Earthquake Source Mechanics

Instructors: Harsha S. Bhat, Pierre Dublanchet, Carolina Giorgetti

Fridays 14:00–17:00, Sept 2026–Jan 2027 (Room E409)

This document is a compilation of handwritten notes, PDF's etc prepared by myself, Harsha S. Bhat, over the years. Claude was used to transcribe, collate and summarize some aspects of this document. Thus it is quite likely that errors have crept in and I am in the process of cleaning it up. I am solely responsible for these errors.

Before we dive into the mathematical machinery of continuum and earthquake source mechanics — the index notation, the tensors, and the rest of the drama that comes with them — I strongly encourage you to first read two short historical sketches. Neither requires any of the mathematics developed in this course; they exist to give you a sense of where these ideas came from, who developed them, and why, before we spend the rest of today building the tools themselves.

- A brief history of solid mechanics (pp 1–15): PDF
- Une Histoire de la compréhension des Tremblements de Terre: PDF
- Wikipedia Biography of Emmy Noether: PDF

Twenty minutes with the first two will do more for your intuition than the three hours of equations will.

Contents

Syllabus

This course provides an introduction to the physical and mechanical processes governing earthquake generation and seismic radiation. Starting from continuum mechanics and elastodynamics, students develop the mathematical tools needed to describe stress, strain, elastic waves, Green's functions, and seismic sources. The course then examines moment tensors, radiation patterns, seismic moment, source spectra, and kinematic source models. Building on these foundations, it introduces fracture mechanics, friction laws, laboratory observations, and the dynamics of earthquake rupture, including supershear and slow-slip phenomena. Particular attention is given to the energy budget of earthquakes and to the connection between individual earthquake sources and statistical patterns of seismicity. The course combines theoretical developments, classical models, observational evidence, and laboratory experiments to provide a unified framework for understanding the earthquake source process.

Lecture 1: Continuum Mechanics & Math Primer

- Why continuum mechanics matters for understanding earthquake sources
- Vectors & tensors — index notation, coordinate transforms, isotropic tensors
- Stress — traction vector, Cauchy's stress tensor, symmetry, principal stresses
- Strain — displacement field, infinitesimal strain tensor, compatibility
- Constitutive law — Hooke's law, Lamé parameters, a first look at Navier's equation

Lecture 2: Elastodynamics I: The Wave Equation

- Quick recap of Navier's equation from last time
- Getting to the elastic wave equation from momentum balance and Hooke's law
- P & S waves — Helmholtz decomposition, wave speeds α and β
- Plane waves — solutions, polarization, particle motion
- Energy density & flux in elastic waves, a first look at geometrical spreading

Lecture 3: Elastodynamics II: Green's Functions & Representation Theorem

- Quick recap
- The point-force Green's function in a full space — near, intermediate, and far-field terms
- Betti's reciprocal theorem
- Deriving the representation theorem, and the fault as a displacement discontinuity
- The moment tensor and the equivalent body-force picture, the double-couple

Lecture 4: Far-Field Radiation & the Seismic Source

- Quick recap of the moment tensor from last time
- Far-field P/S radiation patterns from a double couple
- Seismic moment M_0, and how it relates to magnitude scales
- The source time function and moment-rate function
- The ω^2 model — corner frequency, estimating static stress drop

Lecture 5: Kinematic Source Models

- Quick recap
- Dislocation theory — Steketee, Burridge–Knopoff, static dislocation solutions
- The Haskell model — a propagating dislocation source, rise time, rupture velocity, unilateral vs. bilateral rupture
- Finite-fault parameterization — slip distributions, multiple time-window representations
- A first look at inversion — the forward problem, non-uniqueness, why source inversion is ill-posed

Lecture 6: Fracture Mechanics Fundamentals

- Quick recap, and why we're moving from kinematics to dynamics
- The Griffith energy balance and energy release rate G
- The crack-tip stress field and stress intensity factor K
- Fracture energy G_c and the process-zone idea
- Static vs. dynamic fracture criteria, the Griffith criterion

Lecture 7: Friction Laws

- Quick recap
- Coulomb friction — static/kinetic friction, Byerlee's law
- Slip-weakening friction — Ida (1972), Palmer–Rice, the critical slip distance D_c, the breakdown zone
- Rate-and-state friction — Dieterich, Ruina, the $(a - b)$ parameter, state evolution laws
- Stability — a linear stability analysis, nucleation size

Lecture 8: Laboratory Friction

Lecture 9: The Dynamic Rupture Problem

- Quick recap
- Setting up the elastodynamic crack problem and the boundary integral idea
- The Kostrov self-similar rupture solution
- Crack-like vs. pulse-like rupture, Heaton's self-healing pulse
- The energy budget — radiated energy, fracture energy, frictional heat, how it all partitions

Lecture 10: Supershear Rupture

- Quick recap
- Sub-Rayleigh vs. supershear regimes, the limiting speeds — Rayleigh, S, P
- The Burridge–Andrews mechanism — daughter-crack nucleation
- The critical stress ratio S and Mach-cone radiation
- What we've observed

Lecture 11: Slow-to-Fast Slip Spectrum

- Quick recap
- What we observe — slow-slip events, tremor, LFEs, VLFEs
- Rate-and-state in the conditionally stable regime
- Mechanisms — near-neutral stability, fluid effects, transition zones
- Pulling it together: the slip spectrum, from creep to earthquake

Lecture 12: Energy Budget & Statistical Seismicity

- Why we're zooming out, from single events to populations of earthquakes
- Pulling the energy budget together — E_s, E_f, E_h, E_d, apparent stress, radiation efficiency
- How E_s is actually estimated from seismograms — energy-moment scaling, energy magnitude, and the uncertainties involved
- The Gutenberg–Richter frequency-magnitude relation and the b-value
- Omori–Utsu aftershock decay, productivity, a nod to ETAS, and catalog completeness

Lecture 13: Laboratory Earthquakes

- Photoelasticity experiments
- Triaxial saw-cut experiments

1 Vectors & Tensors: Index Notation

1.1 The Summation Convention

Suppose $\mathbf{x}$ and $\mathbf{y}$ are vectors, and A and B are matrices. Common combinations, in components:

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n=3} x_i y_i, \qquad [A\mathbf{x}]_i = \sum_{j=1}^{n} A_{ij} x_j, \qquad [AB]_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}.$$

Notice the curious thing: *every sum goes with an index that is repeated twice; non-repeated indices are not summed.* We adopt the **Einstein summation convention**: the summation symbol $\sum$ is dropped, and a repeated index implies summation. (An index may not appear more than twice on one side of an equality.) Using the convention,

$$\mathbf{x} \cdot \mathbf{y} = x_i y_i, \qquad [A\mathbf{x}]_i = A_{ij} x_j, \qquad [AB]_{ij} = A_{ik} B_{kj}.$$

Summary. If an index occurs once, it must occur once in every term of the equation, and the equation holds for each value of that index. If an index appears twice, it is summed over all values — it is a *dummy index* and can be renamed at will. If an index appears three or more times in a single term, the expression is wrong.

1.2 The Kronecker Delta δ_{ij}

We define

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases},$$

i.e. the components of the identity tensor I. Since $I\mathbf{y} = \mathbf{y}$, we have the **sifting property**

$$\delta_{ij} y_j = y_i.$$

In words: if one index of δ_{ij} is summed, the effect is to swap it for the other index.

Example: Fix $i = 1$ and write out the sum over $j = 1, 2, 3$ explicitly (this is exactly what the summation convention is hiding):

$$\delta_{1j} y_j = \delta_{11} y_1 + \delta_{12} y_2 + \delta_{13} y_3 = (1)y_1 + (0)y_2 + (0)y_3 = y_1.$$

Only the $j = 1$ term survives, because δ_{1j} is zero unless $j = 1$. The same happens for $i = 2$ and $i = 3$: in each case, out of the three terms in the sum, only the one where j matches i survives, and it contributes y_i exactly. This is why summing over one index of δ_{ij} has the net effect of *renaming* the other index from j to i — it is a bookkeeping device, not a computation.

1.3 The Permutation Symbol ϵ_{ijk}

To express the cross product

$$\mathbf{x} \times \mathbf{y} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{vmatrix} \tag{1}$$

in index notation we use the **permutation symbol** (also called the alternating symbol or Levi-Civita symbol),

$$\epsilon_{ijk} = \begin{cases} 1 & (ijk) \text{ an even permutation of } (1,2,3) \\ -1 & (ijk) \text{ an odd permutation of } (1,2,3) \\ 0 & \text{otherwise (repeated index)} \end{cases}$$

so that

$$[\mathbf{x} \times \mathbf{y}]_i = \epsilon_{ijk} x_j y_k.$$

Example: Take $i = 1$ and write out the sum over $j, k = 1, 2, 3$ (nine terms in principle, since both j and k range over 1,2,3):

$$[\mathbf{x} \times \mathbf{y}]_1 = \epsilon_{1jk} x_j y_k = \sum_{j=1}^{3} \sum_{k=1}^{3} \epsilon_{1jk} x_j y_k.$$

Of these nine terms, $\epsilon_{1jk} = 0$ whenever j or k equals 1 (repeated index) or when $j = k$ (repeated index again). The only surviving terms are $(j, k) = (2, 3)$ and $(j, k) = (3, 2)$:

$$[\mathbf{x} \times \mathbf{y}]_1 = \epsilon_{123} x_2 y_3 + \epsilon_{132} x_3 y_2 = (1) x_2 y_3 + (-1) x_3 y_2 = x_2 y_3 - x_3 y_2,$$

which is exactly the first component of $\mathbf{x} \times \mathbf{y}$ from the determinant formula above. The same bookkeeping gives components 2 and 3.

Symmetries:

$$\epsilon_{ijk} = \epsilon_{kij} = \epsilon_{jki} \quad \text{(cyclic)}, \qquad = -\epsilon_{jik} = -\epsilon_{ikj} = -\epsilon_{kji} \quad \text{(pairwise swap)}.$$

(These all follow directly from the definition: swapping any two of the three letters in (ijk) flips the parity of the permutation, hence flips the sign; swapping *two pairs* — a cyclic shift — restores the original sign.)

1.4 The ϵ–δ Identity

An important *contracted epsilon identity*:

$$\underbrace{\epsilon_{ijk} \epsilon_{klm}}_{\text{sum over } k} = \delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}.$$

Worked Example 1 (Vector triple product)

$$\begin{aligned} [\mathbf{a} \times (\mathbf{b} \times \mathbf{c})]_i &= \epsilon_{ijk} a_j (\mathbf{b} \times \mathbf{c})_k = \epsilon_{ijk} a_j \epsilon_{klm} b_l c_m = (\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl}) a_j b_l c_m \\ &= b_i a_m c_m - c_i a_j b_j = [(\mathbf{a} \cdot \mathbf{c}) \mathbf{b} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c}]_i. \end{aligned}$$

1.5 Symmetry and Anti-symmetry

We expect $\mathbf{a} \times \mathbf{a} = \mathbf{0}$. In index notation, $[\mathbf{a} \times \mathbf{a}]_i = \epsilon_{ijk} a_j a_k$. The term $a_j a_k$ is *symmetric* in (j, k) while ϵ_{ijk} is *anti-symmetric* in (j, k):

$$\epsilon_{ijk} a_j a_k = \tfrac{1}{2} (\epsilon_{ijk} a_j a_k + \epsilon_{ijk} a_j a_k) = \tfrac{1}{2} (\epsilon_{ijk} a_j a_k + \epsilon_{ikj} a_k a_j) = \tfrac{1}{2} (\epsilon_{ijk} a_j a_k - \epsilon_{ijk} a_j a_k) = 0.$$

Why the middle step is allowed. We used $\epsilon_{ijk} a_j a_k = \epsilon_{ikj} a_k a_j$ — this is simply relabeling the dummy indices $j \leftrightarrow k$ (allowed, since both are summed over the same range and appear nowhere else in the term), *not* an assumption. Then $\epsilon_{ikj} = -\epsilon_{ijk}$ (pairwise swap of the last two letters) and $a_k a_j = a_j a_k$ (ordinary multiplication commutes), so $\epsilon_{ikj} a_k a_j = -\epsilon_{ijk} a_j a_k$. Substituting this back is what produces the "$+X = -X$" step, forcing $X = 0$.

The general trick: the contraction of a symmetric object with an anti-symmetric object over the same pair of indices always vanishes. This will be used repeatedly — in particular, it is exactly the reasoning behind $\mathrm{curl}(\mathrm{grad}\, \phi) = 0$ below, and it reappears later when we symmetrize the elastic modulus tensor C_{ijkl} (Section 4.3). Symmetric objects include $a_i a_j$, δ_{ij}, and $\partial^2 \phi / \partial x_i \partial x_j$.

1.6 Vector Derivatives

With $\nabla = [\partial/\partial x_1, \partial/\partial x_2, \partial/\partial x_3]$, for scalar $\phi(\mathbf{x})$ and vector $\mathbf{u}(\mathbf{x})$:

$$[\nabla\phi]_i = \frac{\partial\phi}{\partial x_i}, \qquad \nabla\cdot\mathbf{u} = \frac{\partial u_i}{\partial x_i}, \qquad [\nabla\times\mathbf{u}]_i = \epsilon_{ijk}\frac{\partial u_k}{\partial x_j}.$$

Worked Example 2 (Product rules)

$$\nabla\cdot(\phi\mathbf{u}) = \frac{\partial(\phi u_i)}{\partial x_i} = \frac{\partial\phi}{\partial x_i}u_i + \phi\frac{\partial u_i}{\partial x_i} = \mathbf{u}\cdot\nabla\phi + \phi\nabla\cdot\mathbf{u},$$

$$[\nabla\times(\phi\mathbf{u})]_i = \epsilon_{ijk}\frac{\partial}{\partial x_j}(\phi u_k) = \epsilon_{ijk}\left(u_k\frac{\partial\phi}{\partial x_j} + \phi\frac{\partial u_k}{\partial x_j}\right) = [\nabla\phi\times\mathbf{u} + \phi\nabla\times\mathbf{u}]_i.$$

Worked Example 3 ($\mathrm{curl}(\mathrm{grad}\,\phi) = 0$)

$$[\nabla\times(\nabla\phi)]_i = \epsilon_{ijk}\frac{\partial}{\partial x_j}\left(\frac{\partial\phi}{\partial x_k}\right) = \epsilon_{ijk}\frac{\partial^2\phi}{\partial x_j\partial x_k} = 0 \quad \textit{(symmetric/anti-symmetric in } j,k\textit{)}.$$

1.7 Derivatives of the Position Vector

Since $\partial x_i/\partial x_j = \delta_{ij}$ (directly from the definition of a partial derivative: $\partial x_1/\partial x_1 = 1$, $\partial x_1/\partial x_2 = 0$, etc.),

$$\nabla\cdot\mathbf{x} = \frac{\partial x_i}{\partial x_i} = \delta_{ii} = 3.$$

Here δ_{ii} is *both* indices repeated, so it is a full sum with no free index left over: $\delta_{ii} = \delta_{11} + \delta_{22} + \delta_{33} = 1 + 1 + 1 = 3$ (this is just the trace of the 3×3 identity matrix). For the curl,

$$[\nabla\times\mathbf{x}]_i = \epsilon_{ijk}\frac{\partial x_k}{\partial x_j} = \epsilon_{ijk}\delta_{jk}.$$

Now use the sifting property on δ_{jk} to collapse the sum over k (replacing every k with j): $\epsilon_{ijk}\delta_{jk} = \epsilon_{ijj}$. But ϵ_{ijj} has a *repeated* index (j appears twice within the permutation symbol itself, in positions 2 and 3) — by definition $\epsilon_{ijk} = 0$ whenever any two of its three indices coincide. Hence $[\nabla\times\mathbf{x}]_i = 0$ for every i, i.e. $\nabla\times\mathbf{x} = \mathbf{0}$.

1.8 Magnitude of a Vector

Worked Example 4 ($\nabla|\mathbf{x}|$) *Forcing the magnitude into a scalar product* $|\mathbf{x}| = (x_jx_j)^{1/2}$,

$$[\nabla|\mathbf{x}|]_i = \frac{\partial}{\partial x_i}(x_jx_j)^{1/2} = \tfrac{1}{2}(x_kx_k)^{-1/2}\,2x_j\frac{\partial x_j}{\partial x_i} = (x_kx_k)^{-1/2}x_j\delta_{ij} = (x_kx_k)^{-1/2}x_i,$$

i.e. $\nabla|\mathbf{x}| = \mathbf{x}/|\mathbf{x}|$*, the unit vector in the direction of* $\mathbf{x}$*. (Note: the dummy index* j *was renamed to* k *in the intermediate step to avoid a forbidden triple repetition.)*

1.9 Integral Theorems

Divergence theorem.

$$\int_V \nabla\cdot\mathbf{T}\,dV = \int_S \mathbf{T}\cdot\mathbf{n}\,dS \qquad\Longleftrightarrow\qquad \int_V \frac{\partial T_i}{\partial x_i}\,dV = \int_S T_in_i\,dS,$$

where V is enclosed by surface $S = \partial V$ and $\mathbf{n}$ is the outward unit normal.

Corollary, what is $\int_V \nabla\phi\, dV$ for a scalar ϕ? We cannot apply the divergence theorem directly to $\nabla\phi$, since the divergence theorem as stated needs a vector to dot with $\mathbf{n}$, not a bare gradient. The trick is to *manufacture* a vector out of ϕ using an arbitrary constant vector $\mathbf{a}$ ("arbitrary" meaning: it does not depend on $\mathbf{x}$, so it can be pulled outside any integral over V or S). Define the vector field $f_i = \phi\, a_i$. Applying the divergence theorem (in index form) to f_i:

$$\int_V \frac{\partial f_i}{\partial x_i}\, dV = \int_S f_i n_i\, dS \qquad \Longrightarrow \qquad \int_V \frac{\partial(\phi a_i)}{\partial x_i}\, dV = \int_S \phi a_i n_i\, dS.$$

Since a_i is constant, $\partial(\phi a_i)/\partial x_i = a_i\, \partial\phi/\partial x_i$, and a_i can be pulled outside both integrals:

$$a_i \int_V \frac{\partial \phi}{\partial x_i}\, dV = a_i \int_S \phi\, n_i\, dS \qquad \Longrightarrow \qquad a_i \left(\int_V \frac{\partial \phi}{\partial x_i}\, dV - \int_S \phi\, n_i\, dS \right) = 0.$$

Now the key step: this must hold for *every choice* of the fixed vector a_i (we never used any special property of $\mathbf{a}$ to derive it). Choose $\mathbf{a} = \mathbf{e}_1$ (i.e. $a_1 = 1, a_2 = 0, a_3 = 0$): the bracket's first component must vanish. Choose $\mathbf{a} = \mathbf{e}_2$, then $\mathbf{a} = \mathbf{e}_3$: each component of the bracket must vanish in turn. Hence the bracket is zero *component by component*, independent of $\mathbf{a}$:

$$\int_V \frac{\partial \phi}{\partial x_i}\, dV = \int_S \phi\, n_i\, dS \qquad \Longleftrightarrow \qquad \int_V \nabla\phi\, dV = \int_S \phi \mathbf{n}\, dS.$$

This "insert a fixed vector, then argue it was arbitrary" trick generalizes: for *any* index-notation expression $(\star)$ (scalar, vector-component, or tensor-component — it does not matter, since the argument never used what $(\star)$ was), the identical steps give the general corollary

$$\int_V \frac{\partial}{\partial x_i}(\star)\, dV = \int_S (\star)\, n_i\, dS.$$

We will use exactly this corollary, with $(\star) = \sigma_{ij}$, when deriving the equations of motion in Section 2.

Stokes' theorem.

$$\int_S \nabla \times \mathbf{T}\, d\mathbf{A} = \oint_C \mathbf{T} \cdot d\mathbf{l} \qquad \Longleftrightarrow \qquad \int_S \epsilon_{ijk} \frac{\partial T_k}{\partial x_j} n_i\, dS = \oint_C T_i s_i\, dC,$$

with $C = \partial S$, tangent $\mathbf{s}$, and $d\mathbf{A} = \mathbf{n}\, dS$.

1.10 Tensors

A vector $\mathbf{a}$ has components (a_1, a_2, a_3) *with respect to a basis* $\mathbf{e}_i$, $\mathbf{a} = a_i \mathbf{e}_i$, and rotating the coordinate axes leaves the vector itself (the arrow) unchanged — only its components and the basis change together. A **second-order tensor** is likewise a basis-dependent object,

$$A = A_{ij}\, \mathbf{e}_i \otimes \mathbf{e}_j,$$

which, with respect to a fixed orthonormal basis, we manipulate via its components A_{ij} much like a matrix.

Where the rotation rule comes from. Suppose a rotation takes the old basis $\mathbf{e}_i$ to a new basis $\mathbf{e}'_i$, and define $R_{ij} = \mathbf{e}'_i \cdot \mathbf{e}_j$ (the component of the new i-th basis vector along the old j-th axis). Any vector $\mathbf{v}$ has components v_i in the old frame and v'_i in the new frame, related by

$$v'_i = \mathbf{v} \cdot \mathbf{e}'_i = (v_j \mathbf{e}_j) \cdot \mathbf{e}'_i = v_j (\mathbf{e}_j \cdot \mathbf{e}'_i) = R_{ij} v_j,$$

i.e. $v'_i = R_{ij} v_j$ — a vector transforms by one factor of R.

Now think of a second-order tensor A concretely as *a linear rule that turns one vector into another*: $u_i = A_{ij} v_j$ (e.g. stress turning a normal $\mathbf{n}$ into a traction $\mathbf{t}$). The physical relationship between $\mathbf{u}$ and $\mathbf{v}$ cannot depend on which coordinate frame we describe it in, so in the new frame the *same* relationship must read $u'_i = A'_{ij} v'_j$ for some new component array A'_{ij}. We now find A'_{ij} by transforming both sides of $u_i = A_{ij} v_j$:

$$u'_i = R_{ik} u_k = R_{ik} (A_{kl} v_l).$$

We want to re-express v_l in terms of the new components v'_j. Since $v'_j = R_{jl} v_l$ and R is a rotation (orthogonal matrix,

$R^{-1} = R^T$, i.e. $R_{jl}R_{ml} = \delta_{jm}$), we can invert: $R_{jl}v'_j = R_{jl}R_{jm}v_m = \delta_{lm}v_m = v_l$, i.e. $v_l = R_{jl}v'_j$. Substituting,

$$u'_i = R_{ik}A_{kl}\big(R_{jl}v'_j\big) = \big(R_{ik}A_{kl}R_{jl}\big)v'_j.$$

Comparing with $u'_i = A'_{ij}v'_j$ (true for *every* vector v'_j, so we may match coefficients) gives the transformation rule

$$\boxed{A'_{ij} = R_{ik}A_{kl}R_{jl}}.$$

This is exactly matrix multiplication $[A'] = [R][A][R]^T$, written out in index notation. We will use precisely this rule for the stress tensor in Section 2.6.

Common second-order tensors: stress and strain.

Derivatives of tensors. The gradient of a vector is a second-order tensor, and the divergence of a second-order tensor is a vector:

$$[\nabla\mathbf{a}]_{ij} = \frac{\partial a_i}{\partial x_j}, \qquad [\nabla\cdot A]_i = \frac{\partial A_{ij}}{\partial x_j}, \qquad \int_V \nabla\cdot A\,dV = \int_{\partial V} A\mathbf{n}\,dS.$$

Worked Example 5 (Balance of linear momentum, from tensor calculus) *If $\mathbf{t} = \sigma\mathbf{n}$ is the traction on a surface with outward normal $\mathbf{n}$ (i.e. $t_i = \sigma_{ij}n_j$), then equilibrium of a body of density ρ under gravity $\mathbf{b}$ requires*

$$\int_{\partial V}\mathbf{t}\,dS + \int_V \rho\mathbf{b}\,dV = 0 \implies \int_{\partial V}\sigma\mathbf{n}\,dS + \int_V \rho\mathbf{b}\,dV = 0 \xrightarrow{\text{div. thm}} \int_V (\nabla\cdot\sigma + \rho\mathbf{b})\,dV = 0.$$

Since this holds for any *sub-body V, it localizes to $\nabla\cdot\sigma + \rho\mathbf{b} = 0$: the equilibrium (balance of linear momentum) equation. We will re-derive this more physically, from first principles, in Section 2.*

2 Stress, and Balance of Linear & Angular Momentum

2.1 Surface Traction and Body Force

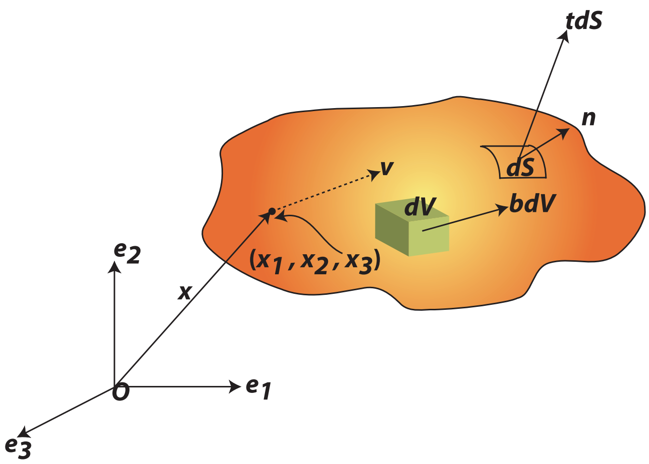


Figure 1: Coordinate system, body force, $\mathbf{b}$, and traction, $\mathbf{t}$.

Forces, $\mathbf{P}$, act on a body in two ways: (1) on the boundary — fluid pressure, wind loads, contact forces; (2) in the interior — body forces such as gravity or electromagnetic forces.

The **surface traction vector $\mathbf{T}$** is the force per unit area of the deformed solid:

$$\mathbf{T} = \lim_{dA\to 0}\frac{d\mathbf{P}}{dA}, \qquad \mathbf{P} = \int_S \mathbf{T}\,dA.$$

It has normal and shear components: $\mathbf{T}_n = (\mathbf{T}\cdot\mathbf{n})\mathbf{n}, \quad \mathbf{T}_s = \mathbf{T} - \mathbf{T}_n$.

The **body force vector** is $\mathbf{b} = \frac{1}{\rho}\lim_{dV\to 0}\frac{d\mathbf{P}}{dV}$, so that $\mathbf{P} = \int_V \rho\mathbf{b}\,dV$.

2.2 Internal Traction and Newton's Third Law

Every internal plane in a solid is subjected to a distribution of traction (Leibniz & Bernoulli, late 1600s). Imagine cutting a body into two parts: each part must remain in static equilibrium, which is only possible if forces act on the new (cut) surface. These are quantified by the internal traction vector $\mathbf{T}(\mathbf{n}) = \lim_{dA\to 0} d\mathbf{P}(\mathbf{n})/dA$ where $\mathbf{n}$ is the unit normal to the surface pointing outwards. The total force acting on any internal volume of a deformed solid is

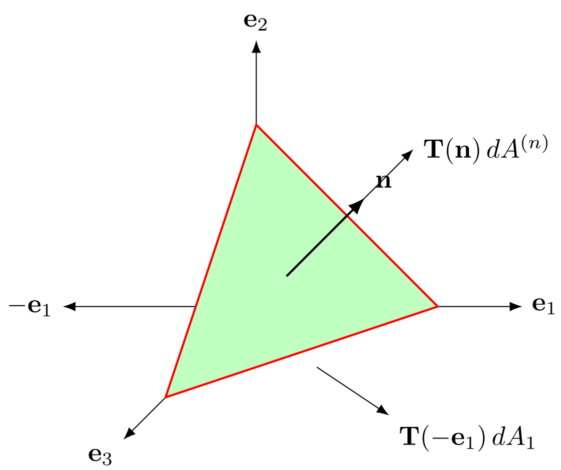
$$\mathbf{P} = \int_V \mathbf{T}(\mathbf{n})\,dA + \int_V \rho\mathbf{b}\,dV.$$

Newton's third law requires

$$\mathbf{T}(\mathbf{n}) = -\mathbf{T}(-\mathbf{n}).$$

Tractions acting on different planes through the same point are related to each other so as to satisfy $\mathbf{F} = m\mathbf{a}$.

2.3 The Tetrahedron Argument: Cauchy's Stress Tensor



The base and slanted faces of the tetrahedron have unit outward normals $-\mathbf{e}_1, -\mathbf{e}_2, -\mathbf{e}_3$ (coordinate planes) and $\mathbf{n}$ (slant face, area $dA^{(n)}$). Newton's second law, $\mathbf{F} = m\mathbf{a}$, applied to the tetrahedron:

$$\mathbf{T}(\mathbf{n})\,dA^{(n)} + \mathbf{T}(-\mathbf{e}_1)\,dA_1 + \mathbf{T}(-\mathbf{e}_2)\,dA_2 + \mathbf{T}(-\mathbf{e}_3)\,dA_3 + \rho\mathbf{b}\,dV = \rho\,dV\,\mathbf{a}.$$

Using $\mathbf{T}(-\mathbf{e}_i) = -\mathbf{T}(\mathbf{e}_i)$ (Newton's 3rd law) and dividing through by $dA^{(n)}$:

$$\mathbf{T}(\mathbf{n}) - \mathbf{T}(\mathbf{e}_1)\frac{dA_1}{dA^{(n)}} - \mathbf{T}(\mathbf{e}_2)\frac{dA_2}{dA^{(n)}} - \mathbf{T}(\mathbf{e}_3)\frac{dA_3}{dA^{(n)}} + \rho\mathbf{b}\frac{dV}{dA^{(n)}} = \rho\frac{dV}{dA^{(n)}}\mathbf{a}.$$

In the limit $dA^{(n)} \to 0$: $dA_i/dA^{(n)} = n_i$ and $dV/dA^{(n)} \to 0$, so

$$\mathbf{T}(\mathbf{n}) = \mathbf{T}(\mathbf{e}_1)\,n_1 + \mathbf{T}(\mathbf{e}_2)\,n_2 + \mathbf{T}(\mathbf{e}_3)\,n_3, \qquad \text{i.e.} \qquad T_i(\mathbf{n}) = T_i(\mathbf{e}_1)n_1 + T_i(\mathbf{e}_2)n_2 + T_i(\mathbf{e}_3)n_3.$$

The significance: the tractions acting on the three coordinate planes are sufficient to characterize the internal state of stress at a point — this is Cauchy's stress tensor. Define

$$\boxed{\sigma_{ij} = T_j(\mathbf{e}_i)} \qquad \text{(direction of normal = } i\text{, direction of traction component = } j\text{),}$$

so that

$$\mathbf{T}(\mathbf{n}) = \mathbf{n}\cdot\sigma \qquad \Longleftrightarrow \qquad T_i(\mathbf{n}) = n_j\sigma_{ji}.$$

Proof that $dA_i/dA^{(n)} = n_i$, worked out in full. Let the tetrahedron have its right-angle vertex at the origin, with the three edges along the axes meeting the slant face at the points $\mathbf{a} = a\mathbf{e}_1$, $\mathbf{b} = b\mathbf{e}_2$, $\mathbf{c} = c\mathbf{e}_3$ (so $a, b, c > 0$ are the edge lengths). The three coordinate-plane faces are right triangles with legs b, c (in the $\mathbf{e}_2$-$\mathbf{e}_3$ plane, area dA_1), a, c (area dA_2), and a, b (area dA_3):

$$dA_1 = \tfrac{1}{2}bc, \qquad dA_2 = \tfrac{1}{2}ac, \qquad dA_3 = \tfrac{1}{2}ab.$$

For the slant face, recall that for a triangle with two edge vectors $\mathbf{p}, \mathbf{q}$ emanating from one vertex, the vector area is $\frac{1}{2}\mathbf{p}\times\mathbf{q}$ (magnitude = area, direction = normal, by the right-hand rule). Taking the two edges of the slant face that emanate from the vertex $\mathbf{b}$, namely $(\mathbf{c} - \mathbf{b})$ and $(\mathbf{a} - \mathbf{b})$, the vector area of the slant face is

$$(dA^{(n)})\,\mathbf{n} = \tfrac{1}{2}(\mathbf{c} - \mathbf{b}) \times (\mathbf{a} - \mathbf{b}).$$

Expand this cross product term by term, using $\mathbf{c} - \mathbf{b} = c\mathbf{e}_3 - b\mathbf{e}_2$ and $\mathbf{a} - \mathbf{b} = a\mathbf{e}_1 - b\mathbf{e}_2$:

$$(\mathbf{c} - \mathbf{b}) \times (\mathbf{a} - \mathbf{b}) = (c\mathbf{e}_3 - b\mathbf{e}_2) \times (a\mathbf{e}_1 - b\mathbf{e}_2)$$

$$= ca(\mathbf{e}_3 \times \mathbf{e}_1) - cb(\mathbf{e}_3 \times \mathbf{e}_2) - ba(\mathbf{e}_2 \times \mathbf{e}_1) + b^2(\mathbf{e}_2 \times \mathbf{e}_2).$$

Using the standard right-handed basis identities $\mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$, $\mathbf{e}_3 \times \mathbf{e}_2 = -\mathbf{e}_1$, $\mathbf{e}_2 \times \mathbf{e}_1 = -\mathbf{e}_3$, and $\mathbf{e}_2 \times \mathbf{e}_2 = \mathbf{0}$:

$$= ca\,\mathbf{e}_2 - cb(-\mathbf{e}_1) - ba(-\mathbf{e}_3) + 0 = bc\,\mathbf{e}_1 + ac\,\mathbf{e}_2 + ab\,\mathbf{e}_3.$$

So

$$(dA^{(n)})\,\mathbf{n} = \tfrac{1}{2}\big(bc\,\mathbf{e}_1 + ac\,\mathbf{e}_2 + ab\,\mathbf{e}_3\big) = \big(\tfrac{1}{2}bc\big)\mathbf{e}_1 + \big(\tfrac{1}{2}ac\big)\mathbf{e}_2 + \big(\tfrac{1}{2}ab\big)\mathbf{e}_3$$
$$= dA_1\mathbf{e}_1 + dA_2\mathbf{e}_2 + dA_3\mathbf{e}_3.$$

Dividing both sides by the scalar $dA^{(n)}$ and matching the coefficients of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ component by component gives

$$\mathbf{n} = \frac{dA_1}{dA^{(n)}}\mathbf{e}_1 + \frac{dA_2}{dA^{(n)}}\mathbf{e}_2 + \frac{dA_3}{dA^{(n)}}\mathbf{e}_3, \qquad \text{i.e.} \qquad n_i = \frac{dA_i}{dA^{(n)}} \quad (i = 1, 2, 3). \qquad \blacksquare$$

2.4 Balance of Linear Momentum

Let $\mathbf{P}$ and $\mathbf{H}$ be linear and angular momentum (about origin O) of the matter occupying volume V:

$$\mathbf{P} = \int_V \rho\mathbf{v}\,dV, \qquad \mathbf{H} = \int_V \rho\,\mathbf{x} \times \mathbf{v}\,dV.$$

Newtonian mechanics: $\mathbf{F} = d\mathbf{P}/dt$, $\mathbf{M} = d\mathbf{H}/dt$ (statics: $\mathbf{F} = \mathbf{M} = 0$). With surface tractions and body forces,

$$\mathbf{F} = \int_S \mathbf{t}\,dS + \int_V \mathbf{b}\,dV, \qquad \mathbf{M} = \int_S \mathbf{x} \times \mathbf{t}\,dS + \int_V \mathbf{x} \times \mathbf{b}\,dV,$$

and $d\mathbf{P}/dt = \int_V \rho\dot{\mathbf{v}}\,dV \equiv \int_V \rho\mathbf{a}\,dV$, $d\mathbf{H}/dt = \int_V \rho\,\mathbf{x} \times \dot{\mathbf{v}}\,dV$ (ignoring momentum flux across S, valid below melting point, where $\mathbf{a} = \dot{\mathbf{v}}$ is acceleration). Newton's second law $\mathbf{F} = d\mathbf{P}/dt$ then reads, component by component ($i = 1, 2, 3$),

$$\int_S t_i\,dS + \int_V b_i\,dV = \int_V \rho a_i\,dV.$$

Step 1: replace the traction with the stress tensor. Substitute Cauchy's result $t_i = \sigma_{ij}n_j$ (derived in Section 2.3 above) into the surface integral:

$$\int_S \sigma_{ij}n_j\,dS + \int_V b_i\,dV = \int_V \rho a_i\,dV.$$

Step 2: convert the surface integral to a volume integral. For each fixed value of i, the quantity σ_{ij} (as j ranges over 1,2,3) is an ordinary vector field, so the divergence theorem applies directly:

$$\int_S \sigma_{ij}n_j\,dS = \int_V \frac{\partial \sigma_{ij}}{\partial x_j}\,dV.$$

Substituting this in, every term is now a volume integral over the *same* volume V:

$$\int_V \frac{\partial \sigma_{ij}}{\partial x_j}\,dV + \int_V b_i\,dV = \int_V \rho a_i\,dV \qquad \Longrightarrow \qquad \int_V \left[\frac{\partial \sigma_{ij}}{\partial x_j} + b_i - \rho a_i\right] dV = 0.$$

Step 3: localize. This must hold for *every* choice of sub-volume V we might have chosen to isolate — V was an arbitrary piece of the body, not a special one. If the bracketed integrand were nonzero (positive, say) at some point $\mathbf{x}_0$, then by continuity it would stay positive in some small neighborhood of $\mathbf{x}_0$; choosing V to be a tiny volume entirely inside that neighborhood would then give a strictly positive integral, contradicting that the integral is zero for *every* V. The same argument rules out the integrand being negative. Hence the integrand must vanish at every point, giving the **equations of motion** (pointwise, not just in integrated/average form):

$$\boxed{\frac{\partial \sigma_{ij}}{\partial x_j} + b_i = \rho a_i}, \qquad (i = 1, 2, 3).$$

This "integral over an arbitrary volume vanishes ⇒ the integrand vanishes pointwise" localization argument is a standard move in continuum mechanics — we will use it again immediately, to get stress symmetry from angular momentum balance.

2.5 Balance of Angular Momentum: Stress Symmetry

Setting up the moment balance. The moment (torque) of a force $\mathbf{t}\, dS$ acting at position $\mathbf{x}$, about the origin, is the cross product $\mathbf{x} \times \mathbf{t}\, dS$, with components $\epsilon_{kij} x_i t_j \, dS$. It is more convenient here to work with the antisymmetric *pair* $(x_i t_j - x_j t_i)$ directly rather than introducing yet another permutation symbol; the two carry the same information (a cross product's three components and an antisymmetric 3×3 array's three independent off-diagonal entries are just two ways of recording the same thing). Angular momentum balance ($\mathbf{M} = d\mathbf{H}/dt$, in this antisymmetric-pair form) reads

$$\int_S (x_i t_j - x_j t_i)\, dS + \int_V (x_i b_j - x_j b_i)\, dV = \int_V \rho (x_i a_j - x_j a_i)\, dV. \tag{$*$}$$

Our goal is to substitute $t_j = \sigma_{jk} n_k$, use the divergence theorem, and see what falls out.

Step 1: substitute the traction. Using $t_j = \sigma_{jk} n_k$ and $t_i = \sigma_{ik} n_k$, the surface-integral term of $(*)$ becomes

$$\int_S (x_i \sigma_{jk} n_k - x_j \sigma_{ik} n_k)\, dS.$$

Step 2: apply the divergence theorem. For fixed i, j, the quantity in parentheses is a vector field in the free index k (contracted with n_k), so the divergence theorem converts the surface integral to a volume integral of a k-derivative:

$$\int_S (x_i \sigma_{jk} - x_j \sigma_{ik}) n_k \, dS = \int_V \frac{\partial}{\partial x_k}\big(x_i \sigma_{jk} - x_j \sigma_{ik}\big)\, dV.$$

Step 3: expand the derivative with the product rule. Take the two terms separately. For the first,

$$\frac{\partial}{\partial x_k}(x_i \sigma_{jk}) = \frac{\partial x_i}{\partial x_k}\sigma_{jk} + x_i \frac{\partial \sigma_{jk}}{\partial x_k} = \delta_{ik}\sigma_{jk} + x_i \frac{\partial \sigma_{jk}}{\partial x_k}.$$

Now apply the sifting property to $\delta_{ik}\sigma_{jk}$: summing over k, only the $k = i$ term survives, replacing k by i: $\delta_{ik}\sigma_{jk} = \sigma_{ji}$. So

$$\frac{\partial}{\partial x_k}(x_i \sigma_{jk}) = \sigma_{ji} + x_i \frac{\partial \sigma_{jk}}{\partial x_k}.$$

By the identical steps with $i \leftrightarrow j$ swapped,

$$\frac{\partial}{\partial x_k}(x_j \sigma_{ik}) = \sigma_{ij} + x_j \frac{\partial \sigma_{ik}}{\partial x_k}.$$

Subtracting,

$$\frac{\partial}{\partial x_k}\big(x_i \sigma_{jk} - x_j \sigma_{ik}\big) = (\sigma_{ji} - \sigma_{ij}) + x_i \frac{\partial \sigma_{jk}}{\partial x_k} - x_j \frac{\partial \sigma_{ik}}{\partial x_k}.$$

Step 4: bring in the linear-momentum result. We already derived, in the previous subsection, that $\partial \sigma_{mn}/\partial x_n + b_m = \rho a_m$ for *any* free index m (it holds for $m = 1, 2, 3$ individually) — so we are free to relabel $m \to j$ (dummy $n \to k$) or

$m \to i$ (dummy $n \to k$) as needed:

$$\frac{\partial \sigma_{jk}}{\partial x_k} = \rho a_j - b_j, \qquad \frac{\partial \sigma_{ik}}{\partial x_k} = \rho a_i - b_i.$$

Substituting these into the boxed expression from Step 3:

$$\frac{\partial}{\partial x_k}\big(x_i\sigma_{jk} - x_j\sigma_{ik}\big) = (\sigma_{ji} - \sigma_{ij}) + x_i(\rho a_j - b_j) - x_j(\rho a_i - b_i)$$
$$= (\sigma_{ji} - \sigma_{ij}) + \rho(x_i a_j - x_j a_i) - (x_i b_j - x_j b_i).$$

Step 5: put it back into $(*)$. The surface-integral term of $(*)$ is therefore

$$\int_S (x_i t_j - x_j t_i)\, dS = \int_V \Big[(\sigma_{ji} - \sigma_{ij}) + \rho(x_i a_j - x_j a_i) - (x_i b_j - x_j b_i)\Big]\, dV.$$

Substitute this back into $(*)$:

$$\int_V \Big[(\sigma_{ji} - \sigma_{ij}) + \rho(x_i a_j - x_j a_i) - (x_i b_j - x_j b_i)\Big]\, dV + \int_V (x_i b_j - x_j b_i)\, dV = \int_V \rho(x_i a_j - x_j a_i)\, dV.$$

Step 6: cancel matching terms. The $-(x_i b_j - x_j b_i)$ inside the first integral cancels exactly against the $+\int_V (x_i b_j - x_j b_i)\, dV$ term next to it. The $\rho(x_i a_j - x_j a_i)$ term inside the first integral is identical to the term on the right-hand side, so it cancels too (subtract it from both sides). What survives is simply

$$\int_V (\sigma_{ji} - \sigma_{ij})\, dV = 0.$$

Step 7: localize. Exactly as in the linear-momentum case, this integral vanishes for *every* choice of sub-volume V, so by the same continuity argument the integrand itself must vanish at every point:

$$\boxed{\sigma_{ij} = \sigma_{ji}}: \qquad \textbf{the stress tensor is symmetric.}$$

Only 6 of the 9 components of σ_{ij} are independent: $\sigma_{11}, \sigma_{22}, \sigma_{33}$ (normal stresses) and $\sigma_{12} = \sigma_{21}, \sigma_{13} = \sigma_{31}, \sigma_{23} = \sigma_{32}$ (shear stresses).

2.6 Tensor Transformation of Stress

This is a direct application of the general tensor rotation rule derived in Section 1.10, specialized to the stress tensor. Let $[Q]$ be the rotation matrix between coordinate systems (x_1, x_2, x_3) and (x_1', x_2', x_3') (so Q_{ij} is the same kind of object as R_{ij} before, and $[Q]^{-1} = [Q]^T$ since it is a rotation). We know the traction vector transforms as an ordinary vector: $\{T'\} = [Q]\{T\}$, and likewise $\{n'\} = [Q]\{n\}$, i.e. $\{n\} = [Q]^T\{n'\}$ (inverting, using $[Q]^{-1} = [Q]^T$).

Derivation, step by step. Start from Cauchy's relation in the old frame, $\{T\} = [\sigma]\{n\}$, and substitute $\{n\} = [Q]^T\{n'\}$:

$$\{T\} = [\sigma][Q]^T\{n'\}.$$

Now transform the left-hand side to the new frame using $\{T'\} = [Q]\{T\}$, i.e. $\{T\} = [Q]^T\{T'\}$:

$$[Q]^T\{T'\} = [\sigma][Q]^T\{n'\}.$$

Multiply both sides on the left by $[Q]$ (using $[Q][Q]^T = [I]$):

$$\{T'\} = [Q][\sigma][Q]^T\{n'\}.$$

But in the new frame, Cauchy's relation must take the *same* form: $\{T'\} = [\sigma']\{n'\}$, for some new stress-component array

$[\sigma']$. Comparing the last two equations (true for every $\{n'\}$, so we may match coefficients):

$$\boxed{[\sigma'] = [Q][\sigma][Q]^T}, \qquad \sigma'_{kl} = Q_{ik}\sigma_{ij}Q_{jl} = Q_{ik}Q_{jl}\,\sigma_{ij}.$$

(In the last step we simply reordered the scalar factors σ_{ij} and Q_{jl}, which commute, to match the earlier form $A'_{ij} = R_{ik}A_{kl}R_{jl}$ from Section 1.10 — stress transforms exactly like any other second-order tensor.)

2.7 Principal Stresses

Symmetry of the stress tensor, together with tensor transformation, implies that at each point $\mathbf{x}$ there exist three mutually perpendicular directions along which there is no shear stress: the **principal directions**, with corresponding **principal stresses** $\sigma_I \leq \sigma_{II} \leq \sigma_{III}$. We ask: for what directions $\mathbf{n}$ is $\mathbf{T} = \sigma\mathbf{n}$ aligned with $\mathbf{n}$ (i.e. purely normal, no shear)? Aligned means $\mathbf{T} = \sigma\mathbf{n}$ for some *scalar* σ (reusing the symbol σ for the eigenvalue is standard, if a little confusing at first), i.e.

$$\sigma_{ij}n_j = \sigma n_i \qquad \Longleftrightarrow \qquad (\sigma_{ij} - \sigma\delta_{ij})n_j = 0 \qquad (\text{using } n_i = \delta_{ij}n_j).$$

This is exactly the statement that $\mathbf{n}$ is an **eigenvector** of the matrix $[\sigma_{ij}]$ with **eigenvalue** σ — the same eigenvalue problem you already know from linear algebra, here applied to the stress matrix. A nonzero $\mathbf{n}$ solving this homogeneous linear system exists only if the matrix $[\sigma_{ij} - \sigma\delta_{ij}]$ is singular, i.e.

$$\det[\sigma_{ij} - \sigma\delta_{ij}] = 0.$$

Expanding the determinant. Written out, with $\sigma_{12} = \sigma_{21}$ etc. (symmetry already used),

$$\det\begin{pmatrix} \sigma_{11}-\sigma & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22}-\sigma & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33}-\sigma \end{pmatrix} = 0.$$

Expand by cofactors along the first row:

$$(\sigma_{11}-\sigma)\big[(\sigma_{22}-\sigma)(\sigma_{33}-\sigma) - \sigma_{23}^2\big] - \sigma_{12}\big[\sigma_{12}(\sigma_{33}-\sigma) - \sigma_{23}\sigma_{13}\big]$$
$$+\sigma_{13}\big[\sigma_{12}\sigma_{23} - (\sigma_{22}-\sigma)\sigma_{13}\big] = 0.$$

Multiplying out and collecting powers of σ (a mechanical but lengthy algebra exercise — try it for a specific numerical matrix, as in the worked example below, to see it concretely) produces a cubic

$$-\sigma^3 + I_1\sigma^2 + I_2\sigma + I_3 = 0,$$

with the **stress invariants**

$$I_1 = \sigma_{ii} = \sigma_{11} + \sigma_{22} + \sigma_{33}, \qquad I_2 = -\tfrac{1}{2}I_1^2 + \tfrac{1}{2}\sigma_{ij}\sigma_{ji}, \qquad I_3 = \det[\sigma_{ij}].$$

They are called *invariants* because they do not depend on which coordinate frame we used to write down σ_{ij} in the first place — physically, the principal stresses themselves (roots of this cubic) obviously cannot depend on our choice of axes, so neither can the cubic's coefficients. This is easy to check directly for I_1: in a rotated frame, using the transformation rule from Section 2.6,

$$\sigma'_{kk} = Q_{ik}Q_{jk}\,\sigma_{ij}.$$

Since Q is a rotation matrix, $Q_{ik}Q_{jk} = \delta_{ij}$ (summing over k is the statement that the rows of Q are orthonormal). So $\sigma'_{kk} = \delta_{ij}\sigma_{ij} = \sigma_{ii}$ (sifting) — the trace is exactly the same number in every frame.

Worked Example 6 *For the plane-stress state* $\sigma_{ij} = \left(\begin{smallmatrix} \sigma_{11} & \sigma_{12} & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ 0 & 0 & 0 \end{smallmatrix}\right)$ *(with* $\sigma_{12} = \sigma_{21}$*), first compute the invariants directly from their definitions.* $I_1 = \sigma_{11} + \sigma_{22} + 0 = \sigma_{11} + \sigma_{22}$*. For* I_2*:* $I_1^2 = (\sigma_{11} + \sigma_{22})^2 = \sigma_{11}^2 + 2\sigma_{11}\sigma_{22} + \sigma_{22}^2$*, and* $\sigma_{ij}\sigma_{ji} = \sigma_{11}^2 + \sigma_{22}^2 + 2\sigma_{12}^2$ *(summing all nine products* $\sigma_{ij}\sigma_{ji}$*, only these four survive since the third row/column is zero), so*

$$I_2 = -\tfrac{1}{2}(\sigma_{11}^2 + 2\sigma_{11}\sigma_{22} + \sigma_{22}^2) + \tfrac{1}{2}(\sigma_{11}^2 + \sigma_{22}^2 + 2\sigma_{12}^2) = -\sigma_{11}\sigma_{22} + \sigma_{12}^2.$$

$I_3 = \det[\sigma_{ij}] = 0$ *directly (third row is all zeros). The cubic becomes*

$$-\sigma^3 + (\sigma_{11} + \sigma_{22})\sigma^2 + (\sigma_{12}^2 - \sigma_{11}\sigma_{22})\sigma = 0.$$

Factor out σ*:*

$$\sigma\Big[-\sigma^2 + (\sigma_{11} + \sigma_{22})\sigma + (\sigma_{12}^2 - \sigma_{11}\sigma_{22})\Big] = 0.$$

One root is immediately $\sigma_I = 0$ *(the out-of-plane direction, as expected for plane stress). The other two roots solve the quadratic* $\sigma^2 - (\sigma_{11} + \sigma_{22})\sigma - (\sigma_{12}^2 - \sigma_{11}\sigma_{22}) = 0$ *(multiplying through by* -1*). By the quadratic formula,*

$$\sigma_{II,III} = \frac{(\sigma_{11} + \sigma_{22}) \pm \sqrt{(\sigma_{11} + \sigma_{22})^2 + 4(\sigma_{12}^2 - \sigma_{11}\sigma_{22})}}{2}.$$

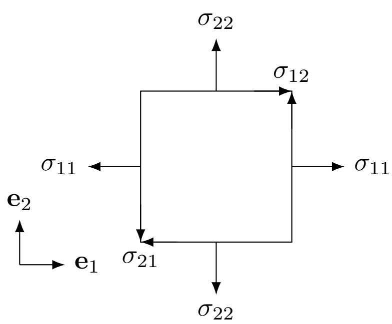
Simplify the term under the square root:

$$(\sigma_{11} + \sigma_{22})^2 + 4\sigma_{12}^2 - 4\sigma_{11}\sigma_{22} = \sigma_{11}^2 + 2\sigma_{11}\sigma_{22} + \sigma_{22}^2 + 4\sigma_{12}^2 - 4\sigma_{11}\sigma_{22} = (\sigma_{11} - \sigma_{22})^2 + 4\sigma_{12}^2,$$

so, dividing the $\pm$ *term by 2 inside the root,*

$$\boxed{\sigma_I = 0, \qquad \sigma_{II,III} = \frac{\sigma_{11} + \sigma_{22}}{2} \mp \sqrt{\left(\frac{\sigma_{11} - \sigma_{22}}{2}\right)^2 + \sigma_{12}^2}.}$$

(The labeling of which root is σ_{II} *vs.* σ_{III} *just depends on ordering them so that* $\sigma_I \le \sigma_{II} \le \sigma_{III}$ *for the particular numbers at hand.)*



2.8 Mohr's Circle

Consider a 2D stress state $\sigma_{ij} = \left(\begin{smallmatrix}\sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22}\end{smallmatrix}\right)$, and a plane whose outward normal $\mathbf{n} = (-\sin\theta, \cos\theta, 0)$ makes angle θ with the x_2-axis; let $\mathbf{s} = (\cos\theta, \sin\theta, 0)$ be the corresponding in-plane tangent (perpendicular to $\mathbf{n}$, obtained by rotating $\mathbf{n}$ by $90°$). We want the normal traction $\sigma_n = \mathbf{n} \cdot \mathbf{T}$ and shear traction $\tau_{ns} = \mathbf{s} \cdot \mathbf{T}$ on this plane, as functions of θ.

Step 1: compute the traction components. $T_i = \sigma_{ij} n_j$, so

$$T_1 = \sigma_{11} n_1 + \sigma_{12} n_2 = -\sigma_{11}\sin\theta + \sigma_{12}\cos\theta, \qquad T_2 = \sigma_{12} n_1 + \sigma_{22} n_2 = -\sigma_{12}\sin\theta + \sigma_{22}\cos\theta.$$

Step 2: project onto $\mathbf{n}$ to get σ_n.

$$\sigma_n = n_1 T_1 + n_2 T_2 = (-\sin\theta)(-\sigma_{11}\sin\theta + \sigma_{12}\cos\theta) + (\cos\theta)(-\sigma_{12}\sin\theta + \sigma_{22}\cos\theta).$$

Expanding,

$$\begin{aligned}\sigma_n &= \sigma_{11}\sin^2\theta - \sigma_{12}\sin\theta\cos\theta - \sigma_{12}\sin\theta\cos\theta + \sigma_{22}\cos^2\theta \\ &= \sigma_{11}\sin^2\theta + \sigma_{22}\cos^2\theta - 2\sigma_{12}\sin\theta\cos\theta.\end{aligned}$$

Now use the double-angle identities $\sin^2\theta = \frac{1-\cos 2\theta}{2}$, $\cos^2\theta = \frac{1+\cos 2\theta}{2}$, $\sin\theta\cos\theta = \frac{\sin 2\theta}{2}$:

$$\sigma_n = \sigma_{11}\frac{1 - \cos 2\theta}{2} + \sigma_{22}\frac{1 + \cos 2\theta}{2} - 2\sigma_{12}\frac{\sin 2\theta}{2} = \frac{\sigma_{11} + \sigma_{22}}{2} + \frac{\sigma_{22} - \sigma_{11}}{2}\cos 2\theta - \sigma_{12}\sin 2\theta.$$

Step 3: project onto s to get τ_{ns}.

$$\tau_{ns} = s_1T_1 + s_2T_2 = \cos\theta(-\sigma_{11}\sin\theta + \sigma_{12}\cos\theta) + \sin\theta(-\sigma_{12}\sin\theta + \sigma_{22}\cos\theta).$$

Expanding,

$$\begin{aligned}\tau_{ns} &= -\sigma_{11}\sin\theta\cos\theta + \sigma_{12}\cos^2\theta - \sigma_{12}\sin^2\theta + \sigma_{22}\sin\theta\cos\theta \\ &= (\sigma_{22}-\sigma_{11})\sin\theta\cos\theta + \sigma_{12}(\cos^2\theta - \sin^2\theta).\end{aligned}$$

Using $\sin\theta\cos\theta = \frac{1}{2}\sin 2\theta$ and $\cos^2\theta - \sin^2\theta = \cos 2\theta$:

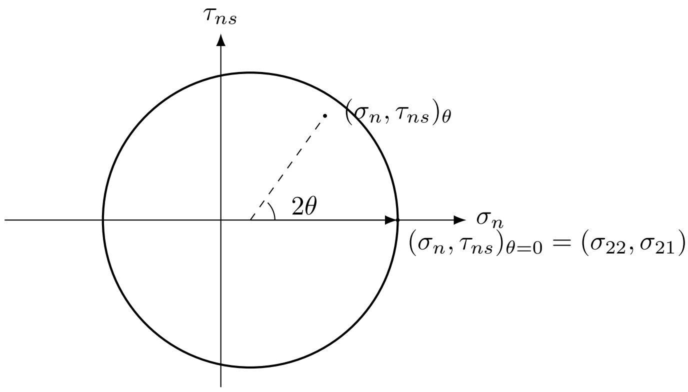
$$\tau_{ns} = \frac{\sigma_{22}-\sigma_{11}}{2}\sin 2\theta + \sigma_{12}\cos 2\theta.$$

Collecting the two results:

$$\boxed{\sigma_n = \frac{\sigma_{11}+\sigma_{22}}{2} + \frac{\sigma_{22}-\sigma_{11}}{2}\cos 2\theta - \sigma_{12}\sin 2\theta,}$$

$$\boxed{\tau_{ns} = \frac{\sigma_{22}-\sigma_{11}}{2}\sin 2\theta + \sigma_{12}\cos 2\theta.}$$

As θ varies, the point (σ_n, τ_{ns}) traces a circle centered at $\left(\frac{\sigma_{11}+\sigma_{22}}{2}, 0\right)$ with radius $R = \sqrt{\left(\frac{\sigma_{22}-\sigma_{11}}{2}\right)^2 + \sigma_{12}^2}$: both σ_n and τ_{ns} are linear combinations of $\cos 2\theta$, $\sin 2\theta$, and a constant, which is exactly the parametric form of a circle. A rotation of θ anticlockwise in physical space corresponds to a rotation of 2θ anticlockwise in the Mohr plane.



Worked Example 7 *Consider a 2D medium under uniform stress $\sigma_{ij} = \left(\begin{smallmatrix}\sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22}\end{smallmatrix}\right)$ (with $\sigma_{12} = \sigma_{21}$), and a horizontal surface S (a potential fault) with unit normal $\mathbf{n} = (0, 1)$. Then $T_i = \sigma_{ij}n_j = \sigma_{i2}$, so $T_1 = \sigma_{12}$, $T_2 = \sigma_{22}$: here the traction components happen to have the same magnitude as the corresponding stress components — this is* not *true in general, as the next example shows.*

Worked Example 8 *Let $S : x_2 - \sin x_1 = 0$. The unit normal to $f(x_1, x_2) = 0$ is $\mathbf{n} = \nabla f/|\nabla f|$, giving $\mathbf{n} = (1+\cos^2 x_1)^{-1/2}(-\cos x_1,\, 1)$. Then $T_i = \sigma_{ij}n_j$, and the normal traction $T_{(n)i} = (T_kn_k)n_i$ works out to*

$$\begin{aligned}T_{(n)1} &= \left(\frac{-\cos x_1}{(1+\cos^2 x_1)^{3/2}}\right)\left[\sigma_{11}\cos^2 x_1 + 2\sigma_{12}\cos x_1 - \sigma_{22}\right], \\ T_{(n)2} &= \left(\frac{1}{(1+\cos^2 x_1)^{3/2}}\right)\left[\sigma_{11}\cos^2 x_1 - 2\sigma_{12}\cos x_1 + \sigma_{22}\right].\end{aligned}$$

Sanity check: *at $x_1 = \pi/2$, $\mathbf{n} = (0, 1)$ as in the previous example, so $T_{(n)1} = 0$ and $T_{(n)2} = \sigma_{22}$ should hold — and indeed the formulas above give exactly this.*

3 Strain

3.1 The Deformation Gradient

Let a material point at position $\mathbf{x}$ move to $\mathbf{y} = \mathbf{x} + \mathbf{u}(\mathbf{x}, t)$ under deformation, i.e. $y_i = x_i + u_i(x_1, x_2, x_3, t)$. Then

$$\frac{\partial y_i}{\partial x_j} = \frac{\partial x_i}{\partial x_j} + \frac{\partial u_i}{\partial x_j} \equiv F_{ij}$$

$$\implies \qquad y_{i,j} = \delta_{ij} + u_{i,j} \equiv F_{ij} \quad \text{(the \textbf{deformation gradient tensor})}.$$

For a nearby material point at $\mathbf{x} + d\mathbf{x}$ mapping to $\mathbf{y} + d\mathbf{y}$: $dy_i = (x_i + dx_i) + u_i(x_k + dx_k) - u_i(x_k) - x_i$. For small $d\mathbf{x}$, $u_i(x_k + dx_k) \approx u_i(x_k) + \frac{\partial u_i}{\partial x_k} dx_k$, so

$$dy_i = dx_i + \frac{\partial u_i}{\partial x_k} dx_k = dx_k \delta_{ij} + \frac{\partial u_i}{\partial x_k} dx_k \qquad \implies \qquad \boxed{dy_i = F_{ik}\, dx_k.}$$

3.2 The Lagrange Strain Tensor

We define the **Lagrange strain tensor** as $E_{ij} = \frac{1}{2}(F_{ki}F_{kj} - \delta_{ij})$ (motivated below by its physical meaning). Let us multiply this out completely, substituting $F_{ki} = \delta_{ki} + u_{k,i}$ and $F_{kj} = \delta_{kj} + u_{k,j}$ (note: from the deformation-gradient definition above, with the *first* subscript playing the role of the free "row" index).

Step 1: expand the product.

$$F_{ki}F_{kj} = (\delta_{ki} + u_{k,i})(\delta_{kj} + u_{k,j}) = \delta_{ki}\delta_{kj} + \delta_{ki}u_{k,j} + u_{k,i}\delta_{kj} + u_{k,i}u_{k,j}.$$

(This is just ordinary $(p + q)(r + s) = pr + ps + qr + qs$ expansion; k is summed in every term.)

Step 2: simplify each term using the sifting property (summing over k).

- $\delta_{ki}\delta_{kj}$: sum over k; only $k = i$ survives in the first factor, forcing δ_{ij} in what remains: $\delta_{ki}\delta_{kj} = \delta_{ij}$.
- $\delta_{ki}u_{k,j}$: sum over k; only $k = i$ survives: $\delta_{ki}u_{k,j} = u_{i,j}$.
- $u_{k,i}\delta_{kj}$: sum over k; only $k = j$ survives: $u_{k,i}\delta_{kj} = u_{j,i}$.
- $u_{k,i}u_{k,j}$: no δ to simplify — this term stays as a sum over k.

So

$$F_{ki}F_{kj} = \delta_{ij} + u_{i,j} + u_{j,i} + u_{k,i}u_{k,j}.$$

Step 3: substitute back and simplify.

$$E_{ij} = \tfrac{1}{2}(F_{ki}F_{kj} - \delta_{ij}) = \tfrac{1}{2}\big[\delta_{ij} + u_{i,j} + u_{j,i} + u_{k,i}u_{k,j} - \delta_{ij}\big] = \tfrac{1}{2}\big[u_{i,j} + u_{j,i} + u_{k,i}u_{k,j}\big].$$

Physical meaning: for a line element of original length ℓ_0 stretched (and rotated) to length $\ell = \ell_0 + \delta\ell$, the Lagrange strain is $E_L \equiv \frac{\ell^2 - \ell_0^2}{2\ell_0^2} = \frac{1}{2}\left(\frac{\ell^2}{\ell_0^2} - 1\right)$, and since $\ell^2 - \ell_0^2 = \delta\ell^2 + 2\ell_0\delta\ell$,

$$E_L = \frac{\delta\ell^2}{2\ell_0^2} + \frac{\delta\ell}{\ell_0}.$$

Small strain approximation. If the displacement gradients are small, say $|u_{i,j}| \sim \epsilon \ll 1$ for every i, j (using ϵ loosely here as "a small number," not the permutation symbol), then the linear terms $u_{i,j}, u_{j,i}$ in E_{ij} are of order ϵ, while the product term $u_{k,i}u_{k,j}$ is of order ϵ^2 — a product of two small quantities is second-order small, i.e. negligible compared to the linear terms when $\epsilon \ll 1$ (e.g. if $\epsilon = 0.001$, then $\epsilon^2 = 0.000001$, a thousand times smaller still). Dropping this quadratic term,

$$\boxed{\epsilon_{ij} = \tfrac{1}{2}(u_{i,j} + u_{j,i})} \qquad \text{(infinitesimal strain tensor)}.$$

Exercise (all equal for $\ell = \ell_0 + \delta\ell$, $\delta\ell \ll \ell_0$): engineering strain $\epsilon = (\ell - \ell_0)/\ell_0$; Cauchy strain $\epsilon = \frac{1}{2}(\ell^2 - \ell_0^2)/\ell^2$; log strain $\epsilon = \ln(\ell/\ell_0)$.

3.3 Deformation Tensor: Strain + Spin

A general algebraic fact. *Any* array A_{ij} (not necessarily strain-related) can be split into a symmetric part and an anti-symmetric part, by the trivial identity

$$A_{ij} = \tfrac{1}{2}A_{ij} + \tfrac{1}{2}A_{ij} = \tfrac{1}{2}A_{ij} + \tfrac{1}{2}A_{ji} + \tfrac{1}{2}A_{ij} - \tfrac{1}{2}A_{ji} = \underbrace{\tfrac{1}{2}(A_{ij} + A_{ji})}_{\text{symmetric}} + \underbrace{\tfrac{1}{2}(A_{ij} - A_{ji})}_{\text{antisymmetric}}$$

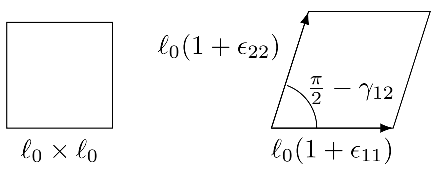
(we simply added and subtracted $\frac{1}{2}A_{ji}$, which changes nothing, and regrouped). Call the first piece $S_{ij} = \frac{1}{2}(A_{ij} + A_{ji})$: swapping i, j leaves it unchanged ($S_{ji} = S_{ij}$), so it is symmetric. Call the second piece $W_{ij} = \frac{1}{2}(A_{ij} - A_{ji})$: swapping i, j flips its sign ($W_{ji} = -W_{ij}$), so it is antisymmetric. This decomposition is always possible and always unique.

Applying this general fact to the displacement gradient $A_{ij} = u_{i,j} = F_{ij}$:

$$F_{ij} = u_{i,j} = \underbrace{\tfrac{1}{2}(u_{i,j} + u_{j,i})}_{\epsilon_{ij},\ \text{symmetric}} + \underbrace{\tfrac{1}{2}(u_{i,j} - u_{j,i})}_{\omega_{ij},\ \text{anti-symmetric (spin tensor)}} .$$

The symmetric part is exactly the infinitesimal strain tensor defined above; the antisymmetric part is new, and is examined next.

Engineering strain (used in FEM codes etc.): $\gamma_{ij} = \epsilon_{ij}$ for $i = j$; $\gamma_{ij} = 2\epsilon_{ij}$ for $i \neq j$.



3.4 The Spin (Vorticity) Tensor

$\omega_{ij} = -\omega_{ji}$ is skew-symmetric and represents rotation through a small angle $\phi = \sqrt{\frac{1}{2}\omega_{ij}\omega_{ij}}$ about axis $q_i = \epsilon_{ijk}\omega_{kj}/(2\phi)$. Recalling $dy_i = F_{ij}dx_j = dx_i + (\epsilon_{ij} + \omega_{ij})dx_j$: **the deformation tensor** $F_{ij} = \epsilon_{ij} + \omega_{ij}$ **represents a stretch and a rotation.**

3.5 Strain Compatibility

In 2D there are 2 displacement components (u_1, u_2) but 3 independent strain components ($\epsilon_{11}, \epsilon_{22}, \epsilon_{12}$), each obtained by differentiating the *same* two functions u_1, u_2:

$$\epsilon_{11} = \frac{\partial u_1}{\partial x_1}, \qquad \epsilon_{22} = \frac{\partial u_2}{\partial x_2}, \qquad \epsilon_{12} = \tfrac{1}{2}\left(\frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1}\right).$$

Since 3 functions are all built from only 2 underlying displacement functions, they cannot be prescribed independently — there must be a hidden algebraic relation among their derivatives. To find it, differentiate each strain component twice, choosing the derivatives so that the same third derivatives of u_1, u_2 appear and can be matched up:

$$\frac{\partial^2\epsilon_{11}}{\partial x_2^2} = \frac{\partial^3 u_1}{\partial x_1 \partial x_2^2}, \qquad \frac{\partial^2\epsilon_{22}}{\partial x_1^2} = \frac{\partial^3 u_2}{\partial x_2 \partial x_1^2}, \qquad \frac{\partial^2\epsilon_{12}}{\partial x_1 \partial x_2} = \tfrac{1}{2}\left(\frac{\partial^3 u_1}{\partial x_1 \partial x_2^2} + \frac{\partial^3 u_2}{\partial x_1^2 \partial x_2}\right).$$

(The last one differentiates ϵ_{12} once with respect to x_1 and once with respect to x_2, applied to each of its two terms.) Now form the specific combination $\dfrac{\partial^2\epsilon_{11}}{\partial x_2^2} - 2\dfrac{\partial^2\epsilon_{12}}{\partial x_1 \partial x_2} + \dfrac{\partial^2\epsilon_{22}}{\partial x_1^2}$ and substitute:

$$\frac{\partial^3 u_1}{\partial x_1 \partial x_2^2} - 2 \cdot \tfrac{1}{2}\left(\frac{\partial^3 u_1}{\partial x_1 \partial x_2^2} + \frac{\partial^3 u_2}{\partial x_1^2 \partial x_2}\right) + \frac{\partial^3 u_2}{\partial x_2 \partial x_1^2} = \frac{\partial^3 u_1}{\partial x_1 \partial x_2^2} - \frac{\partial^3 u_1}{\partial x_1 \partial x_2^2} - \frac{\partial^3 u_2}{\partial x_1^2 \partial x_2} + \frac{\partial^3 u_2}{\partial x_1^2 \partial x_2} = 0,$$

since each pair of terms is identical (mixed partial derivatives commute, e.g. $\partial^3 u_2/\partial x_2 \partial x_1^2 = \partial^3 u_2/\partial x_1^2 \partial x_2$) and cancels. This proves the **2D compatibility equation**:

$$\boxed{\frac{\partial^2 \epsilon_{11}}{\partial x_2^2} - 2\frac{\partial^2 \epsilon_{12}}{\partial x_1 \partial x_2} + \frac{\partial^2 \epsilon_{22}}{\partial x_1^2} = 0.}$$

If you are handed three arbitrary functions and call them $\epsilon_{11}, \epsilon_{22}, \epsilon_{12}$, they will *not* generally satisfy this identity — and if they don't, no displacement field (u_1, u_2) can have produced them. Compatibility is precisely the condition that lets you go backwards from a candidate strain field to an actual displacement field.

In general (Saint-Venant/Beltrami compatibility):

$$\epsilon_{ikr}\epsilon_{jls}\,\epsilon_{ij,kl} = 0 \qquad \Longleftrightarrow \qquad \epsilon_{ij,kl} + \epsilon_{kl,ij} = \epsilon_{ik,jl} + \epsilon_{jl,ik}, \qquad \text{i.e.} \qquad \nabla \times (\nabla \times \epsilon) = 0.$$

Proof, worked out. Suppose a displacement field $\mathbf{u}$ *does* exist, with $u_{i,j} = \epsilon_{ij} + \omega_{ij}$ as in the previous subsection; we derive a necessary condition on ϵ_{ij} alone (with ω_{ij} eliminated), which is exactly the compatibility equation.

Step 1: differentiate the spin tensor. From $\omega_{ij} = \frac{1}{2}(u_{i,j} - u_{j,i})$,

$$\omega_{ij,k} = \tfrac{1}{2}(u_{i,jk} - u_{j,ik}).$$

Add and subtract $\frac{1}{2}u_{k,ji}$ (a legal move — it changes nothing):

$$\omega_{ij,k} = \tfrac{1}{2}\big(u_{i,jk} + u_{k,ji}\big) - \tfrac{1}{2}\big(u_{j,ik} + u_{k,ji}\big).$$

Now recognize each bracket as (twice) a strain derivative. Since $\epsilon_{ik} = \frac{1}{2}(u_{i,k} + u_{k,i})$, differentiating with respect to x_j gives $\epsilon_{ik,j} = \frac{1}{2}(u_{i,kj} + u_{k,ij})$; and since partial derivatives commute ($u_{i,kj} = u_{i,jk}$, $u_{k,ij} = u_{k,ji}$), this is exactly $\frac{1}{2}(u_{i,jk} + u_{k,ji})$ — the first bracket. Similarly the second bracket, $\frac{1}{2}(u_{j,ik} + u_{k,ji})$, equals $\epsilon_{kj,i}$ (from $\epsilon_{kj} = \frac{1}{2}(u_{k,j} + u_{j,k})$, differentiated w.r.t. x_i, using $u_{k,ji} = u_{k,ij}$). So

$$\boxed{\omega_{ij,k} = \epsilon_{ik,j} - \epsilon_{kj,i}.}$$

This already says something useful: the derivatives of the (a priori unknown) spin tensor are completely determined by derivatives of the strain tensor.

Step 2: use that ω_{ij} has well-defined second derivatives. If ω_{ij} is twice continuously differentiable, then (as for any such function) its mixed second partials commute: $\omega_{ij,kl} = \omega_{ij,lk}$ (differentiating first with respect to x_l then x_k gives the same result as the other order). Differentiate the boxed result of Step 1 with respect to x_l:

$$\omega_{ij,kl} = \epsilon_{ik,jl} - \epsilon_{kj,il}.$$

Swap the roles of k and l in the same formula (i.e. differentiate $\omega_{ij,l} = \epsilon_{il,j} - \epsilon_{lj,i}$ with respect to x_k instead):

$$\omega_{ij,lk} = \epsilon_{il,jk} - \epsilon_{lj,ik}.$$

Since $\omega_{ij,kl} = \omega_{ij,lk}$, the right-hand sides must be equal:

$$\epsilon_{ik,jl} - \epsilon_{kj,il} = \epsilon_{il,jk} - \epsilon_{lj,ik}.$$

Rearranging (moving the negative terms to the other side),

$$\epsilon_{ik,jl} + \epsilon_{lj,ik} = \epsilon_{il,jk} + \epsilon_{kj,il},$$

which, relabeling the dummy/free letters to match the standard form, is exactly the Saint-Venant compatibility relation

$$\epsilon_{ij,kl} + \epsilon_{kl,ij} = \epsilon_{ik,jl} + \epsilon_{jl,ik}$$

quoted above. This is a *necessary* condition for a displacement field to exist (we derived it by assuming one does); it can also be shown to be *sufficient* (given on a simply-connected domain), though that direction is more involved and is not needed here.

4 Constitutive Law: Hooke's Law

4.1 General Requirements

Any stress–strain law must:

1. Obey the laws of thermodynamics: the first law requires work done by stresses to be stored as recoverable internal energy or dissipated as heat; the second law requires that a closed deformation cycle (same start/end strain and internal energy, constant temperature or no heat exchange) do non-negative total work.
2. Satisfy **objectivity** (material frame indifference): under a rigid rotation R_{ij} applied quasi-statically to both the solid and its loads, the constitutive law must predict $\sigma_{ij}^{(1)} = R_{ik}\sigma_{kl}^{(0)}R_{lj}$.

4.2 Isotropic Linear Elastic Behavior

For a uniaxial test at sufficiently low stress, most materials show: reversible deformation; strain depending only on the applied stress (not rate or history); a linear stress–strain relation (true for small strain regardless of which stress/strain measure is used); and, for an **isotropic** material, no dependence of the stress–strain curve on specimen orientation. If additionally isotropic and homogeneous, uniform heating causes the specimen to change size but not shape.

4.3 Hooke's Law for Isotropic Materials

In terms of Young's modulus E and Poisson's ratio ν:

$$\begin{pmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{33}\\ \sigma_{23}\\ \sigma_{13}\\ \sigma_{12}\end{pmatrix} = \frac{E}{(1+\nu)(1-2\nu)}\begin{pmatrix}1-\nu & \nu & \nu & 0 & 0 & 0\\ \nu & 1-\nu & \nu & 0 & 0 & 0\\ \nu & \nu & 1-\nu & 0 & 0 & 0\\ 0 & 0 & 0 & \frac{1-2\nu}{2} & 0 & 0\\ 0 & 0 & 0 & 0 & \frac{1-2\nu}{2} & 0\\ 0 & 0 & 0 & 0 & 0 & \frac{1-2\nu}{2}\end{pmatrix}\begin{pmatrix}\epsilon_{11}\\ \epsilon_{22}\\ \epsilon_{33}\\ 2\epsilon_{23}\\ 2\epsilon_{13}\\ 2\epsilon_{12}\end{pmatrix}.$$

More conveniently, in index notation:

$$\boxed{\sigma_{ij} = \frac{E}{1+\nu}\left[\epsilon_{ij} + \frac{\nu}{1-2\nu}\epsilon_{kk}\delta_{ij}\right]}$$

Health warning: note the factor of 2 on the shear-strain entries above; not all texts/FEM codes use this convention, though it makes no difference for isotropic materials.

Checking the index-notation form against the matrix form. Let's verify two entries explicitly.

Normal component, $i = j = 1$. The index formula gives (recalling $\epsilon_{kk} = \epsilon_{11} + \epsilon_{22} + \epsilon_{33}$ and $\delta_{11} = 1$):

$$\sigma_{11} = \frac{E}{1+\nu}\left[\epsilon_{11} + \frac{\nu}{1-2\nu}(\epsilon_{11} + \epsilon_{22} + \epsilon_{33})\right].$$

Put everything over the common denominator $(1 - 2\nu)$:

$$\begin{aligned}\sigma_{11} &= \frac{E}{(1+\nu)(1-2\nu)}\Big[(1-2\nu)\epsilon_{11} + \nu(\epsilon_{11} + \epsilon_{22} + \epsilon_{33})\Big]\\ &= \frac{E}{(1+\nu)(1-2\nu)}\Big[(1-2\nu+\nu)\epsilon_{11} + \nu\epsilon_{22} + \nu\epsilon_{33}\Big],\end{aligned}$$

i.e.

$$\sigma_{11} = \frac{E}{(1+\nu)(1-2\nu)}\Big[(1-\nu)\epsilon_{11} + \nu\epsilon_{22} + \nu\epsilon_{33}\Big],$$

which is exactly the first row of the matrix equation above. ✓

Shear component, $i = 1, j = 2$ (so $i \neq j$, hence $\delta_{ij} = 0$). The index formula gives, since $\delta_{12} = 0$ kills the volumetric term entirely,

$$\sigma_{12} = \frac{E}{1+\nu}\,\epsilon_{12}.$$

Compare with the matrix form's last row: $\sigma_{12} = \dfrac{E}{(1+\nu)(1-2\nu)} \cdot \dfrac{1-2\nu}{2} \cdot (2\epsilon_{12}) = \dfrac{E}{1+\nu}\epsilon_{12}$. ✓ (The factor of 2 in the matrix form's strain vector, $2\epsilon_{12}$, exactly cancels the $\frac{1}{2}$ in the matrix's shear-row coefficient, reproducing the same coefficient $E/(1+\nu)$ as the compact index form — this is the origin of the "health warning" above.)

4.4 Elastic Modulus and Compliance Tensors

We can rewrite Hooke's law $\sigma_{ij} = \frac{E}{1+\nu}[\epsilon_{ij} + \frac{\nu}{1-2\nu}\epsilon_{kk}\delta_{ij}]$ in the manifestly linear form $\sigma_{ij} = C_{ijkl}\epsilon_{kl}$, where C_{ijkl} is a fourth-order tensor (16 free-index slots collapsed to a rank-4 object) that does not depend on ϵ at all — it is a fixed property of the material, and σ_{ij} is obtained from ϵ_{kl} by a single linear (tensor) map.

Deriving C_{ijkl}. We need to write both terms of Hooke's law with ϵ_{kl} (not ϵ_{ij} or ϵ_{kk}) appearing explicitly, so that we can factor it out. Using the sifting property in reverse,

$$\epsilon_{ij} = \delta_{ik}\delta_{jl}\,\epsilon_{kl}, \qquad \epsilon_{kk}\delta_{ij} = \delta_{ij}\delta_{kl}\,\epsilon_{kl}$$

(the first just relabels ϵ_{ij} using dummy sums over k, l that collapse straight back via the two δ's; the second does the same for the trace ϵ_{kk}). Substituting,

$$\sigma_{ij} = \frac{E}{1+\nu}\left[\delta_{ik}\delta_{jl} + \frac{\nu}{1-2\nu}\delta_{ij}\delta_{kl}\right]\epsilon_{kl}.$$

This already has the form $\sigma_{ij} = C_{ijkl}\epsilon_{kl}$ with $C_{ijkl} = \frac{E}{1+\nu}\delta_{ik}\delta_{jl} + \frac{E\nu}{(1+\nu)(1-2\nu)}\delta_{ij}\delta_{kl}$. However, this candidate C_{ijkl} is *not symmetric under* $k \leftrightarrow l$ (swap k, l in $\delta_{ik}\delta_{jl}$ and you get $\delta_{il}\delta_{jk} \neq \delta_{ik}\delta_{jl}$ in general), even though the strain $\epsilon_{kl} = \epsilon_{lk}$ it multiplies *is* symmetric. By exactly the symmetric/anti-symmetric argument from Section 1.5: since ϵ_{kl} is symmetric, only the part of C_{ijkl} that is symmetric in (k, l) actually contributes to the contraction $C_{ijkl}\epsilon_{kl}$ — any antisymmetric-in-(k, l) piece would be contracted away to zero. So we may freely replace the first term by its symmetrization in (k, l) without changing the physics:

$$\delta_{ik}\delta_{jl}\,\epsilon_{kl} = \delta_{il}\delta_{jk}\,\epsilon_{kl}$$

(just relabel the dummy indices $k \leftrightarrow l$ inside the sum — allowed since both are summed and appear nowhere else — and use $\epsilon_{lk} = \epsilon_{kl}$), so

$$\delta_{ik}\delta_{jl}\,\epsilon_{kl} = \tfrac{1}{2}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})\,\epsilon_{kl}.$$

Using this symmetrized version gives the standard, manifestly (k, l)-symmetric elastic modulus tensor:

$$\boxed{\sigma_{ij} = C_{ijkl}\epsilon_{kl}, \qquad C_{ijkl} = \frac{E}{2(1+\nu)}(\delta_{il}\delta_{jk} + \delta_{ik}\delta_{jl}) + \frac{E}{(1+\nu)(1-2\nu)}\delta_{ij}\delta_{kl}.}$$

Inverting this relation (solving for strain in terms of stress; the derivation is the analogous algebra run in reverse and is not repeated here) gives the compliance tensor S_{ijkl}:

$$\epsilon_{ij} = S_{ijkl}\sigma_{kl}, \qquad S_{ijkl} = \frac{1+\nu}{2E}(\delta_{il}\delta_{jk} + \delta_{ik}\delta_{jl}) - \frac{\nu}{E}\delta_{ij}\delta_{kl}.$$

4.5 Other Elastic Constants: Bulk, Shear, and Lamé Moduli

Rather than simply quoting these, let's derive each one directly from Hooke's law.

Bulk modulus K. By definition, K relates the mean (hydrostatic) stress to the volumetric strain: $\sigma_{\text{mean}} \equiv \frac{1}{3}\sigma_{kk} = K\,\epsilon_{kk}$ (recall $\epsilon_{kk} = \Delta V/V$ is the fractional volume change, for small strain). To get σ_{kk}, take the trace of Hooke's law: set

$j = i$ in $\sigma_{ij} = \frac{E}{1+\nu}[\epsilon_{ij} + \frac{\nu}{1-2\nu}\epsilon_{kk}\delta_{ij}]$ and sum over i:

$$\sigma_{ii} = \frac{E}{1+\nu}\left[\epsilon_{ii} + \frac{\nu}{1-2\nu}\epsilon_{kk}\,\delta_{ii}\right].$$

Here ϵ_{ii} and ϵ_{kk} are the *same* quantity (just different dummy-index names for the same sum, $\epsilon_{11}+\epsilon_{22}+\epsilon_{33}$), and $\delta_{ii} = 3$ (derived earlier). So

$$\sigma_{ii} = \frac{E}{1+\nu}\left[\epsilon_{kk} + \frac{3\nu}{1-2\nu}\epsilon_{kk}\right] = \frac{E}{1+\nu}\,\epsilon_{kk}\left[\frac{1-2\nu+3\nu}{1-2\nu}\right] = \frac{E}{1+\nu}\,\epsilon_{kk}\cdot\frac{1+\nu}{1-2\nu} = \frac{E}{1-2\nu}\,\epsilon_{kk}.$$

Dividing by 3, $\frac{1}{3}\sigma_{kk} = \dfrac{E}{3(1-2\nu)}\,\epsilon_{kk}$, and comparing with the definition $\sigma_{\text{mean}} = K\epsilon_{kk}$:

$$\boxed{K = \frac{E}{3(1-2\nu)}.}$$

Shear modulus μ. By definition, pure shear stress and shear strain are related by $\sigma_{ij} = 2\mu\,\epsilon_{ij}$ for $i \neq j$ (equivalently $\sigma_{ij} = \mu\gamma_{ij}$ using the engineering shear strain $\gamma_{ij} = 2\epsilon_{ij}$ from Section 3.3). But we already computed, in checking the shear component above, that Hooke's law gives $\sigma_{ij} = \frac{E}{1+\nu}\epsilon_{ij}$ for $i \neq j$ (the δ_{ij} term vanishes). Comparing $2\mu\,\epsilon_{ij}$ with $\frac{E}{1+\nu}\epsilon_{ij}$:

$$2\mu = \frac{E}{1+\nu} \qquad \Longrightarrow \qquad \boxed{\mu = \frac{E}{2(1+\nu)}.}$$

Lamé modulus λ. Hooke's law is often written in the alternative "Lamé form" $\sigma_{ij} = \lambda\epsilon_{kk}\delta_{ij} + 2\mu\epsilon_{ij}$. Comparing term by term with our boxed Hooke's law $\sigma_{ij} = \frac{E}{1+\nu}\epsilon_{ij} + \frac{E\nu}{(1+\nu)(1-2\nu)}\epsilon_{kk}\delta_{ij}$: the ϵ_{ij} coefficient must match 2μ (already confirmed above, $2\mu = E/(1+\nu)$ ✓), and the $\epsilon_{kk}\delta_{ij}$ coefficient must match λ:

$$\boxed{\lambda = \frac{\nu E}{(1+\nu)(1-2\nu)}.}$$

Physical interpretation.

- E: slope of the stress–strain curve in uniaxial tension; dimensions of stress; a measure of stiffness ($E > 0$ for a stable material). E.g. steel: $E \approx 210 \times 10^9\ \mathrm{N/m^2}$.
- ν: ratio of lateral to longitudinal strain in uniaxial tension; dimensionless, typically 0.2–0.49 (≈ 0.3 for most metals); a measure of compressibility. $\nu = 0.5$: incompressible. $\nu = 0$: no lateral contraction (cork). $\nu < 0$: *auxetic* materials (bar widens under stretching). Stability requires $-1 < \nu < 1/2$.
- Shear modulus: resistance to volume-preserving shear deformation; usually somewhat smaller than E.

By the end of this lecture, we have assembled the following toolkit for earthquake source mechanics:

- **History:** continuum mechanics rests on three pillars — momentum balance (Euler), strain–displacement geometry (Bernoulli), and constitutive relations (Euler, Green, Kelvin) — unified through the continuum hypothesis.
- **Index notation & tensors:** the summation convention, δ_{ij}, ϵ_{ijk}, the ϵ–δ identity, symmetry/anti-symmetry, vector calculus in index form, integral theorems, and the basic algebra of second-order tensors (including coordinate rotation).
- **Stress:** the traction vector, Cauchy's tetrahedron argument giving $\sigma_{ij} = T_j(\mathbf{e}_i)$, balance of linear momentum ($\partial\sigma_{ij}/\partial x_j + b_i = \rho a_i$), balance of angular momentum ($\sigma_{ij} = \sigma_{ji}$), principal stresses, and Mohr's circle.
- **Strain:** the deformation gradient F_{ij}, the (Lagrange and infinitesimal) strain tensor ϵ_{ij}, its decomposition from the spin tensor ω_{ij}, and the compatibility conditions relating strain components.
- **Constitutive law:** thermodynamic and objectivity requirements, and Hooke's law for isotropic linear elasticity, $\sigma_{ij} = C_{ijkl}\epsilon_{kl}$, in terms of (E, ν) or equivalently (λ, μ, K).

Table 1: Elastic constants for representative isotropic materials (Ashby & Jones, 1997, and manufacturer data)

Material	ρ (kg m^{-3})	E (GPa)	ν
Tungsten carbide	14000–17000	450–650	0.22
Tungsten	13400	410	0.30
Nickel	8900	215	0.31
Low alloy steels	7800	200–210	0.30
Iron	7900	196	0.30
Copper	8900	124	0.34
Titanium	4500	116	0.30
Aluminum and alloys	2600–2900	69–79	0.35
Concrete	2400–2500	45–50	0.30
Wood (parallel grain)	400–800	9–16	0.20
PMMA	1200	3.4	0.35–0.4
Natural rubbers	830–910	0.01–0.1	0.49
Representative crustal rocks			
Granite	2600–2700	40–60	0.20–0.25
Basalt	2700–3000	60–90	0.25
Gabbro	2900–3100	70–100	0.25
Sandstone	2000–2600	10–25	0.15–0.25
Limestone	2300–2700	30–70	0.25–0.30

Table 2: Relations between elastic constants

Given	Shear mod. μ	Young's mod. E	Poisson ratio ν	Bulk mod. K
λ, μ	—	$\dfrac{\mu(3\lambda+2\mu)}{\lambda+\mu}$	$\dfrac{\lambda}{2(\lambda+\mu)}$	$\lambda+\dfrac{2\mu}{3}$
λ, ν	$\dfrac{\lambda(1-2\nu)}{2\nu}$	$\dfrac{\lambda(1+\nu)(1-2\nu)}{\nu}$	—	$\dfrac{\lambda(1+\nu)}{3\nu}$
μ, ν	—	$2\mu(1+\nu)$	—	$\dfrac{2\mu(1+\nu)}{3(1-2\nu)}$
E, ν	$\dfrac{E}{2(1+\nu)}$	—	—	$\dfrac{E}{3(1-2\nu)}$

5 Recap: Momentum Balance and Hooke's Law

Two key results from Lecture 1 are all we need to work in this lecture; we simply restate them, without re-deriving.

Balance of linear momentum (from the Cauchy tetrahedron argument and its localization):

$$\boxed{\frac{\partial \sigma_{ij}}{\partial x_j} + b_i = \rho a_i}, \qquad (i = 1, 2, 3).$$

Hooke's law for an isotropic linear elastic solid (Section 4.3), most conveniently in its *Lamé form*:

$$\boxed{\sigma_{ij} = \lambda\,\epsilon_{kk}\delta_{ij} + 2\mu\,\epsilon_{ij}}, \qquad \epsilon_{ij} = \tfrac{1}{2}(u_{i,j} + u_{j,i}),$$

with λ, μ the Lamé parameters (μ doubling as the shear modulus). Note what we do *not* yet have: these are two separate statements, one purely kinematic/dynamical (momentum balance says nothing about the material), the other purely material (Hooke's law says nothing about motion). Today we combine them into a single equation for the displacement field $\mathbf{u}(\mathbf{x}, t)$ alone, and show that equation supports two families of waves.

6 Deriving Navier's Equation

6.1 Substituting Hooke's Law into the Equations of Motion

The idea is simple: momentum balance is a statement about σ_{ij}; Hooke's law tells us what σ_{ij} *is*, in terms of $\mathbf{u}$. Substituting one into the other eliminates σ_{ij} entirely, leaving one equation in the one unknown field $\mathbf{u}(\mathbf{x}, t)$.

Step 1: differentiate Hooke's law. We need $\partial\sigma_{ij}/\partial x_j$, so differentiate the boxed Lamé form with respect to x_j (summed, as momentum balance requires):

$$\frac{\partial \sigma_{ij}}{\partial x_j} = \lambda\,\delta_{ij}\frac{\partial \epsilon_{kk}}{\partial x_j} + 2\mu\,\frac{\partial \epsilon_{ij}}{\partial x_j}.$$

In the first term, apply the sifting property of δ_{ij} (summed over j): $\delta_{ij}\,\partial\epsilon_{kk}/\partial x_j = \partial\epsilon_{kk}/\partial x_i$. So

$$\frac{\partial \sigma_{ij}}{\partial x_j} = \lambda\,\frac{\partial \epsilon_{kk}}{\partial x_i} + 2\mu\,\frac{\partial \epsilon_{ij}}{\partial x_j}.$$

Step 2: write both terms in terms of $\mathbf{u}$. Since $\epsilon_{kk} = u_{k,k}$ (the trace of the strain tensor is the displacement divergence),

$$\frac{\partial \epsilon_{kk}}{\partial x_i} = u_{k,ki}.$$

For the second term, substitute $\epsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i})$ and differentiate:

$$\frac{\partial \epsilon_{ij}}{\partial x_j} = \tfrac{1}{2}\big(u_{i,jj} + u_{j,ij}\big).$$

Mixed partial derivatives commute, so $u_{j,ij} = u_{j,ji}$; relabeling the dummy index $j \to k$ (allowed — it is summed and appears nowhere else) gives $u_{j,ji} = u_{k,ki}$, the *same* quantity computed in Step 2 above. So

$$\frac{\partial \epsilon_{ij}}{\partial x_j} = \tfrac{1}{2}\big(u_{i,jj} + u_{k,ki}\big).$$

Step 3: assemble. Substituting both pieces back:

$$\frac{\partial \sigma_{ij}}{\partial x_j} = \lambda\,u_{k,ki} + 2\mu\cdot\tfrac{1}{2}\big(u_{i,jj} + u_{k,ki}\big) = (\lambda + \mu)\,u_{k,ki} + \mu\,u_{i,jj}.$$

Substitute into momentum balance, $\partial\sigma_{ij}/\partial x_j + b_i = \rho a_i$, writing the acceleration as $a_i = \ddot{u}_i \equiv \partial^2 u_i/\partial t^2$ (there is no large-scale motion of the medium; the only motion is the small oscillation $\mathbf{u}$ itself describes):

$$\boxed{(\lambda+\mu)\, u_{k,ki} + \mu\, u_{i,jj} + b_i = \rho\, \ddot{u}_i} \qquad \textbf{(Navier's equation).}$$

In vector notation, using $u_{k,ki} = [\nabla(\nabla\cdot\mathbf{u})]_i$ and $u_{i,jj} = [\nabla^2\mathbf{u}]_i$ (the Laplacian, applied component-wise):

$$(\lambda+\mu)\,\nabla(\nabla\cdot\mathbf{u}) + \mu\nabla^2\mathbf{u} + \mathbf{b} = \rho\,\ddot{\mathbf{u}}.$$

This is a single vector equation (three scalar equations, $i = 1, 2, 3$) in the one unknown field $\mathbf{u}(\mathbf{x}, t)$: the complete statement of linear elastodynamics for a homogeneous, isotropic solid.

6.2 An Equivalent Curl Form

A useful rewriting uses the vector identity $\nabla^2\mathbf{u} = \nabla(\nabla\cdot\mathbf{u}) - \nabla\times(\nabla\times\mathbf{u})$ (true for any vector field, not special to elasticity). Substituting into the boxed vector form above (source-free, $\mathbf{b} = 0$):

$$(\lambda+\mu)\nabla(\nabla\cdot\mathbf{u}) + \mu\big[\nabla(\nabla\cdot\mathbf{u}) - \nabla\times(\nabla\times\mathbf{u})\big] = \rho\ddot{\mathbf{u}},$$

so

$$\boxed{(\lambda+2\mu)\,\nabla(\nabla\cdot\mathbf{u}) - \mu\,\nabla\times(\nabla\times\mathbf{u}) = \rho\,\ddot{\mathbf{u}}.}$$

This is the form that separates most cleanly into two independent wave equations, which is what we do next.

7 From Navier's Equation to P and S Waves

7.1 Helmholtz Decomposition

Claim. Any sufficiently smooth vector field $\mathbf{u}(\mathbf{x})$ can be written as

$$\mathbf{u} = \nabla\phi + \nabla\times\boldsymbol{\psi}, \qquad \nabla\cdot\boldsymbol{\psi} = 0,$$

for some scalar field $\phi(\mathbf{x}, t)$ (the *dilatational*, or P-wave, potential) and vector field $\boldsymbol{\psi}(\mathbf{x}, t)$ (the *shear*, or S-wave, vector potential), subject to the condition $\nabla\cdot\boldsymbol{\psi} = 0$. (This removes a redundancy: $\boldsymbol{\psi}$ has 3 components, but $\nabla\times\boldsymbol{\psi}$ only needs 2 independent degrees of freedom to match the transverse part of $\mathbf{u}$, so one constraint is imposed by hand; it costs nothing physically, since only $\nabla\times\boldsymbol{\psi}$, never $\boldsymbol{\psi}$ itself, appears in what follows.) We do not prove existence/uniqueness here (it is a standard result, Helmholtz's theorem) — we use it as an ansatz and check that it works.

Geometrical intuition: At any point, a vector field can only do one of two things: spread — flow outward or inward, like water at a source or a drain — or swirl, circulating around that point like a whirlpool. There's no third kind of motion. Divergence measures the spreading; curl measures the swirling.

Why this split is useful. Compute the divergence and curl of $\mathbf{u}$ under the decomposition:

$$\nabla\cdot\mathbf{u} = \nabla\cdot\nabla\phi + \nabla\cdot(\nabla\times\boldsymbol{\psi}) = \nabla^2\phi + 0 = \nabla^2\phi$$

(using $\nabla\cdot(\nabla\times(\text{anything})) \equiv 0$, a general vector identity), and

$$\nabla\times\mathbf{u} = \nabla\times\nabla\phi + \nabla\times(\nabla\times\boldsymbol{\psi}) = 0 + \big[\nabla(\nabla\cdot\boldsymbol{\psi}) - \nabla^2\boldsymbol{\psi}\big] = -\nabla^2\boldsymbol{\psi}$$

(using $\nabla\times\nabla(\text{anything}) \equiv 0$, and the gauge condition $\nabla\cdot\boldsymbol{\psi} = 0$). So: ϕ alone controls the dilatation ($\nabla\cdot\mathbf{u}$), and $\boldsymbol{\psi}$ alone controls the rotation ($\nabla\times\mathbf{u}$) — exactly the strain/spin split of Lecture 1, Section 3.3, now applied to the wave problem.

7.2 Two Decoupled Wave Equations

Substitute $\mathbf{u} = \nabla\phi + \nabla\times\boldsymbol{\psi}$ into the boxed curl form of Navier's equation, using $\nabla\cdot\mathbf{u} = \nabla^2\phi$ and $\nabla\times\mathbf{u} = -\nabla^2\boldsymbol{\psi}$ just derived:

$$(\lambda + 2\mu)\,\nabla(\nabla^2\phi) - \mu\,\nabla\times(-\nabla^2\boldsymbol{\psi}) = \rho\left(\nabla\ddot{\phi} + \nabla\times\ddot{\boldsymbol{\psi}}\right).$$

Group the gradient terms and the curl terms separately:

$$\nabla\Big[(\lambda+2\mu)\nabla^2\phi - \rho\ddot{\phi}\Big] + \nabla\times\Big[\mu\nabla^2\boldsymbol{\psi} - \rho\ddot{\boldsymbol{\psi}}\Big] = 0.$$

A pure gradient field and a pure curl field cannot cancel each other in general — this is the same independence that underlies the Helmholtz decomposition's uniqueness — so each bracket must vanish separately:

$$\boxed{\nabla^2\phi = \frac{1}{\alpha^2}\ddot{\phi}}, \qquad \alpha^2 \equiv \frac{\lambda+2\mu}{\rho},$$

$$\boxed{\nabla^2\boldsymbol{\psi} = \frac{1}{\beta^2}\ddot{\boldsymbol{\psi}}}, \qquad \beta^2 \equiv \frac{\mu}{\rho}.$$

Two *scalar* wave equations (the second holding component-wise), completely decoupled. ϕ propagates at speed α; $\boldsymbol{\psi}$ at speed β. Since $\lambda, \mu > 0$ for a stable material, $\alpha > \beta$ always: the dilatational wave is always faster. (Numerically $\alpha/\beta = \sqrt{2(1-\nu)/(1-2\nu)}$, which is $\sqrt{3} \approx 1.73$ at $\nu = 1/4$ — a convenient rule of thumb.)

8 Interlude: The One-Dimensional Wave Equation

Before jumping to the full three-dimensional solutions, it is worth pausing on the simplest possible version of the equation we just derived, and asking directly: what does a solution actually *look like*?

8.1 The Geometrical Picture First

Imagine plucking a taut string, or clapping your hands and watching the disturbance travel away. Two everyday observations about a disturbance that has propagated some distance:

- its *shape* doesn't change as it travels — a sharp clap arrives elsewhere still sounding like a sharp clap, not smeared out (ignoring dissipation and geometrical spreading, both set aside for the moment);
- it moves at a *constant speed*, fixed by the medium, regardless of how sharp or gentle the disturbance was.

That is the entire physical content of the wave equation: it is the equation whose solutions are rigid translations of an arbitrary shape, at a fixed speed set by material properties alone. We now confirm this mathematically in the simplest setting, then generalize.

8.2 Reducing to One Dimension

Restrict to a disturbance varying only along x_1 (and t), i.e. $\phi = \phi(x_1, t)$. Then $\nabla^2\phi$ collapses to $\partial^2\phi/\partial x_1^2$, and the boxed wave equation of Section 7 becomes

$$\boxed{\frac{\partial^2\phi}{\partial x_1^2} = \frac{1}{\alpha^2}\frac{\partial^2\phi}{\partial t^2}} \qquad \textbf{(1D wave equation).}$$

8.3 d'Alembert's Solution

Step 1: change variables. Introduce the **characteristic coordinates** $\xi = x_1 - \alpha t$, $\eta = x_1 + \alpha t$. By the chain rule,

$$\frac{\partial}{\partial x_1} = \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}, \qquad \frac{\partial}{\partial t} = -\alpha\frac{\partial}{\partial \xi} + \alpha\frac{\partial}{\partial \eta}.$$

Step 2: rewrite the wave equation. Applying each operator twice,

$$\frac{\partial^2 \phi}{\partial x_1^2} = \phi_{,\xi\xi} + 2\phi_{,\xi\eta} + \phi_{,\eta\eta}, \qquad \frac{\partial^2 \phi}{\partial t^2} = \alpha^2\big(\phi_{,\xi\xi} - 2\phi_{,\xi\eta} + \phi_{,\eta\eta}\big).$$

Substituting into the boxed 1D wave equation and dividing by α^2, the $\phi_{,\xi\xi}$ and $\phi_{,\eta\eta}$ terms cancel identically, leaving

$$4\phi_{,\xi\eta} = 0 \qquad \Longrightarrow \qquad \boxed{\frac{\partial^2 \phi}{\partial \xi \partial \eta} = 0.}$$

Step 3: integrate. $\partial \phi_{,\xi}/\partial \eta = 0$ means $\phi_{,\xi}$ does not depend on η, so $\phi_{,\xi} = f'(\xi)$ for some function f' of ξ alone. Integrating in ξ: $\phi = f(\xi) +$ (a function of η alone) $\equiv f(\xi) + g(\eta)$. Converting back to x_1, t:

$$\boxed{\phi(x_1, t) = f(x_1 - \alpha t) + g(x_1 + \alpha t)} \qquad \textbf{(d'Alembert's solution).}$$

f and g are *arbitrary* (twice-differentiable) functions, fixed only by initial/boundary conditions — the wave equation constrains how a shape evolves, never what the shape is.

Historical Aside: Jean le Rond d'Alembert (1717–1783)

Jean le Rond d'Alembert was born in Paris in 1717, the illegitimate son of a society hostess and an artillery officer. Abandoned as an infant on the steps of the church of Saint-Jean-le-Rond — which gave him his name — he was raised by a glazier's wife while his biological father secretly funded his education. Trained in law and then medicine, he found neither congenial and taught himself mathematics instead, entering the Paris Academy of Sciences in 1741. His 1743 *Traité de dynamique* introduced what is now called d'Alembert's principle, recasting dynamics as a form of statics, and he went on to make foundational contributions to fluid mechanics and partial differential equations. It was his 1747 memoir on vibrating strings that produced the first appearance in print of the wave equation and its general solution — the very result derived above. Alongside Diderot, he also edited the great *Encyclopédie*, writing most of its mathematical entries. Famously combative, he feuded for decades with Clairaut and later with Euler, each accusing the other of stealing credit; contemporaries described him as brilliant but constitutionally unable to accept that he might be wrong. In later life he turned increasingly to literature and philosophy, growing frustrated that mathematics — his one true passion — had become difficult for him to pursue. He died in Paris in 1783 and, as a known religious skeptic, was buried in an unmarked grave.

8.4 What This Means

Fix your eye on a single value of the argument, $x_1 - \alpha t = c$ (constant): this traces out a point moving to the *right* at speed α. So $f(x_1 - \alpha t)$ is exactly the initial shape $f(x_1)$, translated rigidly rightward at speed α, undistorted. Likewise $g(x_1 + \alpha t)$ is a rigid left-traveling disturbance. *Every* solution of the 1D wave equation is some superposition of a right-going and a left-going disturbance, both moving at the same fixed speed α regardless of shape — this shape-independence of the speed is the **non-dispersive** property, seen here directly in physical space, with no need to even introduce a frequency.

We now generalize this shape-preserving, rigidly-translating disturbance to three dimensions — which is exactly what the plane-wave ansatz below captures, with the 1D picture's translating $f(x_1 - \alpha t)$ becoming a translating *plane* of constant phase, perpendicular to a propagation direction $\boldsymbol{k}$.

9 Plane Wave Solutions

9.1 P Waves: Longitudinal Motion

Generalizing the traveling-wave picture of Section 8 to three dimensions, try a plane-wave solution for the potential,

$$\phi(\mathbf{x},t) = A\exp\big[i(\boldsymbol{k}\cdot\mathbf{x}-\omega t)\big],$$

where $\boldsymbol{k}$ is the **wavevector** (direction of propagation, magnitude $|\boldsymbol{k}| = 2\pi/\text{wavelength}$) and ω is the **angular frequency** — not to be confused with the symbol for spin tensor ω_{ij} of Lecture 1, but the two never appear in the same context, and a bare scalar ω versus a two-index object ω_{ij} should keep them apart. Substituting into $\nabla^2\phi = \ddot{\phi}/\alpha^2$: each ∇ brings down a factor $i\boldsymbol{k}$, each $\partial/\partial t$ a factor $-i\omega$, so

$$-|\boldsymbol{k}|^2\phi = \frac{-\omega^2}{\alpha^2}\phi \qquad \Longrightarrow \qquad \boxed{\omega = \alpha|\boldsymbol{k}|}$$

(the **dispersion relation**: linear in $|\boldsymbol{k}|$, i.e. *non-dispersive* — every frequency travels at the same speed α, the same fact we already saw directly in physical space in Section 8, now recovered in frequency–wavenumber language).

Particle motion. The displacement itself is $\mathbf{u}^P = \nabla\phi$; differentiating the plane-wave ansatz,

$$\mathbf{u}^P = i\boldsymbol{k}\,A\exp\big[i(\boldsymbol{k}\cdot\mathbf{x}-\omega t)\big] = i\boldsymbol{k}\,\phi.$$

The particle displacement points along $\boldsymbol{k}$, i.e. *parallel to the direction of propagation*: a **longitudinal** wave, alternating compression and dilatation along the ray path — hence **P** for *primary* (it is the faster wave, so always arrives first) or *pressure*.

9.2 S Waves: Transverse Motion, SH/SV Polarization

Similarly try $\boldsymbol{\psi}(\mathbf{x},t) = \boldsymbol{B}\exp[i(\boldsymbol{k}\cdot\mathbf{x}-\omega t)]$ for a constant vector amplitude $\boldsymbol{B}$. The condition $\nabla\cdot\boldsymbol{\psi} = 0$ becomes, after the same substitution, $i\boldsymbol{k}\cdot\boldsymbol{B} = 0$, i.e. $\boldsymbol{B}\perp\boldsymbol{k}$. Substituting into $\nabla^2\boldsymbol{\psi} = \ddot{\boldsymbol{\psi}}/\beta^2$ gives, by identical algebra to the P-wave case,

$$\boxed{\omega = \beta|\boldsymbol{k}|}$$

(again non-dispersive, now at the slower speed β).

Particle motion. The displacement is $\mathbf{u}^S = \nabla\times\boldsymbol{\psi} = i\boldsymbol{k}\times\boldsymbol{\psi}$. A cross product with $\boldsymbol{k}$ is, by definition, always perpendicular to $\boldsymbol{k}$: $\mathbf{u}^S\perp\boldsymbol{k}$, a **transverse** wave — hence **S** for *secondary* (arrives after P) or *shear* (transverse motion is exactly a shearing of the medium with no volume change; check $\nabla\cdot\mathbf{u}^S = \nabla\cdot(\nabla\times\boldsymbol{\psi})\equiv 0$, consistent with Section 7).

Since $\boldsymbol{B}$ has two free components once constrained to the plane perpendicular to $\boldsymbol{k}$, every propagation direction supports two independent S-wave polarizations.

Constructing the SH/SV basis explicitly. The plane perpendicular to $\boldsymbol{k}$ is two-dimensional, so any admissible $\boldsymbol{B}$ is some combination of two orthonormal directions spanning that plane — but which two? The choice is not forced mathematically; physically, the natural choice is dictated by the one direction every problem singles out: vertical, $\mathbf{e}_3$ (because of the Earth's free surface, which we introduce in a later lecture). Define

$$\hat{\boldsymbol{d}}_{SH} \equiv \frac{\mathbf{e}_3\times\boldsymbol{k}}{|\mathbf{e}_3\times\boldsymbol{k}|}, \qquad \hat{\boldsymbol{d}}_{SV} \equiv \hat{\boldsymbol{k}}\times\hat{\boldsymbol{d}}_{SH},$$

where $\hat{\boldsymbol{k}} = \boldsymbol{k}/|\boldsymbol{k}|$ (this construction breaks down exactly at vertical incidence, $\boldsymbol{k}\parallel\mathbf{e}_3$, where $\mathbf{e}_3\times\boldsymbol{k} = \mathbf{0}$ — unsurprising, since straight-up propagation has no preferred horizontal direction to single out an SH axis; any horizontal direction is equally valid there by symmetry). By construction $\hat{\boldsymbol{d}}_{SH}\perp\boldsymbol{k}$ and $\hat{\boldsymbol{d}}_{SH}\perp\mathbf{e}_3$ (both immediate from the defining cross product), so $\hat{\boldsymbol{d}}_{SH}$ is *purely horizontal*, perpendicular to the vertical plane containing both $\boldsymbol{k}$ and $\mathbf{e}_3$. And $\hat{\boldsymbol{d}}_{SV}$, being perpendicular to both $\hat{\boldsymbol{k}}$ and $\hat{\boldsymbol{d}}_{SH}$, must lie *in* that vertical plane while still staying perpendicular to $\boldsymbol{k}$ itself. Together $\{\hat{\boldsymbol{d}}_{SH},\hat{\boldsymbol{d}}_{SV}\}$ form an orthonormal basis for the plane $\perp\boldsymbol{k}$, so the general S-wave amplitude is $\boldsymbol{B} = B_{SH}\hat{\boldsymbol{d}}_{SH} + B_{SV}\hat{\boldsymbol{d}}_{SV}$ for two independent scalar amplitudes — the "two free components" above, now with concrete geometric meaning.

A picture worth holding onto. Take the common case of $\boldsymbol{k}$ lying in the x_1–x_3 vertical plane ($k_2 = 0$), parameterized by an **angle of incidence** θ from the vertical, $\hat{\boldsymbol{k}} = (\sin\theta,\ 0,\ \cos\theta)$. Working out the cross products above gives, explicitly, $\hat{\boldsymbol{d}}_{SH} = (0, 1, 0) = \mathbf{e}_2$ for *every* θ, and $\hat{\boldsymbol{d}}_{SV} = (-\cos\theta,\ 0,\ \sin\theta)$. Two things fall out immediately:

- **SH** is always exactly $\mathbf{e}_2$, regardless of incidence angle — motion purely *out of* the vertical plane of propagation, like plucking a string oriented along x_2 while the wave itself travels within the x_1–x_3 plane. Nothing about the vertical-plane geometry (incidence angle, layering, a free surface) ever touches this component, which is exactly why it decouples completely from everything else in the problem.
- **SV** rotates with θ: at normal incidence ($\theta = 0$, $\boldsymbol{k}$ straight down) it works out to purely *horizontal* motion; at grazing incidence ($\theta \to 90°$) it becomes purely *vertical* motion — the opposite of what intuition might first suggest. In between it is a mixture, always tilted to stay exactly perpendicular to $\boldsymbol{k}$.

This is the real reason the split is useful, not just a bookkeeping convenience: SH motion is entirely decoupled from the vertical-plane geometry, while P and SV (both living in the x_1–x_3 plane) are coupled to each other the moment a wave meets a horizontal boundary — a fact we use directly once we study reflection and transmission at an interface, a topic for later, once we go beyond the homogeneous full space assumed throughout today.

10 Energy Density and Flux in Elastic Waves

10.1 Kinetic and Strain Energy Density

Define the **kinetic energy density** and **strain (potential) energy density** (both per unit volume):

$$T \equiv \tfrac{1}{2}\rho\,\dot{u}_i\dot{u}_i, \qquad W \equiv \tfrac{1}{2}\sigma_{ij}\epsilon_{ij} = \tfrac{1}{2}C_{ijkl}\epsilon_{ij}\epsilon_{kl}.$$

(W is the work done per unit volume in straining the material quasi-statically from zero to ϵ_{ij}, by the usual $\int \sigma\, d\epsilon = \frac{1}{2}\sigma\epsilon$ argument for a linear stress–strain relation — the same logic as the area under a linear spring's force–extension line.)

10.2 The Local Energy Balance: Umov–Poynting Flux

Step 1: multiply momentum balance by velocity. Take $\partial\sigma_{ij}/\partial x_j + b_i = \rho\ddot{u}_i$, multiply every term by $\dot{u}_i$, and sum over i:

$$\dot{u}_i\,\frac{\partial\sigma_{ij}}{\partial x_j} + b_i\dot{u}_i = \rho\,\dot{u}_i\ddot{u}_i.$$

The right-hand side is a perfect time derivative: $\rho\dot{u}_i\ddot{u}_i = \dfrac{\partial}{\partial t}\left(\frac{1}{2}\rho\dot{u}_i\dot{u}_i\right) = \dot{T}$.

Step 2: rewrite the stress term with the product rule, in reverse.

$$\frac{\partial}{\partial x_j}(\sigma_{ij}\dot{u}_i) = \frac{\partial\sigma_{ij}}{\partial x_j}\dot{u}_i + \sigma_{ij}\frac{\partial\dot{u}_i}{\partial x_j} \qquad\Longrightarrow\qquad \dot{u}_i\frac{\partial\sigma_{ij}}{\partial x_j} = \frac{\partial}{\partial x_j}(\sigma_{ij}\dot{u}_i) - \sigma_{ij}\dot{u}_{i,j}.$$

Since σ_{ij} is symmetric (Lecture 1), only the symmetric part of $\dot{u}_{i,j}$ survives the contraction with it (the symmetric/antisymmetric argument of Lecture 1, Section 1.5): $\sigma_{ij}\dot{u}_{i,j} = \sigma_{ij}\dot{\epsilon}_{ij}$. And because Hooke's law is linear and time-independent, together with the major symmetry $C_{ijkl} = C_{klij}$, this equals exactly $\dot{W}$: writing $\sigma_{ij} = C_{ijkl}\epsilon_{kl}$ and differentiating $W = \frac{1}{2}C_{ijkl}\epsilon_{ij}\epsilon_{kl}$ term by term gives $\dot{W} = \sigma_{ij}\dot{\epsilon}_{ij}$ directly (the two terms of the product rule turn out equal to each other by the major symmetry, each contributing half).

Step 3: assemble. Substituting both simplifications back:

$$\frac{\partial}{\partial x_j}(\sigma_{ij}\dot{u}_i) - \dot{W} + b_i\dot{u}_i = \dot{T} \qquad\Longrightarrow\qquad \boxed{\frac{\partial}{\partial t}(T + W) = \frac{\partial}{\partial x_j}(\sigma_{ij}\dot{u}_i) + b_i\dot{u}_i.}$$

This is a **local energy balance**: the rate of change of mechanical energy density (kinetic + strain) at a point equals the divergence of an **energy flux vector**

$$S_j \equiv \sigma_{ij}\dot{u}_i$$

(the **Umov–Poynting flux** for elastic waves, the mechanical analogue of the electromagnetic Poynting vector), plus the power delivered by any body force. Physically, S_j is the rate at which material on one side of a surface does work on material on the other side, per unit area — the traction $\sigma_{ij}n_j$ of Lecture 1, dotted with the velocity there.

A first look at geometrical spreading. For a point source radiating energy steadily into a full space, the total power crossing any sphere of radius r centered on the source must be the same at every r (energy conservation, no dissipation) — but the sphere's area grows as r^2, so the energy *flux* (power per unit area) must fall as $1/r^2$. Since energy density in a wave scales as (displacement rate)^2, this means displacement amplitude itself falls as $1/r$: the familiar geometrical-spreading factor we will use, without further comment, for the rest of the course.

By the end of this lecture, we have assembled the following toolkit:

- **Navier's equation:** $(\lambda + \mu)\nabla(\nabla \cdot \mathbf{u}) + \mu\nabla^2\mathbf{u} + \mathbf{b} = \rho\ddot{\mathbf{u}}$, obtained by substituting Hooke's law into momentum balance — the single governing equation of linear elastodynamics.
- **Helmholtz decomposition:** $\mathbf{u} = \nabla\phi + \nabla \times \boldsymbol{\psi}$ splits Navier's equation into two independent scalar wave equations, for a dilatational potential ϕ (speed $\alpha = \sqrt{(\lambda + 2\mu)/\rho}$) and a shear vector potential $\boldsymbol{\psi}$ (speed $\beta = \sqrt{\mu/\rho}$).
- **P and S waves:** plane-wave solutions give non-dispersive dispersion relations $\omega = \alpha|\boldsymbol{k}|$, $\omega = \beta|\boldsymbol{k}|$; P-wave particle motion is longitudinal (parallel to $\boldsymbol{k}$), S-wave motion transverse (perpendicular to $\boldsymbol{k}$, split into SH/SV).
- **Energy:** kinetic density $T = \frac{1}{2}\rho\dot{u}_i\dot{u}_i$, strain density $W = \frac{1}{2}\sigma_{ij}\epsilon_{ij}$, related by the local balance $\partial_t(T + W) = \partial_j(\sigma_{ij}\dot{u}_i) + b_i\dot{u}_i$, whose flux term underlies the $1/r$ geometrical spreading of wave amplitude.
