.4 Application II: Stress intensity factor for a flat ellipsoidal crack

In late 1960s, John Willis used eigenstrain method solving a class of crack and contact problems in anisotropic space.

In the following, we illustrate Willis' solution procedure in the case of a 3D ellipsoidal crack in an isotropic space.

Consider an ellisoidal crack embbeded in an infinite space. Suppose that the crack region Ω is:

$$\Omega:\ \frac{x_1^2}{a_1^2}+\frac{x_2^2}{a_2^2}\le 1,\quad \text{and}\ \ x_3 = 0\,. \tag{5.56}$$

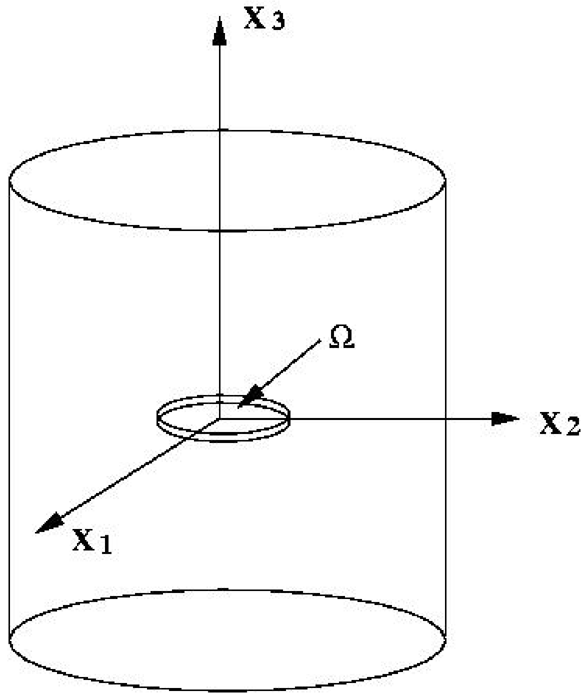


Figure 5.3. A three-dimensional ellipsoidal crack

For simplicity, we assume that the crack opening has the following form:

$$[u_3] = b\sqrt{1 - \frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2}}\chi(\Omega) \tag{5.57}$$

where parameter b is the Burger's vecter, and $\chi(\Omega)$ is the characteristic function of crack region, which can be defined as interpreted as

$$\chi(\Omega) = H(\Omega - \mathbf{x}) = \begin{cases} 1, & \forall \mathbf{x} \in \Omega \\ 0, & \forall \mathbf{x} \in \mathbb{R}^3/\Omega \end{cases} \tag{5.58}$$

where $H(\cdot)$ is the Heavyside function.

This is equivalent to prescrib the following eigenstrain on the crack region,

$$\epsilon_{33}^* = b\sqrt{1 - \frac{x_1^2}{a_1^2} - \frac{x_2^2}{a_2^2}}\delta(\Omega - \mathbf{x}) \ . \tag{5.59}$$

Therefore,

$$\int\int\int_{-\infty}^{\infty} \epsilon^*(\mathbf{x}') \exp\Big(-i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{x}')\Big) d\mathbf{x}^{'} =$$
$$\int\int_{\Omega} b\sqrt{1-\frac{x_1'^2}{a_1^2}-\frac{x_2'^2}{a_2^2}} \exp(-i\xi_3 x_3 - i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{x}')\Big) dx_1' dx_2' \quad (5.60)$$

where in the second line, all vectors become 2D vectors, i.e. $\boldsymbol{\xi} = \xi_1\mathbf{e}_1 + \xi_2\mathbf{e}_2$ and $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$.

Employ the fundamental formula of micro-elasticity,

$$\begin{aligned} u_i(\mathbf{x}) &= \frac{i}{(2\pi)^3}\int_{-\infty}^{\infty}\Big\{\int_{-\infty}^{\infty} C_{j\ell mn}\epsilon^*(\mathbf{x}')\xi_\ell N_{ij}(\boldsymbol{\xi})D^{-1}(\boldsymbol{\xi}) \\ &\quad \exp\Big(-i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{x}')\Big)d\boldsymbol{\xi}\Big\}\, d\mathbf{x}' \end{aligned} \quad (5.61)$$

Changing the dummy indices $i \to k, j \to p, m \to 3, n \to 3, \ell \to q$, we have

$$\begin{aligned} u_k(x_1,x_2,0) &= \frac{ib}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{\Omega} C_{pq33}\sqrt{1-\frac{x'^2}{a_1^2}-\frac{x_2'^2}{a_2^2}}\,\frac{\xi_q N_{kp}(\boldsymbol{\xi})}{D(\boldsymbol{\xi})} \\ &\quad \cdot\exp\Big(-i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{x}^{'})d\Omega_{x'}d\boldsymbol{\xi} \end{aligned} \quad (5.62)$$

and

$$\begin{aligned} u_{k,\ell}(x_1,x_2,0) &= \frac{b}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{\Omega} C_{pq33}\sqrt{1-\frac{x'^2}{a_1^2}-\frac{x_2'^2}{a_2^2}}\,\frac{\xi_q\xi_\ell N_{kp}(\boldsymbol{\xi})}{D(\boldsymbol{\xi})} \\ &\quad \cdot\exp\Big(-i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{x}^{'})d\Omega_{x'}d\boldsymbol{\xi} \end{aligned} \quad (5.63)$$

subsequently,

$$\begin{aligned} \sigma_{ij} &= C_{ijk\ell}u_{k,\ell} = \frac{b}{(2\pi)^3}\int_{-\infty}^{\infty}\frac{C_{ijk\ell}C_{pq33}\xi_q\xi_\ell N_{kp}(\boldsymbol{\xi})}{D(\boldsymbol{\xi})} \\ &\quad \int_{\Omega}\sqrt{1-\frac{x'^2}{a_1^2}+\frac{x'^2}{a_2^2}}\exp\Big(-i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{x}^{'})\Big)dx_1^{'}dx_2^{'} \end{aligned} \quad (5.64)$$

We first calculate the inverse Fourier transform along ξ_3, i.e. evaluating the following integral,

$$\int_{-\infty}^{\infty}\frac{C_{ijk\ell}C_{pq33}\xi_q\xi_\ell N_{kp}(\boldsymbol{\xi})}{D(\boldsymbol{\xi})}\exp(-\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{x}^{'})d\xi_3\,. \quad (5.65)$$

For isotropic materials,

$$\frac{N_{kp}(\boldsymbol{\xi})}{D(\boldsymbol{\xi})} = \frac{[(\lambda+2\mu)\delta_{kp}\xi^2 - (\lambda+\mu)\xi_k\xi_p]}{\mu(\lambda+2\mu)\xi^4} \quad (5.66)$$

where the denominator may be decomposed into

$$\xi^4 = \left(\xi_1^2 + \xi_2^2 + \xi_3^2\right)^2 = \left(\xi_3 - i\sqrt{\xi_1^2 + \xi_2^2}\right)^2 \left(\xi_3 + i\sqrt{\xi_1^2 + \xi_2^2}\right)^2 \quad (5.67)$$

Since the problem is symmetric, we only consider the upper halp space $(x_3 > 0)$. Because the convergence requirement of Fourier transform, we are only interested in the root with a negative imaginary part, i.e.

$$\xi_3^N = -i\sqrt{\xi_1^2 + \xi_2^2} \quad (5.68)$$

which is a double root as shown in Eq. (5.67).

Suppose z_j is a n-th pole of $f(z)$, its residue is then

$$Residue\ at\ (z = z_j) = \frac{1}{(n-1)!} \lim_{z \to z_j} \frac{d^n}{dz^{n-1}} \Big[(z - z_j)^n f(z)\Big] \quad (5.69)$$

Therefore, the integrand inside (5.65) is

$$F_{ijm} = C_{ijk\ell} C_{pq33} \frac{\partial}{\partial \xi_3} \left\{ (\xi_3 - \xi_3^N)^2 \frac{\xi_q \xi_\ell N_{kp}(\boldsymbol{\xi})}{D(\boldsymbol{\xi})} \exp(-i\boldsymbol{\xi} \cdot (\mathbf{x} - \mathbf{x}^{'}) \right\} \quad (5.70)$$

After some tedious calculation, we find that at $x_3 = 0$,

$$F_{333} = -i \frac{\mu(\lambda + \mu)}{(\lambda + 2\mu)} \sqrt{\xi_1^2 + \xi_2^2}\,. \quad (5.71)$$

Hence,

$$\sigma_{33}(x_1, x_2, 0) = -\frac{b\mu(\lambda + \mu)}{4\pi^2(\lambda + 2\mu)} \int\int_\Omega \sqrt{1 - \frac{x_1'^2}{a_1^2} - \frac{x_2'^2}{a_2^2}}$$
$$\left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \xi \exp(-i\boldsymbol{\xi} \cdot (\mathbf{x} - \mathbf{x}') d\xi_1 d\xi_2 \right\} dx_1' dx_2' \quad (5.72)$$

where $\xi = \sqrt{\xi_1^2 + \xi_2^2}$.

Let $y_1 = x_1/a_1, y_2 = x_2/a_2$; $\zeta_1 = a_1\xi_1, \zeta_2 = a_2\xi_2$; and $\eta_1 = \zeta_1/\zeta, \eta_2 = \zeta_2/\zeta$, where $\zeta = \sqrt{\zeta_1^2 + \zeta_2^2}$. Then

$$\boldsymbol{\xi} \cdot (\mathbf{x} - \mathbf{x}') = \boldsymbol{\zeta} \cdot (\mathbf{y} - \mathbf{y}') \quad (5.73)$$

$$dx_1' dx_2' d\xi_1 d\xi_2 = dy_1' dy_2' d\zeta_1 d\zeta_2 \quad (5.74)$$

$$\sqrt{1 - \frac{x_1'^2}{a_1^2} - \frac{x_2'^2}{a_2^2}} = \sqrt{1 - y_1'^2 - y_2'^2} = \sqrt{1 - y'^2} \quad (5.75)$$

$$\xi = \sqrt{\xi_1^2 + \xi_2^2} = \zeta \sqrt{\frac{\eta_1^2}{a_1^2} + \frac{\eta_2^2}{a_2^2}} \quad (5.76)$$

Thus in Eq. (5.72)

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\xi\exp(-i\boldsymbol{\xi}\cdot(\mathbf{x}-\mathbf{x}')d\xi_1 d\xi_2\\
=&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\zeta\sqrt{\frac{\eta_1^2}{a_1^2}+\frac{\eta_2^2}{a_2^2}}\exp\Big(-i\zeta\boldsymbol{\eta}\cdot(\mathbf{y}-\mathbf{y}')\Big)d\zeta_1 d\zeta_2\\
=&\int_0^{2\pi}\int_0^{\infty}\zeta^2\sqrt{\Big(\frac{\zeta_1}{a_1}\Big)^2+\Big(\frac{\zeta_2}{a_2}\Big)^2}\exp\Big(-i\zeta\boldsymbol{\eta}\cdot(\mathbf{y}-\mathbf{y}^{'})\Big)d\zeta d\phi \quad (5.77)
\end{aligned}$$

Denote $g=-\boldsymbol{\eta}\cdot\mathbf{y}$. The above integral becomes

$$\begin{aligned}
&\int_0^{2\pi}\int_0^{\infty}\zeta^2\sqrt{\Big(\frac{\zeta_1}{a_1}\Big)^2+\Big(\frac{\zeta_2}{a_2}\Big)^2}\exp\Big(-i\zeta\boldsymbol{\eta}\cdot(\mathbf{y}-\mathbf{y}^{'})\Big)d\zeta d\phi\\
=&\ -\frac{\partial^2}{\partial g^2}\int_0^{2\pi}\sqrt{\Big(\frac{\zeta_1}{a_1}\Big)^2+\Big(\frac{\zeta_2}{a_2}\Big)^2}\exp(i\zeta(g+\boldsymbol{\eta}\cdot\mathbf{y}')d\eta d\phi\\
=&\ -\frac{\partial^2}{\partial g^2}\int_0^{2\pi}\sqrt{\Big(\frac{\zeta_1}{a_1}\Big)^2+\Big(\frac{\zeta_2}{a_2}\Big)^2}\left\{\Big(-\frac{\partial^2}{\partial g^2}\Big)\int_0^{\infty}\exp(i\zeta(g+\boldsymbol{\eta}\cdot\mathbf{y}')d\zeta\right\}d\phi\\
=&\int_0^{2\pi}\sqrt{\Big(\frac{\zeta_1}{a_1}\Big)^2+\Big(\frac{\zeta_2}{a_2}\Big)^2}\Big(\frac{\partial^2}{\partial g^2}\frac{-i}{g+\boldsymbol{\eta}\cdot\mathbf{y}^{'}}\Big)d\phi \qquad (5.78)
\end{aligned}$$

Denote $\boldsymbol{\eta}\cdot\mathbf{y}'=y'\cos(\theta-\phi)$ and consider following integral identity,

$$\int_0^{2\pi}\frac{d(\theta-\phi)}{g+y'\cos(\theta-\phi)}=\frac{2\pi}{\sqrt{g^2-y'^2}}\ . \qquad (5.79)$$

$$\begin{aligned}
\sigma_{33}\Big|_{x_3=0} &= \frac{ib\mu(\lambda+\mu)}{2\pi(\lambda+\mu)}\int_0^{2\pi}\sqrt{\Big(\frac{\eta_1}{a_1}\Big)^2+\Big(\frac{\eta_2}{a_2}\Big)^2}\left\{\frac{\partial^2}{\partial g^2}\int_\Omega\frac{y'\sqrt{1-y'^2}dy^{'}}{g+y^{'}\cos(\theta-\phi)}dy_1^{'}dy_2^{'}\right\}\\
&= \frac{ib\mu(\lambda+\mu)}{2\pi(\lambda+\mu)}\int_0^{2\pi}\sqrt{\Big(\frac{\eta_1}{a_1}\Big)^2+\Big(\frac{\eta_2}{a_2}\Big)^2}\left\{\frac{\partial^2}{\partial g^2}\int_0^1\frac{y'\sqrt{1-y'^2}dy^{'}}{\sqrt{g^2-y'^2}}\right\} \qquad (5.80)
\end{aligned}$$

Let

$$I=\int_0^1\frac{y'\sqrt{1-y'^2}dy^{'}}{\sqrt{g^2-y'^2}}$$

Change of variable

$$y^{'2}=1-\frac{g^2-1}{4}\Big(w-\frac{1}{w}\Big)^2 \qquad (5.81)$$

One can show that

$$\frac{\partial^2 I}{\partial g^2}=-\frac{1}{2}\ln\frac{g+1}{g-1}+\frac{g}{g^2-1}\ . \qquad (5.82)$$

Interior solution ($y < 1$):
When $y < 1$ $x_3 = 0$, it is crack region. Obviously $|g| = |\boldsymbol{\eta} \cdot \mathbf{y}| < 1$. Since

$$\frac{g+1}{g-1} = -\Big(\frac{1+g}{1-g}\Big) = \exp(-i\pi)\Big(\frac{1+g}{1-g}\Big)$$

then,

$$\frac{\partial^2 I}{\partial g^2} = -\frac{1}{2}\ln\Big|\frac{1+g}{1-g}\Big| - i\frac{\pi}{2} + \frac{g}{g^2-1}\ . \tag{5.83}$$

Both $\ln|(1+g)/(1-g)|$ and $g/(g^2-1)$ are odd function of ϕ, whereas $\Big(\cos^2\phi/a_1^2 + \sin^2\phi/a_2^2\Big)^{1/2}$ is an even function of ϕ.

Hence when $y < 1$

$$\begin{aligned}\sigma_{33}(x_1, x_2, 0) &= -\frac{b\mu(\lambda+\mu)}{4(\lambda+2\mu)}\int_0^{2\pi}\Big(\frac{\cos^2\phi}{a_1^2} + \frac{\sin^2\phi}{a_2^2}\Big)^{1/2} d\phi \\ &= -\frac{b\mu E(k)}{2a_2(1-\nu)}\end{aligned} \tag{5.84}$$

where

$$E(k) = \int_0^{\pi/2}(1 - k^2\sin^2\phi)^{1/2} d\phi,\ \ k^2 := \frac{a_1^2 - a_2^2}{a_1^2}\ . \tag{5.85}$$

If

$$\sigma_{33}(\Omega) = -\sigma_{33}^0 = -\frac{b\mu E(k)}{2a_2(1-\nu)} \tag{5.86}$$

it then links the Burgers' vector with the prescribed stress on the crack surfaces,

$$b = \frac{2(1-\nu)a_2\sigma_{33}^0}{\mu E(k)}\ . \tag{5.87}$$

This suggests that the type of prescribed eigestrain is equivalent to prescribed constant stress on crack surfaces.

Exterior solution:
We are only interested the asymptotic solution, i.e. $y \to 1$. When $y \to 1$, the term $|g/(g^2-1)| > ln|(g+1)/(g-1)| \to \infty$ is the leading term of asymptotic expansion. Therefore

$$\sigma_{33}(x_1, x_2, 0) = \frac{ib\mu(\lambda+\mu)}{2\pi(\lambda+\mu)}\int_0^{2\pi}\sqrt{\Big(\frac{\eta_1}{a_1}\Big)^2 + \Big(\frac{\eta_2}{a_2}\Big)^2}\,\frac{g d\phi}{g^2-1} + \mathcal{O}(1) \tag{5.88}$$

Let

$$f(\boldsymbol{\eta}) = g\Big(\frac{\eta_1^2}{a_1^2} + \frac{\eta_2^2}{a_2^2}\Big)^{1/2},\ \text{ and } \hat{\mathbf{y}} = \frac{\mathbf{y}}{y}\ . \tag{5.89}$$

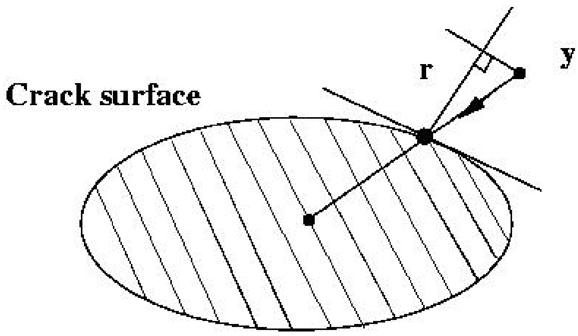


Figure 5.4. The shortest distance between the crack surface and a point

$$
\begin{aligned}
\sigma_{33}(x_1,x_2,0) &= \frac{ib\mu(\lambda+\mu)}{2\pi(\lambda+\mu)}\int_0^{2\pi}\frac{f(\boldsymbol{\eta}-f(\hat{\mathbf{y}})}{g^2-1}d\phi+\frac{1}{2\pi}\int_0^{2\pi}\frac{f(\hat{\mathbf{y}})}{g^2-1}d\phi \\
&= \frac{1}{2\pi}\int_0^{2\pi}\frac{f(\hat{\mathbf{y}})}{g^2-1}d\phi+\mathcal{O}(1) \qquad (5.90)
\end{aligned}
$$

Assume that $g=-\boldsymbol{\eta}\cdot\mathbf{y}=y\cos\psi$. Then

$$
\int_0^{2\pi}\frac{d\phi}{g^2-1}\int_0^{2\pi}\frac{d(\phi-\psi)}{y^2\cos^2(\phi-\psi)-1}=\frac{-2\pi}{\sqrt{1-y^2}}=\frac{2i\pi}{\sqrt{y^2-1}}\,. \qquad (5.91)
$$

and

$$
\sigma_{33}(x_1,x_2,0) = \frac{b\mu(\lambda+\mu)}{(\lambda+2\mu)}\frac{\hat{\mathbf{y}}\cdot\mathbf{y}}{\sqrt{y^2-1}}\left(\frac{\hat{y}_1}{a_1^2}+\frac{\hat{y}_2}{a_2^2}\right)^{1/2}\Big|_{y\to\hat{y}} \qquad (5.92)
$$

The stress intensity factor is defined as

$$
k_1 := \lim_{r\to 0}(2\pi r)^{1/2}\sigma_{33} \qquad (5.93)
$$

For an ellipsoidal crack,

$$
r=\frac{(y-1)y^2}{\left(\dfrac{x_1^2}{a_1^2}+\dfrac{x_2^2}{a_2^2}\right)^{1/2}} \qquad (5.94)
$$

and

$$
k_1 = \frac{\sqrt{2\pi}b\mu(\lambda+\mu)}{(\lambda+2\mu)}\frac{\sqrt{y-1}}{\sqrt{y^2-1}}\left(\frac{x_1^2}{a_1^4}+\frac{x_2^2}{a_2^4}\right)^{1/4},\quad y\to 1 \qquad (5.95)
$$

Substituting $b=(2(1-\nu)a_2\sigma_{33}^0/(\mu E(k))$ into the above expression, one has

$$
k_1=\frac{\sqrt{\pi}a_2\sigma_{33}^0}{E(k)}\left(\frac{x_1^2}{a_1^4}+\frac{x_2^2}{a_2^4}\right)^{1/4}. \qquad (5.96)
$$

5.5 Isotropic inclusion-Eshelby's solution

From 1957 to 1961, J.D.Eshelby published three landmark scientific papers systematically solving inclusion problem in an elastic medium.

Eshelby's ellipsoidal inclusion problem is stated as follows: *Find induced displacement, strain, and stress fields by an ellipsoidal incluseion, Ω, embedded in an isotropic unbounded elastic medium, in which a uniform eigenstrain is prescribed, i.e.*

$$\epsilon^*{}_{ij}(\mathbf{x}) = \begin{cases} \epsilon^*{}_{ij}, & \mathbf{x} \in \Omega \\ 0, & \mathbf{x} \in \mathbb{R}^3/\Omega \end{cases} \tag{5.97}$$

Using the fundamental formula of micro-elasticity,

$$\begin{aligned} u_i(\mathbf{x}) &= -\int_{-\infty}^{\infty} C_{j\ell mn}\epsilon^*{}_{mn}(\mathbf{y})G^{\infty}_{ij,\ell}(\mathbf{x}-\mathbf{y})d\Omega_{\mathbf{y}} \\ &= -\epsilon^*{}_{mn}\int_{\Omega} C_{j\ell mn}G^{\infty}_{ij,\ell}(\mathbf{x}-\mathbf{y})d\Omega_{\mathbf{y}} \end{aligned}$$

For isotropic elastic materials,

$$\begin{aligned} C_{j\ell mn}G^{\infty}_{ij,\ell}(\mathbf{z}) &= \frac{-1}{8\pi(1-\nu)}\Big\{(1-2\nu)\frac{\delta_{mi}z_n+\delta_{ni}z_m-\delta_{mn}z_i}{z^3} + 3\frac{z_m z_n z_i}{z^5}\Big\} \\ &= \frac{g_{imn}(\boldsymbol{\ell})}{8\pi(1-\nu)|\mathbf{z}|^2} \end{aligned} \tag{5.98}$$

where $\mathbf{z} = \mathbf{x} - \mathbf{y}$ and $\boldsymbol{\ell} = -\mathbf{z}/|\mathbf{z}|$, and

$$g_{imn}(\boldsymbol{\ell}) = (1-2\nu)(\delta_{mi}\ell_n + \delta_{ni}\ell_m - \delta_{mn}\ell_i) + 3\ell_m\ell_n\ell_i \tag{5.99}$$

5.5.1 Interior solution

Consider $\mathbf{x} \in \Omega$. Let $z = |\mathbf{z}$. Take a radon decomposition centering around a the point $\mathbf{x}$ $d\Omega_{\mathbf{y}} = dzdS = z^2dzdw$, where dw is the volume angle on S^2. We can rewrite displacement field as

$$\begin{aligned} u_i(\mathbf{x}) &= \frac{-\epsilon^*{}_{mn}}{8\pi(1-\nu)}\int_{\Omega} g_{imn}(\boldsymbol{\ell})\frac{d\Omega_{\mathbf{y}}}{|\mathbf{x}-\mathbf{y}|^2} \\ &= \frac{-\epsilon^*{}_{mn}}{8\pi(1-\nu)}\int_0^r\int_{S^2} g_{imn}(\boldsymbol{\ell})dzdw \\ &= \frac{-\epsilon^*{}_{mn}}{8\pi(1-\nu)}\int_{S^2} r(\boldsymbol{\ell})g_{imn}(\boldsymbol{\ell})dzdw \end{aligned} \tag{5.100}$$

where vector $\mathbf{r} = \mathbf{y} - \mathbf{x}$, $\mathbf{y} \in \partial\Omega$ and the scalar $r(\boldsymbol{\ell})$ is the distance between the point $\mathbf{x}$ and a point $\mathbf{y}$ on the surface of the ellipsoidal in the direction of $\mathbf{r}$. In other words, $r(\boldsymbol{\ell})$ is the distance between $\mathbf{x}$ and the interseption point of straight line $\mathbf{y} = \mathbf{x} + \mathbf{r}$, $\mathbf{y} \in \mathbb{R}^3$ and the surface of the ellipsoidal. To find such

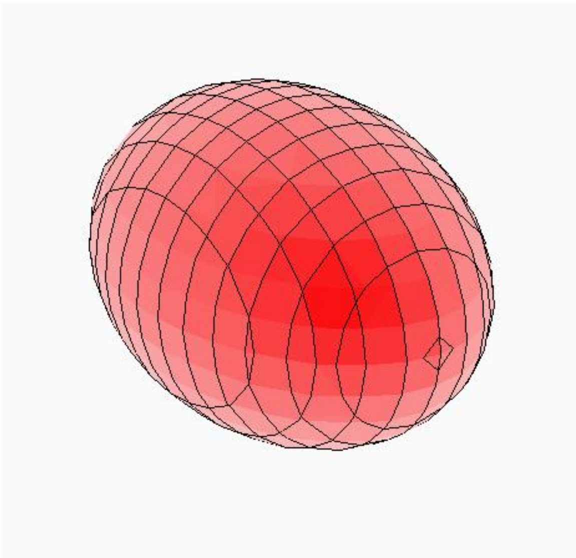


Figure 5.5. An ellipsoidal inclusion

interception point along a fixed direction of $\boldsymbol{\ell}$. We assume that the interception point is marked as $\mathbf{x}'$. Since it must be on both the straight line, $\mathbf{x}' = \mathbf{x} + \mathbf{r}$, i.e.

$$\begin{cases} x_1' = x_1 + r\ell_1 \\ x_2' = x_2 + r\ell_2 \\ x_3' = x_3 + r\ell_3 \end{cases} \tag{5.101}$$

and on the surface of the ellipsoidal

$$\frac{x_1'^2}{a_1^2} + \frac{x_2'^2}{a_2^2} + \frac{x_3'^2}{a_3^2} = 1 \tag{5.102}$$

One can substitute (5.139) into (5.140). For fixed point $\mathbf{x}$ and a fixed direction $\boldsymbol{\ell}$, it yields a quadratic equation,

$$\frac{(x_1 + r\ell_1)^2}{a_1^2} + \frac{(x_2 + r\ell_2)^2}{a_2^2} + \frac{(x_3 + r\ell_3)^2}{a_3^2} = 1 \tag{5.103}$$

of unknown variable, $r(\boldsymbol{\ell})$. More explicitly,

$$\begin{aligned} r^2\Big(\frac{\ell_1^2}{a_1^2} + \frac{\ell_2^2}{a_2^2} + \frac{\ell_3^2}{a_3^2}\Big) + 2r\Big(\frac{x_1\ell_1}{a_1^2} + \frac{x_2\ell_2}{a_2^2} + \frac{x_3\ell_3}{a_3^2}\Big) \\ + \Big[\Big(\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2}\Big) - 1\Big] = 0, \quad \Rightarrow \quad gr^2 + 2rf - e = 0 \end{aligned} \tag{5.104}$$

where

$$g := \left(\frac{\ell_1^2}{a_1^2} + \frac{\ell_2^2}{a_2^2} + \frac{\ell_3^2}{a_3^2}\right) \tag{5.105}$$

$$f := \left(\frac{x_1\ell_1}{a_1^2} + \frac{x_2\ell_2}{a_2^2} + \frac{x_3\ell_3}{a_3^2}\right) \tag{5.106}$$

$$e := 1 - \left(\frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2}\right) \tag{5.107}$$

Eq. (5.142) has two roots,

$$r(\boldsymbol{\ell}) = -\frac{f}{g} \pm \left(\frac{f^2}{g^2} + \frac{e}{g}\right)^{1/2} \tag{5.108}$$

Since $\left(\frac{f^2}{g^2} + \frac{e}{g}\right)^{1/2}$ is even in $\boldsymbol{\ell}$, while $g_{imn}(\boldsymbol{\ell})$ is odd in $\boldsymbol{\ell}$,

$$\int_{S^2} \left(\frac{f^2}{g^2} + \frac{e}{g}\right)^{1/2} g_{imn}(\boldsymbol{\ell}) dw = 0 \tag{5.109}$$

Let $\lambda_1 = \ell_1/a_1^2$,$\lambda_2 = \ell_2/a_2^2$ and $\lambda_3 = \ell_3/a_3^2$. We have

$$\begin{aligned} u_i(\mathbf{x}) &= \frac{\epsilon^*{}_{mn}}{8\pi(1-\nu)} \oint_{S^2} \frac{f}{g} g_{imn}(\boldsymbol{\ell}) dw \\ &= \frac{\epsilon^*{}_{mn}}{8\pi(1-\nu)} \oint_{S^2} \left(\frac{x_\ell \lambda_l}{g}\right) g_{imn}(\boldsymbol{\ell}) dw \\ &= \frac{\epsilon^*{}_{mn} x_\ell}{8\pi(1-\nu)} \oint_{S^2} \left(\frac{\lambda_\ell}{g}\right) g_{imn}(\boldsymbol{\ell}) dw \end{aligned} \tag{5.110}$$

Then

$$\begin{aligned} u_{i,j}(\mathbf{x}) &= \frac{\epsilon^*{}_{mn}\delta_{\ell j}}{8\pi(1-\nu)} \oint_{S^2} \left(\frac{\lambda_\ell}{g}\right) g_{imn}(\boldsymbol{\ell}) dw \\ &= \frac{\epsilon^*{}_{mn}}{8\pi(1-\nu)} \oint_{S^2} \left(\frac{\lambda_j}{g}\right) g_{imn}(\boldsymbol{\ell}) dw \end{aligned} \tag{5.111}$$

One can find induced elastic strain field by symmetrizing the elastic distortion,

$$\epsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i}) = \frac{\epsilon^*{}_{mn}}{16\pi(1-\nu)} \oint_{S^2} \frac{\lambda_i g_{jmn} + \lambda_j g_{imn}}{g} dw \tag{5.112}$$

where $\lambda_i = \dfrac{\ell_i}{a_i^2}$ is the component of the normalized vector $\boldsymbol{\lambda} = \lambda_i \mathbf{e}_i$ and $g = \boldsymbol{\lambda} \cdot \boldsymbol{\lambda} = \lambda^2$.

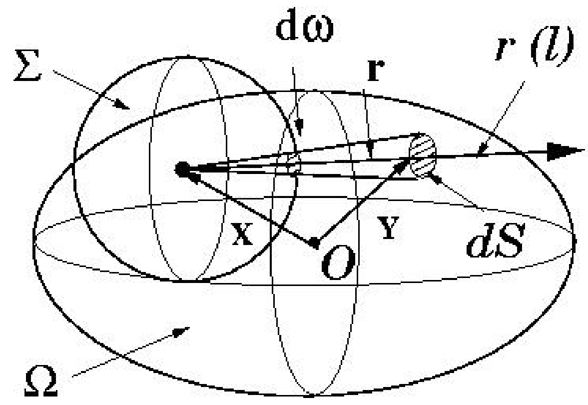


Figure 5.6. Illustration of integration scheme over an ellipsoidal

Consider

$$g_{ijk}(\boldsymbol{\ell}) = (1-2\nu)(\delta_{ij} l_k + \delta_{i\ell}\ell_k - \delta_{jk}\ell_i) + 3\ell_i\ell_j\ell_k = g_{ik\ell}(\boldsymbol{\ell}) \tag{5.113}$$

The last two indices of the third order tensor g_{ijk} is symmetric. We can then define a fourth order symmetric tensor,

$$S^{\Omega}_{ijmn} := \frac{1}{16\pi(1-\nu)} \oint_{S^2} \frac{\lambda_i g_{jmn} + \lambda_j g_{imn}}{g} dw \tag{5.114}$$

This leads to the long anticipated result,

$$\epsilon_{ij}(\mathbf{x}) = \Big(\text{or } \epsilon^d_{ij}(\mathbf{x})\Big) = S^{\Omega}_{ijmn} \epsilon^*{}_{mn} \tag{5.115}$$

It is obvious that

$$S^{\Omega}_{ijmn} = S^{\Omega}_{ijnm} = S^{\Omega}_{jimn}$$

where the superscript indicates that the Eshelby tensor is for induced strain field inside the ellipsoidal, Ω.

REMARK 5.5.1 *The most amazing fact of this result is that the induced strain field and stress field inside the inclusion are uniform, and the Eshelby tensor for any ellipsoidal shape of inclusion is a constant tensor.*

Define the following elliptic integrals

$$
\begin{aligned}
I_I(0) &= \int_{S^2} \frac{\ell_i^2 dw}{a_i^2 g} = 2\pi a_1 a_2 a_3 \int_0^\infty \frac{ds}{(a_I^2+s)\Delta(s)} && (5.116)\\
I_{IJ}(0) &= 3\int_{S^2} \frac{\ell_i^2 l_j^2 dw}{a_i^2 a_j^2 g}\\
&= 2\pi a_1 a_2 a_3 \int_0^\infty \frac{ds}{(a_I^2+s)(a_J^2+s)\Delta(s)} && (5.117)\\
J_{IJ}(0) &= a_I^2 I_{IJ} - I_J && (5.118)
\end{aligned}
$$

where $\Delta(s) = \sqrt{(a_1^2+s)(a_2^2+s)(a_3^2+s)}$ and argument (0) indicating the lower limit of the elliptic integrals are zero.

One can show that Eshelby tensor can be explicitly expressed by these integrals through the following identity,

$$
\begin{aligned}
8\pi(1-\nu)S^{\Omega}_{ijk\ell} &= \delta_{ij}\delta_{k\ell}(2\nu I_I(0) + J_{IK}(0)) + (\delta_{ik}\delta_{k\ell} + \delta_{jk}\delta_{i\ell})\\
&\quad \Big((1-\nu)(I_k(0) + I_L(0)) + J_{IJ}(0)\Big) && (5.119)
\end{aligned}
$$

where the upper case indices are not summed with lower case indices.

EXAMPLE 5.2 *To compute* S^{Ω}_{1111}, *we consider*

$$
\begin{aligned}
8\pi(1-\nu)S^{\Omega}_{1111} &= 2\nu I_1(0) + J_{11}(0) + 2(1-\nu)2I_1(0) + 2J_{11}(0)\\
&= (4-2\nu)I_1(0) + 3J_{11}(0)(a_1^2 I_{11}(0) - I_1(0))\\
&= (1-2\nu)I_1(0) + 3a_1^2 I_{11}(0) && (5.120)
\end{aligned}
$$

which leads to

$$
S^{\Omega}_{1111} = \frac{3a_1^2}{8\pi(1-\nu)} I_{11}(0) + \frac{(1-2\nu)}{8\pi(1-\nu)} I_1(0) \qquad (5.121)
$$

The integral $I_I(0)$ and $I_{IJ}(0)$ can be expressed in terms of standard elliptic integrals. For example, assuming $a_1 > a_2 > a_3$, we have

$$
\begin{aligned}
I_1(0) &= \frac{4\pi a_1 a_2 a_3}{(a_1^2-a_2^2)(a_1^2-a_2^2)^{1/2}}\{F(\theta,k) - E(\theta,k)\}\\
I_3(0) &= \frac{4\pi a_1 a_2 a_3}{(a_2^2-a_3^2)(a_1^2-a_3^2)^{1/2}}\left\{\frac{a_2(a_1^2-a_3^2)^{1/2}}{a_1 a_3} - E(\theta,k)\right\}
\end{aligned}
$$

where

$$
\begin{aligned}
F(\theta,k) &= \int_0^\theta \frac{dt}{(1-k^2\sin^2 t)^{1/2}}\\
E(\theta,k) &= \int_0^\theta (1-k^2\sin^2 t)^{1/2} dt && (5.122)
\end{aligned}
$$

and $\theta = \sin^{-1}(1 - a_3^2/a_1^2)^{1/2}$, $k = \Big[(a_1^2 - a_2^2)/(a_1^2 - a_3^2)\Big]^{1/2}$.

In applications, the following invariant formulas are very useful,

$$\begin{aligned}
&I_1(0) + I_2(0) + I_3(0) = 4\pi \\
&3I_{11}(0) + I_{12}(0) + I_{13}(0) = 4\pi/a_1^2 \\
&3a_1^2 I_{11}(0) + a_2^2 I_{12}(0) + a_3^2 I_{13}(0) = 3I_1 \\
&I_{12}(0) = (I_2(0) - I_1(0))/(a_1^2 - a_2^2)
\end{aligned}$$

When the ellipsoidal becomes a sphere, Eshelby tensor become simple numbers. Let $a_1 = a_2 = a_3 = a$. We have

$$\begin{aligned}
I_I^s &= \frac{4\pi}{3} \\
I_{I,J}^s &= \frac{4\pi}{5}\frac{1}{a^2} \\
J_{IJ}^s &= -\frac{8\pi}{15}
\end{aligned}$$

and hence

$$S^{\Omega}_{ijk\ell} = \Big(\frac{5\nu - 1}{15(1-\nu)}\Big)\delta_{ij}\delta_{k\ell} + \frac{2(4-5\nu)}{15(1-\nu)}(\delta_{ik}\delta_{j\ell} + \delta_{jk}\delta_{i\ell}) \tag{5.123}$$

A remarkable property of the Eshelby tensor for spherical inclusion is that it does not depend on its size, i.e. it does not depend on its radius a. This implies that no matter how large or how small spherical inclusions are, they share the same Eshlby tensor. In other words, there is no embeded length scale or scaling factor for spherical inclusion. This property will lead to some remarkable consequences in ensuing homoginization process.

For other specified shape of ellipsoidal inclusions, readers may consult Mura's book for detailed information. A systematic documentation on Eshelby's tensor in various cases can be found in Mura [1987].

5.6 Exterior Solution of Ellipsoidal Inclusion

For $\mathbf{x} \notin \Omega$, the exterior disturbance displacement and strain fields due to eigenstrain distribution had been also found by Eshelby, though evaluation of the induced exterior displacement fields and strain fields are often difficult.

Suppose that eigenstrain distribution inside the ellipsoid is constant. For any point $\mathbf{x} \in \mathbb{R}^3$, we have

$$u_i(\mathbf{x}) = -C_{jkmn}{\epsilon^*}_{mn}\int_{\Omega} G_{ij,k}(\mathbf{x} - \mathbf{x}')d\Omega_{\mathbf{x}'} \tag{5.124}$$

where

$$
\begin{aligned}
C_{j\ell mn}G^{\infty}_{ij,\ell}(\mathbf{x}-\mathbf{x}') &= \frac{-1}{8\pi(1-\nu)}\cdot \\
&\Big\{(1-2\nu)\frac{\delta_{mi}(x_n-x_n')+\delta_{ni}(x_m-x_m')-\delta_{mn}(x_i-x_i')}{|\bar{\mathbf{x}}|^3} \\
&+3\frac{(x_m-x_m')(x_n-x_n')(x_i-x_i')}{|\bar{\mathbf{x}}|^5}\Big\} \\
&= \frac{-1}{8\pi(1-\nu)}\Big\{\frac{\partial^3}{\partial x_i\partial x_m\partial x_n}|\bar{\mathbf{x}}|-2(1-\nu)\Big[\frac{\partial}{\partial x_n}\frac{\delta_{mi}}{|\bar{\mathbf{x}}|} \\
&\frac{\partial}{\partial x_m}\frac{\delta_{ni}}{|\bar{\mathbf{x}}|}\Big]-2\nu\delta_{mn}\frac{\partial}{\partial x_i}\frac{1}{|\bar{\mathbf{x}}|}\Big\}
\end{aligned}
\tag{5.125}
$$

Introduce the following potential functions,

$$
\psi(\mathbf{x}) = \int_\Omega |\mathbf{x}-\mathbf{x}'|d\Omega_{\mathbf{x}'} \tag{5.126}
$$

$$
\phi(\mathbf{x}) = \int_\Omega \frac{1}{|\mathbf{x}-\mathbf{x}'|}d\Omega_{\mathbf{x}'} \tag{5.127}
$$

where $\psi(\mathbf{x})$ is the biharmonic potential, whereas $\phi(\mathbf{x})$ is the Newtonian potential. This is because of the fact

$$
\nabla^4\psi = 2\nabla^2\phi = \begin{cases} -8\pi & \mathbf{x}\in\Omega \\ 0 & \mathbf{x}\in \mathbb{R}^3/\Omega \end{cases} \tag{5.128}
$$

To verify Eq. (5.166), one can show first

$$
\begin{aligned}
\nabla^2\psi &= \frac{\partial^2}{\partial x_\ell^2}\int_\Omega |\mathbf{x}-\mathbf{x}'|d\Omega_{\mathbf{x}'} \\
&= \int_\Omega\Big\{\frac{\delta_{\ell\ell}}{|\bar{\mathbf{x}}|}-\frac{(x_\ell-x_\ell')(x_\ell-x_\ell')}{|\bar{\mathbf{x}}|^3}\Big\}d\Omega_{\mathbf{x}'} \\
&= \int_\Omega \frac{2}{|\bar{\mathbf{x}}|}d\Omega_{\mathbf{x}'} = 2\phi(\mathbf{x})
\end{aligned}
\tag{5.129}
$$

Subsequently,

$$
\begin{aligned}
\nabla^4\psi &= \nabla^2\nabla^2\psi = 2\nabla^2\phi \\
&= 2\int_\Omega \nabla\frac{1}{|\bar{\mathbf{x}}|}d\Omega = 8\pi\int_\Omega\nabla^2\frac{1}{4\pi|\bar{\mathbf{x}}|}d\Omega \\
&= 8\pi\int_\Omega \nabla^2 G^L(\mathbf{x}-\mathbf{x}')d\Omega_{\mathbf{x}'}
\end{aligned}
\tag{5.130}
$$

where $G^L(\mathbf{x}-\mathbf{x}')$ is the Green's function for three-dimensional Laplace equation, i.e.

$$\nabla^2 G^L(\mathbf{x}-\mathbf{x}') + \delta(\mathbf{x}-\mathbf{x}') = 0 \tag{5.131}$$

Consequently,

$$\begin{aligned}\nabla^4\psi &= 2\nabla^2\phi = 8\pi\int_\Omega \delta(\mathbf{x}-\mathbf{x}')d\Omega_{\mathbf{x}'} \\ &= \begin{cases} -8\pi & \mathbf{x}\in\Omega \\ \\ 0 & \mathbf{x}\in\mathbb{R}^3/\Omega \end{cases}\end{aligned} \tag{5.132}$$

We can then express induced displacement as

$$\begin{aligned}u_i(\mathbf{x}) &= -\int_\Omega \epsilon^*{}_{mn}C_{j\ell mn}G_{ij,\ell}(\mathbf{x}-\mathbf{x}')d\Omega_{\mathbf{x}'} \\ &= \frac{\epsilon^*{}_{mn}}{8\pi(1-\nu)}\Big\{\frac{\partial^3}{\partial x_i\partial x_m\partial x_n}\psi - 2(1-\nu)\Big(\delta_{mi}\frac{\partial}{\partial x_n}+\delta_{ni}\frac{\partial}{\partial x_m}\Big)\phi \\ &\quad -2\nu\delta_{mn}\frac{\partial}{\partial x_i}\phi\Big\}\end{aligned} \tag{5.133}$$

Similarily for elastic distortion field and strain field,

$$\begin{aligned}u_{i,j}(\mathbf{x}) &= \frac{\epsilon^*{}_{mn}}{8\pi(1-\nu)}\Big(\psi_{,mnij} - 2(1-\nu)(\delta_{mi}\phi_{,nj}+\delta_{ni}\phi_{,mj}) \\ &\quad -2\nu\delta_{mn}\phi_{,ij}\Big)\end{aligned} \tag{5.134}$$

$$\begin{aligned}\epsilon_{ij}(\mathbf{x}) &= \frac{1}{2}(u_{i,j}+u_{j,i}) = \frac{\epsilon^*{}_{mn}}{8\pi(1-\nu)}\{\psi_{,mnij} - 2\nu\delta_{mn}\phi_{,ij} \\ &- (1-\nu)(\delta_{mi}\phi_{,nj}+\delta_{ni}\phi_{,mj}+\delta_{mj}\phi_{,ni}+\delta_{nj}\phi_{,mi}\}\end{aligned} \tag{5.135}$$

One can rewrite the above expression in a succinct manner,

$$\epsilon^d_{ij}(\mathbf{x}) = S^\infty_{ijk\ell}(\mathbf{x})\epsilon^*{}_{k\ell}, \ \ \forall\mathbf{x}\in\mathbb{R}^3/\Omega \tag{5.136}$$

which defines the exterior Eshelby tensor, $S^\infty_{ijk\ell}(\mathbf{x})$.

$$\begin{aligned}S^\infty_{ijk\ell}(\mathbf{x}) &= \frac{1}{8\pi(1-\nu)}\Big(\psi_{,ijk\ell}(\mathbf{x}) - 2\nu\delta_{k\ell}\phi_{,ij}(\mathbf{x}) \\ &\quad -(1-\nu)(\delta_{ki}\phi_{,\ell j}(\mathbf{x})+\delta_{\ell i}\phi_{,kj}(\mathbf{x}) \\ &\quad +\delta_{kj}\phi,\ell i(\mathbf{x})+\delta_{\ell j}\phi_{,ki}(\mathbf{x}))\Big)\end{aligned} \tag{5.137}$$

It depends on where the tensor is being evaluated.

The derivatives of Newtonian potential and biharmonic potential can be also expressed by elliptic integrals. For instance,

$$\begin{aligned}\phi_{ij}(\mathbf{x}) &= -\delta_{ij} I_I(\lambda) - x_i I_{IJ}(\lambda) && (5.138)\\ \psi_{,ijk\ell}(\mathbf{x}) &= -\delta_{ij}(x_k I_{IK}(\lambda))_{,\ell} + (x_i x_j I_{IJ}(\lambda))_{,k\ell} && (5.139)\end{aligned}$$

where

$$\begin{aligned}I_I(\lambda) &= 2\pi a_1 a_2 a_3 \int_\lambda^\infty \frac{ds}{(a_I^2 + s)\Delta(s)} && (5.140)\\ I_{IJ}(\lambda) &= 2\pi a_1 a_2 a_3 \int_\lambda^\infty \frac{ds}{(a_I^2 + s)(a_J^2 + s)\Delta(s)} && (5.141)\\ J_{IJ}(\lambda) &= a_I^2 I_I J(\lambda) - I_J(\lambda) && (5.142)\end{aligned}$$

where λ is zero when $\mathbf{x} \in \Omega$ and λ is the largest positive root of the following equation,

$$\frac{x_1^2}{(a_1^2 + \lambda)} + \frac{x_2^2}{(a_2^2 + \lambda)} + \frac{x_3^2}{(a_3^2 + \lambda)} = 1 \qquad (5.143)$$

A very useful identity that related $S^\infty_{ijk\ell}(\mathbf{x})$ with elliptic integrals is

$$\begin{aligned}8\pi(1-\nu)S^\infty_{ijk\ell}(\mathbf{x}) &= 8\pi(1-\nu)S^\Omega_{ijk\ell}(\lambda)\\ &+(1-\nu)\Big[\delta_{i\ell} x_k I_{K,j}(\lambda) + \delta_{k\ell} I_{K,i}(\lambda)\\ &+\delta_{ik} I_{L,j}(\lambda) + \delta_{jk} x_\ell I_{L,i}(\lambda)\Big]\\ &\delta_{ij} x_k J_{IK,\ell}(\lambda) + (\delta_{ik} x_j + \delta_{jk} x_i) J_{IJ,\ell}(\lambda)\\ &(\delta_{i\ell} x_j + \delta_{j\ell} x_i) J_{IJ,k}(\lambda)\\ &+x_i x_j J_{IJ,k\ell}(\lambda) && (5.144)\end{aligned}$$

where

$$\begin{aligned}8\pi(1-\nu)S^\Omega_{ijk\ell} &= \delta_{ij}\delta_{k\ell}(2\nu I_I(\lambda) + J_{IK}(\lambda)) + (\delta_{ik}\delta_{k\ell} + \delta_{jk}\delta_{i\ell}) \cdot\\ &\Big((1-\nu)(I_k(\lambda) + I_L(\lambda)) + J_{IJ}(\lambda)\Big) && (5.145)\end{aligned}$$

when $\mathbf{x} \in \Omega$, Eq. (5.144) becomes (5.145). Ju and Chen [1994] developed a more simple and explicit way to evaluate exterior Eshelby tensor. From

$$u_i(\mathbf{x}) = -\int_\Omega C_{jkmn} G_{ij,\ell}(\mathbf{x} - \mathbf{y}) \epsilon^*{}_{mn}(\mathbf{y}) d\Omega_{\mathbf{y}} \qquad (5.146)$$

one may derive that

$$\begin{aligned}\epsilon_{ij}(\mathbf{x}) &= -\frac{1}{2}\int_\Omega C_{k\ell mn}\Big(G_{ik,\ell j}(\mathbf{x} - \mathbf{y}) + G_{jk,\ell i}(\mathbf{x} - \mathbf{y})\Big)\epsilon^*{}_{mn}(\mathbf{y}) d\Omega_{\mathbf{y}}\\ &= \int_\Omega \mathcal{G}_{ijmn}(\mathbf{x} - \mathbf{y})\epsilon^*{}_{mn}(\mathbf{y}) d\Omega_{\mathbf{y}} && (5.147)\end{aligned}$$

where

$$\begin{aligned}
\mathcal{G}_{ijmn}(\mathbf{x}-\mathbf{y}) &= -\frac{1}{2}C_{k\ell mn}\Big(G_{ik,\ell j}(\mathbf{x}-\mathbf{y})+G_{jk,\ell i}(\mathbf{x}-\mathbf{y})\Big)\\
&= \frac{1}{8\pi(1-\nu)r^3}\Big[(1-2\nu)(\delta_{im}\delta_{jn}+\delta_{in}\delta_{jm}-\delta_{ij}\delta_{mn})\\
&\quad +3\nu(\delta_{im}\ell_j\ell_n+\delta_{in}\ell_j\ell_m+\delta_{jm}\ell_i\ell_n+\delta_{jn}\ell_j\ell_m)\\
&\quad +3\delta_{ij}\ell_m\ell_n+3(1-2\nu)\delta_{mn}\ell_i\ell_j-15\ell_i\ell_j\ell_m\ell_n\Big] \qquad (5.148)
\end{aligned}$$

where $\mathcal{G}_{ijmn}$ is called the fourth order Green's function (the second derivative of the Green's function).

If $\epsilon^*{}_{mn}(\mathbf{x})$ is constant inside the inclusion, the exterior Eshlby tensor can be defined as

$$\bar{\mathcal{G}}_{ijmn}(\mathbf{x}) := \int_\Omega \mathcal{G}_{ijmn}(\mathbf{x}-\mathbf{y})d\Omega_{\mathbf{y}} = S^\infty_{ijmn} \qquad (5.149)$$

For a spherical inclusion $(a_1=a_2=a_3=a)$, one may find that

$$\phi = \frac{4\pi a^3}{3|\mathbf{x}|}, \text{ and } \psi = \frac{4\pi a^3}{3}\Big(|\mathbf{x}|+\frac{a^2}{5|\mathbf{x}|}\Big)\,. \qquad (5.150)$$

The exterior Eshelby tensor can then be obtained by straighttfoward differentiation,

$$\begin{aligned}
\bar{\mathcal{G}}_{ijmn}(\mathbf{x}) &= \frac{\rho^3}{30(1-\nu)}\Big[(3\rho^2+10\nu-5)\delta_{ij}\delta_{mn}\\
&\quad +(3\rho^2-10\nu+5)(\delta_{im}\delta_{jn}+\delta_{im}\delta_{jn})\\
&\quad +15(1-\rho^2)\delta_{ij}\ell_m\ell_n+15(1-2\nu-\rho^2)\delta_{mn}\ell_i\ell_j\\
&\quad +15(\nu-\rho^2)(\delta_{im}\ell_j\ell_n+\delta_{in}\ell_j\ell_m+\delta_{jm}\ell_i\ell_n+\delta_{jn}\ell_i\ell_m)\\
&\quad +15(7\rho^2-5)\ell_i\ell_j\ell_m\ell_n\Big] = S^\infty_{ijmn} \qquad (5.151)
\end{aligned}$$

where $\rho = a/r$. Note that when $r \to a$, $S^\infty_{ijmn} \not\to S^\Omega_{ijmn}$, which indicates that both disturbance strain field is not continuous across the interface of the matrix and inclusion.

In fact, when $\rho \to 1$,

$$\begin{aligned}
[\epsilon_{ij}] &:= \epsilon^{[ex]}_{ij} - \epsilon^{[in]}_{ij} = (S^\infty_{ijmn}-S^\Omega_{ijmn})\epsilon^*{}_{ij}\\
&= \frac{-1}{(1-\nu)}\Big[\nu\delta_{mn}\ell_i\ell_j+\frac{1}{2}\Big(\delta_{im}\ell_j\ell_n+\delta_{in}\ell_j\ell_m\\
&\quad +\delta_{jm}\ell_i\ell_n+\delta_{jn}\ell_i\ell_m\Big)-\ell_i\ell_j\ell_m\ell_n\Big]\epsilon_{mn} \qquad (5.152)
\end{aligned}$$

which is the weak discontinuity at the interface between matrix and inculsion.

5.7 Jock Eshelby (II): Lessons from J.D.Eshelby

The measure of your education is what you remember 15 years afterward, says one wiseacre. Well, it's been a little more than 15 years, and I don't think that I learned anything at the time, but the lectures I had from Professor J.D. (Jock) Eshelby still leave a mark.

Undergraduate students in materials science at Sheffield University were barely aware of the towering stature of this man, in the intellectual sense anyway. If you don't know who he was or what contributions he has made, then you probably have some serious holes in your own materials education, but you can still read on. A few Britishisms must be explained, though. First, the term "Jock" is used in the United Kingdom not for an athlete, as in the United States, but is a nickname commonly accorded to Scotsmen living in England; the U.S. sense could never apply to Jock Eshelby. Second the term "Faculty" in England is equivalent to a college in a U.S. University. Third, a professorship in the United Kingdom is a distinguished academic rank that has almost no equivalent in the United States. The closest would be a "leading professorship".

Way back then, Sheffield had a Faculty of Materials, with departments of Metallurgy, Ceramics, Glasses, Polymers, and the theory of materials. The department of the theory of materials was arguably a little top heavy. It had two professors, Eshelby and B.A.Bilby (whose name you should also know), one other lecturer, and a computer programmer. In a good year it had one undergraduate student.

Eselby taught courses in elasticity and solid state bonding to the undergraduates in all of the departments, and his lecturing style was not particularly student-friendly. He did not work from notes. He would walk into the lecture hall, apparently already half-way through this lectur, pick upthe chalk, and start writing on the board. Whether he was trying to show us how to solve Schrodinger's equation or develp the strain compatibility relations, the technique was always the same. He would clear a patch of board and start deriving a theorem. Running out of space, he would clear another patch, not necessarily connected with the first, and fill that up. Eventually, small pieces of the theorem would be scattered more-or-less at random across the chalkboard, stochastically mixed with the detritus of the previous lecture, and with random parts missing–erased to make space for more. It did not help that his writing was arocious, and his speech sounded as though he had filled his cheeks with marbles before starting. On one occation, one of my classmates managed to ge the professor's attention (a challenge) and asked him if he could possibly write a little more clearly. For a few lines, the writing was four times as large, but still as ilegible as before. Several lectures ended with Eshelby's discovery that he had misderived the theorm in question–a significant risk if you try to do it without notes, even if you are a bona fide genius. When this happened, he would stand back and survey the board. Agter a few moments, he would announce something like, "Well, there's a sign error there. You can correct it and work through to the result for yourselves." As if.

As time went by, our horror at his teaching style gave way to an understanding that the man was, in fact, a genius. Eccentric, yes but a genius. Apparently addicted to cheap cigars, he would smoke them down to the smaalest butt, then draw a cherry pipe out of his pocket, and stuff the remains of the cigar into it, tob e smoked until not a scrap of tobacco was left. He cared little for what people thought of hom, I think, and did not pay much attention to the politics of academia and the scientific community. This resultd in anunconscionalbe delay in his being elevated to the rank of Fellow of the Royal Society, which does seem to have been a sore point. In one memorable lecture, he described all of the current theories on a particular topic, listing the names of their authors on an uncharacteristically cleared chalkboard. He then described what was wrong with each of their work, condemning the weak-mindedness of these "so-called scientists" in quite direct terms. Having disposed of their failed logic, he then wrote the magical letters "FRS" after each of the names. He was elected an FRS himself that year and did not repeat the performance as far as I can gather.

Eshelby's impact on material sience is far, far out of proportion to the numbers of his publications. In total, he published less than 20 papers over his entire career (This is not true by the way. Eshelby published alomost 50-to-60 papers in his lifetime, but the point is valid: this days, you can see a lot of mediocre people published hundreds of junks, and good papers can not be published–Li's comment), but each of them is a classic. A fine demonstraion of the futility of today's obsessiion with publication-counting as a means of career assessment. Eshelby's work is characterized by real physical insight, complemented by elegant mathematical analysis (He was a professor of applied mathematics at Sheffield, in addition to being a professor of the theory of materials.) In contrast with his lectures, his written work is a modl of clarity. Although he was a powerful mathematician, he felt that we should only engage in "mathematical weightlifting" if we could not reason our way to the desired result through simple physical logic. Goodness knows what he would have made of today's computer simulation techniques. I think he would probably have thought of them as the last desperate resort after both physical reasoning and mathematical analysis failed.

An insight into Eshelby's motivations was provided to us in an informal moment one day, sitting in the small but splendid museum of glassware belonging to the Faculty of Materials, in a traditional British tea break. The usually unapproachable Esheby was unusually affable that day–perhaps he had just receibed word of his FRS election–but we fell into conversation and one undergraduate student adked him what had led to his being a "pure theoretician". He told us the story of a formative experience in his life. It seems that as a young teenager he had made a calculation of the thermal shock resistance of a piece of glass. This resulted from his mother's always using a thick cork pad beneath a coffee table. She explained the reason to him and he set to work calculationg the effect of the anticipated thermal shock. A short while later, he came to his mother and announced that he had completed his analysis, and that table would withstand a sudden local rise to the boiling point of water. His mother, being a wise woman, advised him that the obvious experiment would not be forthcoming and that he was forbidden from performing it himself. Well, curiosity and the budding scientific mind got the better of his youthful judgement one day when he was alone in the house. He boiled a pan of water and place it at the center of the prized coffee table. In his own words, "Well, cracks flew in every direction, and I suddently received a discouragement that from performing experiments that has lasted me the rest of my life."

True to the creed of the theoretician, however, he refused to allow that the analysis was flawed, and instead blamed the experiment. " Of course, I knew immediately what was wrong. The d***d thing hadn't been annealed properly. It was FULL of residual stress!"

By all accounts, this attack on the quality of the prized table did not endear him to his mother. Let all theorists beware of blaming the experiment lest they suffer similarly.

—By Alex King(From MRS Bulletin, July, 1999)

5.8 Exercises

PROBELM 5.1 *Show that the integral*

$$\int_{V_0} \exp\{i\boldsymbol{\xi}\cdot\mathbf{x}\}dV_x = 4\pi\sqrt{\frac{\pi}{2}}a^3\frac{J_{3/2}(\eta)}{\eta^{3/2}} \tag{5.153}$$

where V_0 is a sphere with radius a; $\eta = a|\xi|$, and $|\xi| = \sqrt{\xi_1^2+\xi_2^2+\xi_3^2}$.

Hint:

(1)Consider the identity

$$\nabla_{\mathbf{x}}\exp\Big(i\boldsymbol{\xi}\cdot\mathbf{x}\Big) = i\boldsymbol{\xi}\exp\Big(i\boldsymbol{\xi}\cdot\mathbf{x}\Big)$$

(2)

$$\int_0^1 t\sin(a|\xi|t)dt = \Gamma(1)(a|\xi|)^{-1/2}J_{3/2}(a|\xi|)$$

where $\Gamma(1) = \sqrt{\dfrac{\pi}{2}}$, $J_{3/2}(\eta)$ is the Bessel function of the first kind.

PROBELM 5.2 *Derive the displacement field inside an inclusion in which prescribed eigenstrain is a linear functin of coordinates, i.e. Example 5.1.*

PROBELM 5.3 *Derive Green's functin for plane strain problem by solving the following Navier equations,*

$$\sigma_{\beta\alpha,\beta} + \delta(\mathbf{x} - \mathbf{y})\delta_{\alpha\gamma} = 0 \tag{5.154}$$

where γ is the direction that the concentrated force point at.
Assume that the 2D elastic tensor is

$$C_{\alpha\beta\zeta\eta} = \lambda\delta_{\alpha\beta}\delta_{\zeta\eta} + \mu(\delta_{\alpha\zeta}\delta_{\beta\eta} + \delta_{\alpha\eta}\delta_{\beta\zeta}), \ \ \alpha, \beta, \zeta, \eta = 1, 2 \tag{5.155}$$

define 2D permutaion symbol

$$e_{\alpha\beta}: \ \ e_{11} = 0, \ e_{12} = 1, \ e_{21} = -1, \ e_{22} = 0 \tag{5.156}$$

The corresponding e-δ identities are:

$$(1) \qquad e_{\alpha\beta} = \begin{vmatrix} \delta_{\alpha 1} & \delta_{\alpha 2} \\ \delta_{\beta 1} & \delta_{\beta 2} \end{vmatrix}$$

$$(2) \qquad e_{\alpha\zeta}e_{\beta\eta} = \delta_{\alpha\beta}\delta_{\zeta\eta} - \delta_{\alpha\eta}\delta_{\beta\zeta}$$

$$(3) \qquad e_{\alpha\eta}e_{\beta\eta} = \delta_{\alpha\beta} \tag{5.157}$$

$$(4) \qquad e_{\alpha\eta}e_{\alpha\eta} = \delta_{\alpha\alpha} = 2! \tag{5.158}$$

Hints:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\exp(i(\xi_1 x_1 + \xi_2 x_2))}{\xi_1^2 + \xi_2^2} d\xi_1\xi_2 = -2\pi \ln R \tag{5.159}$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\xi_\alpha\xi_\beta}{\xi^4} \exp(i\boldsymbol{\xi}\cdot\mathbf{x})d\boldsymbol{\xi} = -\pi\delta_{\alpha\beta}\ln R - \pi\frac{x_\alpha x_\beta}{R^2} \tag{5.160}$$

where $R = \sqrt{x_1^2 + x_2^2}$.

PROBELM 5.4 *Let Ω be the half plane ($x_2 = 0, \ \ x_1 < 0$), and ϵ^*_{21} be prescribed as*

$$\epsilon^*_{21}(\mathbf{x}) = \frac{b}{2}\delta(x_2)H(-x_1) \tag{5.161}$$

Show

$$u_1(\mathbf{x}) = \frac{b}{2\pi}\tan^{-1}\Big(\frac{x_2}{x_1}\Big) + \frac{b}{4\pi}\Big(\frac{1}{1-\nu}\Big)\frac{x_1 x_2}{x_1^2 + x_2^2} \tag{5.162}$$

where ν is the Poisson's ratio.
Hint: (Mura's book page 17)

$$\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\xi_2}{\xi_1(\xi_1^2 + \xi_2^2)} \exp\{i(\xi_1 x_1 + \xi_2 x_2)\}d\xi_1 d\xi_2 &= 2\pi\tan^{-1}\Big(\frac{x_2}{x_1}\Big) \\
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\xi_1\xi_2}{(\xi_1^2 + \xi_2^2)^2} \exp\{i(\xi_1 x_1 + \xi_2 x_2)\}d\xi_1 d\xi_2 &= -\frac{\pi x_1 x_2}{x_1^2 + x_2^2}
\end{aligned}$$

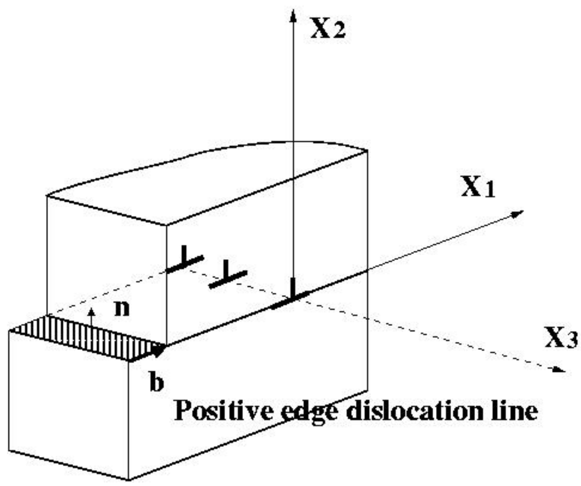


Figure 5.7. A straight edge dislocation

PROBELM 5.5 *Verify the following Hilbert transform formulas*

$$\mathcal{H}\Big(\frac{1}{(x^2+a^2)}\Big) = \frac{x}{a(x^2+a^2)} \tag{5.163}$$

$$\mathcal{H}(\sin(bx)) = -\cos(bx) \tag{5.164}$$

Hints: use Cauchy's residue theorm.

PROBELM 5.6 *Derive Eqs. (5.131), (5.132) and (5.134). Start from (5.111).*

Hints:

Hirth and Lothe [1992] Theory of Dislocations, Reprint Edition, Krieger Publishing Co. pages 228,235-237

Cottrell, A.H. [1953] Dislocations and plastic flow in crystals, Oxford University Press. pages 62-64, 98

PROBELM 5.7 *The 2D Green's function for plane strain problem is*

$$G_{\alpha\beta}(\mathbf{x}-\mathbf{x}') = \frac{1}{8\pi}\frac{1}{\mu(1-\nu)}\left\{\frac{(x_\alpha - x'_\alpha)(x_\beta - x'_\beta)}{R^2} - (3-4\nu)\delta_{\alpha\beta}\ln R\right\} \quad \alpha,\beta = 1,2 \tag{5.165}$$

where $R = \sqrt{(x_1 - x'_1)^2 + (x_2 - x'_2)^2}$.

Consider the following elliptical inclusion problem,

$$\epsilon^*_{\alpha\beta}(\mathbf{x}) = \begin{cases} \epsilon^*_{\alpha\beta}; & \forall\ \mathbf{x} \in \Omega \\ \\ 0; & \forall\ \mathbf{x} \in \mathbb{R}^2/\Omega \end{cases} \tag{5.166}$$

where $\epsilon^*_{\alpha\beta}$ *is a constant tensor, and* $\Omega := \left\{ \mathbf{x} \;\middle|\; \dfrac{x_1^2}{a_1^2} + \dfrac{x_2^2}{a_2^2} \leq 1 \right\}$.

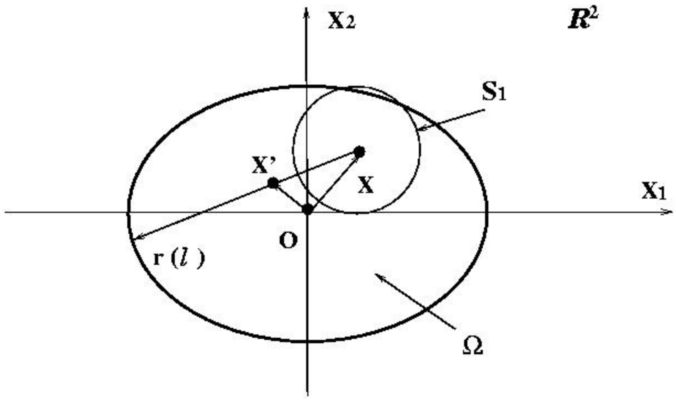


Figure 5.8. 2D elliptical inclusion

Find the Eshelby tensor for interior problem ($\mathbf{x} \in \Omega$)*. Hint (see Li (2000) pages 5606-5607).*

PROBELM 5.8 *Consider a spherical inclusion with radius a. Use identities*

$$\oint_{S^2} \ell_i \ell_j dS = \frac{4\pi a^2}{3}\delta_{ij} \tag{5.167}$$

$$\int_{S^2} \ell_i \ell_j \ell_m \ell_n dS = \frac{4\pi a^2}{15}(\delta_{ij}\delta_{mn} + \delta_{im}\delta_{jn} + \delta_{in}\delta_{jm}) \tag{5.168}$$

to show that

$$\begin{aligned} S^{\Omega}_{ijmn} &= \frac{1}{16\pi(1-\nu)} \oint_{S^2} \frac{\lambda_i g_{jmn} + \lambda_j g_{imn}}{g} dS \\ &= \frac{5\nu - 1}{15(1-\nu)}\delta_{ij}\delta_{mn} + \frac{2(4-5\nu)}{15(1-\nu)}(\delta_{im}\delta_{jn} + \delta_{jm}\delta_{in}) \end{aligned} \tag{5.169}$$

where $g_{ijk} = (1-2\nu)(\delta_{ij}\ell_k + \delta_{ik}\ell_j - \delta_{jk}\ell_i) + 3\ell_i\ell_j\ell_k$, $g = \ell_i\ell_i/a^2 = a^{-2}$, *and* $\lambda_i = \ell_i/a^2$.

PROBELM 5.9 *Show that*

$$\begin{aligned}\mathcal{G}(\mathbf{x}-\mathbf{y}) &= -\frac{1}{2}C_{klmn}\Big(G_{ik,lj}(\mathbf{x}-\mathbf{y})+G_{jk,li}(\mathbf{x}-\mathbf{y})\Big)\\ &= \frac{1}{8\pi(1-\nu)r^3}\Big[(1-2\nu)(\delta_{im}\delta_{jn}+\delta_{in}\delta_{jm}-\delta_{ij}\delta_{mn})\\ &\quad +3\nu(\delta_{im}l_jl_n+\delta_{in}l_jl_m+\delta_{jm}l_il_n+\delta_{jn}l_jl_m)\\ &\quad +3\delta_{ij}l_ml_n+3(1-2\nu)\delta_{mn}l_il_j-15l_il_jl_ml_n\Big] \qquad (5.170)\end{aligned}$$

where $\mathcal{G}_{ijmn}$ is called the fourth order Green's function (the second derivative of the Green's function).

Chapter 6

EFFECTIVE ELASTIC MODULUS

We now present Eshebly's equivalent eigestrain theory and its related engineering homogenization methods.

6.1 Effective elastic moduli for composites of dilute suspension

First, we apply the engineering homogenization theory to composites whose second phase concentration or other phase concentrations are small in comparison with the concentration of the matrix. In literature, we usually refer this as the composite with inhomogeneities of dilute suspension.

6.1.1 Basic equations for average stress and strain

Consider a solid with multiple phases of inhomogeneities, $\alpha = 1, 2, \cdots, n$. The elastic tensor and compliance tensor in the matrix is denoted as $\mathbf{C}$ and $\mathbf{D}$, and the elastic tensors and compliance tensors in the heterogeneous phases are denoted as $\mathbf{C}^{\alpha}$ and $\mathbf{D}^{\alpha}$ where $\alpha = 1, 2 \cdots, n$.

Define the averge stress and average strain in the matrix and in the inclusions,

$$< \boldsymbol{\sigma} >_M \;:=\; \frac{1}{M}\int_M \boldsymbol{\sigma} dV, \quad < \boldsymbol{\epsilon} >_M := \frac{1}{M}\int_M \boldsymbol{\epsilon} dV \tag{6.1}$$

$$< \boldsymbol{\sigma} >_\alpha \;:=\; \frac{1}{\Omega_\alpha}\int_{\Omega_\alpha} \boldsymbol{\sigma} dV, \quad < \boldsymbol{\epsilon} >_\alpha := \frac{1}{\Omega_\alpha}\int_{\Omega_\alpha} \boldsymbol{\epsilon} dV \tag{6.2}$$

By definition,

$$
\begin{aligned}
<\boldsymbol{\sigma}> &= \frac{1}{V}\int_V \boldsymbol{\sigma} dV = \frac{1}{V}\int_{M\cup\Omega_\alpha} \boldsymbol{\sigma} dV \\
&= \frac{1}{V}\Big[\frac{M}{M}\int_M \boldsymbol{\sigma} dV + \sum_{\alpha=1}^{n}\frac{\Omega_\alpha}{\Omega_\alpha}\int_{\Omega_\alpha}\boldsymbol{\sigma} dV\Big] \\
&= f_0 <\boldsymbol{\sigma}>_M + \sum_\alpha f_\alpha <\boldsymbol{\sigma}>_\alpha \qquad (6.3)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
f_0 <\boldsymbol{\sigma}>_M &= <\boldsymbol{\sigma}> - \sum_\alpha f_\alpha <\boldsymbol{\sigma}>_\alpha \\
&= \bar{\boldsymbol{C}} :<\boldsymbol{\epsilon}> - \sum_\alpha f_\alpha \mathbf{C}_\alpha :<\boldsymbol{\epsilon}>_\alpha \qquad (6.4)
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
f_0 <\boldsymbol{\sigma}>_M &= f_0 \mathbf{C} :<\boldsymbol{\epsilon}>_M = \mathbf{C} : \Big[\frac{M}{V}\frac{1}{M}\int_M \boldsymbol{\epsilon} dV\Big] \\
&= \mathbf{C} : \Big[\frac{1}{V}\int_{V/\cup\Omega_\alpha} \boldsymbol{\epsilon} dV\Big] \\
&= \mathbf{C} : \Big[\frac{1}{V}\int_V \boldsymbol{\epsilon} dV - \sum_\alpha \frac{\Omega_\alpha}{V}\frac{1}{\Omega_\alpha}\int_{\Omega_\alpha}\boldsymbol{\epsilon} dV\Big] \\
&= \mathbf{C} : \Big(<\boldsymbol{\epsilon}> - \sum_\alpha f_\alpha <\boldsymbol{\epsilon}>_\alpha\Big) \qquad (6.5)
\end{aligned}
$$

Combining Eqs. (6.4) and (6.5) yields

$$
\Big(\bar{\mathbf{C}} - \mathbf{C}\Big) :<\boldsymbol{\epsilon}>= \sum_\alpha f_\alpha \Big(\mathbf{C}^\alpha - \mathbf{C}\Big) :<\boldsymbol{\epsilon}>_\alpha \qquad (6.6)
$$

If the prescribed displacement boundary condition is applied, it may be also written

$$
\Big(\bar{\mathbf{C}} - \mathbf{C}\Big) : \boldsymbol{\epsilon}^0 = \sum_\alpha f_\alpha \Big(\mathbf{C}^\alpha - \mathbf{C}\Big) :<\boldsymbol{\epsilon}>_\alpha \qquad (6.7)
$$

Following a similar steps, one can show that

$$
\begin{aligned}
<\boldsymbol{\epsilon}> &= \frac{1}{V}\int_V \boldsymbol{\epsilon} dV = \frac{1}{V}\int_{M\cup\Omega_\alpha} \boldsymbol{\epsilon} dV \\
&= f_0 <\boldsymbol{\epsilon}>_M + \sum_\alpha f_\alpha <\boldsymbol{\epsilon}>_\alpha \qquad (6.8)
\end{aligned}
$$

Therefore,

$$\begin{aligned} f_0 < \boldsymbol{\epsilon} >_M &= < \boldsymbol{\epsilon} > - \sum_\alpha f_\alpha < \boldsymbol{\epsilon} >_\alpha \\ &= \bar{\mathbf{D}} :< \boldsymbol{\sigma} > - \sum_\alpha f_\alpha \mathbf{D}^\alpha :< \boldsymbol{\sigma} >_\alpha \end{aligned} \tag{6.9}$$

and

$$\begin{aligned} f_0 < \boldsymbol{\epsilon} >_M &= f_0 \mathbf{D} :< \boldsymbol{\sigma} >_M = \mathbf{D} : \Big[\frac{M}{V} \frac{1}{M} \int_M \boldsymbol{\sigma} dV \Big] \\ &= \mathbf{D} : \Big[\frac{1}{V} \int_V \boldsymbol{\sigma} dV - \sum_\alpha \frac{\Omega_\alpha}{V} \frac{1}{\Omega_\alpha} \int_{\Omega_\alpha} \boldsymbol{\sigma} dV \Big] \\ &= \mathbf{D} : \Big(< \boldsymbol{\sigma} > - \sum_\alpha f_\alpha < \boldsymbol{\sigma} >_\alpha \Big) \end{aligned} \tag{6.10}$$

Combining Eqs. (6.9) and (6.10) yields

$$\Big(\bar{\mathbf{D}} - \mathbf{D} \Big) :< \boldsymbol{\sigma} >= \sum_\alpha f_\alpha \Big(\mathbf{D}^\alpha - \mathbf{D} \Big) :< \boldsymbol{\sigma} >_\alpha \tag{6.11}$$

If the traction boundary condition is applied, it may be written

$$\Big(\bar{\mathbf{D}} - \mathbf{D} \Big) : \boldsymbol{\sigma}^0 = \sum_\alpha f_\alpha \Big(\mathbf{D}^\alpha - \mathbf{D} \Big) :< \boldsymbol{\sigma} >_\alpha \tag{6.12}$$

We name Eqs. (6.6) and (6.11) as the basic equations of average stress/strain fields.

6.1.2 Homeogenization: Equivalent stress/strain conditions

Consider the prescribed macro stress boundary condition,

$$\mathbf{t} = \mathbf{n} \cdot \boldsymbol{\sigma}^0, \quad \forall\, \mathbf{x} \in \partial V$$

Based on the averaging theorem, $< \boldsymbol{\sigma} >= \boldsymbol{\sigma}^0$.

One may note that the remote background strain as

$$\boldsymbol{\epsilon}^0 = \mathbf{D} : \boldsymbol{\sigma}^0 = \mathbf{D} :< \boldsymbol{\sigma} > \neq < \boldsymbol{\epsilon} > \tag{6.13}$$

Similarly, for prescribed macro-strain boundary condition,

$$\mathbf{u}(\mathbf{x}) = \mathbf{x} \cdot \boldsymbol{\epsilon}^0, \quad \mathbf{x} \in \partial V$$

the averaging theorem asserts that in this case

$$\boldsymbol{\epsilon}^0 = < \boldsymbol{\epsilon} > \ .$$

The background stress,

$$\boldsymbol{\sigma}^0 = \mathbf{C} :< \boldsymbol{\epsilon} > \neq < \boldsymbol{\sigma} > \ . \tag{6.14}$$

Suppose that there are $\alpha = 1, 2, \cdots, n$ distinct inhomogenous phases. $\forall \mathbf{x} \in \Omega_\alpha$, the stress and strain equivalent conditions are

$$\mathbf{C}^\alpha : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d) = \mathbf{C} : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d - \boldsymbol{\epsilon}^*) \tag{6.15}$$

or

$$\mathbf{D}^\alpha : (\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d) = \mathbf{C} : (\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d - \boldsymbol{\sigma}^*) \tag{6.16}$$

Then one can find the average stress and strain fields inside each inclusion,

$$< \boldsymbol{\epsilon} >_\alpha \ = \ \mathbf{A}^\alpha : \boldsymbol{\epsilon}^* \tag{6.17}$$
$$< \boldsymbol{\sigma} >_\alpha \ = \ \mathbf{B}^\alpha : \boldsymbol{\sigma}^* \tag{6.18}$$

where

$$\mathbf{A}^\alpha \ = \ (\mathbf{C} - \mathbf{C}^\alpha)^{-1} : \mathbf{C} \tag{6.19}$$
$$\mathbf{B}^\alpha \ = \ (\mathbf{B} - \mathbf{B}^\alpha)^{-1} : \mathbf{B} \tag{6.20}$$

Since the inclusion population is small, one can neglect the interaction among inclusions. The disturbance field inside each inclusion can then be related to eigenstrain fields,

$$\boldsymbol{\epsilon}^d \ = \ \bar{\mathbf{S}}^\alpha : \boldsymbol{\epsilon}^*, \quad \forall \mathbf{x} \in \Omega^\alpha \tag{6.21}$$
$$\boldsymbol{\sigma}^d \ = \ \bar{\mathbf{T}}^\alpha : \boldsymbol{\sigma}^*, \quad \forall \mathbf{x} \in \Omega^\alpha \tag{6.22}$$

Subsequently, one can decide how much the eigenstrain or eigenstress have to be prescribed by the following conditions,

$$\boldsymbol{\epsilon}^* \ = \ (\mathbf{A}^\alpha - \mathbf{S}^\alpha)^{-1} : \boldsymbol{\epsilon}^0 \tag{6.23}$$
$$\boldsymbol{\sigma}^* \ = \ (\mathbf{B}^\alpha - \mathbf{T}^\alpha)^{-1} : \boldsymbol{\sigma}^0 \tag{6.24}$$

Therefore the average strain/stress inside the α-th phase inclusion may be expressed by eigenstrain/eigenstress, i.e.

$$< \boldsymbol{\epsilon} >_\alpha \ = \ \mathbf{A}^\alpha : \boldsymbol{\epsilon}^* = \mathbf{A}^\alpha : (\mathbf{A}^\alpha - \mathbf{S}^\alpha)^{-1} \boldsymbol{\epsilon}^0 \tag{6.25}$$
$$< \boldsymbol{\sigma} >_\alpha \ = \ \mathbf{B}^\alpha : \boldsymbol{\sigma}^* = \mathbf{B}^\alpha : (\mathbf{B}^\alpha - \mathbf{T}^\alpha)^{-1} \boldsymbol{\sigma}^0 \tag{6.26}$$

Subsequently, one can relate the average strain and average stress in the α-th inclusion (inhomogeneity) with the background strain and background stress through the so-called *concentration tensors*,

$$< \boldsymbol{\epsilon} >_\alpha \ = \ \mathcal{A}^\alpha : \boldsymbol{\epsilon}^0 \tag{6.27}$$
$$< \boldsymbol{\sigma} >_\alpha \ = \ \mathcal{B}^\alpha : \boldsymbol{\sigma}^0 \tag{6.28}$$

where the concentration tensors are defined as

$$\mathcal{A}^\alpha = \mathbf{A}^\alpha : (\mathbf{A}^\alpha - \mathbf{S}^\alpha)^{-1} \tag{6.29}$$
$$\mathcal{B}^\alpha = \mathbf{B}^\alpha : (\mathbf{B}^\alpha - \mathbf{T}^\alpha)^{-1} \tag{6.30}$$

Since by definition $< \boldsymbol{\sigma} >_\alpha = \mathbf{C}^\alpha :< \boldsymbol{\epsilon} >_\alpha$ and $< \boldsymbol{\epsilon} >_\alpha = \mathbf{D}^\alpha :< \boldsymbol{\sigma} >_\alpha$, one can rewrite Eqs. (6.27) and (6.28) as

$$< \boldsymbol{\sigma} >_\alpha = \begin{cases} \mathbf{C}^\alpha : \mathcal{A}^\alpha : \mathbf{D} : \boldsymbol{\sigma}^0 \\ \mathcal{B}^\alpha : \boldsymbol{\sigma}^0 \end{cases} \tag{6.31}$$

or

$$< \boldsymbol{\epsilon} >_\alpha = \begin{cases} \mathbf{D}^\alpha : \mathcal{B}^\alpha : \mathbf{C} : \boldsymbol{\epsilon}^0 \\ \mathcal{A}^\alpha : \boldsymbol{\epsilon}^0 \end{cases} \tag{6.32}$$

Suppose that prescribed macro-stress boundary condition is applied. Substituting both expressions in Eq. (6.31) into the basic average equation (6.11) yields,

$$(\bar{\mathbf{D}} - \mathbf{D}) : \boldsymbol{\sigma}^0 = \sum_{\alpha=1}^{n} f_\alpha (\mathbf{D}^\alpha - \mathbf{D}) : \begin{cases} \mathbf{C}^\alpha : \mathcal{A}^\alpha : \bar{\mathbf{D}} : \boldsymbol{\sigma}^0 \\ \\ \mathcal{B}^\alpha : \boldsymbol{\sigma}^0 \end{cases} \tag{6.33}$$

Therefore, for prescribed traction boundary condition, we have the following estimate on effective compliance tensor,

$$\bar{\mathbf{D}} = \begin{cases} \mathbf{D} + \displaystyle\sum_{\alpha=1}^{n} f_\alpha (\mathbf{D}^\alpha - \mathbf{D}) : \mathbf{C}^\alpha : \mathcal{A}^\alpha : \bar{\mathbf{D}} \\ \\ \mathbf{D} + \displaystyle\sum_{\alpha=1}^{n} f_\alpha (\mathbf{D}^\alpha - \mathbf{D}) : \mathcal{B}^\alpha \end{cases} \tag{6.34}$$

By considering the identities,

$$(\mathbf{A}^\alpha)^{-1} = (\mathbf{D}^\alpha - \mathbf{D}) : \mathbf{C}^\alpha, \;\text{ and }\; \mathbf{B}^\alpha = (\mathbf{D} - \mathbf{D}^\alpha)^{-1} : \mathbf{D} \tag{6.35}$$

Finally, we obtain

$$\bar{\mathbf{D}} = \begin{cases} \mathbf{D} + \displaystyle\sum_{\alpha=1}^{n} f_\alpha (\mathbf{A}^\alpha - \mathbf{S}^\alpha)^{-1} : \mathbf{D} \\ \\ \mathbf{D} - \displaystyle\sum_{\alpha=1}^{n} f_\alpha \mathbf{D} : (\mathbf{B}^\alpha - \mathbf{T}^\alpha)^{-1} \end{cases} \tag{6.36}$$

If prescribed macro-strain boundary condition is applied, one may substitute the both expressions of (6.58) into the basic average equation (6.6). It leads to

$$(\bar{\mathbf{C}} - \mathbf{C}) : \boldsymbol{\epsilon}^0 = \sum_{\alpha=1}^{n} f_\alpha (\mathbf{C}^\alpha - \mathbf{C}) : \begin{cases} \mathbf{D}^\alpha : \mathcal{B}^\alpha : \mathbf{C} : \boldsymbol{\epsilon}^0 \\ \\ \mathcal{A}^\alpha : \boldsymbol{\epsilon}^0 \end{cases} \tag{6.37}$$

The following estimate on effective elastic tensor may be obtained,

$$\bar{\mathbf{C}} = \begin{cases} \mathbf{C} + \displaystyle\sum_{\alpha=1}^{n} f_\alpha (\mathbf{C}^\alpha - \mathbf{C}) : \mathbf{D}^\alpha : \mathcal{B}^\alpha : \mathbf{C} \\ \\ \mathbf{C} + \displaystyle\sum_{\alpha=1}^{n} f_\alpha (\mathbf{C}^\alpha - \mathbf{C}) : \mathcal{A}^\alpha \end{cases} \tag{6.38}$$

Using the identities,

$$(\mathbf{B}^\alpha)^{-1} = (\mathbf{C}^\alpha - \mathbf{C}) : \mathbf{D}^\alpha, \text{ and } \mathbf{A}^\alpha = -(\mathbf{C}^\alpha - \mathbf{C})^{-1} : \mathbf{C} \tag{6.39}$$

we have the following estimate on effective elastic stiffness tensor

$$\bar{\mathbf{C}} = \begin{cases} \mathbf{C} + \displaystyle\sum_{\alpha=1}^{n} f_\alpha (\mathbf{B}^\alpha - \mathbf{T}^\alpha)^{-1} : \mathbf{C} \\ \\ \mathbf{C} - \displaystyle\sum_{\alpha=1}^{n} f_\alpha : \mathbf{C} : (\mathbf{A}^\alpha - \mathbf{S}^\alpha)^{-1} \end{cases} \tag{6.40}$$

Note that the index α starts from 1, and each α is an inhomogeneous phase.

One of the drawback of dilute distribution homogenization is

$$\bar{\mathbf{D}} : \bar{\mathbf{C}} \neq \mathbf{1} \quad \text{or} \quad \bar{\mathbf{D}} \neq \bar{\mathbf{C}}^{-1}.$$

This can be shown for $\alpha = 1$:

$$\begin{aligned} \bar{\mathbf{D}} : \bar{\mathbf{C}} &= \left(\mathbf{1}^{(4s)} + f_\alpha (\mathbf{A}^\alpha - \mathbf{S}^\alpha)^{-1}\right) : \mathbf{D} : \mathbf{C} : \left(\mathbf{1}^{(4s)} - f_\alpha (\mathbf{A}^\alpha - \mathbf{S}^\alpha)^{-1}\right) \\ &= \mathbf{1}^{(4s)} - f_\alpha^2 (\mathbf{A}^\alpha - \mathbf{S}^\alpha)^{-1} : (\mathbf{A}^\alpha - \mathbf{S}^\alpha)^{-1} \\ &= \mathbf{1}^{(4s)} + \mathcal{O}(f_\alpha^2) \neq \mathbf{1}^{(4s)} \,. \end{aligned} \tag{6.41}$$

Obviously, the effective elastic stiffness is not consistent with the effective compliance tensors.

6.1.3 Elastic moduli in isotropic case

Suppose that there are n different phases of inhomogeneities in a solid. For prescribed traction boundary condition, Eshelby's equivalent strain method yields,

$$\mathbf{D} = \left\{ \mathbf{1} + \sum_{\alpha=1}^{n} f_\alpha (\mathbf{A}^\alpha - \mathbf{S}^\alpha)^{-1} \right\} : \mathbf{D}$$

where $\mathbf{A}^\alpha$ is defined as

$$\mathbf{A}^\alpha = (\mathbf{C} - \mathbf{C}^\alpha)^{-1} : \mathbf{C}$$

Here $\mathbf{C}$ is the elastic tensor of the matrix, which is assumed to be isotropic, i.e. $\mathbf{C} = 3K\mathbf{E}^{(1)} + 2\mu\mathbf{E}^{(2)}$. We can then calculate

$$\mathbf{C} - \mathbf{C}^\alpha = 3(K - K^\alpha)\mathbf{E}^{(1)} + 2(\mu - \mu^\alpha)\mathbf{E}^{(2)}$$

and

$$\begin{aligned}
\mathbf{A}^\alpha &= (\mathbf{C} - \mathbf{C}^\alpha)^{-1} : \mathbf{C} \\
&= \Big(\frac{1}{3(K - K^\alpha)} \mathbf{E}^{(1)} + \frac{1}{2(\mu - \mu^\alpha)} \mathbf{E}^{(2)} \Big) : \Big(3K\mathbf{E}^{(1)} + 2\mu\mathbf{E}^{(2)} \Big) \\
&= \frac{K}{K - K^\alpha} \mathbf{E}^{(1)} + \frac{\mu}{\mu - \mu^\alpha} \mathbf{E}^{(2)}
\end{aligned}$$

Since the composite is isotropic, we use the Eshelby tensor of spherical inclusions, For spherical inclusion, the Eshelby tensor is

$$\begin{aligned}
\mathbf{S}^\Omega &= \frac{5\nu - 1}{15(1 - \nu)} \mathbf{1}^{(2)} \otimes \mathbf{1}^{(2)} + \frac{2(4 - 5\nu)}{15(1 - \nu)} \mathbf{1}^{(4s)} \\
&= \frac{(1 + \nu)}{3(1 - \nu)} \mathbf{E}^{(1)} + \frac{2(4 - 5\nu)}{15(1 - \nu)} \mathbf{E}^{(2)} \\
&= s_1 \mathbf{E}^{(1)} + s_2 \mathbf{E}^{(2)}
\end{aligned}$$

where $s_1 = \dfrac{1 + \nu}{3(1 - \nu)}$ and $s_2 = \dfrac{2(4 - 5\nu)}{15(1 - \nu)}$.

Then

$$\mathbf{A}^\alpha - \mathbf{S}^\alpha = \Big(\frac{K}{(K - K^\alpha)} - s_1^\alpha \Big) \mathbf{E}^{(1)} + \Big(\frac{\mu}{(\mu - \mu^\alpha)} - s_2^\alpha \Big) \mathbf{E}^{(2)}$$

and

$$\Big(\mathbf{A}^\alpha - \mathbf{S}^\alpha \Big)^{-1} = \frac{1}{\dfrac{K}{(K - K^\alpha)} - s_1^\alpha} \mathbf{E}^{(1)} + \frac{1}{\dfrac{\mu}{(\mu - \mu^\alpha)} - s_2^\alpha} \mathbf{E}^{(2)}$$

Hence

$$
\begin{aligned}
\bar{\mathbf{D}} &= \Big(\mathbf{1} + \sum_{\alpha=1}^{n} f_\alpha (\mathbf{A}^\alpha - \mathbf{S}^\alpha)^{-1}\Big) : \mathbf{D} \\
&= \left\{ \mathbf{E}^{(1)} + \mathbf{E}^{(2)} + \sum_{\alpha=1}^{n} \frac{f_\alpha}{\dfrac{K}{(K - K^\alpha)} - s_1^\alpha} \mathbf{E}^{(1)} \right. \\
&\quad \left. + \frac{f_\alpha}{\dfrac{\mu}{(\mu - \mu_\alpha)} - s_2^\alpha} \mathbf{E}^{(2)} \right\} : \Big(\frac{1}{3K} \mathbf{E}^{(1)} + \frac{1}{2\mu} \mathbf{E}^{(2)} \Big)
\end{aligned}
$$

Finally,

$$
\begin{aligned}
\bar{\mathbf{D}} &= \frac{1}{3K} \left(1 + \sum_{\alpha=1}^{n} \frac{f_\alpha}{\dfrac{K}{K - K^\alpha} - s_1^\alpha} \right) \mathbf{E}^{(1)} \\
&\quad + \frac{1}{2\mu} \left(1 + \sum_{\alpha=1}^{n} \frac{f_\alpha}{\dfrac{\mu}{\mu - \mu^\alpha} - s_2^\alpha} \right) \mathbf{E}^{(2)} \qquad (6.42)
\end{aligned}
$$

Assume that $f_\alpha << 1$,

$$
\begin{aligned}
\frac{\bar{K}}{K} &= \left(1 + \sum_{\alpha=1}^{n} \frac{f_\alpha}{\dfrac{K}{K - K^\alpha} - s_1^\alpha} \right)^{-1} \\
&= 1 - \sum_{\alpha=1}^{n} \frac{f_\alpha}{\dfrac{K}{K - K^\alpha} - s_1^\alpha} + \mathcal{O}(f_\alpha^2)
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\bar{\mu}}{\mu} &= \left(1 + \sum_{\alpha=1}^{n} \frac{f_\alpha}{\dfrac{\mu}{\mu - \mu^\alpha} - s_2^\alpha} \right)^{-1} \\
&= 1 - \sum_{\alpha=1}^{n} \frac{f_\alpha}{\dfrac{\mu}{\mu - \mu^\alpha} - s_2^\alpha} + \mathcal{O}(f_\alpha^2)
\end{aligned}
$$

Similarly, by considering remote traction boundary condition, we have the estimate of effective elastic modulus for solids with dilute suspension of inhomogeneities,

$$\mathbf{C} = \left\{ \mathbf{1} + \sum_{\alpha=1}^{n} f_\alpha (\mathbf{B}^\alpha - \mathbf{T}^\alpha)^{-1} \right\} : \mathbf{C}$$

where $\mathbf{B}^\alpha$ is defined as

$$\mathbf{B}^\alpha = (\mathbf{D} - \mathbf{D}^\alpha)^{-1} : \mathbf{D}$$

Here $\mathbf{D}$ is the elastic compliance tensor of the matrix, i.e. $\mathbf{D} = \frac{1}{3K}\mathbf{E}^{(1)} + \frac{1}{2\mu}\mathbf{E}^{(2)}$. We can then calculate

$$\begin{aligned}
\mathbf{B}^\alpha &= (\mathbf{D} - \mathbf{D}^\alpha)^{-1} : \mathbf{D} \\
&= \left(\frac{1}{3}\Big(\frac{1}{K} - \frac{1}{K^\alpha}\Big)\mathbf{E}^{(1)} + \frac{1}{2}\Big(\frac{1}{\mu} - \frac{1}{\mu^\alpha}\Big)\mathbf{E}^{(2)} \right)^{-1} : \mathbf{D} \\
&= \left(\frac{K^\alpha - K}{3KK^\alpha}\mathbf{E}^{(1)} + \frac{\mu^\alpha - \mu}{2\mu\mu^\alpha}\mathbf{E}^{(2)} \right)^{-1} : \Big(\frac{1}{3K}\mathbf{E}^{(1)} + \frac{1}{2\mu}\mathbf{E}^{(2)} \Big) \\
&= -\frac{K^\alpha}{K - K^\alpha}\mathbf{E}^{(1)} - \frac{\mu^\alpha}{\mu - \mu^\alpha}\mathbf{E}^{(2)}
\end{aligned}$$

Subsequently,

$$\begin{aligned}
\mathbf{B}^\alpha - \mathbf{T}^\alpha &= \Big(-\frac{K^\alpha}{(K - K^\alpha)} - (1 - s_1^\alpha) \Big)\mathbf{E}^{(1)} + \Big(-\frac{\mu^\alpha}{(\mu - \mu^\alpha)} - (1 - s_2^\alpha) \Big)\mathbf{E}^{(2)} \\
&= -\Big(\frac{K}{K - K^\alpha} - s_1^\alpha \Big)\mathbf{E}^{(1)} - \Big(\frac{\mu}{\mu - \mu^\alpha} - s_2^\alpha \Big)\mathbf{E}^{(2)}
\end{aligned}$$

and

$$\Big(\mathbf{B}^\alpha - \mathbf{T}^\alpha\Big)^{-1} = -\frac{1}{\dfrac{K}{(K - K^\alpha)} - s_1^\alpha}\mathbf{E}^{(1)} - \frac{1}{\dfrac{\mu}{(\mu - \mu^\alpha)} - s_2^\alpha}\mathbf{E}^{(2)}$$

Finally

$$
\begin{aligned}
\bar{\mathbf{C}} &= \Big(\mathbf{1} + \sum_{\alpha=1}^{n} f_\alpha (\mathbf{B}^\alpha - \mathbf{T}^\alpha)^{-1}\Big) : \mathbf{C} \\
&= \left\{ \Big(1 - \sum_{\alpha=1}^{n} \frac{f_\alpha}{\dfrac{K}{(K-K^\alpha)} - s_1^\alpha}\Big)\mathbf{E}^{(1)} + \Big(1 - \sum_{\alpha=1}^{n} \frac{f_\alpha}{\dfrac{\mu}{(\mu-\mu_\alpha)} - s_2^\alpha}\mathbf{E}^{(2)} \right\} \\
&\quad : (3K\mathbf{E}^{(1)} + 2\mu\mathbf{E}^{(2)}) \\
&= 3K\Big(1 - \sum_{\alpha=1}^{n} \frac{f_\alpha}{\dfrac{K}{(K-K^\alpha)} - s_1^\alpha}\Big)\mathbf{E}^{(1)} + 2\mu\Big(1 - \sum_{\alpha=1}^{n} \frac{f_\alpha}{\dfrac{\mu}{(\mu-\mu_\alpha)} - s_2^\alpha}\Big)\mathbf{E}^{(2)}
\end{aligned}
$$

Therefore

$$
\frac{\bar{K}}{K} = 1 - \sum_{\alpha=1}^{n} \frac{f_\alpha}{\dfrac{K}{K-K^\alpha} - s_1^\alpha}
$$

and

$$
\frac{\bar{\mu}}{\mu} = 1 - \sum_{\alpha=1}^{n} \frac{f_\alpha}{\dfrac{\mu}{\mu-\mu^\alpha} - s_2^\alpha}
$$

It is obviously that these results are different from the results obtained from prescribed traction boundary condition. They are only agreeable to the first order of the volume fraction. In other words, these two results (the results obtained from prescribed stress b.c. and the results obtained from prescribed strain b.c.) are not consistent in the homogenization scheme for dilute inhomogeneity distribution.

6.2 Self-consistent method

As shown above, effective elastic tensor and compliance tensor obtained via homogenization of inhomogeneities of dilute distribution are not reciprocal to each other as supposed to be. As the volume fraction of inhomogeneity increases, the accuracy of dilute suspension homogenization schemes deteriorates, because the interaction among inhomogeneities become strong.

To take into account the interaction among inhomogeneities, a so-called self-consistent homogenization method is proposed, which is largely attributed to a series papers by Hill ([1962],[1963],[1964]). Rodney Hill is a highly intellectual individual, whose writing style is very close to mathematics literature, which is rigorous, terse, and often esoteric.

The following presentation is mainly adopted from Nemat-Nasser and Hori, and it is blended with authors own interpretation, which is more engineering oriented.

There are two main differences between self-consistent homogenization and dilute suspension homogenization.

The first difference is in the treatment of remote (background) strain and stress.

Consider the prescribed macro stress boundary condition,

$$\mathbf{t} = \mathbf{n} \cdot \boldsymbol{\sigma}^0, \quad \forall\, \mathbf{x} \in \partial V$$

Based on the averaging theorem, $< \boldsymbol{\sigma} >= \boldsymbol{\sigma}^0$. In self-consistent homogenization, we define the remote background strain as

$$\boldsymbol{\epsilon}^0 = \bar{\mathbf{D}} : \boldsymbol{\sigma}^0 = \bar{\mathbf{D}} :< \boldsymbol{\sigma} > \tag{6.43}$$

Therefore in this case,

$$\boldsymbol{\epsilon}^0 = \bar{\mathbf{D}} :< \boldsymbol{\sigma} >=< \boldsymbol{\epsilon} > \ .$$

Similarly, for prescribed macro-strain boundary condition,

$$\mathbf{u}(\mathbf{x}) = \mathbf{x} \cdot \boldsymbol{\epsilon}^0, \quad \mathbf{x} \in \partial V$$

the averaging theorem asserts that in this case

$$\boldsymbol{\epsilon}^0 =< \boldsymbol{\epsilon} > \ .$$

If $\boldsymbol{\sigma} = \bar{\mathbf{C}} : \boldsymbol{\epsilon}$, the background stress will be the average stress,

$$\boldsymbol{\sigma}^0 = \bar{\mathbf{C}} :< \boldsymbol{\epsilon} >=< \boldsymbol{\sigma} > \ . \tag{6.44}$$

The second main difference between the self-consistent method and dilute suspension method is that Eshelby's equivalent inclusion principle is applied with respect to the homogenized solid, instead of matrix. Suppose that there are $\alpha = 1, 2, \cdots, n$ distinct inhomogenous phases. $\forall \mathbf{x} \in \Omega_\alpha$,

$$\mathbf{C}^\alpha : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d) = \bar{\mathbf{C}} : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d - \boldsymbol{\epsilon}^*) \tag{6.45}$$

or

$$\mathbf{D}^\alpha : (\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d) = \bar{\mathbf{C}} : (\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d - \boldsymbol{\sigma}^*) \tag{6.46}$$

Moreover, the disturbance field generated by eigenstrain is also calculated with respect to homogenized solid, i.e.

$$\boldsymbol{\epsilon}^d = \bar{\mathbf{S}}^\alpha : \boldsymbol{\epsilon}^*, \quad \forall \mathbf{x} \in \Omega^\alpha \tag{6.47}$$

$$\boldsymbol{\sigma}^d = \bar{\mathbf{T}}^\alpha : \boldsymbol{\sigma}^*, \quad \forall \mathbf{x} \in \Omega^\alpha \tag{6.48}$$

Therefore the average strain/stress inside the α-th phase inclusion may be expressed by eigenstrain/eigenstress, i.e.

$$\begin{aligned} <\boldsymbol{\epsilon}>_\alpha &= \bar{\mathbf{A}}^\alpha : \boldsymbol{\epsilon}^* && (6.49)\\ <\boldsymbol{\sigma}>_\alpha &= \bar{\mathbf{B}}^\alpha : \boldsymbol{\sigma}^* && (6.50)\end{aligned}$$

where

$$\begin{aligned} \bar{\mathbf{A}}^\alpha &= (\bar{\mathbf{C}} - \mathbf{C}^\alpha)^{-1} : \bar{\mathbf{C}} && (6.51)\\ \bar{\mathbf{B}}^\alpha &= (\bar{\mathbf{B}} - \mathbf{B}^\alpha)^{-1} : \bar{\mathbf{B}} && (6.52)\end{aligned}$$

Subsequently, one can relate the average strain and average stress in the α-th inclusion (inhomogeneity) with the background strain and background stress by concentration tensors,

$$\begin{aligned} <\boldsymbol{\epsilon}>_\alpha &= \bar{\mathcal{A}}^\alpha : \boldsymbol{\epsilon}^0 && (6.53)\\ <\boldsymbol{\sigma}>_\alpha &= \bar{\mathcal{B}}^\alpha : \boldsymbol{\sigma}^0 && (6.54)\end{aligned}$$

where the concentration tensors are defined as

$$\begin{aligned} \bar{\mathcal{A}}^\alpha &= \bar{\mathbf{A}}^\alpha : (\bar{\mathbf{A}}^\alpha - \bar{\mathbf{S}}^\alpha)^{-1} && (6.55)\\ \bar{\mathcal{B}}^\alpha &= \bar{\mathbf{B}}^\alpha : (\bar{\mathbf{B}}^\alpha - \bar{\mathbf{T}}^\alpha)^{-1} && (6.56)\end{aligned}$$

Since by definition $<\boldsymbol{\sigma}>_\alpha = \mathbf{C}^\alpha :<\boldsymbol{\epsilon}>_\alpha$ and $<\boldsymbol{\epsilon}>_\alpha = \mathbf{D}^\alpha :<\boldsymbol{\sigma}>_\alpha$, one can rewrite Eqs. (6.53) and (6.54) as

$$<\boldsymbol{\sigma}>_\alpha = \begin{cases} \mathbf{C}^\alpha : \bar{\mathcal{A}}^\alpha : \bar{\mathbf{D}} : \boldsymbol{\sigma}^0 \\ \bar{\mathcal{B}}^\alpha : \boldsymbol{\sigma}^0 \end{cases} \qquad (6.57)$$

or

$$<\boldsymbol{\epsilon}>_\alpha = \begin{cases} \mathbf{D}^\alpha : \bar{\mathcal{B}}^\alpha : \bar{\mathbf{C}} : \boldsymbol{\epsilon}^0 \\ \bar{\mathcal{A}}^\alpha : \boldsymbol{\epsilon}^0 \end{cases} \qquad (6.58)$$

Note that the relationships $\boldsymbol{\epsilon}^0 =<\boldsymbol{\epsilon}>$ and $\sigma^0 =<\boldsymbol{\sigma}>$ are used.

Suppose that prescribed macro-stress boundary condition is applied. Substituting Eqs. (6.57) and (6.58) into the basic average equation (6.11) yields,

$$(\bar{\mathbf{D}} - \mathbf{D}) : \boldsymbol{\sigma}^0 = \sum_{\alpha=1}^{n} f_\alpha (\mathbf{D}^\alpha - \mathbf{D}) : \begin{cases} \mathbf{C}^\alpha : \bar{\mathcal{A}}^\alpha : \bar{\mathbf{D}} : \boldsymbol{\sigma}^0 \\ \\ \bar{\mathcal{B}}^\alpha : \boldsymbol{\sigma}^0 \end{cases} \qquad (6.59)$$

Therefore, self-consustent method gives the following estimate on effective compliance tensor,

$$\bar{\mathbf{D}} = \begin{cases} \mathbf{D} + \displaystyle\sum_{\alpha=1}^{n} f_\alpha (\mathbf{D}^\alpha - \mathbf{D}) : \mathbf{C}^\alpha : \bar{\mathcal{A}}^\alpha : \bar{\mathbf{D}} \\ \\ \mathbf{D} + \displaystyle\sum_{\alpha=1}^{n} f_\alpha (\mathbf{D}^\alpha - \mathbf{D}) : \bar{\mathcal{B}}^\alpha \end{cases} \qquad (6.60)$$

If prescribed macro-strain boundary condition is applied, one may substitute Eqs. (6.57) and (6.58) into the basic average equation (6.6). It leads to

$$(\bar{\mathbf{C}} - \mathbf{C}) : \boldsymbol{\epsilon}^0 = \sum_{\alpha=1}^{n} f_\alpha (\mathbf{C}^\alpha - \mathbf{C}) : \begin{cases} \mathbf{D}^\alpha : \bar{\mathcal{B}}^\alpha : \bar{\mathbf{C}} : \boldsymbol{\epsilon}^0 \\ \bar{\mathcal{A}}^\alpha : \boldsymbol{\epsilon}^0 \end{cases} \tag{6.61}$$

Hence self-consustent method gives the following estimate on effective elastic tensor,

$$\bar{\mathbf{C}} = \begin{cases} \mathbf{C} + \displaystyle\sum_{\alpha=1}^{n} f_\alpha (\mathbf{C}^\alpha - \mathbf{C}) : \mathbf{D}^\alpha : \bar{\mathcal{B}}^\alpha : \bar{\mathbf{C}} \\ \mathbf{C} + \displaystyle\sum_{\alpha=1}^{n} f_\alpha (\mathbf{C}^\alpha - \mathbf{C}) : \bar{\mathcal{A}}^\alpha \end{cases} \tag{6.62}$$

Note that the index α starts from 1, and each α is an inhomogeneous phase.

We now show that

$$\bar{\mathbf{D}} : \bar{\mathbf{C}} = \mathbf{1} \quad \text{or} \quad \bar{\mathbf{D}} = \bar{\mathbf{C}}^{-1}.$$

Consider

$$\begin{aligned} \mathbf{D} &= \mathbf{D} : \mathbf{1} = \mathbf{D} : \bar{\mathbf{C}} : \bar{\mathbf{C}}^{-1} \\ &= \mathbf{D} : \Big[\mathbf{C} + \sum_{\alpha=1}^{n} f_\alpha (\mathbf{C}^\alpha - \mathbf{C}) : \bar{\mathcal{A}}^\alpha\Big] : \bar{\mathbf{C}}^{-1} \\ &= \bar{\mathbf{C}}^{-1} + \sum_{\alpha=1}^{n} f_\alpha \mathbf{D} : (\mathbf{C}^\alpha - \mathbf{C}) : \mathcal{A}^\alpha : \bar{\mathbf{C}}^{-1} \end{aligned} \tag{6.63}$$

Since,

$$\begin{aligned} \mathbf{D} : (\mathbf{C}^\alpha - \mathbf{C}) &= \mathbf{D} : \mathbf{C}^\alpha - \mathbf{1} \\ &= -\mathbf{1} + \mathbf{D} : \mathbf{C}^\alpha \\ &= -(\mathbf{D}^\alpha - \mathbf{D}) : \mathbf{C}^\alpha \end{aligned}$$

The last line of (6.63) may be rewritten as

$$\mathbf{D} = \bar{\mathbf{C}}^{-1} - \sum_{\alpha=1}^{n} f_\alpha (\mathbf{D}^\alpha - \mathbf{D}) : \mathbf{C}^\alpha : \mathcal{A}^\alpha : \bar{\mathbf{C}}^{-1} \tag{6.64}$$

which leads to

$$\bar{\mathbf{C}}^{-1} = \mathbf{C}^{-1} + \sum_{\alpha=1}^{n} f_\alpha (\mathbf{D}^\alpha - \mathbf{D}) : \mathbf{C}^\alpha : \bar{\mathcal{A}}^\alpha : \bar{\mathbf{C}}^{-1} \tag{6.65}$$

Compare (6.65) with the first line of Eq. (6.60). One can conclude that

$$\bar{\mathbf{C}}^{-1} = \bar{\mathbf{D}} \tag{6.66}$$

Similar arguments can be made to show that $\bar{\mathbf{D}}^{-1} = \bar{\mathbf{C}}$.

EXAMPLE 6.1 *For isotropic composites, the effective moduli obtained from self-consistent scheme can be further simplified.*

Consider

$$\bar{\mathbf{C}} = \mathbf{C} + \sum_{\alpha=1}^{n} f_\alpha(\mathbf{C}^\alpha - \mathbf{C}) : \bar{\mathcal{A}}^\alpha \tag{6.67}$$

Step 1.

$$\mathbf{C} = 3K\mathbf{E}^{(1)} + 2\mu\mathbf{E}^{(2)}, \text{ and } (\mathbf{C}^\alpha - \mathbf{C}) = 3(K^\alpha - K)\mathbf{E}^{(1)} + 2(\mu^{(\alpha)} - \mu)\mathbf{E}^{(2)}$$

Step 2:

$$\begin{aligned}
\bar{\mathbf{A}}^\alpha &= (\bar{\mathbf{C}} - \mathbf{C}^\alpha)^{-1} : \bar{\mathbf{C}} \\
&= \Big(\frac{1}{3(\bar{K} - K^\alpha)}\mathbf{E}^{(1)} + \frac{1}{2(\bar{\mu} - \mu^\alpha)}\mathbf{E}^{(2)}\Big) : (3\bar{K}\mathbf{E}^{(1)} + 2\bar{\mu}\mathbf{E}^{(2)}) \\
&= \frac{\bar{K}}{\bar{K} - K^\alpha}\mathbf{E}^{(1)} + \frac{\bar{\mu}}{\bar{\mu} - \mu^\alpha}\mathbf{E}^{(2)}
\end{aligned}$$

Then,

$$\begin{aligned}
\bar{\mathcal{A}}^\alpha &= \bar{\mathbf{A}}^\alpha(\bar{\mathbf{A}}^\alpha - \bar{\mathbf{S}}^\alpha)^{-1} \\
&= \Big[\frac{\bar{K}}{\bar{K} - K^\alpha}\mathbf{E}^{(1)} + \frac{\bar{\mu}}{\bar{\mu} - \mu^\alpha}\mathbf{E}^{(2)}\Big]\Big[\Big(\frac{\bar{K}}{\bar{K} - K^\alpha} - \bar{s}_1\Big)^{-1}\mathbf{E}^{(1)} \\
&\quad + \Big(\frac{\bar{\mu}}{\bar{\mu} - \mu^\alpha} - \bar{s}_2\Big)^{-1}\mathbf{E}^{(2)}\Big] \\
&= \frac{\bar{K}}{\bar{K} - (\bar{K} - K^\alpha)\bar{s}_1}\mathbf{E}^{(1)} + \frac{\bar{\mu}}{\bar{\mu} - (\bar{\mu} - \mu^\alpha)\bar{s}_2}\mathbf{E}^{(2)}
\end{aligned}$$

Therefore,

$$\begin{aligned}
\bar{\mathbf{C}} &= 3\bar{K}\mathbf{E}^{(1)} + 2\bar{\mu}\mathbf{E}^{(2)} \\
&= \mathbf{C} + \sum_{\alpha=1}^{n} f_\alpha(\mathbf{C}^\alpha - \mathbf{C}) : \bar{\mathcal{A}}^\alpha \\
&= 3\Big(K + \sum_\alpha f_\alpha \frac{(K^\alpha - K)\bar{K}}{\bar{K} + (\bar{K} - K^\alpha)\bar{s}_1}\Big)\mathbf{E}^{(1)} \\
&\quad + 2\Big(\mu + \sum_\alpha f_\alpha \frac{(\mu^\alpha - \mu)\bar{\mu}}{\bar{\mu} + (\bar{\mu} - \mu^\alpha)\bar{s}_2}\Big)\mathbf{E}^{(2)}
\end{aligned}$$

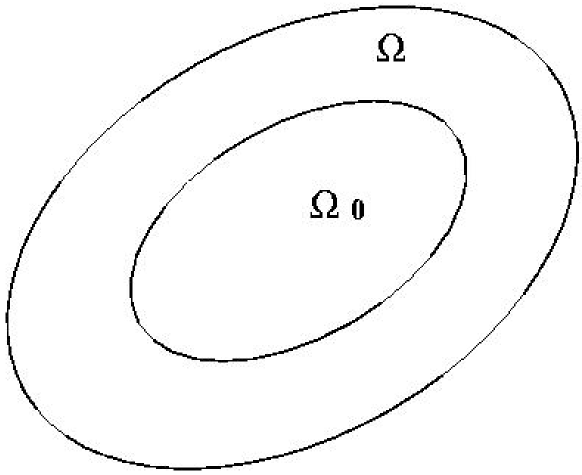


Figure 6.1. Schematic illustration of Mori-Tanaka lemma

which lead to

$$\frac{\bar{K}}{K} = 1 + \sum_{\alpha=1}^{n} f_\alpha \Big(\frac{K^\alpha}{K} - 1\Big)\Big(1 + (\frac{K^\alpha}{\bar{K}} - 1)\bar{s}_1\Big)^{-1} \tag{6.68}$$

$$\frac{\bar{\mu}}{\mu} = 1 + \sum_{\alpha=1}^{n} f_\alpha \Big(\frac{\mu^\alpha}{\mu} - 1\Big)\Big(1 + (\frac{\mu^\alpha}{\bar{\mu}} - 1)\bar{s}_2\Big)^{-1} \tag{6.69}$$

Note that $\nu = \dfrac{3K - 2\mu}{2(3K + \mu)}$.

6.3 Mori-Tanaka methods

6.3.1 Tanaka-Mori lemma

In 1972, a less than two-page technical note was published in *Journal of Elasticity* by Tanaka and Mori (Tanaka and Mori [1972]), which revealed an importance consequence of the scalability of the Eshelby tensor.

That result is the well-known Tanaka-Mori lemma, and it then leads a very effecient homogenization procedure called *Mori-Tanaka method*. Today, the Mori-Tanaka method is one the most popular homogenization methods used in composite industry. Its applications include abraided composite, nano-composites, and reinforce fiber composites.

We start with the Tanaka-Mori lemma first.

LEMMA 6.2 (TANAKA AND MORI) *Consider two coaxial, similar ellipsoidal domains,* Ω_0, Ω $(\Omega_0 \subset \Omega)$,

$$\begin{aligned} \Omega_0 &= \left\{ \mathbf{x} \,\Big|\, \frac{x_1^2}{a_1^2} + \frac{x_2^2}{a_2^2} + \frac{x_3^2}{a_3^2} \leq 1 \right\} \\ \Omega &= \left\{ \mathbf{x} \,\Big|\, \frac{x_1^2}{b_1^2} + \frac{x_2^2}{b_2^2} + \frac{x_3^2}{b_3^2} \leq 1 \right\} \end{aligned} \tag{6.70}$$

where

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} + \frac{a_3}{b_3} = k$$

Assume that a uniform eigenstrain state, $\epsilon_{ij}^*(\mathbf{x})$, *is prescribed in the smaller ellipsoidal region, i.e.*

$$\epsilon_{ij}^*(\mathbf{x}) = \begin{cases} \epsilon_{ij}^* & \mathbf{x} \in \Omega_0 \\ 0 & \mathbf{x} \in \mathbb{R}^3/\Omega_0 \end{cases}$$

The the average disturbance strain field is zero, i.e

$$< \boldsymbol{\epsilon} >_{\Omega-\Omega_0} = \frac{1}{\Omega - \Omega_0} \int_{\Omega-\Omega_0} \epsilon_{ij}(\mathbf{x}) d\Omega = 0 \, . \tag{6.71}$$

Proof:

Suppose that there are three coaixial, similar ellipsoidals, $\Omega_0 \subset \Omega_1 \subset \Omega_2$ in an infinite homogeneous medium, and a uniform eigenstrain is presecibed in Ω_0, i.e.

$$\epsilon_{ij}^*(\mathbf{x}) = \begin{cases} \epsilon_{ij}^* & \mathbf{x} \in \Omega_0 \\ 0 & \mathbf{x} \in \mathbb{R}^3/\Omega_0 \end{cases}$$

The disturbance displacement field can be then written as

$$u_i(\mathbf{x}) = -\int_{\Omega_0} \epsilon_{mn}^* C_{k\ell mn} G_{ik,\ell}(\mathbf{x} - \mathbf{x}') d\mathbf{x}' \tag{6.72}$$

and the disturbance strain field is

$$\epsilon_{ij}(\mathbf{x}) = -\int_{\Omega_0} \epsilon_{mn}^* \frac{C_{k\ell mn}}{2} \Big(G_{ik,\ell j}(\mathbf{x} - \mathbf{x}') + G_{jk,\ell i}(\mathbf{x} - \mathbf{x}') \Big) d\mathbf{x}' \tag{6.73}$$

where $C_{k\ell mn}$ is the elastic tensor, $G_{ik}(\mathbf{x} - \mathbf{x}')$ is the Green's function in the infinite domain, and

$$\begin{aligned} S_{k\ell mn} &= -\int_{\Omega_0} \frac{C_{k\ell mn}}{2} \Big(G_{ik,\ell j} + G_{jk,\ell i} \Big) d\mathbf{x}' \\ &= \begin{cases} S_{k\ell mn}^{\Omega_0}, & \mathbf{x} \in \Omega_0 \\ \\ S_{k\ell mn}^{\infty}, & \mathbf{x} \in \mathbb{R}^3/\Omega_0 \end{cases} \end{aligned} \tag{6.74}$$

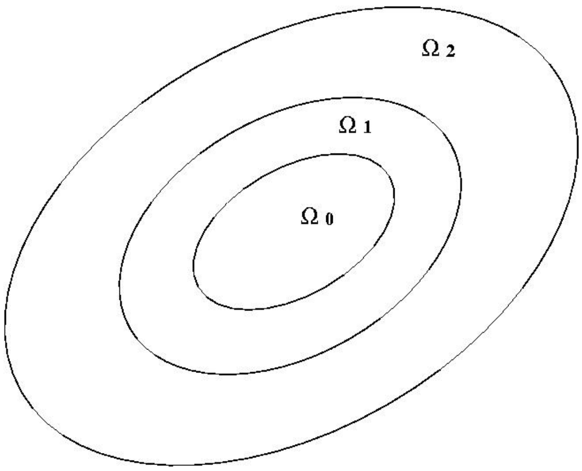


Figure 6.2. Schematic illustration of the Proof of Mori-Tanaka lemma

is the Eshelby tensor.

Now consider the average strain in the region $\Omega_1 - \Omega_2$.

$$\int_{\Omega_2-\Omega_1} \epsilon_{ij}(\mathbf{x})d\mathbf{x} = \int_{\Omega_2-\Omega_1}\Big[\epsilon^*_{mn}\int_{\Omega_0} -\frac{C_{k\ell mn}}{2}\Big(G_{ik,\ell j}(\mathbf{x}-\mathbf{x}')+G_{jk,\ell i}(\mathbf{x}-\mathbf{x}')\Big)d\mathbf{x}'\Big]d\mathbf{x}$$

Since $\mathbf{x} \in \Omega_2 - \Omega_1$, the integrand does contain singularity in either integration domains, Ω_0 and $\Omega_2 - \Omega_1$. We can then change the order of the integration,

$$\begin{aligned}
& \int_{\Omega_2-\Omega_1}\Big[\epsilon^*_{mn}\int_{\Omega_0} -\frac{C_{k\ell mn}}{2}\Big(G_{ik,\ell j}(\mathbf{x}-\mathbf{x}')+G_{jk,\ell i}(\mathbf{x}-\mathbf{x}')\Big)d\mathbf{x}'\Big]d\mathbf{x} \\
= & \int_{\Omega_0}\Big[\epsilon^*_{mn}\int_{\Omega_2-\Omega_1} -\frac{C_{k\ell mn}}{2}\Big(G_{ik,\ell j}(\mathbf{x}-\mathbf{x}')+G_{jk,\ell i}(\mathbf{x}-\mathbf{x}')\Big)d\mathbf{x}'\Big]d\mathbf{x} \\
= & \int_{\Omega_0}\Big[\epsilon^*_{mn}\int_{\Omega_2} -\frac{C_{k\ell mn}}{2}\Big(G_{ik,\ell j}(\mathbf{x}-\mathbf{x}')+G_{jk,\ell i}(\mathbf{x}-\mathbf{x}')\Big)d\mathbf{x}'\Big]d\mathbf{x} \\
& -\int_{\Omega_0}\Big[\epsilon^*_{mn}\int_{\Omega_1} -\frac{C_{k\ell mn}}{2}\Big(G_{ik,\ell j}(\mathbf{x}-\mathbf{x}')+G_{jk,\ell i}(\mathbf{x}-\mathbf{x}')\Big)d\mathbf{x}'\Big]d\mathbf{x} \\
= & \ \epsilon^*_{mn}\int_{\Omega_0}\Big[S^{\Omega_2}_{k\ell mn} - S^{\Omega_1}_{k\ell mn}\Big]d\mathbf{x}' = \epsilon^*_{mn}\Omega_0\Big[S^{\Omega_2}_{k\ell mn} - S^{\Omega_1}_{k\ell mn}\Big]
\end{aligned} \tag{6.75}$$

Since Eshelby tensor only depends on the material property and the aspect ratio of the ellipsoidals,

$$\epsilon^*_{mn}\Omega_0\Big[S^{\Omega_2}_{k\ell mn} - S^{\Omega_1}_{k\ell mn}\Big] = 0 \tag{6.76}$$

if Ω_2, Ω_1 are similar. Hence,

$$\int_{\Omega_2-\Omega_1} \epsilon_{ij}(\mathbf{x}) d\Omega_{\mathbf{x}} = 0 \ . \tag{6.77}$$

Let $\Omega_1 \to \Omega_0$ and $\Omega_2 \to \Omega$. We have the desired result,

$$\int_{\Omega_2-\Omega_1} \epsilon_{ij}(\mathbf{x}) d\Omega_{\mathbf{x}} = \int_{\Omega-\Omega_0} \epsilon_{ij}(\mathbf{x}) d\Omega_{\mathbf{x}} = 0 \ . \tag{6.78}$$

♣

REMARK 6.3.1 *1 It is also true that the average disturbance stress field is also zero*

$$\int_{\Omega-\Omega_0} \sigma_{ij} d\Omega_{\mathbf{x}} = 0 \ . \tag{6.79}$$

2 *Eq. (6.75) is valid as long as $\Omega_1 \subset \Omega_2$. They don't need to be confocal, but they definitely need to be similar, and they may need to be coaxial (some people questioned nececity of this requirment too, the real issue is : does Eshelby tensor depend on coordinates ?).*

3 *This result can be generalized into the cases that the inclusion Ω_0 is not ellipsoidal and the eigenstrain distribution in Ω_0 is not uniform.*

6.3.2 Mori-Tanaka's two-phase model

In this section, we present a straightforward application of Tanaka-Mori lemma for a two-phase double inclusion problem.

We assume that there are only two phases in an RVE, and both the RVE and the inhomogeneity have the shape of ellipsoidal. The are coaxial and similar in shape.

Suppose in the far field, there are constant stress and strain fields, $\boldsymbol{\sigma}^0$ and $\boldsymbol{\epsilon}^0$. Due the presence of inhomogeneity, the total strain and stress fields consist of two parts: constant far fields and perturbed fields, i.e.

$$\boldsymbol{\epsilon}(\mathbf{x}) = \boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d(\mathbf{x}), \quad \forall \mathbf{x} \in V \tag{6.80}$$

$$\boldsymbol{\sigma}(\mathbf{x}) = \boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d(\mathbf{x}), \quad \forall \mathbf{x} \in V \tag{6.81}$$

Inside the inclusion, $\mathbf{x} \in \Omega$, the disturbance field may be expressed in terms of eigenstrain

$$\boldsymbol{\epsilon}^d = \mathbf{S}^\Omega : \epsilon^*, \quad \Rightarrow \quad \boldsymbol{\epsilon}(\mathbf{x}) = \boldsymbol{\epsilon}^0 + \mathbf{S}^\Omega : \boldsymbol{\epsilon}^*, \quad \forall \mathbf{x} \in \Omega \tag{6.82}$$

Therefore,

$$< \boldsymbol{\epsilon} >_\Omega = \boldsymbol{\epsilon}^0 + \mathbf{S}^\Omega : \boldsymbol{\epsilon}^* \tag{6.83}$$

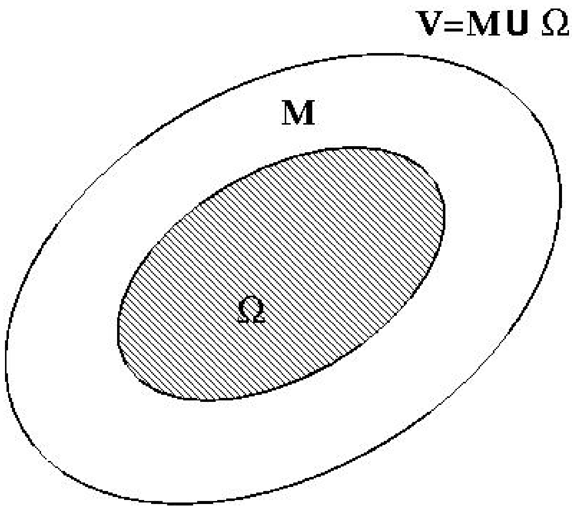


Figure 6.3. Schematic illustration of two-phase model

Recall that the homogenization condition (Eshelby's equivalent principle),

$$\mathbf{C}^{\Omega} : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d) = \mathbf{C} : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d - \boldsymbol{\epsilon}^*), \tag{6.84}$$

let to

$$\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d = \mathbf{A}^{\Omega} : \boldsymbol{\epsilon}^*, \ \ \forall\, \mathbf{x} \in \Omega \tag{6.85}$$

where $\mathbf{A}^{\Omega} = (\mathbf{C} - \mathbf{C}^{\Omega})^{-1} : \mathbf{C}$. Combining with $\boldsymbol{\epsilon}^d = \mathbf{S}^{\Omega} : \boldsymbol{\epsilon}^*$, one can find that

$$\boldsymbol{\epsilon}^* = (\mathbf{A}^{\Omega} - \mathbf{S}^{\Omega})^{-1} : \boldsymbol{\epsilon}^0 \tag{6.86}$$

Substitute (6.86) back to (6.83). We finally have

$$< \boldsymbol{\epsilon} >_{\Omega} = \Big(\mathbf{1}^{(4s)} + \mathbf{S}^{\Omega} : (\mathbf{A}^{\Omega} - \mathbf{S}^{\Omega})^{-1}\Big) : \boldsymbol{\epsilon}^0 \tag{6.87}$$

The average stress inside the inclusion can be also evaluated by considering homogenization condition and (6.86)

$$\begin{aligned} < \boldsymbol{\sigma} >_{\Omega} &= \mathbf{C} : \Big(\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d - \boldsymbol{\epsilon}^*\Big) = \mathbf{C} : \Big(\boldsymbol{\epsilon}^0 + (\mathbf{S}^{\Omega} - \mathbf{1}^{(4s)})\boldsymbol{\epsilon}^*\Big) \\ &= \mathbf{C} : \Big(\mathbf{1}^{(4s)} + (\mathbf{S}^{\Omega} - \mathbf{1}^{(4s)}(\mathbf{A}^{\Omega} - \mathbf{S}^{\Omega})^{-1}\Big) : \boldsymbol{\epsilon}^0 \,. \end{aligned} \tag{6.88}$$

One the other hand, by the Tanaka-Mori lemma, the average strain in the matrix is

$$< \boldsymbol{\epsilon} >_{M} = < \boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d >_{M} = \boldsymbol{\epsilon}^0 \tag{6.89}$$

and hence

$$< \boldsymbol{\sigma} >_M = \mathbf{C} : \boldsymbol{\epsilon}^0 \; . \tag{6.90}$$

Let f be the volume fraction of the inhomogeneity. We then have the following balance equations for average strain and stress

$$\begin{aligned} < \boldsymbol{\epsilon} >_V &= (1-f) < \boldsymbol{\epsilon} >_M + f < \boldsymbol{\epsilon} >_\Omega & (6.91)\\ < \boldsymbol{\sigma} > &= (1-f) < \boldsymbol{\sigma} >_M + f < \boldsymbol{\sigma} >_\Omega & (6.92) \end{aligned}$$

One can readily find that

$$\begin{aligned} < \boldsymbol{\epsilon} >_V &= (1-f)\boldsymbol{\epsilon}^0 + f(\boldsymbol{\epsilon}^0 + \mathbf{S}^\Omega : \boldsymbol{\epsilon}^*\Big) \\ &= \boldsymbol{\epsilon}^0 + f\mathbf{S}^\Omega(\mathbf{A}^\Omega - \mathbf{S}^\Omega) : \epsilon^0 \\ &= \Big(\mathbf{1}^{(4s)} + f\mathbf{S}^\Omega(\mathbf{A}^\Omega - \mathbf{S}^\Omega)^{-1}\Big) : \epsilon^0 \end{aligned} \tag{6.93}$$

and

$$\begin{aligned} < \boldsymbol{\sigma} >_V &= (1-f)\mathbf{C} : \boldsymbol{\epsilon}^0 + f\mathbf{C} : \Big(\mathbf{1}^{(4s)} + (\mathbf{S}^\Omega - \mathbf{1}^{(4s)}(\mathbf{A}^\Omega - \mathbf{S}^\Omega)^{-1}\Big) : \boldsymbol{\epsilon}^0 \; . \\ &= \mathbf{C} : \Big(\mathbf{1}^{(4s)} + f(\mathbf{S}^\Omega - \mathbf{1}^{(4s)}(\mathbf{A}^\Omega - \mathbf{S}^\Omega)^{-1}\Big) : \boldsymbol{\epsilon}^0 \; . \end{aligned} \tag{6.94}$$

By definition,

$$< \boldsymbol{\sigma} >_V = \mathbf{C} :< \boldsymbol{\epsilon} >_V \tag{6.95}$$

It leads to

$$\mathbf{C} : \Big(\mathbf{1}^{(4s)} + f(\mathbf{S}^\Omega - \mathbf{1}^{(4s)})(\mathbf{A}^\Omega - \mathbf{S}^\Omega)^{-1}\Big) : \boldsymbol{\epsilon}^0 = \bar{\mathbf{C}} : \Big(\mathbf{1}^{(4s)} + f\mathbf{S}^\Omega(\mathbf{A}^\Omega - \mathbf{S}^\Omega)^{-1}\Big) : \epsilon^0$$

Finally, the effective elastic tensor is obtained

$$\bar{\mathbf{C}} = \mathbf{C} : \Big(\mathbf{1}^{(4s)} + f(\mathbf{S}^\Omega - \mathbf{1}^{(4s)})(\mathbf{A}^\Omega - \mathbf{S}^\Omega)^{-1}\Big) : \Big(\mathbf{1}^{(4s)} + f\mathbf{S}^\Omega(\mathbf{A}^\Omega - \mathbf{S}^\Omega)^{-1}\Big)^{-1} \tag{6.96}$$

6.3.3 Mori-Tanaka mean field theory

In previous homogenization procedures, the disturbance strain and stress fields due to an inhomogeneity are approximated by Eshelby's single inclusion solution in an infinte space.

In real applications, an RVE is finite, and it is subjected with remote boundary conditions, e.g. prescribed traction condition or prescribed displacement condition, i.e.

$$\mathbf{u} = \mathbf{x} \cdot \boldsymbol{\epsilon}^0, \;\; \mathbf{x} \in \partial V \tag{6.97}$$

or

$$\mathbf{t} = \mathbf{n} \cdot \boldsymbol{\sigma}^0, \;\; \mathbf{x} \in \partial V \tag{6.98}$$

Let $\boldsymbol{\epsilon}^{pt}$ and $\boldsymbol{\sigma}^{pt}$ representing perturbed strain and stress fields due to Eshelby's single inclusion solution in an infinite medium. If we let

$$\begin{aligned} \boldsymbol{\epsilon}(\mathbf{x}) &= \boldsymbol{\epsilon}_0 + \boldsymbol{\epsilon}^d = \boldsymbol{\epsilon}_0 + \boldsymbol{\epsilon}^{pt} && (6.99) \\ \boldsymbol{\sigma}(\mathbf{x}) &= \boldsymbol{\sigma}_0 + \boldsymbol{\sigma}^d = \boldsymbol{\sigma}_0 + \boldsymbol{\sigma}^{pt} && (6.100) \end{aligned}$$

Obviously,

$$\boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^{pt} \neq \boldsymbol{\sigma}^0, \quad \text{or} \quad \boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^{pt} \neq \boldsymbol{\epsilon}^0, \ \ \forall \, \mathbf{x} \in \partial V \qquad (6.101)$$

Therefore, either boundary condition (6.98) and (6.97) will not be satisfied. This is because a finite size RVE will cause additional interaction between matrix and inclusions, interaction between the boudary and inclusions, and interaction among inclusions themself. Note that $\boldsymbol{\epsilon}^{pt}, \boldsymbol{\sigma}^{pt}, \rightarrow \ 0$ only when $|\mathbf{x}| \rightarrow \infty$.

To take into account the effects of a finite size RVE, additional stress and strain fields are need to faithfully represent total stress and strain distribution in an RVE, i.e.

$$\begin{aligned} \boldsymbol{\sigma} &= \boldsymbol{\sigma}^0 + \tilde{\boldsymbol{\sigma}} + \boldsymbol{\sigma}^{pt} && (6.102) \\ \boldsymbol{\epsilon} &= \boldsymbol{\epsilon}^0 + \tilde{\boldsymbol{\epsilon}} + \boldsymbol{\epsilon}^{pt} && (6.103) \end{aligned}$$

where $\tilde{\boldsymbol{\sigma}}$ and $\tilde{\boldsymbol{\epsilon}}$ are the so-called image stress and image strain.

In literature, especially literatures on dislocations, additional stress and strain fields that accommodate the stress solution of a infinite space to satisfy boundary conditions are called image stress and image strain fields, because in practice some of these stress and strain fields are found by placing certain image external sources to achieve their objectives.

Nevertheless, the homogenization problem in a finite REV becomes complicated, because in general it is very difficult to know the precise distribution of image stress and image strain fields. To circumvent this difficulty, Mori and Tanaka [1973] proposed the following mean field assumption, which is an ingenous and very successful method.

Mori & Tanaka's theory was later refined in a landmark paper by G. Weng (Weng [1990]). The following presentation is an adaption of Weng's formulation. Suppose that in an RVE there are many inhomogeneities, or the density of inhomogeneities are statistically stable. Then the strain or stress field in the matrix may be written as

$$\begin{aligned} \boldsymbol{\epsilon}(\mathbf{x}) &= \boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d, \ \forall \mathbf{x} \in M \ \ \Rightarrow \ \ <\boldsymbol{\epsilon}>_M = \boldsymbol{\epsilon}^0 + <\boldsymbol{\epsilon}^d>_M; \\ \boldsymbol{\sigma}(\mathbf{x}) &= \boldsymbol{\sigma}^0 + \boldsymbol{\sigma}^d, \ \forall \mathbf{x} \in M \ \ \Rightarrow \ \ <\boldsymbol{\sigma}>_M = \boldsymbol{\sigma}^0 + <\boldsymbol{\sigma}^d>_M; \end{aligned}$$

In general we don't know the precise disturbance fields in a matrix, i.e., $\boldsymbol{\epsilon}^d_M$ or $\boldsymbol{\sigma}^d_M$.

Consider the matrix is the dominate phase in a composite. We denote the average field in the matrix, $<\boldsymbol{\epsilon}>_M$ or $<\boldsymbol{\sigma}>_M$, as the **mean field**, which include boundary effects and effects of interactions of many other inclusions.

Now we add an inclusion into the average ensemble—the RVE. After the inclusion is added, we call the field as the new field in contrast with the old field before the inclusion is being added. Therefore, in the matrix,

$$<\boldsymbol{\epsilon}^{new}>_M=<\boldsymbol{\epsilon}^{old}>_M+<\boldsymbol{\epsilon}^{pt}>_M+<\boldsymbol{\epsilon}^{im}>_M,\ \ \forall \mathbf{x}\in M \qquad (6.104)$$

where $\boldsymbol{\epsilon}^{pt}$ and $\boldsymbol{\epsilon}^{im}$ are the inclusion solution for infinite space and the corresponding image strain solution due to the finite RVE.

By the Tanaka-Mori lemma, $<\boldsymbol{\epsilon}^{pt}>_M=0$. Mori and Tanaka then further argued that since there have been so many inclusions inside the RVE, the average effects of the image strain or image stress field for a single inclusion may be negligible without alter the mean field of value of the RVE, i.e. $<\boldsymbol{\epsilon}^{im}>_M=0$, which is the essence of Mori-Tanaka mean field theory. Note that $<\boldsymbol{\epsilon}^{old}>_M$ does take into account the average effects of the image stress/strain fields all other inclusions.

Therefore, we have

$$<\boldsymbol{\epsilon}^{new}>_M=<\boldsymbol{\epsilon}^{old}>_M=<\boldsymbol{\epsilon}>_M\ \ ,\forall \mathbf{x}\in M\ . \qquad (6.105)$$

Inside the inclusion, we still neglect the effects of image strain or image stress field of the newly added inclusion, we then have

$$\begin{aligned}<\boldsymbol{\epsilon}>_\Omega &= <\boldsymbol{\epsilon}>_M+<\boldsymbol{\epsilon}^{pt}>_\Omega+<\boldsymbol{\epsilon}^{im}>_\Omega\\ &= <\boldsymbol{\epsilon}>_M+<\boldsymbol{\epsilon}^{pt}>_\Omega\\ &= <\boldsymbol{\epsilon}>_M+\mathbf{S}^\Omega:\boldsymbol{\epsilon}^*,\ \forall \mathbf{x}\in\Omega \end{aligned} \qquad (6.106)$$

Similarly, for the stress field,

$$\begin{aligned}<\boldsymbol{\sigma}^{new}>_M &= <\boldsymbol{\sigma}^{old}>_M=<\boldsymbol{\sigma}>_M,\ \ \mathbf{x}\in M\\ <\boldsymbol{\sigma}>_\Omega &= <\boldsymbol{\sigma}>_M+\mathbf{T}^\Omega:\boldsymbol{\sigma}^*,\ \ \mathbf{x}\in\Omega \end{aligned} \qquad (6.107)$$

Based on Eshelby's equivalence homogenization conditions,

$$\mathbf{C}^\Omega:<\boldsymbol{\epsilon}>_\Omega=\mathbf{C}:\Big(<\boldsymbol{\epsilon}>_\Omega-\boldsymbol{\epsilon}^*\Big) \qquad (6.108)$$

or

$$\mathbf{D}^\Omega:<\boldsymbol{\sigma}>_\Omega=\mathbf{D}:\Big(<\boldsymbol{\sigma}>_\Omega-\boldsymbol{\sigma}^*\Big) \qquad (6.109)$$

One may obtain

$$\begin{aligned}<\boldsymbol{\epsilon}>_\Omega &= \mathbf{A}^\Omega:\boldsymbol{\epsilon}^* \ \Rightarrow\ <\boldsymbol{\epsilon}>_M+<\boldsymbol{\epsilon}^{pt}>_\Omega=\mathbf{A}^\Omega:\boldsymbol{\epsilon}^*\\ \text{or}\ \ <\boldsymbol{\sigma}>_\Omega &= \mathbf{B}^\Omega:\boldsymbol{\sigma}^* \ \Rightarrow\ <\boldsymbol{\sigma}>_M+<\boldsymbol{\sigma}^{pt}>_\Omega=\mathbf{A}^\Omega:\boldsymbol{\sigma}^*\end{aligned}$$

where $\mathbf{A}^\Omega := (\mathbf{C} - \mathbf{C}^\Omega)^{-1} : \mathbf{C}$ and $\mathbf{B}^\Omega := (\mathbf{D} - \mathbf{D}^\Omega)^{-1} : \mathbf{D}$.

Subsequently, one can obtain that

$$\begin{aligned} <\boldsymbol{\epsilon}>_\Omega &= \mathcal{A}^{dil}_\Omega :<\boldsymbol{\epsilon}>_M \\ \text{or} \quad <\boldsymbol{\sigma}>_\Omega &= \mathcal{B}^{dil}_\Omega :<\boldsymbol{\sigma}>_M \end{aligned} \tag{6.110}$$

according to different boundary conditions or different homonization schems.

In passing, we note that that the concentration tensors may be written in different forms,

$$\begin{aligned} \mathcal{A}^{dil}_\Omega &= \mathbf{A}^\Omega : (\mathbf{A}^\Omega - \mathbf{S}^\Omega)^{-1} = \Big[(\mathbf{A}^\Omega - \mathbf{S}^\Omega) : \mathbf{A}^{\Omega-1}\Big]^{-1} \\ &= \Big[\mathbf{1} - \mathbf{S}^\Omega : \mathbf{A}^{\Omega-1}\Big]^{-1} \\ &= \Big[\mathbf{1} - \mathbf{S}^\Omega : \mathbf{C}^{-1} : (\mathbf{C} - \mathbf{C}^\Omega)\Big]^{-1} \\ &= \Big[\mathbf{1} + \mathbf{P}^\Omega : (\mathbf{C}^\Omega - \mathbf{C})\Big]^{-1} \end{aligned} \tag{6.111}$$

and

$$\begin{aligned} \mathcal{B}^{dil}_\Omega &= \mathbf{B}^\Omega : (\mathbf{B}^\Omega - \mathbf{T}^\Omega)^{-1} = \Big[(\mathbf{B}^\Omega - \mathbf{T}^\Omega) : \mathbf{A}^{\Omega-1}\Big]^{-1} \\ &= \Big[\mathbf{1} - \mathbf{T}^\Omega : \mathbf{B}^{\Omega-1}\Big]^{-1} \\ &= \Big[\mathbf{1} - \mathbf{T}^\Omega : \mathbf{D}^{-1} : (\mathbf{D} - \mathbf{D}^\Omega)\Big]^{-1} \\ &= \Big[\mathbf{1} + \mathbf{Q}^\Omega : (\mathbf{D}^\Omega - \mathbf{D})\Big]^{-1} \end{aligned} \tag{6.112}$$

where

$$\mathbf{P}^\Omega = \mathbf{S}^\Omega : \mathbf{C}^{-1} \tag{6.113}$$

$$\mathbf{Q}^\Omega = \mathbf{T}^\Omega : \mathbf{D}^{-1} \tag{6.114}$$

are called polarization tensors.

Since $\mathbf{C} - \mathbf{C}^M = \mathbf{0}$ and $\mathbf{D} - \mathbf{D}^M = \mathbf{0}$, it is easy to see that both

$${\mathcal{A}^{dil}}_M = \mathbf{1} \ \text{ and } \ {\mathcal{B}^{dil}}_M = \mathbf{1}\ . \tag{6.115}$$

By definition,

$$<\boldsymbol{\epsilon}> = (1-f) <\boldsymbol{\epsilon}>_M + f <\boldsymbol{\epsilon}>_\Omega \tag{6.116}$$

$$<\boldsymbol{\sigma}> = (1-f) <\boldsymbol{\sigma}>_M + f <\boldsymbol{\sigma}>_\Omega \tag{6.117}$$

From (6.116) and (6.117), we may find that

$$
\begin{aligned}
<\boldsymbol{\epsilon}>_M &= \Big[(1-f)\mathbf{1} + f\mathcal{A}^{dil}{}_\Omega\Big]^{-1} :<\boldsymbol{\epsilon}> \\
&= \Big[f_M\mathcal{A}^{dil}{}_M + f_\Omega\mathcal{A}^{dil}{}_\Omega\Big]^{-1} :<\boldsymbol{\epsilon}>= \tilde{\mathbf{A}}_0 :<\boldsymbol{\epsilon}> \qquad (6.118)\\
<\boldsymbol{\sigma}>_M &= \Big[f_M\mathcal{B}^{dil}{}_M + f_\Omega\mathcal{B}^{dil}{}_\Omega\Big]^{-1} :<\boldsymbol{\sigma}>= \tilde{\mathbf{B}}_0 :<\boldsymbol{\sigma}> \qquad (6.119)
\end{aligned}
$$

where

$$
\begin{aligned}
\tilde{\mathbf{A}}_0 &:= \Big[f_M\mathcal{A}^{dil}{}_M + f_\Omega\mathcal{A}^{dil}{}_\Omega\Big]^{-1} \qquad (6.120)\\
\tilde{\mathbf{B}}_0 &:= \Big[f_M\mathcal{B}^{dil}{}_M + f_\Omega\mathcal{B}^{dil}{}_\Omega\Big]^{-1} \qquad (6.121)
\end{aligned}
$$

Accordingly,

$$
\begin{aligned}
<\boldsymbol{\epsilon}>_\Omega &= \mathcal{A}^{dil}{}_\Omega :<\epsilon>_M = \mathcal{A}^{dil}{}_\Omega : \tilde{\mathbf{A}}_0 :<\boldsymbol{\epsilon}> \qquad (6.122)\\
<\boldsymbol{\sigma}>_\Omega &= \mathcal{B}^{dil}{}_\Omega :<\boldsymbol{\sigma}>_M = \mathcal{B}^{dil}{}_\Omega : \tilde{\mathbf{B}}_0 :<\boldsymbol{\sigma}> \qquad (6.123)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
<\boldsymbol{\sigma}> &= f_M <\boldsymbol{\sigma}>_M + f_\Omega <\boldsymbol{\sigma}>_\Omega \\
&= f_M\mathbf{C}^0 <\boldsymbol{\epsilon}>_M + f_\Omega\mathbf{C}^\Omega <\boldsymbol{\epsilon}>_\Omega \\
&= f_M\mathbf{C}^0 <\boldsymbol{\epsilon}>_M + f_\Omega\mathbf{C}^\Omega\mathcal{A}^{dil}{}_\Omega <\boldsymbol{\epsilon}>_M \\
&= \Big(f_M\mathbf{C}^0 + f_\Omega\mathbf{C}^\Omega\mathcal{A}^{dil}{}_\Omega\Big)\tilde{\mathbf{A}}_0 <\boldsymbol{\epsilon}> \\
&= \bar{\mathbf{C}} :<\boldsymbol{\epsilon}> \qquad (6.124)
\end{aligned}
$$

and

$$
\begin{aligned}
<\boldsymbol{\epsilon}> &= f_M <\epsilon>_M + f_\Omega <\boldsymbol{\epsilon}>_\Omega \\
&= f_M\mathbf{D}^0 <\boldsymbol{\sigma}>_M + f_\Omega\mathbf{D}^\Omega :<\boldsymbol{\sigma}>_\Omega \\
&= f_M\mathbf{D}^0 <\boldsymbol{\sigma}>_M + f_\Omega\mathbf{D}^\Omega : \mathcal{B}^{dil}{}_\Omega :<\boldsymbol{\sigma}>_M \\
&= \Big(f_M\mathbf{D}^0 + f_\Omega\mathbf{D}^\Omega : \mathcal{B}^{dil}{}_\Omega\Big)\tilde{\mathbf{B}}_0 :<\boldsymbol{\epsilon}> \\
&= \bar{\mathbf{D}} :<\boldsymbol{\sigma}> \qquad (6.125)
\end{aligned}
$$

Recall that $\mathcal{A}^{dil}{}_M = \mathcal{B}^{dil}{}_M = \mathbf{1}$. We have

$$
\begin{aligned}
\bar{\mathbf{C}} &= \Big(f_M\mathbf{C}^0 : \mathcal{A}^{dil}{}_M + f_\Omega\mathbf{C}^\Omega : \mathcal{A}^{dil}{}_\Omega\Big) : \Big(f_M\mathcal{A}^{dil}{}_M + f_\Omega\mathcal{A}^{dil}{}_\Omega\Big)^{-1} \\
\bar{\mathbf{D}} &= \Big(f_M\mathbf{D}^0 : \mathcal{B}^{dil}{}_M + f_\Omega\mathbf{D}^\Omega : \mathcal{B}^{dil}{}_\Omega\Big) : \Big(f_M\mathcal{B}^{dil}{}_M + f_\Omega\mathcal{B}^{dil}{}_\Omega\Big)^{-1}
\end{aligned}
$$

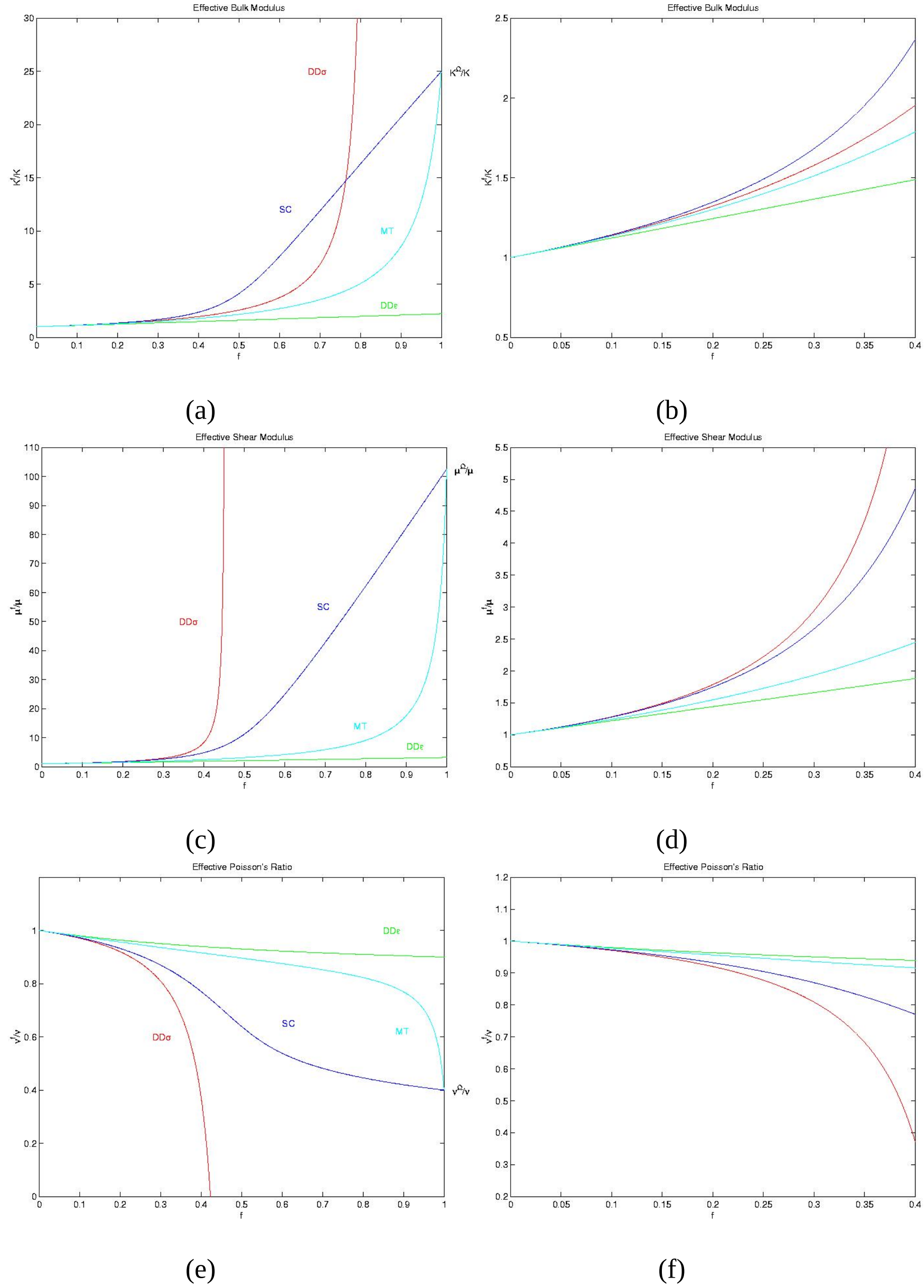


Figure 6.4. Comaprison of effective bulk modulus among various homogenization methods: dilute distribution (DD & DT), self-consistent, and Mori-Tanaka

In general, for a solid with n+1 phases (from $\alpha = 0$ to $\alpha = 0$), the Mori-Tanaka mean field theory gives the following estimates,

$$
\begin{aligned}
\bar{\mathbf{C}} &= \Big(\sum_{\alpha=0}^{n} f_\alpha \mathbf{C}^\alpha : \mathcal{A}^{dil}{}_\alpha\Big) : \Big(\sum_{\alpha=0}^{n} f_\alpha \mathcal{A}^{dil}{}_\alpha\Big)^{-1} \\
\bar{\mathbf{D}} &= \Big(\sum_{\alpha=0}^{n} f_\alpha \mathbf{D}^\alpha : \mathcal{B}^{dil}{}_\alpha\Big) : \Big(\sum_{\alpha=0}^{n} f_\alpha \mathcal{B}^{dil}{}_\alpha\Big)^{-1} \qquad (6.126)
\end{aligned}
$$

where the pahse $\alpha = 0$ represents the matrix, and non-zero α represents the inhomogeneous phases.

Figure 6.5. Rodney Hill

6.4 Rodney Hill

Rodney Hill was born on the 11th June 1921 at Stourton, near Leeds, in Yorkshirt. He comes from a family with deep roots in the practical and culture tradtions of the West Riding, although with no known mathematical ability in an earlier generation. Rodney's father, Harold Harrison Hill, had been an only child and he was educated at the University of Leeds, gaining an M. A. for postgraduate work in history. He also took an external London degree in economics. After wartime service in the Royal Navy he became a schoolmaster, and was eventually senior History Master at Leeds Boy's Modern School.

Rodney's mother had been a student at Leeds School of Art. Rodney himself was also an only child, in an immediate home background which encouraged scholarship and self-sufficiency.

Rodney entered Leeds Grammar School with a scholarship in 1932, and there gave regular prize-wining evidence of all-round intellectual ability not only in mathematics, but equally in art, English literature, and other Arts subjects. During this period he taught himeself to play the piano, and became proficient at chess in which he was later to represent Cambridge University and town. Thus were developing the powers of accurate observation and analysis to be brought to bear on the mathematics and physics which became his formal specialism from the age of 15. The customary large-team games did not attract him as school, but Rodney enjoyed the one-to-one sports of squash, fencing, and golf. He left school as Head of House, and in December 1938 he was awarded an Open Major Scholarship at Pembroke College, Cambridge. However, it needed the State and County Scholarships gained in the preceding summer to make a financially independent undergraduate.

Hill went up to Cambridge to read Mathematics in October 1939, againt a background of external events which must have seemed the least auspicious since the very founding of the University. Major Scholars were expected to take Part II of the Tripos in two years instead of three by omitting all first-year courses. This imposed a heavy workload, to be carried under spartan conditions created by wartime restrictions such as blackout and rationing combined with antique College plumbing. For example, there was no running hot water, the nearest bath was courts away, and the winter allocation of one sack of coal per week fuelled a fire in one's room only in the evenings. Hill was not deflected by the adverse general situation from his aim of a first-class honours degree, and he became a Wrangler in June 1941. This entitled him to take Part III of the mathematical Tripos, in the applied mathematical part of which quantum mechanics figured prominently at the time. However, he felt obliged to war-work, and so lost the opportunity for advanced training which those lecture courses would have provided.

....

Problems brought to the Theoretical Research Branch were distributed initially according to specialisms of the more senior members, some of whom had acquired relevant experience at Woolwich Arsenal. Those problems which were quite new in context tended to go to the young inexperienced graduates newly arrived from university. This was indeed a baptism of fire for them, but it was a test which was to reveal Hill's true metier. One of his initial assigments was the deep penetration of very thick armour by Munoroe jets and high-velocity shells with tungsten-carbide cores. This required a mechanics of plastic deformation with unlimited magnitude, and thus was aroused Hill's interest in the field in which he later became perhapse the foremost world authority. At this stage, however, he had no prior knowledge of the physics and metallurgy of plasticity, and little of stress, strain or the tensors which the mathematics would eventually require. There was no useful textbook, but G. I. Taylor had written one or two helpful reports on shaped charges and Munros jets. Nevertheless, working at first with Mott and Pack, Hill was soon able to show, for example, that penetration by a tungsten–carbide core with pure ogival head would be seriously degraded if too much of the tip were ground conical (the British practice for manufacturing convenience). The demonstration was achieved not only theoretically, but also in field trials planned by Hill in collaboration with an experimental group under Dr. Charles Sykes, F.R.S.

The problems at Fort Halstead called for simple but effective mathematics guided by physical intuiation and a willingness to communicate with others, including non-mathematicians and experimentalists. There was not time for complicated mathematics, there were no electronic computers to assist it, and the experimental data were ususally too crude to warrant it anyway. He acquired a lasting taste for a pragmatic blend of rigour, elgance, and simple realism in the application of mathematics.

The sense of purpose discovered at this time was noticed by colleagues as a cheerful and sparking earnestness. Popular relaxations among the group at Cambridge had included music, books, and lightning chess. At Fort Halstead ballroom dancing was added as a consuming passion for some, and Hill was not slow to find that he had medal-winning ability in this new enthusiasm. He met his future wife, Jeanne Wickens, early in 1945. She had been transferred to work at Fort Halstead from the bombing range at Shoeburyness. Previously she had trained as a dancer and teacher of ballet, but war cut short a promising career. They were married in Cambridge in 1946, and they have one duaghter, born in 1955. The strength of his wife's support could already be detected in the Preface to Hill's first book, and the passage of years has happily reinforced this bond.

By this time the applied mechanics of both solid and fluids was being forced to push the boat out onto a sea of nonlinear problems, and away from the haven linearity in which much pre-war work had lingerd. The trend was evident not only in England, of course, but in other countries too. Hill found himself in demand as the sole adviser on continuum plasticity in England, not only concerning problems arising from the interests at Fort Halstead, but also fot new theories of metal-working processes needed by engineers in the steel industry. He obtained a Cambridge Ph.D. in 1948 for a Thesis entitled "Theoretical studies of the plastic deformation of metals". From the Ph.D. Thesis grew a much more extensive monograph on "The Mathematical Theory of Plasticity", published at the Clarendon Press, Oxford, in 1950. This very rapidly established Hill as an international authority. The final draft was written in his spare time, i.e. in the evenings and weekends. He was then still only in his 28th year, and it is timely to recall a remark from the review of the book in Engineering: "The author has done his work so well that it is difficult to see how it could be bettered. The book should rank for many years as an authoritative source of reference." This prognostication was fully borne out. The book was in print at Oxford for 21 years, Japanese and Russian translations have been made, and total sales currently approach 13,000.

The Journal of Mechanics and Physics of Solids was launched with the encouragement of the infan Pergamon Press in 1952. Hill suggested the title and the general aim of a forum for effective applied mathematics, linked with experimenation, in engineering science. From the onwards the Journal has been regarded as among the foremost in its general field, and unique in flavor. Hill served as Eidtor-in-Chief untill handing over in 1968 to H.G. Hopkins.

The University of Nottinggham had received its Chater, and independence from London, in 1948, and was shortly to embark on two decades substantial expansion. Professor H. R. Pitt was appointed in 1950 to head the existing Mathematics Department, and he was soon instrumental in securing the creation of a new Chair of Applied Mathematics. Rodney Hill applied, and was offered the post in 1953 while still on 31. It was his responsibility to modernize the teaching of applied mathematics. Hill took over some existing course himself, and instigated new ones with the aim of encouraging research students. His undergraduate lectures were characterized by conciseness and tendency to brevity. He would never exceed the time limit. But those stidents who took the trouble to write down what he said, in addition to what was written on the blackboard, found after reflection that they had a first-calss and substantial set of notes.

It may only have been a coincidence that emergence of interest in the so-called rational continuum mechanics was taking place in some American and British universities at this time. Hill's writings demonstrate an independent view of these development, and no taste at all for axiomatics. He was beginning to lay down the basis of general studies of non-uniqueness and instability in continua which were to prov highly influential over the next two decades, and which in due course brought further students and able collaborators.

The University of Cambridge conferred the degree of Sc. D. upon Rodney Hill in 1959. The highest honour to which any British scientist aspires followed in 1961, when he was elected a Fellow of the Royal Society. This gave much pleasure to his colleagues at Nottingham and to his friends elsewhere.

In 1963 Hill was elected to a Berkeley Bye-Fellowship at Gonville and Caius College, Cambridge. This he held for 6 years until the University conferred a personal Readership in Mechanics of Solids. Thus he became a member of the teaching staff of the Department of Applied Mathematics and Theoretical Physics, and in 1972 a personal Professorship was conferred.

During this Cambridge period (he is still at Cambridge under semi-retirement—Li's comment), properties of heterogeneous media (including fibre composites), single crystals, continuum plasticity, and an independent reformulation of rubber elasticity were explored, ...

His standards of scholarship and intellectual honesty are the highest. He is ready in his appreciation of the good work of others; and he has been sharp in candid criticism of misguided thinking or slack presentation (especially by those mature enough to know better) if he thought the subject-matter would be best served thereby—as some celebrated footnotes and book reviews testify.

The outward character of the man is not unlike his papers: physically tall and slim, with the long fingers of a pianist, and having a quiet but compelling presence. His unusally deep reserve has meant that casual social gatherings and conferences have held less interest and been less rewarding for him than for others.

—– By Geoffery Hopkins and Michael Sewell
From *Mechanics of Solids* Pergamon Press

6.5 Exercises

PROBELM 6.1 *Consider a n-phase composite material, and each phase has its own elastic tensor* $\mathbf{C}^\alpha$*, compliance tensor* $\mathbf{D}^\alpha$*; and matrix has elastic tensor,* $\mathbf{C}$*, and compliance tensor,* $\mathbf{D}$*. Assume that in the representative volume element (RVE), each phase only appears as one ellipsoidal inclusion. Under dilute distribution assumption, the corresponding Eshelby tensor and conjugate Eshelby tensor for each phase are* $\mathbf{S}^\alpha$ *and* $\mathbf{T}^\alpha$ *respectively. Denote*

$$\mathbf{A}^\alpha = (\mathbf{C} - \mathbf{C}^\alpha)^{-1} : \mathbf{C} \qquad (6.127)$$

$$\mathbf{B}^\alpha = (\mathbf{D} - \mathbf{D}^\alpha)^{-1} : \mathbf{D} \qquad (6.128)$$

Show

$$\mathbf{C}^\alpha : \mathbf{A}^\alpha : (\mathbf{A}^\alpha - \mathbf{S}^\alpha)^{-1} : \mathbf{D} = \mathbf{B}^\alpha : (\mathbf{B}^\alpha - \mathbf{T}^\alpha)^{-1} \qquad (6.129)$$

$$\mathbf{D}^\alpha : \mathbf{B}^\alpha : (\mathbf{B}^\alpha - \mathbf{T}^\alpha)^{-1} : \mathbf{C} = \mathbf{A}^\alpha : (\mathbf{A}^\alpha - \mathbf{S}^\alpha)^{-1} \qquad (6.130)$$

PROBELM 6.2 *For an isotropic two phase material. Assume the inhomogeneity phase is random distributed spherical cavities (*$\mu_I = 0; K_I = 0$*), and the matrix is an incompressible masterial (*$K \to \infty$*). Use the self-consistent scheme,*

$$\frac{\bar{K}}{K} = 1 + \sum_{\alpha=1}^{n} f_\alpha \Big(\frac{K^\alpha}{K} - 1\Big)\Big(1 + (\frac{K^\alpha}{\bar{K}} - 1)\bar{s}_1\Big)^{-1} \qquad (6.131)$$

$$\frac{\bar{\mu}}{\mu} = 1 + \sum_{\alpha=1}^{n} f_\alpha \Big(\frac{\mu^\alpha}{\mu} - 1\Big)\Big(1 + (\frac{\mu^\alpha}{\bar{\mu}} - 1)\bar{s}_2\Big)^{-1} \qquad (6.132)$$

where

$$\bar{s}_1 = \frac{1+\bar{\nu}}{3(1-\bar{\nu})} \tag{6.133}$$

$$\bar{s}_2 = \frac{2(4-5\bar{\nu})}{15(1-\bar{\nu})} \tag{6.134}$$

to find the effective bulk modulus, $\bar{K}$, and the effective shear modulus, $\bar{\mu}$.

Hint:

J.R. Willis, "Variational and related methods for the overall properties of composite", in Advance in Applied Mechanics, Edited by C.-S. Yih (pages 45-46), (1981), Academic Press, New York.

B. Budiansky, "On the elastic moduli of some heterogeneous materials", Journal of Mechanics and Physics of Solids, Vol. 13, (1965), pages 223-227.

PROBELM 6.3 *Assume that in an RVE there are n+1 phases, $\alpha = 0, 1, \cdots, n$ Mori-Tanaka mean theory states that*

$$\bar{\mathbf{D}} = \sum_{\alpha=0}^{n} f_\alpha \mathbf{D}_\alpha : \mathcal{B}^{dil}{}_\alpha : \Big(\sum_{\alpha=0}^{n} f_\alpha \mathcal{B}^{dil}{}_\alpha\Big)^{-1} \tag{6.135}$$

$$\bar{\mathbf{C}} = \sum_{\alpha=0}^{n} f_\alpha \mathbf{C}_\alpha : \mathcal{A}^{dil}{}_\alpha : \Big(\sum_{\alpha=0}^{n} f_\alpha \mathcal{A}^{dil}{}_\alpha\Big)^{-1} \tag{6.136}$$

Show that Mori-Tanaka scheme is self-consistent, i.e.

$$\bar{\mathbf{C}} = \bar{\mathbf{D}}^{-1} \tag{6.137}$$

Hint: First show that

$$\mathbf{C}^\alpha : \mathcal{A}^{dil}{}_\alpha = \mathcal{B}^{dil}{}_\alpha : \mathbf{C}^0 \tag{6.138}$$

$$\mathbf{D}^\alpha : \mathcal{B}^{dil}{}_\alpha = \mathcal{A}^{dil}{}_\alpha : \mathbf{D}^0 \tag{6.139}$$

PROBELM 6.4 *Consider a two-phase composite with randomly distributed spherical inclusions. The ratios of material constants between inhomogeneity and matrix are*

$$\frac{K^\Omega}{K} = 25, \text{ and } K^\Omega = 750MP_a \tag{6.140}$$

$$\frac{\nu^\Omega}{\nu} = 4, \text{ and } \nu^\Omega = 0.4 \tag{6.141}$$

Plot the ratio of $\dfrac{\bar{K}}{K}$, $\dfrac{\bar{\mu}}{\mu}$, and $\dfrac{\bar{\nu}}{\nu}$ verses the volume fraction of inhomogeneity, f, by using homogenization methods under the assumption of dilute suspension (both prescribed traction and prescribed displacement), self-consistent method, and Mori-Tanaka mean field method.

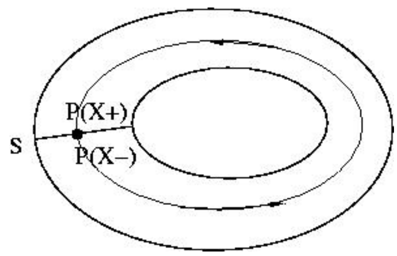


Figure 6.6. Definition of the Volterra dislocation

Chapter 7

INTRODUCTION OF DISLOCATION THEORY

In material science, a dislocation may be defined as a disturbed region between two substantially perfect parts of a crystal. In elasticity theory, a dislocation is defined as the strong discontinuity of the displacement field. In this Chapter, we shall first study dislocation theory within the framework of linear elasticity, and then we shall examine dislocation theory by considering lattice structure, i.e. we shall study the Peierls-Nabarro model and a screw dislocation solution in the framework of molecular dynamics. At the end of this Chapter, we shall discuss one of the most important applications of dislocation theory: dislocations in thin films.

7.1 Screw dislocation

A multiply-connected region is defined as a region that it at least contains one irreducible circuit, i.e. a closed curve that can not be contracted to a single point without passing out of the region (see Fig. 6.6). Consider a multiply-connected region $\mathcal{V}$. A Volterra dislocation is defined as the displacement or rotation discontinuity over the line segment S (2D) or surface S (3D), i.e.

$$\begin{aligned} \Big[\mathbf{u}\Big] &= \mathbf{u}(\mathbf{P}^+) - \mathbf{u}(\mathbf{P}^-) = \mathbf{b} + \mathbf{d} \times \mathbf{x} \\ \Big[\boldsymbol{\omega}\Big] &= \boldsymbol{\omega}(\mathbf{P}^+) - \boldsymbol{\omega}(\mathbf{P}^-) = \mathbf{d} \end{aligned} \tag{7.1}$$

where $\mathbf{b}$ is the Burgers vector that can be defined as

$$\mathbf{b} = \oint_C \Big(\mathbf{E}(\mathbf{y}) + (\mathbf{x} - \mathbf{y}) \times \Big[\nabla \times \mathbf{E}(\mathbf{y})\Big]^T\Big) d\mathbf{y} \tag{7.2}$$

and

$$\mathbf{d} = -\oint_C \Big(\nabla \times \mathbf{E}(\mathbf{y})\Big)^T d\mathbf{y} \tag{7.3}$$

and $\mathbf{E}$ is the strain tensor.

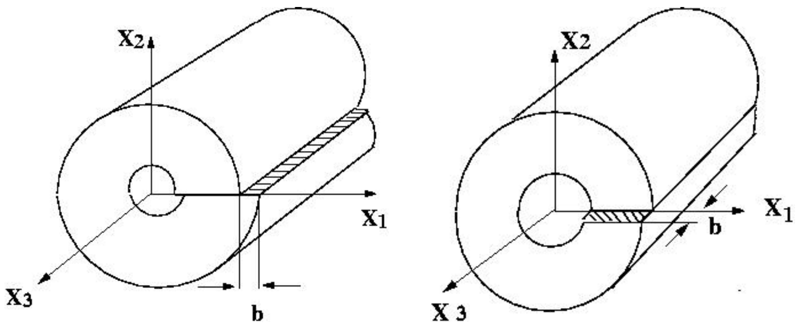


Figure 7.1. Illustraions of dislocations: (a)edge dislocation, and (b) screw dislocation

Historically, there is another type of dislocation: the Somigliana dislocations that are defined as

$$\left[\mathbf{u}\right] = \mathbf{u}^+ - \mathbf{u}^- = \mathbf{b}, \ \forall \mathbf{x} \in S \tag{7.4}$$

$$\left[\mathbf{t}\right] = \mathbf{t}^+ - \mathbf{t}^- = 0, \ \forall \mathbf{x} \in S \tag{7.5}$$

That is the traction is required to be continuous across the slip plane. However, the solution of such boundary-value problem is difficult, and people have not found any important applications of such dislocation model.

7.1.1 The solution of screw dislocation

We first derive the solution for the screw dislocation. The kinematics of the screw dislocation belong to that of anti-plane problem:

$$u_1 = 0, \ \ u_2 = 0, \ \ \text{and} \ \ u_3 = w(x, y) \ . \tag{7.6}$$

All the strain components are zero, except the out-plane shear strains

$$\epsilon_{xz} = \frac{1}{2}\frac{\partial w}{\partial x}, \quad \epsilon_{yz} = \frac{1}{2}\frac{\partial w}{\partial y} \ . \tag{7.7}$$

The corresponding non-zero shear stresses are

$$\sigma_{xz} = \mu \frac{\partial w}{\partial x} \tag{7.8}$$

$$\sigma_{xy} = \mu \frac{\partial w}{\partial y} \tag{7.9}$$

The non-trivial equilibrim equation

$$\frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{yz}}{\partial y} + \frac{\partial \sigma_{zz}}{\partial z} = 0 \tag{7.10}$$

leads to the governing equation

$$\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} = \nabla^2 w = 0 \ . \tag{7.11}$$

We denote the displacement jump in w at $y = 0$ and $x > 0$ as b_z i.e. $\mathbf{b} = b_z \mathbf{e}_z$, and the jump condition may be expressed as

$$\lim_{\eta \to 0, x > 0} \Big(w(x, -\eta) - w(x, \eta) \Big) = [w(x, 0)] = b_z, \ \ \eta > 0 \tag{7.12}$$

Use the polar coordinate,

$$\nabla^2 w = \Big(\frac{\partial^2}{\partial^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} \Big) w = 0 \ . \tag{7.13}$$

Separation of variables and let

$$w(r, \theta) = f(r) g(\theta) \tag{7.14}$$

we have

$$\frac{r^2}{f(r)} \Big(\frac{d^2 f}{dr^2} + \frac{1}{r}\frac{df}{dr} \Big) + \frac{1}{g(\theta)}\frac{d^2 g}{d\theta^2} = 0 \ . \tag{7.15}$$

We then end with two ordinary differential equations,

$$\begin{cases} \dfrac{d^2 f}{dr^2} + \dfrac{1}{r}\dfrac{df}{dr} - \dfrac{n^2 f}{r^2} = 0 \\ \\ \dfrac{d^2 g}{d\theta^2} + n^2 g(\theta) = 0 \end{cases} \tag{7.16}$$

If $n = 0$, one may find that

$$g(\theta) \ = \ A + B\theta \tag{7.17}$$

$$f(r) \ = \ C \ln r + D \tag{7.18}$$

For $n \neq 0$,

$$g(\theta) \ = \ C_n \cos n\theta + D_n \sin n\theta \tag{7.19}$$

$$f(r) \ = \ E_n r^n + F_n r^{-n} \tag{7.20}$$

This is true because

$$\Big(\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - \frac{n^2}{r^2} \Big) r^n = \Big(n(n-1) + n - n^2 \Big) r^{n-2} \equiv 0 \ . \tag{7.21}$$

Because the displacement, w, has to be finite, we can only consider the case $n = 0$. Again, because the convergence requirement for displacement field, $C = 0$; and because of jump condition, $A = 0$.

By absorbing the constant D into the constant B, the displacement field is

$$w(r,\theta) = B\theta \tag{7.22}$$

Use the jump condition,

$$w(r,2\pi) - w(r,0) = b \tag{7.23}$$

one may find that $2\pi B = b$ and hence

$$B = \frac{b}{2\pi} \tag{7.24}$$

Finally,

$$w(r,\theta) = \frac{\theta b}{2\pi} = \frac{b}{2\pi}\arctan\Big(\frac{y}{x}\Big) \tag{7.25}$$

and

$$\frac{\partial w}{\partial x} = -\frac{b}{2\pi}\frac{y}{x^2+y^2} = -\frac{b\sin\theta}{2\pi r} \tag{7.26}$$

$$\frac{\partial w}{\partial y} = \frac{b}{2\pi}\frac{x}{x^2+y^2} = \frac{b\cos\theta}{2\pi r} \tag{7.27}$$

Consequently, the non-zero stress components are

$$\sigma_{xz} = -\Big(\frac{b\mu}{2\pi}\Big)\frac{y}{x^2+y^2} \tag{7.28}$$

$$\sigma_{yz} = \Big(\frac{b\mu}{2\pi}\Big)\frac{x}{x^2+y^2} \tag{7.29}$$

In the cylindrical coordinate,

$$\begin{bmatrix} \sigma_{rr} & \sigma_{r\theta} & \sigma_{rz} \\ \sigma_{\theta r} & \sigma_{\theta\theta} & \sigma_{\theta z} \\ \sigma_{zr} & \sigma_{z\theta} & \sigma_{zz} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & \sigma_{xz} \\ 0 & 0 & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & 0 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The non-zero stress components are

$$\sigma_{rz} = \cos\theta\sigma_{xz} + \sin\theta\sigma_{yz} = 0 \tag{7.30}$$

$$\sigma_{\theta z} = -\sin\theta\sigma_{xz} + \cos\theta\sigma_{yz} = \frac{b\mu}{2\pi r}\,. \tag{7.31}$$

In the following, we calculate the self-energy of the screw dislocation in a hollow cylinder with inner radius r_0 and outer radius R. Note that the self-energy of a dislocation is defined as the strain energy contribution from stress-strain field of the dislocation solution in an unbounded region.

Assume that the length of the hollow cylinder is L. The energy per unit length in z-direction is,

$$
\begin{aligned}
\frac{W}{L} &= \frac{1}{L}\int_V \frac{\sigma_{z\theta}^2}{2\mu} dV = \frac{1}{V}\int_0^L \int_0^{2\pi} \int_{r_0}^{R} \frac{\sigma_{z\theta}^2}{2\mu} r dr d\theta z \\
&= \frac{b^2\mu}{4\pi}\int_{r_0}^{R} \frac{d}{r} = \frac{b^2\mu}{4\pi}\ln\frac{R}{r_0}\,. \qquad (7.32)
\end{aligned}
$$

First, as $R \to \infty$, $W/L \to \infty$. This shows that the self-energy of the dislocation depends on the size of the crystal. On the other hand, for a finite size crystal, the dislocation solution of unbounded domain does not hold true because the image stress caused by the boundary.

Assume that the dislocation is far away from the boundary, the boundary effecrts are abated inside, one may choose the dimension of the crystal, say ℓ as R; in polycrystallines, one may choose the size of a grain as R, where the dislocation resides.

Second, as $r_0 \to 0$, $W/L \to -\infty$. This abnormality is due to the limitation of linear elasticity model. Within five atomic spacing of a dislocation core, the linear elasticity model is no longer valid. In general, the length of the Bergurs vector is close to the lattice spacing. Therefore, in practice, we usually choose $r_0 = 5b$ or $r_0 = b/\alpha$, $0 < \alpha < 1$ such that the elastic self-energy equals to

$$
\frac{W}{L} = \frac{\mu b^2}{4\pi}\ln\frac{\ell}{5b}, \quad \text{or} \quad \frac{W}{L} = \frac{\mu b^2}{4\pi}\ln\frac{\alpha\ell}{b}\,. \qquad (7.33)
$$

By defnition, the self-energy should include the core energy, i.e.

$$
W^{self} = W^{elas} + W^{core} \qquad (7.34)
$$

The core energy is relatively small, but may not be negligible, because it is 10% to 20 % of the elastic self-energy. It may be relatively small, but can not be neglected. Overall, the linear elasticity theory gives a good estimate of self-energy. In Sec. 4 of this Chapter, we shall discuss the Peierls-Nabarro model, which provides a means to estimate the core energy.

7.1.2 Image stress of a screw dislocation in a half space

Consider a crystal occupying a half space $x \leq 0$. Consider a screw dislocation located at the position $x = -\ell$ (see Fig. 7.2). The screw dislocation in an unbounded space gives the following stress distrubution,

$$
\sigma_{xz}^{\infty}(x,y) = -\frac{b\mu}{2\pi}\frac{y}{(x+\ell)^2+y^2} \qquad (7.35)
$$

$$
\sigma_{yz}^{\infty}(x,y) = \frac{b\mu}{2\pi}\frac{(x+\ell)}{(x+\ell)^2+y^2}\,. \qquad (7.36)
$$

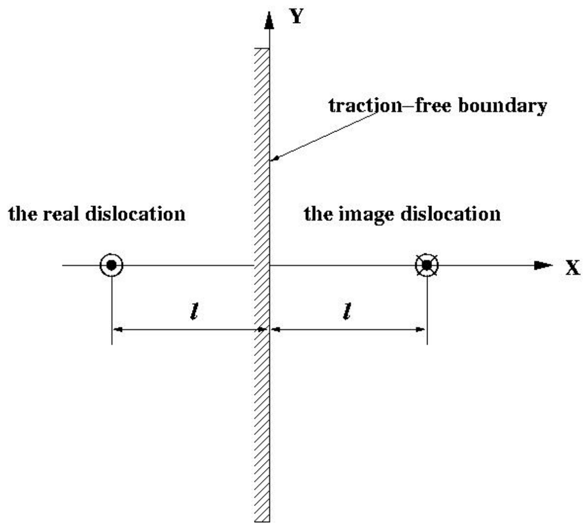


Figure 7.2. An image screw dislocation

This solution does not satisfy the traction-free boundary condition at $x = 0$, because

$$\sigma^{\infty}_{xz}(0, y) = -\frac{b\mu}{2\pi}\frac{y}{\ell^2 + y^2} \neq 0\,. \tag{7.37}$$

To enforce the traction-free boundary condition, we place a fictitious screw dislocation with the Bergurs vector, $b' = -b$, at the position $x = \ell$, and it generates the following so-called image stress distribution:

$$\sigma^{I}_{xz}(x, y) = \frac{b\mu}{2\pi}\frac{y}{(x-\ell)^2 + y^2} \tag{7.38}$$

$$\sigma^{I}_{yz}(x, y) = -\frac{b\mu}{2\pi}\frac{(x-\ell)}{(x-\ell)^2 + y^2}\,. \tag{7.39}$$

The total stress distribution is then the superposition of the solution in the infinite space and the solution of of image stress distribution, i.e. $\sigma^t_{ij} = \sigma^{\infty}_{ij} + \sigma^I_{ij}$, where the superscript, t, ∞, and I denote the total stress solution, the solution obtained in the infinite space, and the image stress solution.

By anti-symmetry, the traction-free boundary condition at $x = 0$ is then enfored,

$$\sigma^t_{xz}(0, y) = \sigma^{\infty}_{xz}(0, y) + \sigma^I_{xz}(0, y) = -\frac{by}{2\pi}\frac{y}{\ell^2 + y^2} + \frac{by}{2\pi}\frac{y}{\ell^2 + y^2} \equiv 0\,. \tag{7.40}$$

REMARK 7.1.1 **1.** *Note that the image stresses at $x = -\ell$ and $y = 0$, i.e. the position of the real dislocation, are*

$$\sigma^I_{xz}(-\ell, 0) = 0, \quad \sigma^I_{yz}(-\ell, 0) = \frac{b\mu}{4\pi\ell} \,. \tag{7.41}$$

2. *When $|\mathbf{x}|, |\mathbf{y}| >> \ell$,*

$$\sigma^t_{xz}(x, y) \approx 0, \ \text{ and } \ \sigma^t_{yz}(x, y) \approx 0, \tag{7.42}$$

which means that outside the region of $\{(x, y) \,\Big|\, (x + \ell)^2 + y^2 \leq \ell^2\}$, the total stress is almost negligible.

7.1.3 Eshelby's twist: screw dislocation in a finite whisker

Consider a screw dislocation in a finite cylinder (whisker). One may find that the solution of a single screw dislocation in an infinite space actually satisfies the lateral boundary conditions of the problem:

$$\sigma_{z\theta} = \frac{\mu b}{4\pi r}, \ \forall r \leq R \tag{7.43}$$

$$\sigma_{rr} = \sigma_{r\theta} = \sigma_{rz} = 0, \quad 0 \leq r \leq R \tag{7.44}$$

However, there is one problem there are resulting moments or torques at the two open ends of the cylinder, i.e.

$$\begin{aligned} M_z &= \int_0^R \int_0^{2\pi} r\sigma_{\theta z} r dr d\theta \\ &= 2\pi \frac{\mu b}{2\pi} \int_0^R r dr = \frac{\mu b R^2}{2} \,. \end{aligned} \tag{7.45}$$

To negate the end moment, we superpose two ends moments with the opposite direction of $M_z^{'} = -M_z$ such that the total moments at the two ends of the cylinder become zero, and then based on Saint-Venatet's principle we can declare the validity of the solution.

The superposed two-end moments will result the following stress distribution that can be calculated by the elementary torsion formula,

$$\sigma^{'}_{\theta z} = \frac{M_z^{'} r}{J} = -\frac{\mu b r}{\pi R^2} \tag{7.46}$$

In the last equation, we used the fact that the polar moment of a circular region is $J = \pi R^4/2$.

Then the stress distribution in a whisker is

$$\sigma_{\theta z} = \frac{\mu b}{2\pi r} - \frac{\mu b r}{\pi R} \,. \tag{7.47}$$

where the extra term $-(\mu b r)/R$ may be viewed as an equivalent image stress steming from the superposed boundary moment.

7.2 Edge dislocation

The edge dislocation problem can be solved as a plane strain problem. Introduce the Airy stress function, such that

$$\sigma_{xx} = \frac{\partial^2 \psi}{\partial y^2},\ \sigma_{yy} = \frac{\partial^2 \psi}{\partial x^2},\ \text{and}\ \sigma_{xy} = -\frac{\partial^2 \psi}{\partial x \partial y}\ . \tag{7.48}$$

The in-plane equilibrium equation,

$$\frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{yx}}{\partial y} = 0, \tag{7.49}$$

leads to the following bi-harmonic equation,

$$\nabla^2 \nabla^2 \psi = 0\ . \tag{7.50}$$

Let $\phi = \sigma_{xx} + \sigma_{yy} = \nabla^2 \psi$. Then

$$\begin{aligned} &\nabla^2 \nabla^2 \psi = \nabla^2 \phi = 0\ ,\ \ \text{and in the polar coordinate :} \\ &\Big(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}\Big)\phi = 0\ . \end{aligned} \tag{7.51}$$

Based on the general solution obtained in the previous subsection, ϕ has the following form,

$$\begin{aligned} \phi(r,\theta) \quad = \quad & (\alpha_0 + \beta_0 \ln r) + \sum_{n=1}^{\infty}\Big(\alpha_n r^n + \beta_n r^{-n}\Big)\sin n\theta \\ & + \sum_{n=1}^{\infty}\Big(\gamma_n r^n + \delta_n r^{-n}\Big)\cos n\theta \end{aligned} \tag{7.52}$$

Because the defect configuration, for an edge dislocation, the region right above around the dislocation core should be in compression, whereas the region right below the dislocation core should be in tension, i.e.

$$\phi(r_0, \pi/2) = \phi_{min},\ \ \text{and}\ \ \phi(r_0, -\pi/2) = \phi_{max}\ \ . \tag{7.53}$$

In consideration with the convergence at remote region, i.e. ($\phi \to 0, r \to \infty$), the right choice of the solution should be $n = 1$ and

$$\phi = \beta_1 r^{-1} \sin\theta\ . \tag{7.54}$$

Then,

$$\Big(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2}\Big)\psi = \beta_1 r^{-1} \sin\theta\ . \tag{7.55}$$

Let $\psi = h(r)\sin\theta$. One may find that

$$\Big(\frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr} - \frac{1}{r^2}\Big)h = \frac{d}{dr}\Big(\frac{1}{r}\frac{d}{dr}(rh)\Big) = \beta_1 r^{-1}\ . \tag{7.56}$$

By straightforward integration, one can verify that a particular solution is

$$\psi_e = \frac{\beta_1}{2} r\sin\theta\ln r \ = \frac{\beta_1 y}{4}\ln(x^2+y^2)\ . \tag{7.57}$$

Consider the jump condition,

$$\lim_{\eta\to 0} - \int_{-\infty}^{\infty}\Big[\epsilon_{xx}(x,\eta) - \epsilon_{xx}(x,-\eta)\Big]dx = b\ . \tag{7.58}$$

One can determine the constant β_1,

$$\beta_1 = -\frac{\mu b}{\pi(1-\nu)} \ \ \Rightarrow\ \ \psi_e = -\frac{\nu b y}{4\pi(1-\nu)}\ln(x^2+y^2)\ . \tag{7.59}$$

One can then find stress components

$$\sigma_{xx} = -\frac{\mu b}{2\pi(1-\nu)}\frac{y(3x^2+y^2)}{(x^2+y^2)^2} \tag{7.60}$$

$$\sigma_{yy} = \frac{\mu b}{2\pi(1-\nu)}\frac{y(x^2-y^2)}{(x^2+y^2)^2} \tag{7.61}$$

$$\sigma_{xy} = \frac{\mu b}{2\pi(1-\nu)}\frac{x(x^2-y^2)}{(x^2+y^2)^2},\ \ \text{and} \tag{7.62}$$

$$\sigma_{zz} = \nu(\sigma_{xx}+\sigma_{yy}) \tag{7.63}$$

or in the polar coordinate

$$\sigma_{rr} = \sigma_{\theta\theta} = -\frac{\mu b\sin\theta}{2\pi(1-\nu)r} \tag{7.64}$$

$$\sigma_{r\theta} = \frac{\mu b\cos\theta}{2\pi(1-\nu)r}\ \ \sigma_{zz} = \nu(\sigma_{rr}+\sigma_{\theta\theta}) = -\frac{\mu b\nu\sin\theta}{\pi(1-\nu)r}\ . \tag{7.65}$$

It is then easy to find the strain fields by simply applying Hooke's law of plane strain condition,

$$\epsilon_{xx} = \frac{by}{2\pi}\frac{(\mu y^2 + (2\lambda+3\mu)x^2}{(\lambda+2\mu)(x^2+y^2)^2} \tag{7.66}$$

$$\epsilon_{yy} = -\frac{by}{2\pi}\frac{((2\lambda+\mu)x^2 - \mu y^2)}{(\lambda+2\mu)(x^2+y^2)^2} \tag{7.67}$$

$$\epsilon_{xy} = -\frac{b}{2\pi(1-\nu)}\frac{x(x^2-y^2)}{(x^2+y^2)^2} \tag{7.68}$$

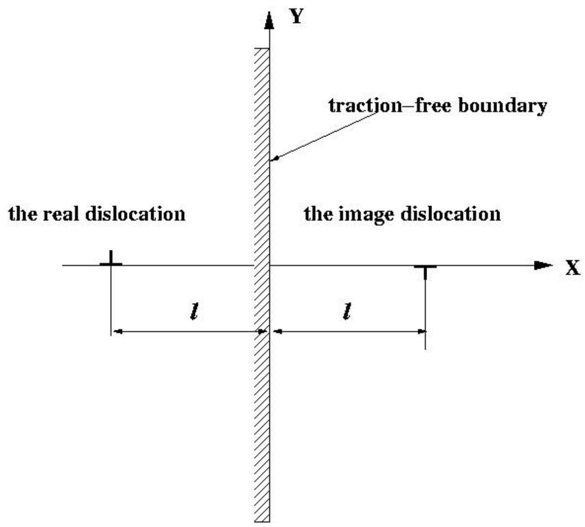


Figure 7.3. An image edge dislocation

By neglecting all the integration constants, a straightforward integration of the above strain components gives

$$u(x,y) = -\frac{b}{2\pi}\left[\tan^{-1}\frac{y}{x} + \frac{\lambda+\mu}{\lambda+2\mu}\frac{xy}{x^2+y^2}\right] \tag{7.69}$$

$$v(x,y) = -\frac{b}{2\pi}\left[-\frac{\mu}{2(\lambda+2\mu)}\ln(x^2+y^2) + \frac{\lambda+\mu}{\lambda+2\mu}\frac{y^2}{x^2+y^2}\right] \tag{7.70}$$

7.2.1 Image stress for an edge dislocation

The solution of the image stress distribution for an edge dislocation is more complicated than that of a screw dislocation.

Consider an edge dislocation being placed at $x = -\ell$ inside a half space $(x < 0)$. The solution obtained from the unbounded space,

$$\sigma_{xx}^{\infty} = -\frac{\mu b}{2\pi(1-\nu)}\frac{y(3(x+\ell)^2+y^2}{((x+\ell)^2+y^2)^2} \tag{7.71}$$

$$\sigma_{yy}^{\infty} = \frac{\mu b}{2\pi(1-\nu)}\frac{y((x+\ell)^2-y^2}{((x+\ell)^2+y^2)^2} \tag{7.72}$$

$$\sigma_{xy}^{\infty} = \frac{\mu b}{2\pi(1-\nu)}\frac{(x+\ell)((x+\ell)^2-y^2}{((x+\ell)^2+y^2)^2} \tag{7.73}$$

will not satisfy the traction-free boundary condition at $x = 0$ i.e. $\sigma_{xx}(0, y) \neq 0$ and $\sigma_{xy}(0, y) \neq 0$.

If we place a fictitous dislocation at $x = \ell$ with the opposite Burgers vector. The induced image stress fields,

$$\sigma^I_{xx} = \frac{\mu b}{2\pi(1-\nu)} \frac{y(3(x-\ell)^2 + y^2}{((x-\ell)^2 + y^2)^2} \tag{7.74}$$

$$\sigma^I_{yy} = \frac{\mu b}{2\pi(1-\nu)} \frac{y((x-\ell)^2 - y^2}{((x-\ell)^2 + y^2)^2} \tag{7.75}$$

$$\sigma^I_{xy} = \frac{\mu b}{2\pi(1-\nu)} \frac{(x-\ell)((x-\ell)^2 - y^2}{((x-\ell)^2 + y^2)^2} \tag{7.76}$$

will cancel the normal stress on traction-free surface, i.e. $\sigma^\infty_{xx}(0, y) + \sigma^I_{xx}(0, y) = 0$, but it can not cancel the shear stress at $x = 0$. In fact,

$$\sigma^\infty_{xy}(0, y) + \sigma^I_{xy}(0, y) = \frac{\mu b}{\pi(1-\nu)} \frac{\ell(\ell^2 - y^2)}{(\ell^2 + y^2)^2} \neq 0\ . \tag{7.77}$$

To cancel the shear stress on traction-free surface, one has to superpose another stress field, such that the third stress fields satisfy the condition,

$$\sigma^{'''}_{xx}(0, y) = 0, \text{ and } \sigma^{'''}_{xy}(0, y) = -\frac{\mu b}{\pi(1-\nu)} \frac{\ell(\ell^2 - y^2)}{(\ell^2 + y^2)^2}\ . \tag{7.78}$$

Consider the Airy stress function, $\Psi(x, y)$, which satisfies the bi-harmonic equation,

$$\nabla^2 \nabla^2 \Psi = 0\ . \tag{7.79}$$

Introduce the Fourier-sine and the Fourier-cosin transforms,

$$\bar{f}_s(\xi) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(y) \sin(\xi y) dy, \quad f(y) = \int_0^\infty \bar{f}_s(\xi) \sin(\xi y) d\xi; \tag{7.80}$$

$$\bar{f}_s(\xi) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(y) \cos(\xi y) dy, \quad f(y) = \int_0^\infty \bar{f}_c(\xi) \cos(\xi y) d\xi \tag{7.81}$$

Since σ_{xy} must be even in y, the Airy stress function, Ψ, is anti-symmetric in y. We apply the Fourier-sine transform to Eq. (7.79), and it yields a ordinary differential equation,

$$\frac{d^4 \bar{\Psi}_s}{dx^4} - 2\xi^2 \frac{d^2 \bar{\Psi}_s}{dx^2} + \xi^4 \bar{\Psi}_s = 0. \tag{7.82}$$

Solving (7.82) yields the following solution,

$$\bar{\Psi}_s(x, \xi) = (a_0(\xi) + a_1(\xi)) \exp(\xi x) + (b_0(\xi) + b_1(\xi)) \exp(-\xi x)\ . \tag{7.83}$$

The boundary conditions,

$$\mathbf{1}.\;\; x \to -\infty,\;\; \bar{\Psi}_s \to 0,\;\; \Rightarrow b_0 = b_1 = 0; \tag{7.84}$$

$$\mathbf{2}.\;\; x = 0,\;\; \sigma_{xx}(0,y) = 0,\quad \Rightarrow a_0 = 0\;. \tag{7.85}$$

Therefore, $\bar{\Psi}_s(x,\xi) = a_1(\xi)x\exp(\xi x)$, and

$$\Psi(x,y) = \frac{1}{\pi}\int_{\infty}^{\infty} a_1(\xi)x\exp(\xi x)\sin(\xi y)d\xi\;. \tag{7.86}$$

Using the boundary condition for the shear stress,

$$\begin{aligned} -\sigma'''_{xy}(0,y) &= \Big(\frac{\partial^2\Psi}{\partial x\partial y}\Big)\Big|_{x=0} = \int_0^{\infty} a_1(\xi)\xi\cos(\xi y)dy \\ &= \frac{\mu b}{\pi(1-\nu)}\frac{\ell(\ell^2-y^2)}{(\ell^2+y^2)^2} \end{aligned} \tag{7.87}$$

and the definition of the Fourier-cosin transform, one may find that

$$\begin{aligned} a_1(\xi)\xi &= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\mu b}{\pi(1-\nu)}\frac{\ell(\ell^2-y^2)}{(\ell^2+y^2)^2}\cos(\xi y)dy \\ &= \frac{\mu b}{\pi^2(1-\nu)}\int_{-\infty}^{\infty}\frac{\ell(\ell^2-y^2)}{(\ell^2+y^2)^2}\exp(i\xi y)dy\;. \end{aligned} \tag{7.88}$$

The last line is because of $\displaystyle\int_{-\infty}^{\infty}\frac{\ell(\ell^2-y^2)}{(\ell^2+y^2)^2}\sin(\xi y)dy = 0$.

Use the residue theorem to evaluate the integra,

$$\begin{aligned} &\int_{-\infty}^{\infty}\frac{\ell(\ell^2-y^2)}{(\ell^2+y^2)^2}\exp(i\xi y)dy = 2\pi i\sum Res\;F(y_N)\Big|_{y_N=i\ell} \\ &= 2\pi i\Big(-\frac{i\xi\ell}{2}\exp(-\xi\ell)\Big) = \pi\xi\ell exp(-\xi\ell)\;. \end{aligned} \tag{7.89}$$

Wer then find that

$$a_1(\xi) = \frac{\mu b\ell}{\pi(1-\nu)}\exp(-\xi\ell) \tag{7.90}$$

so that

$$\begin{aligned} \Psi(x,y) &= \frac{\mu b\ell}{\pi(1-\nu)}\int_0^{\infty} x\exp\xi(x-\ell)\sin\xi y d\xi \\ &= \frac{\mu b\ell xy}{\pi(1-\nu)[(x-\ell)^2+y^2]} \end{aligned} \tag{7.91}$$

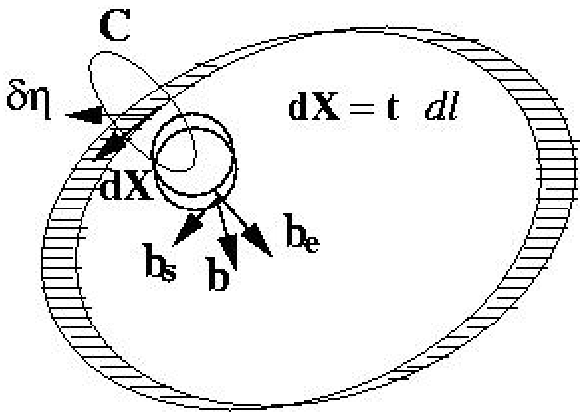


Figure 7.4. A virtual displacement of a dislocation loop

and

$$\sigma^{'''}_{xy} = -\frac{\mu b\ell}{\pi(1-\nu)}\Big(\frac{(\ell^2-x^2)y^2}{[(x-\ell)^2+y^2]^2}+\frac{y^2(3x^2-(y+\ell)^2)}{[(x-\ell)^2+y^2]^3}\Big) \quad (7.92)$$

$$\sigma^{'''}_{xx} = -\frac{2\mu b\ell xy}{\pi(1-\nu)r^6}[3(\ell-x)^2-y^2] \quad (7.93)$$

Indeed, it can be found that

$$\sigma^{'''}_{xy}(0,y)=\frac{\mu b\ell}{\pi(1-\nu)}\frac{\ell^2-y^2}{(\ell^2+y^2)^2}, \text{ and } \sigma^{'''}_{xx}(0,y)=0\,. \quad (7.94)$$

Moreover, since $\sigma^{'''}_{xy}(\ell,0)=0$, the shear stress acting on the real dislocation due the traction-free boundary is the stress applied by the image dislocation (the second dislocation), i.e.

$$\sigma^{I}_{xy}(-\ell,0)+\sigma^{'''}_{xy}(-\ell,0)=\frac{\mu b}{4\pi(1-\nu)}\,. \quad (7.95)$$

7.3 The Peach-Koehle force

Consider a dislocation loop undergoing a virtual displacement $\delta\boldsymbol{\eta}$ (see Fig. 7.4). An infinitesimal dislocation line segment, $d\mathbf{X}$ will sweep through an area,

$$d\mathbf{A}=d\mathbf{X}\times\delta\boldsymbol{\eta}\,. \quad (7.96)$$

Note that the direction of $d\mathbf{A}$ is its out-normal.

All the atoms on this area will be sujected a discontinuous jump with the direction and the magnitude of the local Burgers vector, $\mathbf{b}$. The traction forces on the infinitesimal area can be expressed as $\boldsymbol{\sigma} \cdot d\mathbf{A}$. Be precise, it is

$$\boldsymbol{\sigma} \cdot d\mathbf{A} = \boldsymbol{\sigma} \cdot (d\mathbf{X} \times \delta\boldsymbol{\eta}) \tag{7.97}$$

If we assume that the work done by the stresses relates to the decreases of the potential energy of the dislocation,

$$d(\delta E) = -\mathbf{b} \cdot \boldsymbol{\sigma} \cdot (d\mathbf{X} \times \delta\boldsymbol{\eta}) \tag{7.98}$$

The change of the total energy due to the virtual displacement field is

$$\delta E = -\int_{\mathcal{L}} \mathbf{b} \cdot \boldsymbol{\sigma} \cdot (d\mathbf{X} \times \delta\boldsymbol{\eta}) = -\int_{\mathcal{L}} (\boldsymbol{\sigma} \cdot \mathbf{b}) \times \mathbf{t} d\ell \cdot \delta\boldsymbol{\eta} \tag{7.99}$$

where $d\mathbf{X} = \mathbf{t} d\ell$.

By definition, the decrease of the potential energy under the virtual displacement field is the external virtual work done along the dislocation loop, i.e.

$$\delta E = -\mathbf{F} \cdot \boldsymbol{\eta} = -\int_{\mathcal{L}} \mathbf{F}_\ell d\ell \cdot \delta\boldsymbol{\eta} \ , \tag{7.100}$$

where $\mathbf{F}_\ell$ is the force per unit length along the dislocation loop.

Hence, we derived the celebrated Peach-Koehle equation,

$$\mathbf{F} = \int_{\mathcal{L}} \Big(\boldsymbol{\sigma} \cdot \mathbf{b}\Big) \times \mathbf{t} d\ell, \ \text{ and } \ \mathbf{F}_\ell = \Big(\boldsymbol{\sigma} \cdot \mathbf{b}\Big) \times \mathbf{t} \ . \tag{7.101}$$

where $\mathbf{F}_\ell$ is the force per unit length. In the case of straight dislocation line, we often denote it as $\dfrac{\mathbf{F}}{L}$.

Now, let's look at a few examples.

To simplify the computation, we denote

$$\mathbf{g} := \boldsymbol{\sigma} \cdot \mathbf{b}. \tag{7.102}$$

Then the Peach-Koehle force formula can be conveniently written into a matrix form,

$$\mathbf{F}_\ell = \mathbf{g} \times \mathbf{t} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ g_1 & g_2 & g_3 \\ t_1 & t_2 & t_3 \end{vmatrix} \ . \tag{7.103}$$

EXAMPLE 7.1 *This example is illustrated in Fig. 7.5. We are examing the external forces exerted on a straight screw dislocation.*

Let $x = 1, y = 2, z = 3$. In this case, the unit vector of the dislocation line is $\mathbf{t} = \mathbf{e}_z$, the Burgers vector is $\mathbf{b} = b\mathbf{e}_z$, and the stresses other than self-stress are

$$\boldsymbol{\sigma} = \sigma_{xz}\mathbf{e}_x \otimes \mathbf{e}_z + \sigma_{zx}\mathbf{e}_z \otimes \mathbf{e}_x + \sigma_{yz}\mathbf{e}_y \otimes \mathbf{e}_z + \sigma_{zy}\mathbf{e}_z \otimes \mathbf{e}_y \ . \tag{7.104}$$

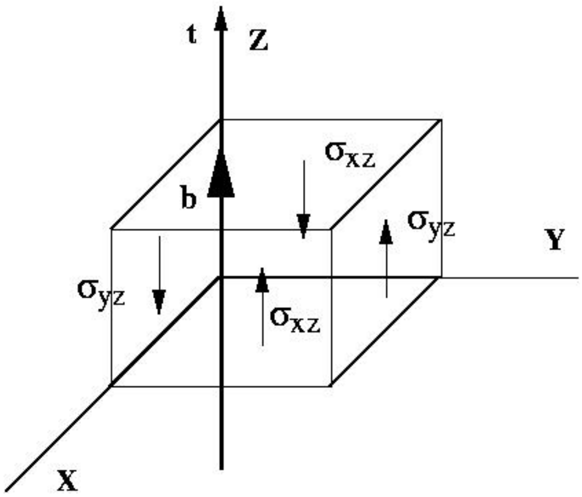


Figure 7.5. A straight screw dislocation.

and

$$g_x = \sigma_{xz} b, \;\; g_y = \sigma_{yz} b, \;\; g_z = 0 \; . \tag{7.105}$$

Hence

$$\mathbf{F}_\ell = \mathbf{g} \times \mathbf{u} = \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ \sigma_{xz} b & \sigma_{yz} b & 0 \\ 0 & 0 & 1 \end{vmatrix} = \sigma_{yz} b \mathbf{e}_x - \sigma_{xz} b \mathbf{e}_y \; . \tag{7.106}$$

To interprete the meanings of this expression, we would say that the shear stress, σ_{xy}, moves the dislocation line to $+x$ direction, whereas shear stress, σ_{xz} moves the dislocation line towards the negative direction of Y-axis, i.e. -Y direction.

EXAMPLE 7.2 *In the second example, we consider a straight edge dislocation. This example is illustrated in Fig. 7.6. In this example, again $\mathbf{u} = \mathbf{e}_z$, but $\mathbf{b} = b\mathbf{e}_x$, and*

$$\boldsymbol{\sigma} = \sigma_{xx} \mathbf{e}_x \otimes \mathbf{e}_x + \sigma_{xy} \mathbf{e}_x \otimes \mathbf{e}_y + \sigma_{yz} \mathbf{e}_y \otimes \mathbf{e}_x \; . \tag{7.107}$$

Thus,

$$g_x = \sigma_{xx} b, \;\; g_y = \sigma_{yx} b, \;\; \text{and} \;\; g_z = 0 \; , \tag{7.108}$$

and

$$\mathbf{F}_\ell = \mathbf{g} \times \mathbf{u} = \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ \sigma_{xx} b & \sigma_{xy} b & 0 \\ 0 & 0 & 1 \end{vmatrix} = \sigma_{xy} b \mathbf{e}_x - \sigma_{xx} b \mathbf{e}_y \; . \tag{7.109}$$

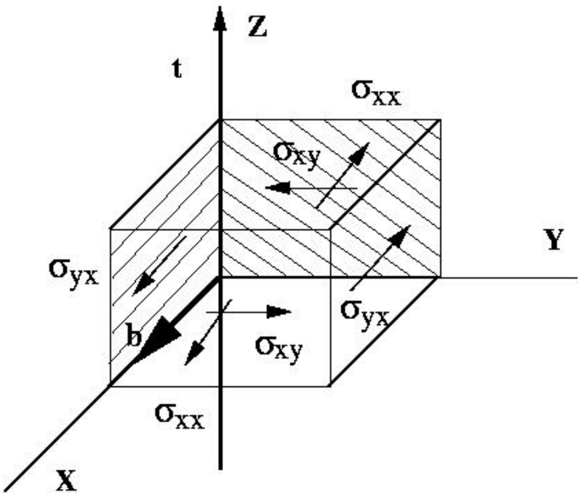


Figure 7.6. A straight edge dislocation.

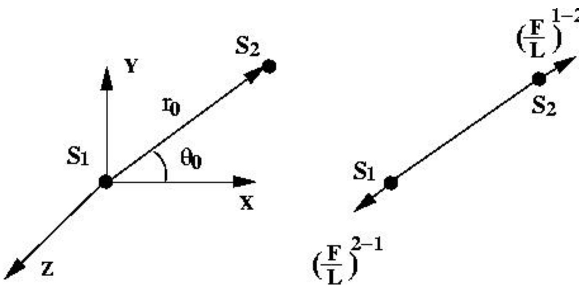


Figure 7.7. Interactions of two parallel screw dislocations

This is to say that the shear stress, σ_{xy}, will move the dislocation line along the slip plane in the positive direction of X-axis. On the other hand, the normal stress, σ_{xx}, will make the dislocation line tranlating along its own direction. This is an unconservative motion, because if the motion is addmissible, one has to remove material at one end of dislocation line and add material (atoms) at the other end of the dislocation line. In literature, we refer such dislocation movement as "climbing".

From Eq. (7.109), one may find that if $\sigma_{xx} < 0$, which means the material is under compression, the Peach-Koehle force will squeeze the dislocation line up in Y-axis, and when $\sigma_{xx} > 0$ it will pull the material apart and let dislocation line climbing down.

EXAMPLE 7.3 *In this example, we consider the interactions between two parallel screw dislocations along the Z-axis,* $\mathbf{t} = \mathbf{e}_z$*,* S_1 *and* S_2*. They have different Burgers vectors, i.e.* $\mathbf{b}_1 = b_1\mathbf{e}_z$ *and* $\mathbf{b}_2 = b_2\mathbf{e}_z$*. For the dislocation,* S_1*, the stress field is*

$$\sigma_{xz}^I = -\frac{\mu b_1}{2\pi}\frac{\sin\theta}{r}, \quad \sigma_{yz}^I = \frac{\mu b_1}{2\pi}\frac{\cos\theta}{r} \quad ; \tag{7.110}$$

and for the dislocation, S_2*, the stress field is*

$$\sigma_{xz}^{II} = -\frac{\mu b_2}{2\pi}\frac{(y-y_0)}{(x-x_0)^2+(y-y_0)^2} \quad , \tag{7.111}$$

$$\sigma_{yz}^{II} = \frac{\mu b_2}{2\pi}\frac{(x-x_0)}{(x-x_0)^2+(y-y_0)^2} \quad . \tag{7.112}$$

In this case, the Peach-Koehle force equation is

$$\mathbf{F}_\ell = \sigma_{yz}\mathbf{e}_x - \sigma_{xz}\mathbf{e}_y \ . \tag{7.113}$$

1. *Calculate the force,* $\mathbf{F}_\ell^{1\to 2}$*, which is the force exeretd on the dislocation,* S_2*, by the dislocation,* S_1*. Let* $r = r_0$ *and* $\theta = \theta_0$ *in (7.110) and substitute them into (7.113). We have*

$$\begin{aligned}
\mathbf{F}_\ell^{1\to 2} &= \sigma_{yz}^I\Big|_{x_0,y_0} b_2\mathbf{e}_x - \sigma_{xz}^I\Big|_{x_0,y_0} b_2\mathbf{e}_y \\
&= \frac{\mu b_1 b_2}{2\pi}\frac{\cos\theta_0}{r_0}\mathbf{e}_x + \frac{\mu b_1 b_2}{2\pi}\frac{\sin\theta_0}{r_0}\mathbf{e}_y \\
&= \frac{\mu b_1 b_2}{2\pi r_0}\Big(\cos\theta_0\mathbf{e}_x + \sin\theta_0\mathbf{e}_y\Big) = \frac{\mu b_1 b_2}{2\pi r_0}\bar{\mathbf{r}}_0 \ ,
\end{aligned} \tag{7.114}$$

where $\bar{\mathbf{r}}_0 = \mathbf{r}_0/|\mathbf{r}_0|$ *is the unit vector in* $\mathbf{r}_0$ *direction.*

2. *Calculate the force exerted on the dislocation* $S-1$ *by the dislocation* S_2*. In this case, we let* $x = 0, y = 0$ *in (7.111) and (7.112) and substitute them into (7.113),*

$$\begin{aligned}
\mathbf{F}_\ell^{2\to 1} &= \sigma_{yz}^{II}\Big|_{0,0} b_1\mathbf{e}_x - \sigma_{xz}^{II}\Big|_{0,0} b_1\mathbf{e}_y \\
&= -\frac{\mu b_1 b_2}{2\pi}\frac{\cos\theta_0}{r_0}\mathbf{e}_x - \frac{\mu b_1 b_2}{2\pi}\frac{\sin\theta_0}{r_0}\mathbf{e}_y \\
&= -\frac{\mu b_1 b_2}{2\pi r_0}\Big(\cos\theta_0\mathbf{e}_x + \sin\theta_0\mathbf{e}_y\Big) = -\frac{\mu b_1 b_2}{2\pi r_0}\bar{\mathbf{r}}_0 \ .
\end{aligned} \tag{7.115}$$

It is obvious that $\mathbf{F}_\ell^{1\to 2} = -\mathbf{F}_\ell^{2\to 1}$ *(see Fig. 7.7).*

We then conclude that when $\mathbf{b}_1$ *and* $\mathbf{b}_2$ *are along the same direction, the two screw dislocation repel each other, if* $b_1 b_2 < 0$*, i.e.* $\mathbf{b}_1$ *and* $\mathbf{b}_2$ *are in opposite direction, then the two screw dislocations attract to each other.*

REMARK 7.3.1 *[Biot-Savart analogy]*
In electro-magnetics, if there are two parallel wires having electric current passing through, the interaction force between the two wires can be calcualted by the well-known Boit-Savart law,

$$\mathbf{F}^i_\ell = \frac{I_i}{c}\Big(\mathbf{t}\times\mathbf{B}_j\Big),\quad i\neq j \text{ and } i,j=1,2 \tag{7.116}$$

where $\mathbf{F}^i_\ell$ is the force exerted on the wire i by the magentic field generated by the wire j; I_i is the electric current density in the wire i, while $\mathbf{B}_j$ is the magnetic induction flux density generated by the wire j, and c is the light speed in the medium.

In the Peach-Koehle equation, if we define $\mathbf{G}_j = \boldsymbol{\sigma}_j\cdot\mathbf{t}$, then

$$\mathbf{g} = \boldsymbol{\sigma}_j\cdot\mathbf{b}_i = \boldsymbol{\sigma}_j\cdot\mathbf{t}b_i = \mathbf{G}_j b_i\ . \tag{7.117}$$

We can rewrite the Peach-Koehle force as

$$\mathbf{F}^i_\ell = -b_i\mathbf{t}\times\mathbf{G}_j\ . \tag{7.118}$$

It has a similar form with the Biot-Savart law. Since b_i is the analogy of I_i/c, we may call the strength of a Burgers vector as the dislocation current density. By the same token, we may call the stress projection due to the dislocation line $E_j,\ j=1,2$ as the stress induction flux.

The only difference between (7.116) and (7.117) are is the minus sign in (7.117). This is because in electro-maganetics. Two wires with the same (opposite) electric current direction attract (repel) to each other, whereas two screw dislocation lines having the same (opposite) dislocation current direction repel (attract) to each other.

7.4 Configuration force: Eshelby's energy-momentum tensor

Assume that if the solid that contains the edge dislocation ($\mathbf{b} = b\mathbf{e}_x$) is under external hydrostatic pressure, $\sigma_{11} = \sigma_{22} = \sigma_{33} = -p$, this will cause the edge dislocation climbing. While an edge dislocation climbs, it does not produce volumetric strain, thus, σ_{11} never does work any work in the process. Therefore, there is actually no real force acting on the dislocation.

Therefore, there is no actual force acting on the dislocation. Then the "virtual force" [1] defined as the decrease of the potential energy change due to the change of the dislocation position,

$$\mathbf{F}_\eta = -\frac{\partial W}{\partial\boldsymbol{\eta}}, \tag{7.119}$$

[1]Do not confusion this with the statically admissible virtual forces in continuum mechanics.

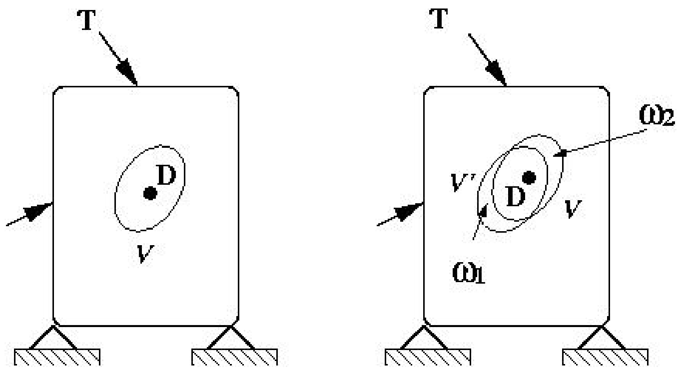


Figure 7.8. Eshelby's argument on configuration force

is really a force due to the change of material's configuration.

Configuration mechanics has been an active research subject since Eshelby's pineeor contribution on configuration force study. In this section, we outline the basic theory of configuration mechanics, and introduce Eshelby's energy-momentum tensor.

In order to evaluate the configuration force acting on a defect, we first calculate the change of potential energy due to the change of configuration.

To do this, we follow the Eshelby's famous thought experiment. The setting of Eshelby's thought experiment is a solid that is subjected external forces or displacement constraints at boundary. Inside the solid, there is a point defect denoted as D, and we link the defect D with its local configuration by embedding it into an arbitrarily chosen local volume V. We define the local configuration as the relative position of D inside V. We denote the boundary of the local volume as $\mathcal{L} = \partial V$ (see Fig. 7.8(a)).

The basic idea of Eshelby's thought experiment is to change the global configuration or the defect position, while comparing the energy change in a local configuration.

The following is the adaptation of Eshelby's imaginary operation, which mainly consists of four steps (I reshuffled the order):

(1) We first change the global configuration, or the position of the defect by amount of $\delta\mathbf{X}$ in the material configuration. We denote the original local volume containing D as $V^{'}$. When the defect, D, moves its new materials position, we still choose the same local struction, or local configuration (but a different sets of material points), to identify it, i.e. we surround the defect D with local volume V, which has the same local configuration as $V^{'}$. It means the relative position of D is the same with respect to V as it was before with respect to $V^{'}$. The comparison is made under the same local structure,

Eshelby called the local configuration of V is a replica of the original local configuration $V^{'}$.

Under this condition, the material virtual displacement field represents a change of configuration. One may observe this in Fig. 7.8(b).

(2) Before calculating the difference of the energy stored inside $V^{'}$ and V, we would like to cleairfy the following point: since the defect changes its position $+\delta\mathbf{X}$, this may change the self-stress field as well as image stress field of the defect, and consequently the energy density at each point. However, the change of energy density due to the defect movement is at order $\delta X_i \delta X_i \sim (\delta X)^2$, and it is a second order effect that can be neglected if $\delta\mathbf{X}$ is infinitesimal. Therefore, we can calculate strain energy stored inside V and $V^{'}$ without taking into account the effects of the defect's movement.

(3) We then calculate the energy difference in two local volume V' and V, which have the same local structure with respect to the defect, due to the variation in global material location,

$$\delta E_1 = \int_{V'} W dV - \int_V W dV \; . \tag{7.120}$$

From Fig. 7.8, one may observe that the area difference between V' and V is $\omega_1 - \omega_2$, i.e. adding the area ω_1 and removing the area ω_2. Hence the stored strain energy difference is

$$\delta E_1 = \int_{\omega_1} W dV - \int_{\omega_2} W dV \; . \tag{7.121}$$

Since $\delta\mathbf{X}$ is infinitesimal,

$$\omega_1 - \omega_2 = \int_{\omega_1 - \omega_2} dA = -\delta \boldsymbol{X} \cdot \int_{\mathcal{L}} ds \mathbf{n}$$

where $dA = -\delta\mathbf{X} \cdot \mathbf{n} ds$. shown in Fig. 7.9. Therefore,

$$\delta E_1 = -\delta\mathbf{X} \cdot \int_{\mathcal{L}} W d\ell \mathbf{n} = -\delta X_\ell \int_{\mathcal{L}} W ds n_\ell \; . \tag{7.122}$$

Note that in this step, all the operations are performed in the material configuration. We are comparing the energy difference between two adjacent local material volumes differing a translation.

(3) During a configuration change, the defect moves $+\delta\mathbf{X}$ from its original material position to the new material position, it will cause the relative material virtual displacement,

$$\delta u_i = \frac{\partial u_i}{\partial X_j} \delta X_j \; . \tag{7.123}$$

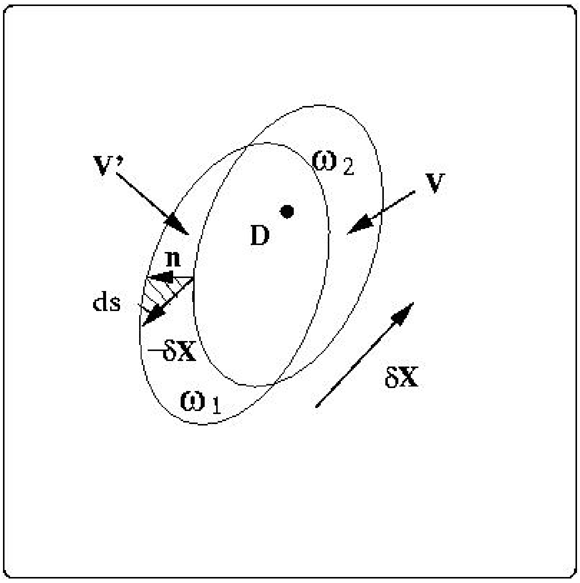


Figure 7.9. Eshelby's imaginary operation

This is to say that if there is no displacements along $\partial V'$, the displacement on ∂V is $\delta u_i = \dfrac{\partial u_i}{\partial X_j}\delta X_j \ \forall \mathbf{X} \in \mathcal{L}$. Then the difference of the work done to the environment of the two local configurations is:

$$\begin{aligned} \delta W^{ext} &= \int_{\mathcal{L}'} 0 \cdot T_i ds - \int_{\mathcal{L}} \delta u_i T_i ds = -\int_{\mathcal{L}} \delta u_i \sigma_{ij} n_j ds \\ &= -\int_{\mathcal{L}} u_{i,k}\sigma_{ij} n_j ds \delta X_k \ , \end{aligned} \tag{7.124}$$

which will cause the decrease of the potential energy of the local configuration, i.e. $\delta E_2 = -\delta W^{ext}$.

Then the total variation due to the change of configuration is,

$$\begin{aligned} \delta E &= \delta E_1 + \delta E_2 = -\delta X_\ell \left\{ \oint_{\mathcal{L}} (W n_\ell - u_{i,\ell}\sigma_{ij} n_j) ds \right\} \\ &= = -\delta X_\ell \left\{ \oint_{\mathcal{L}} \Big(W \delta_{\ell k} - u_{i,\ell}\sigma_{ik} \Big) n_k ds \right\} \end{aligned}$$

To honor the tradtion, the force on the defect is defined to be minus the rate of increase of the total potential energy of the system, i.e.

$$\delta E = -\mathbf{F}^{inh} \cdot \delta \mathbf{X} = \frac{\partial E}{\partial X_\ell} \delta X_\ell \tag{7.125}$$

Therefore the force acting on the inhomogeneity is

$$\mathbb{F}_\ell^{inh} = \oint_{\mathcal{L}} \Big(W\delta_{\ell k} - u_{i,\ell}\sigma_{ik} \Big) n_k ds \ . \tag{7.126}$$

In two-dimensional space, the special case, $\ell = 1$, is Rice's celebrated J-integral,

$$\mathbb{F}_1^{inh} = J = \oint_{\mathcal{L}} \Big(W dx_2 - u_{i,1}\sigma_{ik} n_k ds \Big) \ , \tag{7.127}$$

which can be interpreted as the driving force of a crack that grows along x-axis.

The integrand of (7.126) is Eshelby's another celebrate tensor: the energy-momentum tensor. The name comes from the fact that the tensor is obtained by tranlating or giving a motion to the energy of a local configuration. We denote it as

$$P_{\ell k} = W\delta_{\ell k} - u_{i,\ell}\sigma_{ik} \ . \tag{7.128}$$

Just like the Peach-Koehle force, Eshelby's energy momentum tensor was inspired by an electromagnetic analogy as well. As Eshelby pointed out, "*the archetypal energy-momentum tensor is Maxwell's stress tensor in electromagnetics.*" We juxtapose the two for comparison,

$$\mathbb{P}^E = W\mathbf{1}^{(2)} - \mathbf{E} \otimes \mathbf{D} \tag{7.129}$$

$$\mathbb{P}^M = W\mathbf{1}^{(2)} - \nabla \mathbf{u} \otimes \boldsymbol{\sigma} \ . \tag{7.130}$$

where the supercripts, E and M, denote mechanical and electrical energy-momentum tensors respectively.

In the following, we show that the energy-momentum tensor is divergence-free in homogeneous solid, which is in essence the path-independence of the J-integral.

The straightforward differentiation gives,

$$\begin{aligned} \frac{\partial P_{\ell k}}{\partial x_k} &= \frac{\partial W}{\partial \epsilon_{mn}} \frac{\partial \epsilon_{mn}}{\partial x_k} \delta_{\ell k} - u_{i,\ell k}\sigma_{ik} - u_{i,\ell}\sigma_{ik,k} \\ &= \sigma_{mn} u_{m,nk} \delta_{\ell k} - u_{i,\ell k}\sigma_{ik} \\ &= \sigma_{mn} u_{m,n\ell} - \sigma_{ik} u_{i,k\ell} = 0 \ . \end{aligned} \tag{7.131}$$

Therefore, for homogenous solids,

$$\mathbb{F}_\ell = \oint_{\mathcal{L}} P_{\ell k} n_k ds = \oint_{\mathcal{L}} \Big(W\delta_{\ell k} - u_{i,\ell}\sigma_{ik}) n_k ds = 0 \ . \tag{7.132}$$

For inhomogeneous solids, the above statement is no longer true, this is because,

$$\frac{\partial C_{ijmn}(\mathbf{x})}{\partial x_k} \neq 0,$$

and

$$\frac{\partial W}{\partial x_k} \neq \sigma_{mn} u_{m,nk} \ .$$

Suppose that there is a defect at a material point ξ_i, we assume that this may be captured by an equivalent inhomogeneous elastic stiffness tensor $C_{ijk\ell}(\mathbf{X} - \boldsymbol{\xi})$, i.e.

$$C_{ijk\ell}(\mathbf{X} - \boldsymbol{\xi}) = \begin{cases} C^0_{ijk\ell}, & \forall \mathbf{X} \neq \boldsymbol{\xi} \\ C_{ijk\ell}(\boldsymbol{\xi}), & \forall \mathbf{X} = \boldsymbol{\xi} \end{cases} \tag{7.133}$$

where

$$C_{ijk\ell}(\boldsymbol{\xi}) = C^0_{ijk\ell} - \frac{\partial^2 W^*}{\partial \epsilon_{ij} \partial \epsilon_{k\ell}} \ \text{ and } \ W^* = \frac{1}{2} C^{ijk\ell} \epsilon^*_{ij} \epsilon^*_{k\ell} \tag{7.134}$$

and ϵ^*_{ij} is the character eigenstrain of the defect.

Therefore, the total strain energy of the inhomogeneous body is

$$E = \frac{1}{2} \int_V C_{ijk\ell}(\mathbf{X} - \boldsymbol{\xi}) \epsilon_{ij} \epsilon_{k\ell} dV \tag{7.135}$$

By the definition,

$$\begin{aligned} \mathbb{F}^{inh}_n &= -\frac{\partial E}{\partial \xi_n} = -\frac{1}{2} \int_V \frac{\partial C_{ijk\ell}}{\partial \xi_n} \epsilon_{ij} \epsilon_{k\ell} dV \\ &= \frac{1}{2} \int_V C_{ijk\ell,m} (\delta_{mn} - u_{m,n}) \epsilon_{ij} \epsilon_{k\ell} dV \approx \frac{1}{2} \int_V C_{ijk\ell,n} \epsilon_{ij} \epsilon_{k\ell} dV \\ &= \frac{1}{2} \int_V \Big[\Big(C_{ijk\ell} \epsilon_{ij} \epsilon_{k\ell} \Big)_{,n} - 2 C_{ijk\ell} u_{i,j} u_{k,\ell n} \Big] dV \end{aligned}$$

Consider $C_{ijk\ell} u_{i,j} = \sigma_{k\ell}$ and integration by parts for the second term of the integrand.

$$\begin{aligned} \mathbb{F}^{inh}_n &= \int_V \Big[\Big(\frac{1}{2} C_{ijk\ell} \epsilon_{ij} \epsilon_{k\ell} \Big)_{,n} - (\sigma_{k\ell} u_{k,n})_\ell + \sigma_{k,\ell\ell} u_{k,n} \Big] dV \\ &= \int_V \Big[\Big(\frac{1}{2} C_{ijk\ell} \epsilon_{ij} \epsilon_{k\ell} \Big)_{,n} - (\sigma_{k\ell} u_{k,n})_{,\ell} \Big] dV \\ &= \oint_{\mathcal{L}} \Big(W \delta_{n\ell} - u_{k,n} \sigma_{k\ell} \Big) n_\ell ds = \oint_{\mathcal{L}} P_{n\ell} n_\ell ds \ . \end{aligned} \tag{7.136}$$

EXAMPLE 7.4 *The asymptotic stress fields for a mode III crack is*

$$\sigma_{13} = -\frac{K_{III}}{\sqrt{2\pi r}} \sin \frac{\theta}{2}, \quad \sigma_{23} = \frac{K_{III}}{\sqrt{2\pi r}} \cos \frac{\theta}{2} \quad . \tag{7.137}$$

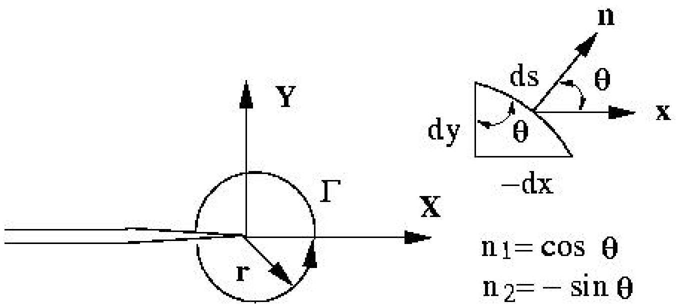


Figure 7.10. Contour for J-integral around a crack tip

We choose the integration contour $\Gamma : x_1 = r\cos\theta,\ x_2 = r\sin\theta,\quad -\pi \le \theta \le \pi$ *.*

The J-integral reads as follows,

$$\begin{aligned} J &= \oint_\Gamma \Big(W dx_2 - \frac{\partial u_i}{\partial x_1}\sigma_{ik} n_k ds \Big) \\ &= \int_{-\pi}^{\pi} \Big(W r\cos\theta - \frac{\partial u_3}{\partial x_1}(\sigma_{31}n_1 + \sigma_{32}n_2) r d\theta \Big) \end{aligned} \tag{7.138}$$

Consider $n_1 = \cos\theta, n_2 = -\sin\theta, \dfrac{\partial u_3}{\partial x_1} = 2\epsilon_{31} = \dfrac{\sigma_{31}}{\mu}$*, and* $W = K_{III}^2/(4\mu\pi r)$.

$$\begin{aligned} J &= \int_{-\pi}^{\pi} \frac{K_{III}^2}{4\mu\pi}\cos\theta d\theta - \frac{K_{III}^2}{2\pi\mu}\int_{-\pi}^{\pi}\Big(\sin^2\frac{\theta}{2}\cos\theta - \sin\frac{\theta}{2}\cos\frac{\theta}{2}\sin\theta\Big) d\theta \\ &= \frac{K_{III}^2}{2\pi\mu}\int_{-\pi}^{\pi}\Big(2\sin^2\frac{\theta}{2}\cos^2\theta - \sin^2\frac{\theta}{2}\Big(\cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2}\Big)\Big) d\theta \\ &= \frac{K_{III}^2}{2\pi\mu}\int_{-\pi}^{\pi}\sin^2\frac{\theta}{2} d\theta = \frac{K_{III}^2}{2\mu} \end{aligned} \tag{7.139}$$

7.5 Continuum theory of dislocation

One of the popular meso-scale simulations in solids is the discrete dislocation dynamics, which is often referred in the literature as DD. Since Kubin and Devincre's pioneer work, numerical simulations of dislocation dynamics has become an indispensible part of multiscale simulations. The current trend is to develop con-current multiscale simulations to couple the atomistic molecular dynamics (MD) simulations with continuum based dislocation dynamics (DD) simulations. In this section, we shall briefly introduce the basic concepts and theories of dislocation dynamics.

7.5.1 Volterra and Mura's formulas

We begin the discussions with the displacement and the stress fields of the curved dislocations. The general theory of curved dislocations in anisotropic media was developed by Volterra [1907], De Wit [1960, and Mura [1963,1968]. The special case of curved dislocation in an isotropic medium was attributed to Burgers [1924] and Peach & Koehler [1950]. The presentation in this book is an adaptation of Mura's work with contemprorary flavor.

Before we proceed to derive the Volterra and Mura's formulas, it is expedient to lay out some useful formulas. Consider a simply connected region, $\Omega \in \mathbb{R}^3$, with a smooth boundary. Define a characteristic function,

$$\chi(\mathbf{x}) = \begin{cases} 1, & \mathbf{x} \in \Omega \\ 0, & \mathbf{x} \notin \Omega \end{cases} \tag{7.140}$$

Consider a (slip) plane S that is characterized by its normal $\mathbf{n}$ and its distance to the origin of the coordinate, s. The Radon transform of $\chi(\mathbf{x})$ will be

$$\int_{-\infty}^{\infty} \chi(\mathbf{x}')\delta(s - \mathbf{n}\cdot\mathbf{x}')d\mathbf{x}' = \int_{S\cap\Omega} dS \tag{7.141}$$

if $\Omega = \mathbb{R}^3$, we have

$$\int_{-\infty}^{\infty} \chi(\mathbf{x}')\delta(s - \mathbf{n}\cdot\mathbf{x}')d\mathbf{x}' = \int_{-\infty}^{\infty} \delta(s - \mathbf{n}\cdot\mathbf{x}')d\mathbf{x}' = \int_{S} dS \tag{7.142}$$

Conceptually, we can generalize the Radon projection formula to a two-dimensional curved surface (2D manifold), S, i.e.

$$\int_{\Omega} f(\mathbf{x}')\delta(s - \mathbf{n}\cdot\mathbf{x}')d\mathbf{x}' = \int_{S\cap\Omega} f(\mathbf{x}')dS' \tag{7.143}$$

$$\int_{-\infty}^{\infty} f(\mathbf{x}')\delta(s - \mathbf{n}\cdot\mathbf{x}')d\mathbf{x}' = \int_{S} f(\mathbf{x}')dS' \tag{7.144}$$

or

$$\int_{\Omega} f(\mathbf{x}')\delta(S - \mathbf{x}')d\mathbf{x}' = \int_{S\cap\Omega} f(\mathbf{x}')dS' \tag{7.145}$$

$$\int_{-\infty}^{\infty} f(\mathbf{x}')\delta(S - \mathbf{x}')\mathbf{x}' = \int_{S} f(\mathbf{x}')dS' \tag{7.146}$$

where $\delta(S-\mathbf{x})$ is an abbrieviation of $\delta(dist(S,\mathbf{x}))$ and $dist(S,\mathbf{x}) = \inf\{|\mathbf{x} - \mathbf{y}|, \forall \mathbf{y} \in S\}$.

Now we consider the following integral,

$$\int_{S} \delta(\mathbf{x} - \mathbf{x}')dS' \tag{7.147}$$

where $\delta(\mathbf{x}-\mathbf{x}')$ is Dirac's delta function in three-dimensional space. Based on 7.146, we have

$$\int_S \delta(\mathbf{x}-\mathbf{x}')dS' = \int_{-\infty}^{\infty} \delta(\mathbf{x}-\mathbf{x}')\delta(S-\mathbf{x}')d\mathbf{x}' = \delta(S-\mathbf{x}) \qquad (7.148)$$

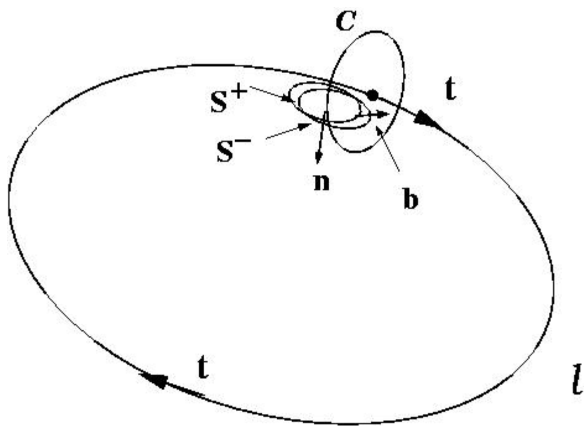


Figure 7.11. Curved dislocation loop $\mathcal{L}$ and the Burgers circuit $\mathcal{C}$.

Assume that there is a dislocation loop embedded in an elastic continuum. To define a dislocation line, we take the tangent at a position $\mathbf{x}$ on the dislocation loop, $\mathbf{t}$, as the local direction of the dislocatin. Obviously, $\mathbf{t}$ lies on the tangent plane at point $\mathbf{x}$. We denote the tangent plane at $\mathbf{x}$ as S. S is also the local slip plane. It is assumed that the upper plane of S (denoted by S^+) slips a distance **b** relative to its lower plane S^-. Choose a circuit around the vector $\mathbf{t}$ in a plane that is perpendicular to $\mathbf{t}$ (or $\mathbf{t}$ is the normal of the plane). Circle the circuit (the Burgers circuit) in a direction that makes **t**as a right-handed rotation vector.

In this definition, both the tangent vector $\mathbf{t}$ and the local Burgers vector, **b** could depend on the spatial location, though in the rest of the presentation, we assume that **b** is a constant vector. Note that the real slip plane may not be the tangent plane at $\mathbf{x}$, it could be a curved surface, but the tangent plane of the slip surface at the interception of Burgers circuit should coincide with the tangent plane of the dislocation loop at point $\mathbf{x}$.

To homogenize such dislocation field, one may assume that the total displacement gradient can be written as two parts,

$$u_{i,j} = \beta_{ij} + \beta^*_{ij} \qquad (7.149)$$

where β_{ij} is elastic distortion, and β^* is equivalent eigen-distortion, or plastic distortion. [2]

The total strain, ϵ_{ij}, elastic strain, e_{ij}, and eigenstrain, $\epsilon^*{}_{ij}$ can be expressed as

$$\begin{aligned}\epsilon_{ij} &= \frac{1}{2}\Big(u_{i,j} + u_{j,i}\Big) && (7.150)\\ e_{ij} &= \frac{1}{2}\Big(\beta_{ij} + \beta_{ji}\Big) && (7.151)\\ \epsilon^*{}_{ij} &= \frac{1}{2}\Big(\beta^*_{ij} + \beta^*_{ji}\Big) && (7.152)\end{aligned}$$

where the eigen-distortion is prescribed as

$$\beta^*_{ji} = -b_i n_j \delta(S - \mathbf{x}) \tag{7.153}$$

where the normal vector, **n**, is pointing from S^+ to S^-.

The eigen-distortion caused by slip b_i of plane S^+ may be wretten as

$$\beta^*_{ji}(\mathbf{x}) = -b_i n_j \delta(S - \mathbf{x}) \tag{7.154}$$

(Question: why is there a minus sign?) Therefore,

$$\epsilon^*{}_{ij} = -\frac{1}{2}\Big(b_i n_j + b_j n_i\Big)\delta(S - \mathbf{x}) \tag{7.155}$$

Therefore,

$$\begin{aligned}u_i(\mathbf{x}) &= -\int_{-\infty}^{\infty} C_{j\ell mn}\epsilon^*{}_{mn}(\mathbf{y})G_{ij,\ell}(\mathbf{x}-\mathbf{y})d\mathbf{y}\\ &= \int_{-\infty}^{\infty} C_{j\ell mn}\epsilon^*{}_{mn}(\mathbf{y})\delta(S-\mathbf{y})G_{ij,\ell}(\mathbf{x}-\mathbf{y})d\mathbf{y}\\ &= \int_S C_{j\ell mn} b_m n_n G_{ij,\ell}(\mathbf{x}-\mathbf{y})dS_{\mathbf{y}} && (7.156)\end{aligned}$$

The above expression was derived by Volterra, and it is called Volterra formula (Volterra [1907]).

Differentiating (7.156) yields

$$u_{i,j}(\mathbf{x}) = \int_S C_{j\ell mn} b_m n_n G_{ij,\ell j}(\mathbf{x}-\mathbf{y})dS_{\mathbf{y}} \tag{7.157}$$

and the elastic distortion becomes

$$\beta_{ji}(\mathbf{x}) = \int_S C_{j\ell mn} b_m n_n G_{ij,\ell j}(\mathbf{x}-\mathbf{y})dS_{\mathbf{y}} + b_i n_j \delta(S - \mathbf{x}) \tag{7.158}$$

[2] There are many attempts to derive plasticity theory from this formualtion.

Mura showed (Mura [1963]) that the above surface integration can be written as a line integration,

$$\beta_{ji}(\mathbf{x}) = \oint_L e_{jnh} C_{pqmn} G_{ip,q}(\mathbf{x}-\mathbf{y}) b_m t_h d\ell_{\mathbf{y}} \tag{7.159}$$

which is termed as Mura's formula.

To prove the equivalency between (7.159) and (7.158), we first consider Stokes' theorem of a third order tensor field, $\mathbf{A} = A_{jih}\mathbf{e}_j \otimes \mathbf{e}_h$.

$$\int_S \mathbf{n} \cdot (\nabla \times \mathbf{A}) dS = \oint \mathbf{t} \cdot \mathbf{A} d\ell \tag{7.160}$$

or in component form

$$\int_S e_{k\ell h} n_k A_{jih,\ell} dS = \oint t_h A_{jih} d\ell \tag{7.161}$$

Let $A_{jih} = e_{jnh} C_{pqmn} b_m G_{ip,q}$. We have

$$\begin{aligned} & \oint_L e_{jnh} C_{pqmn} b_m G_{ip,q}(\mathbf{x}-\mathbf{y}) t_h d\ell_{\mathbf{y}} \\ = & -\int_S e_{k\ell h} n_k \Big(e_{jnh} C_{pqmn} b_m G_{ip,q\ell}(\mathbf{x}-\mathbf{y}) \Big) dS_{\mathbf{y}} \end{aligned} \tag{7.162}$$

where $G_{ip,q\ell} = -\dfrac{\partial}{\partial x'_\ell} G_{ip,q}$. Utilizing the identity $e_{k\ell h} e_{jnh} = \delta_{kj}\delta_{\ell n} - \delta_{kn}\delta_{\ell j}$, one can obtaion

$$\begin{aligned} & -\int_S (\delta_{kj}\delta_{\ell n} - \delta_{kn}\delta_{\ell j}) n_k b_m C_{pqmn} G_{ip,q\ell}(\mathbf{x}-\mathbf{x}') dS' \\ = & -\int_S \Big(n_j b_m C_{pqm\ell} G_{ip,q\ell}(\mathbf{x}-\mathbf{x}') - n_n b_m C_{pqmn} G_{ip,qj}(\mathbf{x}-\mathbf{x}') \Big) dS \\ = & \int_S \Big(n_j b_m \delta_{im} \delta(\mathbf{x}-\mathbf{x}') + n_n b_m C_{pqmn} G_{ip,qj}(\mathbf{x}-\mathbf{x}') \Big) dS' \\ = & \int_\Omega n_j b_i \delta(S-\mathbf{x}') \delta(\mathbf{x}-\mathbf{x}') d\mathbf{x}' + \int_S n_n b_m C_{pqmn} G_{ip,qj}(\mathbf{x}-\mathbf{x}') dS' \\ = & \; n_j b_i \delta(S-\mathbf{x}) + \int_S n_n b_m C_{pqmn} G_{ip,qj}(\mathbf{x}-\mathbf{x}') dS' \end{aligned} \tag{7.163}$$

Finally, we showed that (7.158) is equivalent to (7.159).

7.5.2 The Burgers formula

For isotropic materials, the Volterra formula can be simplified and explicited expressed in terms of elementary line integrals, which are instrumental in contemporary discrete dislocation dynamics formulations.

To derive the Burgers formula, we start from the Volterra formula,

$$u_m(\mathbf{x}) = b_i \int_S C_{ijk\ell} G^{\infty}_{km,\ell}(\mathbf{x} - \mathbf{x}^{'}) dS^{\prime}_j \tag{7.164}$$

where the surface S is the dislocation surface, which is a cap of dislocation line $C = \partial S$, and $dS^{\prime}_j := n_j dS$.

For isotropic materials, both the elastic tensor and the Green's function are quite amieble

$$\begin{aligned} C_{ijk\ell} &= \lambda \delta_{ij}\delta_{k\ell} + \mu(\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}) && (7.165)\\ G^{\infty}_{km}(\mathbf{x}) &= \frac{1}{8\pi\mu}\Big[\delta_{km} r_{,pp} - \frac{\lambda+\mu}{\lambda+2\mu} r_{,km}\Big] \; . && (7.166)\end{aligned}$$

Denote $\mathbf{R} = \mathbf{x} - \mathbf{x}^{'}$ and $R = |\mathbf{x} - \mathbf{x}^{'}| = \sqrt{(x_i - x^{\prime}_i)(x_i - x^{\prime}_i)}$.

Then,

$$\begin{aligned} & C_{ijk\ell} G^{\infty}_{km,\ell}(\mathbf{R}) = (\lambda\delta_{ij}\delta_{k\ell} + \mu(\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}))\frac{1}{8\pi\mu}\Big[\delta_{km} R_{,pp\ell} \\ & -\frac{\lambda+\mu}{\lambda+2\mu} R_{,km\ell}\Big] = \frac{1}{8\pi\mu}\Big\{\frac{\lambda\mu}{\lambda+\mu}\delta_{ij} R_{,ppm} \\ & +\mu(\delta_{im} R_{,ppj} + \delta_{jm} R_{,ppi}) - 2\Big(\frac{\lambda+\mu}{\lambda+2\mu}\Big)\mu R_{,mij}\Big\} \end{aligned} \tag{7.167}$$

Utilizing the identity,

$$\frac{\lambda}{\lambda+2\mu} = 2\frac{(\lambda+\mu)}{\lambda+2\mu} - 1 \; ,$$

one may find that

$$\begin{aligned} b_i C_{ijk\ell} G^{\infty}_{km,\ell}(\mathbf{R}) &= \frac{1}{8\pi\mu}\{\mu b_m R_{,ppj} + \mu(b_\ell R_{,pp\ell}\delta_{jm} - b_j R_{,ppm}) \\ &+ 2\Big(\frac{\lambda+\mu}{\lambda+2\mu}\Big)\mu\Big(b_j R_{,ppm} - b_i R_{,mij}\Big)\Big\} \end{aligned} \tag{7.168}$$

Changing the dummy variable, we can then write

$$\begin{aligned} u_m(\mathbf{x}) &= \frac{1}{8\pi}\int_S b_m R_{,ppj} dS^{'}_j + \frac{1}{8\pi}\int_S \Big(b_\ell R_{,pp\ell} dS^{'}_m - b_\ell R_{,ppm} dS^{'}_\ell\Big) \\ &+ \frac{1}{4\pi}\frac{\lambda+\mu}{\lambda+2\mu} b_j \int_S (R_{,pmp} dS_j - R_{,jmp} dS^{'}_p) \; . \end{aligned} \tag{7.169}$$

Consider Stoke's theorem,

$$\int_S (\nabla \times \mathbf{A}) \cdot d\mathbf{S} = \oint_{\partial S} \mathbf{A} \cdot d\ell \; . \tag{7.170}$$

Let,

$$\nabla = \frac{\partial}{\partial x_m}\mathbf{e}_m, \;\; \mathbf{A} = A_{,\cdots}\mathbf{e}_n, d\mathbf{S} = dS_k\mathbf{e}_k, \text{ and } d\boldsymbol{\ell} = t_k d\ell \mathbf{e}_k = dx_k\mathbf{e}_k \; .$$

A special case of the Stoke's theorem is,

$$\int_S \epsilon_{mnk}\frac{\partial A_{,\cdots}}{\partial x_m}dS_k = \oint_{\partial S} A_{,\cdots}dx_n \; . \tag{7.171}$$

Change the free-index, $n \rightarrow k$,

$$-\int_S \epsilon_{mnk}\frac{\partial A_{,\cdots}}{\partial x_m}dS_n = \oint_{\partial S} A_{,\cdots}dx_k \; . \tag{7.172}$$

We then have

$$\begin{aligned} -\epsilon_{ijk}\epsilon_{mnk}\int_S A_{,\cdots m}dS_n &= \epsilon_{ijk}\oint_{\partial S} A_{,\cdots}dx_k \\ -(\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm})\int_S A_{,\cdots m}dS_n &== \epsilon_{ijk}\oint_{\partial S} A_{,\cdots}dx_k \end{aligned} \tag{7.173}$$

which eventually leads to the desired form,

$$\int_S \Big(A_{,\cdots j}dS_i - A_{,\cdots i}dS_j\Big) = \epsilon_{ijk}\oint_{\partial S} A_{,\cdots}dx_k \; . \tag{7.174}$$

In (7.169), we may view $R_{,pp}$ as $A_{,pp}$ in the second integral and $R_{,mp}$ as $A_{,mp}$ in the third integral and then apply the Stoke's theorem (7.174) to (7.169),

$$\begin{aligned} & b_\ell \int_S \Big(R_{,pp\ell}dS'_m - R_{,ppm}dS'_\ell\Big) = -b_\ell \int_S \Big(R_{,pp\ell'}dS'_m - R_{,ppm'}dS'_\ell\Big) \\ & = -b_\ell \oint_C \epsilon_{m\ell k}R_{,pp}dx'_k \\ & b_j \int_S \Big(R_{,pmp}dS'_j - R_{,pmj}dS'_p\Big) = -b_j \int_S \Big(R_{,pmp'}dS'_j - R_{,pmj'}dS'_p\Big) \\ = \;\; & -b_j \oint_C \epsilon_{jpk}R_{,pm}dx'_k \end{aligned}$$

We derive the Burgers formula,

$$\begin{aligned} u_m(\mathbf{x}) \;\; = \;\; & \frac{1}{8\pi}\int_S b_m R_{,ppj}dS'_j - \frac{1}{8\pi}\int_C b_\ell\epsilon_{m\ell k}R_{,pp}dx'_k \\ - \;\; & \frac{1}{8\pi(1-\nu)}\int_C b_j\epsilon_{jpk}R_{,mp}dx'_k \; . \end{aligned} \tag{7.175}$$

In the last line, the identity $\dfrac{\lambda+\mu}{\lambda+2\mu} = \dfrac{1}{2(1-\nu)}$ is used. Consider the fact that

$$R_{,j} = \frac{x_j - x_j^{'}}{R} = \frac{R_j}{R}, \text{ and } R_{,mp} = \frac{\delta_{mp}}{R} - \frac{R_m R_p}{R^3}$$

hence

$$R_{,pp} = \frac{2}{R} \text{ and } R_{,ppj} = \frac{-2R_j}{R^3} \ .$$

Therefore,

$$\begin{aligned} u_m(\mathbf{x}) &= -\frac{1}{4\pi}\int_S \frac{b_m R_j}{R^3} dS_j^{'} - \frac{1}{4\pi}\oint_C \frac{\epsilon_{m\ell k} b_\ell}{R} dx_k^{'} \\ &\quad -\frac{1}{8\pi(1-\nu)}\oint_C \epsilon_{pjk} b_j \frac{\partial}{\partial x_m}\Big(\frac{R_p}{R}\Big) dx_k^{'} \end{aligned} \tag{7.176}$$

which can be put into an elementary vector form, i.e. the Burgers formula

$$\mathbf{u}(\mathbf{x}) = -\frac{\mathbf{b}}{4\pi}\Omega - \frac{1}{4\pi}\int_C \frac{\mathbf{b}\times d\boldsymbol{\ell}^{'}}{R} - \frac{1}{8\pi(1-\nu)}\nabla \oint_C \frac{\mathbf{b}\times\mathbf{R}\cdot d\boldsymbol{\ell}^{'}}{R} \ . \tag{7.177}$$

In (7.177), $d\boldsymbol{\ell}^{'} = t_k d\ell \mathbf{e}_k = dx_k^{'}\mathbf{e}_k$, and Ω is the so-called solid angle, which is defined as the surface area Ω of a unit sphere covered by the surface's projection onto the sphere. In this case, the angle is subtended by the dislocation surface, S, i.e.

$$\Omega = \int_S \frac{R_j dS_j^{'}}{R^3} = \int_S \frac{\mathbf{n}\cdot d\mathbf{S}^{'}}{R^2} \tag{7.178}$$

where $\mathbf{n} := \mathbf{R}/R$ is a unit vector from the point $\mathbf{x}$ to the dislocation surface, S.

If the surface is a sphere, $d\mathbf{S} = R^2 d\boldsymbol{\omega}$ and

$$\begin{aligned} \Omega &= \oint_{S_2} \frac{R^2\mathbf{n}\cdot d\boldsymbol{\omega}}{R^3} = \oint_{S_2} \mathbf{n}\cdot d\boldsymbol{\omega} \\ &= \oint_{S_2} n_i n_i d\omega = 4\pi \ . \end{aligned} \tag{7.179}$$

7.5.3 Peach-Koehler stress formula for dislocation loop

The objective of this section is to express stress field of a dislocation loop in terms of line integral. Take derivative of the Bergurs' displacement formula,

$$\begin{aligned} u_{m,\ell} &= \frac{1}{8\pi}\int_S b_m R_{,ppj\ell} dS_j^{'} - \frac{1}{8\pi}\oint_C \epsilon_{mnk} b_n R_{,pp\ell} dx_k^{'} \\ &= -\frac{1}{8\pi(1-\nu)}\oint_C \epsilon_{jpk} b_j R_{,mp\ell} dx_k^{'} \end{aligned} \tag{7.180}$$

In the above equation, only the first term is not a line integral. Nevertheless, we claim that

$$\int_S b_m R_{,pp\ell j} dS_j^{'} = -8\pi\delta(S-\mathbf{x}) b_m n_\ell - b_m \oint_C \epsilon_{j\ell k} R_{,ppj} dx_k^{'} \, .$$

Proof:

Apply Stokes' theorem,

$$\oint_C \epsilon_{ijk}\phi dx_k^{'} = \int_S \Big[\phi_{,j} dS_i - \phi_{,i} dS_j^{'}\Big] \tag{7.181}$$

to the above expression,

$$\begin{aligned} \oint_C \epsilon_{i\ell k} R_{,pp} dx_k^{'} &= \int_S \Big(R_{,pp\ell'} dS_j^{'} - R_{,ppj'} dS_\ell^{'} \\ &= \int_S \Big(R_{,ppj} dS_\ell^{'} - R_{,pp\ell} dS_j^{'} \Big) \end{aligned} \tag{7.182}$$

Therefore,

$$\frac{\partial}{\partial x_j} \oint_C \epsilon_{j\ell k} R_{,pp} dx_k^{'} = \int_S \Big[R_{,ppjj} dS_\ell^{'} - R_{,pp\ell j} dS_j^{'} \Big] \tag{7.183}$$

Since

$$G^P(\mathbf{x}-\mathbf{x}^{'}) = \frac{1}{4\pi R}, \text{ and } \nabla^2 G^p = -\delta(\mathbf{x}-\mathbf{x}^{'}),$$

we then have

$$R_{,pp} = \frac{2}{R} = 8\pi G^P(\mathbf{x}-\mathbf{x}^{'}) \text{ and } R_{,ppjj} = 8\pi\nabla^2 G^P(\mathbf{x}-\mathbf{x}^{'}) = -8\pi\delta(\mathbf{x}-\mathbf{x}^{'}).$$

Consequently,

$$b_m \oint_C \epsilon_{j\ell k} R_{,ppj} dx_k^{'} = -8\pi b_m \int_S \delta(\mathbf{x}-\mathbf{x}^{'}) dS_\ell^{'} - b_m \int_S R_{,pp\ell j} dS_j^{'}$$

Use Radon transformation,

$$\begin{aligned} & \int_S \delta(\mathbf{x}-\mathbf{x}^{'}) dS_\ell^{'} = \int_S \delta(\mathbf{x}-\mathbf{x}^{'}) n_\ell dS \\ = & \int_{\mathbb{R}^3} \delta(\mathbf{x}-\mathbf{x}^{'}) n_\ell \delta(S-\mathbf{x}^{'}) d\Omega^{'} = \delta(S-\mathbf{x}) n_\ell \end{aligned} \tag{7.184}$$

Hence, we verfied the claim.

Note that $\beta^*_{m\ell} = -8\pi b_m n_\ell \delta(S-\mathbf{x})$, we again recover Mura's formula

$$\begin{aligned} \beta_{m\ell} &= u_{m,\ell} - \beta^*_{m\ell} = -\frac{1}{8\pi} \oint_C \epsilon_{j\ell k} b_m R_{,ppj} dx_k^{'} \\ &- \frac{1}{8\pi} \oint_C \epsilon_{mnk} b_n R_{,pp\ell} dx_k^{'} - \frac{1}{8\pi(1-\nu)} \oint_C \epsilon_{jpk} b_j R_{,mp\ell} dx_k^{'} \end{aligned} \tag{7.185}$$

Shifting the dummy indices, one may find that

$$\begin{aligned} e_{ij} &= \frac{1}{2}(\beta_{ij}+\beta_{ji}) = \frac{1}{8\pi}\oint_C \Big\{-\frac{1}{2}\Big(\epsilon_{jk\ell}b_i R_{,\ell} + \epsilon_{ik\ell}b_j R_{,\ell} \\ &\quad -\epsilon_{jk\ell}b_\ell R_{,i} - \epsilon_{ik\ell}b_\ell R_{,j}\Big) + \frac{1}{1-\nu}\epsilon_{mnk}b_n R_{,ijm}\Big\}\, dx_k^{'} \end{aligned} \quad (7.186)$$

Repeatly using the e-δ identity $\epsilon_{pij}\epsilon_{pmn} = \delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}$, one has

$$\begin{aligned} \epsilon_{jk\ell}(b_i R_{,\ell} - b_\ell R_{,i}) &= \epsilon_{jk\ell}(\delta_{is}\delta_{\ell t} - \delta_{\ell s}\delta_{it})b_s R_{,t} = \epsilon_{jk\ell}\epsilon_{i\ell p}\epsilon_{stp}b_s R_{,t} \\ &= \epsilon_{pst}\epsilon_{jk\ell}\epsilon_{ip\ell}b_s R_{,t} = \epsilon_{pst}(\delta_{ji}\delta_{kp} - \delta_{jp}\delta_{ki})b_s R_{,t} \\ &= (\epsilon_{kst}\delta_{ji} - \epsilon_{jst}\delta_{ki})b_s R_{,t} \end{aligned} \quad (7.187)$$

Similarly, one may find,

$$\epsilon_{ik\ell}(b_j R_{,\ell} - b_\ell R_{,j}) = (\epsilon_{kst}\delta_{ij} - \epsilon_{ist}\delta_{kj})b_s R_{,t} \quad (7.188)$$

which enable us to write

$$\begin{aligned} e_{ij} &= \frac{1}{8\pi}\oint_C \Big\{-b_s R_{,ppt}\Big[\epsilon_{kst}\delta_{ij} - \frac{1}{2}\epsilon_{ist}\delta_{kj} - \frac{1}{2}\epsilon_{jst}\delta_{ki}\Big] \\ &\quad + \frac{1}{1-\nu}\epsilon_{mnk}b_n R_{,ijm}\Big\}\, dx_k^{'} \end{aligned} \quad (7.189)$$

For linear isotropic elastic materials,

$$\sigma_{ij} = C_{ijk\ell}e_{k\ell}, \ \text{ and } \ C_{ijk\ell} = \lambda\delta_{ij}\delta_{k\ell} + \mu(\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}) \quad (7.190)$$

Finally, one can obtain the Peach-Koehler formula for stress field of a dislocation loop,

$$\begin{aligned} \sigma_{ij} &= \frac{\mu}{4\pi}\oint_C \Big(\frac{b_n}{2}R_{,mpp} + (\epsilon_{jmn}dx_i^{'} + \epsilon_{imn}dx_j^{'}) \\ &\quad + \frac{b_n}{1-\nu}\epsilon_{kmn}(R_{,ijm} - \delta_{ij}R_{,ppm})dx_k^{'}\Big) \end{aligned} \quad (7.191)$$

Considering,

$$\begin{aligned} R_{,ppm} &= -\frac{2R_m}{R^3} = \frac{\partial}{\partial x_m}\Big(\frac{2}{R}\Big) \\ R_{,ijm} &= \nabla_m^{'}\cdot\Big(\nabla_i \otimes \nabla_j R)\Big) \end{aligned} \quad (7.192)$$

One can re-write the Peach-Koehler formula in a vector form,

$$\begin{aligned} \boldsymbol{\sigma} &= \frac{\mu}{4\pi}\oint_C (\mathbf{b}\times\nabla^{'})\frac{1}{R}\otimes d\boldsymbol{\ell}^{'} + \frac{\mu}{4\pi}\oint_C d\boldsymbol{\ell}^{'}\otimes(\mathbf{b}\times\nabla^{'})\frac{1}{R} \\ &= -\frac{\mu}{4\pi(1-\nu)}\oint_C \nabla^{'}\cdot(\mathbf{b}\times d\boldsymbol{\ell}^{'})\cdot(\nabla\otimes\nabla - \mathbf{1}\nabla^2)R\,. \end{aligned} \quad (7.193)$$

7.6 Discrete Dislocation Dynamics (DD)

The first discrete dislocation dynamics simulation was attempted in late 1980s by Lepinous and Kubin [1987] and Ghoniem and Amodeo [1988]. The simulations were conducted then were the interactions among infinitely long straight dislocations. Since 1990s, more realistic DD simulations have been proposed in situations that are involved with more complicated micro-structures. In the following, we shall outline one of the latest formulations of DD simulations.

7.6.1 Galerkin weak form formulation

The Galerkin weak form formulation is proposed by Ghoniem and Sun and their co-workers.

The following presentation is mainly based on a series papers by Ghoniem et al [1990] [2000], and [2004].

In this approach, the formulation focus on simulating one dislocation loop among many different dislocation loops.

To formulate the discrete dislocation dynamics, we employ the virtual work principle. For a given virtual displacement field, $\delta \mathbf{x}$, the virtual work will be balanced on the dislocation loop considered.

The internal virtual work consists of the virtual work done by all the stresses acting on the dislocation loop, which includes the stress fields of all other dislocation loops and the stress field due to external loads, the virtual work done by the self-stress field. The external virtual work is mainly the virtual work done by the friction forces that resist the motion of the dislocation loop.

We first consider the virtual work due to all other internal stresses except the self-stress,

$$\begin{aligned} \delta W_{PK} &= \oint_C d\mathbf{F}_{PK} \cdot \delta \mathbf{x} = \oint_C \Big[(\mathbf{b} \cdot \mathbf{\Sigma}) \times d\boldsymbol{\ell} \Big] \cdot \delta \mathbf{x} \\ &= \oint_C \Big(\mathbf{b} \cdot \mathbf{\Sigma} \times \mathbf{t} \Big) d\ell \cdot \delta \mathbf{x} = \oint_C (\epsilon_{ijk} \Sigma_{jm} b_m t_k \delta x_i) d\ell \end{aligned} \quad (7.194)$$

where $\mathbf{b}$ is the Burgers vector, $\mathbf{t}$ is the tangential vector along the dislocation loop, and

$$\Sigma_{ij} = \sigma^I_{ij} + \sigma^e_{ij} \quad (7.195)$$

Here σ^I_{ij} are the stress fields of all other dislocation loops inside the solid, which can be expressed as

$$\begin{aligned} \sigma^I_{ij} &= \frac{\mu}{4\pi} \oint_C b_n \Big[\frac{1}{2} R_{,mpp} (\epsilon_{jmn} dx'_i + \epsilon_{imn} dx'_j) \\ &\quad + \frac{1}{1-\nu} \epsilon_{kmn} (R_{,ijm} - \delta_{ij} R_{,ppm}) \Big] dx'_k \end{aligned} \quad (7.196)$$

and σ^e_{ij} is the stress field due to externally applied loads.

Denote

$$f_i^{PK} = \epsilon_{ijk}\Sigma_{jm}b_m t_k. \tag{7.197}$$

One may write

$$\delta W_{PK} = \oint_C f_i^{PK} d\ell \delta x_i. \tag{7.198}$$

In principle, the virtual work done by the self stress field can be also expressed by Eq. (7.196). However, in that case, Eq. (7.196) would become a singular integral, which can be evaluated in the sense of Cauchy principal value.

Since the core of a dislocation loop has specific physical meanings, it would be appropriate to treat the virtual work of self-stress field separately. Gavazza and Barnett [1976] expressed the virtual work of the self-stress field of planar curved dislocation loop in terms of a single integral expression,

$$\begin{aligned} \delta W_{self} \;\; = & \;\; \oint_C \Big\{ \Big[E(\mathbf{t}) - \Big(E(\mathbf{t}) + E^{''}(\mathbf{t}) \Big) \ln\Big(\frac{8}{\epsilon\kappa} \Big) \Big] \kappa - J(L,p) \Big\} \mathbf{n} \cdot \delta \mathbf{x} d\ell \\ & + [dU]_{core} \end{aligned} \tag{7.199}$$

where $E(\mathbf{t}) = \frac{1}{2}\sigma_{ij}(\mathbf{t}) b_i n_j$, ϵ is related to the core size, κ is the curvature of the dislocation line, $J(L,p)$ is a non-local interaction term, and $[dU]_{core}$ is the virtual work contribution from the core of the dislocation loop. Since $[dU]_{core}$ is related to the dislocation mobility, this term may be absorbed into the friction force.

Let,

$$\mathcal{E}^{self} = \Big\{ E(\mathbf{t}) - \Big(E(\mathbf{t}) + E^{''}(\mathbf{t}) \Big) \ln\Big(\frac{8}{\epsilon\kappa} \Big) \Big] \kappa - J(L,p) \Big\} \tag{7.200}$$

and

$$f_i^{self} = \mathcal{E} n_i \tag{7.201}$$

The total active forces acting on a dislocation loop are

$$f_i^T = f_i^{PK} + f_i^{self} \tag{7.202}$$

In many cases, it has to include the change of chemical potential induced *Osmotic force*. Since the change in chemical potnetial per vacncy or interstitial will cause the dislocation loop climbing, or causing the none-conservative dislocation loop movement, the *Osmotic force* is usually responsible for the dislocation loop climb (see Hirth, Rhee, and Zbib [1996]).

When a dislocation loop starting to move, it has to overcome the friction forces that resist its motion. The friction forces consist of (1) extrinsic resistances due to alloying, impurity atoms, Peierls stress (this part of force coming from $[dU]_{core}$), etc., and (2) Intrinsic friction forces that are due to the atomistic bond force in a surface separation (fracture) process. Empirically, one can always assume that the friction forces are proportional to the dislocation velocity, such that

$$\delta W^{friction} = \oint_C C_{ik} V_k d\ell \delta x_i = \oint_C \mathbf{C} \cdot \mathbf{V}) d\ell \cdot \delta \boldsymbol{x} \tag{7.203}$$

where

$$\mathbf{V} = \frac{d\mathbf{x}}{dt} \tag{7.204}$$

and $\mathbf{C}$ is called the resistivity matrix, which has three independent components in an isotropic medium (two for glide motion and one for climb motion),

$$[C_{ik}] = \begin{bmatrix} C_1 & 0 & 0 \\ 0 & C_2 & 0 \\ C_1 & 0 & C_3 \end{bmatrix} \tag{7.205}$$

Then the principle of virtual reads

$$\delta W^{int} - \delta W^{fric} = 0, \quad \Rightarrow \quad \oint_C \Big(f_i^T - C_{ik} V_k \Big) d\ell \delta x_i = 0 \; . \tag{7.206}$$

7.6.2 Finite element implementation

Truncating the dislocation loop into N_s segments, and mapping each segment into a one-dimensional parametric space, i.e., $N_I : [\mathbf{x}_{I-1}, \mathbf{x}_I] \quad \rightarrow \quad u \in [0, 1]$. Thereby, for $\mathbf{x} \in N_I$,

$$d\ell = \sqrt{\Big(\frac{\partial x_i}{\partial u} \frac{\partial x_i}{\partial u} \Big)} du \tag{7.207}$$

Consider the finite element discrettization,

$$x_i^h(u,t) = \sum_{m=1}^{N_{DF}} N_{im}(u) q_m(t) \tag{7.208}$$

where $N_{im}(u)$ is the finite element shape function. The discreteized velocity field is

$$V_i^h = x_{i,t}^h = \sum_{m=1}^{N_{DF}} N_{im}(u) q_{m,t}(t) \; . \tag{7.209}$$

Denote the gradient of FEM shape function as $B_{im}(u) := N_{im,u}(u)$. The line integration element will be

$$d\ell = (x_\ell x_\ell)^{1/2} du = \Big(\sum_{p,s=1}^{N_{DF}} q_p q_s B_{\ell p}(u) B_{\ell s}(u) \Big)^{1/2} du \tag{7.210}$$

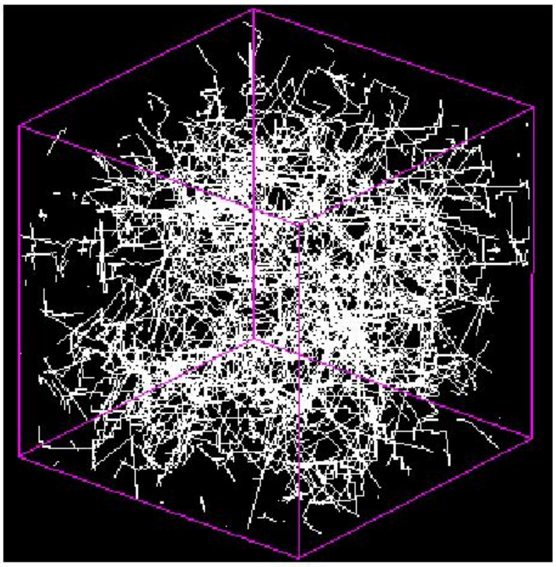


Figure 7.12. Simulations of Discrete Dislocation Dyanmics

We can evaluate the internal stresses acting on the dislocation loop by quadrature integration, i.e.

$$\begin{aligned}\sigma^I_{ij} &= \frac{\mu}{4\pi}\sum_{\gamma=1}^{N_{loop}}\sum_{\beta=1}^{N_s}\sum_{\alpha=1}^{Q_{max}} b_n w_\alpha\Big[\frac{1}{2}R_{,mpp}(\epsilon_{jmn}x_{i,u}+\epsilon_{imn}x_{j,u})\\ &\quad +\frac{1}{1-\nu}\epsilon_{kmn}(R_{,ijm}-\delta_{ij}R_{,ppm})x_{k,u}\Big] \qquad (7.211)\end{aligned}$$

where N_{loop} is the total number of dislocation loops, N_s is the total number of segments in each dislocation loop, and Q_{max} is the total number of quadrature point in a segment, and w_α is the quadrature weight.

Denote each segment of the dislocation loop as L_j. The discretized weak formulation is

$$\begin{aligned}&\sum_{j=1}^{N_s}\sum_{\alpha=1}^{Q_{max}}\sum_{m=1}^{N_{DF}} N_{im}(u)\delta q_m\Big[f_i^T - C_{ik}\sum_{n=1}^{N_{DF}} N_{kn}\dot{q}_n\Big]\\ &\times\Big(\sum_{p,s=1}^{N_{DF}} q_p q_s B_{\ell p}B_{\ell s}\Big)^{1/2} w_\alpha = 0\,. \qquad (7.212)\end{aligned}$$

Define the generalized force vector,

$$f_m^h = \sum_{\alpha=1}^{Q_{max}} f_i^T N_{im}(u) \Big(\sum_{p,s=1}^{N_{DF}} q_p, q_s B_{\ell p} B_{\ell s} \Big)^{1/2} w_\alpha \tag{7.213}$$

and the resistivity matrix $\{\gamma_{mn}\}$, in which

$$\gamma_{mn} = \sum_{\alpha=1}^{Q_{max}} N_{im}(u) C_{ik} N_{kn}(u) \Big(\sum_{p,s=1}^{N_{DF}} q_p, q_s B_{\ell p} B_{\ell s} \Big)^{1/2} w_\alpha \tag{7.214}$$

Then, we can put the dislocation loop weak form into a matrix form,

$$\sum_{j=1}^{N_s} \Big[[\mathbf{f}]_j - [\boldsymbol{\gamma}]_j \Big[\frac{d\mathbf{q}}{dt} \Big]_j \Big]^T \Big[\delta \mathbf{q} \Big]_j = 0 \, , \tag{7.215}$$

which leads to the global matrix formulation,

$$\Big[\Big[\mathbf{F} \Big] - \Big[\boldsymbol{\Gamma} \Big] \Big[\frac{d\mathbf{Q}}{dt} \Big] \Big]^T \Big[\delta \mathbf{Q} \Big] = \mathbf{0} \, , \tag{7.216}$$

where

$$\Big[\mathbf{F} \Big] = \mathbf{A}_{j=1}^{N_s} \Big[\mathbf{f} \Big]_j^{1 \times N_{DF}} \tag{7.217}$$

$$\Big[\boldsymbol{\Gamma} \Big] = \mathbf{A}_{j=1}^{N_s} \Big[\gamma \Big]_j^{N_{DF} \times N_{DF}} \tag{7.218}$$

Sovling (7.216) yields,

$$\Big[\frac{d\mathbf{Q}}{dt} \Big] = \Big[\boldsymbol{\Gamma} \Big]^{-1} \Big[\mathbf{F} \Big] \tag{7.219}$$

Employing any desirable time stepping algorithm, one find the updated dislocation loop configuration or position by

$$\Big[\mathbf{Q} \Big]_{n+1} = \Big[\mathbf{Q} \Big]_n + \Big[\boldsymbol{\Gamma} \Big]_{n+\alpha}^{-1} \Big[\mathbf{F} \Big]_{n+\alpha} \Delta t \tag{7.220}$$

where $0 \leq \alpha \leq 1$.

This is the state of the art discrete dislocation dynamics formulation.

7.7 The Peierls-Nabarro Model

7.7.1 Hilbert transform

The Hilbert transform is a particular case of the Cauchy integral transforms. Let L be a closed smooth contour and $\phi(\zeta)$ be an arbitrary Holder continuous function specified on L and vanishing at infinity. Cauchy integral transforms are the following pair of mutually invertible integrals (e.g. Zhdanov [1984]),

$$\psi(\zeta_0) = \frac{1}{\pi i}\int_L \frac{\phi(\zeta)}{\zeta-\zeta_0}d\zeta \tag{7.221}$$

$$\phi(\zeta_0) = \frac{1}{\pi i}\int_L \frac{\psi(\zeta)}{\zeta-\zeta_0}d\zeta \tag{7.222}$$

One special case of great value for applications is that L real axis, $Im(\psi(\zeta)) = g(x)$, $Re(\psi(\zeta)) = 0$, $Re(\phi(\zeta)) = f(x)$, and $Im(\phi(\zeta)) = 0$. That is $\phi(\zeta) = f(x) + i0$ and $\psi(\zeta) = 0 + ig(x)$. Here $f(x)$ and $g(x)$ are real functions of a real variable x satisfying the Holder condition for any finite x and vanishing at infinity. This special case of Cauchy integral transforms is the so-called *the Hilbert transforms*:

$$g(x) = \mathcal{H}(f(x)) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{f(t)dt}{x-t} \tag{7.223}$$

$$f(x) = -\mathcal{H}(g(x)) = -\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{g(t)dt}{x-t} \tag{7.224}$$

Note the position between x and t and position between ζ and ζ_0.

Hilbert transform table is available in many mathematics handbooks. In general, one can find Hilbert transform via Cauchy's residue theorem.

The following are a few examples:

$$\mathcal{H}\Big(\frac{1}{\pi(b-x)}\Big) = \frac{1}{\pi}\int_{-\infty}^{\infty}\Big(\frac{dt}{\pi(b-t)(x-t)}\Big) = \delta(x-b) \tag{7.225}$$

$$\mathcal{H}\Big(\frac{1}{(x^2+a^2)}\Big) = \frac{1}{\pi}\int_{-\infty}^{\infty}\Big(\frac{dt}{(t^2+a^2)(x-t)}\Big) = \frac{x}{a(x^2+a^2)} \tag{7.226}$$

$$\mathcal{H}\Big(\sin(bx)\Big) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\sin(bt)dt}{(x-t)} = -\cos(bx) \tag{7.227}$$

7.7.2 The Peierls-Nabarro dislocation model

In the early development of dislocation theory, scientists were concerned with two important issues: (1) What is the size of a dislocation for a given Burgers vector? (2)How much force is needed to move a dislocation out of its stable position?

The second question is the so-called dislocation mobility, which is central to the understanding of the ductile material strength. The Peierls-Nabarro dislocation model tries to answer this question.

Before we discuss Peierls-Nabarro model, we first examine the mechanical fields of a straight edge dislocation (displacement fields are given up to a rigid

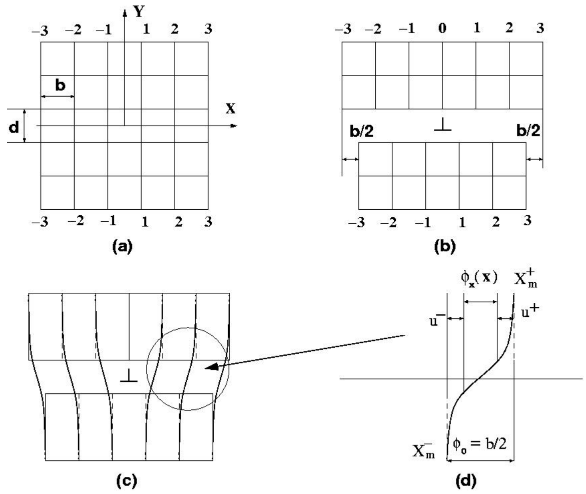


Figure 7.13. The Peierls-Nabarro Model

body displacement) ,

$$u_x = \frac{b}{2\pi}\left[\tan^{-1}\frac{y}{x} + \frac{xy}{2(1-\nu)(x^2+y^2)}\right] \tag{7.228}$$

$$u_y = -\frac{b}{2\pi}\left[\frac{1-2\nu}{4(1-\nu)}ln(x^2+y^2) + \frac{x^2}{2(1-\nu)(x^2+y^2)}\right] \tag{7.229}$$

$$\sigma_{xx} = -\frac{\mu b}{2\pi(1-\nu)}\frac{y(3x^2+y^2)}{(x^2+y^2)^2} \tag{7.230}$$

$$\sigma_{yy} = \frac{\mu b}{2\pi(1-\nu)}\frac{y(x^2-y^2)}{(x^2+y^2)^2} \tag{7.231}$$

$$\sigma_{xy} = \frac{\mu b}{2\pi(1-\nu)}\frac{x(x^2-y^2)}{(x^2+y^2)^2} \tag{7.232}$$

As evident from the above equations, the stress fields are singular at the origin. Therefore the analytical solution presented above is no longer accurate near the core of the dislocation. To remove this singularity inside the dislocation core, Peierls [1940] and Nabarro [1947] included the discrete atomic nature of the material and proposed the following lattice correction model.

The Peierls-Nabarro model(PN model) for a straight edge dislocation is described using two semi-infinite simple cubic crystals as shown in Fig. 5.4. The formal glide plane is $y = 0$. The two elastic half spaces are terminated on the planes $y \geq d/2$ and $y \leq -d/2$. At the middle of glide plane, a non-Hookean slab of width d (atomic spacing) joins the two half spaces. The symmetrical configuration indicated in Fig. 5.4 suggests that this is done by cutting the perfect crystal into two halves along the $y = 0$ plane, and inserting an additional layer of atoms in the upper half of the crystal space, which displaces the upper half crystal moving rigidly a distance $0.5b$ in both positive and negative x-direction, and we then re-weld the two half crystals.

Before the "re-welding", the initial dis-registry (misalignment) in x-direction of two vertical atom layers with respect to the upper and lower half crystal spaces is

$$\phi_x^0(x) := X_m^+ - X_m^- = \begin{cases} \dfrac{b}{2}, & x > 0 \\ \\ -\dfrac{b}{2}, & x < 0 \end{cases} \qquad m = \pm 1, \pm 2, \cdots \pm \infty \tag{7.233}$$

After the re-welding, the misalignment, or the discontinuity, between the atom layer in the upper part of crystal and the same atom layer (m) of the lower part of the crystal becomes

$$\begin{aligned} \phi_x(x) &= x_m^+ - x_m^- = X_m^+ + u^+(x) - (X_m^- + u^-(x)) \\ \phi_x(x) &= \begin{cases} \dfrac{b}{2} + u^+(x) - u^-(x), & x > 0 \\ \\ -\dfrac{b}{2} + u^+(x) - u^-(x), & x < 0 \end{cases} \\ &= \begin{cases} 2u_x(x) + \dfrac{b}{2}, & x > 0 \\ \\ 2u_x(x) - \dfrac{b}{2}, & x < 0 \end{cases} \end{aligned}$$

By antisymmetry, we assume that $u_x(x) = u^+(x) = -u^-(x)$.

At the remote boundary, dis-registry is enforced to be zero, i.e. there is no discontinuity at the remote boundary

$$\phi_x(x) \to 0, \text{ when } x \to \pm\infty \quad \Rightarrow \quad 2u_x(x) \pm \frac{b}{2} = 0, \; x \to \pm\infty \tag{7.234}$$

Therefore, $u_x(\pm\infty) = \mp\dfrac{b}{4}$. This implies that the total displacement along the interface should be

$$u_x(\infty) - u_x(-\infty) = \int_{-\infty}^{\infty} \left(\frac{du_x}{dx}\right)_{x=x'} dx' = -\frac{b}{2} \tag{7.235}$$

Based on Eshelby's interpretation (Eshelby [1949]), one may think that Peierls-Nabarro model deploys a continuous edge dislocation distribution along the cohesive interface with its local Burgers vector density as $b'(x')$ to replace a single dislocation with a Burgers vector $\mathbf{b}$. To make sure that these two dislocation systems are equivalent, we enforce the following condition on net Burgers vector equality,

$$-2\int_{-\infty}^{\infty}\Big(\frac{du_x}{dx}\Big)_{x=x'}dx' = \int_{-\infty}^{\infty} b'(x')dx = b \tag{7.236}$$

From the above relation, one may derive that the distribution density of Burgers vector should be $b'(x') = -2\dfrac{du_x}{dx}(x')$.

The strains near the dislocation core are large, and therefore use of Hooke's law for the stresses is unappropriate. One the other hand, it is relevant to use the periodicity of the lattice, which implies σ_{xy} to be a periodic function of $\phi(x)$. We therefore assume that,

$$\sigma_{xy}(x,0) = C\sin\Big(\frac{2\pi\phi_x}{b}\Big) \tag{7.237}$$

When $\phi_x(x) << 1$,$\sigma_{xy}(x,0) \sim C\dfrac{2\pi\phi_x(x)}{b}$. Under small deformation limit, it is assumed that the cohesive law should comply to Hooke's law as well (is this a good assumption?), i.e.

$$\sigma_{xy}(x,0) = 2\mu\epsilon_{xy} = \frac{\mu\phi_x(x)}{d} = C\frac{2\pi\phi_x(x)}{b} \tag{7.238}$$

which determines the constant $C = \dfrac{\mu b}{2\pi d}$. Note that the shear strain inside the cohesive interface is (see Fig. 5.4)

$$\gamma_{xy} = \frac{\phi_x(x)}{d} \tag{7.239}$$

Thereby, one obtain that

$$\sigma_{xy}(x,0) = \frac{\mu b}{2\pi d}\sin\Big(\pm\pi + \frac{4\pi u_x(x)}{b}\Big) = -\frac{\mu b}{2\pi d}\sin\Big(\frac{4\pi u_x(x)}{b}\Big) \tag{7.240}$$

One can calculate the shear stress inside the cohesive strip due the continuously distributed dislocation via superposition. At $y = 0$,

$$\begin{aligned}\sigma_{xy}(x,0) &= \frac{\mu}{2\pi(1-\nu)}\int_{-\infty}^{\infty}\frac{b'(t)dt}{x-t}\\ &= -\frac{\mu}{\pi(1-\nu)}\int_{-\infty}^{\infty}\frac{(du_x/dx)_{x=t}dt}{x-t}\end{aligned} \tag{7.241}$$

One may also derive the above integral equation based Boussinesq solution of linear elastic half space (e.g. Timoshenko and Goodier [1972]).

Apparently, $\sigma_{xy}(x,0)$ is proportional to the Hilbert transform of du_x/dx. Thereby the inverse Hilbert transform gives

$$\frac{du_x}{dx} = \frac{(1-\nu)}{\mu}\int_{-\infty}^{\infty}\frac{\sigma_{xy}(t,0)dt}{x-t} \tag{7.242}$$

Integrating this yields,

$$u(x) = \frac{(1-\nu)}{\mu}\int_{-\infty}^{\infty}\sigma_{xy}(t,0)\ln|t-x|dt \tag{7.243}$$

Using ((7.240)) and ((7.241)), one can obtain the well-known Peierls-Nabarro integral equation for unknown displacement field, $u_x(x)$,

$$\int_{-\infty}^{\infty}\frac{(du_x/dx)_{x=t}dt}{x-t} = \frac{b(1-\nu)}{2d}\sin\frac{4\pi u_x}{b} \tag{7.244}$$

which is a singular, nonlinear integral equation with unknown function $u_x(x)$.

Luckily, the solution of the above integral equation can be found in closed form [3],

$$u_x(x) = -\frac{b}{2\pi}\tan^{-1}\frac{x}{r_c} \tag{7.245}$$

where $r_c = d/2(1-\nu)$, which is a parameter that characterizes the size of the dislocation core. When $|x| < r_c$, the dis-registry $\phi_x(x) > b/4$. At $x = r_c$, $u_x(r_c) = -b/8$ and $\phi_x(r_c) = b/4$.

Substituting ((7.245)) into ((7.240)) and utilizing the trigonometry identity

$$\tan^{-1}(y) = \sin^{-1}\Big(\frac{y}{\sqrt{1+y^2}}\Big)$$

one can find that

$$\sigma_{xy}(x,0) = \frac{\mu b}{2\pi(1-\nu)}\frac{x}{x^2+r_c^2} \tag{7.246}$$

On the other hand, by virtue of (7.245) the displacement gradient in x-direction is

$$\Big(\frac{du_x}{dx}\Big)_{x=t} = -\frac{b}{2\pi}\frac{r_c}{t^2+r_c^2} \tag{7.247}$$

and the Hilbert transform of the above expression is

$$\mathcal{H}\Big(\frac{du_x}{dx}\Big) = \mathcal{H}\Big(-\frac{br_c}{2\pi}\frac{1}{x^2+r_c^2}\Big) = -\frac{b}{2\pi}\frac{x}{x^2+r_c^2} \tag{7.248}$$

[3] My guess is that the reason why they took sine function as the cohesive law was to match the exact solution of this particular integral equation, which people had known before.

where the following Hilbert transform formula is used,

$$\mathcal{H}\Big(\frac{1}{x^2+r_c^2}\Big) = \frac{1}{r_c}\frac{x}{x^2+r_c^2}$$

Based on ((7.241)),

$$\sigma_{xy}(x,0) = -\frac{\mu}{(1-\nu)}\mathcal{H}\Big(\frac{du_x}{dx}\Big) = \frac{\mu b}{2\pi(1-\nu)}\frac{x}{x^2+r_c^2} \tag{7.249}$$

which is the same as the expression obtained above.

7.7.3 Misfit Energy and the Peierls Force

As we mentioned before, one of the motives to discuss the Peierls-Nabarro dislocation model is to find the critical stress needed in order to move a dislocation from its stable position. This question can not be answered by analyzing a Volterra dislocation.

To find the critical stress to move a dislocation, we first examine the stored elastic energy due to an edge dislocation. The total elastic energy stored induced by an edge dislocation may be divided into two parts: the energy stored inside the elastic crystal and the energy stored inside the cohesive layer. Since the two crystal half spaces maintain substantially perfect lattice structure, most of shear deformation is confined within the cohesive layer. For this reason, we call the energy stored inside the cohesive layer as the misfit energy.

The shear strain, in fact that it is the eigen shear strain because it is the "shear strain" caused by the local jump, inside the cohesive zone is,

$$\gamma_{xy} = \frac{\phi_x(x)}{d} = \frac{2u_x(x)+(b/2)}{d}, \ \ x>0 \tag{7.250}$$

The misfit energy for a pair of atomic planes is,

$$\begin{aligned}\Delta W &= -\frac{1}{2}\int_0^{\gamma_{xy}} \sigma'_{xy}(x,0)d\gamma'_{xy} b\cdot d\\ &= \int_{-b/4}^{u_x} \sigma_{xy} du_x b\cdot d \end{aligned} \tag{7.251}$$

The factor of half is introduced in calculating the misfit energy because it is getting shared between two planes. Note that when $u(x) = -b/4 \rightarrow \gamma_{xy} = 0$. Therefore,

$$\begin{aligned}\Delta W(x) &= \frac{\mu b^2}{2\pi d}\int_{-b/4}^{u_x}\sin\Big(\frac{4\pi u_x}{b}\Big)du_x = \frac{\mu b^3}{8\pi^2 d}\cos\Big(\frac{4\pi u_x}{b}\Big)\Big|_{-b/4}^{u_x}\\ &= \frac{\mu b^3}{8\pi^2 d}\Big(1+\cos\Big(\frac{4\pi u_x}{b}\Big)\Big)\end{aligned} \tag{7.252}$$

Substitute,

$$u_x = -\frac{b}{2\pi}\tan^{-1}\Big(\frac{x}{r_c}\Big) \tag{7.253}$$

to obtain the misfit energy for a pair of atomic planes as,

$$\Delta W = \frac{\mu b^3}{8\pi^2 d}\Big(1+\cos\Big(2\tan^{-1}(\frac{x}{r_c})\Big)\Big) \tag{7.254}$$

Let the distance of the center of the dislocation from the nearest position of symmetry be $\xi = \alpha b$, where α is a variable. Then the position of all the atoms, on the two faces of the slip plane are defined by

$$x_m = \begin{cases} 2m\dfrac{b}{2} & \text{the upper half crystal} \\ (2m-1)\dfrac{b}{2} & \text{the lower half crystal} \end{cases} \tag{7.255}$$

and $m = 0, \pm1, \pm2, \pm3, \cdots$ (see Fig. 7.14).

Then the total misfit energy is the summation,

$$\begin{aligned} W &= \sum_{m=-\infty}^{\infty} \Delta W(2m) + \Delta W(2m-1) \\ &= \sum_{n=0,\pm2,\pm4}^{\cdots} \frac{\mu b^3}{8\pi^2 d}\sum_{n=-\infty}^{+\infty}\Big(1+\cos\Big(2\tan^{-1}(\alpha+0.5n)(\frac{b}{r_c})\Big)\Big) \\ &+ \sum_{n=\pm1,\pm3}^{\cdots} \frac{\mu b^3}{8\pi^2 d}\sum_{n=-\infty}^{+\infty}\Big(1+\cos\Big(2\tan^{-1}(\alpha+0.5n)(\frac{b}{r_c})\Big)\Big) \end{aligned} \tag{7.256}$$

which can be combined into a single expression, i.e. $x = (\alpha + 0.5n)b$ and $n = 0, \pm1, \pm2, ...$ Therefore summing up over all the atomic planes we get the total misfit energy as

$$W = \sum_{n=-\infty}^{+\infty} f(n) = \frac{\mu b^3}{8\pi^2 d}\sum_{n=-\infty}^{+\infty}\Big(1+\cos\Big(2\tan^{-1}(\alpha+0.5n)(\frac{b}{r_c})\Big)\Big) \tag{7.257}$$

This may be transformed using the Poission's summation formula in Harmonic analysis:

$$\sum_{n=-\infty}^{+\infty} f(n) = \sum_{n=-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x)\exp(-i2\pi xn)dx, \tag{7.258}$$

where $f(x)$ is an even function, it reads

$$\sum_{n=-\infty}^{+\infty} f(n) = \sum_{n=-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x)cos(2\pi xn)dx, \tag{7.259}$$

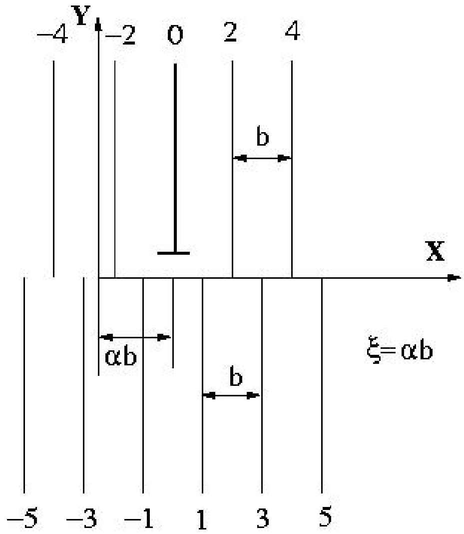


Figure 7.14. The Nabarro counting scheme

where we have used the fact that the function $f(n)$ is even in n. We can rewrite the above relation as,

$$\sum_{n=-\infty}^{+\infty} f(n) = \int_{-\infty}^{+\infty} f(x)dx + 2\sum_{n=1}^{\infty}\int_{-\infty}^{+\infty} f(x)cos(2\pi xn)dx, \tag{7.260}$$

Therefore we can rewrite the total misfit energy from the equation ((7.257)) as,

$$\begin{aligned} W &= \frac{\mu b^3}{8\pi^2 d}\int_{-\infty}^{+\infty}(1+\cos(2\tan^{-1} z))dx \\ &\quad + \frac{\mu b^3}{4\pi^2 d}\sum_{n=1}^{+\infty}\int_{-\infty}^{+\infty}(1+\cos(2\tan^{-1} z))cos\Big(2\pi n\Big(\frac{dz}{(1-\nu)b} - 2\alpha\Big)\Big)dx \end{aligned} \tag{7.261}$$

where $z = (\alpha + \frac{x}{2})\frac{b}{r_c} = 2(1-\nu)(\alpha + \frac{x}{2})\frac{b}{d}$. Therefore $\frac{dz}{dx} = (1-\nu)\frac{b}{d}$ and $dx = \frac{d}{(1-\nu)b}dz$. Using these transformations and that $\cos(2\tan^{-1} z) =$

$\frac{2}{1+z^2} - 1$, we get,

$$
\begin{aligned}
W &= \frac{\mu b^2}{4\pi^2(1-\nu)} \int_{-\infty}^{+\infty} \frac{1}{1+z^2} dz \\
&\quad + \frac{\mu b^2}{2\pi^2(1-\nu)} \sum_{n=1}^{+\infty} \int_{-\infty}^{+\infty} \cos\Big(2\pi n\Big(\frac{dz}{(1-\nu)b} - 2\alpha\Big)\Big) \frac{dz}{1+z^2}
\end{aligned}
\tag{7.262}
$$

The first integral above can be calculated using the Cauchy residual theorem, that is we use the result:

$$
\int_{-\infty}^{+\infty} \frac{1}{1+z^2} dz = 2\pi i Re(\frac{1}{1+z^2}) = \pi
$$

where $Re(.)$ denotes the residual. Therefore the first term of the total misfit energy as $\frac{\mu b^2}{4\pi(1-\nu)}$. The second term in equation ((7.262)) can be further reduced to,

$$
\frac{\mu b^2}{2\pi^2(1-\nu)} \sum_{n=1}^{+\infty} \cos(4\pi n\alpha) \int_{-\infty}^{+\infty} \cos\Big(\frac{2\pi nzd}{(1-\nu)b}\Big) \frac{dz}{1+z^2}
$$

To evaluate this term we again use Cauchy residual theorem. Say $k = \frac{2\pi nd}{(1-\nu)b}$; then the integral in the above equation is equal to,

$$
\int_{-\infty}^{+\infty} \frac{e^{ikz}}{1+z^2} dz
$$

which is equal to πe^{-k}. Therefore we obtain the total misfit energy as,

$$
W = \frac{\mu b^2}{4\pi(1-\nu)} + \frac{\mu b^2}{2\pi^2(1-\nu)} \sum_{n=1}^{+\infty} \pi e^{\frac{-4\pi r_c n}{b}} \cos(4\pi n\alpha) \tag{7.263}
$$

The term in $n = 1$ dominates the sum, therefore we have,

$$
W(\alpha) = \frac{\mu b^2}{4\pi(1-\nu)} + \frac{\mu b^2}{2\pi(1-\nu)} \exp\Big(-\frac{4\pi r_c}{b}\Big) \cos 4\pi\alpha \tag{7.264}
$$

The corresponding force acting on dislocation is given by,

$$
F = -\frac{1}{b}\frac{dW(\alpha)}{d\alpha} \tag{7.265}
$$

Note that the dislocation moves a distance $-\alpha b$.

$dW(\alpha)/d\alpha$ reaches to maximum when $\sin 4\pi\alpha = 1$. From the relation that $\sigma_{xy} = F(b \times 1)$(unit thickness in z-direction), the critical shear stress to move the dislocation by one lattice site is

$$\sigma = \frac{2\mu}{(1-\nu)} \exp\Big(-\frac{4\pi r_c}{b}\Big) \tag{7.266}$$

where F is called the Peierls force and σ is called the Peierls stress, which are required to move a dislocation over a Peierls barrier.

A more physically realistic restoring stress is obtained if we use relative displacement (of the two half planes) instead of the lattice displacement in the above discussion. In the following, a more recent treatment of the PN model is outlined (Joós and Duesbery, 1997) which considers the relative displacement instead of the independent lattice displacements in two half planes. We restrict our attention to the case of a straight edge dislocation. The new model predicts a Peierls stress which differs from the above mentioned expression by a factor of two in both the exponential and the coefficient of the exponential. This approach is also valid for the case of narrow dislocations. By $f(x)$ we define the displacement of the upper half of the crystal with respect to the lower half. If c is a constant, then $f(x-c)$ corresponds to a dislocation translated by c. For a discrete lattice this can be understood like this: If the dislocation is introduced at c, then the atomic planes at a position mb in the upper half of the crystal will experience a displacement of $f(mb-c)$ along the Burgers vector. The total misfit energy in this case can be written as:

$$W(c) = \frac{\mu b^3}{4\pi^2 d} \sum_{m=-\infty}^{+\infty} \Big(1 + \cos\Big(2\tan^{-1}(\frac{mb-c}{r_c})\Big)\Big) \tag{7.267}$$

Note the difference of factor of half in the expression of W from the earlier discussion. This is because we are no longer treating the two half planes independently, but we are using a relative displacement. Using further manipulations and substituting $\Gamma = r_c/b$ and $y = c/b$ we have,

$$W(y) = \frac{\mu b^2}{4\pi^2(1-\nu)} \sum_{m=-\infty}^{+\infty} \frac{\Gamma}{\Gamma^2 + (m-y)^2} \tag{7.268}$$

$W(y)$ is an even periodic function of period 1. Using this information we can express the energy as the sum,

$$W(y) = \frac{a_0}{2} + \sum_{n=1}^{+\infty} a_n \cos 2\pi n y \tag{7.269}$$

Where we can calculate the Fourier coefficients in the usual manner. After substituting the value of these Fourier coefficients, we get the expression for the total misfit energy as,

$$W(y) = \frac{\mu b^2}{4\pi(1-\nu)} + \frac{\mu b^2}{2\pi(1-\nu)} \sum_{n=1}^{+\infty} e^{-2\pi n\Gamma} \cos 2\pi n y \tag{7.270}$$

For the limit of wide dislocations ($\Gamma \gg 1$), only the first exponential term is kept. Then in the limit of wide dislocations we have,

$$W(c) = \frac{\mu b^2}{4\pi(1-\nu)} \Big(1 + 2e^{\frac{-2\pi r_c}{b}} \cos \frac{2\pi c}{b}\Big) \tag{7.271}$$

From which we obtain, (using the relation $\sigma = \max\Big\{\frac{1}{b}\frac{dW}{dc}\Big\}$)

$$\sigma = \frac{\mu}{(1-\nu)} \exp\Big(-\frac{2\pi r_c}{b}\Big) \tag{7.272}$$

Note the difference between the above stress and the one obtained in the equation ((7.266)).

Figure 7.15. Paul Dirac (left), Wolfgang Pauli (middle) and Rudolf Peierls (right) in discussion at the international Conference on Nuclear Physics, Birmingham, 1948

7.7.4 Story of the Peierls-Nabarro Model

The following is an account on the discovery of Peierls-Nabarro model, which was given by the late Professor, Egon Orowan, of Massachusetts Institute of Technology, who was a well-known physicist and material scientist at the time.

"1937 I was invited to work at the University of Birmingham, in the Physics Department which had just taken over by M. L. E. Oliphant (now Sir Mark Oliphant). I felt that it would be urgent to know the width of the dislocation belt and the stress required to move it. The simplest assumption about this was the one made by Taylor, that the stress was zero; however, the extremely high yield stress of many hard materials such as diamond (which could be remarkably free from imperfections and thus could not contain too many dislocations) indicated that the most frequent cause of the hardness of crystalline materials was the high shear stress required to move a dislocation. I found that the width of the dislocation and the stress for moving it could be calculated, with a crude approximation, simply enough by assuming that the shearing force between the opposite shores of the slip plane in a dislocation was a sine function of the relative shear displacement (the initial tangent of the sine, of course, was given by the elastic modulus).

One the other hand, displacement and shear traction at the surface of a half-space were connected by the equations of Boussinesq; equating the stresses and displacements of the sine approximation with those of Boussinesq led to an integral equation which was the solution of the problem It would have taken me days or weeks of study to solve it; fortunately I was a daily guest in the hospitable house of the brilliamnt theoretical physicist Rudolf Peierls. He solved the equation, if I remember well, within a few hours, and he also drove me to a conference at Bristol University in 1939 where I gave a paper and he gave another on the problem he had just solved.

The calculation of the width of the dislocaiton and of the Peierls-Nabarro stress required for moving it was repeated and improved by Nabarro in 1947. The result was puzzling at first: the width calculatied by Nabarro amounted to a few atomic spacings while Peierls obtains an order of magnitude of thousands of spacings. After some research in Birmingham and in Cambridge (where I was wat the time) I discovered the sheet with Peierls's calculations in my desk; Peierls checked it and found that a factor of 2π was accidentally omitted in an exponent, which amounted to a factor of about 1000 in the result.

Of course, the calculation with the sinusoidal approximation is useless in most interesting cases of directinal bonds, in transition metals and the hard non-metallic crystals."

From *The Sorby Centennial Symposium on the History of Metallugy*, MSC, Vol. 27, 1963, pages 368-369.

7.8 Dislocations in the epitaxial thin film

The thin film is the basic configuration structure for integrated circuits, computer memories (RAM), and various sensors, filters, and other electronic devices. Study the mechanical, chemical, and electrical properties of the thin films has particular significance for nano-technologies.

The ancient Greek word $\epsilon\pi\iota$ (*epi*–placed or resting upon) and the word $\tau\alpha\xi\iota\zeta$ (*taxis* – arrangement) are the root of the modern word *epitaxy*, which describes an extremely important phenomenon exhibited by thin films. Epitaxy refers to a single-crystal film formation on top of a crystalline substrate and both have the exactly the same crystal structure as the thin film. 90 % of thin films used in semi-conductor and computer industry, communication industry, and sensor and information industry are epitaxial thin films. To grow various

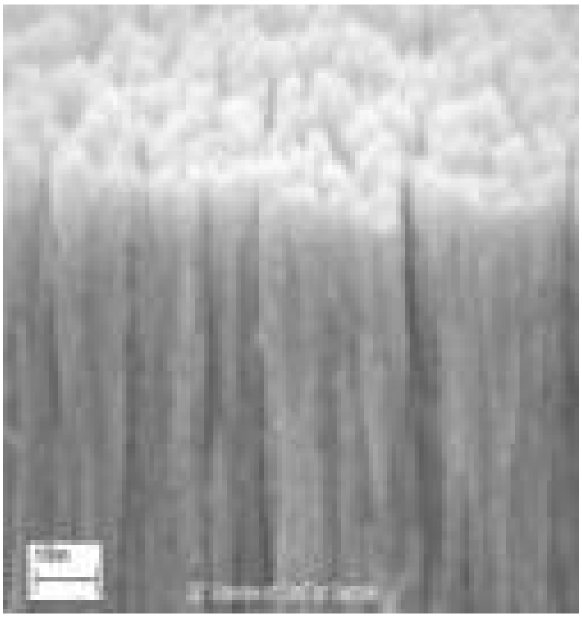


Figure 7.16. An epitaxial thin film.

defect-free epitaxial thin films has been the main challenge in semi-conductor industry in the past half century.

In this section, we shall introduce the two basic dislocation models in thin-film mechanics.

7.8.1 Frenkel & Kontorova model and Frank & van der Merwe model

The Frenkel & Kontorova dislocation model is a one-dimensional dislocation model, which was proposed in 1937. This model was studied in detailed by Frank and van der Merwe [1950ab], and they applied it to study thin film mechanics or epitaxial thin film mechanics.

In Frenkel & Kontorova model, the thin film is modeled as one dimensional monolayer with lattice spacing a_f, and the substrate is modeled as large slab with lattice spacing a_s, and $a_s \neq a_f$ and the lattice misfit is $\Delta = a_f - a_s$ (see Fig. 7.17).

The row of atoms in the thin film are under combined influence of harmonic forces between the nearest neighbours in the monolayer and non-linear interaction forces from substrate. Since the substrate is assumed much larger in dimension than the thin film, it is assumed to be rigid. The interaction between the thin film and substrate, or the force exerted on the thin film by the sub-

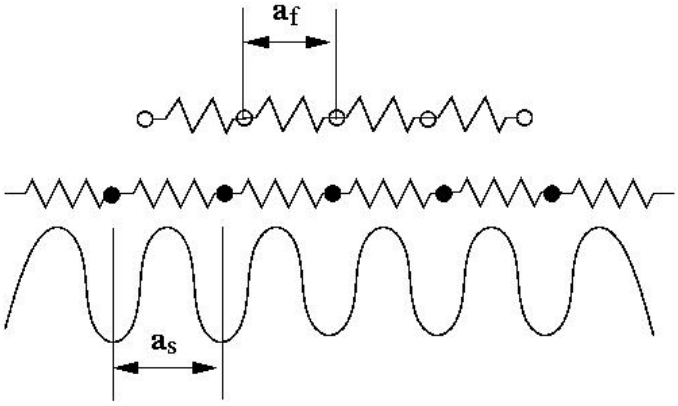


Figure 7.17. Frank-van der Merwe dislocation thin film model

strate is characterized by a sinusoidal potential with the amplitude $\frac{1}{2}W$ (see Fig. 7.17).

Make the position (the open circle in Fig. 7.17) of the m-th atoms in the un-strained monolayer as

$$X_m = ma_f, \quad m = 0, \pm 1, \pm 2, \cdots \tag{7.273}$$

After attach the thin film onto the substrate, the thin film will be stretched to the position

$$x_m^i = ma_s = X_m + u_m^{mis}, \;\; m = 0, \pm 1, \pm 2, \cdots \; . \tag{7.274}$$

where x_m^r is denoted as the reference position of the m-th atom with respect to the aubstrate, and u_m^{mis} is the displacement of the atom due to the lattice misfit,

$$u_m^{mis} = m(a_s - a_f) \; .$$

During actual deformation, the spatial position the m-th atom is

$$x_m = X_m + u_m^{mis} + u_m^e \tag{7.275}$$

or

$$u_m = x_m - x_m = u_m^{mis} + u_m^e \tag{7.276}$$

where u_m^e is the elastic deformation of the atom.

The relative displacement between the two atoms is now

$$u_{m+1} - u_m = (u_{m+1}^e - u_m^e) - (a_f - a_s) \; . \tag{7.277}$$

The total potential energy of the system is

$$\Pi = \frac{1}{2}\sum_m \left\{ \mu(u^e_{m+1} - u^e_m - (a_f - a_s))^2 + W[1 - \cos\frac{2\pi u^e_m}{a}] \right\} \quad (7.278)$$

Let,

$$\xi_m = \frac{u^e_m}{a_s}, \ \text{ and } \ f = \frac{a_f - a_s}{a_s} \quad . \quad (7.279)$$

Hence

$$\Pi = \frac{1}{2}\sum_m \left\{ \mu a^2(\xi_{m+1} - \xi_m - f)^2 + W[1 - \cos(2\pi\zeta_m)] \right\} \quad (7.280)$$

The equilibrium equation is derived from the stationary condition

$$\begin{aligned} &\frac{dE}{d\xi_n} = 0, \ \ n = 0, \pm 1, \pm 2, \cdots \ \Rightarrow \\ &-\mu a^2(\xi_{n+1} - \xi_n + f) + \mu a^2(\xi_n - \xi_{n-1} + f) + W\pi \sin 2\pi\xi_n = 0, \end{aligned} \quad (7.281)$$

i.e.

$$\Delta^2_n \xi = (\xi_{n+1} - 2\xi_n + \xi_{n-1}) = \frac{\pi}{2\ell_0^2}\sin 2\pi\xi_n \quad (7.282)$$

where $\ell_0 = \sqrt{\mu a^2/2W}$.

The dynamics version of Eq. (7.282) is the finite-difference sine-Gordon equation,

$$\Delta^2_n \xi - \frac{m_n}{\mu}\frac{d^2\xi_n}{dt^2} = \frac{\pi}{2\ell_0^2}\sin 2\pi\xi_n \quad (7.283)$$

If $\ell_0 >> 1$, one may use continuous approximation to replace the finite difference equation with a differential equation,

$$\Delta^2_n \xi = \frac{d^2\xi_n}{dX_n^2}a_f^2 + \frac{2}{4!}\frac{d^4\xi}{dX_n^4}a_f^4 + \mathcal{O}(a_f^6) = \frac{d^2\xi}{dn^2} + \mathcal{O}(a_f^4) \quad (7.284)$$

Therefore, if we only consider static deformation, we have the following nonlinear ordinary differential equation

$$\frac{d^2\xi}{dn^2} = \frac{\pi}{2\ell_0^2}\sin 2\pi\xi \ . \quad (7.285)$$

Consider the following boundary conditions,

$$\frac{d\xi}{dn}\Big|_{n=n_0} = \epsilon, \ \text{ and } \ \xi\Big|_{n=n_0} = 0 \ . \quad (7.286)$$

One can integrate (7.282),

$$\left(\frac{d\xi}{dn}\right)^2 - \epsilon^2 = \frac{1}{2\ell_0^2}(1-\cos 2\pi\xi)\ , \tag{7.287}$$

which can be re-arranged as

$$\left(\frac{d\xi}{dn}\right)^2 = \frac{(1+\ell_0^2\epsilon^2}{\ell_0^2}\left(1-\frac{\cos^2\pi\xi}{1+\ell_0^2\epsilon^2}\right) \tag{7.288}$$

Change variable

$$\phi = \pi\zeta - \frac{\pi}{2}\ \ \text{and}\ \ k = (1+\ell_0^2\epsilon^2)^{-1/2}\ . \tag{7.289}$$

One may transfer into the standard form of differential equations that can be solved by using elliptic functions and integrals,

$$\frac{d\phi}{dn} = \pm\left(\frac{\pi}{\ell_0 k}\right)(1-k^2\sin^2\phi)^{1/2} \tag{7.290}$$

Solutions of FKV model:
1. Consider boundary condition

$$\epsilon = 0,\ \ \text{and}\ \ k = 1. \tag{7.291}$$

In this case, Eq. (7.288) is simplified to

$$\frac{d\xi}{dn} = \frac{1}{\ell_0}\sin\pi\xi \tag{7.292}$$

Assume at $n = 0$, $\xi(0) = 0.5$, and then

$$\frac{\pi}{\ell_0}\int_0^n dp = \pi\int_0^\xi \frac{d\zeta}{\sin\pi\zeta} \tag{7.293}$$

which yields the solution

$$\frac{\pi n}{\ell_0} = \ln\tan\left(\frac{\pi\xi}{2}\right) \tag{7.294}$$

Or inversely,

$$\xi = \frac{2}{\pi}\tan^{-1}\left[\exp\left(\frac{\pi n}{\ell_0}\right)\right] \tag{7.295}$$

This solution represents a single dislocation far away from the remote boundary. We plot the positive solution in Fig. 7.18. One may find that at $\xi = 1/2$,

$$\frac{d\xi}{dn} = \frac{1}{\ell_0} \tag{7.296}$$

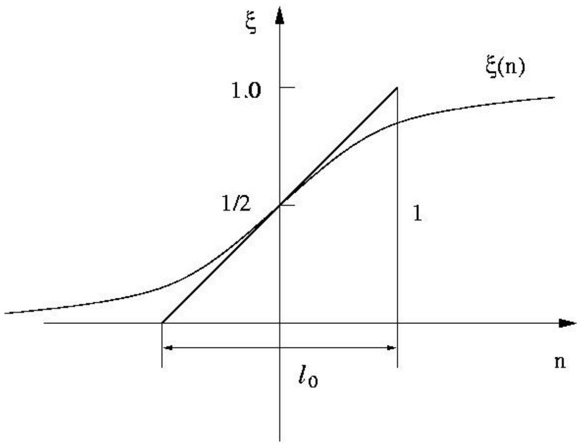


Figure 7.18. A single dislocation solution of FKV model

Since a unit change of ξ means a relative displacement of one lattice spacing a_s, it then implies that in a region of length ℓ_0 number of troughs is one more than the number of atoms, i.e. there is extra plane of atoms in the substrate, which forms a edge dislocation. We call ℓ_0 as the effective length of the dislocation region.

2. General solution

The general static solution of sine-Gordon equation can be expressed by elliptic function,

$$\left(\frac{\pi}{\ell_0 k}\right) = \int_0^{\phi} (1 - k^2 \sin^2 \psi)^{-1/2} d\psi = F(\phi, k) \tag{7.297}$$

where the upper limit ϕ is called the amplitude. The inverse relation of the above elliptic function is

$$\phi = am\left(\frac{\pi n}{\ell_0 k}\right) \tag{7.298}$$

or

$$\xi = \frac{1}{2} + \frac{1}{\pi} am\left(\frac{\pi n}{\ell_0 k}\right) \tag{7.299}$$

and

$$\frac{d\xi}{dn} = \frac{1}{\ell_0 k} dn\left(\frac{\pi n}{\ell_0 k}\right) = \frac{1}{\ell_0 k}(1 - k^2 \cos^2 \pi\xi)^{1/2} \tag{7.300}$$

At $\xi = \xi(0) = 1/2$,

$$\frac{d\xi}{dn} = \frac{1}{\ell_0 k} \tag{7.301}$$

i.e. $\ell_0 k$ is now the effective dislocation length.

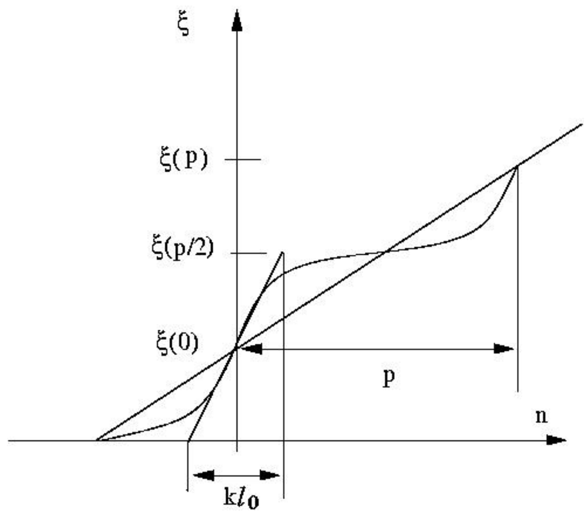


Figure 7.19. The general solution of static sine-Gordon equation ($\frac{d\xi}{dn} \geq 0$).

Assume that $\xi(p) = 1.5$. The general solution of FKV model is depicted on Fig. 7.19. Obviously, the number is the atoms per dislocation,

$$p = \frac{2\ell_0 k E(k)}{\pi} \tag{7.302}$$

where $E(k)$ is the following elliptic integral,

$$E(k) = \int_0^{\pi/2} (1 - k^2 \sin^2 \psi)^{1/2} d\psi \tag{7.303}$$

The general solution indicates that there are many dislocation occuring simutanelously along the chain in periodic fashion. In Fig. 7.20, we show the dislocation pattern created by the general solution.

It would be interesting to examin the stability of Frenkel-Kontorova system. The potential energy of one dislocation

$$\begin{aligned} \Pi &= W\ell_0^2 \sum_{n=0}^{p-1} (\xi_{n+1} - \xi_n - f)^2 + \frac{W}{2} \sum_{n=0}^{p-1} \Big(1 - \cos 2\pi \xi_n) \\ &= W\ell_0^2 \int_0^P \Big(\frac{d\xi}{dn} - f\Big)^2 dn + W \int_0^P \sin^2 \pi\xi dn \end{aligned} \tag{7.304}$$

Consider

$$\frac{d\xi}{dn} = \frac{1}{\ell_0^2 k^2} (1 - k^2 \cos^2 \pi\xi)^{1/2} \tag{7.305}$$

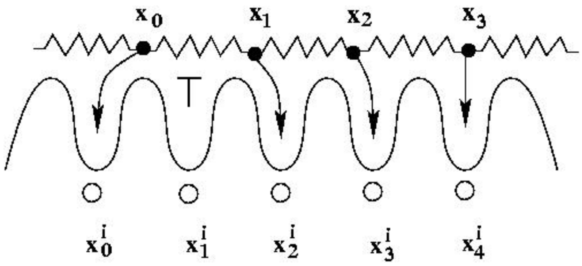


Figure 7.20. Dislocation pattern for $p = 3$.

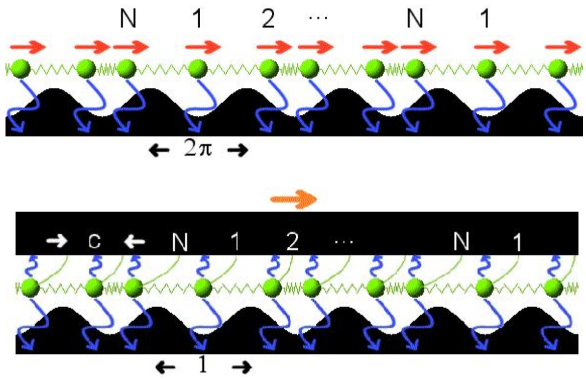


Figure 7.21. Dislocation pattern for $\frac{d\xi}{dn} \leq 0$.

One can write the potential energy per dislocation as

$$\Pi = W\ell_0^2 \left\{ \frac{4E(k)}{\pi k \ell_0} - \frac{2(1-k^2)K(k)}{\pi k \ell_0} - 2f + pf^2 \right\} \tag{7.306}$$

where

$$K(k) = \int_0^{\pi/2} (1 - k^2 \sin^2 \psi)^{-1/2} d\psi$$

One may find that the potential energy consists of contribution from both lattice misfit and dislocation misfit.

To examine the stability, let,

$$\frac{\partial \Pi}{\partial f} = W\ell_0^2 (2 - 2pf) = 0\,. \tag{7.307}$$

We find the critical lattice misfit,

$$f_{cr} = \frac{1}{p} = \frac{\pi}{2\ell_0 k K(k)} \tag{7.308}$$

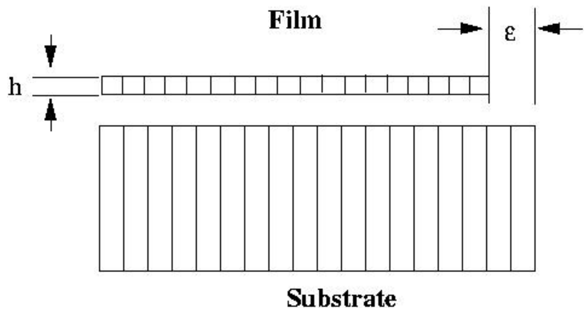


Figure 7.22. Matthews & Blackeslee Model

When $k = 1$,

$$f_{cr} = \frac{1}{p} = \frac{2}{\pi}\Big(\frac{W}{\mu a^2/2}\Big)^{1/2} \tag{7.309}$$

It is beleived that when lattice misfit $f > f_{cr}$, dislocations will spontaneouslly enter or depart from the monolayer chain.

7.8.2 Matthews & Blackesless's equilibrium theory

In 1974, Matthews and Blackeslee proposed their equilibrium theory of dislocation relaxation mechanism for thin film growth. It was an immediate success, and it was soon received widespread attentions. Today, the Matthews theory has become the foundamental theory for epitaxial thin film growth in semi-conductor industry, and it is now viewed an early and integrated part of nano-mechanics.

In the following, we outlined a simple version of the Matthews theory based on Nix's presentation.

Assume that the thin film is under homogeneous bi-axial palne stress load, i.e. in the film, $\epsilon_x = \epsilon_y = \epsilon$ and $\sigma_x = \sigma_y = \dfrac{E}{1-\nu}\epsilon$. The homogeneous misfit strain is due to the lattice misfit, i.e.

$$\epsilon = \frac{a_s - a_f}{a_f} \ \text{ or } \ \epsilon = \frac{a_s - a_f}{a_s} \ . \tag{7.310}$$

The deformation of the substrate may be neglected. For a coherent thin film-substrate system, the strain energy per unit thin film area is (see Fig. 7.22)

$$E = \frac{2\mu(1+\nu)}{(1-\nu)}\epsilon^2 h = M\epsilon^2 h \ . \tag{7.311}$$

When the lattice misfit ϵ increases, it is energetically favorable to have dislocations present to relaxe the lattice misfit strain.

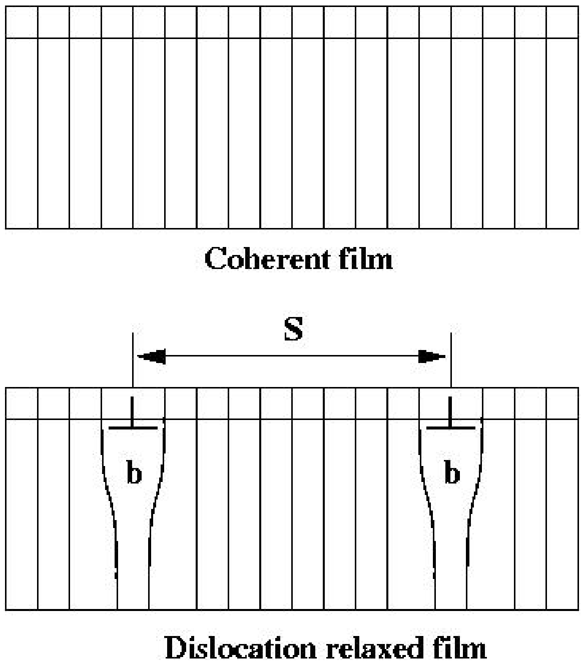


Figure 7.23. Matthews & Blackeslee Model

Consider a simplist sceenario that there is periodically distributed edge dislocations distributed along the interface between the thin film and the substrate. The homogeneous distributed lattice misfit strain will be reduced to $f - b/S$ where S is the spacing between two edges dislocations. Then the elastic energy due to homogeneous deformation is

$$E_h = M\Big(\epsilon - \frac{b}{S}\Big)^2 h \tag{7.312}$$

Since there are two edge dislocations in an area $S \times 1$, the strain energy due to dislocation is

$$E_d = \frac{\mu b^2}{4\pi(1-\nu)} \ln\Big(\frac{\beta h}{b}\Big)\frac{2}{S} \tag{7.313}$$

The total energy is the summation of E_h and E_d,

$$E = M\Big(\epsilon - \frac{b}{S}\Big)^2 h + \frac{\mu b^2}{4\pi(1-\nu)} \ln\Big(\frac{\beta h}{b}\Big)\frac{2}{S} \tag{7.314}$$

The two competing effects will yield an equilibrium point at the bottom of energy well as shown in Fig. 7.24. We are seeking to find an equilibrium

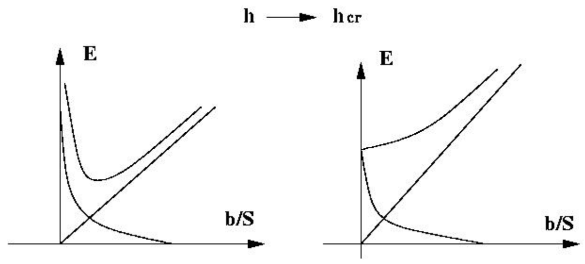


Figure 7.24. Matthews & Blackeslee Model

state that is defect-free, i.e. we are intereted in an equilibrium state at which $b/S = 0$.

Consider the stationary condition,

$$\frac{\partial E}{\partial \frac{1}{S}} = -2Mh\Big(\epsilon - \frac{b}{S}\Big)b + \frac{\mu b^2}{2\pi(1-\nu)} \ln\Big(\frac{\beta h}{b}\Big) \Big|_{h=h_{cr}} = 0 \,. \tag{7.315}$$

We can find a critical thickness, h_{cr}, of the thin film below which the thin film will stay in a coherent state with the substrate that is the thin film is defect-free.

From (7.315), one can find that the critical thickness can be determined from the following non-linear equation,

$$\frac{h_{cr}}{\ln\Big(\dfrac{\beta h_{cr}}{b}\Big)} = \frac{\mu b}{4\pi(1-\nu)M\epsilon} \tag{7.316}$$

Exercise

PROBELM 7.1 *Consider cuboidal region of inelastic strain (eigenstrain) due to solute segregation forming cuboidal precipitates. The precipitate subdomain (or inclusion) has the dimension $2a \times 2a \times 2a$, and the unit cell (U) has the dimension $2L \times 2 : \times 2L$. The eigenstrain is assumed to have a constant value ϵ within each inclusion, and be zero outside the inclusion,*

$$\varepsilon^*_{ij} = \begin{cases} \delta_{ij}\varepsilon, & \forall \ \mathbf{x} \in \Omega; \\ 0; & \forall \ \mathbf{x} \in U/\Omega \end{cases} \tag{7.317}$$

where

$$U = \Big\{\mathbf{x}\Big| - L \le x_i \le L, \ i = 1,2,3\Big\} \tag{7.318}$$

$$\Omega = \Big\{\mathbf{x}\Big| - a \le x_i \le a, \ i = 1,2,3\Big\}, \text{ and } a < L \tag{7.319}$$

Find the disturbed displacement field $u_1(\mathbf{x})$. (Hint:Mura pages: 20-21).

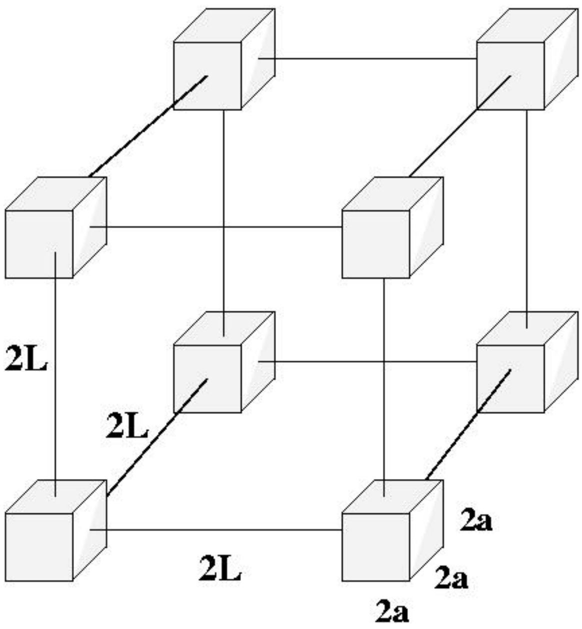


Figure 7.25. Distribution of periodic precipitates

Chapter 8

COMPARISON VARIATIONAL PRINCIPLES

8.1 Review of Variational Calculus

Consider a functional, which is a map,

$$I[y] : H^1([x_0, x_1]) \to \mathbb{R} \tag{8.1}$$

where $I[y]$ is the following integral a map

$$I[y] = \int_{x_0}^{x_1} \Big[p(x)(y^{'})^2 + q(x)y^2 + 2f(x)y \Big] dx \tag{8.2}$$

with prescribed boundary conditions,

$$y(x_0) = y_0, \quad y(x_1) = y_1 \tag{8.3}$$

Assume that $p(x), q(x)$, and $f(x)$ are given continuous functions, i.e. $p(x), q(x)$, and $f(x) \in C^0[x_0, x_1]$, and $p(x) > 0, q(x) > 0$. Let,

$$\tilde{y}(x) = y(x) + \alpha\eta(x) \tag{8.4}$$

as a function that is very close to function, $y(x)$.

We require that $y(x) \in \mathcal{V}$ and $\eta(x) \in \overset{\circ}{\mathcal{V}}$, and

$$\mathcal{V} := \Big\{ y(x) \;\Big|\; y \in H^1([x_0, x_1]), y(x_0) = y_0 \;\text{ and }\; y(x_1) = y_1 \Big\} \tag{8.5}$$

$$\overset{\circ}{\mathcal{V}} := \Big\{ eta(x) \;\Big|\; \eta \in H^1([x_0, x_1]), \eta(x_0) = 0 \;\text{ and }\; \eta(x_1) = 0 \Big\} \tag{8.6}$$

We usually call y as the trial function and $\alpha\eta(x)$ as the test function.

In order to find the function $y(x)$ that yields the extreme value of $I[y]$, we consider the value of $I[\tilde{y}]$,

$$\begin{aligned}
I[y(x)+\alpha\eta(x)] &= \int_{x_0}^{x_1}\Big\{p(x)[y^{'}(x)+\alpha\eta^{'}(x)]^2+q(x)[y(x)+\alpha\eta(x)]^2 \\
&\quad +2f(x)[y(x)+\alpha\eta(x)]\big\}\,dx \\
&= \int_{x_0}^{x_1}\Big[p(x)(y^{'}(x))^2+q(x)y^2(x)+2f(x)y(x)\Big]dx \\
&+ 2\alpha\int_{x_0}^{x_1}\Big[p(x)y^{'}(x)\eta^{'}(x)+q(x)y(x)\eta(x)+f(x)\eta(x)\Big]dx \\
&\quad +\alpha^2\int_{x_0}^{x_1}\Big(p(x)(\eta^{'}(x))^2+q(x)\eta^2(x)\Big)dx \qquad (8.7)
\end{aligned}$$

Thereby,

$$\Delta I = I[y(x)+\alpha\eta(x)]-I[y(x)] = \alpha\delta I+\frac{\alpha^2}{2!}\delta^2 I \qquad (8.8)$$

where

$$\begin{aligned}
\delta I &= 2\int_{x_0}^{x_1}[p(x)y^{'}(x)\eta^{'}(x)+q(x)y(x)\eta(x)+f(x)\eta(x)]dx \qquad (8.9)\\
\delta^2 I &= 2\int_{x_0}^{x_1}\Big[p(x)(\eta^{'}(x))^2+q(x)\eta^2(x)\Big]dx \qquad (8.10)
\end{aligned}$$

We say that

$I[y]$ is stationary at $y=y(x)$ if $\delta I\,\Big|_{y=y(x)}=0$. Since both $p(x),q(x)>0$ and $\delta^2 I>0$, $I[y]$ will reach a minimum at $y=y(x)$.

The first order variation illustrated above is in the sense of Gateaux. The definition of the Gateaux variation is in terms of the so-called Gateaux derivative

$$\delta_G I = D_G I\eta = \lim_{\alpha\to 0}\frac{I(y+\alpha\eta)-I(y)}{\alpha} = \frac{d}{d\alpha}I(y+\alpha\eta)\,\Big|_{\alpha=0} \qquad (8.11)$$

REMARK 8.1.1 *One may compare this with the so-called Frechet derivative, $D_F I[y]\eta$, which is defined as a linear functional such that*

$$\frac{I(y+\eta)-I(y)-D_F I(y)\cdot\eta}{\|\eta\|_V}\Rightarrow 0,\ \text{ as }\ \|\eta\|_V\to 0\,. \qquad (8.12)$$

Gateaus derivative coincides with Fre'chet derivative, if $\delta_F I$ is linear in η and uniformly continuous in η, i.e. $|\delta I(y,\eta)-\delta I(y_0,\eta)|\to 0$, as $y\to y_0$ uniformly $\forall y\in B(y_0)$.

In general, the n-th order Gateaux variation is defined as

$$\delta_G^n I = \frac{d^n}{d\alpha^n} I(y+\alpha\eta)\Big|_{\alpha=0}, \ \forall n \geq 1 \tag{8.13}$$

such that

$$\Delta I = I(y+\alpha\eta) - I(y) = \alpha\delta_G I + \frac{\alpha^2}{2!}\delta_G^2 I + \frac{\alpha^3}{3!}\delta_G^3 I + \frac{\alpha^4}{4!}\delta_G^4 I + \cdots \tag{8.14}$$

In the rest of the book, we omit the subscript G in variation operator. Let $\alpha = 1$. We have

$$\Delta I = I(y+\eta) - I(y) = \delta I + \frac{1}{2!}\delta^2 I + \frac{1}{3!}\delta^3 I + \frac{1}{4!}\delta^4 I + \cdots \tag{8.15}$$

One nice thing about the Gateaux variation is that it is defined based on a scaler differentiation operation. In other words, the variation operation follows the same rule as the differentiation operation in elementary calculus.

This can be seen by examining the first order variation of $I[y]$,

$$\delta I = 2\int_{x_0}^{x_1}\Big[p(x)y^{'}(x)\eta^{'}(x) + q(x)y(x)\eta(x) + f(x)\eta(x)\Big]dx \tag{8.16}$$

Let $\eta(x) = \delta y$. The Gateaux variation becomes,

$$\begin{aligned}\delta I &= 2\int_{x_0}^{x_1}\Big[p(x)y^{'}(x)\delta y^{'} + q(x)y(x)\delta y + f(x)\delta y\Big]dx \\ &= \delta\left\{\int_{x_0}^{x_1}\Big[p(x)(y^{'}(x))^2 + q(x)(y(x))^2 + f(x)y(x)\Big]dx\right\} \\ &= \delta I \ .\end{aligned}$$

This is to say that one can find the first variation of a functional, $I[y]$, by simply differentiating (taking G-derivative) the unknown function according to the same rule of differentiation in calculus. The only difference is: dy is replaced by δy, which is the variation of the unknown function, or in general, a test function satisfying homogeneous boundary conditions, i.e. $\delta y \in \overset{\circ}{\mathcal{V}}$.

Consider the first term in (8.16). Integration by parts yields,

$$\int_{x_0}^{x_1} p(x)y^{'}\eta^{'}dx = [p(x)y^{'}\eta]_{x_0}^{x_1} - \int_{x_0}^{x_1}(p(x)y^{'})\eta dx = -\int_{x_0}^{x_1}(p(x)y^{'})^{'}\eta dx$$

Therefore,

$$\delta I = 2\int_{x_0}^{x_1}\Big[-[p(x)y^{'}(x)]^{'} + q(x)y(x) + f(x)\Big]\eta(x)dx = 0 \tag{8.17}$$

Since this equation must holds for any $\eta(x) \in \overset{\circ}{\mathcal{V}}$, the integrand must vanish, i.e. the solution of the following differential equation

$$-[p(x)y^{'}(x)]^{'}+q(x)y(x)+f(x)=0, \ \ y(x_0)=y_0 \text{ and } y(x_1)=y_1 \ . \quad (8.18)$$

is a minimizer of the functional $I[y]$. Eq. (8.18) is called the Euler-Lagrange equation.

Note that the solution of (8.18), $y^*(x)$ may not be the only minimizer of the functional $I[y]$. In fact, $y^* \in C^1([x_0,x_1])$, and hence Eq. (8.18) is called strong form of the Euler-Lagrange equation. On the other hand, a necessary minimizer only requires that $y \in H^1([x_0,x_1])$, since

$$I = 2\int_{x_0}^{x_1}\Big[p(x)(y^{'}(x))^2 + q(x)y(x)^2\eta(x) + f(x)y(x)\Big]dx \qquad (8.19)$$

and for this purpose we call a function that makes $I[y]$ stationary, but not necessarily satisfy the Euler-Lagrange equation, i.e.,

$$\delta I = 2\int_{x_0}^{x_1}\Big[p(x)y^{'}(x)\delta y^{'} + q(x)y(x)\delta y + f(x)\delta y\Big]dx \qquad (8.20)$$

as the weak solution, since $C^1([x_0,x_1]) \subset H^1([x_0,x_1])$.

In general, consider a functional of the following form,

$$I[y] = \int_{x_0}^{x_1} F(x,y,y^{'})dx, \ \ y(x_0)=y_0 \ \text{ and } \ y(x_1)=y_1 \ . \qquad (8.21)$$

Its first variation is

$$\delta I = \int_{x_0}^{x_1}\left\{\frac{\partial F}{\partial y}\delta y + \frac{\partial F}{\partial y^{'}}\delta y^{'}\right\}dx$$

Integration by parts yields

$$\begin{aligned}\delta I &= \int_{x_0}^{x_1}\Big[\frac{\partial F}{\partial y}\delta y\Big]dx + \frac{\partial F}{\partial y^{'}}\delta y\,\Big|_{x_0}^{x_1} - \int_{x_0}^{x_1}\frac{\partial}{\partial x}\Big[\frac{\partial F}{\partial y^{'}}\Big]\delta y dx \\ &= \int_{x_0}^{x_1}\Big[\frac{\partial F}{\partial y} - \frac{\partial}{\partial x}\Big(\frac{\partial F}{\partial y^{'}}\Big)\Big]\delta y dx \qquad (8.22)\end{aligned}$$

One obtains the Euler-Lagrange equation,

$$E[F]_y = \frac{\partial F}{\partial y} - \frac{\partial}{\partial x}\Big(\frac{\partial F}{\partial y^{'}}\Big) = 0 \ . \qquad (8.23)$$

8.2 Extreme variational principles in linear elasticity

8.2.1 Minimum potential enery principle

Consider a linear elastic solid, V. The total potential energy of the elastic solid is

$$\begin{aligned}\Pi(u_i, u_{i,j}) &= \frac{1}{2}\int_V \sigma_{ij}\epsilon_{ij} dV - \int_V f_i u_i dV - \int_{\Gamma_{\mathbf{t}}} t_i^0 u_i dS \\ &= \frac{1}{2}\int_V C_{ijk\ell} u_{i,j} u_{k,\ell} dV - \int_V f_i u_i dV - \int_{\Gamma_{\mathbf{t}}} t_i^0 u_i dS\end{aligned}$$

The solid is subjected to the following boundary conditions,

$$u_i = u_i^0 = x_j \epsilon_{ij}^0, \quad \forall \mathbf{x} \in \Gamma_{\mathbf{u}} \tag{8.24}$$

$$t_i = n_j \sigma_{ij} = t_i^0 = n_j \sigma_{ij}^0, \quad \forall \mathbf{x} \in \Gamma_{\mathbf{t}} \tag{8.25}$$

where the displacement boundary conditions are essential boundary conditions for ensuing variational principles, because they are the constraints on primary variables u_i and the space of the trial function. Consider trial function $u_i \in \mathcal{V}$,

$$\mathcal{V} := \Big\{ y_i(\mathbf{x}) \;\Big|\; y_i(\mathbf{x}) \in H^1(V), \text{ and } y_i = x_j \epsilon_{ij}^0 \;\; \forall \mathbf{x} \in \Gamma_{\mathbf{u}} \Big\} \tag{8.26}$$

and test function $\delta u_i \in \overset{\circ}{\mathcal{V}}$ where,

$$\overset{\circ}{\mathcal{V}} := \Big\{ \eta_i(\mathbf{x}) \;\Big|\; \eta_i(\mathbf{x}) \in H^1(V), \text{ and } \eta_i(\mathbf{x}) = 0, \;\; \forall \mathbf{x} \in \Gamma_{\mathbf{u}} \Big\} \tag{8.27}$$

which is equivalent to $\delta u_i \in H_c^1(V)$. When $u_i(\mathbf{x}) \in \mathcal{V}$, we say $u_i(\mathbf{x})$ is kinematically addmissible.

A necessary condition that $\Pi(u_i, u_{i,j})$ reaches to an extreme is the stationary condition of its first variation, i.e.

$$\delta\Pi[u_i, u_{i,j}] = \int_V C_{ijk\ell} u_{i,j} \delta u_{k,\ell} dV - \int_V f_i \delta u_i dV - \int_{\Gamma_t} t_i^0 n_i dS = 0 \tag{8.28}$$

which is often called virtual displacement principle in solid mechanics. By the way, the stationary condition in mechanics terms is equilibrium condition. Any y satisfies virtual displacement principle is an equilibrium solution.

On the other hand, Eq.(8.28) is called as the weak formulation of Navier equations in computational mechanics. This can be easily seen via integration by parts,

$$\begin{aligned}
\delta\Pi &= \int_V \sigma_{ij}\delta u_{i,j}dV - \int_V f_i\delta u_i dV - \int_{\Gamma_{\mathbf{t}}} t_i^0 \delta u_i dS \\
&= \int_V \Big((\sigma_{ij}\delta u_i)_{,j} - \sigma_{ij,j}\delta u_i\Big)dV - \int_V f_i\delta u_i dV - \int_{\Gamma_{\mathbf{t}}} t_i^0\delta u_i \\
&= \int_{\partial V} \sigma_{ij}n_j\delta u_i dS - \int_V (\sigma_{ij,j}+f_i)\delta u_i dV - \int_{\Gamma_{\mathbf{t}}} t_i^0\delta u_i dS \\
&= \int_{\Gamma_{\mathbf{t}}} (\sigma_{ij}n_j - t_i^0)\delta u_i dS - \int_V (\sigma_{ij,j}+f_i)\delta u_i dV + \int_{\Gamma_{\mathbf{u}}} \sigma_{ij}n_j\delta u_i dS
\end{aligned}$$

which yields the Navier equation

$$C_{ijk\ell}u_{k,\ell j} + f_i = 0, \tag{8.29}$$

and the natural boundary conditions,

$$\sigma_{ij}n_j = t_i^0 = \sigma_{ij}^0 n_j, \ \ \forall\, \mathbf{x} \in \Gamma_{\mathbf{t}}\ . \tag{8.30}$$

Examine the perturbance of the potential energy $\Delta\Pi(u_i, u_{i,j})$ around an equilibrium configuration,

$$\begin{aligned}
\Delta\Pi &= \Pi(u_i+\delta u_i, u_{i,j}+\delta u_{i,j}) - \Pi(u_i, u_{i,j}) \\
&= \frac{1}{2}\int_V C_{ijk\ell}(u_{i,j}+\delta u_{i,j})(u_{k,\ell}+\delta u_{k,\ell})dV - \int_V f_i(u_i+\delta u_i)dV \\
&\quad - \int_{\Gamma_{\mathbf{t}}} t_i^0(u_i+\delta u_i)dV \\
&\quad -\frac{1}{2}\int_V C_{ijk\ell}u_{i,j}u_{k,\ell}dV - \int_V f_i u_i dV - \int_{\Gamma_{\mathbf{t}}} t_i^0 u_i dV \\
&= \int_V C_{ijk\ell}u_{i,j}\delta u_{k,\ell}dV - \int_V f_i\delta u_i dV - \int_{\Gamma_{\mathbf{t}}} t_i^0\delta u_i dV \\
&\quad +\frac{1}{2}\int_V C_{ijk\ell}\delta u_{i,j}\delta u_{k,\ell}dV \\
&= \delta\Pi + \frac{1}{2!}\delta^2\Pi
\end{aligned} \tag{8.31}$$

For the equilibrium solution $\delta\Pi = 0$, $\Delta\Pi = \frac{1}{2!}\delta^2\Pi > 0$.

This means that for all the kinematically admissible vector fields, $\mathbf{u} = u_i\mathbf{e}_i$, $u_i(\mathbf{x}) \in \mathcal{V}$ the equilibrium solution (real solution ? is the solution unique ? weak solution = strong solution) is the minimizer of total potential energy $\Pi(u_i, u_{i,j})$.

THEOREM 8.1 (MINIMUM POTENTIAL ENERGY PRINCIPLE) *Among all (infinitesimal) kinematically admissible displacement fields, that which is also* *statically admissible (real solution) render the potential energy* Π *an absolute minimum.*

That is $\Pi(\tilde{\mathbf{u}}, \nabla \cdot \tilde{\mathbf{u}}) \leq \Pi(\mathbf{u}, \nabla \cdot \mathbf{u})\ \forall \mathbf{u} \in \mathcal{V}$. Or

$$\Pi(\tilde{\mathbf{u}}, \nabla \cdot \tilde{\mathbf{u}}) = \inf_{\mathbf{u} \in \mathcal{V}} \Pi(\mathbf{u}, \nabla \mathbf{u}) \tag{8.32}$$

If macros strain boundary condition is applied on entire boundary ∂V,

$$\mathbf{u} = \mathbf{x} \cdot \boldsymbol{\epsilon}^0, \ \ \mathbf{x} \in \partial V \tag{8.33}$$

Then $\Gamma_t = \emptyset$ and $\Pi(\mathbf{u}, \nabla \cdot \mathbf{u}) = VW(\nabla \cdot \mathbf{u})$, where

$$W(\nabla \mathbf{u}) := \frac{1}{2V} \int_V C_{ijk\ell} \epsilon_{ij} \epsilon_{k\ell} dV \tag{8.34}$$

The minimum potential energy principle reads as

$$W(\tilde{\boldsymbol{\epsilon}}) = \inf_{\mathbf{u} \in \mathcal{V}} W(\boldsymbol{\epsilon}) \tag{8.35}$$

For the real solution, $\tilde{\mathbf{u}}$,

$$\begin{aligned} W(\tilde{\boldsymbol{\epsilon}}) &= \frac{1}{2V} \int_V \boldsymbol{\sigma} : \boldsymbol{\epsilon} dV = \frac{1}{2} < \tilde{\boldsymbol{\sigma}} >:< \tilde{\boldsymbol{\epsilon}} > \\ &= \frac{1}{2} < \tilde{\boldsymbol{\sigma}} >: \boldsymbol{\epsilon}^0 = \frac{1}{2} \boldsymbol{\epsilon}^0 : \bar{\mathbf{C}} : \boldsymbol{\epsilon}^0 \end{aligned}$$

On the other hand,

$$\begin{aligned} W(\boldsymbol{\epsilon}) &= \frac{1}{2V} \int_V \boldsymbol{\sigma} : \boldsymbol{\epsilon} dV = \frac{1}{2} < \boldsymbol{\sigma} >:< \tilde{\boldsymbol{\epsilon}} > \\ &= \frac{1}{2} < \tilde{\boldsymbol{\sigma}} >: \boldsymbol{\epsilon}^0 = \frac{1}{2} \boldsymbol{\epsilon}^0 : \Big(\sum_{\alpha=0}^{n} \mathbf{C}^\alpha :< \boldsymbol{\epsilon} >_\alpha \Big) \end{aligned}$$

Since $\boldsymbol{\epsilon}^0 \in \mathcal{V}$, we can choose $\boldsymbol{\epsilon}^\alpha = \boldsymbol{\epsilon}^0$. Then we have

$$\frac{1}{2} \boldsymbol{\epsilon}^0 : \bar{\mathbf{C}} : \boldsymbol{\epsilon}^0 \leq \frac{1}{2} \boldsymbol{\epsilon}^0 : \Big(\sum_{\alpha=0}^{n} \mathbf{C}^\alpha : \boldsymbol{\epsilon}^0 \Big)$$

which then leads to

$$\bar{\mathbf{C}} \leq \sum_{\alpha=0}^{n} f_\alpha \mathbf{C}^\alpha \ . \tag{8.36}$$

8.2.2 Minimum complementary potential energy principle

Consider the following complementray potential energy,

$$\Pi^c(\sigma_{ij}) = \frac{1}{2}\int_V D_{ijk\ell}\sigma_{ij}\sigma_{k\ell}dV - \int_{\Gamma_u} u_i^0 \sigma_{ij} n_j dS \tag{8.37}$$

which is a map,

$$\Pi^c : \mathcal{S} \to \mathbb{R} \tag{8.38}$$

where $\mathcal{S}$ is the trial function space

$$\mathcal{S} = \Big\{\sigma_{ij} \;\Big|\; \sigma_{ij} \in H^1(V),\;\; \sigma_{ij,j} = 0 \;\text{ and }\; n_j\sigma_{ij} = t_i^0,\;\; \forall \mathbf{x} \in \Gamma_t\Big\} \tag{8.39}$$

and the test function space is

$$\overset{\circ}{\mathcal{S}} = \Big\{\sigma_{ij} \;\Big|\; \sigma_{ij} \in H^1(V),\;\; \sigma_{ij,j} = 0 \;\text{ and }\; n_j\sigma_{ij} = 0,\;\; \forall \mathbf{x} \in \Gamma_t\Big\} \tag{8.40}$$

Note that in this variational statement, the essential boundary condition becomes

$$n_j\sigma_{ij} = t_i^0,\;\; \forall \mathbf{x} \in \Gamma_t \tag{8.41}$$

whereas the natural boundary condition becomes

$$u_i = \bar{u}_i,\;\; \forall \mathbf{x} \in \Gamma_u\ . \tag{8.42}$$

To study extreme property, we examine complementary potential energy perturbance,

$$\begin{aligned}
\Delta\Pi^c &= \Pi^c(\sigma_{ij} + \delta\sigma_{ij}) - \Pi^c(\sigma_{ij}) \\
&= \Big[\frac{1}{2}\int_V D_{ijk\ell}(\sigma_{ij} + \delta\sigma_{ij})(\sigma_{k\ell} + \delta\sigma_{k\ell})dV - \int_{\Gamma_u} u_i^0(\sigma_{ij} + \delta\sigma_{ij})n_j dS\Big] \\
&\quad -\Big[\frac{1}{2}\int_V D_{ijk\ell}\sigma_{ij}\sigma_{k\ell}dV - \int_{\Gamma_u} u_i^0\sigma_{ij}dS\Big] \\
&= \underbrace{\int_V D_{ijk\ell}\sigma_{ij}\delta\sigma_{k\ell}dV - \int_{\Gamma_u} u_i^0\delta\sigma_{ij}n_j dS}_{=\delta\Pi^c} \\
&\quad + \frac{1}{2}\underbrace{\int_V D_{ijk\ell}\delta\sigma_{ij}\delta\sigma_{k\ell}dV}_{=\delta^2\Pi^c}
\end{aligned}$$

The necessary condition for $\Pi^c(\sigma_{ij})$ attaining extreme value is the stattionary condition,

$$\delta\Pi^c = 0\ .$$

Hence

$$\Delta\Pi^c = \frac{1}{2!}\delta^2\Pi^c > 0 \tag{8.43}$$

since $D_{ijk\ell}$ is positive definite. Thus, $\Pi^c(\sigma_{ij})$ reaches a minimum value at $\sigma_{ij} = \tilde{\sigma}_{ij}$, where $\tilde{\sigma}_{ij}$ renders stationary condition $\delta\Pi^c(\tilde{\sigma}_{ij}) = 0$. This fact is the so-called minimum complementray potential energy principle.

THEOREM 8.2 (MINIMUM COMPLEMENTARY ENERGY PRINCIPLE) *Among all statically admissible stress fields, the actual stress field (whose correrponding strain field satisfies compatibility condition) rensers* Π^c *an absolute minimum, i.e.*

$$\Pi^c(\tilde{\boldsymbol{\sigma}}) \leq \Pi^c(\boldsymbol{\sigma}), \quad \forall \boldsymbol{\sigma} \in \mathcal{S} \tag{8.44}$$

or

$$\Pi^c(\tilde{\boldsymbol{\sigma}}) = \inf_{\boldsymbol{\sigma}\in\mathcal{S}} \Pi^c(\boldsymbol{\sigma}) \tag{8.45}$$

The stationary condition of complementary energy has well-known names, e.g. *virtual force principle* in continuum mechanics, or *the weak form of compatibility condition* in computational mechanics,

$$\delta\Pi^c(\tilde{\sigma}_{ij}) = \int_V D_{ijk\ell}\tilde{\sigma}_{ij}\delta\sigma_{k\ell}dV - \int_{\Gamma_u} u_i^0 \delta\sigma_{ij}n_j dS = 0 \tag{8.46}$$

The above equation can be rewriten as

$$\begin{aligned}&\int_V \tilde{\epsilon}_{ij}\delta\sigma_{ij}dV - \frac{1}{2}\int_V \Big(u_{i,j}\delta\sigma_{ij} + u_{j,i}\delta\sigma_{ij}\Big)dV \\ &+ \int_V u_{i,j}\delta\sigma_{ij}dV - \int_{\Gamma_u} u_i^0\delta\sigma_{ij}n_j dS = 0\end{aligned}$$

Integration by parts yields

$$\begin{aligned}&\int_V \Big(\tilde{\epsilon}_{ij} - \frac{1}{2}(u_{i,j}+u_{j,i})\delta\sigma_{ij}\Big)dV + \int_{\partial V} u_i\delta\sigma_{ij}n_j dS \\ &- \underbrace{\int_V u_i\delta\sigma_{ij,j}dV}_{=0} - \int_{\Gamma_u} u_i^0\delta\sigma_{ij}n_j dS = 0 \\ \Rightarrow \quad &\int_V \Big(\tilde{\epsilon}_{ij} - \frac{1}{2}(u_{i,j}+u_{j,i})\Big)\delta\sigma_{ij}dV + \int_{\Gamma_u}(u_i - u_i^0)\delta\sigma_{ij}n_j dS = 0\,.\end{aligned}$$

which leads to the Euler-Lagrange equation,

$$\tilde{\epsilon}_{ij} = D_{ijk\ell}\tilde{\sigma}_{k\ell} = \frac{1}{2}(u_{i,j}+u_{j,i}) \tag{8.47}$$

$$\Rightarrow \quad \tilde{\epsilon}_{ij,k\ell} + \tilde{\epsilon}_{k\ell,ij} - \tilde{\epsilon}_{ik,j\ell} - \tilde{\epsilon}_{j\ell,ik} = 0\,. \tag{8.48}$$

and the natural boundary condition

$$u_i = u_i^0, \;\; \forall \mathbf{x} \in \Gamma_u \tag{8.49}$$

Consider prescribed macro-stress boundary condition, $\mathbf{n}\cdot\boldsymbol{\sigma} = \mathbf{t}^0, \forall \mathbf{x} \in \partial V$, $\Gamma_u = \emptyset$. In this case, $\Gamma_u = \emptyset$. Therefore,

$$\Pi^c = \frac{1}{2}\int_V D_{ijk\ell}\sigma_{ij}\sigma_{k\ell}dV = W_c(\boldsymbol{\sigma})V \tag{8.50}$$

where

$$W_c := \frac{1}{2V}\int_V D_{ijk\ell}\sigma_{ij}\sigma_{k\ell}dV \tag{8.51}$$

is the complementary energy density.

The minimum complementary potential energy principle then gives

$$W_c(\tilde{\boldsymbol{\sigma}}) = \inf_{\boldsymbol{\sigma}\in\mathcal{S}} W_c(\boldsymbol{\sigma}) \tag{8.52}$$

Recall,

$$< \boldsymbol{\sigma} : \boldsymbol{\epsilon} > - < \boldsymbol{\sigma} >:< \boldsymbol{\epsilon} >= \frac{1}{V}\int_{\partial V}\Big(\mathbf{u}-\mathbf{x}\cdot < \nabla\otimes\mathbf{u} >\Big)\Big(\mathbf{n}\cdot(\boldsymbol{\sigma}- < \boldsymbol{\sigma} >)\Big)dS$$

The real complementary energy density becomes

$$\begin{aligned} W_c(\tilde{\boldsymbol{\sigma}}) &= \frac{1}{2} < \tilde{\boldsymbol{\sigma}} : \tilde{\boldsymbol{\epsilon}} >= \frac{1}{2} < \tilde{\boldsymbol{\sigma}} >:< \tilde{\boldsymbol{\epsilon}} > \\ &= \frac{1}{2}\boldsymbol{\sigma}^0 : \bar{\mathbf{D}} : \boldsymbol{\sigma}^0 = \frac{1}{2} < \tilde{\boldsymbol{\sigma}} >: \bar{\mathbf{D}} :< \tilde{\boldsymbol{\sigma}} > \end{aligned} \tag{8.53}$$

Note that under prescribed remote stress boundary condition,

$$< \boldsymbol{\sigma} >= \boldsymbol{\sigma}^0, \;\; \forall\, \boldsymbol{\sigma} \in \mathcal{S}\ .$$

Choose $\boldsymbol{\sigma} = \boldsymbol{\sigma}^0 \in \mathcal{S}$,

$$\begin{aligned} W_c(\boldsymbol{\sigma}) &= \frac{1}{2V}\int_V \boldsymbol{\sigma} : \boldsymbol{\epsilon} dV = \boldsymbol{\sigma}^0 : \frac{1}{2V}\int_V \boldsymbol{\epsilon} dV \\ &= \boldsymbol{\sigma}^0 : \frac{1}{2V}\int_V \sum_{\alpha=0}^{n} \mathbf{D}^\alpha : \boldsymbol{\sigma}^\alpha dV \\ &= \frac{1}{2}\boldsymbol{\sigma}^0 : \sum_{\alpha=0}^{n} \frac{\Omega_\alpha}{V}\mathbf{D}^\alpha : \boldsymbol{\sigma}^0 \\ &= \frac{1}{2}\boldsymbol{\sigma}^0 : \sum_{\alpha=0}^{n} f_\alpha \mathbf{D}^\alpha : \boldsymbol{\sigma}^0 \end{aligned}$$

Therefore,

$$\boldsymbol{\sigma}^0 : \bar{\mathbf{D}} : \boldsymbol{\sigma}^0 < \boldsymbol{\sigma}^0 : \sum_{\alpha=0}^{n} f_\alpha \mathbf{D}^\alpha : \boldsymbol{\sigma}^0 \tag{8.54}$$

Since $\bar{\mathbf{D}} : \bar{\mathbf{C}} = \mathbf{1}^{(4s)}$ and both $\bar{\mathbf{D}}$ and $\bar{\mathbf{C}}$ are positive definite, we then have

$$\Big(\sum_{\alpha=0}^{n} f_\alpha \mathbf{C}_\alpha^{-1}\Big)^{-1} \leq \bar{\mathbf{C}} \tag{8.55}$$

which is called the Reuss bound. It is a lower bound for elastic moduli.

Assume that

$$\begin{aligned} \mathbf{C}^\alpha &= 3K^\alpha \mathbf{E}^{(1)} + 2\mu^\alpha \mathbf{E}^{(2)} \\ \mathbf{C}^{\alpha-1} &= \frac{1}{3K^\alpha}\mathbf{E}^{(1)} + \frac{1}{2\mu^\alpha}\mathbf{E}^{(2)} \end{aligned}$$

One can derive that

$$\begin{aligned} \Big(\sum_{\alpha=0}^{n} f_\alpha \mathbf{C}^\alpha\Big) &= 3\sum_{\alpha=0}^{n} f_\alpha K_\alpha \mathbf{E}^{(1)} + 2\sum_{\alpha=0}^{n} f_\alpha \mu_\alpha \mathbf{E}^{(2)} \\ \Big(\sum_{\alpha=0}^{n} f_\alpha \mathbf{C}^{\alpha-1}\Big)^{-1} &= \frac{3}{\displaystyle\sum_{\alpha=0}^{n} \frac{f_\alpha}{K_\alpha}}\mathbf{E}^{(1)} + \frac{2}{\displaystyle\sum_{\alpha=0}^{n} \frac{f_\alpha}{\mu_\alpha}}\mathbf{E}^{(2)} \end{aligned}$$

Combining Reuss bound with the Voigt bound, we have

$$\Big(\sum_{\alpha=0}^{n} f_\alpha \mathbf{C}^{\alpha-1}\Big) < \bar{\mathbf{C}} < \Big(\sum_{\alpha=0}^{n} f_\alpha \mathbf{C}^\alpha\Big)$$

and consequently,

$$\begin{aligned} \frac{1}{\displaystyle\sum_{\alpha=0}^{n} \frac{f_\alpha}{K_\alpha}} &< \bar{K} < \sum_{\alpha=0}^{n} f_\alpha K_\alpha \\ \frac{1}{\displaystyle\sum_{\alpha=0}^{n} \frac{f_\alpha}{\mu_\alpha}} &< \bar{\mu} < \sum_{\alpha=0}^{n} f_\alpha \mu_\alpha \end{aligned}$$

One can see that the Voigt bound is in fact an arithmetic average and the Reuss bound can be viewed as a geometric average or the harmonic average.

8.3 Hashin-Shtrikman variational principles

In order to narrow the gap between the Voigt bound and the Reuss bound, we need new mathematical tools. One of powerful such tools is the celebrated Hashin-Shtrikman (HS) variational principle. The essence of the HS variational principles is that they are the variational principles specifically designed for composites, or inhomogeneous solids. To measure the differences between homogeneous solids and inhomogeneous solids, a comparison homogenous solid is used to identify the inhomogeneous fields.

Let's first consider a boundary value problem of the original composite (RVE),

$$\begin{aligned} \sigma_{ij,j} &= 0, \\ \sigma_{ij} &= C_{ijk\ell}(\mathbf{x})\epsilon_{k\ell}, \\ U(\boldsymbol{\epsilon}) &= \frac{1}{2}C_{ijk\ell}\epsilon_{ij}\epsilon_{k\ell}, \ \text{ and } \ W(\boldsymbol{\epsilon}) =< U(\boldsymbol{\epsilon}) >_V \\ u_i &= \bar{u}_i, \ \forall \mathbf{x} \in \Gamma_u, \quad (\Gamma_t = \emptyset, \ \ \Gamma_u = \partial V). \end{aligned}$$

Consider a second BVP in a comparison solid,

$$\begin{aligned} \sigma^{(0)}_{ij,j} &= 0, \\ \sigma^{(0)}_{ij} &= C^{(0)}_{ijk\ell}(\mathbf{x})\epsilon^{(0)}_{k\ell}, \\ U^{(0)}(\boldsymbol{\epsilon}^{(0)}) &= \frac{1}{2}C^{(0)}_{ijk\ell}\epsilon^{(0)}_{ij}\epsilon^{(0)}_{k\ell}, \ \text{ and } \ W_0(\boldsymbol{\epsilon}^{(0)}) =< U^{(0)}(\boldsymbol{\epsilon}^{(0)}) >_V \\ u^{(0)}_i &= \bar{u}_i, \ \forall \mathbf{x} \in \Gamma_u, \quad (\Gamma_t = \emptyset, \ \ \Gamma_u = \partial V). \end{aligned}$$

To relate the two BVPs, we introduce the following decomposition in strain field and stress field,

$$u_i = u^{(0)}_i + u^d_i \tag{8.56}$$

$$\epsilon_{ij} = \epsilon^{(0)}_{ij} + \epsilon^d_{ij} \tag{8.57}$$

and

$$\begin{aligned} \sigma_{ij} &= p_{ij} + C^{(0)}_{ijk\ell}\epsilon_{k\ell} \\ &= p_{ij} + C^{(0)}_{ijk\ell}(\epsilon^{(0)}_{ijk\ell} + \epsilon^d_{ijk\ell}) \end{aligned} \tag{8.58}$$

where u^d_i is the disturbance displacement field and p_{ij} is called polarization stress.

A better definition of stress polarization is

$$p_{ij} = \sigma_{ij} - C^{(0)}_{ijk\ell}\epsilon_{k\ell} = (C_{ijk\ell} - C^{(0)}_{ijk\ell})\epsilon_{k\ell} \tag{8.59}$$

which indicates that stress polarization is due to inhomogeneouness of the composite.

Furthermore, since

$$u_i = \bar{u}_i, \ \ \forall \mathbf{x} \in \partial V \ \text{ and } \ u_i^{(0)} = \bar{u}_i, \ \ \forall \mathbf{x} \in \partial V$$

it leads to homogeneous boundary conditions for displacement disturbance field

$$u_i^d = 0, \ \ \forall \mathbf{x} \in \partial V \tag{8.60}$$

In passing, we note that because $u_i^d = 0, \forall \mathbf{x} \in \partial V$ it can be readily to show that the average work done by the disturbance field over any self-equilibrium stress field will be zero, that is

$$\begin{aligned} \int_V \sigma_{ij}\epsilon_{ij}^d dV &= \int_V \sigma_{ij} u_{i,j}^d dV \\ &= \int_{\partial V} u_i^d n_j \sigma_{ij} dS + \int_V u_i^d \sigma_{ij,j} dV = 0 \ . \end{aligned} \tag{8.61}$$

On the other hand, since

$$\sigma_{ij,j} = 0, \quad \sigma_{ij,j}^{(0)} = 0 \ ,$$

one has

$$\sigma_{ij,j} = \sigma_{ij,j}^{(0)} + p_{ij,j} + \left(C_{ijk\ell}^{(0)} \epsilon_{k\ell}^d \right)_{,j} = 0$$

We can see that the stress field can be divided into the homogeneous (or comparison) stress field, $\sigma_{ij}^{(0)}$, and the inhomogeneous stress field,

$$\sigma_{ij} = \sigma_{ij}^0 + t_{ij}, \ \ \text{where } \ t_{ij} = p_{ij} + C_{ijk\ell}^0 \epsilon_{k\ell}^d \tag{8.62}$$

Both homogeneous stress field, σ_{ij}^0, and inhomogeneous stress field, t_{ij} satisfy equilibrium equations, i.e.

$$\sigma_{ij,j}^{(0)} = 0, \ \ t_{ij,j} = 0 \ \ . \tag{8.63}$$

In literature, the inhomogeneous equilibrium equation

$$t_{ij,j} = \left(C_{ijk\ell}^{(0)} \epsilon_{k\ell}^d \right)_{,j} + p_{ij,j} = 0 \tag{8.64}$$

is often called “the subsidiary condition.”

THEOREM 8.3 (HASHIN-SHTRIKMAN) *Let* $u_i^d \in \mathcal{U}$ *and* $p_{ij} \in \mathcal{S}$ *where*

$$\mathcal{U} = \left\{ u_i \ \middle| \ u_i \in H^1(V), u_i = 0, \forall \mathbf{x} \in \partial V \right\} \tag{8.65}$$

$$\mathcal{S} = \left\{ \sigma_{ij} \ \middle| \ \sigma_{ij} \in L^2(V) \right\} \tag{8.66}$$

Consider the following functional,

$$\Pi : \mathcal{S} \times \mathcal{U} \to \mathbb{R},$$

where

$$\Pi(p_{ij}, \epsilon^d_{ij}) = \frac{1}{2}\int_V \Big(C^{(0)}_{ijk\ell}\epsilon^{(0)}_{ij}\epsilon^{(0)}_{k\ell} - \Delta C^{-1}_{ijk\ell} p_{ij} p_{k\ell} + p_{ij}\epsilon^d_{ij} + 2p_{ij}\epsilon^{(0)}_{ij}\Big) dV$$

$$\text{where} \quad \left\{ \begin{array}{lll} \Delta C_{ijk\ell} & = & C_{ijk\ell} - C^{(0)}_{ijk\ell} \\ p_{ij} & = & \Delta C_{ijk\ell}\epsilon_{k\ell} \\ \epsilon^d_{ij} & = & \epsilon_{ij} - \epsilon^{(0)}_{ij} \end{array} \right. \tag{8.67}$$

We have the following variational statements:

1. *The functional* Π *is stationary, i.e.* $\delta\Pi = 0$*, if the inhomogeneous equilibrium equation (subsidiary condition) is satisfied,*

$$\Big(C^{(0)}_{ijk\ell}\epsilon^d_{k\ell} \Big)_{,j} + p_{ij,j} = 0\ ; \tag{8.68}$$

2.

$$\delta^2\Pi > 0, \ \text{ if } \ \Delta\mathbf{C} < 0, \ \ \Pi \to Minimum \tag{8.69}$$

$$\delta^2\Pi < 0, \ \text{ if } \ \Delta\mathbf{C} > 0, \ \ \Pi \to Maximum \tag{8.70}$$

Proof:

$$\begin{aligned} \Delta\Pi &= \Pi(p_{ij} + \delta p_{ij}, \epsilon^d_{ij} + \delta\epsilon^d_{ij}) - \Pi(p_{ij}, \epsilon^d_{ij}) \\ &= \frac{1}{2}\int_V \Big(-2\Delta C^{-1}_{ijk\ell} p_{ij}\delta p_{k\ell} + p_{ij}\delta\epsilon^d_{ij} + \delta p_{ij}\epsilon^d_{ij} + 2\delta p_{ij}\epsilon^{(0)}_{ij} \Big) dV \\ &\quad + \frac{1}{2}\int_V \Big(-\Delta C^{-1}_{ijk\ell}\delta p_{ij}\delta p_{k\ell} + \delta p_{ij}\delta\epsilon^d_{k\ell} \Big) dV = \delta\Pi + \frac{1}{2!}\delta^2\Pi \end{aligned}$$

We first show that the first statement is true.

$$\begin{aligned} \delta\Pi &= \Big(-\frac{1}{2}\Big)\int_V \Big(2\Delta C^{-1}_{ijk\ell} p_{k\ell}\delta p_{ij} - 2\epsilon^{(0)}_{ij}\delta p_{ij} - \epsilon_{ij}\delta p_{ij} - p_{ij}\delta\epsilon^d_{ij} \Big) dV \\ &= \Big(-\frac{1}{2}\Big)\int_V \Big(2\Delta C^{-1}_{ijk\ell} p_{k\ell}\delta p_{ij} - 2\underbrace{(\epsilon_{ij} - \epsilon^{(d)}_{ij})}_{=\epsilon^{(0)}_{ij}}\delta p_{ij} - \epsilon^d_{ij}\delta p_{ij} - p_{ij}\delta\epsilon^d_{ij} \Big) dV \\ &= \Big(-\frac{1}{2}\Big)\int_V 2\underbrace{(\Delta C^{-1}_{ijk\ell} p_{k\ell} - \epsilon_{ij})}_{=0}\delta p_{ij} + \epsilon^d_{ij}\delta p_{ij} - p_{ij}\delta\epsilon^d_{ij} \Big) dV \\ &= \Big(-\frac{1}{2}\Big)\int_V \Big(\epsilon^d_{ij}\delta p_{ij} - p_{ij}\delta\epsilon^d_{ij} \Big) dV \end{aligned} \tag{8.71}$$

If the subsidiary condition is satisfied, i.e.

$$\Big(C^{(0)}_{ijk\ell}\epsilon^d_{k\ell}\Big)_{,j} + p_{ij,j} = 0, \ \text{ or } \ t_{ij,j} = 0\ . \tag{8.72}$$

which leads to

$$\delta t_{ij} = \delta p_{ij} + C^{(0)}_{ijk\ell}\delta\epsilon^d_{k\ell}, \ \text{ and } \ \delta t_{ij,j} = 0\ \ . \tag{8.73}$$

Substituting (8.72) and (8.73) into (8.71) yields

$$\begin{aligned}
\delta\Pi &= \Big(-\frac{1}{2}\Big)\int_V \Big(\epsilon^d_{ij}(\delta t_{ij} - C^{(0)}_{ijk\ell}\delta\epsilon^d_{k\ell})\Big) - \delta\epsilon^d_{ij}(t_{ij} - C^{(0)}_{ijk\ell}\epsilon^d_{k\ell})\Big)dV \\
&= \Big(-\frac{1}{2}\Big)\int_V \left(\Big(\epsilon^d_{ij}\delta t_{ij} - t_{ij}\delta\epsilon^d_{ij}\Big) - \underbrace{\epsilon^d_{ij}C^{(0)}_{ijk\ell}\delta\epsilon^d_{k\ell} + \delta\epsilon^d_{ij}C^{(0)}_{ijk\ell}\epsilon^d_{k\ell}}_{=0,\ because\ \mathbf{C}^{(0)}\ has\ major\ symmetry}\right)dV \\
&= \Big(-\frac{1}{2}\Big)\int_V \Big(u^d_{i,j}\delta t_{ij} - t_{ij}\delta u^d_{i,j}\Big)dV
\end{aligned}$$

Considering the facts

$$\begin{aligned}
\int_V \delta t_{ij}u^d_{i,j}dV &= \int_{\partial V}\delta t_{ij}n_j u^d_i dS - \int_V \delta t_{ij,j}u^d_i dV \equiv 0 \\
\int_V t_{ij}\delta u^d_{i,j}dV &= \int_{\partial V} t_{ij}n_j\delta u^d_i dS - \int_V t_{ij,j}\delta u^d_i dV \equiv 0,
\end{aligned}$$

we just proved that $\delta\Pi = 0$, if $t_{ij,j} = 0$ holds.

Now we examin the extreme conditions. Substituting $\delta p_{ij} = \delta t_{ij} - C^{(0)}_{ijk\ell}\delta\epsilon^d_{k\ell}$ into the second order variation,

$$\begin{aligned}
\delta^2\Pi &= \Big(\frac{1}{2}\Big)\int_V \Big(-\Delta C^{-1}_{ijk\ell}\delta p_{ij}\delta p_{k\ell} + \delta p_{ij}\delta\epsilon^d_{ij}\Big)dV \\
&= \Big(-\frac{1}{2}\Big)\int_V \Big(-\Delta C^{-1}_{ijk\ell}\delta p_{ij}\delta p_{k\ell} + C^{(0)}_{ijk\ell}\delta\epsilon^d_{ij}\epsilon^d_{k\ell} - \delta t_{ij}\epsilon^d_{ij}\Big)dV
\end{aligned}$$

Agan, the last term

$$\int_V \delta t_{ij}\epsilon^d_{ij}dV = 0.$$

Therefore, we have

$$\delta^2\Pi = \Big(-\frac{1}{2}\Big)\int_V \Big(\Delta C^{-1}_{ijk\ell}\delta p_{ij}\delta p_{k\ell} + C^{(0)}_{ijk\ell}\delta\epsilon^d_{ij}\epsilon^d_{k\ell}\Big)dV \tag{8.74}$$

Obviously if $\Delta\mathbf{C} > 0$, $\Delta\Pi = \delta^2\Pi < 0$, therefore, Π achives a maximum value.

On the other hand, if $\Delta\mathbf{C} < 0$, the judgement is not straightforward. Consider a positive integral,

$$I := \int_V C^{(0)}_{ijk\ell}{}^{-1} \delta p_{ij} \delta p_{k\ell} dV > 0 \tag{8.75}$$

Substitute $\delta p_{ij} = \delta t_{ij} - C^{(0)}_{ijk\ell} \delta\epsilon^d_{k\ell}$ into (8.75). It can be readily shown that

$$\begin{aligned} I &= \int_V \Big(C^{(0)}_{ijk\ell}{}^{-1} \delta t_{ij} \delta t_{k\ell} - \underbrace{2\delta t_{ij} \delta\epsilon^d_{k\ell}}_{=0} + C^{(0)}_{ijk\ell} \delta\epsilon^d_{ij} \delta\epsilon^d_{k\ell} \Big) dV \\ &= \int_V \Big(C^{(0)}_{ijk\ell}{}^{-1} \delta t_{ij} \delta t_{k\ell} + C^{(0)}_{ijk\ell} \delta\epsilon^d_{ij} \delta\epsilon^d_{k\ell} \Big) dV \end{aligned}$$

A direct consequency is

$$\int_V C^{(0)}_{ijk\ell}{}^{-1} \delta p_{ij} \delta p_{k\ell} dV > \int_V C^{(0)}_{ijk\ell} \delta\epsilon^d_{ij} \epsilon^d_{k\ell} dV \tag{8.76}$$

which leads the following inequality,

$$\begin{aligned} \delta\Pi &= \Big(-\frac{1}{2}\Big) \int_V \Big(\Delta C^{-1}_{ijk\ell} \delta p_{ij} \delta p_{k\ell} + C^{(0)}_{ijk\ell} \delta\epsilon^d_{ij} \epsilon^d_{k\ell} \Big) dV \\ &\geq \Big(-\frac{1}{2}\Big) \int_V \Big(\Delta C^{-1}_{ijk\ell} + C^{(0)}_{ijk\ell}{}^{-1} \Big) \delta p_{ij} \delta p_{k\ell} dV \end{aligned}$$

Consider

$$\begin{aligned} \Delta\mathbf{C}^{-1} + \mathbf{C}^{(0)^{-1}} &= \Delta\mathbf{C}^{-1} + \mathbf{C}^{(0)^{-1}} : (\mathbf{C} - \mathbf{C}^{(0)}) : (\mathbf{C} - \mathbf{C}^{(0)})^{-1} \\ &= \Delta\mathbf{C}^{-1} + \mathbf{C}^{(0)^{-1}} : \mathbf{C} : \Delta\mathbf{C}^{-1} - \Delta\mathbf{C}^{-1} \\ &= \mathbf{C}^{(0)^{-1}} : \mathbf{C} : \Delta\mathbf{C}^{-1} \; . \end{aligned}$$

One can write that

$$\delta^2\Pi > \Big(-\frac{1}{2}\Big) \int_V \mathbf{p} : \mathbf{C}^{(0)^{-1}} : \mathbf{C} : \Delta\mathbf{C}^{-1} : \mathbf{p} dV \tag{8.77}$$

It is clear now that if $\Delta\mathbf{C}^{-1} < 0$, $\delta^2\Pi > 0$ and hence Π has a global minimum. To sum up, we have the following extreme conditions,

$$\begin{aligned} &\delta^2\Pi < 0, \text{ if } \Delta\mathbf{C} > 0, \ \Pi \rightarrow \text{maximum} \; ; \\ &\delta^2\Pi > 0, \text{ if } \Delta\mathbf{C} < 0, \ \Pi \rightarrow \text{minimum} \; . \end{aligned}$$

♣

Since both σ_{ij} and σ^0_{ij} are self-equlibrium stress field,

$$\begin{aligned}\int_V \sigma_{ij}\epsilon^d_{ij}dV &= \int_V \sigma_{ij}u^d_{i,j}dV = 0\\ \int_V \sigma^{(0)}_{ij}\epsilon^d_{ij}dV &= \int_V \sigma^{(0)}_{ij}u^d_{i,j}dV = 0\end{aligned}$$

because $u^d_i = 0,\ \ \forall \mathbf{x} \in \partial V$.

Therefore the total potential energy of a kinematically admissible field, $u_i \in \mathcal{V}$, can be written as

$$\begin{aligned}\Pi(\boldsymbol{\epsilon}) &= \frac{1}{2}\int_V \sigma_{ij}\epsilon_{ij}dV = \frac{1}{2}\int_V \Big(\sigma_{ij}\epsilon^{(0)}_{ij} - \underbrace{\sigma_{ij}\epsilon^d_{ij}}_{=0}\Big)dV\\ &= \frac{1}{2}\int_V \sigma_{ij}\epsilon^{(0)}_{ij}dV\end{aligned}$$

Consider

$$\begin{aligned}\sigma_{ij}\epsilon^{(0)}_{ij} &= \Big(\sigma^{(0)}_{ij} + p_{ij} + C^{(0)}_{ijk\ell}\epsilon^d_{k\ell}\Big)\epsilon^{(0)}_{ij}\\ &= \sigma^{(0}_{ij}\epsilon^{(0)}_{ij} + p_{ij}\epsilon^{(0)}_{ij} + C^{(0)}_{ijk\ell}\epsilon^d_{k\ell}\epsilon^{(0)}_{ij} + + \underbrace{p_{ij}\epsilon^{(0)}_{ij} - p_{ij}\epsilon^{(0)}_{ij}}_{=0}\\ &= C^{(0)}_{ijk\ell}\epsilon_{k\ell}\epsilon^{(0)}_{ij} + C^{(0)}_{ijk\ell}\epsilon^d_{k\ell}\epsilon^{(0)}_{ij} + +2p_{ij}\epsilon^{(0)}_{ij} - p_{ij}(\epsilon_{ij} - \epsilon^d_{ij})\\ &= C^{(0)}_{ijk\ell}\epsilon^{(0)}_{k\ell}\epsilon^{(0)}_{ij} + \underbrace{C^{(0)}_{ijk\ell}\epsilon^{(0)}_{ij}\epsilon^d_{k\ell}}_{=0} + 2p_{ij}\epsilon^{(0)}_{ij} - p_{ij}\epsilon_{ij} + p_{ij}\epsilon^d_{ij}\end{aligned}$$

Therefore under prescribed remote strain boundary condition,

$$\begin{aligned}\Pi(\boldsymbol{\epsilon}) &= \frac{1}{2}\int_V \sigma_{ij}\epsilon_{ij}dV = W(\boldsymbol{\epsilon})V\\ &= \frac{1}{2}\int_V \Big(C^{(0)}_{ijk\ell}\epsilon^{(0)}_{ij}\epsilon^{(0)}_{k\ell} - \Delta C^{-1}_{ijk\ell}p_{ij}p_{k\ell} + p_{ij}\epsilon^d_{ij} + 2p_{ij}\epsilon^{(0)}_{ij}\Big)dV\\ &= W^{(0)}(\boldsymbol{\epsilon}^{(0)})V + \frac{1}{2}\int_V \Big(-\Delta C^{-1}_{ijk\ell}p_{ij}p_{k\ell} + p_{ij}\epsilon^d_{ij} + 2p_{ij}\epsilon^{(0)}_{ij}\Big)dV\\ &= W^{(0)}(\boldsymbol{\epsilon}^{(0)})V + R_\pi V\end{aligned}$$

where $R_\pi := \frac{1}{2V}\int_V\Big(-\Delta C^{-1}_{ijk\ell}p_{ij}p_{k\ell} + p_{ij}\epsilon^d_{ij} + 2p_{ij}\epsilon^{(0)}_{ij}\Big)dV$.

Based on Hashin-Shtrikman principle, if $\Delta\mathbf{C} > 0$ Π has a global minimum, $W^{(0)}(\boldsymbol{\epsilon}^{(0)}) + \underline{R_\pi}$; whereas if $\Delta\mathbf{C} < 0$, Π has a global maximum, $W^{(0)}(\boldsymbol{\epsilon}^{(0)}) + \bar{R_{pi}}$. Therefore, the Hashin-Shtrikman principle provides the following bound,

$$\underline{R_\pi}(\tilde{\mathbf{p}}, \tilde{\boldsymbol{\epsilon}}^d) \le W(\boldsymbol{\epsilon}) - W^{(0)}(\boldsymbol{\epsilon}^{(0)}) \le \bar{R}_\pi(\tilde{\mathbf{p}}, \tilde{\boldsymbol{\epsilon}}^d) \tag{8.78}$$

8.4 Review of Functional Analysis and Convex Analysis

DEFINITION 8.4 (VECTOR SPACE (LINEAR SPACE)) *Let F be a field, whose elements are referred to as scalars. A vector space over F is a nonempty set V, whose elements are referred to as vectors, together with two operations. The first operation, called addition and denoted by +, assignes to each pair* $(\mathbf{u},\mathbf{v}) \in V \times V$ *of vectors in V a vector* $\mathbf{u}+\mathbf{v}$ *in V. The second operation, called multiplication and denoted by juxtaposition, assigns to each pair* $(r,\mathbf{u}) \in F \times V$ *a vector* $r\mathbf{v} \in V$*. Furthermore, the following properties must be satisfied,*

1 *Associativity of addition*

$$\mathbf{u}+(\mathbf{v}+\mathbf{w}) = (\mathbf{u}+\mathbf{v})+\mathbf{w}, \quad \forall \mathbf{u},\mathbf{v},\mathbf{w} \in V$$

2 *Commutivity of addition*

$$\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}, \quad \forall \mathbf{u},\ \mathbf{v} \in V$$

3 *Existence of a zero vector,* $\mathbf{0} \in V$ *such that*

$$\mathbf{0}+\mathbf{u} = \mathbf{u}+\mathbf{0} = \mathbf{u}, \quad \forall \mathbf{u} \in V$$

4 *Existence of additive inverse: i.e.* $\forall \mathbf{u} \in V$*,* $\exists -\mathbf{u} \in V$*, such that*

$$\mathbf{u}+(-\mathbf{u}) = (-\mathbf{u})+\mathbf{u} = \mathbf{0}$$

5 *Properties of scalar multiplication.* $\forall r,s \in F$ *and* $\mathbf{u},\mathbf{v} \in V$*,*

$$\begin{aligned} r(\mathbf{u}+\mathbf{v}) &= r\mathbf{u}+r\mathbf{v} \\ (r+s)\mathbf{u} &= r\mathbf{u}+r\mathbf{v} \\ rs\mathbf{u} &= r(s\mathbf{u}) \\ 1\mathbf{u} &= \mathbf{u} \end{aligned}$$

REMARK 8.4.1 1 *The first four properties in the definitions of vector space can be summarized that V is an abelian group under addition;*

2 *Any expression of the form*

$$r_1\mathbf{v}_1 + r_2\mathbf{v}_2 + \cdots + r_n\mathbf{v}_n$$

where $r_i \in F$ *and* $\mathbf{v}_i \in V\ \forall i = 1,2,\cdots,n$ *is called a linear combination of the vectors* $\mathbf{v}_1,\mathbf{v}_2,\cdots,\mathbf{v}_n$*, and*

$$r_1\mathbf{v}_1 + r_2\mathbf{v}_2 + \cdots + r_n\mathbf{v}_n \in V$$

3 *The addition operation*

$$V \times V \to V : (\mathbf{u}, \mathbf{v}) \to \mathbf{u} + \mathbf{v} \in V$$

4 *and the scalar multiplication operation,*

$$F \times V \to V : (\alpha, \mathbf{u}) \to \alpha \mathbf{u} \in V$$

are closed.

5 *When the operations*

$$f : (\mathbf{u}, \mathbf{v}) \to \mathbf{u} + \mathbf{v} \in V$$

$$g : (\alpha, \mathbf{u}) \to \alpha \mathbf{u} \in V$$

are continuous, the vector space is called topological vector space.

EXAMPLE 8.5 *Let $F = \mathbb{R}$. The set of all ordered n-tuples, i.e.*

$$\mathbf{u} = (u_1, u_2, \cdots, u_n), \ \ u_i \in \mathbb{R}$$

with addition and scalar miltiplication defined component-wise,

$$(a_1, \cdots, a_n) + (b_1, \cdots, b_n) = (a_1 + b_1, \cdots, a_n + b_n)$$

and

$$\alpha(a_1, \cdots, a_n) = (\alpha a_1, \cdots, \alpha a_n)$$

is a vector space, and it is denoted as $\mathbb{R}^n$. Note that in general vector space (a mathematical concept) is still a primitive set. It may have some algebraic structures, but it does not have topologival structures, or geometric structures, such as distance between two elements.

EXAMPLE 8.6 *Let $F = \mathbb{R}$. The set of all continuous function, $C^0(\mathbb{R})$, i.e.* $\forall f \in C^0(\mathbb{R})$

$$f : X \subset \mathbb{R} \to Y \subset \mathbb{R}$$

and

$$d_Y(f(x), f(y)) < \epsilon, \forall d_X(x, y) < \delta, \ \ \forall \delta > 0 \ .$$

is a vector space under the operations of addition and scalar multiplication, i.e.

$$(f + g)(x) = f(x) + g(x), \ \ f, g \in C^0(\mathbb{R})$$

and

$$\alpha f(x) = \alpha f(x), \ \ \forall \alpha \in \mathbb{R}, \ \ f \in C^0(\mathbb{R})$$

DEFINITION 8.7 (BILINEAR FORM) *Let X be a vector space and X^* is its dual space. A mapping g of $X \times X^*$ into $\mathbb{R}$ is called a bilinear functional or a bilinear form if*

1 *For fixed $\mathbf{y}$, $g(\mathbf{x}, \mathbf{y})$ is a linear functional in $\mathbf{x}$, i.e.*

$$g(\alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z}) = \alpha g(\mathbf{x}, \mathbf{z}) + \beta g(\mathbf{x}, \mathbf{z}), \ \ \forall \mathbf{x}, \mathbf{y} \in X, \ \ \mathbf{z} \in X^*$$

2 *For fixed $\mathbf{x}$, $g(\mathbf{x}, \mathbf{y})$ is a linear functional in $\mathbf{y}$, i.e.*

$$g(\mathbf{x}, \alpha\mathbf{y} + \beta\mathbf{z}) = \alpha g(\mathbf{x}, \mathbf{y}) + \beta g(\mathbf{x}, \mathbf{z}), \ \ \forall \mathbf{x} \in X, \ \ \mathbf{y}, \mathbf{z} \in X^*$$

A bilinear form is denoted as

$$g(\mathbf{x}, \mathbf{y}) :=< \mathbf{x}, \mathbf{y} >$$

DEFINITION 8.8 (INNER PRODUCT) *Choose $X^* = X$. The bilinear form of $X \times X$ is called inner product, denoting $< \cdot, \cdot >$ as $(\cdot, \cdot)$, such that*

$$(\cdot, \cdot) : X \times X \to \mathbb{R}$$

with properties:

1 $(\mathbf{x}, \mathbf{x}) \geq 0, \forall \mathbf{x} \in X$ *and* $(\mathbf{x}, \mathbf{x}) = 0$ *iff* $\mathbf{x} = \mathbf{0}$*;*

2 *Symmetry* $(\mathbf{x}, \mathbf{y}) = (\mathbf{y}, \mathbf{x})$*;*

3 *Linearity*

$$(\alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z}) = \alpha(\mathbf{x}, \mathbf{z}) + \beta(\mathbf{y}, \mathbf{z}),$$

and

$$(\mathbf{x}, \alpha\mathbf{y} + \beta\mathbf{z}) = \alpha(\mathbf{x}, \mathbf{y}) + \beta(\mathbf{x}, \mathbf{z}) \ \ \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in X \ \text{ and } \ \alpha\beta \in \mathbb{R}.$$

EXAMPLE 8.9 *Space E^n. Let $X = \mathbb{R}^n$. For $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ and $\mathbf{y} = (y_1, y_2, \cdots, y_n) \in \mathbb{R}^n$, we define an inner product*

$$(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} x_i y_i$$

This particular inner product space is denoted as $E_n = \{\mathbb{R}^n, (\cdot, \cdot)\}$. It generates a norm,

$$\|\mathbf{x}\|_{\ell_2} := \Big(\sum_{i=1}^{n} x_i x_i\Big)^{1/2} = \sqrt{(\mathbf{x}, \mathbf{x})}$$

This norm is called Euclidean norm on $\mathbb{R}^n$. The space is therefore a normed space as well — called n-dimensional Euclidean space, $E_n = \{\mathbb{R}^n, \| \cdot \|_{\ell_2}\}$. One can show that

$$
\begin{array}{ll}
(i) & \|\mathbf{x}\|_{\ell_2} \geq 0, \ \ \forall \mathbf{x} \in E_n \\
 & \|\mathbf{x}\|_{\ell_2} = 0, \quad \Longleftrightarrow \quad \mathbf{x} = 0\ ; \\
(ii) & \|\alpha \mathbf{x}\|_{\ell_2} = |\alpha| \|\mathbf{x}\|_{\ell_2}, \ \ \forall \mathbf{x} E_n, \ \ \alpha \in \mathbb{R} \\
(iii) & \|\mathbf{x}+\mathbf{y}\|_{\ell_2} \leq \|\mathbf{x}\|_{\ell_2} + \|\mathbf{y}\|_{\ell_2} \ \ \leftarrow \ \ \text{triangle inequality;} \\
(iii) & \|(\mathbf{x},\mathbf{y})\|_{\ell_2} \leq \|\mathbf{x}\|_{\ell_2} \|\mathbf{y}\|_{\ell_2} \ \ \leftarrow \ \ \text{Cauchy} - \text{Schwartz inequality;}
\end{array}
$$

Based on the ℓ_2-norm, one can measure the distance between two vectors in E_n,

$$\rho(\mathbf{x},\mathbf{y}) := \|\mathbf{x}-\mathbf{y}\|_{\ell_2};$$

One can also show that

$$
\begin{array}{ll}
(i) & \rho(\mathbf{x},\mathbf{y}) = \rho(\mathbf{y},\mathbf{x}); \\
(ii) & \rho(\mathbf{x},\mathbf{y}) > 0, \ \ \text{and} \ \ \rho(\mathbf{x},\mathbf{y}) = 0, \ \ \text{iff} \ \ \mathbf{x} = \mathbf{y}; \\
(iii) & \rho(\mathbf{x},\mathbf{y}) \leq \rho(\mathbf{x},\mathbf{z}) + \rho(\mathbf{z},\mathbf{y}), \ \ \forall \mathbf{x},\mathbf{y},\mathbf{z} \in E_n
\end{array}
$$

The distance function $\rho(\mathbf{x},\mathbf{y})$ is called a metric, and the associated vector space is called metric space.

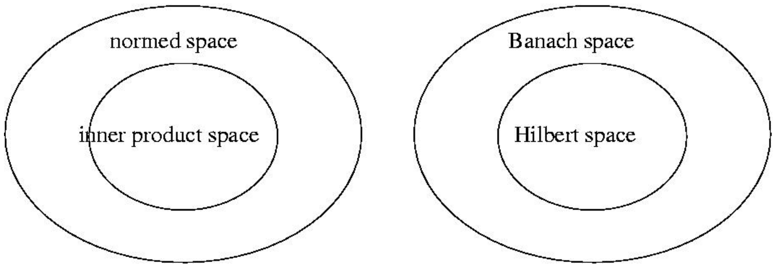


Figure 8.1. Banach space and Hilbert space

REMARK 8.4.2 *1 A normed space or a metric space is not necessarily an inner product space, but an inner product vector space is a normed space, becauce inner product can generate a norm, not vice versa.*

2 A complete normed vector space is called Banach space and a complete inner product space is called Hilber space.

Note that the term completeness means that: A metric space, V, is called complete if every Cauchy sequence $\{f_i\}$ of V has a limit $f \in V$. For a metric space, a Cauchy sequence is one such that $\|v_j - v_k\| \to 0$, as $j, k \to \infty$.

EXAMPLE 8.10 (L^2 SPACE) *Consider a real value function $f(x)$, $x \in [a, b]$. Define an inner product,*

$$(f, g) = \int_a^b f(x)g(x)dx$$

We call the set that contains all $f(x)$ such that

$$\sqrt{\int_a^b f^2(x)dx} < +\infty$$

as space $L^2([a, b])$, where L^2 norm is defined as

$$\|f\|_{L^2([a,b])} = \sqrt{(f, f)} = \sqrt{\int_a^b f^2(x)dx} \tag{8.79}$$

Therefore, $L^2([a, b])$ is an inner product vector space, and of course, normed space (metric space).

EXAMPLE 8.11 (LEBESGUE SPACE ($L^p(\Omega)$)) *Let Ω be an open set in $\mathbb{R}^n$. For $1 < p < \infty$, one can define a L_p-norm for a measurable function f,*

$$\|f\|_{L^p(\Omega)} := \Big(\int_\Omega |f(x)|^p dx\Big)^{1/p}$$

and a $Lebesgue$ space is defined as

$$L^p(\Omega) := \Big\{ f \;\Big|\; \|f\|_{L^p(\Omega)} < \infty \Big\}$$

It has the following properties,

$$\begin{array}{ll}
(i) & \|f\|_{L^p(\Omega)} \geq 0, \;\; \|f\|_{L^p(\Omega)} = 0, \;\; \Rightarrow \;\; f = 0 \;\; \text{almost everywhere;} \\
(ii) & \|cf\|_{L^p(\Omega)} \leq |c|\|f\|_{L^p(\Omega)}, \;\; \forall f \in L^p(\Omega), \;\; c \in \mathbb{R} \\
(iii) & \|f + g\|_{L^p(\Omega)} \leq \|f\|_{L^p(\Omega)} + \|g\|_{L^p(\Omega)} \;\; \leftarrow \text{Minkowski's inequality} \\
(iv) & \text{For } \; 1 \leq p, q \leq \infty, \; \text{such that } \; \dfrac{1}{p} + \dfrac{1}{q} = 1, \\
 & \text{if } f \in L^p(\Omega) \text{ and } gL^q(\Omega), \;\; \text{then for finite } \Omega, f, g \in L^1(\Omega), \; \text{and} \\
 & \|fg\|_{L^1(\Omega} \leq \|f\|_{L^p(\Omega)}\|g\|_{L^q(\Omega)}, \;\; \leftarrow \text{Holder's inequality}
\end{array}$$

In particular, $p = q = 2$, then $f \cdot g \in L^1(\Omega)$ because

$$\int_\Omega |f(x)g(x)|dx \leq \|f\|_{L^2(\Omega)}\|g\|_{L^2(\Omega)}$$

Note that in general $L^p(\Omega)$ is not an inner product space, except $p = 2$. $L^p(\Omega)$ is, nevertheless, a complete normed space, therefore, a Banach space and $L^2(\Omega)$ is a Hilber space.

EXAMPLE 8.12 (SOBOLEV SPACE) *Define Soblev norm*

$$\|f\|_{W^k_p(\Omega)} = \Big(\sum_{\alpha=0}^{k} \|D^\alpha f\|^p_{L^p(\Omega)}\Big)^{1/p}$$

Note that the Sobolev norm is not generated by an inner product in general.

A Sobolev space is defined as

$$W^k_p(\Omega) = \{f \;\Big|\; \|f\|_{W^k_p(\Omega)} < \infty\}$$

For $p = 2$, Sobolev spaces become inner product spaces. In particular,

1 *For $p = 2, k = 0$, $W^0_2(\Omega) = L^2(\Omega)$,*

$$(f, g)_{L^2(\Omega)} = \int_\Omega f(\mathbf{x})g(\mathbf{x})dV$$

2 *For $p = 2, k = 1$, $W^1_2(\Omega) = H^1(\Omega)$,*

$$(f, g)_{H^1(\Omega)} = \int_\Omega \Big[f(\mathbf{x})g(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot \nabla g(\mathbf{x})\Big]dV$$

and

$$\|f\|_{H^1(\Omega} = \sqrt{\int_\Omega \Big[f(\mathbf{x})^2 + \nabla f(\mathbf{x}) \cdot \nabla f(\mathbf{x})\Big]dV}$$

3 *For $p = 2, k = 2$, $W^2_2(\Omega) = H^2(\Omega)$,*

$$(f, g)_{H^2(\Omega)} = \int_\Omega \Big[f(\mathbf{x})g(\mathbf{x}) + \nabla f(\mathbf{x}) \cdot \nabla g(\mathbf{x}) + \nabla \otimes \nabla f(\mathbf{x}) : \nabla \otimes \nabla g(\mathbf{x})\Big]dV$$

and

$$\|f\|_{H^2(\Omega} = \sqrt{\int_\Omega \Big[f(\mathbf{x})^2 + \nabla f(\mathbf{x}) \cdot \nabla f(\mathbf{x}) + \nabla \otimes \nabla f(\mathbf{x}) : \nabla \otimes \nabla f(\mathbf{x})\Big]dV}$$

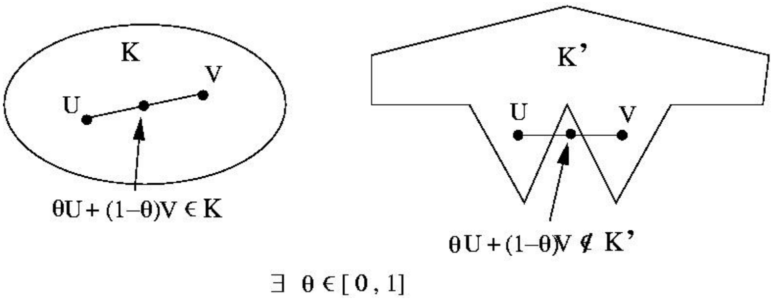


Figure 8.2. Convex set and non-convex set in $\mathbb{R}^2$

8.4.1 Concept of convexity

DEFINITION 8.13 *Let U be a linear vector space over $\mathbb{R}$. A subset (subspace) $\mathcal{K} \subset U$ is said to be convex, if it contains the line segment between any two of its elements, i.e.*

$$\theta \mathbf{u} + (1-\theta)\mathbf{v} \in \mathcal{K}, \ \ \forall \mathbf{u}, \mathbf{v} \in \mathcal{K}$$

where $\theta \in [0, 1]$.

EXAMPLE 8.14 *Let $U = \mathbb{R} \times \mathbb{R}$, and $\mathcal{K} \in U$. We say $\mathcal{K}$ is convex, when $\mathbf{u}_1 = (x_1, x_2)$, $\mathbf{u}_2 = (y_1, y_2) \in \mathcal{K}$, then $\theta\mathbf{u}_1 + (1-\theta)\mathbf{u}_2 \in \mathcal{K}$, $\theta \in [0, 1]$. We say $\mathcal{K}$ is not convex, for any $\mathbf{u}_1$, $\mathbf{u}_2 \in \mathcal{K}$, if $\exists \mathbf{u}_\theta \in \theta\mathbf{u}_1 + (1-\theta)\mathbf{u}_2$ but $\mathbf{u}_\theta \notin \mathcal{K}$. A graphic illustration is demonstrated in Fig. (8.2).*

DEFINITION 8.15 (CONVEX AND CONCAVE FUNCTIONALS) *1 A functional $P : \mathcal{U} \to \mathbb{R}$ is said to be convex on $\mathcal{U}$ if*

$$P(\theta\mathbf{u}_1 + (1-\theta)\mathbf{u}_2) \leq \theta P(\mathbf{u}_1) + (1-\theta)P(\mathbf{u}_2), \ \ \forall \mathbf{u}_1, \mathbf{u}_2 \in U, \ \ \forall \theta \in [0, 1]$$

whenever the right-hand side is defined.

2 *P is said to be strictly convex if the strict form of the inequality holds for any $\mathbf{u}_1 \neq \mathbf{u}_2$;*

3 *P is said to be concave if $-P$ is convex.*

EXAMPLE 8.16 *Let $\mathcal{U} = \mathbb{R}$ and $P(x) = (x-a)^2$.*

EXAMPLE 8.17 *Consider a 1D elastic string, $I = [0, \ell]$. Let $\mathcal{U} = \mathcal{E}$ and $\mathcal{U}^* = \mathcal{S}$ where*

$$\mathcal{E}\{\epsilon \;\Big|\; \epsilon \in L^\alpha(I), \epsilon = \frac{du}{dx}\}$$

$$\mathcal{S}\{\sigma \;\Big|\; \sigma \in L^\beta(I), \frac{d\sigma}{dx} = 0\}$$

$$1 < \alpha, \beta < \infty, \ \text{ and } \ \frac{1}{\alpha} + \frac{1}{\beta} = 1.$$

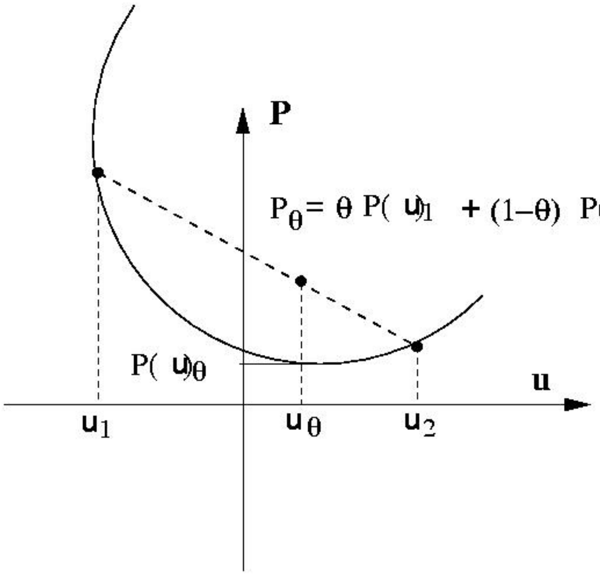


Figure 8.3. An example of convex function.

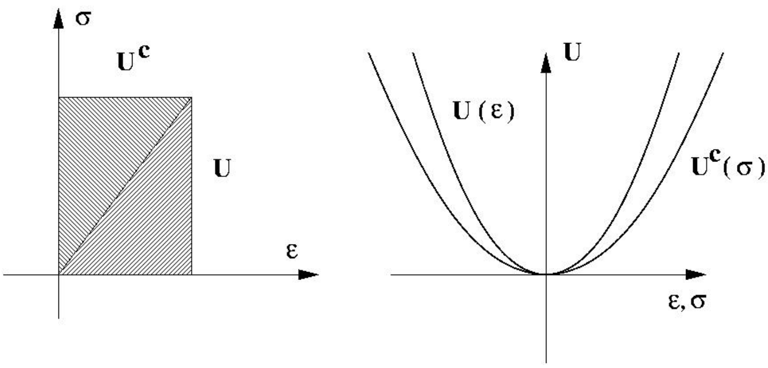


Figure 8.4. Strain energy density and complementray strain energy

Define

$$
\begin{aligned}
&U : E \rightarrow \mathbb{R}, \qquad\quad U(\epsilon) = \int_0^{\epsilon} \sigma(\tilde{\epsilon})d\tilde{\epsilon} \\
&U^c : E^* = S \rightarrow \mathbb{R}, \;\; U^c(\sigma) = \int_0^{\sigma} \epsilon(\tilde{\sigma})d\tilde{\sigma}
\end{aligned}
$$

Both strain energy density and complementary strain energy density are convex, and they are plotted in Fig. (8.4).

8.4.2 Gâteaux variation and convex functional

The Gâteaus variation of a functional in a linear space is the generalized directional derivative of a real-value function in vector calculus.

DEFINITION 8.18 (GÂTEAUX VARIATION) *1 Let $P : \mathcal{U} \to \mathbb{R}$ be a real-valued functional and $\mathcal{U}_a \subset \mathcal{U}$ a subspace. For a given $\bar{\mathbf{u}} \in \mathcal{U}_a$, if the limit,*

$$\delta U(\bar{\mathbf{u}}, \mathbf{u}) := \lim_{\lambda \to 0^+} \frac{P(\bar{\mathbf{u}} + \lambda \mathbf{u}) - P(\bar{\mathbf{u}})}{\lambda}, \quad \forall \mathbf{u} \in \mathcal{U}_a$$

exists as $\lambda \to 0^+$ (i.e. $\lambda \to 0, \lambda > 0$), then $\delta P(\bar{\mathbf{u}}; \mathbf{u}) \in \mathbb{R}$ is called the Gâteaus variation of P at $\bar{\mathbf{u}}$ in the direction of $\mathbf{u}$.

2 If the Gâteau variation is a linear operator in $\mathbf{u}$ such that

$$\delta P(\bar{u}, \mathbf{u}) =< \mathbf{u}, DP(\bar{\mathbf{u}}) >, \quad \forall \mathbf{u} \in \mathcal{U}_a$$

we say that P is Gâteaux differentiable at $\bar{\mathbf{u}}$. The linear operator $DP(\bar{\mathbf{u}}) : \mathcal{U}_a \to \mathcal{U}^*$, which generally depends on $\bar{\mathbf{u}}$, is called the Gâteaux derivative of P at $\bar{\mathbf{u}}$.

3 The functional $P : \mathcal{U} \to \mathbb{R}$ is said to be Gâteaux differentiable on $\mathcal{U}_a$ if it is Gâteaux differentiable at each $\mathbf{u} \in \mathcal{U}_a$.

Note that

$$\begin{aligned} \delta P(\bar{\mathbf{u}}, \mathbf{u}) &= \frac{d}{d\lambda} P(\bar{\mathbf{u}} + \lambda \mathbf{u}) \Big|_{\lambda=0} \\ \frac{\delta P}{\delta u} &:= DP(\bar{\mathbf{u}}) \end{aligned}$$

Question: *why are convex functionals so special ?* The following theorem answers this question:

THEOREM 8.19 *If $P : \mathcal{U}_k \subset \mathcal{U} \to \mathbb{R}$ is Gâteaux differentiable, then, the following statements are equivalent to each other*

$$\begin{aligned} &(S1) \qquad P : \mathcal{U}_k \subset \mathcal{U} \to \mathbb{R} \text{ is convex;} \\ &(S2) \qquad P(\mathbf{v}) - P(\mathbf{u}) \geq < \mathbf{v} - \mathbf{u}, DP(\mathbf{u}) >, \ \forall \mathbf{v}, \mathbf{u} \in \mathcal{U}_k \\ &(S3) \qquad < \mathbf{v} - \mathbf{u}, DP(\mathbf{v}) - DP(\mathbf{u}) > \geq 0 \ , \ \ \forall \mathbf{v}, \mathbf{u} \in \mathcal{U}_k \end{aligned}$$

REMARK 8.4.3 *The statement (S3) shows that Gâteaux derivative of a convex function is a monotone operator of $\mathcal{U}$ into $\mathcal{U}^*$. By the mean value theorem,*

$$< \mathbf{v} - \mathbf{u}, DP(\mathbf{v}) - DP(\mathbf{u}) > = < \mathbf{v} - \mathbf{u}, D^2 P(\bar{\mathbf{u}}) \cdot (\mathbf{v} - \mathbf{u}) > \geq 0$$

where $\bar{\mathbf{u}} = \mathbf{v} + \theta(\mathbf{v} - \mathbf{u}), \ \theta \in [0, 1]$.

Hencea, a sufficient condition for P being convex on $\mathcal{U}$ is that

$$D^2 P(\mathbf{u}) \geq 0, \ \ \forall \mathbf{u} \in \mathcal{U}_k$$

Recall the total potential energy for a linear elastic solid is

$$\begin{aligned}
\Pi(\mathbf{u}, \nabla\mathbf{u}) &= \int_V U(\boldsymbol{\epsilon}) dV - \int_{\Gamma_t} t_i^0 u_i dS \\
\delta\Pi(\mathbf{u}, \nabla\mathbf{u}) &= \int_V \frac{\partial U}{\partial \epsilon_{ij}} \delta\epsilon_{ij} dV - \int_{\Gamma_t} t_i^0 \delta u_i dS \\
\delta^2\Pi(\mathbf{u}, \nabla\mathbf{u}) &= \int_V \frac{\partial^2 U}{\partial \epsilon_{ij} \partial \epsilon_{k\ell}} \delta\epsilon_{ij} \delta\epsilon_{k\ell} dV = \int_V C_{ijk\ell} \delta\epsilon_{ij} \delta\epsilon_{k\ell} dV \geq 0 \ .
\end{aligned}$$

This is to say that if elastic tensor is positive definite, the elastic potential energy is convex. Similar statement can be made for complementary potential energy, if the compliance tensor is positive definite.

8.4.3 Primal variational problems

We consider the following primal variational problems:
Let $P : \mathcal{U}_\kappa \subset U \to \mathbb{R}$ be a given functional.

1 The infimum (or $\inf$) primal variational problems is to find a global minimizer $\tilde{\mathbf{u}} \in \mathcal{U}_\kappa$ such that

$$\left(\mathcal{P}_{inf}\right) : \quad P(\tilde{\mathbf{u}}) = \inf P(\mathbf{u}), \ \ \forall \mathbf{u} \in \mathcal{U}_\kappa$$

2 The supremum (or sup) primal problem is to find a global maximizer $\tilde{\mathbf{u}} \in \mathcal{U}_\kappa$ such that

$$\left(\mathcal{P}_{sup}\right) : \quad P(\tilde{\mathbf{u}}) = \sup P(\mathbf{u}), \ \ \forall \mathbf{u} \in \mathcal{U}_\kappa$$

3 The stationar (or sta) primal variational problem is to find a stationary point $\mathbf{u} \in \mathcal{U}_\kappa$ such that

$$\left(\mathcal{P}_{sta}\right) : \quad P(\tilde{\mathbf{u}}) = \mathrm{sta}\, P(\mathbf{u}), \ \ \forall \mathbf{u} \in \mathcal{U}_\kappa$$

REMARK 8.4.4 *1 A stationary point is also called critical point. The critical point condition,*

$$\delta P(\tilde{\mathbf{u}}, \mathbf{u}) = 0, \quad \forall \mathbf{u} \in \mathcal{U}_\kappa$$

leads to the Euler-Lagrange equation.

2 *The problem* $(\mathcal{P}_{\inf})$ *is called realisable if there exists a vector* $\tilde{\mathbf{u}} \in \mathcal{U}_\kappa$ *such that the infimum of* P *is achieved at* $\tilde{\mathbf{u}}$ *and is not* $+\infty$*. Then* $\tilde{\mathbf{u}}$ *is called the minimizer of* $(\mathcal{P}_{inf})$ *and we write* $P(\tilde{\mathbf{u}}) = \min_{\mathbf{u}\in\mathcal{U}_\kappa} P(\mathbf{u})$.

Similarly, a vector $\tilde{\mathbf{u}} \in \mathcal{U}_\kappa$ *is called the maximizer of* $(\mathcal{P}_{sup})$ *if the supermum is achieved at* $\tilde{\mathbf{u}}$ *and is not* $+\infty$*. We write* $\mathcal{P}(\tilde{\nu}) = \max_{\mathbf{u}\in\mathcal{U}_\kappa}(\mathbf{u})$.

EXAMPLE 8.20 *The real-value function,* $P(x) = \exp(x)$ *is convex on* $\mathcal{U} = \mathbb{R}$ *and*

$$\inf P(x) = 0, \quad \sup_{x\in\mathcal{U}} P(x) = +\infty$$

Howeverm on the closed interval, $U_\kappa = [a, b]$ *with* $-\infty < a < b < +\infty$*, the two* $inf-$ *and* $sup-$ *problmes are realisable and*

$$\begin{aligned} \inf_{x\in\mathcal{U}_\kappa} P(x) &= \min_{x\in\mathcal{U}_\kappa} P(x) = P(a) = e^a, \\ \sup_{x\in\mathcal{U}_\kappa} P(x) &= \min_{x\in\mathcal{U}_\kappa} P(x) = P(b) = e^b\,. \end{aligned}$$

8.5 Legendre Transformation and Duality

In continuum mechanics, for a given stored-energy density $U(\boldsymbol{\epsilon})$ such that the strain-stress relation $\boldsymbol{\sigma} = \dfrac{\partial U}{\partial \boldsymbol{\epsilon}}$ is invertible, then one can define so-called complementary energy density of $U^c(\boldsymbol{\sigma})$ by

$$U^c(\boldsymbol{\sigma}) = \boldsymbol{\sigma} : \boldsymbol{\epsilon}(\boldsymbol{\sigma}) - U(\boldsymbol{\epsilon}(\boldsymbol{\sigma})) \tag{8.80}$$

Note that here

$$U = U(\boldsymbol{\epsilon}) : \qquad \mathcal{E} \to \mathbb{R} \tag{8.81}$$

$$U^c = U^c(\boldsymbol{\sigma}) : \qquad \mathcal{S} \to \mathbb{R} \tag{8.82}$$

$$< \boldsymbol{\epsilon}, \boldsymbol{\sigma} >= \boldsymbol{\sigma} : \boldsymbol{\epsilon} : \qquad \mathcal{E} \times \mathcal{E}^* \to \mathbb{R} \tag{8.83}$$

where the space $\mathcal{S}$ may be viewed as $\mathcal{E}^*$.

In mathematics, this is the well-known Legendre transformation. Generally speaking, the classical Legendre transformation can be viewed as a conversion of one continuous real-valued function into another one. If the transformation is reversible, then we say that each function is the dual of the other. The reversible Legendre transformation is also called the Legendre conjugate transformation, or simply the Legendre transformation.

Let $E = \mathbb{R}^n = E^*$. The element $\boldsymbol{\epsilon} = \{\epsilon_i\} \in E$ and $\boldsymbol{\sigma} = \{\sigma_i\} \in E^*, (i = 1, 2, \cdots, n)$ are vectors in $\mathbb{R}^n$. The bilinear form

$$< \boldsymbol{\epsilon}, \boldsymbol{\sigma} >= \boldsymbol{\epsilon} \cdot \boldsymbol{\sigma} = \sum_{i=1}^{n} \epsilon_i \boldsymbol{\sigma}_i \tag{8.84}$$

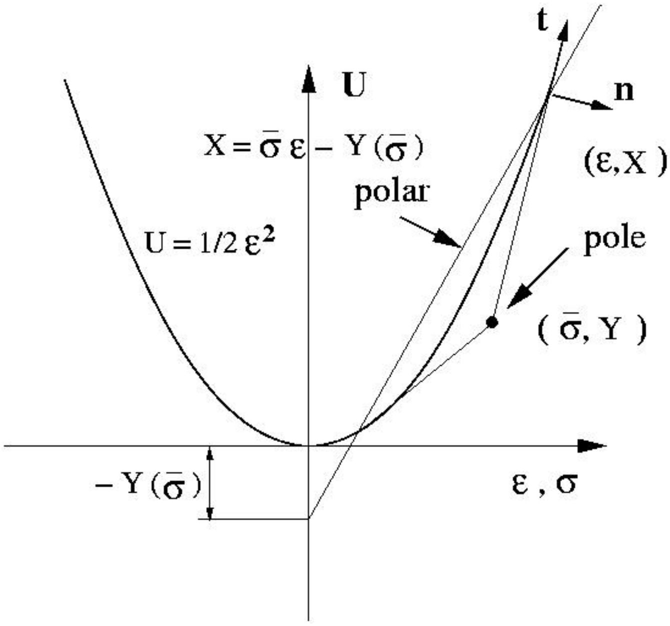


Figure 8.5. Duality between the pole and polar

is then the inner product on $\mathbb{R}^n$.

Let $U : E \to \mathbb{R}$ be a real-valued function. Its graph,

$$\{(\boldsymbol{\epsilon}, X) \in \mathbb{R}^{n+1} \;\Big|\; X = U(\boldsymbol{\epsilon})\}$$

is a manifold (or hypersurface) in $\mathbb{R}^{n+1}$.

Let any particular point $(\boldsymbol{\sigma}, Y) \in \mathbb{R}^{n+1}$ be called the pole. Then the linear function

$$X(\boldsymbol{\epsilon}) = \boldsymbol{\epsilon} \cdot \boldsymbol{\sigma} - Y \tag{8.85}$$

is called the polar, which is a hyperplane in $\mathbb{R}^{n+1}$.

Thus, given a pole at a finite point, the polar is well-defined by (8.85), Conversely, given a polar of finite slope, a finite pole can be read off from Eq. (8.85). This correspondence is called the duality between points and planes.

The duality comes to live when the graphi of a paraboloid is blended into the picture.

Theorem 8.21 (Duality between the pole and polar) *(T1) If the pole is outside the paraboloid, the points of contact of tangents drawn from the pole to the paraboloid lie on the polar.*

(T2) If the pole is inside of the paraboloid, the polar lies outside it.

Proof:

We only prove the theroem in $\mathbb{R}^2$, which has the full flavor of a rigorous proof.

We first show (T1). The tangential vector from the pole to the paraboloid is

$$\mathbf{t} = (\bar{\sigma} - \epsilon, Y(\bar{\sigma}) - X)$$

the normal vector of graph $G = U - \frac{1}{2}\epsilon^2 = 0$ is

$$\mathbf{n} = \Big(\frac{\partial G}{\partial \epsilon}, \frac{\partial G}{\partial U}\Big) = (-\epsilon, 1)$$

We want show that the contact point is in the polar : $X(\epsilon) = \bar{\sigma}\epsilon - Y(\bar{\sigma})$.

Consider the condition $\mathbf{t} \cdot \mathbf{n} = 0$.

$$\begin{aligned}
\mathbf{t} \cdot \mathbf{n} &= (\bar{\sigma} - \epsilon, Y - X)(-\epsilon, 1) \\
&= -\epsilon\bar{\sigma} + \epsilon^2 + Y - X \\
&= -\epsilon\bar{\sigma} + 2X + Y - X = -\epsilon\bar{\sigma} + X + Y = 0
\end{aligned}$$

We just showed that $X = \epsilon\bar{\sigma} - Y$.

We now show (T2). Suppose the pole is inside the paraboloid. We want to show that the polar is outside the paraboloid region.

Assume that part of the polar is inside or no the paraboloid, i.e.

$$X \geq \frac{1}{2}\epsilon^2$$

Since the pole is also inside the paraboloid, i.e.

$$Y(\bar{\sigma}) > \frac{1}{2}\bar{\sigma}^2$$

Therefore,

$$\begin{aligned}
X + Y(\bar{\sigma}) &> \frac{1}{2}\Big(\bar{\sigma}^2 + \epsilon^2\Big) \\
\bar{\sigma}\epsilon &> \frac{1}{2}\Big(\bar{\sigma}^2 + \epsilon^2\Big) \\
0 &> \frac{1}{2}\Big(\bar{\sigma}^2 - 2\bar{\sigma}\epsilon + \epsilon^2\Big) = \frac{1}{2}(\bar{\sigma} - \epsilon)^2 > 0
\end{aligned}$$

which leads to contradiction. Hence, polar must be outside the paraboloid, if the pole is inside the paraboloid. ♣

DEFINITION 8.22 (REGULAR POINT AND REGULAR DOMAIN) *Let* $U : E \to \mathbb{R}$ *be a piecewise* C^2 *function.*

(D1) A regular point of the function $U(\boldsymbol{\epsilon})$ is a point $\boldsymbol{\epsilon} \in E$ where the determinant of the Hessian matrix $D^2U = \{\dfrac{\partial^2 U}{\partial \epsilon_i \partial \epsilon_j}\}$ satisfies,

$$det \left\{ \frac{\partial^2 U}{\partial \epsilon_i \partial \epsilon_j} \neq 0, \ \ \text{or} \ \ \pm \infty \right\}$$

(D2) A regular domain, denoted by E_r is a continuous subset of regular points.

Now we let $U^c : \mathbb{R}^n \to \mathbb{R}$ be a given continuous function such that the graph,

$$G_{U^c} = \{(\boldsymbol{\sigma}, Y) \in \mathbb{R}^n \ \Big| \ Y = U^c(\boldsymbol{\sigma}), \boldsymbol{\sigma} \in \mathbb{R}^n\}$$

of U^c is a continuous surface in $\mathbb{R}^{n+1}$.

When the pole, $(\boldsymbol{\sigma}, Y)$, moves on the graph of U^c, each point on G_{U^c} is corresponding to a polar hyperplane. The collective of these polars hyperplanes will envelop another continuous surface, the graph of $X = U(\boldsymbol{\epsilon})$, described as $U : \mathbb{R}^n \to \mathbb{R}$, which is the conjugate Legendre pair of $U^c(\sigma)$. This is the geometric interpretation of Legendre transformation. In other words, the correspondence between the functions $U(\boldsymbol{\epsilon})$ and $U^c(\boldsymbol{\sigma})$ is called Legendre transformation.

Now we state the important Legendre Dulaity theorem.

THEOREM 8.23 (LEGENDRE DUALITY THEOREM) *Let $U(\boldsymbol{\epsilon}) \in C^2(E)$. If $E_r \subset E$ is an open, finite subset of the regular domain of U and $E_r^* \subset \mathbb{R}^n$ is the range of the mapping $DU : E_r \to E^*$. Then there exists a unique C^2 function U^c on E^*, which is dual to U on E_r in the sense that the Legendre duality relates*

$$U(\boldsymbol{\epsilon}) + U^c(\boldsymbol{\sigma}) = \boldsymbol{\sigma} \cdot \boldsymbol{\epsilon} \Leftrightarrow \boldsymbol{\sigma} = \partial U(\boldsymbol{\epsilon}), \Leftrightarrow \boldsymbol{\epsilon} = \partial U^c(\sigma)$$

hold. Moreover, for $(\boldsymbol{\epsilon}, \boldsymbol{\sigma}) \in E_r \times E_r^*$ satisfying above relationship,

$$\frac{\partial^2 U}{\partial \epsilon_i \partial \epsilon_k} \frac{\partial^2 U^c}{\partial \sigma_k \partial \sigma_j} = \delta_{ij} \ .$$

The proof of this theorem is basically application of implicit function theorem. It is omitted here. The readers who are interested in the proof may consult Gao [2000].

Now we move to the essentail technical ingradient of convex analysis.

THEOREM 8.24 (DUALITY BETWEEN THE REGULAR MANIFOLDS) *Let U and U^c be Legendre dual functions over the duality domain E and E^* respectively.*

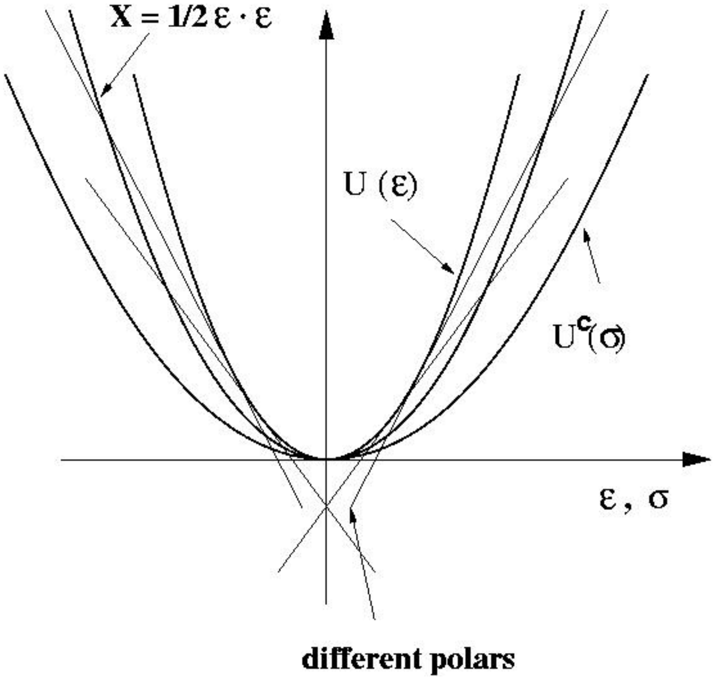


Figure 8.6. Geometric interpretation of Legendre transformation

(S1) If U is convex on E, U^c is convex on E^* and

$$U^c(\boldsymbol{\sigma}) = \max_{\boldsymbol{\epsilon}\in E}\{\boldsymbol{\sigma}\cdot\boldsymbol{\epsilon} - U(\boldsymbol{\epsilon})\}$$

(S2) If U is concave on E, U^c is concave on E^* and

$$U^c = \min_{\boldsymbol{\epsilon}\in E}\{\boldsymbol{\sigma}\boldsymbol{\epsilon} - U(\boldsymbol{\epsilon})\}$$

Proof;

For simplicity, we only prove it for case $E \subset \mathbb{R}$, which contains the enssential substance of a general, rigorous proof.

Since $\sigma = \dfrac{\partial U}{\partial \epsilon}$, by Taylor expansion,

$$\sigma = \frac{\partial U}{\partial \epsilon}\Big|_{\epsilon=\bar{\epsilon}} + \frac{\partial^2 U}{\partial \epsilon^2}\Big(*\Big)(\epsilon - \bar{\epsilon}) \tag{8.86}$$

where

$$\frac{\partial^2 U}{\partial \epsilon^2}\Big(*\Big) = \frac{\partial^2 U}{\partial \epsilon^2}\Big|_{\epsilon=\bar{\epsilon}+\theta\Delta\epsilon}$$

and $0 \leq \theta \leq 1$.

Eq. (8.86) can be rewritten as

$$(\sigma - \bar{\sigma}) = +\frac{\partial^2 U}{\partial \epsilon^2}\Big(*\Big)(\epsilon - \bar{\epsilon}) \tag{8.87}$$

By the same token, because of $\epsilon = \dfrac{\partial U^c}{\partial \sigma}$, one can have

$$(\epsilon - \bar{\epsilon}) = +\frac{\partial^2 U}{\partial \epsilon^2}\Big(*\Big)(\sigma - \bar{\sigma}) \tag{8.88}$$

where

$$\frac{\partial^2 U^c}{\partial \sigma^2}\Big(*\Big) = \frac{\partial^2 U}{\partial \sigma^2}\Big|_{\epsilon = \bar{\sigma} + \theta \Delta \sigma}$$

and $0 \leq \theta \leq 1$. Therefore,

$$\begin{aligned}(\sigma - \bar{\sigma})(\epsilon - \bar{\epsilon}) &= \frac{\partial^2 U}{\partial \epsilon^2}\Big(*\Big)(\epsilon - \bar{\epsilon})^2 \\ &= \frac{\partial^2 U^c}{\partial \sigma^2}\Big(*\Big)(\sigma - \bar{\sigma})^2 \end{aligned} \tag{8.89}$$

Eq. (8.89) indicates that if $\dfrac{\partial^2 U}{\partial \epsilon^2}\Big(*\Big)$ is positive definite, $\dfrac{\partial^2 U^c}{\partial \sigma^2}\Big(*\Big)$ is also positive; whereas if $\dfrac{\partial^2 U}{\partial \epsilon^2}\Big(*\Big)$ is negative definite, $\dfrac{\partial^2 U^c}{\partial \sigma^2}\Big(*\Big)$ is also negative definite, or both being indefinite.

To prove the Legendre inequality, we consider a special 1D example, $U(\epsilon) = \dfrac{1}{2}k_0\epsilon^2, k_0 > 0$.

For a given point $\bar{\epsilon}$ on horizontal axis, the associated stress $\bar{\sigma} = k_0\bar{\epsilon}$ is the slope of the polar, the straight line $X = \bar{\sigma} - Y$, which is tangent to the graphy of U at $\bar{\epsilon}$ (see Fig. (8.7).

Therefore, point $(\bar{\epsilon}, U(\bar{\epsilon})$ is in both polar $X = \bar{\sigma}\epsilon - Y$ and on $U = 1/2k_0\epsilon^2$, which is to say that $X(\bar{\epsilon}) = U(\bar{\epsilon})$ and

$$Y = \bar{\sigma}\bar{\epsilon} - U(\bar{\epsilon}) =: U^c(\bar{\sigma})$$

For any given $\epsilon \in E_r$, we define a continuous function,

$$y(\epsilon) = \bar{\sigma}\epsilon - U(\epsilon)$$

we want to show that $Y = U^c(\bar{\sigma}) \geq y(\epsilon)$.

Since the polar $X(\epsilon)$ is always below the parabola $(U(\epsilon) \geq X(\epsilon)$,

$$\begin{aligned} U(\epsilon) - X \geq 0 \quad &\Rightarrow \quad U(\epsilon) - (\bar{\sigma}\epsilon - Y) \geq 0 \\ &\Rightarrow \quad Y \geq \bar{\sigma}\epsilon - U(\epsilon) \end{aligned}$$

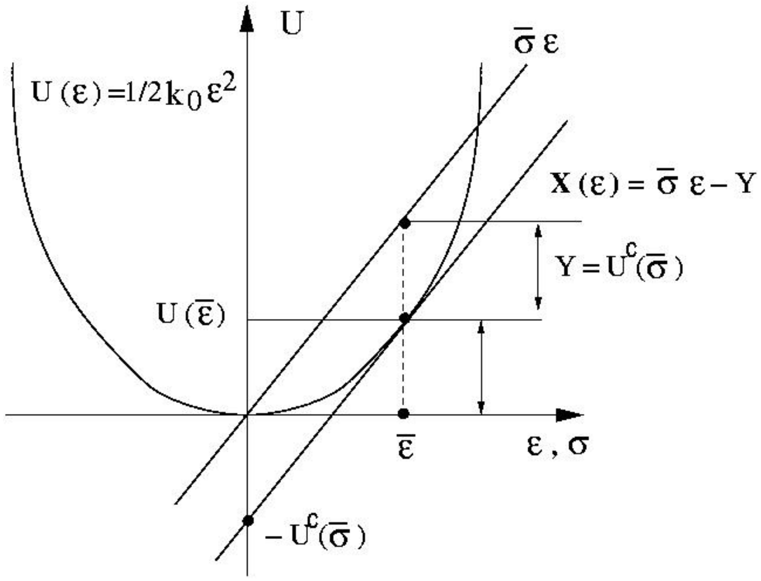


Figure 8.7. Legendre transformation

Since $U(\epsilon)$ is convex, $y(\epsilon)$ is then concave because $\dfrac{\partial^2 y}{\partial \epsilon^2} < 0$. It then takes its maximum value at $\bar{\epsilon}$ because $y^{'}(\bar{\epsilon}) = 0$. That is

$$Y = U^c(\bar{\sigma}) = \max_{\epsilon \in E_r} \{\bar{\sigma}\epsilon - U(\epsilon)\} \tag{8.90}$$

One drop the bar on σ, because domain of $\bar{\sigma}$ is the same as σ.

Similarly, for concave function, one can show that

$$U^c(\sigma) = \min_{\epsilon \in E_r} \{\sigma\epsilon - U(\epsilon)\}$$

♣

REMARK 8.5.1 *In the infinite-dimensional space E, Eq. (8.90) is called Legendre-Fenchel transformation, and it reads as*

$$U^*(\boldsymbol{\sigma}) = \sup_{\boldsymbol{\epsilon} \in E} \{\boldsymbol{\sigma} \cdot \boldsymbol{\epsilon} - U(\boldsymbol{\epsilon})\}$$

where the superscript $*$ replaces the superscript c meaning as the dual function.

Accordingly, if U is concave, its Legendre-Fenchel conjugate is defined as

$$U^*(\boldsymbol{\sigma}) = \inf_{\boldsymbol{\epsilon} \in E} \{\boldsymbol{\sigma} \cdot \boldsymbol{\epsilon} - U(\boldsymbol{\epsilon})\}$$

The reason we add the name Fenchel is because when U is defined as

$$U : E \to \mathbb{R} \bigcup \{+\infty\}$$

the transformation

$$U^*(\boldsymbol{\sigma}) = \sup_{\boldsymbol{\epsilon} \in E} \{\boldsymbol{\sigma} \cdot \boldsymbol{\epsilon} - U(\boldsymbol{\epsilon})\}$$

is called the Fenchel transformation.

8.6 Legendre-Fenchel transformation in linear elasticity

In a classical paper (Hill [1965]), Hill illustrated the Legendre-Fenchel transformation in linear elastic system and extend the use of classical minimum potential energy principle and minimum complementary energy principle to micromechanics.

Consider the prescribed displacement boundary condition (prescribed macro strain condition),

$$\mathbf{u}^0 = \mathbf{x} \cdot \boldsymbol{\epsilon}^0, \ \ \forall \mathbf{x} \in \partial V$$

Under such condition, we have shown previously that

$$\boldsymbol{\epsilon}^0 =< \boldsymbol{\epsilon} >=< \tilde{\boldsymbol{\epsilon}} >, \ \ \forall \boldsymbol{\epsilon} \in \mathcal{E}$$

where $\mathcal{E}$ is the space of compatible strain.

Therefore, the potential energy and complementary energy take the form

$$\begin{aligned} \Pi^c &= \frac{1}{2} \int_\Omega D_{ijk\ell} \sigma_{ij} \sigma_{k\ell} dV - \int_{\partial V} x_k \epsilon^0_{ki} \sigma_{ij} n_j dS \\ &= \frac{1}{2} \int_\Omega D_{ijk\ell} \sigma_{ij} \sigma_{k\ell} dV - \int_V \Big[\delta_{kj} \epsilon^0_{ki} \sigma_{ij} + \underbrace{x_k \epsilon^0_{ki} \sigma_{ij,j}}_{=0} \Big] dV \\ &= \frac{1}{2} \int_\Omega D_{ijk\ell} \sigma_{ij} \sigma_{k\ell} dV - \int_V \epsilon^0_{ij} \sigma_{ij} dV \end{aligned}$$

Based on minimum complementary energy principle, for any statically admissible stress field, $\forall \boldsymbol{\sigma} \in \mathcal{S}$,

$$\begin{aligned} \Pi^c(\boldsymbol{\sigma}) &\geq \frac{1}{2} \int_V D_{ijk\ell} \tilde{\sigma}_{ij} \tilde{\sigma}_{k\ell} - \int_{\partial V} u^0_i \tilde{\sigma}_{ij} n_j dS \\ &= \frac{1}{2} \int_V D_{ijk\ell} \tilde{\sigma}_{ij} \tilde{\sigma}_{k\ell} - \int_{\partial V} \epsilon^0_{ij} \tilde{\sigma}_{ij} dV \\ &= -\frac{1}{2} \int_C C_{ijk\ell} \tilde{\epsilon}_{ij} \tilde{\epsilon}_{k\ell} dV \end{aligned}$$

where $\tilde{\boldsymbol{\sigma}}$ and $\tilde{\boldsymbol{\epsilon}}$ is the real solution. In the last line the equality under prescribed macros strain,

$$\int_V \epsilon^0_{ij} \tilde{\sigma}_{ij} dV = \int_V \tilde{\epsilon}_{ij} \tilde{\sigma}_{ij} dV$$

is used.

Therefore,

$$\frac{1}{2}\int_V C_{ijk\ell}\tilde{\epsilon}_{ij}\tilde{\epsilon}_{k\ell}dV \geq \int_V \sigma_{ij}\epsilon^0_{ij}dV - \frac{1}{2}\int_V D_{ijk\ell}\sigma_{ij}\sigma_{k\ell}dV$$

which is essentially

$$W(\tilde{\epsilon}) = \sup_{\{\boldsymbol{\sigma}\in\mathcal{S}\}} \left\{\boldsymbol{\epsilon}^0 :< \boldsymbol{\sigma} > -W^c(\boldsymbol{\sigma})\right\} \tag{8.91}$$

where

$$\begin{aligned} W(\boldsymbol{\epsilon}) &= \frac{1}{2V}\int_V C_{ijk\ell}\epsilon_{ij}\epsilon_{k\ell}dV \\ W^c(\boldsymbol{\sigma}) &= \frac{1}{2V}\int_V C_{ijk\ell}\sigma ij\sigma_{k\ell}dV \end{aligned}$$

One may further tighten the bound

$$W(\tilde{\epsilon}) = \sup_{\{<\boldsymbol{\sigma}>:\boldsymbol{\sigma}\in\mathcal{S}\}} \left\{\boldsymbol{\epsilon}^0 :< \boldsymbol{\sigma} > -W^c(\tilde{\boldsymbol{\sigma}})\right\} \tag{8.92}$$

REMARK 8.6.1 **1.** *Note that Eq. (8.91) looks like Legendre-Fenchel transformation. However, there is a subtle difference.*

If W is a convex functional of $\boldsymbol{\epsilon} \in \mathcal{E}$, the Legendre-Fenchel transformation assures that

$$W^c(\boldsymbol{\sigma}) = \sup_{\{\boldsymbol{\epsilon}\in\mathcal{E}\}} \{\boldsymbol{\sigma} : \boldsymbol{\epsilon} - W(\boldsymbol{\epsilon})\}$$

If the space $\mathcal{E} = \mathcal{E}^{**}$ is reflexive (all the $L^p(V)$ spaces are reflexive, see Rudin [1991]), the inverse Legendre-Fenchel transformation exists,

$$W(\boldsymbol{\epsilon}) = (W^c)^c(\boldsymbol{\epsilon}) = W^{cc}(\boldsymbol{\epsilon}) = \sup_{\{\boldsymbol{\sigma}\in\mathcal{S}\}} \{\boldsymbol{\epsilon} : \boldsymbol{\sigma} - W^c(\boldsymbol{\sigma})\}$$

2. *Choose*

$$< \boldsymbol{\sigma} >= \sum_{\alpha=0}^{n} \frac{1}{V}\int_{V_\alpha}\Big(\mathbf{C}^\alpha : \boldsymbol{\epsilon}^0\Big)dV = \sum_{\alpha=0}^{n} f_\alpha \mathbf{C} : \boldsymbol{\epsilon}^0 \ .$$

One can show that

$$\begin{aligned} \frac{1}{2}\boldsymbol{\epsilon}^0 : \bar{\mathbf{C}} : \boldsymbol{\epsilon}^0 \quad &\geq \quad \boldsymbol{\epsilon}^0 : \left\{ \Big(\sum_{\alpha=0}^{n} f_\alpha \mathbf{C}^\alpha\Big) \right. \\ &\left. -\frac{1}{2}\Big(\sum_{\alpha=0}^{n} f_\alpha \mathbf{C}^\alpha\Big) : \Big(\sum_{\alpha=0}^{n} f_\alpha \mathbf{C}^\alpha\Big) : \Big(\sum_{\alpha=0}^{n} f_\alpha \mathbf{C}^\alpha\Big) \right\} : \boldsymbol{\epsilon}^0 \end{aligned}$$

Hence

$$\bar{\mathbf{C}} \geq 2\Big(\sum_{\alpha=0}^{n} f_\alpha \mathbf{C}^\alpha\Big) : \left\{ \mathbf{1}^{(4s)} - \Big(\sum_{\alpha=0}^{n} f_\alpha \mathbf{C}^\alpha\Big) : \Big(\sum_{\alpha=0}^{n} f_\alpha \mathbf{C}^\alpha\Big) \right\}$$

which is referred to as the Sachs bound.

8.7 Talbot-Willis variational principles

In a series papers (Talbot and Willis [1985],[1987]), Talbot and Willis generalized Hashin-Shtrikman variational principles to well-behaviored nonlinear media.

Consider a composite with nonlinear strain potential energy density, $U(\boldsymbol{\epsilon})$,

$$\begin{aligned} \nabla \cdot \boldsymbol{\sigma} &= \mathbf{0}, \\ \boldsymbol{\sigma} &= \partial_{\boldsymbol{\epsilon}} U, \\ \boldsymbol{\epsilon} &= \frac{1}{2}(\nabla \otimes \mathbf{u} + (\nabla \otimes \mathbf{u})^T) \\ \mathbf{u} &= \mathbf{x} \cdot \bar{\boldsymbol{\epsilon}}, \ \ \forall \mathbf{x} \in \partial V \ \ (\Gamma_t = \emptyset) \end{aligned}$$

Consider a homogeneous composite,

$$\nabla \cdot \boldsymbol{\sigma}^0 = \mathbf{0}, \tag{8.93}$$

$$\boldsymbol{\sigma}^0 = \partial_{\boldsymbol{\epsilon}^0} U^0, \tag{8.94}$$

$$\boldsymbol{\epsilon}^0 = \frac{1}{2}(\nabla \otimes \mathbf{u}^0 + (\nabla \otimes \mathbf{u}^0)^T) \tag{8.95}$$

$$\mathbf{u}^0 = \mathbf{x} \cdot \bar{\boldsymbol{\epsilon}}, \ \ \forall \mathbf{x} \in \partial V \ \ (\Gamma_t = \emptyset) \tag{8.96}$$

Compare the differences in potential energy density, $\mathcal{U}(\boldsymbol{\epsilon}) = U(\boldsymbol{\epsilon}) - U^0(\boldsymbol{\epsilon})$. We define

$$\mathcal{U}_p(\boldsymbol{\epsilon}) := U(\boldsymbol{\epsilon}) - U^0(\boldsymbol{\epsilon}), \ \ \partial^2_{\boldsymbol{\epsilon}} \mathcal{U} > 0 \tag{8.97}$$

$$\mathcal{U}^p(\boldsymbol{\epsilon}) := U(\boldsymbol{\epsilon}) - U^0(\boldsymbol{\epsilon}), \ \ \partial^2_{\boldsymbol{\epsilon}} \mathcal{U} < 0 \tag{8.98}$$

Assume the following kinematic decomposition,

$$\mathbf{u} = \mathbf{u}^{(0)} + \mathbf{u}^{(d)} \tag{8.99}$$

$$\boldsymbol{\epsilon} = \boldsymbol{\epsilon}^{(0)} + \boldsymbol{\epsilon}^d \tag{8.100}$$

Assume that the stress and strain fields in the comparison solid are uniqluy determined by the boundary condition. The total potential energy difference is a functional of ϵ^d, i.e.

$$\Pi_p(\boldsymbol{\epsilon}^d) = W(\boldsymbol{\epsilon}^d) - W_0(\boldsymbol{\epsilon}^d) = \frac{1}{V}\int_V \mathcal{U}_p(\boldsymbol{\epsilon}^d) dV \tag{8.101}$$

$$\Pi^p(\boldsymbol{\epsilon}^d) = W(\boldsymbol{\epsilon}^d) - W_0(\boldsymbol{\epsilon}^d) = \frac{1}{V}\int_V \mathcal{U}^p(\boldsymbol{\epsilon}^d) dV \tag{8.102}$$

where

$$\begin{aligned} W(\boldsymbol{\epsilon}^d) &= \frac{1}{2V}\int_V U(\boldsymbol{\epsilon}^d)dV \\ W_0(\boldsymbol{\epsilon}^d) &= \frac{1}{2V}\int_V U_0(\boldsymbol{\epsilon}^d)dV \end{aligned}$$

Obviously, $\Pi_p(\boldsymbol{\epsilon})$ is convex and $\Pi^p(\boldsymbol{\epsilon})$ is concave.

Define stress polarization

$$p_{ij} = \frac{\partial \mathcal{U}}{\partial \epsilon^d_{ij}} \tag{8.103}$$

Subsequently, we can form the following Legendre-Fenchel transformation,

$$\Pi^*_p = \sup_{\epsilon^d \in E}\Big\{ < \mathbf{p} : \boldsymbol{\epsilon}^d > - \Pi_p(\boldsymbol{\epsilon}^d)\Big\} \tag{8.104}$$

$$\Pi^{p*} = \inf_{\epsilon^d \in E}\Big\{ < \mathbf{p} : \boldsymbol{\epsilon}^d > - \Pi^p(\boldsymbol{\epsilon}^d)\Big\} \tag{8.105}$$

where

$$< \mathbf{p} : \boldsymbol{\epsilon}^d > = \frac{1}{V}\int_V \mathbf{p} : \boldsymbol{\epsilon}^d dV$$

and

$$\begin{aligned} E &:= \Big\{ \epsilon_{ij} \,\Big|\, \epsilon_{ij} \in L^2(V), \epsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i}), \text{ and } u_i \in \overset{\circ}{\mathcal{V}} \Big\} \\ \mathcal{V} &:= \Big\{ u_i \,\Big|\, u_i \in L^2(V), W(u_{i,j}), W_0(u_{i,j}) < \infty,\ u_i = x_j \epsilon^0_{ij},\ \forall \mathbf{x} \in \partial V \Big\} \\ \overset{\circ}{\mathcal{V}} &:= \Big\{ u_i \,\Big|\, u_i \in L^2(V), W(u_{i,j}), W_0(u_{i,j}) < \infty,\ u_i = 0,\ \forall \mathbf{x} \in \partial V \Big\} \end{aligned}$$

In fact, in plain terms, Eqs. (8.104) and (8.105) are just

$$\Pi^*_p(\mathbf{p}) = (W - W_0)^*(\mathbf{p}) = \sup_{\{\boldsymbol{\epsilon}^d \in E\}}\Big\{ < \mathbf{p} : \boldsymbol{\epsilon}^d > - (W - W_0)(\boldsymbol{\epsilon}^d)\Big\} \tag{8.106}$$

$$\text{when } \partial^2\mathcal{U} > 0, \ \text{ and } \ (W - W^0) \text{ is convex,}$$

$$\Pi^{p*}(\mathbf{p}) = (W - W_0)^*(\mathbf{p}) = \inf_{\{\boldsymbol{\epsilon}^d \in E\}}\Big\{ < \mathbf{p} : \boldsymbol{\epsilon}^d > - (W - W_0)(\boldsymbol{\epsilon}^d)\Big\} \tag{8.107}$$

$$\text{when } \partial^2\mathcal{U} < 0, \ \text{ and } \ (W - W^0) \text{ is concave.}$$

$(\mathbf{1}.)$ Assume $\partial^2\mathcal{U} > 0$. From Eq. (8.106)

$$\begin{aligned} &\Pi^*_p(\mathbf{p}) \geq \Big(< \mathbf{p} : \boldsymbol{\epsilon}^d > - W(\boldsymbol{\epsilon}^d) + W_0(\boldsymbol{\epsilon}^d)\Big) \\ \Rightarrow\ &W(\boldsymbol{\epsilon}^d) \geq \{ < \mathbf{p} : \boldsymbol{\epsilon}^d > + W_0(\boldsymbol{\epsilon}^d)\} - \Pi^*_p(\mathbf{p}) \end{aligned}$$

Take an infimum through the both sides of the inequality,

$$\inf_{\{\boldsymbol{\epsilon}^d \in E\}} W(\boldsymbol{\epsilon}^d) \geq \inf_{\boldsymbol{\epsilon}^d \in E} \{< \mathbf{p} : \boldsymbol{\epsilon}^d > + W_0(\boldsymbol{\epsilon}^d)\} - \Pi_p^*(\mathbf{p}) \tag{8.108}$$

$(\mathbf{2}.)$ Assume $\partial^2 \mathcal{U} < 0$. From Eq. (8.107)

$$\begin{aligned} \Pi^{p*}(\mathbf{p}) &\leq \Big(< \mathbf{p} : \boldsymbol{\epsilon}^d > - W(\boldsymbol{\epsilon}^d) + W_0(\boldsymbol{\epsilon}^d)\Big) \\ \Rightarrow \quad W(\boldsymbol{\epsilon}^d) &\leq \{< \mathbf{p} : \boldsymbol{\epsilon}^d > + W_0(\boldsymbol{\epsilon}^d)\} - \Pi_p^*(\mathbf{p}) \end{aligned}$$

Take an infimum through the both sides of the above inequality

$$\inf_{\{\boldsymbol{\epsilon}^d \in E\}} W(\boldsymbol{\epsilon}^d) \leq \inf_{\{\boldsymbol{\epsilon}^d \in E\}} \{< \mathbf{p} : \boldsymbol{\epsilon}^d > + W_0(\boldsymbol{\epsilon}^d)\} - \Pi^{p*}(\mathbf{p}) \tag{8.109}$$

The prime variational principle is

$$(\text{The primal problem}) \;\; \mathcal{P} : \inf_{\{\boldsymbol{\epsilon}^d \in E\}} W(\boldsymbol{\epsilon}^d)$$

Combining Eqs. (8.108) and (8.109), we have the original form of Talbot-Willis variational princinple

$$\begin{aligned} &\inf_{\{\boldsymbol{\epsilon}^d \in E\}} \{< \mathbf{p} : \boldsymbol{\epsilon}^d > + W_0(\boldsymbol{\epsilon}^d)\} - \Pi_p^*(\mathbf{p}) \\ &\qquad \leq \inf_{\{\boldsymbol{\epsilon}^d \in E\}} W(\boldsymbol{\epsilon}^d) \leq \\ &\inf_{\{\boldsymbol{\epsilon}^d \in E\}} \{< \mathbf{p} : \boldsymbol{\epsilon}^d > + W_0(\boldsymbol{\epsilon}^d)\} - \Pi^{p*}(\mathbf{p}) \end{aligned} \tag{8.110}$$

which is the generalization of Hashin-Shtrikman principle.

If both the original composite and the comparison solid are linear elastic materials, we easily calculate,

$$\begin{aligned} \Pi_p^*(\mathbf{p}) \;\Big(\text{or } \Pi^{p*}(\mathbf{p})\Big) &= \frac{1}{V}\int_V \Big(\epsilon_{ij}^d p_{ij} - \frac{1}{2}\Delta C_{ijk\ell}\epsilon_{ij}\epsilon_{k\ell}\Big) dV \\ &= \frac{1}{V}\int_V \Big((\epsilon_{ij} - \epsilon_{ij}^0) p_{ij} - \frac{1}{2} p_{ij}\epsilon_{ij}\Big) dV \\ &= \frac{1}{2V}\int_V \Big(\epsilon_{ij} p_{ij} - 2\epsilon_{ij}^0 p_{ij}\Big) dV \\ &= \frac{1}{2V}\int_V \Big(\Delta C_{ijk\ell}^{-1} p_{ij} p_{k\ell} - 2\epsilon_{ij}^0 p_{ij}\Big) dV \end{aligned}$$

Denote

$$\underline{I}(\boldsymbol{\epsilon}^d, \mathbf{p}) = \inf_{\boldsymbol{\epsilon}^d \in E} \{< \mathbf{p} : \boldsymbol{\epsilon}^d > + W_0(\boldsymbol{\epsilon}^d)\} - \Pi_p^*(\mathbf{p}) \tag{8.111}$$

$$\bar{I}(\boldsymbol{\epsilon}^d, \mathbf{p}) = \inf_{\boldsymbol{\epsilon}^d \in E} \{< \mathbf{p} : \boldsymbol{\epsilon}^d > + W_0(\boldsymbol{\epsilon}^d)\} - \Pi^{p*}(\mathbf{p}) \tag{8.112}$$

We can find that

$$
\begin{aligned}
\underline{I}\ (\text{or } \bar{I}) &= \frac{1}{V}\int_V \Big(p_{ij}\epsilon^d_{ij} + \frac{1}{2}C^0_{ijk\ell}\epsilon_{k\ell}(\epsilon^0_{ij}+\epsilon^d_{ij}) \\
&\quad -\frac{1}{2}\Delta C^{-1}_{ijk\ell}p_{ij}p_{k\ell} + \epsilon^0_{ij}p_{ij}\Big)dV \\
&= \underbrace{\frac{1}{2V}\int_V \Big(p_{ij} + C^0_{ijkl}\epsilon_{k\ell}\Big)\epsilon^d_{ij}dV}_{=0} \\
&\quad +\frac{1}{2V}\int_V C^0_{ijk\ell}(\epsilon^0_{k\ell}+\epsilon^d_{k\ell})\epsilon^0_{ij}dV \\
&\quad +\frac{1}{V}\int_V \Big(\frac{1}{2}\epsilon^d_{ij}p_{ij} - \frac{1}{2}\Delta C^{-1}_{ijk\ell}p_{ij}p_{k\ell} + \epsilon^0_{ij}p_{ij}\Big)dV \\
&= \underbrace{\frac{1}{2V}\int_V C^0_{ijk\ell}\epsilon^0_{k\ell}\epsilon^d_{ij}dV}_{=0} \\
&\quad +\frac{1}{V}\int_V \Big(\frac{1}{2}C^0_{ijk\ell}\epsilon^0_{ij}\epsilon^0_{k\ell} + \frac{1}{2}\epsilon^d_{ij}p_{ij} - \frac{1}{2}\Delta C^{-1}_{ijk\ell}p_{ij}p_{k\ell} + \epsilon^0_{ij}p_{ij}\Big)dV
\end{aligned}
$$

Hence

$$
\begin{aligned}
\underline{I},\ (\text{or } \bar{I}) &= \frac{1}{V}\int_V \Big(\frac{1}{2}C^0_{ijk\ell}\epsilon^0_{ij}\epsilon^0_{k\ell} + \frac{1}{2}\epsilon^d_{ij}p_{ij} - \frac{1}{2}\Delta C^{-1}_{ijk\ell}p_{ij}p_{k\ell} + \epsilon^0_{ij}p_{ij}\Big)dV \\
&= W_0(\boldsymbol{\epsilon}^0) + \underline{R_\pi}\ (\text{or } \bar{R}_\pi)
\end{aligned}
$$

where

$$
\underline{R_\pi},\ (\text{or}\bar{R}_\pi) = \frac{1}{2V}\int_V \Big(-\Delta C^{-1}_{ijk\ell}p_{ij}p_{k\ell} + p_{ij}\epsilon^d_{ij} + 2p_{ij}\epsilon^0_{ij}\Big)dV
$$

We then recover the Hashin-Shtrikman variational principle

$$
\underline{R_\pi}(\mathbf{p},\boldsymbol{\epsilon}^d) \le \inf_{\boldsymbol{\epsilon}^d\in E} W(\boldsymbol{\epsilon}^d) - W_0(\boldsymbol{\epsilon}^0) \le \bar{R}_\pi(\mathbf{p},\boldsymbol{\epsilon}^d)
$$

8.8 Exercises

PROBELM 8.1 *Consider a functional*

$$
P : H^1([a,b]) \to \mathbb{R}
$$

where

$$
P(u) = \int_a^b \sqrt{1+[u'(x)]^2}dx\ .
$$

with essential boundary condition $u(a)=\bar{u}_a$ *and* $u(b)=\bar{u}_b$.

Find the first variation, second variation, and Gâteaux derivative. Derive associated the Euler-Lagrange equation.

PROBELM 8.2 *Let $\Gamma_u = \emptyset$, $\partial V = \Gamma_t$, and $f_i = 0$. Assume that the RVE has the prescribed traction boundary condition,*

$$\mathbf{n} \cdot \bar{\boldsymbol{\sigma}} = \mathbf{t}^0(\mathbf{x}), \quad \forall \mathbf{x} \in \partial V \tag{8.113}$$

where $\bar{\boldsymbol{\sigma}} > 0$ is a constant tensor.

Show that

$$W^c(\tilde{\boldsymbol{\sigma}}) = \sup_{\{<\boldsymbol{\epsilon}>\big|\boldsymbol{\epsilon}\in\mathcal{E}\}} \Big\{ \bar{\boldsymbol{\sigma}} : \langle \boldsymbol{\epsilon} \rangle - \tilde{W}(<\boldsymbol{\epsilon}>) \Big\} \tag{8.114}$$

where $\mathcal{E} := \{\epsilon_{ij} \;\Big|\; \epsilon_{ij,kl} + \epsilon_{kl,ij} - \epsilon_{ik,jl} - \epsilon_{jl,ik} = 0, \;\; and \;\; \epsilon_{ij} \in L^2(V)\}$,

$$W^c(\tilde{\boldsymbol{\sigma}}) \;\; := \;\; \frac{1}{2V}\int_V D_{ijkl}\tilde{\sigma}_{ij}\tilde{\sigma}_{kl}dV = \frac{1}{2}\int_V \tilde{\epsilon}_{ij}\bar{\sigma}_{kl}dV \tag{8.115}$$

$$\tilde{W}(<\boldsymbol{\epsilon}>) \;\; := \;\; \frac{1}{2V} \inf_{\{\frac{1}{V}\int_V \boldsymbol{\epsilon}dV=<\boldsymbol{\epsilon}>,\ \boldsymbol{\epsilon}\in\mathcal{E}\}} \int_V C_{ijkl}\epsilon_{ij}\epsilon_{kl}dV \tag{8.116}$$

Note that $\tilde{\sigma}_{ij}$ and $\tilde{\epsilon}_{ij}$ are the real solutions.

PROBELM 8.3 *Let $\Gamma_u = \emptyset$ and $\partial\Omega = \Gamma_t$. Consider the following the boundary-value problem,*

$$\sigma_{ij,j} = 0, \quad \forall \mathbf{x} \in \Omega \tag{8.117}$$

$$n_j\sigma_{ij} = t_i^0, \quad \forall \mathbf{x} \in \Gamma_t, \quad \text{and} \quad \Gamma_u = \emptyset \tag{8.118}$$

$$\epsilon_{ij} = \frac{1}{2}\Big(u_{i,j} + u_{j,i}\Big) \tag{8.119}$$

$$\epsilon_{ij} = \frac{\partial U_c}{\partial \sigma_{ij}}, \quad U_c(\boldsymbol{\sigma}) := \frac{1}{2}D_{ijk\ell}\sigma_{ij}\sigma_{k\ell}. \tag{8.120}$$

Consider a comparison elastic solid with compliance tensor, $D^0_{ijk\ell}$ and

$$\sigma^{(0)}_{ij,j} = 0, \quad \forall \mathbf{x} \in \Omega \tag{8.121}$$

$$n_j\sigma^{(0)}_{ij} = t_i^0, \quad \forall \mathbf{x} \in \Gamma_t, \quad \text{and} \quad \Gamma_u = \emptyset \tag{8.122}$$

$$\epsilon^{(0)}_{ij} = \frac{1}{2}\Big(u^{(0)}_{i,j} + u^{(0)}_{j,i}\Big) \tag{8.123}$$

$$\epsilon^{(0)}_{ij} = \frac{\partial U^{(0)}_0}{\partial \sigma^{(0)}_{ij}}, \quad U^{(0)}_0(\boldsymbol{\sigma}) := \frac{1}{2}D^{(0)}_{ijk\ell}\sigma^{(0)}_{ij}\sigma^{(0)}_{k\ell}\;. \tag{8.124}$$

Let

$$\sigma_{ij} = \sigma_{ij}^{(0)} + \sigma_{ij}^{d} \tag{8.125}$$

$$\epsilon_{ij} = D_{ijk\ell}^{(0)}\sigma_{k\ell} + q_{ij} \tag{8.126}$$

where σ_{ij}^{d} *is called disturbance stress, and* q_{ij} *is called polarization strain (eigenstrain).*

They are connected by the following subsidary conditions: **1.** *the weak form of subsidiary condition (complementary virtual work principle),*

$$\int_\Omega \epsilon_{ij}\sigma_{ij}^{d} d\Omega = 0 \tag{8.127}$$

or **2.** *the strong form of subsidiary condition*

$$\epsilon^{'} := D_{ijk\ell}^{(0)}\sigma_{k\ell}^{d} + q_{ij}\ , \quad \mathcal{C}(\epsilon_{ij}^{'}) = \epsilon_{ij,k\ell}^{'} + \epsilon_{k\ell,ij}^{'} - \epsilon_{ik,j\ell}^{'} - \epsilon_{i\ell,jk}^{'} = 0, \quad \forall \mathbf{x} \in \Omega \tag{8.128}$$

Consider the following variational problem

$$(\text{The primal problem :}) \quad \mathcal{P}: \qquad \inf_{\boldsymbol{\sigma}^d \in S(\Omega)} \Pi_c(\boldsymbol{\sigma}^d) \tag{8.129}$$

or

$$(\text{The primal problem :}) \quad \mathcal{P}: \qquad \inf_{\boldsymbol{\sigma}^d \in S(\Omega)} W_c(\boldsymbol{\sigma}^d) \tag{8.130}$$

where

$$W_c(\boldsymbol{\sigma}^d) := \frac{1}{2|\Omega|}\int_\Omega D_{ijk\ell}\sigma_{ij}\sigma_{k\ell} d\Omega = \int_\Omega D_{ijk\ell}(\sigma_{ij}^{(0)} + \sigma_{ij}^{d})(\sigma_{k\ell}^{(0)} + \sigma_{k\ell}^{d}) d\Omega, \tag{8.131}$$

$\Pi_c(\boldsymbol{\sigma}^d) = \Omega W_c(\boldsymbol{\sigma}^d)$ *and*

$$S := \left\{ \boldsymbol{\sigma} \;\middle|\; n_j\sigma_{ij} = 0, \ \forall \mathbf{x} \in \Gamma_t, \ \text{and} \ \sigma_{ij} \in C^0(\Omega) \right\} \tag{8.132}$$

Derive Hashin-Shtrikman variational principle.

Hints:

Z. Hasin and S. Shtrikman [1962], "On some variational principles in anisotropic and nonhomogebeous elasticity," Journal of Mechanics and Physics of Solids, **10***, pp. 335-342.*

D. R. S. Talbot and J. R. Willis [1985], "Variational principles for inhomogeneous non-linear media," IMA Journal of Applied Mathematics, **35***, 39-54.*

Chapter 9

BOUNDS ON EFFECTIVE PROPERTIES

9.1 Hashin-Shtrikman bounds

Consider prescribed macro strain boundary condition for both the composite and the comparison solid,

$$\begin{aligned}\mathbf{u} &= \bar{\mathbf{u}} = \mathbf{x}\cdot\bar{\boldsymbol{\epsilon}}, \ \ \forall \mathbf{x}\in\partial V \quad (\Gamma_t = \emptyset)\\ \mathbf{u}^0 &= \bar{\mathbf{u}} = \mathbf{x}\cdot\bar{\boldsymbol{\epsilon}}, \ \ \forall \mathbf{x}\in\partial V \quad (\Gamma_t = \emptyset)\end{aligned}$$

by the averaging theorem $\bar{\boldsymbol{\epsilon}} =< \boldsymbol{\epsilon} >$.

Under such condition, Hashin-Shtrikman variational principles are

$$\underbrace{\underline{I}}_{\Delta\mathbf{C}>0} \leq \inf_{\boldsymbol{\epsilon}^d\in E} W(\boldsymbol{\epsilon}^d) \leq \underbrace{\bar{I}}_{\Delta\mathbf{C}<0} \tag{9.1}$$

where $\Delta\mathbf{C} = \mathbf{C} - \mathbf{C}^{(0)}$, and

$$\underline{I}\ \left(\text{or } \bar{I}\right) = W_0(\boldsymbol{\epsilon}^0) - \frac{1}{2V}\int_V \Big[\Delta C^{-1}_{ijk\ell}p_{ij}p_{k\ell} - p_{ij}\epsilon^d_{ij} - 2p_{ij}\epsilon^0_{ij}\Big]dV \tag{9.2}$$

Assume that there are n-phase in the composite (including the matrix). In each phase (inclusion), the elastic tensor as well as stress polarization tensor is constant, i.e.

$$\mathbf{C}(\mathbf{x}) = \sum_{r=1}^{n} \mathbf{C}^r H(\Omega_r) \tag{9.3}$$

$$\mathbf{p}(\mathbf{x}) = \sum_{r=1}^{n} \mathbf{p}^r H(\Omega_r) \tag{9.4}$$

where $H(\cdot)$ is the Heaviside function, and Ω_r is the domain of each phase,

$$H(\Omega_r) = \begin{cases} 1, & \forall \mathbf{x} \in \Omega_r \\ \\ 0, & \forall \mathbf{x} \notin \Omega_r \end{cases}$$

We now calculate each term in (9.1).

1

$$\begin{aligned} \inf_{\boldsymbol{\epsilon}^d \in E} W(\boldsymbol{\epsilon}^d) &= \frac{1}{2V} \int_V \boldsymbol{\sigma} : \boldsymbol{\epsilon} dV = \frac{1}{2} < \boldsymbol{\sigma} >:< \boldsymbol{\epsilon} > \\ &= \frac{1}{2} < \boldsymbol{\epsilon} >: \bar{\mathbf{C}} :< \boldsymbol{\epsilon} >= \frac{1}{2} \bar{\boldsymbol{\epsilon}} : \bar{\mathbf{C}} : \bar{\boldsymbol{\epsilon}} \end{aligned} \tag{9.5}$$

2

$$\begin{aligned} W_0(\boldsymbol{\epsilon}^0) &= \frac{1}{2V} \int_V \boldsymbol{\sigma}^0 : \boldsymbol{\epsilon}^0 dV = \frac{1}{2} < \boldsymbol{\sigma}^0 >:< \boldsymbol{\epsilon}^0 > \\ &= \frac{1}{2} < \boldsymbol{\epsilon}^0 >: \mathbf{C}^0 :< \boldsymbol{\epsilon}^0 >= \frac{1}{2} \bar{\boldsymbol{\epsilon}} : \mathbf{C}^0 : \bar{\boldsymbol{\epsilon}} \end{aligned} \tag{9.6}$$

3

$$\begin{aligned} \frac{1}{2V} \int_V \mathbf{p}^r : \Delta \mathbf{C}^{-1} : \mathbf{p} dV &= \frac{1}{2} \sum_{r=1}^{n} \frac{1}{V} \int_{\Omega_r} \mathbf{p}^r : \mathbf{C}_r^{-1} : \mathbf{p}^r dV \\ &= \frac{1}{2} \sum_{r=1}^{n} f_r \mathbf{p}^r : \Delta \mathbf{C}_r^{-1} : \mathbf{p}^r \end{aligned} \tag{9.7}$$

4

$$\frac{1}{V} \int_V \mathbf{p} : \boldsymbol{\epsilon}^0 dV = \Big(\frac{1}{V} \int_V \mathbf{p} dV \Big) : \bar{\boldsymbol{\epsilon}} =< \mathbf{p} >: \bar{\boldsymbol{\epsilon}} = \sum_{r=1}^{n} f_r \mathbf{p}_r : \bar{\boldsymbol{\epsilon}} \tag{9.8}$$

5

$$\frac{1}{2V} \int_V \mathbf{p} : \boldsymbol{\epsilon}^d dV = -\frac{1}{2} \sum_{r=1}^{n} f_r \mathbf{p}^r : \mathbf{P}^r : \Big(\mathbf{p}^r - < \mathbf{p} > \Big) \tag{9.9}$$

where

$$\mathbf{P}^r := \int_{\Omega_r} \boldsymbol{\Gamma}^\infty (\mathbf{x}' - \mathbf{x}) dV_{\mathbf{x}'}$$

and

$$\Gamma^\infty_{ijk\ell} := -\frac{1}{4} \Big(G^\infty_{ki,j\ell}(\mathbf{x}' - \mathbf{x}) + G^\infty_{kj,i\ell}(\mathbf{x}' - \mathbf{x}) + G^\infty_{\ell i,jk}(\mathbf{x}' - \mathbf{x}) + G^\infty_{\ell j,ik}(\mathbf{x}' - \mathbf{x}) \Big)$$

How to integrate

$$\frac{1}{2V}\int_V \mathbf{p} : \boldsymbol{\epsilon}^d dV = ? \tag{9.10}$$

Consider the subsidiary condition,

$$C^{(0)}_{ijk\ell} u^d_{k,\ell j} + p_{ij,j} = 0 \tag{9.11}$$

We solve u^d_k in terms of p_{ij} by using Green's function method. Conisder the Green's function of the comparison solid in an infinite medium, i.e.

$$C^{(0)}_{ijk\ell} G^{\infty}_{km,\ell j} + \delta_{im}\delta(\mathbf{x}-\mathbf{x}') = 0, \;\; \forall\, \mathbf{x}, \mathbf{x}' \in \mathbb{R}^3$$

Multiplying $G_{im}(\mathbf{x}'-\mathbf{x})$ with (9.11) and integrating it over V, one has

$$\int_V \Big[C^{(0)}_{ijk\ell} u^d_{k,\ell} + p_{ij}\Big]_{,j} G^{\infty}_{im}(\mathbf{x}'-\mathbf{x}) dV_{\mathbf{x}'} = 0$$

Let $t_{ij} = C^{(0)}_{ijk\ell} u^d_{k,\ell}$. Integration by parts yields,

$$\begin{aligned}
& \int_{\partial V} G^{\infty}_{im}(\mathbf{x}'-\mathbf{x}) \Big[\underbrace{C^{(0)}_{ijk\ell} u^d_{k,\ell}}_{t_{ij}} + p_{ij}\Big] n_j dS \\
& - \int_V \frac{\partial}{\partial x'_j} G^{\infty}_{im}(\mathbf{x}'-\mathbf{x}) \Big[C^{(0)}_{ijk\ell} u^d_{k,\ell} + p_{ij}\Big] dV \\
= \; & \int_{\partial V} G^{\infty}_{im}(\mathbf{x}-\mathbf{x}')[t_{ij}+p_{ij}] n_j dS - \int_{\partial V} \Big[\frac{\partial}{\partial x'_j} G^{\infty}_{im}(\mathbf{x}'-\mathbf{x})\Big] \Big[\underbrace{C^{(0)}_{ijk\ell} u^d_k n_\ell}_{=0}\Big] dS \\
& + \int_V \Big[\frac{\partial^2}{\partial x'_j \partial x'_\ell} G^{\infty}_{im}\Big] \Big[C^{(0)}_{ijk\ell} u^d_k n_\ell\Big] dV - \int_V \frac{\partial}{\partial x'_j} G^{\infty}_{ij}(\mathbf{x}'-\mathbf{x}) p_{ij}(\mathbf{x}') dV \\
= \; & \int_{\partial V} G^{\infty}_{im}(\mathbf{x}'-\mathbf{x})[t_{ij}+p_{ij}] n_j dS - \int_V \frac{\partial}{\partial x'_j} G^{\infty}_{ij}(\mathbf{x}'-\mathbf{x}) p_{ij}(\mathbf{x}') dV \\
& + \int_V \underbrace{C^{(0)}_{ijk\ell} G^{\infty}_{km,j\ell}(\mathbf{x}'-\mathbf{x})}_{-\delta_{im}\delta(\mathbf{x}'-\mathbf{x})} u^d_i(\mathbf{x}') dV
\end{aligned}$$

because of major symmetry of $\mathbf{C}^{(0)}$, one can interchange indices $k \to i$ and $j \to \ell$.

Therefore,

$$u^d_m(\mathbf{x}) = \int_{\partial V} G^{\infty}(\mathbf{x}'-\mathbf{x})[t_{ij}(\mathbf{x}') + p_{ij}(\mathbf{x}')] n_j dS - \int_V G^{\infty}_{im,j}(\mathbf{x}'-\mathbf{x}) p_{ij}(\mathbf{x}') dV \tag{9.12}$$

Since $\mathbf{u}^d = 0,\ \forall\, \mathbf{x} \in \partial V$, t_{ij} oscillate around zero. Then its average $< C^{(0)}_{ijk\ell} u^d_{k,\ell} >_{\partial V}$ along the boundary should be very small. We assume that

$$< C^{(0)}_{ijk\ell} u^d_{k,\ell} >_{\partial V} \approx 0$$

Now the only term remaining is

$$\int_{\partial V} G^\infty_{im}(\mathbf{x}' - \mathbf{x}) p_{ij}(\mathbf{x}') dS$$

To essence of the additional manipulation is to modify the volume integral in (9.12) in order to drop out the surface integral in (9.12). To do accomplish this goal, we consider identity,

$$< p_{ij} >_{,j} = 0 \quad \Rightarrow \quad \int_V < p_{ij} >_{,j} G^\infty_{im}(\mathbf{x}' - \mathbf{x}) dV = 0$$

Integration by parts yields,

$$\int_V < p_{ij} >_{,j} G^\infty_{im}(\mathbf{x}' - \mathbf{x}) dV = \int_{\partial V} < p_{ij} > n_j G^\infty_{im}(\mathbf{x}' - \mathbf{x}) dS$$
$$- \int_V G^\infty_{im,j}(\mathbf{x}' - \mathbf{x}) < p_{ij} > dV = 0 \tag{9.13}$$

Thus substracting (9.13) from (9.12) will be affect the value of (9.12),

$$\begin{aligned} u^d_m(\mathbf{x}) &= \int_{\partial V} G^\infty_{im}(\mathbf{x}' - \mathbf{x})[t_{ij}(\mathbf{x}^{'}) + (p_{ij}(\mathbf{x}^{'}) - < p_{ij} >)] n_j dS \\ & \quad - \int_V G^\infty_{im,j}(\mathbf{x}' - \mathbf{x})(p_{ij}(\mathbf{x}^{'}) - < p_{ij} >) dV \end{aligned} \tag{9.14}$$

Now $p_{ij} - < p_{ij} >$ also oscillates around zero, since its mean is zero, i.e. $< p_{ij} - < p_{ij} > > = 0$. We can then neglect the boundary term, and finally we have

$$u^d_m(\mathbf{x}) \approx - \int_V G^\infty_{im,j}(\mathbf{x}' - \mathbf{x})(p_{ij}(\mathbf{x}') - < p_{ij} >) dV_{\mathbf{x}'} \tag{9.15}$$

The gradient of the disturbance displacement field is

$$u^d_{m,\ell}(\mathbf{x}) = \int_V G^\infty_{im,j\ell}(\mathbf{x}' - \mathbf{x})(p_{ij}(\mathbf{x}^{'}) - < p_{ij} >) dV_{\mathbf{x}'}$$

Hence

$$\epsilon^d_{m\ell}(\mathbf{x}) = \frac{1}{2} \int_V \Big[G^\infty_{im,j\ell} + G^\infty_{i\ell,jm} \Big] (\mathbf{x}^{'} - \mathbf{x}) \Big(p_{ij}(\mathbf{x}^{'}) - < p_{ij} > \Big) dV_{\mathbf{x}'} \tag{9.16}$$

Since p_{ij} is symmetric, we can also write that

$$
\begin{aligned}
\epsilon^d_{m\ell}(\mathbf{x}) &= \frac{1}{4}\int_V \Big[G^\infty_{im,j\ell} + G^\infty_{i\ell,jm} + G^\infty_{jm,i\ell} + G^\infty_{j\ell,im}\Big](\mathbf{x}' - \mathbf{x}) \\
&\cdot \Big(p_{ij}(\mathbf{x}') - < p_{ij} >\Big) dV_{\mathbf{x}'} \\
&= -\int_V \Gamma^\infty_{m\ell ij}(\mathbf{x}' - \mathbf{x})\Big(p_{ij}(\mathbf{x}') - < p_{ij} >\Big) dV_{\mathbf{x}'} \\
&= -\int_V \mathbf{\Gamma}^\infty(\mathbf{x}' - \mathbf{x}) : \Big(\mathbf{p}(\mathbf{x}') - < \mathbf{p} >\Big) dV_{\mathbf{x}'} \qquad (9.17)
\end{aligned}
$$

where

$$
\Gamma^\infty_{m\ell ij}(\mathbf{y} - \mathbf{x}) := -\frac{1}{4}\Big[G^\infty_{im,j\ell} + G^\infty_{i\ell,jm} + G^\infty_{jm,i\ell} + G^\infty_{j\ell,im}\Big](\mathbf{y} - \mathbf{x}) \quad (9.18)
$$

Consider a bounded and simply-connected region, $\Omega \in V$. We define a new tensor, $\mathbf{P}$,

$$
\mathbf{P}^\Omega(\mathbf{x}) := \int_\Omega \mathbf{\Gamma}^\infty(\mathbf{y} - \mathbf{x}) dV_{\mathbf{y}}, \quad \forall \mathbf{x} \in \Omega \qquad (9.19)
$$

and in components form,

$$
\begin{aligned}
P^\Omega_{ijk\ell}(\mathbf{x}) &= \int_\Omega \Gamma^\infty_{ijk\ell}(\mathbf{y} - \mathbf{x}) dV_{\mathbf{y}} \\
&= -\frac{1}{4}\int_\Omega \Big[G^\infty_{im,j\ell} + G^\infty_{i\ell,jm} + G^\infty_{jm,i\ell} + G^\infty_{j\ell,im}\Big](\mathbf{y} - \mathbf{x}) dV_{\mathbf{y}}
\end{aligned}
\qquad (9.20)
$$

One may verify that when Ω is an ellipsoidal $\mathbf{P}^\Omega$ is constant. In fact, if one recalls the general definition of Eshelby tensor, for $\mathbf{x} \in \Omega$,

$$
\begin{aligned}
S^\Omega_{ijk\ell} &= \int_\Omega \mathcal{G}_{ijk\ell}(\mathbf{y} - \mathbf{x}) dV_{\mathbf{y}} \qquad (9.21) \\
&= -\frac{1}{2}\int_\Omega C_{mnk\ell}\Big[G^\infty_{im,nj} + G^\infty_{jm,ni}\Big](\mathbf{y} - \mathbf{x}) dV_{\mathbf{y}} \\
&= -\frac{1}{4}\int_\Omega \Big[G^\infty_{im,nj} + G^\infty_{jm,ni} + G^\infty_{in,mj} + G^\infty_{jn,mj}\Big](\mathbf{y} - \mathbf{x}) C_{mnk\ell} dV_{\mathbf{y}} \\
&= \int_\Omega \Gamma^\infty_{ijmn}(\mathbf{y} - \mathbf{x}) C_{mnk\ell} dV_{\mathbf{y}} \\
&= P^\Omega_{ijmn} C_{mnk\ell} \qquad (9.22)
\end{aligned}
$$

Now we come back to evaluate (9.10). Let stress polarization $\mathbf{p}(\mathbf{x})$ is piece-wise constant, i.e.

$$\begin{aligned} \mathbf{p}(\mathbf{x}) &= \sum_{r=1}^{n} \mathbf{p}_r H(\Omega_r) \\ < \mathbf{p} > &= \sum_{r=1}^{n} f_r \mathbf{p}_r \end{aligned}$$

Therefore,

$$\begin{aligned} \frac{1}{2V}\int_V \mathbf{p} : \boldsymbol{\epsilon}^d dV &= \frac{1}{2V}\int_V \Big(\sum_{r=1}^{n} p_r H(\Omega_r)\Big) : \\ & \Big(-\int_{V'} \Gamma^{\infty}(\mathbf{x}^{'} - \mathbf{x}) : \Big[\mathbf{p}(\mathbf{x}^{'}) - < \mathbf{p} >\Big]\Big) dV_{\mathbf{x}'} dV_{\mathbf{x}} \end{aligned}$$

Consider $\mathbf{x} \in \Omega_s$. $\mathbf{p}_s - < \mathbf{p} >$ is constant inside Ω_s. Thus,

$$\begin{aligned} & \int_{V'} \mathbf{\Gamma}^{\infty}(\mathbf{x}^{'} - \mathbf{x}) : \Big(\mathbf{p}_r - < \mathbf{p} >\Big) dV_{\mathbf{x}'} \\ = & \Big(\int_{\Omega_s} \mathbf{\Gamma}^{\infty}(\mathbf{x}^{'} - \mathbf{x}) dV_{\mathbf{x}'} + \int_{V'-\Omega_s} \mathbf{\Gamma}^{\infty}(\mathbf{x}^{'} - \mathbf{x}) dV_{\mathbf{x}'}\Big) : \Big(\mathbf{p}_r - < \mathbf{p} >\Big) dV_{\mathbf{x}'} \end{aligned}$$

Assume that the RVE is a gigantic spherical ball and all Ω_r are spherical inclusions. By Mori-Tanaka lemma,

$$\int_{V'-\Omega_s} \mathbf{\Gamma}^{\infty}(\mathbf{x}^{'} - \mathbf{x}) dV_{\mathbf{x}'} = 0$$

In fact, for $\mathbf{x} \in \Omega_s$

$$\int_{V} \mathbf{\Gamma}^{\infty}(\mathbf{x}^{'} - \mathbf{x}) dV_{\mathbf{x}'} = \int_{\Omega_s} \mathbf{\Gamma}^{\infty}(\mathbf{x}^{'} - \mathbf{x}) dV_{\mathbf{x}'}$$

because the integral over a spherical ball does not dependent on the size of inlcusion (recall $\mathbf{P} = \mathbf{S} : \mathbf{D}$).

Hence,

$$\begin{aligned} \frac{1}{2V}\int_V \mathbf{p} : \boldsymbol{\epsilon}^d dV &= -\frac{1}{2V}\sum_{r=1}^{n}\sum_{s=1}^{n}\Big\{\int_{\Omega_r}\Big(\mathbf{p}_r H(\Omega_r(\mathbf{x})\Big) \\ & : \int_{\Omega_s} \mathbf{\Gamma}^{\infty}(\mathbf{x}^{'} - \mathbf{x}) dV_{\mathbf{x}'} : \Big(\mathbf{p}_s H(\Omega_s(\mathbf{x})\Big) dV_{\mathbf{x}}\Big\} \\ & + \frac{1}{2V}\sum_{r=1}^{n}\Big\{\int_{\Omega_r}\Big(\mathbf{p}_r H(\Omega_r(\mathbf{x})\Big) \\ & : \int_{\Omega_s} \mathbf{\Gamma}^{\infty}(\mathbf{x}^{'} - \mathbf{x}) dV_{\mathbf{x}'} :< \mathbf{p} >\Big) dV_{\mathbf{x}}\Big\} \end{aligned}$$

Consider

$$H(\Omega_r(\mathbf{x}))H(\Omega_s(\mathbf{x})) = \begin{cases} 1 & r = s \\ \\ 0 & r \neq s \end{cases} \tag{9.23}$$

and let

$$\mathbf{P}^r := \int_{\Omega_r} \mathbf{\Gamma}^\infty(\mathbf{x}' - \mathbf{x})dV_{\mathbf{x}'}$$

We then have

$$\begin{aligned} \frac{1}{2V}\int_V \mathbf{p} : \boldsymbol{\epsilon}^d dV &= -\frac{1}{2V}\sum_{r=1}^n \int_{\Omega_r} dV_{\mathbf{x}} \mathbf{p}_r : \mathbf{P}^r : \mathbf{p}_r \\ &\quad + \frac{1}{2V}\sum_{r=1}^n \int_{\Omega_r} dV_{\mathbf{x}} \mathbf{p}_r : \mathbf{P}^r :< \mathbf{p} > \\ &= -\frac{1}{2}\sum_{r=1}^n f_r \mathbf{p}_r : \mathbf{P}^r : (\mathbf{p}_r - < \mathbf{p} >) \\ &= -\frac{1}{2}\sum_{r=1}^n f_r p_{ij}^r P_{ijk\ell}^r \Big(p_{k\ell}^r - < p_{k\ell} > \Big) \end{aligned}$$

where $< p_{k\ell} >= \sum_{r=1}^n f_r p_{k\ell}^r$.

REMARK 9.1.1 *Recall that by using Radon transform one can write,*

$$\delta(\mathbf{x}) = -\frac{1}{8\pi^2}\int_{|\boldsymbol{\xi}|=1} \delta''(\xi_n x_n)dS$$

and consequently,

$$G_{ij}^\infty(\mathbf{x}) = \frac{1}{8\pi^2}\int_{|\boldsymbol{\xi}|=1} K_{ij}^{-1}(\boldsymbol{\xi})\delta(\xi_n x_n)dS$$

and for isotropic materials,

$$K_{ij}^{-1}(\boldsymbol{\xi}) = \frac{1}{\mu}\Big[\delta_{ij} - \frac{(\lambda+\mu)\xi_i\xi_j}{(\lambda+2\mu)}\Big]$$

Therefor,

$$G_{ij,k\ell}^\infty(\mathbf{x}) = \frac{1}{8\pi^2}\int_{|\boldsymbol{\xi}|=1} K_{ij}^{-1}(\boldsymbol{\xi})\xi_k\xi_\ell\delta''(\xi_n x_n)dS$$

By definition,

$$\begin{aligned} \Gamma_{ijk\ell}^\infty(\mathbf{x}-\mathbf{y}) &:= -\frac{1}{4}\Big[G_{ik,j\ell}^\infty + G_{i\ell,jk}^\infty + G_{jk,i\ell}^\infty + G_{j\ell,ik}^\infty\Big](\mathbf{x}-\mathbf{y}) \\ &= -\frac{1}{8\pi^2}\int_{|\boldsymbol{\xi}|=1} K_{ij}^{-1}(\boldsymbol{\xi})\xi_k\xi_\ell\delta''\Big(\xi_n(x_n - y_n)\Big)dS \end{aligned}$$

because indices i & j and k & ℓ are symmetric ($K_{ij}^{-1}(\boldsymbol{\xi})$ is symmetric).

To this end, we are in a position to establish Hashin-Shtrikman bounds. Before proceeding to derive Hashin-Shtrikman bound, we first evaluate $\mathbf{P}$ tensor, which can be written as

$$\mathbf{P} = \mathbf{S} : \mathbf{D}^{(0)}$$

For spherical inclusion,

$$\mathbf{S} = s_1^{(0)}\mathbf{E}^{(1)} + s_2^{(0)}\mathbf{E}^2$$

where

$$s_1 = \frac{1+\nu^{(0)}}{3(1-\nu^{(0)})}, \ \ s_2 = \frac{2(4-5\nu^{(0)})}{15(1-\nu^{(0)})}$$

and for isotropic comparison solid,

$$\mathbf{D}^{(0)} = \frac{1}{3K^{(0)}}\mathbf{E}^{(1)} + \frac{1}{2G^{(0)}}\mathbf{E}^{(2)}$$

Therefore,

$$\begin{aligned}
\mathbf{P} &= \frac{s_1^{(0)}}{3K^{(0)}}\mathbf{E}^{(1)} + \frac{s^{(0)}}{2G^{(0)}}\mathbf{E}^{(2)} \\
&= \frac{1+\nu^{(0)}}{9K^{(0)}(1-\nu^{(0)})}\mathbf{E}^{(1)} + \frac{(4-5\nu^{(0)})}{15G^{(0)}(1-\nu^{(0)})}\mathbf{E}^{(2)} \\
&= \frac{1}{2G^{(0)}(1-\nu^{(0)})}\left\{-\frac{1}{15}\mathbf{1}^{(2)}\otimes\mathbf{1}^{(2)} + \frac{2(4-5\nu^{(0)})}{15}\mathbf{1}^{(4s)}\right\} \quad (9.24)
\end{aligned}$$

Consider $\nu^{(0)} = \dfrac{3K^{(0)}-2G^{(0)}}{2(3K^{(0)}+G^{(0)})}$. One can also have

$$\mathbf{P} = \frac{1}{3K^{(0)}+4G^{(0)}}\mathbf{E}^{(1)} + \frac{3(K^{(0)}+2G^{(0)})}{5G^{(0)}(3K^{(0)}+4G^{(0)})}\mathbf{E}^{(2)}$$

For simplicity, we only illustrate Hashin-Shtrikman bound for a two-phase composite. Consider a two-phase well order composite, which implies that $K_2 > K_1$ and $G_2 > G_1$.

Step 1. Let

$$K_0 = K_1, \ \ K = K_2, \ \ \text{and} \ \ G_0 = G_1, \ \ G = G_2 \ .$$

Obviously that

$$\Delta\mathbf{C} = \mathbf{C} - \mathbf{C}^{(0)} = 3(K_2 - K_1)\mathbf{E}^{(1)} + 2(G_2 - G_1)\mathbf{E}^{(2)} > 0$$

Choose a special stress polarization distribution,

$$p_{ij}^{(1)} = 0, \text{ and } p_{ij}^{(2)} = p\delta_{ij} \ .$$

and remote macro strain distribution

$$\bar{\epsilon}_{ij} = \bar{\epsilon}\delta_{ij}$$

We now calculate each terms in $\underline{I}$.

1

$$\begin{aligned} \inf_{\boldsymbol{\epsilon}^d \in E} W(\boldsymbol{\epsilon}^d) &= \frac{1}{2}\bar{C}_{ijk\ell}\Big(\bar{\epsilon}\delta_{ij}\Big)\Big(\bar{\epsilon}\delta_{k\ell}\Big) \\ &= \frac{1}{2}\Big[3\bar{K}E^{(1)}_{ijk\ell} + 2\bar{G}E^{(2)}_{ijk\ell}\Big](\bar{\epsilon})^2\delta_{ij}\delta_{k\ell} \\ &= \frac{9}{2}\bar{K}\bar{\epsilon}^2 \end{aligned}$$

Note that $E^{(1)}_{ijk\ell}\delta_{ij}\delta_{k\ell} = 3$ and $E^{(2)}_{ijk\ell}\delta_{ij}\delta_{k\ell} = 0$.

2

$$\begin{aligned} W_0(\boldsymbol{\epsilon}^0) &= \frac{1}{2}C^{(0)}_{ijk\ell}(\bar{\epsilon}\delta_{ij})(\bar{\epsilon}\delta_{k\ell}) \\ &= \frac{1}{2}\Big[3K_1E^{(1)}_{ijk\ell} + 2G_2E^{(2)}_{ijk\ell}\Big](\bar{\epsilon})^2\delta_{ij}\delta_{k\ell} \\ &= \frac{9}{2}K_1\bar{\epsilon}^2 \end{aligned}$$

3

$$\frac{1}{V}\int_V \mathbf{p} : \boldsymbol{\epsilon}^{(0)} dV = f_1\mathbf{p}_1 : \bar{\boldsymbol{\epsilon}} + f_2\mathbf{p}_2 : \bar{\boldsymbol{\epsilon}} = 3f_2p\bar{\epsilon}$$

4 Because $p_{ij}^{(1)} = 0$ and $p_{ij}^{(2)} = p\delta_{ij}$,

$$\begin{aligned} &\frac{1}{2V}\int_V \mathbf{p} : \Delta\mathbf{C}^{-1} : \mathbf{p} dV = \frac{1}{2}\sum_{r=1}^{2} f_r\mathbf{p}_r : \Delta\mathbf{C}_r^{-1} : \mathbf{p}_r \\ &= \frac{1}{2}\Big(\frac{f_2}{3(K_2 - K_1)}\mathbf{E}^{(1)} + \frac{f_2}{2(G_2 - G_1)}\mathbf{E}^{(2)}\Big)p^2\delta_{ij}\delta_{k\ell} \\ &= \frac{f_2 p^2}{2(K_2 - K_1)} \end{aligned}$$

5 Because $< p_{k\ell} >= f_2 p\delta_{k\ell}$,

$$\begin{aligned}
\frac{1}{2V}\int_V p_{ij}\epsilon^d_{ij}dV &= -\frac{1}{2}\sum_{r=1}^{2} f_r P^r_{ijk\ell}p^r_{ij}p^r_{k\ell} + \frac{1}{2}\sum_{r=1}^{2} f_r P^r_{ijk\ell}p^r_{ij} < p^r_{k\ell} > \\
&= -\frac{f_2}{2}\Big(\frac{3p^2}{3K_1+4G_1}\Big) + \frac{1}{2}\frac{3f_2^2p^2}{3K_1+4G_1} \\
&= -\frac{1}{2}\frac{3f_1f_2p^2}{3K_1+4G_1} = -\frac{1}{2}\frac{f_1f_2p^2}{K_1+\dfrac{4}{3}G_1}
\end{aligned}$$

Therefore, when $\Delta\mathbf{C} > 0$,

$$\underline{I}(p) = \frac{9}{2}K_1\bar{\epsilon}^2 - \frac{f_2p^2}{2(K_2-K_1)} - \frac{1}{2}\frac{f_1f_2p^2}{K_1+\frac{4}{3}G_1} + 3f_2p\bar{\epsilon} \le \frac{9}{2}\bar{K}\bar{\epsilon}^2 \tag{9.25}$$

To find $\min_p \underline{I}$, we check the stationary condition,

$$\begin{aligned}
\frac{\partial \underline{I}}{\partial p} = 0 \quad &\Rightarrow \quad -\frac{f_2p}{(K_2-K_1)} - \frac{f_1f_2p}{K_1+\frac{4}{3}G_1} + 3\bar{\epsilon}f_2 = 0 \\
&\Rightarrow \quad p_{sta} = \frac{3\bar{\epsilon}}{\dfrac{1}{K_2-K_1} + \dfrac{f_1}{K_1+\frac{4}{3}G_1}}
\end{aligned} \tag{9.26}$$

Substituting (9.26) into (9.25) yields a lower bound on bulk modulus

$$\bar{K} \ge K_1 + \frac{f_2}{\dfrac{1}{K_2-K_1} + \dfrac{f_1}{K_1+\frac{4}{3}G_1}} \tag{9.27}$$

Step 2: Let

$$K_0 = K_2, K = K_1, \ \text{ and } \ G_0 = G_2, G = G_1$$

and choose

$$p^{(1)}_{ij} = p\delta_{ij}, \ \ p^{(2)}_{ij} = 0 \ .$$

One can find an upper bound,

$$\bar{I}(p) = \frac{9}{2}K_2\bar{\epsilon}^2 - \frac{f_1p^2}{2(K_1-K_2)} - \frac{1}{2}\frac{f_1f_2p^2}{K_2+\frac{4}{3}G_2} + 3f_1p\bar{\epsilon} \ge \frac{9}{2}\bar{K}\bar{\epsilon}^2 \tag{9.28}$$

To find the maximum value of $\bar{I}(p)$, we examine the stationary condition,

$$\frac{\partial \bar{I}}{\partial p} = 0, \quad \Rightarrow \quad p_{sta} = \frac{3\bar{\epsilon}}{\dfrac{1}{K_1-K_2} + \dfrac{f_2}{K_2+\frac{4}{3}G_2}} \tag{9.29}$$

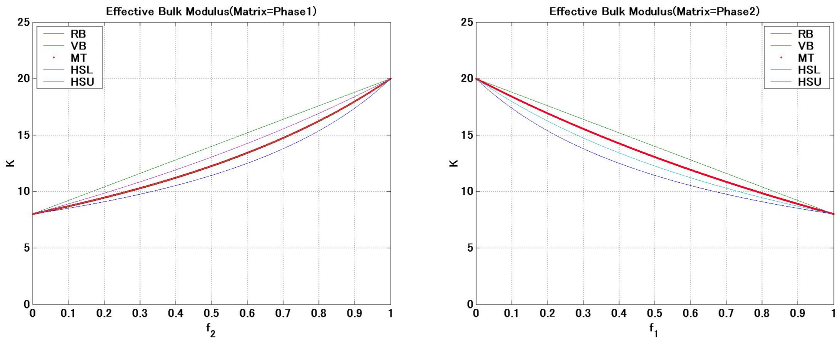


Figure 9.1. Variational Bounds for Bulk Modulus: (a) Medium one, and (b) Medium two.

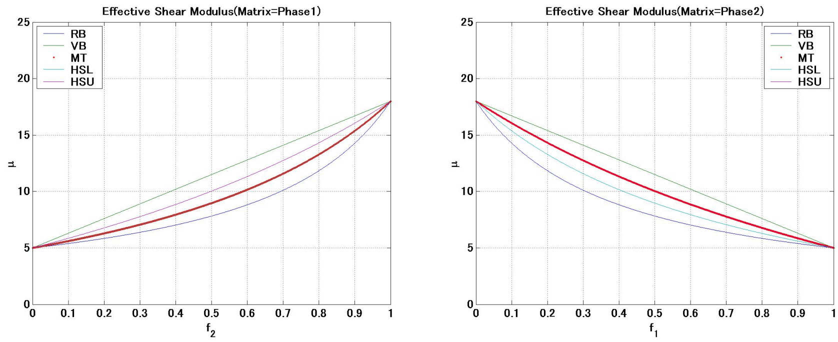


Figure 9.2. Variational Bounds for Shear Modulus: (a) Medium one, and (b) Medium two.

Substituting (9.29) into (9.28), one will find that

$$\bar{K} \leq K_2 + \frac{f_2}{\dfrac{1}{K_1 - K_2} + \dfrac{f_2}{K_2 + \frac{4}{3}G_2}} \tag{9.30}$$

By combining (9.27) and (9.30), we will have the Hashin-Shtrikman bound on bulk modulus,

$$K_1 + \frac{f_2}{\dfrac{1}{K_2 - K_1} + \dfrac{f_1}{K_1 + \frac{4}{3}G_1}} \leq \bar{K} \leq K_2 + \frac{f_1}{\dfrac{1}{K_1 - K_2} + \dfrac{f_2}{K_2 + \frac{4}{3}G_2}} \tag{9.31}$$

It is readily to show that the following Hashin-Shtrikman bounds are held for shear modulus,

$$G_1 + \frac{f_2}{\dfrac{1}{G_2 - G_1} + \dfrac{6(K_1 + 2G_1)f_1}{5(3K_1 + 4G_1)G_1}} \leq \bar{G} \leq G_2 + \frac{f_1}{\dfrac{1}{G_1 - G_2} + \dfrac{6(K_2 + 2G_2)f_2}{5(3K_2 + 4G_2)G_2}} \tag{9.32}$$

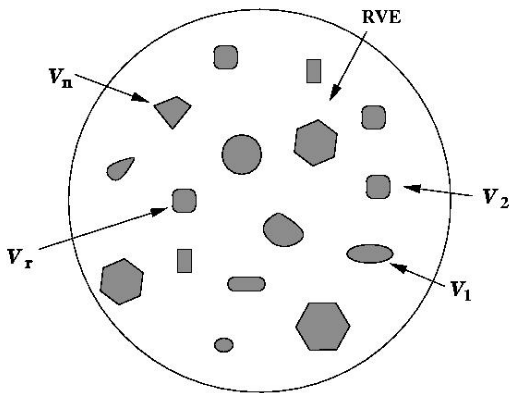


Figure 9.3. A compsite with n-phases

9.2 Microstructure Characterization

9.2.1 Preliminary

In this section, a few important concepts about statistical evaluate of a random heterogeneous material shall be discussed, or formally defined. First, we assume that any sample of a random heterogenous material is a realization of a specific random or stochatic process. Mathematically speaking, a realization is an event, α, that belongs to a sample space, $\mathcal{S}$. Second, an *ensemble* is the collection of all the possible realizations of a random medium generalized by a specific stochastic process.

Consider a sample space $\mathcal{S}$ over which a probability density function, $p(\alpha)$, is defined, $\alpha \in \mathcal{S}$. Then any particular property, f, of a composite (such as mass density, volume fraction density) is a function of α, and its *ensemble* average can defined as

$$< f >= \int_{\mathcal{S}} f(\alpha) p(\alpha) d\alpha \tag{9.33}$$

Of particular interest is the indicator function, Suppose that there is a n-phase ramdom medium (composite), $\mathcal{V} \in \mathbb{R}^d$. The total volume of $\mathcal{V}$ is partition into n-disjoint random sets or phases. The phase 1 occupies the set $\mathcal{V}_1$, and, in general, the phase r occupies the $\mathcal{V}_r, r = 1, 2, \cdots, n$. The measure of set $\mathcal{V}_r$ is denoted as volume fraction, $f_r = meas(\mathcal{V}_r)$. Obviously, the set $\{\mathcal{V}_r\}$ is a subdivision, i.e.

$$\begin{aligned} \bigcup_{r=1}^{n} \mathcal{V}_r(\alpha) &= V, \\ \mathcal{V}_i \cap \mathcal{V}_j &= \emptyset, \text{ if } i \neq j \end{aligned}$$

The indicator function for the phase, r, is defined as

$$I^{(r)}(\mathbf{x},\alpha) = \begin{cases} 1, & \text{if } \mathbf{x} \in \mathcal{V}_r(\alpha) \\ 0, & \text{otherwise} \end{cases} \tag{9.34}$$

The indicator function is a partition of unity,

$$\sum_{r=1}^{n} I^{(r)}(\mathbf{x},\alpha) = 1 \ .$$

In many mathmatical literature, the indicator function is also called as charateristic function.

The expectation or probability of finding phase R ar a chosen point, $\mathbf{x}$, is then denoted as

$$\mathcal{S}_1^r(\mathbf{x}) :=< I^{(r)}(\mathbf{x}) >= \int_{\mathcal{S}} I^{(r)}(\mathbf{x},\alpha)p(\alpha)d\alpha = \mathcal{P}\left\{ I^{(r)}(\mathbf{x}) = 1 \right\} \tag{9.35}$$

In the literature, the function, $\mathcal{S}_1^{(r)}$, is referred to as the one-point probability function for phase, r, since it gives the probability to find phase r at position $\mathbf{x}$. It is also referred to as the one-point correlation function for the phase indicator function, $I^{(r)}$.

In general, the expectation, or probablity, to find the phase, r, at different n points simulatenously is referred to as the n-point probability function, which is defined as

$$\mathcal{S}_n^{(r)}(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) :=< I^{(r)}(\mathbf{x}_1) I^{(r)}(\mathbf{x}_2) \cdots I^{(r)}(\mathbf{x}_n) > \tag{9.36}$$

Here the subscript, n, indicates that this is a n-point probability function, and the superscript, (r), denotes that this is a n-point correlation function for phase r.

One can further generalize the above concept of correlation function to the probability of finding any subset of points n_i of the n points in phase i and another subset of points n_j of the n points in phase j as

$$\mathcal{S}_n^{(ij)}(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) :=< I^{(i)}(\mathbf{x}_1) \cdots I^{(i)}(\mathbf{x}_{n_i}) I^{(j)}(\mathbf{x}_{n_i+1}) \cdots I^{(j)}(\mathbf{x}_n) > \tag{9.37}$$

For instance, a two-point correlation function that represents the probability to find the phase, r, in $\mathbf{x}_1$ and the phase, s, in $\mathbf{x}_2$ is defined as

$$\mathcal{S}_2^{(rs)}(\mathbf{x}_1, \mathbf{x}_2) :=< I^{(r)}(\mathbf{x}_1) I^{(s)}(\mathbf{x}_2) > \tag{9.38}$$

Consider a n-phase composite. Its mass density can be expressed as

$$\rho(\mathbf{x}) = \sum_{r=1}^{n} \rho_r I^{(r)}(\mathbf{x}) \tag{9.39}$$

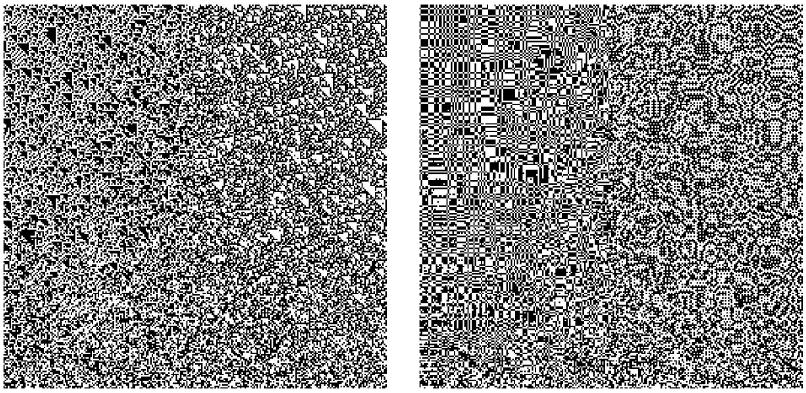


Figure 9.4. Examples of statistically inhomogeneous mdeia

Then the expectation of the density function is

$$
\begin{aligned}
< \rho(\mathbf{x}) > &= \int_{\mathcal{S}} \sum_{r=1}^{n} \rho_r I^{(r)}(\mathbf{x}, \alpha) p(\alpha) d\alpha \\
&= \sum_{r=1}^{n} \rho_r \mathcal{S}_1^{(r)}(\mathbf{x})
\end{aligned}
$$

The expectation of the product of $\rho(\mathbf{x}_1)$ and $\rho(\mathbf{x}_2)$ is

$$
\begin{aligned}
< \rho(\mathbf{x}_1)\rho(\mathbf{x}_2) > &= \int_{\mathcal{S}} \Big(\sum_{r=1}^{n} \rho_r I^{(r)}(\mathbf{x}_1, \alpha) \Big) \Big(\sum_{s=1}^{n} \rho_s I^{(s)}(\mathbf{x}_2, \alpha) \Big) p(\alpha) d\alpha \\
&= \sum_{r=1}^{n} \sum_{s=1}^{n} \rho_r \rho_s S_2^{(rs)}(\mathbf{x}_1, \mathbf{x}_2)
\end{aligned}
$$

9.2.2 Symmetry and Ergodicity

If a n-point probability function, $S_n^{(r)}$ depends on the absolute positions, $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n$, explicitely, i.e.

$$
S_n^{(r)} = S_n^{(i)}(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) = \int_{\mathcal{S}} I^{(i)}(\mathbf{x}_1, \alpha) I^{(i)}(\mathbf{x}_2, \alpha) \cdots I^{(i)}(\mathbf{x}_n, \alpha) p(\alpha) d\alpha,
$$

we say that the medium is strictly statistically inhomogeneous. Examples of statistically inhomogeneous media are shown in Fig. (9.5)

We say that a system is statically homogeneous, or when a stochastic spatial distribution is homogeneous, if $S_n^{(i)}(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)$ is invariant under translation, i.e. $\forall\, \mathbf{y} \in \mathbb{R}^d$,

$$\begin{aligned} S_n^{(i)}(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n) &= S_n^{(i)}(\mathbf{x}_1 + \mathbf{y}, \mathbf{x}_2 + \mathbf{y}, \cdots, \mathbf{x}_n + \mathbf{y}) \\ &= S_n^{(i)}(\mathbf{x}_{12}, \mathbf{x}_{13}, \cdots, \mathbf{x}_{1n}) \quad (\Leftarrow \mathbf{y} = -\mathbf{x}_1) \end{aligned} \tag{9.40}$$

where $\mathbf{x}_{jk} = \mathbf{x}_k - \mathbf{x}_j$. Obviously, in this case, $V = \mathbb{R}^d$ and $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n \in \mathbb{R}^d$.

When a system is statistically homogeneous, or when a stochastic spatial distribution is homogeneous, one can relate ensemble (time) average to the volume (spatial) average. This is because that material properties in every regions of the space are similar, and hence any realization of a statistical ensemble must contain the all statistical information or details as the rest of other realizations do, provided that the spatial realization space is large enough to render a stable statistical interpretation.

This suggests an ergodic hypotheis: *The result of averaging over all realizations of the ensemble is equivalent to averaging over the volume of one realization in an infinite-volume limit.*

Under the ergodic assumption, the complete probabilistic information can be obtained from a single realization of an infinite domain. By letting

$$\alpha = \mathbf{y}, \;\; p(\alpha) = \frac{1}{V}, \;\; \text{and} \;\; d\alpha = dV_{\mathbf{y}}$$

the ergodic hypothsis enables us to replace ensemble averaging with volume averaging in the limit that the volume tends to infinity, i.e.

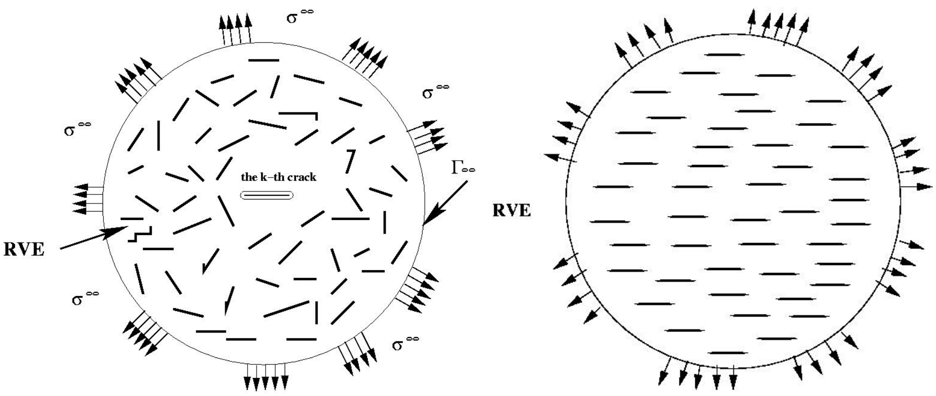
$$S_n^{(i)} = \lim_{V \to \infty} \frac{1}{V} \int_V I^{(i)}(\mathbf{y}) I^{(i)}(\mathbf{y} + \mathbf{x}_{12}) \cdots I^{(i)}(\mathbf{y} + \mathbf{x}_{1n}) d\mathbf{y}$$

We refer to such systems as ergodic media.

REMARK 9.2.1 *Ergodicity is a mathematics term, meaning “ space filling”. Ergodic theory has its origins from the work of Boltzmann in statistical physics. Ergodic theory in statistical mechanics refers to where time- and space-distribution averages are equal. Steinhaus (1983, pp. 237-239) gives a practical analogy to ergodic theory as to keeping one’s feet dry ("in most cases," "stormy weather excepted") when walking along a shoreline without having to constantly turn one’s head to anticipate incoming waves. The mathematical origins of ergodic theory are due to von Neumann, Birkhoff, and Koopman*

In practice, instead of using the infinite spatial space, if a domain is much larger than a basic spatial mechanical element, we usually take it as the spatial sampling space that is the so-called *representative volume element* (RVE).

One can see that for statistically homogeneous media, the n-point probability function do not depend on their absolute positions, but on their relative



(a) (b)

Figure 9.5. Examples of homogeneous isotropic (a) and homogeneous anisotropic media.

displacement. Therefore, there is no preferred origin in the system. In Eq. (9.40), $\mathbf{x} = \mathbf{x}_1$ is chosen as the origin of the coordinate.

For one-point probability function (or one-point correlation function), we then have

$$S_1^{(r)} := \frac{1}{V}\int_V I^{(r)}(\mathbf{x},\mathbf{y})dV_{\mathbf{y}} = \frac{1}{V}\int_V H(\mathcal{V}_r)dV_{\mathbf{y}} = \frac{1}{V}\int_{\mathcal{V}_r} dV_{\mathbf{y}} = f_r \quad (9.41)$$

which is the volume fraction of the phase r.

If the n-point probability function of a medium is both translation and rotation invariant, the medium is called isotropic homogeneous. It means that the n-point correlation function only depend on the distance among the particles. For instance,

$$\begin{aligned} S_2^{(r)}(\mathbf{x}_1,\mathbf{x}_2) &= S_2^{(r)}(x_{12}) \\ S_3^{(r)}(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3) &= S_2^{(r)}(x_{12},x_{13}) \end{aligned}$$

where $x_{kj} = \|\mathbf{x}_j - \mathbf{x}_k\|$.

9.2.3 Applications

EXAMPLE 9.1 *Consider Voigt bound and Reuss bound,*

$$\Big(\sum_{r=1}^{n} f_r \mathbf{C}^{r-1}\Big) \le \bar{\mathbf{C}} \le \Big(\sum_{r=1}^{n} f_r \mathbf{C}^{r}\Big)$$

Both these two bounds only require information of volume fraction of each phase. Since volume fraction,

$$f_r = S_1^{(r)}(\mathbf{x}),$$

is, by definition, the one-point probability function (or correlation function), both Voigt bound and Reuss bound are called as one-point bound.

EXAMPLE 9.2 *To evaluate Hashin-Shtrikman bound, we may let*

$$\mathbf{p}(\mathbf{x}) = \sum_{r=1}^{n} \mathbf{p}_r I^{(r)}(\mathbf{x})$$

where $\mathbf{p}_r$ is a constant second order tensor.
Then

$$< \mathbf{p} >= \sum_{r=1}^{n} \mathbf{p}_r < I^{(r)}(\mathbf{x}) >= \sum_{r=1}^{n} \mathbf{p}_r S_1^{(r)}(\mathbf{x}) = \sum_{r=1}^{n} f_r \mathbf{p}_r$$

Therefore,

$$\begin{aligned}
\frac{1}{V}\int_V \mathbf{p} : \boldsymbol{\epsilon}^d dV &= \frac{1}{2V}\int_V \Big(\sum_{r=1}^{n} \mathbf{p}_r I^{(r)}(\mathbf{x})\Big) : \\
&\Big(-\int_{V'} \boldsymbol{\Gamma}^{\infty}(\mathbf{x}' - \mathbf{x}) : \Big[\mathbf{p}(\mathbf{x}^{'} - < \mathbf{p} >\Big]\Big) dV_{\mathbf{x}'} dV_{\mathbf{x}} \\
= & -\frac{1}{2V}\int_V \sum_{r=1}^{n} \mathbf{p}_r I^{(r)}(\mathbf{x}) : \Big(\int_{V''} \boldsymbol{\Gamma}^{\infty}(\mathbf{x}^{''}) \\
& : \Big[\sum_{s=1}^{n} \mathbf{p}_s I^{(s)}(\mathbf{x} + \mathbf{x}^{''}) - \sum_{s=1}^{n} f_s \mathbf{p}_s\Big]\Big) dV_{\mathbf{x}''} dV_{\mathbf{x}} \\
= & -\frac{1}{2}\sum_{r=1}^{n}\sum_{s=1}^{n} \mathbf{p}_r : \Big(\int_{V''} \boldsymbol{\Gamma}^{\infty}(\mathbf{x}^{''}) dV_{\mathbf{x}''}\Big) \\
& : \Big\{\frac{1}{V}\int_V \Big(I^{(r)}(\mathbf{x}) I^{(s)}(\mathbf{x} + \mathbf{x}^{''}) - I^{(r)} f_s\Big) \mathbf{p}_s dV_{\mathbf{x}}\Big\} \\
= & -\frac{1}{2}\sum_{r=1}^{n}\sum_{s=1}^{n} \mathbf{p}_r : \Big(\int_{V''} \boldsymbol{\Gamma}^{\infty}(\mathbf{x}^{''}) dV_{\mathbf{x}''}\Big) \\
& : \Big(S_2^{(rs)}(\mathbf{x}, \mathbf{x} + \mathbf{x}^{''}) - f_r f_s\Big) \mathbf{p}_s
\end{aligned}$$

Assume that the composite possesses no long-range interaction. The mathematical implication is that

$$S_2^{(rs)}(\mathbf{x}, \mathbf{x}^{'}) = S_1^{(r)}(\mathbf{x}) S_1^{(s)}(\mathbf{x}^{'}), \ \text{ when } \ \|\mathbf{x} - \mathbf{x}^{'}\| >> 1$$

because the probability of two independent events occur simulatenously should equal to the product of the probability of two single events.

One the other hand, when $\|\mathbf{x}-\mathbf{x}'\| \le R_r$ or R_s. There can be only one phase exists within the region, hence

$$S_2^{(rs)}(\mathbf{x},\mathbf{x}') = S_1^{(r)}\delta_{rs}$$

To sum up

$$S_2^{(rs)}(\mathbf{x},\mathbf{x}') = \begin{cases} f_r\delta_{rs} & \|\mathbf{x}-\mathbf{x}'\| \le R_r \\ \\ f_r f_s & \|\mathbf{x}-\mathbf{x}'\| > R_r \end{cases}$$

Again, we end with the relationship,

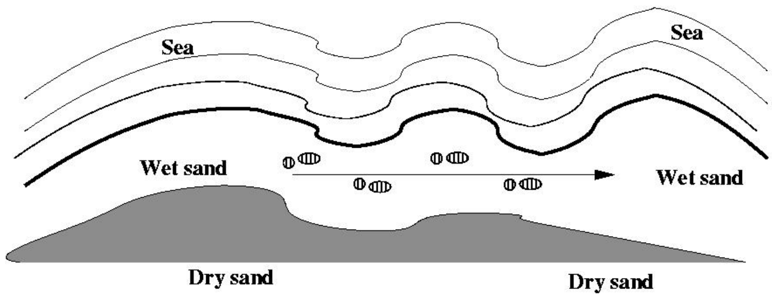
$$\begin{aligned} \frac{1}{2V}\int_V \mathbf{p} : \boldsymbol{\epsilon}^d dV &= -\frac{1}{2}\sum_{r=1}^{n}\sum_{s=1}^{n} p_{ij}^r P_{ijk\ell}^r \Big(f_r\delta_{rs} - f_r f_s\Big) p_{k\ell}^s \\ &= -\frac{1}{2}\sum_{r=1}^{n} f_r p_{ij}^r P_{ijk\ell}^r \Big(\sum_{s=1}^{n}(\delta_{rs}p_{k\ell}^s - f_s p_{k\ell}^s\Big) \\ &= -\frac{1}{2}\sum_{r=1}^{n} f_r p_{ij}^r P_{ijk\ell}^r \Big(p_{k\ell}^r - < p_{k\ell} >\Big) \end{aligned}$$

which was derived previously by using the argument of Mori-Tanaka theorem.

As shown above, the evaluation of Hashin-Shtrikman bounds is intimately related with the evaluation of two-point probability function, or two-point correlation function, $S_2^{(rs)}$. It is this reason that Hashin-Shtrikman bounds are called two-point bounds.

9.2.4 Ergodic principle

The intuistive concept of Ergodicity was popularized by Hugo Steinhaus. Steinhaus wrote in his well-known book *Mathematical Snapshots*,



"When strolling along a sandy beach in shores most people choose the wet strip left by retreating waves, which is hard and smooth enough to make the walk more comfortable than the dry part of the beach. On the other hand, to avoid their shoes and socks being soaked they must constantky watch the play the surf licking the strip. This steady twisting of the neck becomes disagreeable after a few minutes. There is, however, a remedy. Instead of looking sidewise one keeps looking straight ahead; in every instant he sees the instantaneous water edge and he directs his steps tangentially; he walks along a line touching the edge in a single point without cutting contact lies far enough away to render the variations small and easily accounted for: neither looking to the left, nor sudden jumping to the right is necessary.

The background for the behavior I recommend here (after having tried it) is the ' **ergodic principle** *': the distribution of water tongues licking the shore in a fixed point observed during a long time is the same as the distributions shown in a fixed moment by a long portion of the water edge — the principle involved is the identity of time-distribution and space-distribution. To apply it here the walker has to limit his observation to the part of the shore he will cover in the next minute — in most cases such tactics keep him on the safe side without leading him out of the wet strip of the beach. $\cdots\cdots$"*

I thought that some explanation may be needed to correctly understand Steinhaus' analogy:

What Steinhaus was trying to say is that consider an infinite set of good weather day, if a person comes to a beach every afternoon at 2:00 clock he may find that at a particular spot (fixed spatial location) the sea water line on the beach is a stochastic event and all the measurement on water line on each day consist of a statistic ensemble. We assume that there is a statistical average value for the sea water line on that spot, which is the average in time. The ergodic principle suggests that if a system is both homogeneous in space and in time, one can then find that average without measuring water line at 2:00 pm on infinite days. Instead, he can just walk along a path that is tangential to the water (shore) line on the beach, which is also assumed to "infinite". By *doing so, the average position along his path on the beach may be equal to the statistical average of the time ensemble.*

Note that we do not consider the the surge or recede of sea water line due to the effect of tide. Hence, the person who is in charge the measurement has to come to the beach every afternoon at the same time (e.g. 2:00 pm), provided that the weather is always good.

9.3 Exercises

PROBELM 9.1 *Show that for a spherical inlusion, $\Omega \subset V$,*

$$\begin{aligned} \mathbf{P} &:= \int_\Omega \mathbf{\Gamma}^\infty(\mathbf{y}-\mathbf{x})dV_{\mathbf{y}} \\ &= \frac{1}{4\pi}\int_{|\boldsymbol{\xi}|=1} \tilde{\mathbf{\Gamma}}^\infty(\boldsymbol{\xi})dS \end{aligned} \tag{9.42}$$

PROBELM 9.2 *Consider a well-order two phase composite ($K_2 > K_1$ and $G_2 > G_1$). Derive the Hashin-Shtrikman bounds for shear modulus,*

$$G_1 + \cfrac{f_2}{\cfrac{1}{G_2 - G_1} + \cfrac{6(K_1+2G_1)f_1}{5(3K_1+4G_1)G_1}} \le \bar{G} \le G_2 + \cfrac{f_1}{\cfrac{1}{G_1 - G_2} + \cfrac{6(K_2+2G_2)f_2}{5(3K_2+4G_2)G_2}} \tag{9.43}$$

Assume that $K_1 = 8GP_a$ & $G_1 = 5GP_a$ and $K_2 = 20.0GP_a$ & $G_2 = 18GP_a$. Plot the Voigt bound, Ruess bound, Mori-Tanaka, and Hashin-Shtrikman bounds for both bulk modulus and shear modulus for comparison.

Hints:

Hashin, Z. and Shtrikman, S. [1961], "Note on a variational approach to the theory of composite elastic materials," The Frabklin Institute Laboratories, pp. 336-341.

Hashin, Z. and Shtrikman, S. [1962a], "On some variational principles in anisotropic and non-homogeneous elasticity," Journal of Mechanics and Physics of Solids, Vol. 10, pp. 335-342.

Hashin, Z. and Shtrikman, S. [1962b], "A variational approach to the theory of the elastic behavior of polycrystals," Journal of Mechanics and Physics of Solids, Vol. 10, pp. 343-352.

EXAMPLE 9.3 *Consider a two-phase fiber reinforced composite as shown in Figure (9.6) . Use two-dimensional Hashin-Shtrikman bounds to find the in-plane (or transverse) bulk modulus and shear modulus.*

Hints:

Hashin, Z. [1965] "On elastic behaviour of fibre reinforced materials of arbitrary transverse phase geometry," Journal of Mechanics and Physics of Solids, Vol. 13, pp. 119-113.

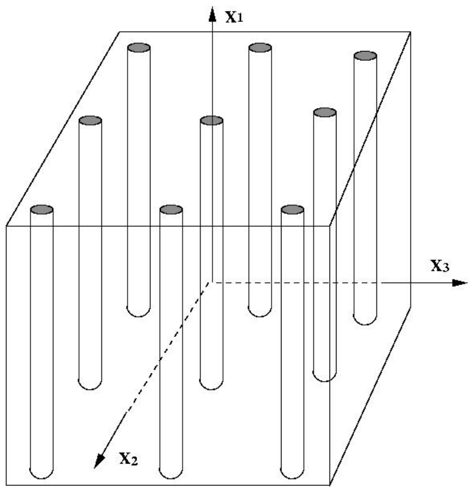


Figure 9.6. Cylindrical fibre-reinforced composite

Torquato, S. [2002] Random Heterogeneous Materials, Springer, New York, pp. 328-337.

Christensen, R. M. [1979],

Mechanics of Composite Materials, Chapter III;

Chapter 10

PERIODIC MICROSTRUCTURE

In engineering applications, often times, we encounter situations where materials have periodic structure. Such examples are various composites with periodic structure, reticulated structures (see Fig. (10.1), DNA, masonary structures, so forth. In fact, at very fine scale, most metals may be regarded as composites with periodic structure because of their lattice structures. There are mainly two types of methodologies in analysis: (1) equivalent eigenstrain

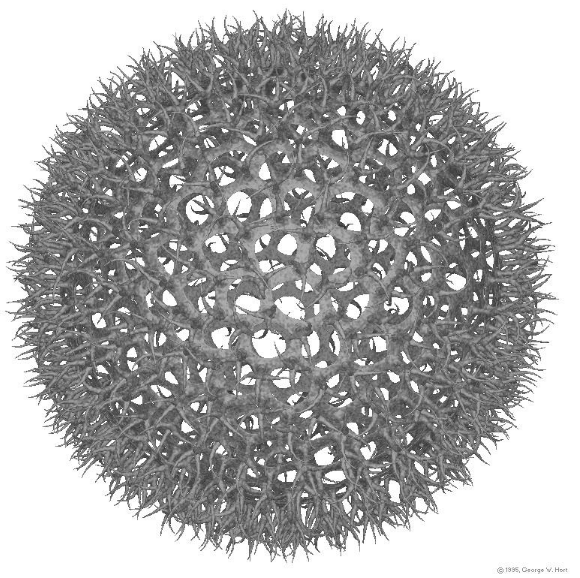


Figure 10.1. An example of periodic reticulated structure

approach, and (2) asymptotic homogenization. We first start with equivalent eigenstrain approach.

10.1 Unit cell and Fourier series

Conisder a rectangular unit cell defined as

$$Y := \left\{ \mathbf{x} \,\middle|\, -a_j \le x_j \le a_j, \;\; j = 1, 2, 3 \right\} \tag{10.1}$$

where a_j is the half length of the unit cell in j-th direction.

For materials with periodic structures, material properties should be periodic functions, i.e.

$$\mathbf{C}(\mathbf{x} + \mathbf{d}) = \mathbf{C}(\mathbf{x})$$

where $\mathbf{d} = \sum_{j=1}^{3} 2m_j a_j \mathbf{e}_j$, $j = 1, 2, 3$. Here m_j are arbitrary integers. The vector, $\mathbf{d}$, is not the minimum periodicity, unless $m_j = 1$.

Under certain conditions, it is possible that displacement field may be periodic as well, i.e.

$$\mathbf{u}(\mathbf{x} + \mathbf{d}) = \mathbf{u}(\mathbf{x})$$

An immediate consequence is that strain field is periodic,

$$\boldsymbol{\epsilon}(\mathbf{x} + \mathbf{d}) = \boldsymbol{\epsilon}(\mathbf{x})$$

Nevertheless, periodic strain field does not necessarily produce periodic displacement field. For instance, a constant strain field is periodic,

$$\boldsymbol{\epsilon}(\mathbf{x} + \mathbf{d}) = \boldsymbol{\epsilon}(\mathbf{x}) = \boldsymbol{\epsilon}^0, \quad \forall \mathbf{d} \in \mathbb{R}^3,$$

but it does not generate a periodic displacement field, instead $\mathbf{u}(\mathbf{x}) = \mathbf{x} \cdot \boldsymbol{\epsilon}^0$, and $\mathbf{u}(\mathbf{x} + \mathbf{d}) \neq \mathbf{u}(\mathbf{x})$.

A convenient mathematical tool to treat periodic functions is Fourier series. Define a vector,

$$\boldsymbol{\xi} = \xi_j \mathbf{e}_j, \;\; \text{and} \;\; \xi_j = \frac{n_j \pi}{a_j}, \;\; n_j = 0, \pm 1, \pm 2, \cdots, \cdots$$

and a countable set,

$$\Lambda = \left\{ \boldsymbol{\xi} = \xi_j \mathbf{e}_j \,\middle|\, \xi_j \frac{n_j \pi}{a_j}, n_j = 0, \pm 1, \pm 2, \cdots, \cdots, \right\} \tag{10.2}$$

For any real function, $f(\mathbf{x}) \in C^1(Y)$, $f(\mathbf{x})$ can be expanded into Fourier series,

$$f(\mathbf{x}) = \sum_{\boldsymbol{\xi} \in \Lambda} \mathcal{F}[f](\boldsymbol{\xi}) \exp(\mathrm{i} \boldsymbol{\xi} \cdot \mathbf{x}), \quad \mathrm{i} = \sqrt{-1}, \tag{10.3}$$

where the Fourier coefficient is

$$\mathcal{F}[f](\boldsymbol{\xi}) = \frac{1}{|Y|}\int_Y \mathbf{u}(\mathbf{x})\exp(-\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi})dV_{\mathbf{x}}$$

where $|Y|$ is the volume of the unit cell. For a rectangular unit cell, $|Y| = 8a_1a_2a_3$.

Recall the definition of Fourier series in an 1D interval, $[-a, a]$,

$$\begin{aligned} f(x) &= \sum_{n=-\infty}^{\infty} \mathcal{F}[f](\xi)\exp\Big(\mathrm{i}\frac{n\pi}{a}x\Big), \quad n = 0, \pm 1, \pm 2, \cdots, \\ \mathcal{F}[f] &= \frac{1}{2a}\int_{-a}^{a} f(x)\exp(-\mathrm{i}\frac{n\pi}{a}x)dx \end{aligned}$$

and the orthonormal condition

$$\frac{1}{2a}\int_{-a}^{a}\exp(\mathrm{i}x\xi_m)\exp(-\mathrm{i}x\xi_n)dx = \delta_{mn}$$

where $\xi_n = \dfrac{n\pi}{a}$ and $\xi_m = \dfrac{m\pi}{a}$.

Accordingly, 3D orthonormal condition is

$$\frac{1}{|Y|}\int_Y \exp(\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi})\exp(-\mathrm{i}\mathbf{x}\cdot\boldsymbol{\zeta})dV_{\mathbf{x}} = \begin{cases} 1 & \boldsymbol{\xi} = \boldsymbol{\zeta} \\ 0 & \boldsymbol{\xi} \neq \boldsymbol{\zeta} \end{cases}$$

where $\boldsymbol{\xi}, \boldsymbol{\zeta} \in \Lambda$, i.e.

$$\boldsymbol{\xi} = \xi_j\mathbf{e}_j = \frac{n_j\pi}{a_j}\mathbf{e}_j \ \text{ and } \ \boldsymbol{\zeta} = \zeta_k\mathbf{e}_k = \frac{n_k\pi}{a_k}\mathbf{e}_k \ .$$

10.1.1 Fourier transform of displacement field and strain field

Suppose that displacement field is periodic. We may exoand displacement field into Fourier series

$$\mathbf{u}(\mathbf{x}) = \sum_{\boldsymbol{\xi}\in\Lambda} \mathcal{F}[\mathbf{u}](\boldsymbol{\xi})\exp(\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi}) \tag{10.4}$$

where

$$\mathcal{F}[\mathbf{u}](\boldsymbol{\xi}) = \frac{1}{|Y|}\int_Y \mathbf{u}(\mathbf{x})\exp(-\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi})dV_{\mathbf{x}}$$

or in component form

$$\mathcal{F}[u_i](\boldsymbol{\xi}) = \frac{1}{|Y|}\int_Y u_i(\mathbf{x})\exp(-\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi})dV_{\mathbf{x}}$$

REMARK 10.1.1 *In literature, the following expression is often used,*

$$\mathbf{u}(\mathbf{x}) = \sum_{\boldsymbol{\xi}\in\Lambda'} \mathcal{F}[\mathbf{u}](\boldsymbol{\xi}) \exp(\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi})$$

where

$$\Lambda' = \left\{ \boldsymbol{\xi} = \xi_j \mathbf{e}_j \;\middle|\; \xi_j = \frac{n_j \pi}{a_j}, \;\; j = \pm 1, \pm 2, \cdots, \cdots \right\}$$

Note that the difference between index set Λ' *and* Λ *is that* $n_j \neq 0$, *or* $\boldsymbol{\xi} \neq 0$.

When $\boldsymbol{\xi} = 0$,

$$\mathcal{F}[\mathbf{u}](0) = \frac{1}{|Y|}\int_Y \mathbf{u}(x) dV_{\mathbf{x}}$$

which is the average displacement field.

On the other hand, if the composite undergoes a rigid body translation, $\mathbf{u}(\mathbf{x}) = \mathbf{u}^0$, *which is not periodic, one may find that*

$$\mathcal{F}[\mathbf{u}](0) = \mathbf{u}^0$$

Obviously, $u = u^0 \notin L^1(\mathbb{R})$ *nor* $u = u^0 \in L^2(\mathbb{R})$. *Convergence issue may rise in mathematical manipulation. Anyway, rigid body translation is a trivial physical motion, we neglect its contribution in Fourier transform by restricting* $\boldsymbol{\xi} \in \Lambda'$.

Now, we consider the Fourier transform of displacement gradient,

$$\nabla \otimes \mathbf{u}(\mathbf{x}) = \sum_{\boldsymbol{\xi}\in\Lambda} \mathcal{F}[\nabla \otimes \mathbf{u}](\boldsymbol{\xi}) \exp(\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi}) \tag{10.5}$$

and

$$\mathcal{F}[\nabla \otimes \mathbf{u}](\boldsymbol{\xi}) = \frac{1}{|Y|}\int_Y \nabla \otimes \mathbf{u}(\mathbf{x}) \exp(-\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi}) dV_{\mathbf{x}}$$

On the other hand, from (10.4), one may find that

$$\begin{aligned} \nabla \otimes \mathbf{u}(\mathbf{x}) &= \sum_{\boldsymbol{\xi}\in\Lambda} \nabla \exp(\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi}) \otimes \mathcal{F}[\mathbf{u}](\boldsymbol{\xi}) && (10.6) \\ &= \mathrm{i}\sum_{\boldsymbol{\xi}\in\Lambda} \boldsymbol{\xi} \otimes \mathcal{F}[\mathbf{u}](\boldsymbol{\xi}) \exp(\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi}) && (10.7) \end{aligned}$$

Comparing (10.5) with (10.7), we have

$$\mathcal{F}[\nabla \otimes \mathbf{u}](\boldsymbol{\xi}) = \mathrm{i}\boldsymbol{\xi} \otimes \mathcal{F}[\mathbf{u}](\boldsymbol{\xi}) \,.$$

Moreover, we may write Fourier series transform of strain field as

$$\boldsymbol{\epsilon}(\mathbf{x}) = \frac{\mathrm{i}}{2} \sum_{\boldsymbol{\xi} \in \Lambda} \Big(\boldsymbol{\xi} \otimes \mathcal{F}[\mathbf{u}](\boldsymbol{\xi}) + \mathcal{F}[\mathbf{u}](\boldsymbol{\xi}) \otimes \boldsymbol{\xi} \Big) \exp(\mathrm{i}\mathbf{x} \cdot \boldsymbol{\xi}) \tag{10.8}$$

From (10.8), we can deduce that

$$\mathcal{F}[\boldsymbol{\epsilon}](\boldsymbol{\xi}) = \frac{\mathrm{i}}{2} \Big(\boldsymbol{\xi} \otimes \mathcal{F}[\mathbf{u}](\boldsymbol{\xi}) + \mathcal{F}[\mathbf{u}](\boldsymbol{\xi}) \otimes \boldsymbol{\xi} \Big)$$

Hence

$$\mathcal{F}[\boldsymbol{\epsilon}](0) = \frac{1}{|Y|} \int_Y \boldsymbol{\epsilon}(\mathbf{x}) dV_{\mathbf{x}} = 0$$

which implies that the average of a periodic strain field is a null field.

10.1.2 Fourier series transform of stress field

Consider a periodic elastic stiffness tensor, $\mathbf{C}(\mathbf{x} + \mathbf{d}) = \mathbf{C}(\mathbf{x})$, which may be expanded into Fourier series,

$$\mathbf{C}(\mathbf{x}) = \sum_{\boldsymbol{\xi} \in \Lambda} \mathcal{F}[\mathbf{C}](\boldsymbol{\xi}) \exp(\mathrm{i}\mathbf{x} \cdot \boldsymbol{\xi}) \tag{10.9}$$

where

$$\mathcal{F}[\mathbf{C}] = \frac{1}{|Y|} \int_Y \mathbf{C}(\mathbf{x}) \exp(-\mathrm{i}\mathbf{x} \cdot \boldsymbol{\xi}) dV_{\mathbf{x}}$$

The corresponding stress field may then be written as

$$\begin{aligned} \boldsymbol{\sigma}(\mathbf{x}) &= \mathbf{C}(\mathbf{x}) : \boldsymbol{\epsilon}(\mathbf{x}) \\ &= \left\{ \sum_{\boldsymbol{\xi} \in \Lambda} \mathcal{F}[\mathbf{C}](\boldsymbol{\xi}) \exp(\mathrm{i}\mathbf{x} \cdot \boldsymbol{\xi}) \right\} : \left\{ \sum_{\boldsymbol{\zeta} \in \Lambda} \mathcal{F}[\boldsymbol{\epsilon}](\boldsymbol{\zeta}) \exp(\mathrm{i}\mathbf{x} \cdot \boldsymbol{\zeta}) \right\} \end{aligned}$$

Let $\boldsymbol{\eta} = \boldsymbol{\xi} + \boldsymbol{\zeta}$ or $\boldsymbol{\xi} = \boldsymbol{\eta} - \boldsymbol{\zeta}$. We have

$$\boldsymbol{\sigma}(\mathbf{x}) = \sum_{\boldsymbol{\eta} \in \Lambda} \left(\sum_{\boldsymbol{\zeta} \in \Lambda} \mathcal{F}[\mathbf{C}](\boldsymbol{\eta} - \boldsymbol{\zeta}) : \mathcal{F}[\boldsymbol{\epsilon}](\boldsymbol{\zeta}) \right) \exp(\mathrm{i}\mathbf{x} \cdot \boldsymbol{\eta})$$

and it is straightforward that

$$\mathcal{F}[\boldsymbol{\sigma}](\boldsymbol{\eta}) = \sum_{\boldsymbol{\zeta} \in \Lambda} \mathcal{F}[\mathbf{C}](\boldsymbol{\eta} - \boldsymbol{\zeta}) : \mathcal{F}[\boldsymbol{\epsilon}](\boldsymbol{\zeta})$$

If $\mathbf{C} = \mathbf{C}^0$ is a constant fourth order tensor,

$$\mathcal{F}[\mathbf{C}](\boldsymbol{\eta} - \boldsymbol{\zeta}) = \mathbf{C}^0, \;\; \boldsymbol{\eta} = \boldsymbol{\zeta}, \quad \text{and} \quad \mathcal{F}[\mathbf{C}](\boldsymbol{\eta} - \boldsymbol{\zeta}) = 0, \;\; \boldsymbol{\eta} \neq \boldsymbol{\zeta},$$

There is only term left,

$$\mathcal{F}[\boldsymbol{\sigma}](\boldsymbol{\eta}) = \mathcal{F}[\mathbf{C}](0) : \mathcal{F}[\boldsymbol{\epsilon}](\boldsymbol{\zeta}) = \mathbf{C}^0 : \mathcal{F}[\epsilon](\boldsymbol{\eta}) \qquad \text{when} \quad \boldsymbol{\eta} = \boldsymbol{\zeta}.$$

Therefore,

$$\begin{aligned}
\boldsymbol{\sigma}(\mathbf{x}) &= \sum_{\boldsymbol{\eta}\in\Lambda} \mathbf{C}^0 : \mathcal{F}[\boldsymbol{\epsilon}](\boldsymbol{\eta}) \exp(\mathrm{i}\mathbf{x}\cdot\boldsymbol{\eta}) \\
&= \frac{\mathrm{i}}{2} \sum_{\boldsymbol{\eta}\in\Lambda} \mathbf{C}^0 : \Big(\boldsymbol{\eta} \otimes \mathcal{F}[\mathbf{u}](\boldsymbol{\eta}) + \mathcal{F}[\mathbf{u}](\boldsymbol{\eta}) \otimes \boldsymbol{\eta}\Big) \exp(\mathrm{i}\mathbf{x}\cdot\boldsymbol{\eta})
\end{aligned}$$

Last, we evaluate Fourier expansion,

$$\nabla\cdot\boldsymbol{\sigma} = \sum_{\boldsymbol{\xi}\in\Lambda} \mathcal{F}[\nabla\cdot\boldsymbol{\sigma}](\boldsymbol{\xi}) \exp(\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi})$$

Via integration by parts,

$$\begin{aligned}
\mathcal{F}[\nabla\cdot\boldsymbol{\sigma}](\boldsymbol{\xi}) &= \frac{1}{|Y|}\int_Y \nabla\cdot\boldsymbol{\sigma}(\mathbf{x})\exp(-\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi}) dV_{bx} \\
&= \frac{1}{|Y|}\int_Y \Big\{\nabla\cdot\Big(\boldsymbol{\sigma}(\mathbf{x})\exp(-\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi})\Big) - \boldsymbol{\sigma}\cdot\Big(\nabla\exp(-\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi})\Big)\Big\} dV_{\mathbf{x}} \\
&= \frac{1}{Y}\Big\{\int_{\partial Y} \mathbf{n}\cdot\boldsymbol{\sigma}(\mathbf{x})\exp(-\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi}) dS \\
&\quad + i\boldsymbol{\xi}\int_Y \boldsymbol{\sigma}(\mathbf{x})\exp(-\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi}) dV_{\mathbf{x}}\Big\} \\
&= \mathrm{i}\boldsymbol{\xi}\int_Y \boldsymbol{\sigma}(\mathbf{x})\exp(-\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi}) dV_{\mathbf{x}}
\end{aligned}$$

because

$$\int_{\partial Y} \mathbf{n}\cdot\boldsymbol{\sigma}(\mathbf{x})\exp(-\mathrm{i}\mathbf{x}\cdot\boldsymbol{\xi}) dS = 0$$

by periodicity. In particular, when $\boldsymbol{\xi} = 0$,

$$\int_{\partial Y} \mathbf{n}\cdot\boldsymbol{\sigma}(\mathbf{x}) dS = 0$$

which stems from the fact that unit cell is in equilibrium.

10.2 Eigenstrain homogenization

Let $\mathbf{C}^M$ and $\mathbf{D}^M$ be elastic stiffness and compliance tensors in the matrix, $\mathbf{C}^\Omega$, $\mathbf{D}^\Omega$ be the effective stiffness and compliance tensors in the second phase, which is assumed to be distributed periodically in the composite. We are looking for finding effective stiffness and compliance tensors, $\bar{\mathbf{C}}$ and $\bar{\mathbf{D}}$.

Consider prescribed macro-strain boundary condition,

$$\boldsymbol{\epsilon} = \mathbf{x} \cdot \boldsymbol{\epsilon}^0, \quad \forall \mathbf{x} \partial V$$

The total strain may be written as

$$\epsilon_{ij} = \epsilon^0_{ij} + \epsilon^d_{ij}, \quad \forall \ \mathbf{x} \in V$$

The stress fields in the matrix and in the second phase are

$$\begin{aligned} \sigma^M_{ij} &= C^M_{ijk\ell}(\epsilon^0_{ij} + \epsilon^d_{ij}), \quad \forall \mathbf{x} \in M = Y/\Omega \\ \sigma^\Omega_{ij} &= C^\Omega_{ijk\ell}(\epsilon^0_{ij} + \epsilon^d_{ij}), \quad \forall \mathbf{x} \in \Omega \end{aligned}$$

They satisfy the equilibrium equations,

$$\sigma^M_{ij,j} = \ = 0, \quad \forall \ \mathbf{x} \in M \tag{10.10}$$

$$\sigma^\Omega_{ij,j} = 0, \quad \forall \ \mathbf{x} \in \Omega \tag{10.11}$$

and continuity condition at interface,

$$u^{d+}_i = u^{d-}_i, \ \ \forall \ \mathbf{x} \in \partial \Omega$$

Consider a eigenstrain field,

$$\epsilon^*_{ij}(\mathbf{x}) = \epsilon^*_{ij}(\mathbf{x}) H(\Omega)$$

Eshelby's equivalent inclusion principle reads as

$$\sigma^\Omega_{ij} = C^\Omega_{ijk\ell}(\epsilon^0_{k\ell}\epsilon^d_{k\ell}) = C^M_{ijk\ell}(\epsilon^0_{k\ell} + \epsilon^d_{k\ell} - \epsilon^*_{k\ell}) \tag{10.12}$$

Substituting (10.12) into (10.11) yields

$$C^M_{ijk\ell}(\epsilon^0_{k\ell} + \epsilon^d_{k\ell} - \epsilon^*_{k\ell})_{,j} = 0, \quad \Rightarrow \quad C^M_{ijk\ell} u^d_{k,\ell j} = C^M_{ijk\ell}\epsilon^*_{k\ell,j} \tag{10.13}$$

Let,

$$\epsilon^*_{k\ell}(\mathbf{x}) = \sum_{\boldsymbol{\xi} \in \Lambda'} \mathcal{F}[\epsilon^*_{k\ell}](\boldsymbol{\xi}) \exp(\mathrm{i}\boldsymbol{\xi} \cdot \mathbf{x}) = \sum_{\boldsymbol{\xi} \in \Lambda'} \hat{\epsilon}^*_{k\ell} \exp(\mathrm{i}\boldsymbol{\xi} \cdot \mathbf{x}) \tag{10.14}$$

where

$$\hat{\epsilon}^*_{k\ell} = \frac{1}{Y}\int_Y \epsilon^*_{k\ell} \exp(-\mathrm{i}\boldsymbol{\xi} \cdot \mathbf{x}) dV_{\mathbf{x}} = \frac{1}{Y}\int_\Omega \epsilon^*_{k\ell} \exp(-\mathrm{i}\boldsymbol{\xi} \cdot \mathbf{x}) dV_{\mathbf{x}}$$

and

$$u_i(\mathbf{x}) = \mathcal{F}[u_i](\boldsymbol{\xi}) \exp(i\mathbf{x} \cdot \mathbf{x}) \exp(i\boldsymbol{\xi} \cdot \mathbf{x}) = \sum_{\boldsymbol{\xi} \in \Lambda'} \hat{u}_i(\boldsymbol{\xi}) \exp(\mathrm{i}\boldsymbol{\xi} \cdot \mathbf{x}) \tag{10.15}$$

where

$$\hat{u}_i(\boldsymbol{\xi}) = \frac{1}{|Y|}\int_Y u_i(\mathbf{x})\exp(-\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x})dV_{\mathbf{x}}$$

Note that uniform eigenstrain is excluded because it induces a divergent displacement field, i.e.

$$u_i^*(\mathbf{x}) = \epsilon_{ij}^{*0}x_j \quad \to \infty \quad \text{as } \mathbf{x}\to\infty$$

Substituting (10.14) and (10.15) into (10.13), we have

$$-C_{ijk\ell}^M \hat{u}_k \xi_\ell \xi_j = \mathrm{i}C_{ijk\ell}^M \hat{\epsilon}_{k\ell}^* \xi_j \tag{10.16}$$

Denote $K_{ik}(\boldsymbol{\xi}) = C_{ijk\ell}^M \xi_\ell \xi_j$ and $K_{ik}^{-1}(\boldsymbol{\xi}) = N_{ik}(\boldsymbol{\xi})/D(\boldsymbol{\xi})$.

$$\mathcal{F}[u_i](\boldsymbol{\xi}) := \hat{u}_i(\boldsymbol{\xi}) = -\mathrm{i}\frac{N_{ik}(\boldsymbol{\xi})}{D(\boldsymbol{\xi})}C_{k\ell mn}^M \epsilon_{mn}^* \xi_\ell \tag{10.17}$$

Recall,

$$\epsilon_{ij}^d(\mathbf{x}) = \frac{\mathrm{i}}{2}\sum_{\boldsymbol{\xi}\in\Lambda'}\Big(\xi_i\mathcal{F}[u_j](\boldsymbol{\xi}) + \mathcal{F}[u_i^d](\boldsymbol{\xi})\xi_j\Big)\exp(i\boldsymbol{\xi}\cdot\mathbf{x})$$

One can write

$$\begin{aligned}
\epsilon_{ij}^d &= \sum_{\boldsymbol{\xi}\in\Lambda'}\frac{1}{2}\Big(\xi_i\xi_\ell\frac{N_{jk}(\boldsymbol{\xi})}{D(\boldsymbol{\xi})}C_{k\ell mn}^M + \xi_j\xi_\ell\frac{N_{ik}(\boldsymbol{\xi})}{D(\boldsymbol{\xi})}C_{k\ell mn}^M\Big)\hat{\epsilon}_{mn}^*\exp(i\boldsymbol{\xi}\cdot\mathbf{x}) \\
&= \sum_{\boldsymbol{\xi}\in\Lambda'} g_{ijmn}(\boldsymbol{\xi})\hat{\epsilon}_{mn}^*\exp(\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x}) \\
&= \frac{1}{|Y|}\sum_{\boldsymbol{\xi}\in\Lambda'} g_{ijmn}(\boldsymbol{\xi})\int_Y \epsilon_{mn}^*(\mathbf{x}')\exp(-\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x}')dV_{\mathbf{x}'}\exp(\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x})
\end{aligned}$$

where a new fourth order tensor g_{ijmn} is defined as

$$g_{ijmn}(\boldsymbol{\xi}) = \frac{1}{2}\Big(\xi_i N_{jk}(\boldsymbol{\xi}) + \xi_j N_{ik}(\boldsymbol{\xi})\Big)\frac{C_{k\ell mn}^M \xi_\ell}{D(\boldsymbol{\xi})} \tag{10.18}$$

For isotropic materials,

$$\begin{aligned}
g_{ijk\ell}(\boldsymbol{\xi}) &= \frac{1}{2\xi^2}\Big[\xi_j(\delta_{i\ell}\xi_k + \delta_{ik}\xi_\ell) + \xi_i(\delta_{j\ell}\xi_k + \delta_{jk}\xi_\ell)\Big] \\
&\quad -\frac{1}{1-\nu}\frac{\xi_i\xi_j\xi_k\xi_\ell}{\xi^4} + \frac{\nu}{1-\nu}\frac{\xi_i\xi_j}{\xi^2}\delta_{k\ell}
\end{aligned} \tag{10.19}$$

Consider the dilute homoegenization scheme,

$$\mathbf{C}^{\Omega} : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d) = \mathbf{C}^M : (\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d - \boldsymbol{\epsilon}^*) \ .$$

We have

$$\boldsymbol{\epsilon}^0 + \boldsymbol{\epsilon}^d = (\mathbf{C}^M - \mathbf{C}^{\Omega})^{-1} : \mathbf{C}^M : \boldsymbol{\epsilon}^*$$

and subsequently,

$$\boldsymbol{\epsilon}^0 = \mathbf{A}^{\Omega} : \boldsymbol{\epsilon}^*(\mathbf{x}) - \boldsymbol{\epsilon}^d(\mathbf{x})$$

This leads to the following integral equation,

$$\begin{aligned}&\epsilon^0_{ij} - A^{\Omega}_{ijmn}\epsilon^*_{mn}(\mathbf{x})\\ &+ \sum_{\boldsymbol{\xi}\in\Lambda'} g_{ijmn}(\boldsymbol{\xi})\frac{1}{|Y|}\int_{\Omega}\epsilon^*_{mn}(\mathbf{x}')\exp(\mathrm{i}(\mathbf{x}-\mathbf{x}')\cdot\boldsymbol{\xi})dV_{\mathbf{x}'} = 0\end{aligned} \tag{10.20}$$

This equation is difficult to solve. Calculate the average $\displaystyle\frac{1}{|\Omega|}\int_{\Omega}(10.20)dV_{\mathbf{x}}$ in the inclusion. One has

$$\begin{aligned}\boldsymbol{\epsilon}^0 \ = \ & \mathbf{A}^{\Omega} : \bar{\boldsymbol{\epsilon}}^* - \sum_{\boldsymbol{\epsilon}\in\Lambda'} \mathbf{g}(\boldsymbol{\xi}) : \Big(\frac{1}{|\Omega|}\int_{\Omega}\exp(\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x})dV_{\mathbf{x}}\Big)\\ &\cdot\Big(\frac{1}{|Y|}\int_{\Omega}\boldsymbol{\epsilon}^*(\mathbf{x}')\exp(-\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x}')dV_{\mathbf{x}'}\Big)\end{aligned}$$

Define a scalar function,

$$g_0(\boldsymbol{\xi}) = \frac{1}{|\Omega|}\int_{\Omega}\exp(\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x})dV_{\mathbf{x}} \tag{10.21}$$

The eigenstrain integral equation may be written as

$$\begin{aligned}&\epsilon^0_{ij} - A^{\Omega}_{ijmn}\bar{\epsilon}^*_{mn} + \sum_{\boldsymbol{\xi}\in\Lambda'} g_{ijmn}(\boldsymbol{\xi})g_0(\boldsymbol{\xi})\\ &\cdot\Big(\frac{1}{|Y|}\int_{\Omega}\epsilon^*_{mn}(\mathbf{x}')\exp(-\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x}')dV_{\mathbf{x}'}\Big) = 0 \ .\end{aligned} \tag{10.22}$$

For prescribed macros stress boundary condition, one may be able to show that

$$\begin{aligned}&\bar{\epsilon}_{ij} - A^{\Omega}_{ijmn}\bar{\epsilon}^*_{mn} + \sum_{\boldsymbol{\xi}\in\Lambda'} g_{ijmn}(\boldsymbol{\xi})g_0(\boldsymbol{\xi})\\ &\cdot\Big(\frac{1}{|Y|}\int_{\Omega}\epsilon^*_{mn}(\mathbf{x}')\exp(-\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x}')dV_{\mathbf{x}'}\Big) = 0 \ .\end{aligned} \tag{10.23}$$

where $\bar{\epsilon}_{ij} = D^M_{ijmn}\sigma^0_{mn}$.

The simplest approach to solve (10.22) is to replace $\boldsymbol{\epsilon}^*(\mathbf{x})$ by its volume average, i.e., $\boldsymbol{\epsilon}^*(\mathbf{x}) \approx \bar{\epsilon}^*$. Therefore,

$$\begin{aligned}
\boldsymbol{\epsilon}^0 &= \mathbf{A}^\Omega : \bar{\boldsymbol{\epsilon}}^* - \sum_{\boldsymbol{\xi}\in\Lambda'} \mathbf{g}(\boldsymbol{\xi}) g_0(\boldsymbol{\xi}) \Big(\frac{1}{|Y|} \int_\Omega \exp(-\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x}') dV_{\mathbf{x}'} \Big) : \bar{\boldsymbol{\epsilon}}^* \\
&= \mathbf{A}^\Omega : \bar{\boldsymbol{\epsilon}}^* - \sum_{\boldsymbol{\xi}\in\Lambda'} f g_0(\boldsymbol{\xi}) g_0(-\boldsymbol{\xi}) \mathbf{g}(\boldsymbol{\xi}) : \bar{\boldsymbol{\epsilon}}^* \\
&= \mathbf{A}^\Omega : \bar{\boldsymbol{\epsilon}}^* - \sum_{\boldsymbol{\xi}\in\Lambda'} f G(\boldsymbol{\xi}) \mathbf{g}(\boldsymbol{\xi}) : \bar{\boldsymbol{\epsilon}}^*
\end{aligned}$$

where $G(\boldsymbol{\xi}) = g_0(\boldsymbol{\xi}) g_0(-\boldsymbol{\xi})$.

Define Eshelby tensor for periodic inhomogeneities,

$$S^\Omega_{ijmn} = \sum_{\boldsymbol{\xi}\in\Lambda'} f G(\boldsymbol{\xi}) g_{ijmn}(\boldsymbol{\xi}) \tag{10.24}$$

We recover the relationship between remote strain and eigenstrain (average eigenstrain be more precise),

$$\epsilon^0_{ij} = \Big(A^\Omega_{ijmn} - S_{ijmn} \Big) \bar{\epsilon}^*_{mn}$$

To this end, the homogenization of a composite with periodic microstructure can follow the same route as the homogenization of a composite with randomly distributed inhomogeneities, if one can find the corresponding Eshelby tensor. The key to evaluate Eshelby tensor is to find function, $G(\boldsymbol{\xi})$.

EXAMPLE 10.1 *Calculate $G(\boldsymbol{\xi})$ for a one-dimensional periodic unit cell as shown in Fig. (10.2).*

One can show that

$$\begin{aligned}
g_0(\xi) &= \frac{1}{2a} \int_{-a}^{a} \exp(i\xi x) dx \\
&= \frac{1}{2a} \frac{1}{i\xi} \exp(i\xi x) \Big|_{-a}^{a} \\
&= \frac{1}{2a\xi \mathrm{i}} \Big[\Big(\cos(\xi a) + \mathrm{i}\sin(\xi a) \Big) - \Big(\cos(\xi a) - \mathrm{i}\sin(\xi a) \Big) \Big] \\
&= \frac{1}{a\xi} \sin(\xi a)
\end{aligned}$$

It is obvious that

$$g_0(-\xi) = g_0(\xi)$$

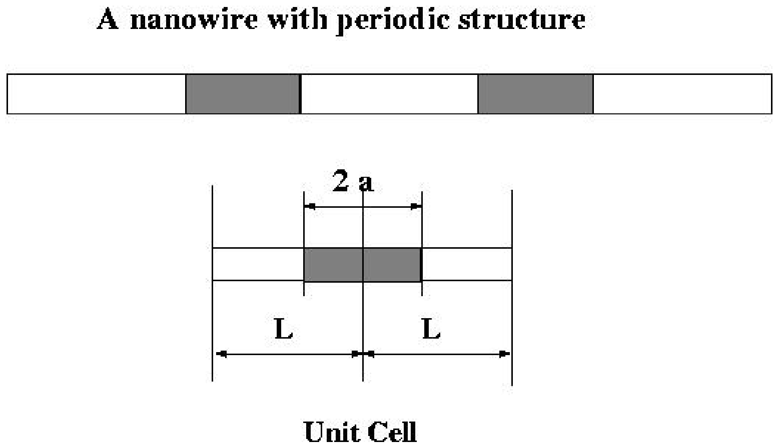


Figure 10.2. An 1D model for a nanowire with periodic structure

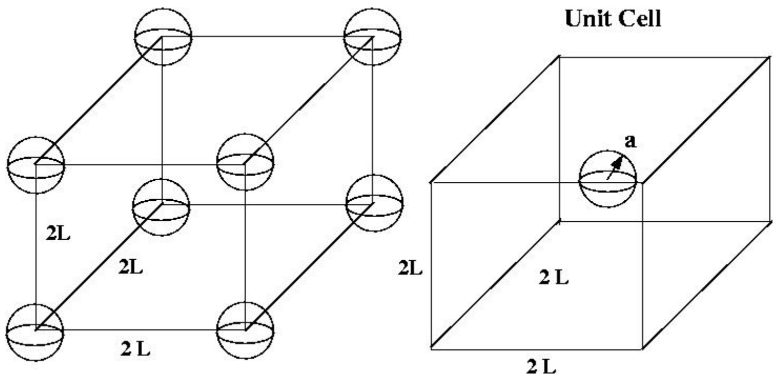


Figure 10.3. Periodic distribution of spherical percipitates.

Hence

$$G(\xi) = \frac{1}{a^2\xi^2}\sin^2(\xi a)$$

EXAMPLE 10.2 *In this example, we consider a spherical percipitate distribution in a cubic lattice as shown in Fig. (10.3). The unit cell in this case is a* $2L \times 2L \times 2L$ *cubic region. There is a spherical ball with radius* $r = a$ *inside the unit cell.*

Recall

$$\int_\Omega \exp(-\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x})d\Omega = (2\pi)^{3/2}a^3\frac{J_{3/2}(\eta)}{\eta^{3/2}}$$

where

$$\begin{aligned}\eta &= a|\xi| = a\sqrt{\xi_1^2+\xi_2^2+\xi_3^2}\\ &= a\sqrt{\Big(\frac{n_1\pi}{L}\Big)^2+\Big(\frac{n_2\pi}{L}\Big)^2+\Big(\frac{n_3\pi}{L}\Big)^2}\\ &= \frac{\pi a}{L}\sqrt{n_1^2+n_2^2+n_3^2} = \frac{\pi a}{L}|\mathbf{n}|\end{aligned}$$

Considering,

$$J_{3/2}(\eta) = \Big(\frac{2}{\pi\eta}\Big)^{1/2}(\eta^{-1}\sin\eta - \cos\eta) = \sqrt{\frac{2}{\pi}}\frac{1}{\eta^{3/2}}(\sin\eta - \eta\cos\eta)$$

one may write

$$\frac{1}{|\Omega|}\int_\Omega \exp(-\mathrm{i}\boldsymbol{\xi}\cdot\mathbf{x})d\Omega = \frac{3}{\eta^3}(\sin\eta - \eta\cos\eta)$$

and

$$G(\xi) = \frac{9}{a^6|\xi|^6}\Big[\sin(a|\xi|) - a|\xi|\cos(a|\xi|)\Big]^2 .$$

One may find that for bcc precipitate cluster,

$$g_0(-\boldsymbol{\xi}) = \frac{3}{\eta^3}(\sin\eta - \eta\cos\eta)\Big(1+\exp(-i\boldsymbol{\xi}\cdot\mathbf{c})\Big)$$

and for fcc precipitate cluster,

$$g_0(-\boldsymbol{\xi}) = \frac{3}{\eta^3}(\sin\eta-\eta\cos\eta)\Big(1+\exp(-i\boldsymbol{\xi}\cdot\mathbf{c}_1)+\exp(-i\boldsymbol{\xi}\cdot\mathbf{c}_2)+\exp(-i\boldsymbol{\xi}\cdot\mathbf{c}_3)\Big)$$

as shown in Fig. (10.4)

10.3 Introduction to Asymptotic Homogenization

The asymptotic method of homogenization is a systematic tool to find effective material properties or effective coefficients of a homogenized differential equation.

The main technique of asymptotic homogenization is the use of multiple-scale expansion. Often times, it involves with singular purturbation technique.

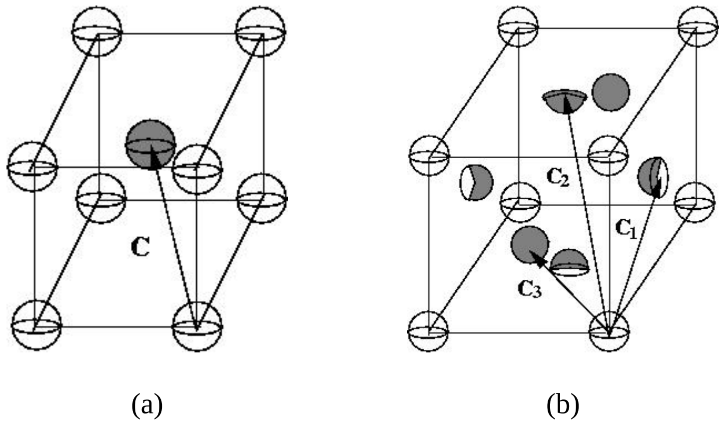


Figure 10.4. Cluster of peripitates in various unit cells: (a) b.c.c. cluster, and (b) f.c.c cluster .

10.3.1 One-dimensional model problem

Consider an 1D model,

$$\frac{d}{dx}\Big(E\frac{du}{dx}\Big) = 0, \quad 0 < x < L \tag{10.25}$$

This equation can be viewed as either the deformation of 1D elastic bar, or 1D steady-state heat diffusion, etc.

Assume that the medium has periodic micro-structure that is varying at microscale, ℓ, which is the characteristic length of a unit cell. Therefore, the coefficient, E, is a periodic function of spatial variable. We also assume that at the interface of two different media in the unit cell the following continuity conditions hold,

$$[u] = 0, \quad \Big[E\frac{du}{dx}\Big] = 0\ .$$

This 1D model problem has a very simple differential equation. An exact solution is possible. In general, for multiple dimension problems or nonlinear problems, analytical solutions may not be possible.

An important characteristics of this problem is the existence of two vastly different length scales: the microscale ℓ, which characterizes the dimension of the unit cell, and the macroscale L, which characterizes the global variations of external force or boundary data.

Suppose that one is more interested in the average variation over a region which is much greater than the typical period and less interested in the detailed variation over a local region. One may ask oneself that

> Can one bypass the details to find an equation governing the variations on the global scale L ?

We define a small paramter $\epsilon = \dfrac{\ell}{L}$. Obviously, $\epsilon << 1$. To separate the effect of two scales, we introduce two coordinates: a fast coordinate and a slow coordinate, which are defined as

$$y \quad \text{and} \quad x = \epsilon y \tag{10.26}$$

You may suggest that the slow coordinate is slowed by small parameter, ϵ. Or vice versa,

$$x \quad \text{and} \quad y = \frac{x}{\epsilon} \tag{10.27}$$

You may suggest that the fast coordinate is speed up by a large paramter $\dfrac{1}{\epsilon}$.

Then, the field variable u may be expressed in a two-scale representation: $u = u(x, y)$ By using chain rule, we may write

$$\frac{d}{dy} = \frac{\partial}{\partial y} + \epsilon \frac{\partial}{\partial x} \tag{10.28}$$

or vice versa,

$$\frac{d}{dx} = \frac{\partial}{\partial x} + \frac{1}{\epsilon}\frac{\partial}{\partial y} \tag{10.29}$$

One can then rewrite Eq. (10.25) as

$$\frac{d}{dy}\Big(E(y)\frac{du}{dy}\Big) = 0, \;\; 0 < y < L \tag{10.30}$$

It is clear that the coefficent has to be a periodic fundtion of fast coordinate, i.e. $E = E(y)$.

Consider the following muti-scale expansion,

$$u(x, y) = u_0(x, y) + \epsilon u_1(x, y) + \epsilon^2 u_2(x, y) + \cdots \tag{10.31}$$

where $u_i(x, y)$ represents activity at i-th scale.

Applying (10.28) to (10.30) leads to the following partial differential equation,

$$\Big(\frac{\partial}{\partial y} + \epsilon\frac{\partial}{\partial x}\Big)\Big[E(y)\Big(\frac{\partial u_0}{\partial y} + \epsilon\Big[\frac{\partial u_0}{\partial x} + \frac{\partial u_1}{\partial y}\Big] + \epsilon^2\Big[\frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y}\Big] \cdots\cdots\Big)\Big] = 0$$

A complete equilibrium implies that equilibrium holds in each scale,

$$\begin{aligned}
\epsilon^0 :& \quad \frac{\partial}{\partial y}\Big[E(y)\frac{\partial u_0}{\partial y}\Big] = 0;\\
\epsilon^1 :& \quad \frac{\partial}{\partial y}\Big[E(y)\Big(\frac{\partial u_0}{\partial x} + \frac{\partial u_1}{\partial y}\Big)\Big] + E(y)\frac{\partial^2 u_0}{\partial x \partial y} = 0;\\
\epsilon^2 :& \quad \frac{\partial}{\partial y}\Big[E(y)\Big(\frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y}\Big)\Big] + E(y)\Big(\frac{\partial^2 u_0}{\partial x^2} + \frac{\partial u^2 u_1}{\partial x \partial y}\Big) = 0;\\
& \quad \cdots\cdots
\end{aligned}$$

We first solve the zero-th order equation,

$$\frac{\partial}{\partial y}\Big(E(y)\frac{\partial u_0}{\partial y}\Big) = 0 \tag{10.32}$$

which only involves with the lowest scale field variable, $u_0(x, y)$.

Integrate (10.32) once,

$$E(y)\frac{\partial u_0}{\partial y} = A_1(x)$$

where $A_1(x)$ is a integration constant.

Integrating second time, we have

$$u_0(x, y) = A_1(x)\int_{y_0}^{y}\frac{d\tilde{y}}{E(\tilde{y})} + A_2(x)$$

Since $u_0(x, y)$ is periodic,

$$u_0(x, y_0) = u_0(x, y_0 + \ell) \ \ \Rightarrow \ \ A_2(x) = A_1(x)\int_{y_0}^{y_0+\ell}\frac{d\tilde{y}}{E(\tilde{y})} + A_2(x)$$

which implies that $A_1(x) = 0$.

This suggests that the leading-order displacement field only depends on the macro-scale variable,

$$u_0 = A_2(x) = u_0(x) \tag{10.33}$$

Now let's examine the first order differential equation,

$$\frac{\partial}{\partial y}\Big[E(y)\Big(\frac{\partial u_0}{\partial x} + \frac{\partial u_1}{\partial y}\Big)\Big] + E(y)\frac{\partial^2 u_0}{\partial x \partial y} = 0 \tag{10.34}$$

Based on (10.33), the last term in (10.34) vanishes.

To solve (10.34), we introduce the following partial separation of variable,

$$u_1(x, y) = Q(x, y)\frac{\partial u_0}{\partial x} + \bar{u}_1(x)$$

where $Q(x,y)$ is an unknown function.

Substitute the above expression into (10.34),

$$\frac{\partial}{\partial y}\left\{E(y)\Big(\frac{\partial u_0}{\partial x}+\frac{\partial Q}{\partial y}\frac{\partial u_0}{\partial x}\Big)\right\}=$$
$$\frac{\partial u_0}{\partial x}\frac{\partial}{\partial y}\left\{E(y)\Big(1+\frac{\partial Q}{\partial y}\Big)\right\}=0\ .$$

This leads to the so-called inhomogeneous canonical cell problem for unknown function, $Q(x,y)$,

$$\frac{\partial}{\partial y}\left\{E(y)\Big(1+\frac{\partial Q}{\partial y}\Big)\right\}=0,\ \ \forall y\in(y_0,y_0+\ell) \tag{10.35}$$

$$[Q]=0,\ \ \text{and}\ \ \Big[E(y)\Big(1+\frac{\partial Q}{\partial y}\Big)\Big]=0\ ,\forall x \text{ at interface.} \tag{10.36}$$

Integrate (10.35) once,

$$E(y)\Big(1+\frac{\partial Q}{\partial y}\Big)=D_1(x)$$
$$\text{or}\ \ \frac{\partial Q}{\partial y}=-1+\frac{D_1(x)}{E(y)}$$

where $D_1(x)$ is an integration constant.

Integrate second times,

$$Q(x,y)=-y+D_1(x)\int_{y_0}^{y}\frac{d\tilde{y}}{E(\tilde{y})}+D_2(x) \tag{10.37}$$

where $D_2(x)$ is another integration constant.

Since $Q(x,y)$ is y-periodic,

$$Q(x,y_0)=Q(x,y_0+\ell)$$

It leads to

$$-y_0+D_2(x)=-(y_0+\ell)+D_1(x)\int_{y_0}^{y_0+\ell}\frac{d\tilde{y}}{E(\tilde{y})}+D_2(x) \tag{10.38}$$

Eq. (10.38) is called the solvability condition for inhomogeneous problem for Q or u_1.

We then find that

$$D_1(x)=\frac{1}{\dfrac{1}{\ell}\displaystyle\int_{y_0}^{y_0+\ell}\frac{d\tilde{y}}{E(\tilde{y})}} \tag{10.39}$$

and hence

$$Q(x,y) = -y + \frac{\displaystyle\int_{y_0}^{y} \frac{d\tilde{y}}{E(\tilde{y})}}{\displaystyle\frac{1}{\ell}\int_{y_0}^{y_0+\ell} \frac{d\tilde{y}}{E(\tilde{y})}} + D_2(x) \tag{10.40}$$

Therefore,

$$u_1(x,y) = \left(-y + \frac{\displaystyle\int_{y_0}^{y} \frac{d\tilde{y}}{E(\tilde{y})}}{\displaystyle\frac{1}{\ell}\int_{y_0}^{y_0+\ell} \frac{d\tilde{y}}{E(\tilde{y})}} + D_2(x)\right)\frac{\partial u_0}{\partial x} + \bar{u}_1(x) \tag{10.41}$$

$$\frac{\partial u_1}{\partial y} = -\frac{\partial u_0}{\partial x} + \frac{1}{E(y)\Big(\displaystyle\frac{1}{\ell}\int_{y_0}^{y_0+\ell} \frac{d\tilde{y}}{E(\tilde{y})}\Big)}\frac{\partial u_0}{\partial x} \tag{10.42}$$

Next, we consider the differential equation at the second scale,

$$\epsilon^2 : \quad \frac{\partial}{\partial y}\Big[E(y)\Big(\frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y}\Big)\Big] + E(y)\Big(\frac{\partial^2 u_0}{\partial x^2} + \frac{\partial u^2 u_1}{\partial x \partial y}\Big) = 0\,. \tag{10.43}$$

Consider

$$\frac{\partial^2 u_1}{\partial x \partial y} = -\frac{\partial^2 u_0}{\partial x^2} + \frac{1}{E(y)\Big(\displaystyle\frac{1}{\ell}\int_{y_0}^{y_0+\ell} \frac{d\tilde{y}}{E(\tilde{y})}\Big)}\frac{\partial^2 u_0}{\partial x^2}$$

Eq. (10.43) becomes

$$\underbrace{\frac{\partial}{\partial y}\Big[E(y)\Big(\frac{\partial u_1}{\partial x} + \frac{\partial u_2}{\partial y}\Big)\Big]}_{function\ of\ y} + \underbrace{\frac{1}{\Big(\displaystyle\frac{1}{\ell}\int_{y_0}^{y_0+\ell} \frac{d\tilde{y}}{E(\tilde{y})}\Big)}\frac{\partial^2 u_0}{\partial x^2}}_{function\ of\ x} = 0 \tag{10.44}$$

Hence,

$$\frac{1}{\Big(\displaystyle\frac{1}{\ell}\int_{y_0}^{y_0+\ell} \frac{d\tilde{y}}{E(\tilde{y})}\Big)}\frac{\partial^2 u_0}{\partial x^2} = 0$$

or

$$\frac{\partial}{\partial x}\left\{\frac{1}{\Big(\displaystyle\frac{1}{\ell}\int_{y_0}^{y_0+\ell} \frac{d\tilde{y}}{E(\tilde{y})}\Big)}\frac{\partial u_0}{\partial x}\right\} = 0\,. \tag{10.45}$$

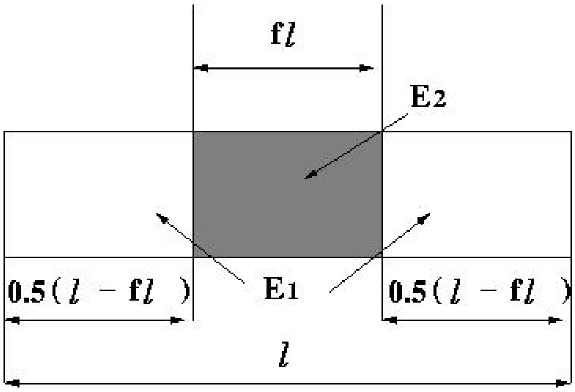


Figure 10.5. One-dimensional unit cell

This is the homoegenized differential equation that governs the macroscale variation of the mean displacement field.

Compare the mean-field differential equation to the original differential equation,

$$\frac{d}{dy}\Big(E(y)\frac{du}{dy}\Big) = 0$$

We conclude that the effective coefficient for the homogenized differential equation is

$$E_e = \frac{1}{\Big(\dfrac{1}{\ell}\displaystyle\int_{y_0}^{y_0+\ell}\frac{d\tilde{y}}{E(\tilde{y})}\Big)} = \Big\langle \frac{1}{E}\Big\rangle^{-1} \tag{10.46}$$

which is the harmonic mean of $E(y)$, or the estimate from Reuss bound.

Consider the unit cell shown in Fig. (10.5). One may find that

$$\begin{aligned}\frac{1}{\ell}\int_{y_0}^{y_0+\ell}\frac{dt}{E(t)} &= \frac{2}{\ell}\int_0^{\frac{1-f\ell}{2}}\frac{dt}{E_1} + \frac{1}{\ell}\int_0^{f\ell}\frac{dt}{E_2}\\ &= \frac{\ell - f\ell}{\ell}\frac{1}{E_1} + \frac{f}{E_2} = \frac{(1-f)E_2 + fE_1}{E_1E_2}\end{aligned}$$

and

$$E_e = \frac{1}{\dfrac{1}{\ell}\displaystyle\int_{y_0}^{y_0+\ell}\frac{dt}{E(t)}} = \frac{E_1E_2}{(1-f)E_2 + fE_1} \tag{10.47}$$

The homogenized differential equation is,

$$\frac{d}{dx}\Big(E_e\frac{du_0}{dx}\Big) = 0\ . \tag{10.48}$$

To sum up, asymptotic homogenization consists of the following steps:

Summary of Asymptotic Homogenization

1 The objective of the homogenization is to find the average coefficients of the homogenized differential equation and find its solution;

2 Identify the micro- and macroscales;

3 Introduce multiple-scale variables and expansions, and deduce cell boundary-value problems (BVPs) at successive orders. The leading-order cell problem is homogeneous, i.e. $u_0 = u_0(\mathbf{x})$;

4 Use linearity (or separation of variables) to express the next-order solution in terms of the leading-order solution and deduce an inhomogeneous canonical cell BVP;

5 Require the solvability of the inhomogeneous cell problem;

6 Find the differential equation that governs the macro-scale variation of the mean displacement or the evolution of the leading-order solution which includes the constitutive coeffocoients of the differential equation.

10.3.2 A multiple dimension example

Consider a 3D example,

$$A^\epsilon u = f, \quad \forall \mathbf{x} \in \Omega \tag{10.49}$$

$$u_\epsilon = 0, \quad \forall \mathbf{x} \in \partial\Omega \tag{10.50}$$

where

$$A^\epsilon = -\frac{\partial}{\partial x_i}\Big(a_{ij}(\frac{x}{\epsilon})\frac{\partial}{\partial x_j}\Big)$$

where $x = (x_1, x_2, x_3)$.

Define the fast coordinate,

$$y = \frac{x}{\epsilon}$$

as if y is speed-up by the large paramter $\dfrac{1}{\epsilon}$. We then can express the field variable as a function of two independent scales, $u_\epsilon(x) = u(x, y)$.

From chain rule, we have

$$\frac{\partial}{\partial x_i} = \frac{\partial}{\partial x_i} + \frac{\partial}{\partial y_i}\frac{\partial y_i}{\partial x_i} = \frac{\partial}{\partial x_i} + \frac{1}{\epsilon}\frac{\partial}{\partial y_i}$$

We can then expand the differential operator, A^ϵ, as

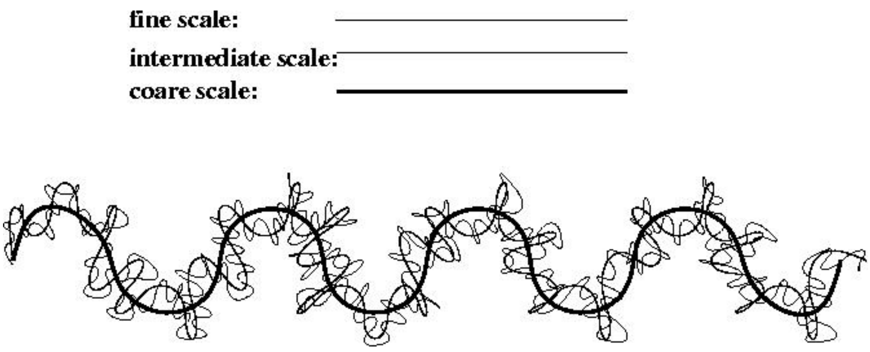


Figure 10.6. Illustration of multiscale phenomena

$$
\begin{aligned}
A^{\epsilon} &= -\Big(\frac{\partial}{\partial x_i} + \frac{1}{\epsilon}\frac{\partial}{\partial y_i}\Big)\Big[a_{ij}(y)\Big(\frac{\partial}{\partial x_i} + \frac{1}{\epsilon}\frac{\partial}{\partial y_i}\Big)\Big] \\
&= -\epsilon^{-2}\Big[\frac{\partial}{\partial x_i}\Big(a_{ij}\frac{\partial}{\partial y_j}\Big)\Big] - \epsilon^{-1}\Big[\frac{\partial}{\partial x_i}a_{ij}(y)\frac{\partial}{\partial y_j} + \frac{\partial}{\partial y_i}a_{ij}(y)\frac{\partial}{\partial x_i}\Big] \\
&\quad -\epsilon^{0}\Big[\frac{\partial}{\partial x_i}a_{ij}(y)\frac{\partial}{\partial x_i}\Big] \\
&= \epsilon^{-2}A_1 + \epsilon^{-1}A_2 + \epsilon^{0}A_3
\end{aligned}
\tag{10.51}
$$

where

$$
\begin{aligned}
A_1 &= -\Big[\frac{\partial}{\partial x_i}\Big(a_{ij}\frac{\partial}{\partial y_j}\Big)\Big] \\
A_2 &= -\Big[\frac{\partial}{\partial x_i}a_{ij}(y)\frac{\partial}{\partial y_j} + \frac{\partial}{\partial y_i}a_{ij}(y)\frac{\partial}{\partial x_i}\Big] \\
A_3 &= -\Big[\frac{\partial}{\partial x_i}a_{ij}(y)\frac{\partial}{\partial x_i}\Big]
\end{aligned}
$$

Now we consider multiple scale expansion,

$$
u_{\epsilon}(x) = u_0(x,y) + \epsilon u_1(x,y) + \epsilon^2 u_2(x,y) + \cdots \tag{10.52}
$$

which decomposes or separates the activities at different scales.
Substituting both (10.52) and (10.51) into (10.49), we have

$$
\begin{aligned}
&\Big(\epsilon^{-2}A_1 + \epsilon^{-1}A_2 + \epsilon^{0}A_3\Big)\Big(u_0 + \epsilon u_1 + \epsilon^2 u_2 + \cdots\Big) = f \\
&\epsilon^{-2}A_1 u_0 + \epsilon^{-1}\Big(A_1 u_1 + A_2 u_0\Big) + \epsilon^{0}(A_1 u_2 + A_2 u_1 + A_3 u_0) \\
&+ \cdots = f
\end{aligned}
\tag{10.53}
$$

The total state equilibrium is equivalent to equilibrium states in each every scale. That is

$$\begin{aligned}
\epsilon^{-2}: &\quad A_1 u_0 = 0; && (10.54)\\
\epsilon^{-1}: &\quad A_1 u_1 + A_2 u_0 = 0; && (10.55)\\
\epsilon^{0}: &\quad A_1 u_2 + A_2 u_1 + A_3 u_0 = f && (10.56)\\
&\quad \cdots\cdots
\end{aligned}$$

If one can solve differential equations at each scale, one can find out both local detailed information as well as global information.

As far as homogenization concern, we are looking for a homogenized differential equation that carries the overall information of fine scale.

Before we proceed further, we prove the following lemma.

LEMMA 10.3 *If the differential equation,*

$$A_1 u = F, \quad \forall \mathbf{y} \in Y$$

has a unique Y-periodic solution, the following equation holds

$$< F >= \frac{1}{|Y|} \int_Y F(y) dV_y = 0 \qquad (10.57)$$

where $y = (y_1, y_2, y_3)$.

Proof:

By the assumption, one can assume that both u and F are Y-periodic, and

$$\begin{aligned}
F(y) &= \sum_{\xi \in \Lambda} \mathcal{F}[F](\xi) \exp(i\xi y) && (10.58)\\
u(x,y) &= \sum_{\xi \in \Lambda} \mathcal{F}[u](\xi) \exp(i\xi y) && (10.59)
\end{aligned}$$

Hence

$$\begin{aligned}
A_1 u &= -\Big(\frac{\partial}{\partial y_i} a_{ij}(y) \frac{\partial}{\partial y_j}\Big) u\\
&= -\sum_{\xi \in \Lambda} \xi_j \Big(\frac{\partial a_{ij}}{\partial y_i} + i\xi_i\Big) \mathcal{F}[u](\xi) \exp(i\xi y)
\end{aligned}$$

Based on $A_1 u = F$, one has

$$\begin{aligned}
&-\sum_{\xi \in \Lambda} \mathrm{i}\xi_j \Big(\frac{\partial a_{ij}}{\partial y_i} + i\xi_i\Big) \mathcal{F}[u] \exp(i\xi y) = \sum_{\xi \in \Lambda} \mathcal{F}[F](\xi) \exp(i\xi y)\\
\Rightarrow \quad & \mathcal{F}[F](\xi) = -\mathrm{i}\xi_j \Big(\frac{\partial a_{ij}}{\partial y_i} + i\xi_i\Big)
\end{aligned}$$

Therefore,

$$\mathcal{F}[F] = 0 \;\;\Rightarrow\;\; \mathcal{F}[F](0) = \frac{1}{|Y|}\int_Y F dV_y = 0 \; .$$

♣

To this end, we start to solve differential equations at each scale. At scale ϵ^{-2}, we have

$$A_1 u_0 = -\frac{\partial}{\partial y_i}\Big(a_{ij}(y)\frac{\partial}{\partial y_j}\Big)u_0 = 0$$

We claim that

$$u_0 = u_0(x) \; .$$

That is the leading-order expansion is only the function of slow scale variable.

Since u_0 is Y-peridoic, we have

$$u_0 = \sum_{\xi\in\Lambda}\mathcal{F}[u_0](\xi)\exp(\mathrm{i}\xi y) \; .$$

Consequently,

$$A_1 u_0 = 0 \;\Rightarrow\; -\sum_{\xi\in\Lambda}\mathrm{i}\xi_j\Big(\frac{\partial a_{ij}}{\partial y_i} + \mathrm{i}\xi_i\Big)\mathcal{F}[u](\xi)\exp(\mathrm{i}\xi y) = 0 \; .$$

Then for $\xi \neq 0$, it is necessary

$$\mathcal{F}[u](\xi) = 0 \; . \tag{10.60}$$

Assume that

$$u_0 = c(x)Q(y) + \bar{u}_0(x)$$

Eq. (10.60) becomes

$$\begin{aligned}\mathcal{F}[u](\xi) &= \frac{1}{|Y|}\int_Y \Big(c(x)Q(y) + \bar{u}_0(x)\Big)\exp(-\mathrm{i}\xi y)dV_y \\ &= \frac{1}{|Y|}\int_Y \Big(c(x)Q(y)\Big)\exp(-\mathrm{i}\xi y)dV_y = 0 \end{aligned} \tag{10.61}$$

because $\displaystyle\int_Y \bar{u}_0(x)\exp(-\mathrm{i}\xi y)dV_y = 0$ when $\xi \neq 0$.

The only possibility that (10.61) holds is that $Q(y) = 1$ or $Q(y) = 0$. In either case, $u_0 = u_0(x)$. We proved our claim.

Next, we consider the differential equation at scale ϵ^{-1}:

$$A_1 u_1 + A_2 u_0 = 0 \; .$$

One can show that

$$A_2 u_0 = -\Big[\frac{\partial}{\partial x_i}\Big(a_{ij}(y)\Big)\frac{\partial}{\partial y_j} + \frac{\partial}{\partial y_i}\Big(a_{ij}(y)\Big)\frac{\partial}{\partial x_j}\Big]u_0(x) = -\frac{\partial a_{ij}}{\partial y_i}\frac{\partial u_0}{\partial x_j}$$

Hence

$$A_1 u_1 = \frac{\partial a_{ij}}{\partial y_i}\frac{\partial u_0}{\partial x_j} \tag{10.62}$$

This suggets the following separation of variable,

$$u_1(x,y) = U_k(y)\frac{\partial u_0}{\partial x_k} + \bar{u}_1(x) \tag{10.63}$$

and subsequently,

$$\begin{aligned} A_1 u_1 &= \Big(A_1 U_k(y)\Big)\frac{\partial u_0}{\partial x_k} \\ &= -\frac{\partial}{\partial y_i}\Big(a_{ij}(y)\Big)\frac{\partial U_k}{\partial y_j}\frac{\partial u_0}{\partial x_k} \end{aligned} \tag{10.64}$$

Combining (10.62) and (10.64), we find the canonical equation for a unit cell problem,

$$\frac{\partial a_{ik}}{\partial x_i} + \frac{\partial}{\partial y_i}\Big(a_{ij}(y)\Big)\frac{\partial U_k}{\partial y_j} = 0\ . \tag{10.65}$$

with the possible boundary conditions at interface of different phases,

$$\Big[U_k\Big] = 0, \quad \text{and} \quad \Big[\Big(a_{ik} + a_{ij}\frac{\partial U_k}{\partial x_j}\Big)n_i\Big] = 0 \tag{10.66}$$

We now consider the differential equation at ϵ^0 scale,

$$A_1 u_2 + A_2 u_1 + A_3 u_0 = f$$

which can be rewritten as

$$A_1 u_2 = f - \Big(A_2 u_1 + A_3 u_0\Big) \tag{10.67}$$

The condition that equation (10.67) has a unique periodic solution is that

$$< f - (A_2 u_1 + A_3 u_0) >= 0$$

That is

$$\frac{1}{|Y|}\int_Y \Big(A_2 u_1 + A_3 u_0\Big)dy = f \tag{10.68}$$

Consider

$$\begin{aligned} u_0 &= u_0(x) \\ u_1 &= U_j \frac{\partial u_0}{y_j} + \bar{u}_1(x) \end{aligned}$$

One can show that

$$\begin{aligned} A_3 u_0 &= -a_{ij} \frac{\partial^2 u_0}{\partial x_i \partial x_j} && (10.69) \\ A_2 u_1 &= -\Big[\frac{\partial}{\partial x_i} \Big(a_{ij}(y) \frac{\partial}{\partial y_j} + \frac{\partial}{\partial y_i} \Big(a_{ij}(y) \frac{\partial}{\partial x_j} \Big] \Big(U_k(y) \frac{\partial u_0}{\partial x_k} + \bar{u}_1 \Big) \\ &= -a_{ij} \frac{\partial U_k}{\partial y_j} \frac{\partial^2 u_0}{\partial x_i \partial x_k} - \frac{\partial}{\partial y_i} \Big(a_{ij}(y) U_k(y) \Big) \frac{\partial^2 u_0}{\partial x_j \partial x_k} \\ &\quad - \frac{\partial}{\partial y_i} \Big(a_{ij}(y) \Big) \frac{\partial \bar{u}_1}{\partial x_j} && (10.70) \end{aligned}$$

Change the dummy indices $j \leftrightarrow k$ in the first term of (10.70). We can write that

$$\begin{aligned} A_2 u_1 + A_3 u_0 &= -\Big(a_{ij} + a_{ik} \frac{\partial U_j}{\partial x_k} \Big) \frac{\partial^2 u_0}{\partial x_i \partial x_j} - \frac{\partial}{\partial y_i} \Big(a_{ij}(y) U_k(y) \frac{\partial^2 u_0}{\partial x_j \partial x_k} \Big) \\ &\quad - \frac{\partial}{\partial y_i} \Big(a_{ij}(y) \frac{\partial \bar{u}_1}{\partial x_j} \Big) \end{aligned}$$

Via divergence theorem,

$$\begin{aligned} \frac{1}{|Y|} \int_Y (A_2 u_1 + A_3 u_0) dy &= -\frac{1}{|Y|} \int_Y \Big(a_{ij} + a_{ik} \frac{\partial U_j}{\partial x_k} \Big) dV_y \frac{\partial^2 u_0}{\partial x_i \partial x_j} \\ &\quad - \Big[(a_{ij}(y) U_k(y) u_{0,jk}(x) \Big] n_i - \Big[a_{ij}(y) \bar{u}_{1,j} \Big] n_i \end{aligned}$$

By periodicity, the boundary terms will vanish. We then have

$$-\frac{1}{|Y|} \int_Y \Big(a_{ij} + a_{ik} \frac{\partial U_j}{\partial x_k} \Big) dV_y \frac{\partial^2 u_0}{\partial x_i \partial x_j} = f$$

Denote the effective coefficients as

$$\bar{a}_{ij} = \frac{1}{|Y|} \int_Y \Big(a_{ij} + a_{ik} \frac{\partial U_j}{\partial x_k} \Big) dV_y \tag{10.71}$$

and homogenized differential operator

$$A^H = -\frac{\partial}{\partial x_i} \Big(\bar{a}_{ij} \frac{\partial}{\partial x_j} \Big) \tag{10.72}$$

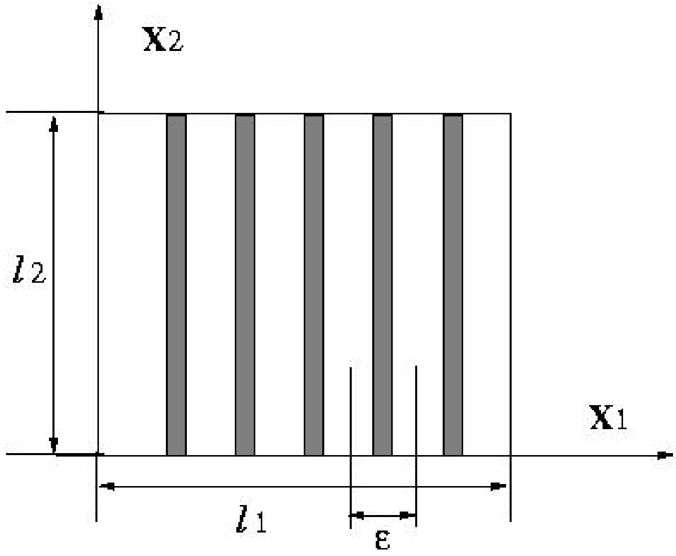


Figure 10.7. An unilateral composite with periodic straucture.

we finally derived the homoegenized boundary-value problem,

$$A^H u_0 = 0, \ \forall x \in \Omega \tag{10.73}$$

$$u_0 = 0, \ \forall x \in \partial\Omega \tag{10.74}$$

EXAMPLE 10.4 *Consider a 2D steady-state heat transfer problem (see Fig. (10.7)),*

$$\frac{\partial}{\partial x_\alpha}\Big(\lambda_{\alpha\beta}\Big(\frac{x_1}{\epsilon}\Big)\frac{\partial T^\epsilon}{\partial x_\beta}\Big) = 0, \ \forall x \in D \tag{10.75}$$

where $T^\epsilon(x)$ is temperature field and $\lambda_{\alpha\beta}$ are heat conduction coefficients. We assume that the region $D = \Big\{(x_1, x_2)\ \Big|\ 0 \le x_1 \le \ell_1, \text{ and } 0 \le x_2 \le \ell_2\Big\}$ is thermally insulated in horizontal boundaries, i.e.

$$q_2 = \lambda_{2\beta}\frac{\partial T^\epsilon}{\partial x_\beta} = 0, \ \forall x_2 = 0, \text{ and } x_2 = \ell_2 \tag{10.76}$$

Along the vertical boundaries of the region D, the heat flows are prescribed,

$$q_1 = \lambda_{1\beta}\frac{\partial T^\epsilon}{\partial x_\beta} = \mp q_0, \ \forall x_1 = 0, \text{ and } x_1 = \ell_1 \tag{10.77}$$

Consider multiple expansion,

$$T^\epsilon(x) = T_0(x) + \epsilon T_1(x, y_1) + \cdots$$

and the following separation of variable,

$$T_1(x, y_1) = U_\alpha(y_1)\frac{\partial T_0(x)}{\partial x_\alpha}, \quad \alpha = 1, 2$$

Note that first we assume that the mean temperature at this scale is zero, i.e. $\bar{T}_1(x) = 0$*; and* $U_\alpha(y_1)$ *are Y-periodic functions that are the following 1D canonical cell problem,*

$$-\frac{d}{dy_1}\Big(\lambda_{11}(y_1)\frac{dU_\alpha(y_1)}{dy_1}\Big) = \frac{d\lambda_{1\alpha}}{dy_1}, \ \forall\, y_1 \in Y \tag{10.78}$$

$$\Big[U_\alpha\Big] = 0, \ \text{ and } \Big[\lambda_{11}\frac{dU_\alpha}{dy_1}\Big] = 0, \ \forall y_1 \text{ at interface.} \tag{10.79}$$

Integrate (10.78),

$$\begin{aligned} -\lambda_{11}(y_1)\frac{dU_\alpha(y_1)}{dy_1} &= \lambda_{1\alpha}(y_1) - C_\alpha \\ \Rightarrow \ \frac{dU_\alpha(y_1)}{dy_1} &= -\frac{\lambda_{1\alpha}(y_1)}{\lambda_{11}(y_1)} + \frac{C_\alpha}{\lambda_{11}(y_1)} \end{aligned}$$

where C_α *are constants (note that they are not functions of x !).*

Integrate second time,

$$U_\alpha(y_1) = -\int_0^{y_1} \lambda_{1\alpha}(\xi)\lambda_{11}^{-1}(\xi)d\xi + C_\alpha\int_0^{y_1}\lambda_{11}^{-1}(\xi)d\xi + D_\alpha$$

Note that we choose $D_\alpha = 0$*, because the average temperature at scale* ϵ^{-1} *is assumed to be zero.*

The solvability condition of the canonical cell problem requires $U_\alpha(y_1)$ *as a Y-periodic function, i.e.*

$$U_\alpha(0) = U_\alpha(\ell)$$

This condition allows us to determine the constants C_α*,*

$$C_\alpha = \frac{\displaystyle\int_0^\ell \lambda_{1\alpha}(\xi)\lambda_{11}^{-1}(\xi)d\xi}{\int_0^\ell \lambda_{11}^{-1}(\xi)d\xi} \tag{10.80}$$

In specific,

$$\begin{aligned} C_1 &= \Big(\int_0^\ell \lambda_{11}^{-1}(\xi)d\xi\Big)^{-1} \\ C_2 &= \frac{\displaystyle\int_0^\ell \lambda_{12}(\xi)\lambda_{11}^{-1}(\xi)d\xi}{\displaystyle\int_0^\ell \lambda_{11}^{-1}(\xi)d\xi} \end{aligned}$$

Consequently, we find the closed form solution for canonical cell problem,

$$U_1(y_1) = -y_1 + \frac{\displaystyle\int_0^{y_1} \lambda_{11}^{-1}(\xi)d\xi}{\displaystyle\int_0^\ell \lambda_{11}^{-1}(\xi)d\xi} \tag{10.81}$$

$$\begin{aligned} U_2(y_1) &= -\int_0^{y_1} \lambda_{12}(\xi)\lambda_{11}^{-1}(\xi)d\xi \\ &\quad + \frac{\displaystyle\int_0^\ell \lambda_{12}(\xi)\lambda_{11}^{-1}(\xi)d\xi}{\displaystyle\int_0^\ell \lambda_{11}^{-1}(\xi)d\xi}\Big(\int_0^{y_1} \lambda_{11}^{-1}(\xi)d\xi\Big) \end{aligned} \tag{10.82}$$

Define the effective heat conduction coefficients,

$$\bar{\lambda}_{ij} := \frac{1}{|Y|}\int_Y \Big(a_{ij} + a_{ik}\frac{\partial U_j}{\partial x_k}\Big)dy\ .$$

It is easy to find that

$$\begin{aligned} \bar{\lambda}_{11} &= \frac{1}{\ell}\int_0^\ell \Big(\lambda_{11}(\xi) + \lambda_{11}(\xi)\frac{\partial U_1}{\partial y_1}(\xi)\Big)d\xi \\ &= \frac{1}{\ell}\int_0^\ell \Big(\lambda_{11} - \lambda_{11} + C_1\Big)dy = \frac{1}{\ell}C_1 \\ &= (\frac{1}{\ell}\Big(\int_0^\ell \lambda_{11}^{-1}(\xi)d\xi\Big)^{-1} \end{aligned}$$

and

$$\begin{aligned}
\bar{\lambda}_{12} &= \frac{1}{\ell}\int_0^{\ell}\Big(\lambda_{12}(\xi)+\lambda_{11}(\xi)\frac{\partial U_2}{\partial y_1}(\xi)\Big)d\xi \\
&= \frac{1}{\ell}\int_0^{\ell}\Big(\lambda_{12}(\xi)-\lambda_{12}(\xi)+C_2\Big)d\xi \\
&= \frac{\dfrac{1}{\ell}\displaystyle\int_0^{\ell}\lambda_{12}(\xi)\lambda_{11}^{-1}(\xi)d\xi}{\dfrac{1}{\ell}\displaystyle\int_0^{\ell}\lambda_{11}^{-1}(\xi)d\xi} = \bar{\lambda}_{21}
\end{aligned}$$

and

$$\begin{aligned}
\bar{\lambda}_{22} &= \frac{1}{\ell}\int_0^{\ell}\Big(\lambda_{22}(\xi)+\lambda_{21}(\xi)\frac{\partial U_2}{\partial y_1}(\xi)\Big)d\xi &\quad (10.83)\\
&= \frac{1}{\ell}\int_0^{\ell}\Big(\lambda_{22}(\xi)-\lambda_{12}^2\lambda_{11}^{-1}(\xi)+C_2\lambda_{12}\lambda_{11}^{-1}(\xi)\Big)d\xi &\quad (10.84)\\
&= \frac{1}{\ell}\int_0^{\ell}\lambda_{22}(\xi)d\xi-\frac{1}{\ell}\int_0^{\ell}\lambda_{12}(\xi)\lambda_{11}^{-1}(\xi)d\xi \\
&\quad +\frac{1}{\ell}\frac{\Big(\displaystyle\int_0^{\ell}\lambda_{12}(\xi)\lambda_{11}^{-1}(\xi)d\xi\Big)^2}{\displaystyle\int_0^{\ell}\lambda_{11}^{-1}(\xi)d\xi} &\quad (10.85)
\end{aligned}$$

and the homogenized partial differential equation becomes

$$\bar{\lambda}_{11}\frac{\partial^2 T_0}{\partial x_1^2}+2\bar{\lambda}_{12}\frac{\partial^2 T_0}{\partial x_1\partial x_2}+\bar{\lambda}_{22}\frac{\partial^2 T_0}{\partial x_2^2}=0\ .$$

10.4 Variational Characterization

Recall the homogenization of conduction problem,

$$\begin{aligned}
A^{\epsilon}u_{\epsilon} &= f,\ \ \forall x\in\Omega \\
u_{\epsilon} &= 0,\ \ \forall x\in\partial\Omega
\end{aligned}$$

Assume that

$$u^{(1)}(x,y)=U_k(y)\frac{\partial u^{(0)}(x)}{\partial x_k} \tag{10.86}$$

One can derive the following governing equations for the canonical cell problem,

$$\frac{\partial}{\partial y_k}\Big(a_{kj}+a_{k\ell}\frac{\partial U_j}{\partial y_\ell}\Big)=0,\ \ \forall y\in Y \tag{10.87}$$

with the proper interface and periodic conditions.

Subsequently, one can derive the effective coefficients for homogenized differential equation,

$$\bar{a}_{ij} = \frac{1}{|Y|}\int_Y \Big(a_{ij} + a_{i\ell}\frac{\partial U_j}{\partial y_\ell}\Big)dy = \frac{1}{|Y|}\int_Y a_{i\ell}\Big(\delta_{\ell j} + \frac{\partial U_j}{\partial y_\ell}\Big)dy \tag{10.88}$$

Based on (10.87), one may find that

$$-\frac{1}{|Y|}\int_Y \Big(a_{kj}(y) + a_{k\ell}(y)\frac{\partial U_j}{\partial y_\ell}\Big)U_i(y)dy = 0$$

Integration by parts yields

$$\begin{aligned} &-\frac{1}{|Y|}\int_{\partial Y}\Big(a_{kj} + a_{k\ell}\frac{\partial U_j}{\partial y_\ell}\Big)U_j(y)n_k dS + \frac{1}{|Y|}\int_Y \Big(a_{kj} + a_{k\ell}\frac{\partial U_j}{\partial y_\ell}\Big)\frac{\partial U_i}{\partial y_k}dy \\ = \ &\frac{1}{|Y|}\int_Y a_{k\ell}\Big(\delta_{\ell j} + \frac{\partial U_j}{\partial y_\ell}\Big)\frac{\partial U_i}{\partial y_k}dy = 0 \end{aligned} \tag{10.89}$$

Adding (10.89) to (10.88), one may find that

$$\begin{aligned} \bar{a}_{ij} &= \frac{1}{|Y|}\int_Y \Big(\delta_{\ell j} + \frac{\partial U_j}{\partial y_\ell}\Big)\Big(a_{i\ell}(y) + a_{k\ell}(y)\frac{\partial U_i}{\partial y_k}\Big)dy \\ &= \frac{1}{|Y|}\int_Y a_{k\ell}(y)\Big(\delta_{ik} + \frac{\partial U_i}{\partial y_k}\Big)\Big(\delta_{\ell j} + \frac{\partial U_j}{\partial y_\ell}\Big)dy \end{aligned} \tag{10.90}$$

Eq. (10.90) links the effective coefficients of the homogenized equation with the variational characters of unit cell problem, which plays a significant role in Tartar's variational principle.

Consider constant vector, $\boldsymbol{\xi} = \xi_i \mathbf{e}_i$, or a flux vector of macro-scale variable. We can form the following quadratic form,

$$\begin{aligned} \bar{a}_{ij}\xi_i\xi_j &= \frac{1}{|Y|}\int_Y \xi_i\xi_j a_{k\ell}(y)\Big(\delta_{ik} + \frac{\partial U_i}{\partial y_k}\Big)\Big(\delta_{\ell j} + \frac{\partial U_j}{\partial y_\ell}\Big)dy \\ &= \frac{1}{|Y|}\int_Y a_{k\ell}(y)\Big(\xi_k + \frac{\partial U_i\xi_i}{\partial y_k}\Big)\Big(\xi_\ell + \frac{\partial U_j\xi_j}{\partial y_\ell}\Big)dy \end{aligned} \tag{10.91}$$

Eq. (10.91) suggests that there exists a functional,

$$J(\mathbf{U}) = \frac{1}{Y}\int_Y a_{ij}(y)\Big(\xi_i + \frac{\partial U_k\xi_k}{\partial y_i}\Big)\Big(\xi_j + \frac{\partial U_\ell\xi_\ell}{\partial y_j}\Big)dy \tag{10.92}$$

such that

$$\bar{a}_{ij}\xi_i\xi_j = \min_{\mathbf{U}\in H^1_{\#}(Y)} J\Big(\mathbf{U}\Big) \tag{10.93}$$

where the function space $H^1_{\#}(Y)$ [1] is defined as $H^1(Y)$ space of Y-periodic functions, i.e.

$$H^1_{\#}(Y) := \Big\{ u \;\Big|\; u \,\text{is Y} - \text{periodic}, \;\text{and}\; u \in H^1(Y) \Big\}$$

that is

$$\int_Y (u^2 + |\nabla u|^2) dy < +\infty$$

To show this, we first show that the Euler-Lagrange equation of $J(\mathbf{U})$ is the governing equation of canonical cell problem.

Assume that a_{ij} is symmetric and real. It subsequently implies that a_{ij} is positive definite. Therefore,

$$\begin{aligned} \delta J &= \frac{1}{|Y|} \int_Y a_{ij}(y) \Big(\frac{\partial \delta U_k \xi_k}{\partial y_i} \Big(\xi_j + \frac{\partial U_\ell \xi_\ell}{\partial y_j} \Big) + \Big(\xi_i + \frac{\partial U_k \xi_k}{\partial y_i} \Big) \frac{\partial \delta U_\ell \xi_\ell}{\partial y_j} \Big) dy \\ &= \frac{2}{|Y|} \int_{\partial Y} a_{ij}(y) \Big(\xi_i + \frac{\partial U_k \xi_k}{\partial y_i} \Big) \delta U_\ell \xi_\ell dS \\ & \quad - \frac{2}{|Y|} \int_Y \frac{\partial}{\partial y_i} \Big(a_{ij}(y) \Big(\xi_i + \frac{\partial U_k \xi_k}{\partial y_i} \Big) \Big) \delta U_\ell \xi_\ell dy = 0 \end{aligned}$$

By periodic conditions

$$\frac{2}{|Y|} \int_{\partial Y} a_{ij}(y) \Big(\xi_i + \frac{\partial U_k \xi_k}{\partial y_i} \Big) \delta U_\ell \xi_\ell dS = 0,$$

it then leads to

$$\delta J = - \frac{2}{|Y|} \int_Y \frac{\partial}{\partial y_i} \Big(a_{ij}(y) \Big(\delta_{ik} + \frac{\partial U_k}{\partial y_i} \Big) \Big) \delta U_\ell \xi_k \xi_\ell dy = 0$$

and hence

$$- \frac{\partial}{\partial y_i} \Big(a_{ij}(y) \Big(\delta_{ik} + \frac{\partial U_k}{\partial y_i} \Big) \Big) \delta U_\ell = 0 \,.$$

Consider $U_k = 0 \in H^1_{\#}(Y)$. One can find an upper bound for effective coefficient, $\bar{a}_{ij}$, i.e.

$$0 < \bar{a}_{ij} \xi_i \xi_j \leq \Big(\frac{1}{|Y|} \int_Y a_{ij}(y) dy \Big) \xi_i \xi_j \tag{10.94}$$

or

$$\bar{a}_{ij} \leq \frac{1}{|Y|} \int_Y a_{ij}(y) dy \tag{10.95}$$

[1] In music, the sign $\#$ is used to indicate that a note is to be raised by a half tone. Similar meaning implies here as well, i.e. a "half level higher" H^1 space.

This is the arithmetic mean or the so-called Voigt bound.

To find the lower bound, we have to enlarge the space $H^1_{\#}(Y)$. Consider function $\zeta_i \in L^2_{\#}(Y)$ and the mean value of ζ_i is zero, i.e.

$$\int_Y \zeta_i(y)dy = 0 \ .$$

It is obvious that

$$\bar{a}_{ij}\xi_i\xi_j \geq \min_{\substack{\boldsymbol{\zeta}\in L^2_{\#}(Y) \text{ and} \\ \int_Y \boldsymbol{\zeta}(y)dy=0}} J_c(\boldsymbol{\zeta}) \tag{10.96}$$

where

$$\begin{aligned} J_c(\boldsymbol{\zeta}) \quad &:= \quad \frac{1}{|Y|}\int_Y a_{ij}(\xi_i + \zeta_i(y))(\xi_j + \zeta_j(y))dy \\ &\quad -2C_k\Big(\int_Y \zeta_k(y)dy - 0\Big) \end{aligned} \tag{10.97}$$

where C_k are Lagrange multipliers.

To find the minimizer in $L^2_{\#}(Y)$, we calculate the first variation of the functional, $J_c(\boldsymbol{\zeta})$,

$$\begin{aligned} \delta J_c \quad &= \quad \frac{2}{|Y|}\int_Y a_{ij}(y)(\xi_i + \zeta_i)\delta\zeta_j dy - 2\delta C_j \frac{1}{|Y|}\int_Y \zeta_j(y)dy \\ &\quad -2C_j\frac{1}{|Y|}\int_Y \delta\zeta_j(y)dy \\ &= \quad \frac{2}{|Y|}\int_Y \Big(a_{ij}(y)(\xi_i+\zeta_i) - C_j\Big)\delta\zeta_j dy - 2\delta C_j\frac{1}{|Y|}\int_Y \zeta_j(y)dy = 0 \end{aligned}$$

which yields Euler-Lagrangian equation and the constrain condition,

$$a_{ij}(\xi_j + \zeta_j) \quad = \quad C_i \tag{10.98}$$

$$\int_Y \zeta_j(y)dy \quad = \quad 0\ . \tag{10.99}$$

Solving (10.98), we have

$$\xi_i + \zeta_i = a^{-1}_{ij}C_j \tag{10.100}$$

Average the above expression over the unit cell and considering the constraint condition (10.99),

$$\xi_i =< a^{-1}_{ij}(y) > C_j \tag{10.101}$$

which solves C_j in terms of ξ_i, i.e.

$$C_j =< a^{-1}_{ji}(y) >^{-1} \xi_i \tag{10.102}$$

The minimizer in $L^2_{\#}(Y)$ under the constraint is then

$$\begin{aligned}
\min_{\substack{\boldsymbol{\zeta}\in L^2_{\#}(Y) \text{ and} \\ \int_Y \boldsymbol{\zeta}(y)dy=0}} J_c(\boldsymbol{\zeta}) &= \frac{1}{|Y|}\int_Y a_{ij}(\xi_i+\zeta_i)(\xi_j+\zeta_j)dy \\
&= \frac{1}{|Y|}\int_Y C_j(\xi_i+\zeta_i)dy = C_j\xi_j \\
&= < a_{ji}^{-1} >_Y^{-1} \xi_i\xi_j \\
&= \Big(\frac{1}{|Y|}\int_Y a_{ij}^{-1}(y)dy\Big)^{-1}\xi_i\xi_j
\end{aligned}$$

From the above estimate, we find a lower bound for effective coefficient, $\bar{a}_{ij}$, i.e.

$$\bar{a}_{ij} \geq \Big(\frac{1}{|Y|}\int_Y a_{ij}^{-1}(y)dy\Big)^{-1}. \tag{10.103}$$

which is the so-called Reuss bound.

10.5 Multiscale Finite Element Method

10.5.1 Asymptotic homogenization of linear elasticity

Consider a composite material with periodic structure and its elastic stiffness tensor satisfies the relation,

$$C_{ijk\ell}\Big(\frac{x}{\epsilon}\Big)\xi_{ij}\xi_{k\ell} = C_{ijk\ell}(y)\xi_{ij}\xi_{k\ell} \geq \alpha\xi_{ij}\xi_{ij}$$

where $\alpha > 0$.

Consider the following boundary value problem,

$$\begin{aligned}
\frac{\partial \sigma_{ij}^{\epsilon}}{\partial x_j} + f_i &= 0, \quad \forall x\in\Omega & (10.104)\\
\sigma_{ij}^{\epsilon} &= C_{ijk\ell}^{\epsilon}u_{k,\ell}^{\epsilon} = C_{ijk\ell}^{\epsilon}e_{k\ell}^{\epsilon} & (10.105)\\
e_{k\ell}^{\epsilon} &= \frac{1}{2}\Big(\frac{\partial u^{\epsilon}}{\partial x_\ell} + \frac{\partial u_\ell^{\epsilon}}{\partial x_k}\Big) & (10.106)\\
\sigma_{ij}^{\epsilon}n_j &= t_i^0, \quad \forall x\in\Gamma_t & (10.107)\\
u_i^{\epsilon} &= \bar{u}_i, \quad \forall x\in\Gamma_u & (10.108)
\end{aligned}$$

Consider multiple scale expansion,

$$u_i^{\epsilon}(x) = u_i^{(0)}(x,y) + \epsilon u_i^{(1)}(x,y) + \epsilon^2 u_i^{(2)}(x,y) + \cdots, \quad y := \frac{x}{\epsilon}$$

Hence

$$
\begin{aligned}
u^{\epsilon}_{k,\ell} &= \frac{\partial}{\partial x_\ell}u^{\epsilon}_k = \Big(\frac{\partial}{\partial x_\ell} + \frac{1}{\epsilon}\frac{\partial}{\partial y_\ell}\Big)\Big(u^0_k + \epsilon u^1_k + \epsilon^2 u^2_k + \cdots\Big) \\
&= \epsilon^{-1} e_{Yk\ell}(\mathbf{u}^{(0)}) + \epsilon^0 (e_{Xk\ell}(\mathbf{u}^{(0)} + e_{Yk\ell}(\mathbf{u}^{(1)})) + \\
&\quad + \epsilon^1 (e_{Xk\ell}(\mathbf{u}^{(1)}) + e_{Yk\ell}(\mathbf{u}^2)) + \cdots
\end{aligned} \tag{10.109}
$$

and

$$
\begin{aligned}
\sigma^{\epsilon}_{ij}(x,y) &= C_{ijk\ell}(y) u^{\epsilon}_{k,\ell} \\
&= C_{ijk\ell}(y)\Big[\epsilon^{-1} u^{(0)}_{Yk,\ell} + \epsilon^0 (u^0_{Xk,\ell} + u^1_{Yk,\ell}) + \epsilon(u^{(1)}_{Xk,\ell} + u^{(2)}_{Yk,\ell}) \\
&\quad + \cdots\Big] \\
&= \epsilon^{-1}\sigma^{(0)}_{ij} + \epsilon^0 \sigma^{(1)}_{ij} + \epsilon^1 \sigma^{(2)}_{ij} + \cdots
\end{aligned} \tag{10.110}
$$

In each scale, the constitutive relations are

$$
\begin{aligned}
\epsilon^{-1}: &\quad \sigma^{(0)}_{ij} = C_{ijk\ell}(y) u^{(0)}_{Yk,\ell}; \\
\epsilon^{0}: &\quad \sigma^{(1)}_{ij} = C_{ijk\ell}(y)(u^{(0)}_{Xk,\ell} + u^{(1)}_{Yk,\ell}); \\
\epsilon^{1}: &\quad \sigma^{(1)}_{ij} = C_{ijk\ell}(y)(u^{(1)}_{Xk,\ell} + u^{(2)}_{Yk,\ell}); \\
&\quad \cdots
\end{aligned}
$$

To derive equilibrium equation at different scales, one may write

$$
\begin{aligned}
\frac{\partial \sigma^{\epsilon}_{ij}}{\partial x_j} &= \Big(\frac{\partial}{\partial x_j} + \frac{1}{\epsilon}\frac{\partial}{\partial y_j}\Big)\sigma^{\epsilon}_{ij} + f_i = 0 \\
&= \Big(\frac{\partial}{\partial x_j} + \frac{1}{\epsilon}\frac{\partial}{\partial y_j}\Big)\Big(\epsilon^{-1}\sigma^{(0)}_{ij} + \epsilon^0\sigma^{(1)}_{ij} + \epsilon^1\sigma^{(2)}_{ij} + \cdots\Big) + f_i = 0
\end{aligned}
$$

Consequently,

$$
\epsilon^{-2}: \quad \frac{\partial \sigma^{(0)}_{ij}}{\partial y_j} = 0; \tag{10.111}
$$

$$
\epsilon^{-1}: \quad \frac{\partial \sigma^{(0)}_{ij}}{\partial x_j} + \frac{\partial \sigma^{(1)}_{ij}}{\partial y_j} = 0; \tag{10.112}
$$

$$
\epsilon^{0}: \quad \frac{\partial \sigma^{(1)}_{ij}}{\partial x_j} + \frac{\partial \sigma^{(2)}_{ij}}{\partial y_j} + f_i = 0; \tag{10.113}
$$

$$
\epsilon^{s-1}: \quad \frac{\partial \sigma^{(s)}_{ij}}{\partial x_j} + \frac{\partial \sigma^{(s+1)}_{ij}}{\partial y_j} = 0; \quad s = 2, 3, \cdots \tag{10.114}
$$

and the boundary conditions are

$$\begin{aligned}\Big(\epsilon^{-1}\sigma_{ij}^{(0)} + \epsilon^0\sigma_{ij}^1 + \epsilon^1\sigma_{ij}^2 + \cdots\Big)n_j &= t_i^0,\ \forall x \in \Gamma_t &\quad (10.115)\\ \Big(u_i^{(0)} + \epsilon^1 u_i^{(1)} + \epsilon^2 u_i^{(2)} + \cdots\Big) &= 0,\ \forall x \in \Gamma_u &\quad (10.116)\end{aligned}$$

The boundary conditions in different scale are

$$\left.\begin{array}{ll}\epsilon^{-1}: & \sigma_{ij}^{(0)}n_j = 0;\\ \epsilon^0: & \sigma_{ij}^{(1)}n_j = t_i^0;\\ \epsilon^1: & \sigma_{ij}^{(2)}n_j = 0;\\ & \cdots\cdots\end{array}\right. \quad \forall x \in \Gamma_t \qquad (10.117)$$

and

$$\left.\begin{array}{ll}\epsilon^0: & u_i^{(0)} = \bar{u}_i;\\ \epsilon^1: & u_i^{(1)} = 0;\\ \epsilon^2: & u_i^{(2)} = 0;\\ & \cdots\cdots\end{array}\right. \quad \forall x \in \Gamma_u \qquad (10.118)$$

We first examine the leading order equilibrium equation and boundary condition,

$$\frac{\partial \sigma_{ij}^{(0)}}{\partial y_j} = 0$$

This yields

$$\sigma_{ij}^{(0)} = \sigma_{ij}^{(0)}(x)$$

On the other hand

$$\sigma_{ij}^{(0)} = C_{ijk\ell}(y)\frac{\partial u_k^{(0)}}{\partial y_\ell}$$

To commodate both conditions, we have to set

$$\sigma_{ij}^{(0)} = 0\,. \qquad (10.119)$$

and

$$u_i^{(0)} = u_i^{(0)}(x) \qquad (10.120)$$

To solve the second order boundary-value problem, the follwing separation of variable is adopted

$$u_i^{(1)}(x,y) = \chi_i^{k\ell}(y)\frac{\partial u_k^{(0)}}{\partial x_\ell}(x) + \bar{u}_i^{(1)}(x) \qquad (10.121)$$

where the unknown vector function, $\chi_i^{k\ell}(y)\mathbf{e}_i$, is often referred to as the *characteristic displacement field*. We further assume that

$$\sigma_{ij}^{(1)}(x,y) = \hat{\sigma}_{ij}^{k\ell}(y)\frac{\partial u_k^0}{\partial x_\ell}(x) \tag{10.122}$$

Consider

$$\frac{\partial u_i^{(1)}}{\partial y_j} = \frac{\partial \chi_i^{k\ell}}{\partial y_j}\frac{\partial u_k^{(0)}}{\partial x_\ell} \tag{10.123}$$

and

$$\sigma_{ij}^{(1)} = C_{ijk\ell}\Big(u_{Xk,\ell}^{(0)} + u_{Yk,\ell}^{(1)}\Big) = C_{ijk\ell}\Big(e_{Xk\ell}^{(0)} + u_{Yk,\ell}^{(1)}\Big)\ . \tag{10.124}$$

We find that

$$\sigma_{ij}^{(1)} = C_{ijk\ell}\Big(T_{k\ell}^{mn} + \frac{\partial \chi_k^{mn}}{\partial y_\ell}\Big)u_{Xm,n}^{(0)} \tag{10.125}$$

where $T_{k\ell}^{mn} = \dfrac{1}{2}\Big(\delta_{km}\delta_{\ell n} + \delta_{kn}\delta_{\ell m}\Big)$, because $T_{k\ell}^{mn}u_{Xm,n}^{(0)} = e_{Xk\ell}^{(0)}$.

Accordingly,

$$\hat{\sigma}_{ij}^{mn} = C_{ijk\ell}\Big(T_{k\ell}^{mn} + \frac{\partial \chi_k^{mn}}{\partial y_\ell}\Big)$$

Then the equilibrium equation on second scale (ϵ^{-1}) provides the governing equation for the canonical cell problem,

$$\frac{\partial \sigma_{ij}^{(1)}}{\partial y_j} = 0, \quad \Rightarrow \quad \frac{\partial \hat{\sigma}_{ij}^{mn}}{\partial y_j}\frac{\partial u_m^{(0)}}{\partial x_n} = 0, \quad \Rightarrow \quad \frac{\partial \hat{\sigma}_{ij}^{mn}}{\partial y_j} = 0\ . \tag{10.126}$$

More explicitely, the governing equation for canonical cell problem is

$$\boxed{\frac{\partial}{\partial y_j}\Big(C_{ijk\ell}\Big[T_{k\ell}^{mn} + \frac{\partial \chi_k^{mn}}{\partial y_\ell}\Big]\Big) = 0, \ \ \forall y \in Y} \tag{10.127}$$

The related interface continuity conditions and periodic conditions are omited here.

Consider the equilibrium equation at third scale (ϵ^0). We have

$$\frac{\partial \sigma_{ij}^{(2)}}{\partial y_j} = -\Big(f_i + \frac{\partial \sigma_{ij}^{(1)}}{\partial x_j}\Big) = F_i, \ \ \forall\, y \in Y$$

The Fredholm alternative condition requires that

$$\frac{1}{|Y|}\int_Y F_i(y)dy = 0\ .$$

This can be shown from the fact that

$$\frac{1}{|Y|}\int_Y \frac{\partial \sigma_{ij}^{(2)}}{\partial y_j}dy = \frac{1}{|Y|}\int_{\partial Y}\sigma_{ij}^{(2)} n_j dS = 0 \ .$$

Thereby,

$$\frac{1}{|Y|}\int_Y \Big(f_i + \frac{\partial \sigma_{ij}^{(1)}}{\partial x_j}\Big)dy = 0, \quad \Rightarrow \quad f_i + \frac{\partial}{\partial x_j} < \sigma_{ij}^{(1)} >_Y = 0 \ .$$

where

$$< \sigma_{ij}^{(1)} >_Y = < \hat{\sigma}_{ij}^{k\ell}(y) >_Y \frac{\partial u_k^{(0)}}{\partial x_\ell} = C^h_{ijk\ell}\frac{\partial u_k^{(0)}}{\partial x_\ell} \qquad (10.128)$$

and the homegenized elastic stiffness tensor is determined by the solution of the canonical cell problem,

$$C^h_{ijk\ell} = \frac{1}{|Y|}\int_Y C_{ijmn}(y)\Big[T^{k\ell}_{mn} + \frac{\partial \chi^{k\ell}_m}{\partial y_\ell}\Big]dy = \Big\langle \hat{\sigma}^{k\ell}_{ij}\Big\rangle_Y \ . \qquad (10.129)$$

The homogenized BVP is,

$$\begin{aligned} < \sigma_{ij} >_{,j} + f_i &= 0, \ \ \forall x \in \Omega && (10.130)\\ < \sigma_{ij} > n_j &= t_i, \ \ \forall x \in \Gamma_t && (10.131)\\ u_i^{(0)} &= \bar{u}_i, \ \ \forall x \in \Gamma_u && (10.132)\end{aligned}$$

10.5.2 Finite element formulation

Choose $v_i \in H^1_{\#}(Y)$. Multipling v_i with the leading order equilibrium equation (10.111) and integrating it over Y, we have

$$\int_Y \frac{\partial \sigma_{ij}^{(0)}}{\partial y_j} v_i d\Omega_y = 0, \quad \forall v_i \in H^1_{\#}(Y)$$

Integration by parts yields,

$$\begin{aligned} &\int_Y \sigma_{ij}^{(0)} n_j v_i dS - \int_Y \sigma_{ij}^{(0)}\frac{\partial v_i}{\partial y_j}dV_y \\ = \ \ & -\int_Y \sigma_{ij}^{(0)}\frac{\partial v_i}{\partial y_j}dV_y = -\int_Y C_{ijk\ell}\frac{\partial u_k^{(0)}}{\partial y_\ell}\frac{\partial v_i}{\partial y_j}d\Omega_y = 0 \ .\end{aligned}$$

Let $v_i(x,y) = u_i^{(0)}(x,y)$. We have

$$\int_Y C_{ijk\ell}\frac{\partial u_k^{(0)}}{\partial y_\ell}\frac{\partial u_i^{(0)}}{\partial y_j}d\Omega_y = 0 \ . \qquad (10.133)$$

Since $C_{ijk\ell}(y)$ is positive definite,

$$\frac{\partial u_i^{(0)}}{\partial y_j} = 0 \,, \quad \Rightarrow \quad u_i^{(0)} = u_i^{(0)}(x)$$

and consequently $\sigma_{ij}^{(0)} = 0$, as we have derived before.

Multiply Eq. (10.127) with a test function, $v_i \in H^1_{\#}(Y)$, and integrate them over Y. Integration by parts yields,

$$\begin{aligned}
&\int_Y \frac{\partial}{\partial y_j}\Big[C_{ijk\ell}\Big(T_{k\ell}^{mn} + \frac{\partial \chi_k^{mn}}{\partial x_\ell}\Big)\Big] v_i dV_y \\
= \quad &\int_{\partial Y}\Big[C_{ijk\ell}\Big(T_{k\ell}^{mn} + \frac{\partial \chi_k^{mn}}{\partial x_\ell}\Big)\Big] n_j v_i dS_y - \int_Y C_{ijk\ell}\Big(T_{k\ell}^{mn} + \frac{\partial \chi_k^{mn}}{\partial x_\ell}\Big)\frac{\partial v_i}{\partial y_j} dV_y \\
= \quad & -\int_Y C_{ijk\ell}\Big(T_{k\ell}^{mn} + \frac{\partial \chi_k^{mn}}{\partial x_\ell}\Big)\frac{\partial v_i}{\partial y_j} dV_y = 0 \,.
\end{aligned}$$

Consider the following parametric vector,

$$\mathbf{P}^{mn} = y_m \delta_{nk} \mathbf{e}_k = P_k^{mn} \mathbf{e}_k \tag{10.134}$$

One can show that

$$T_{k\ell}^{mn} = \frac{1}{2}\Big(\frac{\partial P_\ell^{mn}}{\partial y_k} + \frac{\partial P_k^{mn}}{\partial y_\ell}\Big) = P_{(k,\ell)}^{mn}$$

Therefore, the weak formulation for the canonical cell problem can be written as

$$\frac{1}{|Y|}\int_Y C_{ijk\ell}(y)\Big(P_{(k,\ell)}^{mn} + \chi_{(k,\ell)}^{mn}\Big) v_{(i,j)} dV_y = 0 \,. \tag{10.135}$$

Define the bilinear form

$$a_Y(\mathbf{u}, \mathbf{v}) = \frac{1}{|Y|}\int_Y C_{ijk\ell}(y) u_{(i,j)} v_{(k,\ell)} dV_y \tag{10.136}$$

The finite element formulation of canonical cell problem is:

Find $\boldsymbol{\chi}^{mn} \in H^1_{\#}(Y)$, such that

$$a_Y(\mathbf{P}^{mn} + \boldsymbol{\chi}^{mn}, \mathbf{v}) = 0, \;\; \forall \;\; \mathbf{v} \in H^1_{\#}(Y) \tag{10.137}$$

Once $\chi_{k,\ell}^{mn}$ being determined, the effective elastic stiffness tensor can then be calculated based on definition

$$C_{stk\ell}^H = \frac{1}{|Y|}\int_Y C_{stmn}(y)(P_{m,n}^{k\ell} + \chi_{m,n}^{k\ell}(y)) dV_y \tag{10.138}$$

Consider the fact that

$$T_{st}^{ij} = P_{(s,t)}^{ij} = \frac{1}{2}\Big(\delta_{si}\delta_{tj} + \delta_{sj}\delta_{ti}\Big)$$

It is readily to show that

$$C_{stk\ell}^{H} T_{st}^{ij} = C_{stk\ell}^{H} \frac{1}{2}\Big(\delta_{si}\delta_{tj} + \delta_{sj}\delta_{ti}\Big) == C_{ijk\ell}^{H} \tag{10.139}$$

and

$$\begin{aligned} C_{ijk\ell}^{H} &= \frac{1}{|Y|}\int_{Y} C_{stmn}\Big(P_{m,n}^{k\ell} + \chi_{m,n}^{k\ell}) T_{st}^{ij} dV_y \\ &= \frac{1}{|Y|}\int_{Y} C_{stmn}\Big(P_{m,n}^{k\ell} + \chi_{m,n}^{k\ell}) P_{(s,t)}^{ij} dV_y \\ &= a_Y(\mathbf{P}^{k\ell} + \boldsymbol{\chi}^{k\ell}, \mathbf{P}^{ij}\Big) dV_y \end{aligned} \tag{10.140}$$

Finally, we define another function space,

$$\mathcal{V}_\Omega = \left\{ \mathbf{v}(x), x \in \Omega \;\Big|\; \mathbf{v}(x) \in [H^1(\Omega)]^d, d = dim\{\Omega\}, \text{ and } , \mathbf{v}(x)\Big|_{\Gamma_u} = 0 \right\}$$

The weak formulation for the following macro-level BVP,

$$\begin{aligned} \frac{\partial < \sigma_{ij}^{(1)} >_Y}{\partial x_j} + f_i &= 0, && (10.141) \\ \text{where} \quad < \sigma_{ij}^{(1)} > &= C_{ijk\ell}^{H} u_{(k,\ell)}^{(0)} && (10.142) \\ \text{and} \quad \frac{\partial}{\partial x_j}\Big[C_{ijk\ell}^{H} u_{(k,\ell)}^{(0)}\Big] + f_i &= 0, \;\; \forall x \in \Omega && (10.143) \\ < \sigma_{ij}^{(1)} > n_j &= t_i^0, \;\; \forall\, x \in \Gamma_t && (10.144) \\ u_i^{(0)} &= \bar{u}_i, \;\; \forall\, x \in \Gamma_u && (10.145) \end{aligned}$$

is:

Find $\mathbf{u}^{(0)}(x) \in \mathcal{V}_\Omega$ such that

$$\int_\Omega C_{ijk\ell}^{H} u_{(k,\ell)}^{(0)} v_{(i,j)} dV_x = \int_\Omega f_i v_i dV_x + \int_{\Gamma_t} t_i^0 v_i dS, \;\; \forall\, \mathbf{v} \in \mathcal{V}_\Omega \,. \tag{10.146}$$

where $\mathbf{v} = v_i \mathbf{e}_i$.

Summary of Multiscale Finite Element Method

1. Solve the canonical cell problem on Y first, i.e. find $\boldsymbol{\chi}^{k\ell}(y) \in H^1_{\#}(Y)$ by solving
$$a_Y\Big(\mathbf{P}^{k\ell} + \boldsymbol{\chi}^{k\ell}, \mathbf{v}\Big) = 0, \ \ \forall \mathbf{v} \in H^1_{\#}(Y)$$

2. Calculate macro-scale elastic stiffness tensor
$$C^H_{ijk\ell} = a_Y\Big(\mathbf{P}^{ij} + \boldsymbol{\chi}^{ij}, \mathbf{P}^{k\ell}\Big) \ \text{and} \ a_Y(\mathbf{u}, \mathbf{v}) := \frac{1}{Y}\int_Y C_{ijk\ell}(y) u_{i,j} v_{k,\ell} dV_y$$

3. Solve the macro displacement field, $\mathbf{u}^{(0)}(\mathbf{x}) \in \mathcal{V}_\Omega$,
$$a^H_\Omega(\mathbf{u}^{(0)}, \mathbf{v}) = \Big\langle \mathbf{f}, \mathbf{v} \Big\rangle_\Omega + \Big\langle \mathbf{t}^0, \mathbf{v} \Big\rangle_{\Gamma_t} \ \text{where} \ a^H_\Omega(\mathbf{u}, \mathbf{v}) := \int_\Omega C^H_{ijk\ell} u_{i,j} v_{k,\ell} dV_y$$
where $\mathbf{v}$ is any function in $\mathcal{V}_\Omega$;

4. Calculate the fine (local) scale stress distribution,
$$\sigma^{(1)}_{ij}(x, y) = C_{ijk\ell}(y)\Big(T^{mn}_{k\ell} + \chi^{mn}_{(k,\ell)}(y)\Big)\frac{\partial u^{(0)}_m}{\partial x_n}$$

10.6 G-, H-, and Γ- convergence

Various notions of convergence are introduced in relation to asymptotic homogenization theory, such as Γ-convergence of De Giorgi [1975][1984], the G-convergence of Spagnolo [1968][1976], and the H-convergence of Tartar [1978]. These abstract mathematical notions provide powerful tools to analysis various numerical simulations of homogenization.

The question we would like to answer is: what is the limit in a homogenization process when micro-scale approaches to zero (Fig. (10.8) ? does upscale homogenizations will eventually converge to that limit ?

To answer this questions, we have to first define what do we mean by convergence, or convergence in what sense.

10.6.1 Strong convergence and weak convergence

We first discuss the notion of strong convergence and weak convergence of functions in Banach spaces.

Let Ω be an open set in $\mathbb{R}^d$. For $1 \le p \le +\infty$, the Lebesgue space $L^p(\Omega)$ of all measurable functions u in Ω is a Banach space endowed with the following norm,
$$\|u\|_{L^p(\Omega)} = \Big(\int_\Omega |u|^p dx\Big)^{1/p}, \ \ \forall 1 \le p < +\infty$$

When $p = \infty$, we define the so-called essential supremum
$$\|u\|_{L^\infty(\Omega)} = ess \sup_{x\in\Omega} |u(x)| := \inf_{\substack{Z\in\Omega \\ \mu(Z)=0}} \Big\{ \sup_{x\in\Omega - Z} |u(x)| \Big\}$$

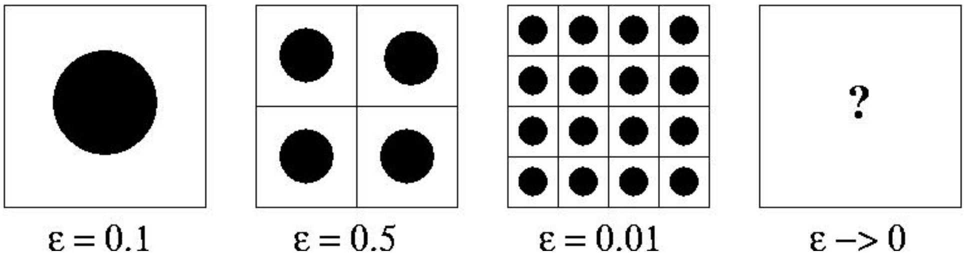


Figure 10.8. Notion of convergence in homogenization

Note that the physical meaning of $L^\infty(\Omega)$ space is that its occupant functions satisfying the condition $|u(x)| < \infty$ almost everywhere in Ω.

We use the short-handed notation, $\epsilon \to 0$ to denote a limit process of a sequence $\epsilon = \{\epsilon_1, \epsilon_2, \cdots \epsilon_n, \cdots\cdots\}$, and $\epsilon_n \to 0$ as $n \to \infty$.

The strong convergence of a function sequence, $u_\epsilon := \{u_{\epsilon_1}, u_{\epsilon_2}, \cdots, u_{\epsilon_n}, \cdots\}$, is measured by the distance in the particular normed space, i.e. a sequence, u_ϵ, is said to converge strongly in $L^p(\Omega)$ to a limit u_0, if

$$\lim_{\epsilon\to 0} \|u_\epsilon - u_0\|_{L^p(\Omega)} = 0 \ .$$

The strong convergence is denoted by an arrow, namely,

$$u_\epsilon \to u_0, \ \text{ in } L^p(\Omega) \ \text{ strongly}$$

On the other hand, the weak convergence is measured by a so-called weighted residual distance, which is associated with a weighting function, or test function in the dual space of the original norm space.

For the weak convergence in Lebesgue space $L^p(\Omega)$, the test function is in its dual space $L^{p'}(\Omega)$ with

$$\frac{1}{p} + \frac{1}{p'} = 1 \ .$$

Therefore, the formal statement of weak convergence in $L^p(\Omega),\ 1 \le p < +\infty$ is as follows: a sequence u_ϵ is said to converge weakly in $L^p(\Omega)$ to a limit u_0, if for any test function $\phi \in L^{p'}(\Omega)$, it satisfies

$$\lim_{\epsilon\to 0} \int_\Omega u_\epsilon(x)\phi(x)dx = \int_\Omega u(x)\phi(x)dx$$

The weak convergence is denoted by a harpoon, namely

$$u_\epsilon \rightharpoonup u_0 \ \text{ in } L^p(\Omega) \text{ weakly} \ .$$

The main interest of the weak convergence is that it is sequentially relative compact on bounded set. This means that for all the bounded sequence, $\|u_\epsilon\|_{L^p(\Omega)} \leq C$, there exists a subsequence $\left(u_\epsilon\right)_{\epsilon'>0}$ and a limit u_0 such that $\left(u_\epsilon\right)_{\epsilon'>0}$ converges weakly to u_0 in $L^p(\Omega),\ 1 < p < \infty$, which is not true for strong convergence.

Intuitively speaking, the strong convergence is more or less the usual pointwise convergence, while the weak convergence is a notion of convergence "in average" (up to a fluctuation of zero-mean).

If Ω is finite, we may choose test function

$$\phi(x) = \frac{1}{\Omega} \in L^{p'}(\Omega)$$

then $u_\epsilon(x) \rightharpoonup u_0(x)$ requires that

$$\lim_{\epsilon \to 0} \int_\Omega u_\epsilon = \frac{1}{\Omega} \int_\Omega u_\epsilon(x)dx = \frac{1}{\Omega} \int_\Omega u_0(x)dx$$

That is $\lim_{\epsilon \to 0} < u_\epsilon >_\Omega = < u_0 >_\Omega$.

We state (without proof) the connection between strong convergence and pointwise convergence. This statement is false for weakly convergence.

THEOREM 10.5 *1 Let Ω be a bounded open set in $\mathbb{R}^d$. Let u_ϵ be a sequence converging strongly to a limit u_0 in $L^p(\Omega)$, $1 \leq p \leq +\infty$, i.e.*

$$u_\epsilon(x) \to u_0(x)$$

Then there exists a subsequence, $u_{\epsilon'} \subset u_\epsilon$, and a function $h(x) \in L^p(\Omega)$ such that,

$$\lim_{\epsilon' \to 0} u_{\epsilon'}(x) = u_0(x), \quad \text{almost everywhere in } \Omega$$
$$|u_\epsilon(x)| \leq h(x), \quad \text{almost everywhere in } \Omega$$

2 Assume that the sequence $u_\epsilon(x)$ is bounded in $L^p(\Omega)$ $(1 < p \leq \infty)$, and

$$\lim_{\epsilon \to 0} u_\epsilon(x) = u_0(x), \quad \text{almost everywhere in } \Omega$$

Then

$$u_\epsilon(x) \to u_0(x) \text{ in } L^q(\Omega)\ (1 \leq q < p)\ strongly\ .$$

To feel the differences between strong convergence and weak convergence, we consider the following example.

EXAMPLE 10.6 *Let* $u_\epsilon(x) = \sin\Big(\dfrac{x}{\epsilon}\Big)$, $p = 2$, *and* $\Omega = (1, 0)$. *Choose test function* $\phi(x) = 1$. *We have*

$$\begin{aligned}\int_0^1 u_\epsilon(x)\phi(x)dx &= \int_0^1 \sin\Big(\frac{x}{\epsilon}\Big)dx \\ &= -\epsilon\cos\Big(\frac{x}{\epsilon}\Big)\Big|_0^1 = \epsilon\Big(1 - \cos\Big(\frac{1}{\epsilon}\Big)\Big)\end{aligned}$$

As $\epsilon \to 0$, $u_\epsilon \rightharpoonup 0$, *weakly in* $L^2(\Omega)$, *i.e. the weak limit of the sequence* $u_\epsilon(x)$ *is zero.*

On the other hand, it seems that $u_\epsilon(x)$ *has no strong limit in* $L^2(\Omega)$. *This is because*

$$\begin{aligned}\|u_\epsilon\|_{L^2(\Omega)} &= \sqrt{\int_0^1 u_\epsilon^2(x)dx} = \sqrt{\int_0^1 \sin^2\Big(\frac{x}{\epsilon}\Big)dx} \\ &= \sqrt{\frac{1}{2}\int_0^1 \Big(1 - \cos\Big(\frac{2x}{\epsilon}\Big)\Big)dx} \\ &= \sqrt{\frac{1}{2}\Big(1 - \frac{\epsilon}{2}\sin^2(\frac{2}{\epsilon}\Big)}\end{aligned}$$

Suppose $u_\epsilon \to f(x)$ *and* $f(x) \in L^2(\Omega)$. *Therefore,*

$$\lim_{\epsilon\to 0}\int_0^1 \Big(\sin\frac{x}{\epsilon} - f(x)\Big)^2 dx = \int_0^1 \sin^2\Big(\frac{x}{\epsilon}\Big)dx - 2\int_0^1 \sin\Big(\frac{x}{\epsilon}\Big)f(x)dx$$
$$+ \int_0^1 f^2(x)dx = \frac{1}{2} + \int_0^1 f^2(x)dx \neq 0\ .$$

because $f(x) \in (L^2)'(\Omega)$.

Moreover, the fact that

$$\lim_{\epsilon\to 0}\int_0^1 \sin^2\Big(\frac{x}{\epsilon}\Big)dx = \frac{1}{2}$$

also indicates that the product of two weakly convergence sequences does not converge to the product of their weak limits. Otherwise,

$$\lim_{\epsilon\to 0}\int_0^1 \sin^2\Big(\frac{x}{\epsilon}\Big)dx = 0$$

because both $\sin\Big(\frac{x}{\epsilon}\Big) \rightharpoonup 0$ *in* $L^2([0, 1])$.

It is worth noting that the product of two strong convergence sequence does converge to the product of the two limits strongly, but it may be in a different Lebesgue space in general.

For instance, if both $u_\epsilon \to u_0$ in $L^2(\Omega)$ strongly and $v_\epsilon \to v_0$ in $L^2(\Omega)$ strongly, then

$$\begin{aligned}
\|u_\epsilon v_\epsilon - u_0 v_0\|_{L^2(\Omega)} &= \|(u_\epsilon - u_0)(v_\epsilon - v_0) + (u_\epsilon - u_0)v_0 + (v_\epsilon - v_0)u_0\|_{L^2(\Omega)} \\
&\leq \Big(\|u_\epsilon - u_0\|_{L^2(\Omega)}\Big)^{1/2}\Big(\|v_\epsilon - v_0\|_{L^2(\Omega)}\Big)^{1/2} \\
&\quad + \|v_0\|^{1/2}_{L^2(\Omega)}\Big(\|u_\epsilon - u_0\|_{L^2(\Omega)}\Big)^{1/2} \\
&\quad + \|u_0\|^{1/2}_{L^2(\Omega)}\Big(\|v_\epsilon - v_0\|_{L^2(\Omega)}\Big)^{1/2}
\end{aligned}$$

Hence

$$u_\epsilon v_\epsilon \to u_0 v_0 \text{ in } L^2(\Omega) \text{ strongly .}$$

Unfortunately, the same is not true for the weakly convergent sequences. In our previous example,

$$u_\epsilon(x) = \sin\Big(\frac{x}{\epsilon}\Big) \to 0 \text{ in } L^2(\Omega) \text{ weakly}$$

but for $u_\epsilon(x) = v_\epsilon(x) = \sin\Big(\frac{x}{\epsilon}\Big)$

$$u_\epsilon(x) v_\epsilon(x) \rightharpoonup \frac{1}{2}\,! \text{ in } L^p(\Omega) \;\; 1 \leq p < +\infty \,.$$

Moreover, in practice, if $u_\epsilon \rightharpoonup u_0$ in $L^p(\Omega)$, and $J(u)$ is a nonlinear functional, say quadratic functional, $J : L^p(\Omega) \to \mathbb{R}$.

It is usually

$$J(u_\epsilon) \not\rightharpoonup J(u_0) \text{ in any sense !}$$

10.6.2 G- Convergence

Consider our model homogenization BVP,

$$\begin{aligned}
L^\epsilon u_\epsilon &= f, \;\; x \in \Omega, \;\; \text{where } L^\epsilon = -\nabla \cdot \mathbf{A}(\frac{x}{\epsilon}) \cdot \nabla \\
u_\epsilon \Big|_{\partial\Omega} &= \bar{u}, \;\; \forall x \in \partial\Omega
\end{aligned}$$

where the heat conduction (or diffusion) coefficient $A_{ij}(y)$ are Y-periodic functions.

Suppose that solution of the above BVP can be found as

$$u_\epsilon(x) = \Big(L^\epsilon\Big)^{-1} f,$$

Obviously, $u_\epsilon \in H^1(\Omega)$ and $f \in H^{-1}(\Omega)$.

Recall the definition of Green's function. We have

$$u_\epsilon(x) = \Big(L^\epsilon\Big)^{-1} f = \int_\Omega G_\epsilon(x-y) f(y) dy$$

Suppose that there exists a weak limit $u_0(x)$ in $H^1(\Omega)$ such that

$$u_\epsilon(x) \rightharpoonup u_0(x) \text{ in } H^1(\Omega) \text{ weakly}$$

and the weak limit $u_0(x)$ has the representation,

$$u_0(x) = \int_\Omega G_0(x-y) f(y) dy =: \Big(L_0\Big)^{-1} f$$

Therefore, the weak convergence of $u_\epsilon(x)$, i.e. $u_\epsilon \rightharpoonup u_0(x)$, implies that

$$\int_{\Omega_x} \int_{\Omega_y} \Big(G_\epsilon(x-y) - G_0(x-y)\Big) f(y) dy dx = 0, \ \ \epsilon \to 0 \tag{10.147}$$

Change the order of integration, (10.147) yields

$$\int_{\Omega_y} f(y) \left(\int_{\Omega_x} \Big(G_\epsilon(x-y) - G_0(x-y)\Big) dx \right) = 0, \text{ as } \epsilon \to 0 \ . \tag{10.148}$$

Equation (10.148) suggests that the weak convergence of Green's function, i.e. $G_\epsilon \rightharpoonup G_0$, which implies a special type of convergence of the differential operator sequence $L^\epsilon = -\nabla \cdot \mathbf{A}^\epsilon \cdot \nabla$. We call the convergence of differential operator sequence L_ϵ as the G-convergence,

$$L^\epsilon \overset{G}{\to} L_0 \tag{10.149}$$

in the sense of

$$G_\epsilon * f \rightharpoonup G_0 * f, \ \text{ in } \ H^1(\Omega) \text{ weakly} \ .$$

Note that the symbol $*$ denotes the standard convolution.

In fact, the convergence of the differential operator sequence, $L^\epsilon = -\nabla \cdot A^\epsilon \cdot \nabla$, may be viewed as the convergence of matrix sequence, A^ϵ_{ij}, to its G-limit A^0_{ij}, or

$$A_{ij}\Big(\frac{x}{\epsilon}\Big) \overset{G}{\to} A^0_{ij}$$

The following definition of G-convergence is provided by Allaire.

Let $\mathcal{M}^s_d$ be the linear space of symmetric real matrices of order d. For any two positive constants $\alpha > 0$ and $\beta > 0$, we define a subspace of $\mathcal{M}^s_d$ made of coercive matrices with coercive inverse, namely,

$$\begin{aligned}\mathcal{M}^s_{\alpha,\beta} \ \ := \ \ & \Big\{ \{M_{ij}\} \in \mathcal{M}^s_d, \text{ such that } \alpha\xi^2 \le M_{ij}\xi_i\xi_j \\ & \text{and } \beta\xi^2 \le M^{-1}_{ij}\xi_i\xi_j, \ \forall \xi \in \mathbb{R}^d \Big\}\end{aligned}$$

Let Ω be a bounded open set in $\mathbb{R}^d$ and define the space $L^\infty(\Omega;\mathcal{M}^s_{\alpha,\beta})$ of admissible symmetric coefficient matrices.

We have the following definition of G-convergence,

DEFINITION 10.7 *A sequence of symmetric matrices, $A^\epsilon \in L^\infty(\Omega,\mathcal{M}^s_{\alpha,\beta})$ is said to be G-convergence to an homogenized, or G-limit, matrix $A^0 \in L^\infty(\Omega,\mathcal{M}^s_{\alpha,\beta})$, if, for any $f \in H^{-1}(\Omega)$., the sequence solution $u_\epsilon(x)$ of the following model problem*

$$\begin{aligned} -\nabla\cdot A^\epsilon \nabla u_\epsilon &= f, \;\; x\in\Omega \\ u_\epsilon &= \bar{u}, \;\; \forall x \in \partial\Omega \end{aligned}$$

converges weakly in $H^1(\Omega)$ to the solution of the homogenized BVP,

$$\begin{aligned} -\nabla\cdot A^0\cdot \nabla u_0 &= f, \;\; x\in\Omega \\ u_0 &= \bar{u}, \;\; \forall x \in \partial\Omega \end{aligned}$$

This definition makes sense because the following compactness theorem,

THEOREM 10.8 *For any sequence $A^\epsilon \in L^\infty(\Omega;\mathcal{M}_{\alpha,\beta})$ of symmetric matrices, there exits a subsequence, $A^{\epsilon'} \subset A^\epsilon$, and a limit $A^0 \in L^\infty(\Omega;\mathcal{M}_{\alpha,\beta})$ such that $A_{\epsilon'}$ G-converges to A^0.*

In the following examples, we want to show the differences between strong convergence, weak convergence, and G-convergence.

EXAMPLE 10.9 *In this example, suppose that we have two objects with the same macroscopic dimensions but different checkerborad microscopic structure.*

The diffusitity matrix coefficients are assumed to be

$$A_{ij} = a\delta_{ij}$$

We denote the diffusitivity in the white region as a_1 and the diffusitivity in the black region as a_2, and $a_2 > a_1$.

We denote the first micro-structure as $\mathcal{S}^\epsilon_1$ and the second micro-structure as $\mathcal{S}^\epsilon_2$.

Obviously, the first sequence $A^\epsilon(\mathcal{S}^\epsilon_1)$ and the second sequence $A^\epsilon(\mathcal{S}^\epsilon_2)$ have the same G-limit, i.e.

$$a^0(\mathcal{S}^\epsilon_1) = a^0(\mathcal{S}^\epsilon_2)\ .$$

As one can see that there is no pointwise convergence possibility, because for a fixed spatial point,

$$|a^0(\mathcal{S}^\epsilon_1) - a^0(\mathcal{S}^\epsilon_2)| = a_2 - a_1 > 0\ .$$

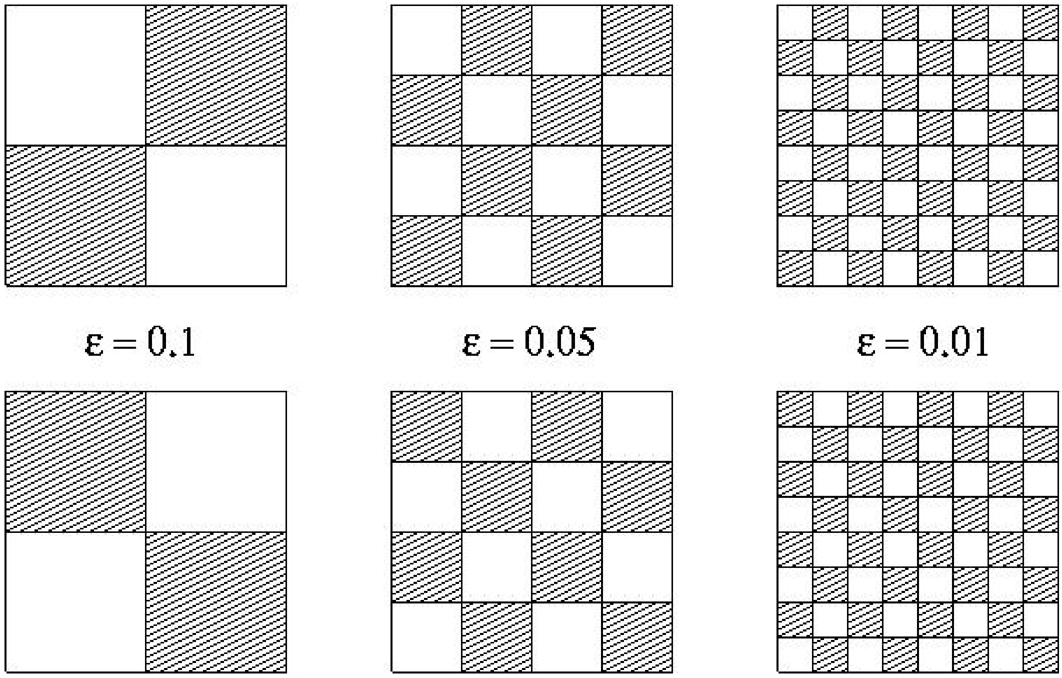


Figure 10.9. The difference between strong convergence and G-convergence

Nevertheless, in this example, indeed, the weak convergence limit of the two layouts are the same

$$< A^\epsilon(\mathcal{S}_1^\epsilon) >_\Omega = < A^\epsilon(\mathcal{S}_2^\epsilon) >_\Omega$$

EXAMPLE 10.10 *In this second example, we would like to show a case that there are two micro-structure layouts with the same weak convergence limits, but different G-limits.*

In this example, we assume that in each unit cell, the black and white areas are the same, therefore the volume fraction of the two phases are the same.

In the layout A, all the "good" material are connected, therefore it is a better arrangement for heat conduction, whereas in the layout B, all the "good" materials are isolated, disconnected, or insulated, it should be very hard for heat to diffuse from one point to another point.

Based on this argument, the two layouts should have different G-limit, and

$$a_0(\mathcal{S}_1^\epsilon) > a_0(\mathcal{S}_2^\epsilon) \; .$$

On the other hand,

$$< a(\mathcal{S}_1^\epsilon) > = < a(\mathcal{S}_2^\epsilon) >$$

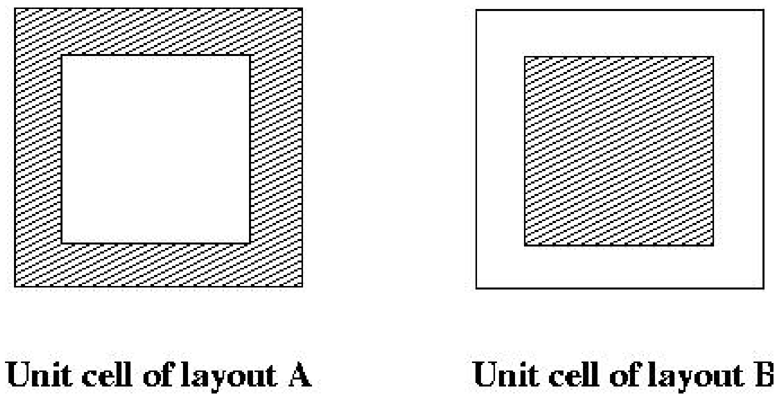


Figure 10.10. The difference between weak convergence and G-convergence

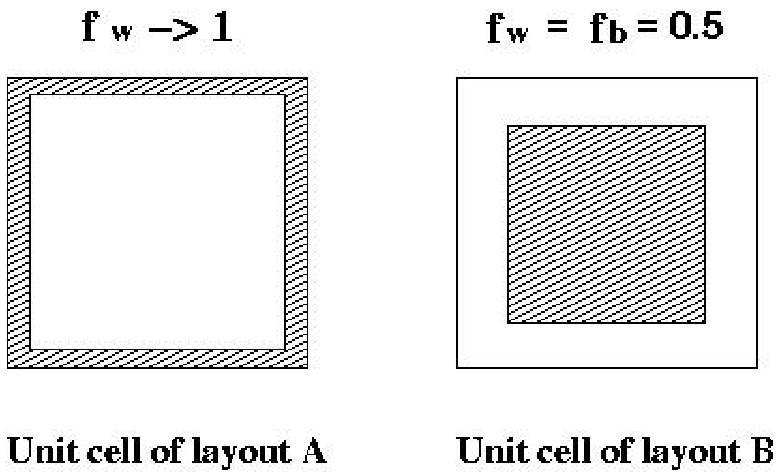


Figure 10.11. The difference between weak convergence and G-convergence

as indicated above.

EXAMPLE 10.11 *In the third example, we would like to show a case in which two microstructure layouts have the same G-limit but different weak convergence limits.*

In this example, we fix the second layout of the previous example. Therefore, we know that the G-limit of the second layout will be bounded by Vogit upper bound and Reuss lower bound, i.e.

$$\text{Reuss bound} = \frac{2a_1 a_2}{a_1 + a_2} \leq a^0\left(\mathcal{S}_2^\epsilon\right) \leq \frac{1}{2}(a_1 + a_2)$$

We know change the first layout by increase the volume fraction of insolated white phase, f_1 *such that* $f_1 \in [0.5,)$ *and* $f_1 \to 1$. *Therefore, the G-limit of the first layout will be bounded by*

$$\frac{1}{\dfrac{f_1}{a_1} + \dfrac{1-f_1}{a_2}} \le a^0(\mathcal{S}_1^\epsilon) \le f_1 a_1 + (1-f_1) a_2 \tag{10.150}$$

Initially when $f_1 = 0.5$ *we have,*

$$\frac{2a_1 a_2}{a_1 + a_2} < a^0\Big(\mathcal{S}_2^\epsilon\Big) < a^0\Big(\mathcal{S}_1^\epsilon\Big) < \frac{1}{2}(a_1 + a_2)$$

If $a_1 << a_2$.

The Reuss bound for the second layout is almost $\approx 2a_1$. *From Eq. (10.150), one can see that as* $f_1 \to 1$, *the Reuss bound (lower bound) of the first layout will become*

$$\frac{1}{\dfrac{f_1}{a_1} + \dfrac{1-f_1}{a_2}} \to a_1, \;\; \text{as} \;\; f_1 \to 1 \; .$$

This suggests that at certain volume fraction, $0.5 < f_1 = f_w < 1.0$, *the G-limits of the two layouts will be the same, i.e.*

$$a^0(\mathcal{S}_1^\epsilon) = a^0(\mathcal{S}_2^\epsilon) \; .$$

At that moment, since $f_w > 0.5 \neq f_2$, *the weak convergence limits of the two layouts will not be the same, i.e.*

$$< a^0(\mathcal{S}_1^\epsilon) >= f_w a_1 + (1-f_w) a_2 \neq < a^0(\mathcal{S}_2^\epsilon) >= 0.5(a_1 + a_2) \; .$$

10.6.3 H- Convergence

H-convergence is a generalization of G-convergence, in which, the differential operator A^ϵ, or its coefficient matrix, does not require to be symmetric anymore.

DEFINITION 10.12 (DEFINITION OF H-CONVERGENCE) *A sequence of matrices* $\mathbf{A}^\epsilon$ *in* $L^\infty(\Omega, M_{\alpha,\beta})$ *is said to converge in the sense of homogenization, or simply H-convergence, to an homogenized limit, or H-limit, matrix* $\mathbf{A}^0 \in L^\infty(\Omega, M_{\alpha,\beta})$ *if, for any right hand side* $f \in H^{-1}(\Omega)$, *the sequence* u_ϵ *of solution of*

$$-\nabla \cdot \mathbf{A}^\epsilon \cdot \nabla u_\epsilon = f(x), \;\; \forall x \in \Omega \tag{10.151}$$

$$u_\epsilon = \bar{u}, \;\; \forall x \in \partial\Omega \tag{10.152}$$

satisfies

$$u_\epsilon(x) \to u_0(x) \ \text{ weakly in } \mathrm{H}^1(\Omega) \tag{10.153}$$

$$\mathbf{A}^\epsilon \cdot \nabla u_\epsilon \to \mathbf{a}^* \cdot \nabla u_0 \ \text{ weakly in } \left[L^2(\Omega)\right]^N \tag{10.154}$$

where u_0 is the solution of the homogenized equation,

$$-\nabla \cdot \mathbf{A}^0 \cdot \nabla u_0 = f(x), \ \ \forall x \in \Omega \tag{10.155}$$

$$u_0 = \bar{u}, \ \ \forall x \in \partial\Omega \tag{10.156}$$

10.6.4 Γ- Convergence

For a large class of elliptical BVPs, each BVP under consideration has one-to-one correspondence to a variational principle. The well-known Lax-Milgram theorem guarantees the equivalence between the two.

Therefore, the convergence of differential operators may imply a possible convergence of the corresponding functional in the related function spaces.

DEFINITION 10.13 (DEFINITION OF Γ-CONVERGENCE) *Let X be a functional space endowed with a norm $\|\cdot\|_d$. Let ϵ be a sequence of positive indexes which goes to zero. Let F_ϵ be a sequence of functional defined on X with values in $\mathbb{R}$. The sequence F_ϵ is said to Γ-convergence to a limit functional F_0 if, for any function $x \in X$,*

1 *all sequences x_ϵ converging to x satisfy*

$$F_0(x) \le \lim_{\epsilon \to 0} \inf_{x \in X} F_\epsilon(x_\epsilon)$$

and

2 *there exists at least one sequence x_ϵ converging to x, such that*

$$F_0(x) = \lim_{\epsilon \to 0} F_\epsilon(x_\epsilon)$$

EXAMPLE 10.14 (AN EXAMPLE OF Γ-CONVERGENCE) *Consider the following diffusion problem, with diffusion coefficient matrix, $\mathbf{A}^\epsilon$ is symmetric and Y-periodic,*

$$-\nabla \cdot \mathbf{A}\Big(\frac{x}{\epsilon}\Big)\nabla u_\epsilon = f, \quad \forall x \in \Omega \tag{10.157}$$

$$u_\epsilon(x) = 0, \ \ \forall\, x \in \partial\Omega \tag{10.158}$$

The BVP (1) and (2) is equivalent to the following variational problem:

Find $u_\epsilon \in H_0^1(\Omega)$ such that

$$\inf_{u \in H_0^1} J(u) = \inf_{u \in H_0^1} \Big(\frac{1}{2} \int_\Omega \nabla u \cdot \mathbf{A}\Big(\frac{x}{\epsilon}\Big) \cdot \nabla u dx - \int_\Omega f u dx \Big)$$

Therefore. the Γ-convergence of $J_\epsilon(u)$ (with respect to the strong topology of $L^2(\Omega)$) is equivalent to the homogenization of the PDE (1)-(2).

10.7 Exercises

PROBELM 10.1 *Show that for isotropic materials the fourth-order tensor,*

$$\begin{aligned} g_{ijk\ell}(\boldsymbol{\xi}) &= \frac{1}{2\xi^2}\Big[\xi_j(\delta_{i\ell}\xi_k + \delta_{ik}\xi_\ell) + \xi_i(\delta_{j\ell}\xi_k + \delta_{jk}\xi_\ell)\Big] \\ &\quad - \frac{1}{1-\nu}\frac{\xi_i\xi_j\xi_k\xi_\ell}{\xi^4} + \frac{\nu}{1-\nu}\frac{\xi_i\xi_j}{\xi^2}\delta_{k\ell}\Big] \,. \end{aligned} \tag{10.159}$$

PROBELM 10.2 *Consider cuboidal region of inelastic strain (eigenstrain) due to solute segregation forming cuboidal precipitates. The precipitate sub-domain (or inclusion) has the dimension $2a \times 2a \times 2a$, and the unit cell (U) has the dimension $2L \times 2L \times 2L$. The eigenstrain is assumed to have a constant value ϵ within each inclusion, and be zero outside the inclusion,*

$$\varepsilon^*_{ij} = \begin{cases} \delta_{ij}\varepsilon, & \forall \mathbf{x} \in \Omega; \\ 0\,, & \forall \mathbf{x} \in U/\Omega, \end{cases} \tag{10.160}$$

where

$$U = \Big\{\mathbf{x}\,\Big|\, -L \le x_i \le L,\ \ i = 1,2,3\Big\} \tag{10.161}$$

$$\Omega = \Big\{\mathbf{x}\,\Big|\, -a \le x_i \le a,\ \ i = 1,2,3\Big\}\ ,\text{and}\quad a < L \tag{10.162}$$

Find :

(a) the disturbed displacement field $u_1(\mathbf{x})$ (Hint: Mura's book pages: 20-21).

(b) $G(\boldsymbol{\xi}) = g_0(\boldsymbol{\xi})g_0(-\boldsymbol{\xi})$.

PROBELM 10.3 *Consider the followin boundary-value problem in a medium with periodic structure,*

$$-\frac{\partial^2 u_\epsilon}{\partial x_i \partial x_i} = f, \quad \forall x \in \Omega \tag{10.163}$$

$$u_\epsilon = 0, \quad \forall x \in \partial\Omega \tag{10.164}$$

$$\frac{\partial u_\epsilon}{\partial n} = 0, \quad \forall x \in \Gamma \tag{10.165}$$

where Γ is the interface between the matrix and inhomogeneous phase.

Show that the homogenized differential equation is

$$-q_{ik}\frac{\partial^2 u_0}{\partial x_k \partial x_k} = f, \quad \forall x\ \in \Omega$$

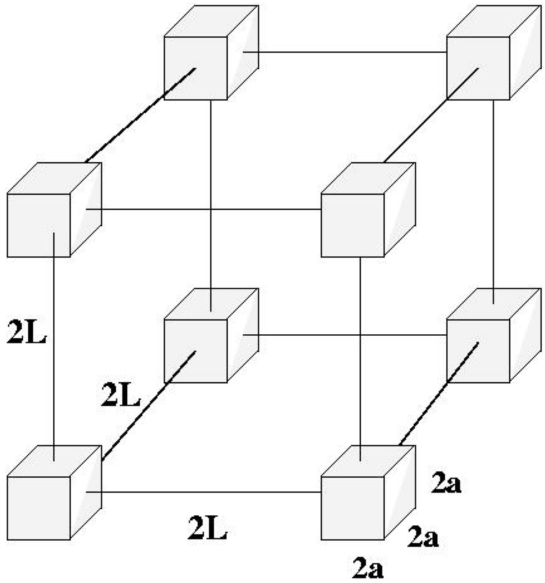


Figure 10.12. Distribution of periodic precipitates

with effective coefficients q_{ik} *defined as*

$$q_{ik} = \frac{1}{|Y|}\int_Y \Big(\delta_{ik} + 2\frac{\partial U_k}{\partial y_i}\Big) dy$$

and the associated canonical cell problem is

$$\begin{aligned} \frac{\partial^2 U_k}{\partial y_i \partial y_i} &= 0, \quad \forall y \in Y && (10.166) \\ \frac{\partial U_k}{\partial y_i} n_i &= n_k, \quad \forall y \in S && (10.167) \end{aligned}$$

10.8 Toshia Mura

This is the biography sketch of Professor Toshio Mura, the sole author of our second text book, " Micromechanics of Defects in Solids". The biography sketch was written more than 10 years ago by Professor Mori (who also made some contributions in micromechanics as well, the Mori-Tanaka theory, for instance, bears his name). Before I copy the biography sketch, I would say few things about professor Mura myself. For the past four and five years, I have the opportunity to study and work with Professor Mura, and I have stayed with

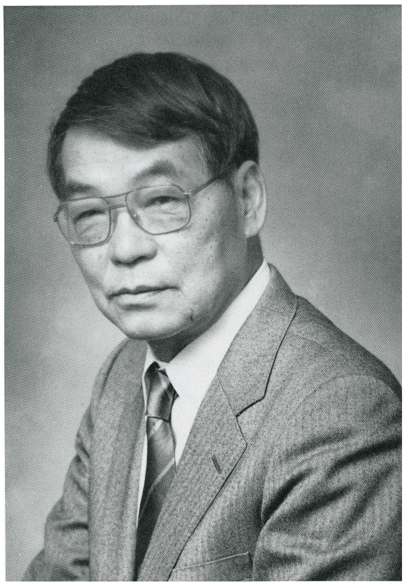


Figure 10.13. Toshio Mura

him in the same office for almost four years (I was a postdoctal fellow then and he was an emeritus professor).

Almost every week, he took me to lunch (because he insisted to pay everytimes, so we can not go out everyday), and I learned a lot of things from Professor Mura, and had many good conversations as well as good memories. Last year, Professor Mura received the Japanese Imperial model—the highest honor bestow by Janpanes emperor and Royal family to scientists and other citizens—for his contribution in micromechanics. I remembered back in 1997, in his retirement party, professor Jan Achenbach said that Professor Mura is one of the “seven samurai” (an international renowed Japanese moive, samurai in Japanese means warrior, previously in Northwstern there were seven famous Mechanics professors: Achenbach, Belytschko, Dundurs, Keer, Mura, Nemat-Nasser, and Bazant). Professor Mura is a theoretician, and has a very “romantic” outlook of the world, (romantic is opposed to the “down-to-earth” mentality of experimentalist) he believes that you are at your most creative stage, when you are in your dream.

Biography sketch of Toshio Mura.

“ Toshio Mura, second son of Shinzo and Chie Fujii, was born in Ono, a small port village of Kanazawa, the capital of Ishikawa Prefecture, Japan, on December 7, 1925. Among the locals, the Fujiis are well known as brewers having a long history in the area. Kanazawa is an old city on the coast of the

Sea of Japan, where traditional culture is proudly maintained and apprecaited. .....

In 944, during the most difficult time of the war, Mura went to the Imperial University of Tokyo to read Aeronautical Engineering. After the war, his department was dissolved and changed to the Department of Applied Mathematics at the University of Tokyo. ....

The title of his Ph.D. dissertation was "Study on Thermal Stresses". His work in the dissertation turned out to be one of the earliest papers on the dynamic wave of thermal stresses.

As a graduate student, Mura also began his teaching career as a mathematics professor at Meiji University, where he met and worked with his lifelong friend, Nobuo Kinoshita. Their joint paper, "On the boundary value problem of elasticity," which was published during his tenure at Mriji University (1956), agitated some Russian mathematicians in the field of integral equations. Had this work been extended, it would have led to the powerful computational technique now known as the boundary element method. .....

At the graduate school, Mura was introduced to his future wife, Sawa, by her sister, Sumi, who had worked in the Department of Aeronautical Engineering. During the courtship, Mura often visited the Ozaki's and Sumi fondly recalls that he praised Sawa's cooking. They married in 1953 and their first daughter, Miyako, was born in 1955.

In 1958, Mura went to Northwestern University's Department of Materials Science, Evanston, Illinois, to work with John O. Brittain. While at this department, Mura conceived the idea of the Periodic Distribution of Dislocations, which was documented in a paper and published later in the Proceddings of the Royal Society of London as a communication by A. H. Cottrell and R. E. Peierls (1964). In this paper, for the first time, the Fourier method was used to obtain the elastic field of dislocations. As seen in his later publications, the Fourier method became Mura's favorite tool to analyze elastic fields.

In 1961 Mura jointed the department of civil engineering at Northwestern University as an assistant professor. The pleasant but stimulating atmosphere, brewed by his colleagues, John Dundurs and Leon Keer, also encouraged him. Dundurs and Mura obtained the elastic fields of dislocations parallel to a cylindrical inhomegeneity (1964). Keer and Mura analyzed a penny-shaped crack with a plastic zone by solving an integral equation, Mura's first paper concerned with a crack (1963).

In 1963, Mura succeeded in expressing the elastic field of a curved dislocation in a line integral, now known as Mura's Formula (1963). The line integral is along the dislocation and contains only the state quantities that characterize the dislocation. This solution was later extended by John R. Willis, who gave the field of a dislocation segment in the form algebraic equations, whcih equired the solution of sextic equation (1970). .... The paper in 1963 is also noteworthy for introducing the concept of a dislocation flux tensor, which is yseful when the dynamic motion of dislocations is examined. The period, during which Mura's Formula was found, coincided with his promotion to Associate Professor of Civil Engineering.

....

The dislocation density and flux tensors were applied to continuum plasticity theory. Believing that a stress appearing within the framework of continuum plasticity was the sum of external and dislocation stresses, Mura published a series papers, in the late 1960s, along these lines that emphasized the distribution and stress of dislocation.

In 1967 Mura became Professor of Civil Engineering. At that time Mura nad J. G. Kunag, his student, obtained the solutions for a pile-up of edge dislocations against the interfacial boundary between different materials.

The pioneering work of J. D. Eshelby, his beloved peer, appears to have inspired and stimulated Mura, as seen in his studies of static and dynamic fields of dislocations in anisotropic media and in dislocation pile-ups. As can be inferred from the preface to his book, Micromechanics of Defectcs in solids, Mura regards Eshelby's work on inclusions and inhomogeneities as being the most important and fundamental.

To Mura the evaluation of the disturbance in elastic fields due to elastic inhomogeneities is the most interesting application of the theory of inclusions. For example, Z. A. Moschovides and Mura solved the stress field caused by two inhomogeneities by applying the equivalent inclusion method with polynomial eigenstrains. A computer program, performing the numerical calculations, complained that the matrice involved for linear equations were singular. Moschovides looked for the bugs that might have caused this complaint, but no bugs were found. The linear equations were carefully examined analytically and the cause of the complaint was found. There existed certain distributions of eigenstrains that yields no elastic field. Rozo Furuhashi, a visiting scholar, and Mura later generalized this finding and showed that impotent inclusions exist in a general sense. The impotent inclusions have eigenstrains defined by derivatives of a continuous vector (displacement) that vanished at the boundary of the inclusions. This anecdote illustrates Mura's teachings: "study and examine a specific subject carefully. If there is anything strange and exciting, you can later generalize it in a broader sense."

Mura also interacted with experimentalists, who eagerly sought his advice and aid on issues of mathematics and mechanics. In particular, Morris E. Fine, and his students in Northwestern's Department of Materials Science and Engineering, benefited from this interaction in their studies of the fatigue of alloys. Mura also gained insight into material properties and structures by the interactions with these materials scientists.

....

In 1986, Mura was elected to membership in the National Academy of Engineering, U.S.A. with the citation, ‘For initiating and promoting micromechnics to bridge the gap between metal physics and engineering mechanics.’ During the same year, he was appointed Walter P. Murphy Professor in the Technological Institute at Northwestern University.

......”

Chapter 11

MICROMECHANICS THEORY OF VOID GROWTH

Damage theory of void growth is central to failure mechanism of ductile materials. In late 1960's and early 1970's, pioneer contribution have been made by several authors, Mclintock [1968], Rice and Tracy [1969], and Gurson [1972], using micro-mechanics techniques to develop damage theory in constitutive modeling of ductile materials.

The homogenization result obtained by Gurson marks a significant milestone in the development of micromechanics, because the outcome of the homogenization is foundamentally different from that of micro-elasticity theory. In micro-elasticity theory, the homogenized consititutive relations are virtually the same as the constitutive relation in micro-scale, i.e., linear elastic constitutive relations or generalized Hook's law. The only differences in constitutive laws at different scales are the magnitude and the spatial distribution of elastic constants. Whereas, in the Gurson model, a completely new constititive relation at macro-level emerges from the homogenization, which represents a new philosophy:

finding new physical laws and new mechanics by doing homogenization. This notion is so attractive, and it has remained the very ideal and ultimate objective of contemporary micromechanics and multiscale simulations.

11.1 Void Growth in Linear Viscous Solids

Consider a linear viscous RVE, whose constitutive behaviors at microscale can be described as the following rate dependent expression,

$$\sigma_{ij} = C_{ijk\ell}\dot{\epsilon}_{k\ell}$$

The viscous coefficients resemble to that of linear elastic tensor,

$$C_{ijk\ell} = \frac{2\eta\nu}{1-2\nu}\delta_{ij}\delta_{k\ell} + \eta(\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk})$$

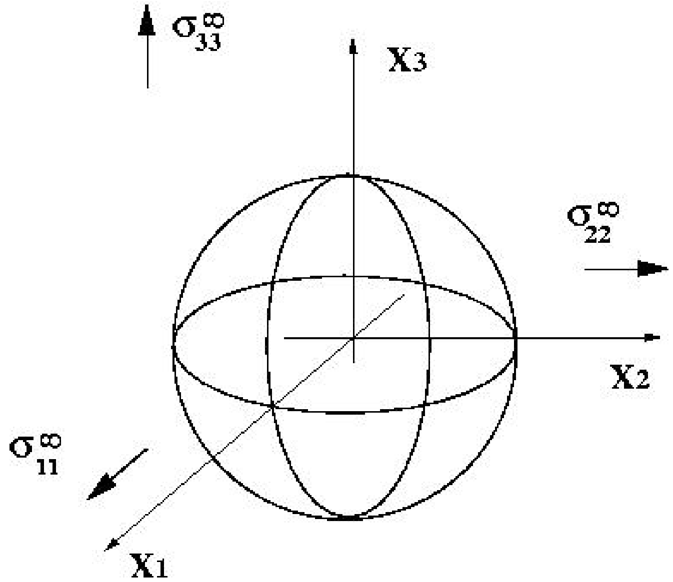


Figure 11.1. A spherical void in the middle of an RVE

In the case of incompressible viscous media,

$$C_{ijk\ell} = 2\eta\Big[\frac{1}{2}\Big(\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{jk}\Big) - \frac{1}{3}\delta_{ij}\delta_{k\ell}\Big] + \frac{2}{3}\eta\delta_{ij}\delta_{k\ell}$$

Consider a spherical void, Ω, inside an RVE with a radius, $R = a$. A uniform triaxial stress state is imposed at the remote boundary of the RVE, i.e.

$$t_i = \sigma^{\infty}_{ij} n_j, \;\; \forall x \in \partial V$$

where $\sigma^{\infty}_{ij} = T\delta_{ij}$.

Applying Eshelby's equivalent eigenstrain principle, the stress inside the void may be written as

$$\sigma_{ij} = C_{ijk\ell}\Big(\dot{\epsilon}^{\infty}_{k\ell} + \dot{\epsilon}^{d}_{k\ell} - \dot{\epsilon}^{*}_{k\ell}\Big)$$

Note that $\dot{\epsilon}^{\infty}_{ij} = D_{ijk\ell}\sigma^{\infty}_{k\ell}$ and $D_{ijk\ell} = C^{-1}_{ijk\ell}$.

Since inside the void, there is no stress $\sigma_{ij} = 0$, we have

$$\dot{\epsilon}_{ij} = \dot{\epsilon}^{\infty}_{ij} + \dot{\epsilon}^{d}_{k\ell} = \dot{\epsilon}^{*}_{ij}$$

This means that eigenstrain rate should be the same as the actual strain rate, which gives the physical meaning for eigenstrain rates. That is the prescribed eigenstrain rate should be the expansion rate of the void.

Moreover, one can find that

$$\sigma_{ij}^{\infty} = C_{ijk\ell}(\dot{\epsilon}_{k\ell}^{*} - \dot{\epsilon}_{k\ell}^{d})$$

By Eshelby's single inclusion solution, one can write

$$\epsilon_{ij}^{d} = S_{ijk\ell}\dot{\epsilon}_{ij}^{*}$$

Therefore,

$$\boldsymbol{\sigma} = \mathbf{C} : (\mathbf{1}^{(4s)} - \mathbf{S}) : \dot{\boldsymbol{\epsilon}} \ .$$

Denote

$$\mathbf{Q} := \mathbf{C} : (\mathbf{1}^{(4s)} - \mathbf{S}) \ .$$

The remote stress can be related with volumetric strain rate of the void, i.e.

$$\sigma_{ii}^{\infty} = Q_{ii11}\dot{\epsilon}_{11}^{*} + Q_{ii22}\dot{\epsilon}_{22}^{*} + Q_{ii33}\dot{\epsilon}_{33}^{*}$$

Consider,

$$\begin{aligned} \mathbf{C} &= 2\eta 1 + \nu + 1 - 2\nu \mathbf{E}^{(1)} + 2\eta \mathbf{E}^{(2)} \\ \mathbf{S} &= s_1 \mathbf{E}^{(1)} + s_2 \mathbf{E}^{(2)} \\ E_{ii11}^{(1)} &= 1, \ \text{and} \ E_{ii11}^{(2)} = 0 \ , \end{aligned}$$

where $s_1 = \dfrac{1+\nu}{3(1-\nu)}$ and $s_2 = \dfrac{2(4-5\nu)}{3(1-\nu)}$.

One may find that

$$Q_{ii11} = Q_{ii22} = Q_{ii33} = 8\eta \frac{(1+\nu)}{3(1-\nu)} \ .$$

By symmetry, it is easy to see that

$$\dot{\epsilon}_{11}^{*} = \dot{\epsilon}_{22}^{*} = \dot{\epsilon}_{33}^{*} = \dot{\epsilon}$$

Consequently, we have

$$T = \frac{8}{3}\frac{\eta(1+\nu)}{(1-\nu)}\dot{\epsilon}$$

Since the volume of the void is,

$$V = \frac{4\pi}{3}a^3 \ \Rightarrow \ \dot{V} = 4\pi a^2 \dot{a},$$

The relative void growth rate will be

$$\frac{\dot{V}}{V} = 3\frac{\dot{a}}{a} = 3\dot{\epsilon}$$

where $\dot{\epsilon} = \frac{\dot{a}}{a}$ is the strain rate in radial direction.

Finally, we link the magnitude remote stress with the void growth rate,

$$T = \frac{8}{9}\frac{\eta(1+\nu)}{1-\nu}\frac{\dot{V}}{V}$$

The above solution was obtained by Budiansky et al in 1981, almost ten years after publication of the McClintock solution and the Gurson model.

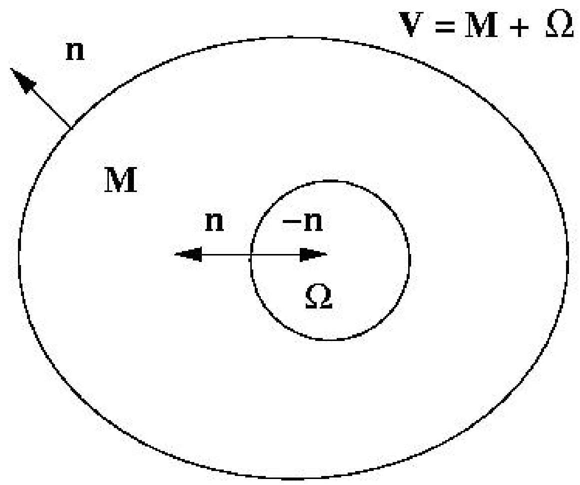


Figure 11.2. A solid with traction-free defect

11.1.1 Averaging theorems for soilds with traction-free defects

Consider an RVE, V, containing a traction-free defect, Ω. That is the traction force $t_i = \sigma_{ij}n_j = 0,\ \forall \mathbf{x} \in \partial\Omega$. Suppose that on the remote boundary condition ∂V the prescribed traction boundary condition is imposed

$$t_i = \sigma_{ij}n_j = \Sigma_{ij}n_j \ \ \forall \mathbf{x} \in \partial V$$

where Σ_{ij} is a contant tensor, and it is often denoted as the macro-stress tensor.

The following averagy theorems hold in the RVE,

$$\mathbf{1}. \qquad < \sigma_{ij} >_V = \Sigma_{ij} \tag{11.1}$$

$$\mathbf{2}. \qquad < \dot{\epsilon}_{ij} >_V = \dot{E}_{ij} + \dot{\epsilon}^{(add)}_{ij} \tag{11.2}$$

$$\text{where} \qquad \dot{\epsilon}^{(add)}_{ij} = \frac{1}{2V}\int_{\partial\Omega}\Big(\dot{u}_i n_j + \dot{u}_j n_i\Big) dS \tag{11.3}$$

and $\dot{E}_{ij} = D_{ijk\ell}\Sigma_{k\ell}$ for the linear viscous solid.

Expressions (11.2) and (11.3) are called additional strain rate formulas [1].

We first show (11.1),

$$
\begin{aligned}
< \sigma_{ij} > &= \frac{1}{V}\int_V \sigma_{ij} dV = \frac{1}{V}\int_V \Big(\sigma_{ip} x_j\Big)_{,p} dV \\
&= \frac{1}{V}\Big(\int_{\partial V} \sigma_{ip} n_p x_j dS - \underbrace{\int_{\partial \Omega} \sigma_{ip} n_p x_j dS}_{=0,\ \text{because}\ \sigma_{ip}n_p=0,\ \forall x \in \partial\Omega} \Big) \\
&= \frac{1}{V}\int_{\partial V} \Sigma_{ip} n_p x_j dS = \Sigma_{ij}
\end{aligned}
$$

We know that under the prescriber traction boundary condition,

$$< \dot{\epsilon}_{ij} > \neq \dot{E}_{ij}$$

To prove the additional strain rate formula, we use the so-called reciprocal theorem of virtual power. Consider two sets of traction boundary conditions and the corresponding velocity fields on the same ilinear viscous RVE, V, the following equality holds,

$$\int_{\partial V \bigcup \partial\Omega^-} t_i^{(1)} \dot{u}_i^{(2)} dS = \int_{\partial V \bigcup \partial\Omega^-} t_i^{(2)} \dot{u}_i^{(1)} dS$$

Let the traction b.c. for the first state be

$$\mathbf{t}^{(1)} = \mathbf{n} \cdot \delta\mathbf{\Sigma},\ \ \forall x \in \partial V \bigcup \partial\Omega^-$$

which yields the following trivial solution,

$$\{\dot{\mathbf{u}}^{(1)}, \dot{\boldsymbol{\epsilon}}^{(1)}, \boldsymbol{\sigma}^{(1)}\} = \{\mathbf{x} \cdot \delta\dot{\mathbf{E}}, \delta\dot{\mathbf{E}}, \delta\mathbf{\Sigma}\}$$

where $\delta\dot{\mathbf{E}} = \mathbf{D} : \delta\mathbf{\Sigma}$.

Let the traction b.c. for the second state as

$$\mathbf{t}^{(2)} = \begin{cases} \mathbf{n} \cdot \mathbf{\Sigma}, & \forall x \in \partial V \\ 0, & \forall x \in \partial\Omega^- \end{cases}$$

and it correspondes to the real solution,

$$\{\dot{\mathbf{u}}^{(2)}, \dot{\boldsymbol{\epsilon}}^{(2)}, \boldsymbol{\sigma}^{(2)}\} = \{\mathbf{u}, \dot{\boldsymbol{\epsilon}}, \delta\boldsymbol{\sigma}\}$$

[1] A similar expression is hold for infinitesimal strain as well.

The reciprocal theorem gives,

$$\int_{\partial V} t_i^{(1)} \dot{u}_i^{(2)} dS + \int_{\partial\Omega^-} t_i^{(1)} \dot{u}_i^{(2)} dS = \int_{\partial V \bigcup \partial\Omega^-} t_i^{(2)} \dot{u}_i^{(1)} dS + \underbrace{\int_{\partial\Omega^-} t_i^{(2)} \dot{u}_i^{(1)} dS}_{=0}$$

$$\int_{\partial V} \Big(\mathbf{n} \cdot \delta\mathbf{\Sigma}\Big) \dot{\mathbf{u}} dS + \int_{\partial\Omega^-} (\mathbf{n} \cdot \delta\mathbf{\Sigma}) \cdot \dot{\mathbf{u}} dS = \int_{\partial V} (\mathbf{n} \cdot \mathbf{\Sigma}) \cdot (\mathbf{x} \cdot \delta\dot{\mathbf{E}}) dS$$

Notice the following facts:

1

$$\mathbf{n} \cdot \delta\mathbf{\Sigma} \cdot \dot{\mathbf{u}} = \delta\mathbf{\Sigma} : \frac{1}{2}(\dot{\mathbf{u}} \otimes \mathbf{n} + \mathbf{n} \otimes \dot{\mathbf{u}})$$

2

$$\Big(\mathbf{n} \cdot \mathbf{\Sigma}\Big) \cdot \Big(\mathbf{x} \cdot \delta\dot{\mathbf{E}}\Big) = \delta\mathbf{\Sigma} : \mathbf{D} : \Big((\mathbf{x} \otimes \mathbf{n}) \cdot \mathbf{\Sigma}\Big)$$

We then have

$$\frac{1}{V}\delta\mathbf{\Sigma} : \{\int_{\partial V} \mathbf{D} : (\mathbf{x} \otimes \mathbf{n}) \cdot \mathbf{\Sigma} dS - \int_{\partial V} \mathbf{n} \otimes \dot{\mathbf{u}} dS - \int_{\partial\Omega} \mathbf{n} \otimes \dot{\mathbf{u}} dS\} = 0\,. \quad (11.4)$$

Consider

1

$$\mathbf{D} : \Big(\frac{1}{V}\int_{\partial V} \mathbf{x} \otimes \mathbf{n} dS\Big) \cdot \mathbf{\Sigma} = \dot{\mathbf{E}}; \quad (11.5)$$

2

$$Sym\,\frac{1}{V}\int_{\partial V} \mathbf{n} \otimes \dot{\mathbf{u}} dS = \frac{1}{V}\int_{V} \dot{\boldsymbol{\epsilon}} dV = < \dot{\boldsymbol{\epsilon}} > \quad (11.6)$$

3

$$\begin{aligned} Sym\,\frac{1}{V}\int_{\partial\Omega^-} \mathbf{n} \otimes \dot{\mathbf{u}} dS &= -\frac{1}{V}\int_{\partial\Omega} \mathbf{n} \otimes \dot{\mathbf{u}} dS \\ = \quad & -\frac{1}{2V}\int_{\partial\Omega}\Big(\mathbf{n} \otimes \dot{\mathbf{u}} + \dot{\mathbf{u}} \otimes \mathbf{n}\Big) dS \end{aligned} \quad (11.7)$$

Substitution (11.5)–(refeq:cond3) into (11.4) gives the following additional formula for strain rate

$$< \dot{\epsilon} > = \dot{\mathbf{E}} + \dot{\epsilon}^{(add)}$$

where

$$\dot{\boldsymbol{\epsilon}}^{(add)} = \frac{1}{2V}\int_{\partial\Omega}\Big(\dot{\mathbf{u}} \otimes \mathbf{n} + \mathbf{n} \otimes \dot{\mathbf{u}}\Big) dS$$

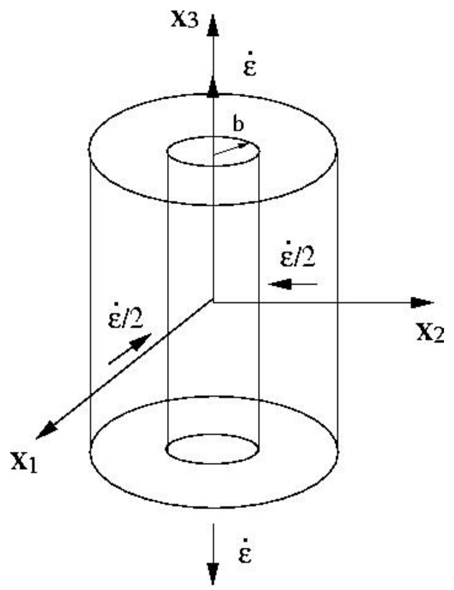


Figure 11.3. A cylindrical void in an inelastic RVE

11.2 The McClintock solution

The McClintock solution is the classic result of void growth in an inelastic RVE, which has been served as the bench mark example in many homogenizations of inelastic solids.

The basic premises of McClintock solution are two: (1) at micro-level, the RVE behaves as a rigid-plastic material, and (2) the RVE is incompressible.

Consider the following flow rule,

$$\dot{\epsilon}^p_{ij} = \dot{\lambda}\frac{\partial f}{\partial s_{ij}}$$

The yield surface is described by J_2 criterion (von Mises criterion),

$$f = J^2 - \frac{Y^2}{3} = \frac{1}{2}s_{ij}s_{ij} - \frac{Y^2}{3} = 0$$

where s_{ij} is the deviatoric stress tensor,

$$s_{ij} = \sigma_{ij} - \frac{1}{3}\sigma_{ii}$$

One can then rewrite the flow rule as

$$\dot{\epsilon}^p_{ij} = \dot{\lambda}\frac{\partial f}{\partial s_{ij}} = \dot{\lambda}s_{ij} \tag{11.8}$$

where the proportionality $\dot{\lambda}$ can be determined by contracting the flow rule with plastic strain rate, i.e.

$$\frac{1}{2}\dot{\epsilon}_{ij}\dot{\epsilon}_{ij} = \dot{\lambda}^2 \frac{1}{2} s_{ij} s_{ij} = \dot{\lambda}^2 \frac{Y^2}{3}$$

One can then solve for $\dot{\lambda}$,

$$\dot{\lambda} = \sqrt{\frac{3}{2}} \frac{\sqrt{\dot{\epsilon}^p_{ij}\dot{\epsilon}^p_{ij}}}{Y} = \frac{3}{2Y}\Big(\frac{2}{\sqrt{3}}\sqrt{I'_2(\dot{\epsilon}^p_{ij})}\Big)^{1/2} = \frac{3}{2}\frac{\dot{\bar{\epsilon}}^p}{Y}$$

where

$$I'_2 := \frac{1}{2}\epsilon^p_{ij}\epsilon^p_{ij} \qquad (11.9)$$

$$\dot{\bar{\epsilon}}^p = \frac{2}{\sqrt{3}} I'_2(\epsilon^p_{ij}) \qquad (11.10)$$

Therefore, the constitutive relation at micro-level are,

$$\dot{\epsilon}^p_{ij} = \frac{3}{2}\frac{\dot{\bar{\epsilon}}^p}{Y} s_{ij}$$

In the cylindrical coordinate,

$$\dot{\bar{\epsilon}}^p = \Big[\frac{2}{3}\Big((\dot{\epsilon}^p_r)^2 + (\dot{\epsilon}^p_\theta)^2 + (\dot{\epsilon}^p_z)^2\Big)\Big]^{1/2}$$

Consider the problem is axisymmetry and independent on z coordinate. The equlibrium equation becomes,

$$\frac{d\sigma_r}{dr} + \frac{\sigma_r - \sigma_\theta}{r} = 0 \ . \qquad (11.11)$$

Assume that the velocity field is

$$u_r = u(r), \ u_\theta = 0, \ \text{and} \ \dot{\epsilon}_z = \dot{\epsilon} = constant \ .$$

Hence,

$$\dot{\epsilon}_r = \frac{d\dot{u}}{dr} \qquad (11.12)$$

$$\dot{\epsilon}_\theta = \frac{\dot{u}}{r} \qquad (11.13)$$

The incompressible condition yields,

$$\dot{\epsilon}_r + \dot{\epsilon}_\theta + \dot{\epsilon}_z = \frac{d\dot{u}}{dr} + \frac{\dot{u}}{r} + \dot{\epsilon}_z = 0 \ .$$

Rewrite the above expression as

$$r\frac{d\dot{u}}{dr} + \dot{u} + r\dot{\epsilon}_z = 0 \;\Rightarrow\; \frac{d}{dr}\Big(r\dot{u}\Big) = -r\dot{\epsilon}_z$$

Integrate over the radial direction from the surface of the void to the interior of the RVE,

$$\int_b^r d\Big(\rho\dot{u}(\rho)\Big) = -\int_b^r \rho\dot{\epsilon}_z d\rho$$

Note that variable ρ is the dummy variable.

Considering $\dot{\epsilon}_z = \dot{\epsilon} = const.$, we have

$$\rho\dot{u}(\rho)\Big|_b^r = -\frac{\rho^2}{2}\dot{\epsilon}_z\Big|_0^r$$

Consequently,

$$r\dot{u}(r) - b\dot{b} = -\Big(r^2 - b^2\Big)\frac{\dot{\epsilon}_z}{2}$$

$$\Rightarrow \quad r\dot{u} = b\dot{b} + \frac{\dot{\epsilon}_z}{2}\Big(r^2 - b^2\Big)$$

Finally

$$\dot{u}(r) = \frac{b^2}{r}\Big(\frac{\dot{b}}{b} + \frac{\dot{\epsilon}_z}{2}\Big) - \frac{\dot{\epsilon}_z r}{2} \tag{11.14}$$

Let,

$$\sigma = \frac{1}{3}(\sigma_r + \sigma_\theta + \sigma_z)\ .$$

We have

$$s_r = \sigma_r - \sigma \tag{11.15}$$

$$s_\theta = \sigma_\theta - \sigma \tag{11.16}$$

The components of the flow rule in an axisymmetric plane are

$$\dot{\epsilon}_r = \dot{\epsilon}_r^p = \frac{3}{2}s_r\frac{\dot{\bar{\epsilon}}^p}{Y} = \frac{3}{2}(\sigma_r - \sigma)\frac{\dot{\bar{\epsilon}}^p}{Y} \tag{11.17}$$

$$\dot{\epsilon}_\theta = \dot{\epsilon}_\theta^p = \frac{3}{2}s_\theta\frac{\dot{\bar{\epsilon}}^p}{Y} = \frac{3}{2}(\sigma_\theta - \sigma)\frac{\dot{\bar{\epsilon}}^p}{Y} \tag{11.18}$$

(11.17) - (11.18) leads to

$$\dot{\epsilon}_\theta - \dot{\epsilon}_r = \frac{3}{2}(\sigma_\theta - \sigma_r)\frac{\dot{\bar{\epsilon}}^p}{Y} \tag{11.19}$$

Utilizing (11.19), it can be found that

$$\frac{\sigma_\theta - \sigma_r}{r} = \frac{2Y}{3r}\frac{\dot{\epsilon}_\theta - \dot{\epsilon}_r}{\dot{\bar{\epsilon}}^p}$$

Therefore the equilibrium equation becomes

$$\frac{d\sigma_r}{dr} + \frac{\sigma_r - \sigma_\theta}{r} = \frac{d\sigma_r}{dr} + \frac{2Y}{3r}\frac{(\dot\epsilon_r - \dot\epsilon_\theta)}{\dot{\bar\epsilon}^p} = 0\,. \tag{11.20}$$

Integrating over the radius direction,

$$\begin{aligned} \frac{1}{Y}\int_b^\infty d\sigma_r &= \frac{2}{3}\int_b^\infty \frac{(\dot\epsilon_\theta - \dot\epsilon_r)}{\dot{\bar\epsilon}^p}\frac{d\rho}{\rho} \\ \Rightarrow \quad \frac{1}{Y}[\sigma_r(\infty) - \sigma_r(b)] &= \frac{2}{3}\int_b^\infty \frac{(\dot\epsilon_\theta - \dot\epsilon_r)}{\dot{\bar\epsilon}^p}\frac{d\rho}{\rho} \end{aligned}$$

Consider the traction boundary condition,

$$\sigma_r(b) = 0, \ \text{ and } \ \sigma_r(\infty) = \sigma^\infty \tag{11.21}$$

We have

$$\frac{\sigma_r(\infty)}{Y} = \frac{2}{3}\int_b^\infty \frac{(\dot\epsilon_\theta - \dot\epsilon_r)}{\dot{\bar\epsilon}^p}\frac{d\rho}{\rho} \tag{11.22}$$

To integrate (11.22), one has to evalute $\dot{\bar\epsilon}^p$ first. Since

$$\dot u(r) = \frac{b^2}{r}\Bigl(\frac{\dot b}{b} + \frac{\dot\epsilon_z}{2}\Bigr) - \frac{\dot\epsilon_z r}{2},$$

direct calculation gives

$$\dot\epsilon_r = \frac{d\dot u_r}{dr} = -\frac{b^2}{r^2}\Bigl(\frac{\dot b}{b} + \frac{\dot\epsilon_z}{2}\Bigr) - \frac{\dot\epsilon_z}{2} \tag{11.23}$$

$$\dot\epsilon_\theta = \frac{\dot u_r}{r} = \frac{b^2}{r^2}\Bigl(\frac{\dot b}{b} + \frac{\dot\epsilon_z}{2}\Bigr) - \frac{\dot\epsilon_z}{2} \tag{11.24}$$

In cylindrical coordinate, the effective strain rate is

$$\begin{aligned} \dot{\bar\epsilon}^p &= \Bigl[\frac{2}{3}\Bigl((\dot\epsilon^p_r)^2 + (\dot\epsilon^p_\theta)^2 + (\dot\epsilon^p_z)^2\Bigr)\Bigr]^{1/2} \\ &= \left\{\frac{2}{3}\Bigl[\Bigl(\frac{b^2}{r^2}\Bigl(\frac{\dot b}{b} + \frac{\dot\epsilon_z}{2}\Bigr) + \frac{\dot\epsilon_z}{2}\Bigr)^2 + \Bigl(\frac{b^2}{r^2}\Bigl(\frac{\dot b}{b} + \frac{\dot\epsilon_z}{2}\Bigr) - \frac{\dot\epsilon_z}{2}\Bigr)^2 + \dot\epsilon_z^2\Bigr]\right\}^{1/2} \\ &= |\dot\epsilon_z|\Bigl[\frac{4}{3}\Bigl(\frac{\dot b}{b\dot\epsilon_z} + \frac{1}{2}\Bigr)^2\Bigl(\frac{b^2}{r^2}\Bigr)^2 + 1\Bigr]^{1/2} \end{aligned}$$

Define

$$x := \frac{b^2}{r^2}\frac{2}{\sqrt3}\Bigl(\frac{\dot b}{b\dot\epsilon_z} + \frac{1}{2}\Bigr) = \alpha\frac{b^2}{r^2} \tag{11.25}$$

where

$$\alpha := \frac{2}{\sqrt3}\Bigl(\frac{\dot b}{b\dot\epsilon_z} + \frac{1}{2}\Bigr) \tag{11.26}$$

Subsquently,

$$\dot{\bar{\epsilon}}_p = \dot{\epsilon}_x (1 + x^2)^{1/2} \tag{11.27}$$

and

$$\dot{\epsilon}_\theta - \dot{\epsilon}_r = 2\frac{b^2}{r^2}\Big(\frac{\dot{b}}{b} + \frac{\dot{\epsilon}_z}{2}\Big) = \sqrt{3}\dot{\epsilon}_z\Big(\frac{2}{\sqrt{3}}\frac{b^2}{r^2}\Big(\frac{\dot{b}}{b\dot{\epsilon}_z} + \frac{1}{2}\Big)\Big) = \sqrt{3}\dot{\epsilon}_z x \tag{11.28}$$

Since

$$dx = -\frac{2b^2}{r^3}\frac{2}{\sqrt{3}}\Big(\frac{\dot{b}}{b\dot{\epsilon}_z} + \frac{1}{2}\Big)dr = -\frac{2}{r}x dr,$$

$$\frac{dr}{r} = -\frac{1}{2}\frac{dx}{x} \ .$$

Make change of variable,

$$x = \alpha \frac{b^2}{r^2} \ ,$$

and

$$r = b, \ x \to \alpha; \ \ r \to \infty, \ x \to 0 \ .$$

We can then integrate (11.22)

$$\begin{aligned}
\frac{\sigma_\infty}{Y} &= \frac{2}{3}\int_b^\infty \frac{(\dot{\epsilon}_\theta - \dot{\epsilon}_r)}{\dot{\bar{\epsilon}}^p}\frac{d\rho}{\rho} = \frac{2}{3}\int_b^\infty \frac{\sqrt{3}\dot{\epsilon}_z x}{\dot{\epsilon}_z\sqrt{1+x^2}}\frac{dr}{r} \\
&= -\frac{1}{\sqrt{3}}\int_\alpha^0 \frac{dx}{\sqrt{1+x^2}} = \frac{1}{\sqrt{3}}\int_0^\alpha \frac{dx}{\sqrt{1+x^2}} \\
&= \frac{1}{\sqrt{3}}\mathrm{arcsinh}x \Big|_0^\alpha = \frac{1}{\sqrt{3}}\mathrm{arcsinh}(\alpha)
\end{aligned}$$

The inverse expression of the above result is

$$\frac{2}{\sqrt{3}}\Big(\frac{\dot{b}}{b\dot{\epsilon}_z} + \frac{1}{2}\Big) = \sinh\Big[\frac{\sqrt{3}\sigma_\infty}{Y}\Big] \tag{11.29}$$

Based on uniaxial tension test, one can measure

$$\tau_0 = \sqrt{J_2'} = \frac{Y}{\sqrt{3}}$$

We obtain the relationship between void growth rate and remote stress value,

$$\frac{\dot{b}}{b} = \frac{\sqrt{3}}{2}\dot{\epsilon}_z \sinh\Big[\frac{\sigma_\infty}{\tau_0}\Big] - \frac{1}{2}\dot{\epsilon}_z \tag{11.30}$$

A few comments about the McClintock solution are as follows:

1 McClintock solution is the only (essential) exact solution available for void growth in nonlinear viscous media;

2 McCintock solution reveals an exponential increase in the void growth rate under the positive remote stress load.

To illustrate the fact, we consider a finite cylindrical void with a heigh, H, and radius b. The volume of the cylinder is

$$\Omega = \pi b^2 H \ \Rightarrow \ \dot{\Omega} = 2\pi b\dot{b}H + \pi b^2 \dot{H}$$

Thereby,

$$\frac{\dot{\Omega}}{\Omega} = 2\frac{\dot{b}}{b} + \dot{\epsilon}_z$$

and hence

$$\frac{\dot{\Omega}}{\Omega} = \frac{\sqrt{3}}{2}\dot{\epsilon}_z \sinh\Big[\frac{\sqrt{3}\sigma_\infty}{Y}\Big] \tag{11.31}$$

Compare (11.31) with Budiansky et al's linear viscous void solution,

$$\frac{\dot{\Omega}}{\Omega} = \frac{9}{8}\frac{1-\nu}{\eta(1+\nu)}\sigma_\infty$$

One may appreciate the significant difference between the two.

3 At the remote boundary, $\mathbf{x} \in \partial V$,

$$\dot{\epsilon}_z = \dot{\epsilon},\ \dot{\epsilon}_r = \dot{\epsilon}_\theta = -\frac{1}{2}\dot{\epsilon}$$

Hence the macro equivalent strain rate is

$$\dot{\epsilon}^\infty_{eq} = \Big[\frac{2}{3}\dot{\epsilon}^\infty_{ij}\dot{\epsilon}^\infty_{ij}\Big]^{1/2} = \Big[\frac{2}{3}\Big(\frac{1}{4}\dot{\epsilon}^2 + \frac{1}{4}\dot{\epsilon}^2 + \dot{\epsilon}^2\Big)\Big]^{1/2} = \dot{\epsilon} \tag{11.32}$$

Bi-axial stress state is applied at the remote boundary, ∂V, i.e.

$$\Sigma_{11} = \Sigma_{22} = \sigma_\infty,\ \ \Sigma_{33} = T, \text{and}\ \ \Sigma_m = \frac{1}{3}\Big(2\Sigma_{11} + \Sigma_{33}\Big)$$

The von Mises criterion becomes

$$\begin{aligned}
\Sigma_{eq} &= \Big[\frac{3}{2}\Sigma_{ij}\Sigma_{ij}\Big]^{1/2} \\
&= \Big[\frac{3}{2}\Big((\Sigma_{11} - \Sigma_m)^2 + (\Sigma_{22} - \Sigma_m)^2 + (\Sigma_{33} - \Sigma_m)^2\Big)\Big]^{1/2} \\
&= |\Sigma_{33} - \Sigma_{11}| \le Y
\end{aligned}$$

The yield surface is $|\Sigma_{33} - \Sigma_{11}| = Y$.

Under such condition, we can rewrite the void growth rate equation as

$$\frac{\dot{\Omega}}{\Omega\dot{\epsilon}} = \sqrt{3}\sinh\Big(\frac{\sqrt{3}\sigma_\infty}{Y}\Big) = \sqrt{3}\sinh\Big(\frac{\sqrt{3}\Sigma_{11}}{|\Sigma_{33} - \Sigma_{11}|}\Big)\ . \tag{11.33}$$

4 Let the total volume of the RVE be

$$V = \Omega + V_{matrix}$$

and

$$\frac{dV}{dt} = \dot{V} = \frac{d\Omega}{dt} + \frac{dV_{matrix}}{dt} = \frac{d\Omega}{dt}$$

because the matrix is incompressible, $\dfrac{dV_{matrix}}{dt} = 0$.

Define the volume fraction of the void as

$$f = \frac{\Omega}{V}\ .$$

Then

$$\begin{aligned} \dot{f} &= \frac{\dot{\Omega}}{V} - \frac{\Omega}{V^2}\dot{V} = \frac{\dot{\Omega}}{V}\Big(\frac{V - \Omega}{V^2}\Big) \\ &= \frac{\dot{\Omega}}{V}(1 - f) = \frac{\dot{\Omega}}{\Omega}f(1 - f) \end{aligned}$$

Finally, we can express the rate of volume fraction as

$$\dot{f} = \frac{\sqrt{3}f(1 - f)}{\dot{\epsilon}_{eq}}\sinh\Big(\frac{\sqrt{3}\Sigma_{11}}{|\Sigma_{33} - \Sigma_{11}|}\Big)$$

11.3 The Gurson model

The significance of McClintock solution it that it links the remote stress, or macro stress, with the void growth rate, and it reveals that in a perfectly plastic RVE, the void growth rate is expenonetially related with the macro-stress. Although, it can be argued that the notion representative volume element is employed in McClintock solution, it does provide new constitutive representation at macro-level.

Not long after the publication of McClintock solution, a young scientist at the time, A. L. Gurson, realized that there is more in the cylindrical void model analyzied by McClinktock. In fact, one can derived the plastic potential at macro-level by homogenized (meaning averaging in space) micro-stress distribution. It was eaxctly what Gurson did his Ph.D. thesis, which has become one of most cited papers in inelastic constitutive modeling and micromechanics.

11.3.1 Gurson's homogenization of cylindrical void in a rigid perfectly-plastic RVE

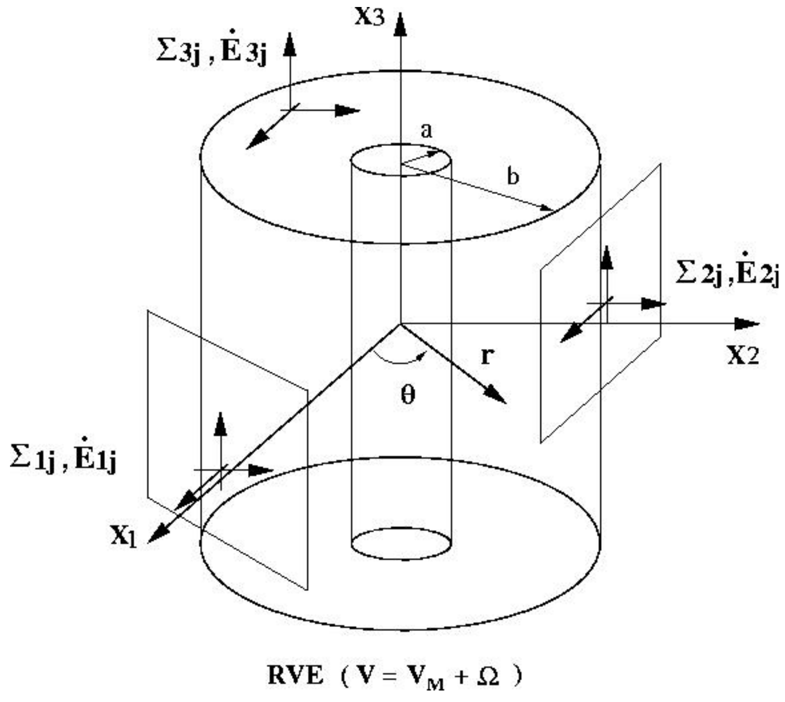


Figure 11.4. A cylindrical void in a rigid-perfectly plastic von Mises RVE

The objective of the Gurson model is to find macroscopic yield potential function in terms of macro-stress and volume fraction of void in an RVE, i.e., we are looking for

$$F(\Sigma_{eq}, \Sigma_m, f) = 0$$

where

$$\Sigma_{eq} = \sqrt{\frac{3}{2}\Sigma'_{ij}\Sigma'_{ij}}, \;\; \Sigma'_{ij} = \Sigma_{ij} - \Sigma_m, \;\; \text{and} \;\; \Sigma_m = \frac{1}{3}\Sigma_{ii}$$

Again, the governing equations in the RVE are,

1 Equilibirum equations:

$$\frac{d\sigma_{rr}}{dr} + \frac{\sigma_{rr} - \sigma_{\theta\theta}}{r} = 0\ .$$

2 von Mises flow rule:

$$s_{ij} = \frac{2}{3}\frac{\sigma_y}{\dot{\epsilon}_{eq}}\dot{\epsilon}_{ij}$$

3 incompressible condition of the matrix:

$$\dot{\epsilon}_{rr} + \dot{\epsilon}_{\theta\theta} + \dot{\epsilon}_{zz} = 0 \ .$$

Consider axisymetric remote (macro-stress) loading,

$$\sigma_{11}\Big|_{\partial V} = \Sigma_{11}, \ \ \sigma_{22}\Big|_{\partial V} = \Sigma_{22}, \ \text{and} \ \ \Sigma_{11} = \Sigma_{22} \tag{11.34}$$

$$\sigma_{33}\Big|_{\partial V} = \Sigma_{33} \tag{11.35}$$

Under axisymmetric loading condition,

$$\begin{aligned} \Sigma_{eq} &= \sqrt{\frac{1}{2}\Big[(\Sigma_{11}-\Sigma_{22})^2 + (\Sigma_{33}-\Sigma_{11})^2 + (\Sigma_{33}-\Sigma_{22})^2\Big]} \\ &= |\Sigma_{33} - \Sigma_{11}|, \end{aligned} \tag{11.36}$$

$$\begin{aligned} \Sigma_m &= \frac{1}{3}(\Sigma_{11}+\Sigma_{22}+\Sigma_{33}) = \frac{1}{3}\Big(\Sigma_{\alpha\alpha} + \Sigma_{33}\Big) \\ &= \Sigma_{11} + \frac{1}{3}\Big(\Sigma_{33} - \Sigma_{11}\Big) = \frac{1}{2}\Sigma_{\alpha\alpha} + \frac{1}{3}\Sigma_{eq} \end{aligned} \tag{11.37}$$

where $\Sigma_{\alpha\alpha} = \Sigma_{11} + \Sigma_{22} = 2\Sigma_{11}$, or $\Sigma_{11} = \Sigma_{22} = \dfrac{1}{2}\Sigma_{\alpha\alpha}$. Therefore, we are essentially looking for the yeilding effects due to Σ_{11} and $\Sigma_{33} - \Sigma_{11}$.

Consider the following axisymmetric kinemetic pattern,

$$\dot{u}_r = \dot{u}(r), \ \ \dot{u}_z(z) = \dot{E}_{33} z \ .$$

Strain rate components are

$$\dot{\epsilon}_{rr} = \frac{d\dot{u}}{dr}, \ \ \dot{\epsilon}_{\theta\theta} = \frac{\dot{u}}{r}, \ \ \dot{\epsilon}_{zz} = \dot{E}_{33} \ .$$

Since the matrix is incompressible,

$$\dot{\epsilon}_{rr} + \dot{\epsilon}_{\theta\theta} + \dot{\epsilon}_{zz} = \frac{d\dot{u}}{dr} + \frac{\dot{u}}{r} + \dot{E}_{33} = 0,$$

one has

$$\int d(r\dot{u}) = -\int \dot{E}_{33} r dr \ \ \Rightarrow \ \ \dot{u}(r) = -\frac{\dot{E}_{33}}{2} r + \frac{A}{r}$$

where A is an unknown constant.

Subsequently,

$$\dot{\epsilon}_{rr} = \frac{d\dot{u}}{dr} = -\frac{\dot{E}_{33}}{2} - \frac{A}{r^2} \tag{11.38}$$

$$\dot{\epsilon}_{\theta\theta} = \frac{\dot{u}}{r} = -\frac{\dot{E}_{33}}{2} + \frac{A}{r^2} \tag{11.39}$$

In fact, the constant A has a clear physical interpretation. Consider a cylindrical void with finite height,

$$\Omega = \pi a^2 H$$

The void growth rate and relative void growth rate are

$$\frac{d\Omega}{dt} = 2\pi a\dot{a}H + \pi a^2 \dot{H} \tag{11.40}$$

$$\frac{\dot{\Omega}}{\Omega} = 2\frac{\dot{a}}{a} + \frac{\dot{H}}{H} \tag{11.41}$$

Since,

$$\frac{\dot{a}}{a} = \dot{\epsilon}_{rr}(a) \ \text{ and } \ \frac{\dot{H}}{H} = \dot{E}_{33},$$

one may find that

$$\frac{\dot{\Omega}}{\Omega} = 2\Big(-\frac{\dot{E}_{33}}{2} - \frac{A}{a^2}\Big) + \dot{E}_{33} = -\frac{2A}{a^2}$$

which leads to

$$A = -\frac{a^2}{2}\frac{\dot{\Omega}}{\Omega}\,. \tag{11.42}$$

That is: A is proportional to the relative void growth rate.

Since the matrix is a rigid-perfectly plastic von-Mises material, it obeys the following flow rule,

$$s_{ij} = \frac{2}{3}\frac{\sigma_y}{\dot{\epsilon}_{eq}}\dot{\epsilon}_{ij}$$

where the effective strain rate can be explicitely expressed as

$$\begin{aligned}
\dot{\epsilon}_{eq} &= \Big(\frac{2}{3}\dot{\epsilon}_{ij}\dot{\epsilon}_{ij}\Big)^{1/2} = \Big[\frac{2}{3}\Big(\dot{\epsilon}^2_{rr} + \dot{\epsilon}^2_{\theta\theta} + \dot{\epsilon}^2_{zz}\Big]^{1/2} \\
&= \left[\frac{2}{3}\Big[\Big(\frac{\dot{E}_{33}}{2} + \frac{A}{r^2}\Big)^2 + \Big(\frac{\dot{E}_{33}}{2} - \frac{A}{r^2}\Big)^2 + \dot{E}^2_{33}\Big]\right] \\
&= \Big(\dot{E}^2_{33} + \frac{4}{3}\frac{A^2}{r^4}\Big)^{1/2} = \dot{E}_{33}\Big(1 + \alpha^2\Big(\frac{a}{r}\Big)^4\Big)^{1/2}
\end{aligned} \tag{11.43}$$

where the parameter, α, is defined as

$$\alpha := \frac{2}{\sqrt{3}}\frac{|A|}{\dot{E}^2_{33}a^2} = \Big|\,\frac{\dot{\Omega}}{\Omega}\,\Big|\,\frac{1}{\sqrt{3}\dot{E}_{33}} \tag{11.44}$$

Therefore, we can write,

$$
\begin{aligned}
s_{rr} &= \frac{2}{3}\frac{\sigma_y}{\dot{E}_{33}\Big(1+\alpha^2\Big(\frac{a}{r}\Big)^4\Big)^{1/2}}\dot{\epsilon}_{rr} = \frac{2}{3}\frac{\sigma_y}{\dot{E}_{33}\Big(1+\alpha^2\Big(\frac{a}{r}\Big)^4\Big)^{1/2}}\Big(-\frac{\dot{E}_{33}}{2}-\frac{A}{r^2}\Big) \\
s_{\theta\theta} &= \frac{2}{3}\frac{\sigma_y}{\dot{E}_{33}\Big(1+\alpha^2\Big(\frac{a}{r}\Big)^4\Big)^{1/2}}\dot{\epsilon}_{\theta\theta} = \frac{2}{3}\frac{\sigma_y}{\dot{E}_{33}\Big(1+\alpha^2\Big(\frac{a}{r}\Big)^4\Big)^{1/2}}\Big(+\frac{\dot{E}_{33}}{2}-\frac{A}{r^2}\Big) \\
s_{zz} &= \frac{2}{3}\frac{\sigma_y}{\dot{E}_{33}\Big(1+\alpha^2\Big(\frac{a}{r}\Big)^4\Big)^{1/2}}\dot{E}_{33} = \frac{2}{3}\frac{\sigma_y}{\Big(1+\alpha^2\Big(\frac{a}{r}\Big)^4\Big)^{1/2}}
\end{aligned}
$$

We can then find that

$$
\begin{aligned}
s_{\theta\theta}-s_{rr} &= \frac{2}{3}\frac{\sigma_y}{\dot{E}_{33}(1+\alpha^2(a/r)^4)^{1/2}}\Big(\frac{A}{r^2}+\frac{A}{r^2}\Big) \\
&= \frac{4}{3}\frac{\sigma_y}{\dot{E}_{33}(1+\alpha^2(a/r)^4)^{1/2}}\Big(\frac{A}{r^2}\Big) \\
&= \sigma_{\theta\theta}-\sigma_{rr}
\end{aligned}
$$

and

$$
\begin{aligned}
s_{zz}-\frac{1}{2}(s_{rr}+s_{\theta\theta}) &= \frac{2}{3}\frac{\sigma_y}{(1+\alpha^2(a/r)^4)^{1/2}} \\
&\quad -\frac{1}{2}\frac{2}{3}\frac{\sigma_y}{\dot{E}_{33}(1+\alpha^2(a/r)^4)^{1/2}}(-\dot{E}_{33}) \\
&= \frac{\sigma_y}{(1+\alpha^2(a/r)^4)^{1/2}} \\
&= \sigma_{zz}-\frac{1}{2}(\sigma_{rr}+\sigma_{\theta\theta})
\end{aligned}
$$

To this end, we are in a position to link the macro-stresses, Σ_{11}, $\Sigma_{33}-\Sigma_{11}$, and void volume fraction, f, together in a macro yield potential.

We first link Σ_{11} and $|\Sigma_{33}-\Sigma_{11}|$ with remote strain rate, $\dot{E}_{ij}$.

Consider the traction boundary conditions on the surface of the void and the surface of the RVE,

$$
\sigma_{rr}(a)=0, \text{ and } \sigma_{rr}(b)=\frac{1}{2}\Sigma_{\alpha\alpha}=\Sigma_{11}
$$

note that $\Sigma_{rr}(b)=\Sigma_{\theta\theta}(b)=\frac{1}{2}\Sigma_{\alpha\alpha}$.

1. Integrating equilibrium equation along the radius direction yields,

$$
\Sigma_{11}=\sigma_{rr}(b)-\sigma_{rr}(a)=\int_a^b\frac{d\sigma_{rr}}{dr}dr=\int_a^b\frac{\sigma_{\theta\theta}-\sigma_{rr}}{r}dr
$$

Since,

$$\frac{\sigma_{\theta\theta} - \sigma_{rr}}{r} = \frac{s_{\theta\theta} - s_{rr}}{r} = \frac{4}{3} \frac{\sigma_y}{\dot{E}_{33}(1 + \alpha^2 (a/r)^4)^{1/2}} \frac{A}{r^3}$$

we have

$$\Sigma_{11} = \frac{4}{3}\sigma_y \int_a^b \frac{A}{\dot{E}_{33}(1 + \alpha^2 (a/r)^4)^{1/2}} \frac{dr}{r^3} \tag{11.45}$$

2. Consider the fact that $\sigma_{11} + \sigma_{22} = \sigma_{rr} + \sigma_{\theta\theta}$, and $\Sigma_{11} = \Sigma_{22} = \frac{1}{2}\Sigma_{\alpha\alpha}$,

$$\begin{aligned}
\Sigma_{33} - \Sigma_{11} &= \Sigma_{33} - \frac{1}{2}(\Sigma_{11} + \Sigma_{22}) = \frac{1}{V}\int_V \Big(\sigma_{zz} - \frac{1}{2}(\sigma_{xx} + \sigma_{yy})\Big) dV \\
&= \frac{1}{V}\int_V \Big(\sigma_{zz} - \frac{1}{2}(\sigma_{rr} + \sigma_{\theta\theta})\Big) dV \\
&= \frac{1}{V}\int_V \Big(s_{zz} - \frac{1}{2}(s_{rr} + s_{\theta\theta})\Big) dV \\
&= \frac{1}{V}\int_{V_M} \Big(s_{zz} - \frac{1}{2}(s_{rr} + s_{\theta\theta})\Big) dV
\end{aligned}$$

Recall that

$$s_{zz} - \frac{1}{2}(s_{rr} + s_{\theta\theta}) = \frac{\sigma_y}{(1 + \alpha^2 \Big(\frac{a}{r}\Big)^4)^{1/2}} \; ,$$

and $dV = r dr d\theta dz$. We have

$$\begin{aligned}
\Sigma_{33} - \Sigma_{11} &= \frac{2\pi H}{\pi b^2 H}\int_a^b \frac{\sigma_y}{\Big(1 + \alpha^2 \Big(\frac{a}{r}\Big)^4\Big)^{1/2}} r dr \\
&= \frac{2\sigma_y}{b^2}\int_a^b \frac{r dr}{\Big(1 + \alpha^2 \Big(\frac{a}{r}\Big)^4\Big)^{1/2}}
\end{aligned} \tag{11.46}$$

Make change of variable,

$$x = \alpha\Big(\frac{a}{r}\Big)^2 : \;\; x \rightarrow [\alpha, f\alpha], \;\; \text{when} \;\; r \rightarrow [a, b] \, .$$

where $f = \dfrac{a^2}{b^2} = \dfrac{\Omega}{V}$.

Therefore,

$$dx = -2\alpha \frac{a^2}{r^3} dr = -\frac{4A}{\sqrt{3}\dot{E}_{33} r^3} dr \;\; \Leftarrow \;\; \alpha = \frac{2A}{\sqrt{3}\dot{E}_{33} a^2}$$

and

$$\frac{Adr}{\dot{E}_{33}r^3} = -\frac{\sqrt{3}}{4}dx, \tag{11.47}$$

$$rdr = -\frac{r^4}{2a^2\alpha}dx = -\frac{a^2\alpha}{2}\frac{dx}{x^2}, \quad \Leftarrow \quad x = \alpha\frac{a^2}{r^2} \tag{11.48}$$

Reconsider (11.45) and $\dfrac{Adr}{\dot{E}_{33}r^3} = -\dfrac{\sqrt{3}}{4}dx$,

$$\begin{aligned}\Sigma_{11} &= \frac{1}{2}\Sigma_{\alpha\alpha} = \frac{4}{3}\sigma_y \int_a^b \frac{1}{(1+x^2)^{1/2}}\frac{Adr}{\dot{E}_{33}r^3} \\ &= -\Big(\frac{4}{3}\sigma_y\Big)\frac{\sqrt{3}}{4}\int_\alpha^{f\alpha}\frac{dx}{\sqrt{1+x^2}}\end{aligned}$$

Thereby,

$$\frac{\Sigma_{\alpha\alpha}}{2} = \frac{\sigma_y}{\sqrt{3}}\int_{f\alpha}^{\alpha}\frac{dx}{\sqrt{1+x^2}} \tag{11.49}$$

We then find that the in-plane hydrostatic stress can be written as

$$\boxed{\frac{\sqrt{3}}{2}\frac{\Sigma_{\alpha\alpha}}{\sigma_y} = \log\Big[\frac{\alpha+\sqrt{1+\alpha^2}}{f\alpha+\sqrt{1+(f\alpha)^2}}\Big]} \tag{11.50}$$

Reconsider Eq. (11.46) and $rdr = -\dfrac{\alpha a^2}{2}\dfrac{dx}{x^2}$,

$$\begin{aligned}\Sigma_{33} - \Sigma_{11} &= \frac{2\sigma_y}{b^2}\int_a^b \frac{rdr}{(1+x^2)^{1/2}} \\ &= \Big(\frac{2\sigma_y}{b^2}\Big)\Big(-\frac{\alpha a^2}{2}\Big)\int_\alpha^{f\alpha}\frac{dx}{x^2\sqrt{1+x^2}} \\ &= f\alpha\sigma_y\int_{f\alpha}^{\alpha}\frac{dx}{x^2\sqrt{1+x^2}}\end{aligned}$$

Carrying the integration, we have

$$\Sigma_{33} - \Sigma_{11} = \sigma_y\Big[\sqrt{1+\alpha^2 f^2} - f\sqrt{1+\alpha^2}\Big]$$

We can then link the deriatoric macro-stress with macro-strain rate and void volume fraction,

$$\boxed{\frac{\Sigma_{eq}}{\sigma_y} = \sqrt{1+f^2\alpha^2} - f\sqrt{1+\alpha^2}} \tag{11.51}$$

Denote that

$$\begin{aligned} A_1 &= \frac{\sqrt{3}}{2}\frac{\Sigma_{\alpha\alpha}}{\sigma_y} \\ A_2 &= \frac{\Sigma_{eq}}{\sigma_y} \\ A_3 &= \alpha + \sqrt{1+\alpha^2} \\ A_4 &= f\alpha + \sqrt{1+f^2\alpha^2} \end{aligned}$$

Then results (11.50) and (11.51) can be rewritten as

$$A_1 = \log\frac{A_3}{A_4} \tag{11.52}$$

$$A_2 = A_4 - fA_3 \tag{11.53}$$

We want to connect A_1 and A_2 by elminating A_3 and A_4.
Rewrite (11.52) and (11.53) as

$$\exp A_1 = \frac{A_3}{A_4} \tag{11.54}$$

$$A_4 = A_2 + fA_3 \tag{11.55}$$

Substituting (11.55) into (11.54) leads to an equation of A_1, A_2, and A_3,

$$\exp A_1 = \frac{A_3}{A_2 + fA_3}$$

which expresses A_3 in terms of A_1 and A_2,

$$\boxed{A_3 = \frac{A_2\exp(A_1)}{1 - f\exp(A_1)}} \tag{11.56}$$

Substituting (11.57) back into (11.55) yields an equation among A_1, A_2, and A_4,

$$A_4 = A_2 + fA_3 = \frac{(1 - f\exp(A_1)A_2 + fA_2\exp(A_1)}{1 - f\exp(A_1)}$$

Solving this equation yields

$$\boxed{A_4 = \frac{A_2}{1 - f\exp(A_1)}} \tag{11.57}$$

Consider the identities,

$$\begin{aligned} A_3^2 &= (\alpha + \sqrt{1+\alpha^2})^2 = \alpha^2 + 2\alpha\sqrt{1+\alpha^2} + (1+\alpha^2) \\ &= 2\alpha(\alpha + \sqrt{1+\alpha^2}) + 1 = 2\alpha A_3 + 1 \end{aligned}$$

and

$$A_4^2 = A_3^2(f\alpha) = 2f\alpha(f\alpha + \sqrt{1+f^2\alpha^2}) + 1 = 2f\alpha A_4 + 1 \tag{11.58}$$

We may find that

$$2\alpha = \frac{A_3^2 - 1}{A_3} \tag{11.59}$$

$$2f\alpha = \frac{A_4^2 - 1}{A_4} \tag{11.60}$$

Combining (11.59) and (11.60), we may find that the following expression,

$$\boxed{2\alpha = \frac{A_3^2 - 1}{A_3} = \frac{A_4^2 - 1}{fA_4}} \tag{11.61}$$

Substituting

$$\begin{aligned} A_3 &= \frac{A_2 \exp(A_1)}{1 - f\exp(A_1)} \\ A_4 &= \frac{A_2}{1 - f\exp(A_1)} \end{aligned}$$

into (11.61), we obtain the following identity,

$$\frac{A_3^2 - 1}{A_3} = \frac{A_4^2 - 1}{fA_4} \Rightarrow \frac{A_2^2 \exp(2A_1) - (1 - f\exp(A_1))^2}{A_2^2 - (1 - f\exp(A_1))^2} = \frac{A_2 \exp(A_1)}{fA_2}$$

Rewrite the above equation,

$$\begin{aligned} & fA_2^2 \exp(2A_1) - f(1 - f\exp(A_1))^2 \\ = \;& A_2^2 \exp(A_1) - \exp(A_1)(1 - f\exp(A_1))^2 \\ \Rightarrow \;& A_2^2 \exp(A_1)(1 - f\exp(A_1)) = (1 - f\exp(A_1))^2(\exp(A_1) - f) \end{aligned}$$

which leads to

$$\begin{aligned} A_2^2 &= (1 - f\exp(A_1))(1 - f\exp(A_1)) \\ &= 1 + f^2 - f\Big[\exp(A_1) + \exp(-A_1)\Big] \\ &= 1 + f^2 - 2f\cosh A_1 \end{aligned}$$

We finally link A_1 and A_2 in a single equation.

Substituting the expressions of A_1 and A_2 into the above equation, we have the desired result,

$$\boxed{F(\Sigma_{eq}, \Sigma_{\alpha\alpha}, f) = \frac{\Sigma_{eq}^2}{\sigma_y^2} + 2f\cosh\Big(\frac{\sqrt{3}}{2}\frac{\Sigma_{\alpha\alpha}}{\sigma_y}\Big) - (1+f^2) = 0\ .} \tag{11.62}$$

On the other hand, if we rewrite (11.50) as,

$$\begin{aligned}
\frac{\sqrt{3}}{2}\frac{\Sigma_{\alpha\alpha}}{\sigma_y} &= \log\Big[\frac{\alpha + \sqrt{1+\alpha^2}}{f\alpha + \sqrt{1+f^2\alpha^2}}\Big] = \mathrm{Arcsinh}(\alpha) - \mathrm{Arcsinh}(f\alpha) \\
&= \mathrm{Arcsinh}(\alpha\sqrt{1+\alpha^2 f^2} - f\alpha\sqrt{1+\alpha^2})
\end{aligned}$$

Therefore,

$$\sinh\Big(\frac{\sqrt{3}}{2}\frac{\Sigma_{\alpha\alpha}}{\sigma_y}\Big) = \alpha(\sqrt{1+f^2\alpha^2} - f\sqrt{1+\alpha^2}) \tag{11.63}$$

Consider

$$\alpha = \frac{\dot{\Omega}}{\Omega}\frac{1}{\sqrt{3}\dot{E}_{33}} \tag{11.64}$$

$$\frac{\Sigma_{eq}}{\sigma_y} = \sqrt{1+f^2\alpha^2} - f\sqrt{1+\alpha^2} \tag{11.65}$$

Eq. (11.63) can be rewritten as

$$\sinh\Big(\frac{\sqrt{3}}{2}\frac{\Sigma_{\alpha\alpha}}{\sigma_y}\Big) = \Big|\ \frac{\dot{\Omega}}{\Omega}\ \Big|\ \frac{1}{\sqrt{3}\dot{E}_{33}}\frac{\Sigma_{eq}}{\sigma_y}$$

or

$$\Big|\ \frac{\dot{\Omega}}{\Omega}\ \Big| = \sqrt{3}\dot{E}_{33}\Big(\frac{\sigma_y}{\Sigma_{eq}}\Big)\sinh\Big(\frac{\sqrt{3}}{2}\frac{\Sigma_{\alpha\alpha}}{\sigma_y}\Big)$$

Considering the fact $\dot{f} = \Big|\ \frac{\dot{\Omega}}{\Omega}\ \Big|\ f(1-f)$, we recover the McClintock solution,

$$\boxed{\dot{f} = \sqrt{3}f(1-f)\dot{E}_{33}\Big(\frac{\sigma_y}{\Sigma_{eq}}\Big)\sinh\Big(\frac{\sqrt{3}}{2}\frac{\Sigma_{\alpha\alpha}}{\sigma_y}\Big)} \tag{11.66}$$

11.3.2 Gurson-Tvergaard-Needleman model

$$\Phi = \Bigg(\frac{\sigma_{eq}}{\sigma_y} \tag{11.67}$$

11.4 Exercise

References

Ablowitz, M. and Fokas, A. S. (1997), *Complex Variables: Introduction and Applications*, Cambridge University Press.

Cottrell, A. H. (1953), *Dislocations and Plastic Flow in Crystals*, Oxford University Press.

Eshelby, J. D. (1957) "The determination of the elastic field of an ellipsoidal inclusion and related problems," *Proceedings of Royal Society of London*, **A 241**, pp. 376-396.

Eshelby, J. D. (1959), "The elastic field outside an ellipsoidal inclusion," *Proceedings of Royal Society of London*, **A 252**, pp. 561-569.

Hashin, Z. and Shtrikman, S. [1961], "Note on a variational approach to the theory of composite elastic materials," *The Franklin Institute Laboratories*, **271**, pp. 336-341.

Hashin, Z. and Shtrikman, S. [1962], "On some variational principlrd in anisotropic and nonhomogeneous elasticity," *Journal of Mechanics and Physics of Solids*, **10**, pp. 335-342.

Hashin, Z. and Shtrikman, S. [1962], "The elastic moduli of heterogeneous materials," *Journal of Applied Mechanics*, pp. 143-148.

Hashin, Z. (1965), "On elastic behaviour of fibre reinforced materials of arbitrary transverse phase geometry," *Journal of Mechanics and Physics of Solids*, **13**, pp. 119-134.

Hirth, J. P. and Lothe, J. (1982). *Theory of Dislocations*, Krieger

Love, A. E. H. (1926). *A Treatise on the Mathematical Theory of Elasticity*, Fourth Edition, Dover Publication, New York.

Malvern, L. E. (1969). *Introduction to the Mechanics of a Continuous Mdedium*, Prentice-Hall.

Mura, T. (1987). *Micromechanics of Defects in Solids*, Second, Revised Edition, Kluwer Academic Pub. Dordrecht/Boston/London

Nemat-Nasser, S. and Hori, M. (1999), *Micromechanics: overall properties of heterogeneous materials*, Elsevier, Amsterdam-Lausanne-New York

Sneddon, I. N. (1950), *Fourier Transforms*, Dover Publication, Inc. New York.

Sokolnikoff, I. S. (1956). *Mathematical Theory of Elasticity*, 2n edition, McGraw-Hill

Talbot, D.R.S. and Willis, J.R. [1985], “Variational principles for inhomogeneous non-linear media,” *IMA Journal of Applied Mathematics*, **35**, pp. 39-54.

Talbot, D.R.S. and Willis, J.R. [1998], “Upper and lower bounds for the overall response of an elastoplastic composite,” *Mechanics of Materials*, **28**, pp. 1-8.

Timoshenko, S. and Goodier, J. N. (1951). *Theory of Elasticity*, 2nd ed. McGraw-Hill

Titchmarsh, E. C. (1948). *Introduction to the Theory of Fourier Integrals*, Oxford.

Zhdanov, M. (1988). *Integral Transforms in Geophysics*, Springer-Verlag, Berlin / Heidelberg / New York