Oct. 07 2005: **Lecture 10:**

Real Eigenvalue Systems; Transformations to Eigenbasis

Reading:
Kreyszig Sections: §7.4 (pp:385–89) , §7.5 (pp:392–96)

Similarity Transformations

A matrix has been discussed as a linear operation on vectors. The matrix itself is defined in terms of the coordinate system of the vectors that it operates on—and that of the vectors it returns.

In many applications, the coordinate system (or laboratory) frame of the vector that gets operated on is the same as the vector gets returned. This is the case for almost all physical properties. For example:

- In an electronical conductor, local current density, $\vec{J}$, is linearly related to the local electric field $\vec{E}$:
$$\underline{\rho}\vec{J} = \vec{E} \tag{10-1}$$
- In a thermal conductor, local heat current density is linearly related to the gradient in temperature:
$$\underline{k}\nabla T = \vec{j_Q} \tag{10-2}$$
- In diamagnetic and paramagnetic materials, the local magnetization, $\vec{B}$ is related to the applied field, $\vec{H}$:
$$\underline{\mu}\vec{H} = \vec{B} \tag{10-3}$$
- In dielectric materials, the local total polarization, $\vec{D}$, is related to the applied electric field:
$$\underline{\kappa}\vec{E} = \vec{D} = \kappa_o\vec{E} + \vec{P} \tag{10-4}$$

When $\vec{x}$ and $\vec{y}$ are vectors representing a physical quantity in Cartesian space (such as force $\vec{F}$, electric field $\vec{E}$, orientation of a plane $\hat{n}$, current $\vec{j}$, etc.) they represent something *physical*. They don't change if we decide to use a different space in which to represent them (such as, exchanging x for y, y for z, z for x; or, if we decide to represent length in nanometers instead of inches, or if we simply decide to rotate the system that describes the vectors. The representation of the vectors themselves may change, but they stand for the same thing.

One interpretation of the operation $\underline{A}\vec{x}$ has been described as geometric transformation on the vector $\vec{x}$. For the case of orthogonal matrices $\underline{A_{orth}}$, geometrical transformations take the forms of rotation, reflection, and/or inversion.

Suppose we have some *physical* relation between two physical vectors in some coordinate system, for instance, the general form of Ohm's law is:

$$\begin{aligned} \vec{J} =& \underline{\sigma}\vec{E} \\ \begin{pmatrix} J_x \\ J_y \\ J_z \end{pmatrix} =& \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_{zz} \end{pmatrix} \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} \end{aligned} \tag{10-5}$$

The matrix (actually it is better to call it a *rank-2 tensor*) $\underline{\sigma}$ is a physical quantity relating the amount of current that flows (in a direction) proportional to the applied electric field (perhaps in a different direction). $\underline{\sigma}$ is the "conductivity tensor" for a particular material.

The physical law in Eq. 10-5 can be expressed as an inverse relationship:

$$\begin{aligned} \vec{E} =& \underline{\rho}\vec{j} \\ \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix} =& \begin{pmatrix} \rho_{xx} & \rho_{xy} & \rho_{xz} \\ \rho_{xy} & \rho_{yy} & \rho_{yz} \\ \rho_{xz} & \rho_{yz} & \rho_{zz} \end{pmatrix} \begin{pmatrix} j_x \\ j_y \\ j_z \end{pmatrix} \end{aligned} \tag{10-6}$$

where the resistivity tensor $\underline{\rho} = \underline{\sigma}^{-1}$.

What happens if we decide to use a new coordinate system (i.e., one that is rotated, reflected, or inverted) to describe the relationship expressed by Ohm's law?

The two vectors must transform from the "old" to the "new" coordinates by:

$$\begin{aligned} \underline{A_{orth}^{old\to new}}\vec{E^{old}} = \vec{E^{new}} \qquad & \underline{A_{orth}^{old\to new}}\vec{j^{old}} = \vec{j^{new}} \\ \underline{A_{orth}^{new\to old}}\vec{E^{new}} = \vec{E^{old}} \qquad & \underline{A_{orth}^{new\to old}}\vec{j^{new}} = \vec{j^{old}} \end{aligned} \tag{10-7}$$

Where is simple proof will show that:

$$\begin{aligned} \underline{A_{orth}^{old\to new}} =& \underline{A_{orth}^{new\to old^{-1}}} \\ \underline{A_{orth}^{new\to old}} =& \underline{A_{orth}^{old\to new^{-1}}} \\ \underline{A_{orth}^{new\to old}} =& \underline{A_{orth}^{old\to new^{T}}} \\ \underline{A_{orth}^{new\to old}} =& \underline{A_{orth}^{old\to new^{T}}} \end{aligned} \tag{10-8}$$

where the last two relations follow from the special properties of orthogonal matrices.

How does the physical law expressed by Eq. 10-5 change in a new coordinate system?

$$\begin{aligned} \text{in old coordinate system: } & \vec{j^{old}} = \underline{\chi^{old}}\vec{E^{old}} \\ \text{in new coordinate system: } & \vec{j^{new}} = \underline{\chi^{new}}\vec{E^{new}} \end{aligned} \tag{10-9}$$

To find the relationship between $\underline{\chi^{old}}$ and $\underline{\chi^{new}}$: For the first equation in 10-9, using the transformations in Eqs. 10-7:

$$\underline{A_{orth}^{new\to old}}\vec{j^{new}} = \underline{\chi^{old}}\,\underline{A_{orth}^{new\to old}}\vec{E^{new}} \tag{10-10}$$

and for the second equation in 10-9:

$$\underline{A_{orth}^{old\to new}}\vec{j^{old}} = \underline{\chi^{new}}\,\underline{A_{orth}^{old\to new}}\vec{E^{old}} \tag{10-11}$$

Left-multiplying by the inverse orthogonal transformations:

$$\begin{aligned} \underline{A_{orth}^{old\to new}\,A_{orth}^{new\to old}}\vec{j^{new}} &= \underline{A_{orth}^{old\to new}}\,\underline{\chi^{old}}\,\underline{A_{orth}^{new\to old}}\vec{E^{new}} \\ \underline{A_{orth}^{new\to old}\,A_{orth}^{old\to new}}\vec{j^{old}} &= \underline{A_{orth}^{new\to old}}\,\underline{\chi^{new}}\,\underline{A_{orth}^{old\to new}}\vec{E^{old}} \end{aligned} \tag{10-12}$$

Because the transformation matrices are inverses, the following relationship between *similar* matrices in the old and new coordinate systems is:

$$\begin{aligned} \underline{\chi^{old}} &= \underline{A_{orth}^{old\to new}}\,\underline{\chi^{new}}\,\underline{A_{orth}^{new\to old}} \\ \underline{\chi^{new}} &= \underline{A_{orth}^{new\to old}}\,\underline{\chi^{old}}\,\underline{A_{orth}^{old\to new}} \end{aligned} \tag{10-13}$$

The $\underline{\chi^{old}}$ is said to be *similar* to $\underline{\chi^{new}}$ and the double multiplication operation in Eq. 10-13 is called *a similarity transformation.*

Stresses and Strains

Stresses and strains are rank-2 tensors that characterize the mechanical state of a material.

A spring is an example of a one-dimensional material—it resists or exerts force in one direction only. A volume of material can exert forces in all three directions simultaneously—and the forces need not be the same in all directions. A volume of material can also be "squeezed" in many different ways: it can be squeezed along any one of the axis or it can be subjected to squeezing (*or smeared*) around any of the axes[2]

[2]Consider a blob of modeling clay—you can deform it by placing between your thumbs and one opposed finger; you can deform it by simultaneously squeezing with two sets of opposable digits; you can "smear" it by pushing and pulling in opposite directions. These are examples of uniaxial, biaxial, and shear stress.

All the ways that a force can be applied to small element of material are now described. A force divided by an area is a *stress*—think of it the areal density of force.

$$\sigma_{ij} = \frac{F_i}{A_j} \qquad (\text{i.e., } \sigma_{xz} = \frac{F_x}{A_z} = \sigma_{xz} = \frac{\vec{F}\cdot\hat{i}}{\vec{A}\cdot\hat{k}}) \tag{10-14}$$

A_j is a plane with its normal in the $\hat{j}$-direction (or the projection of the area of a plane $\vec{A}$ in the direction parallel to $\hat{j}$)

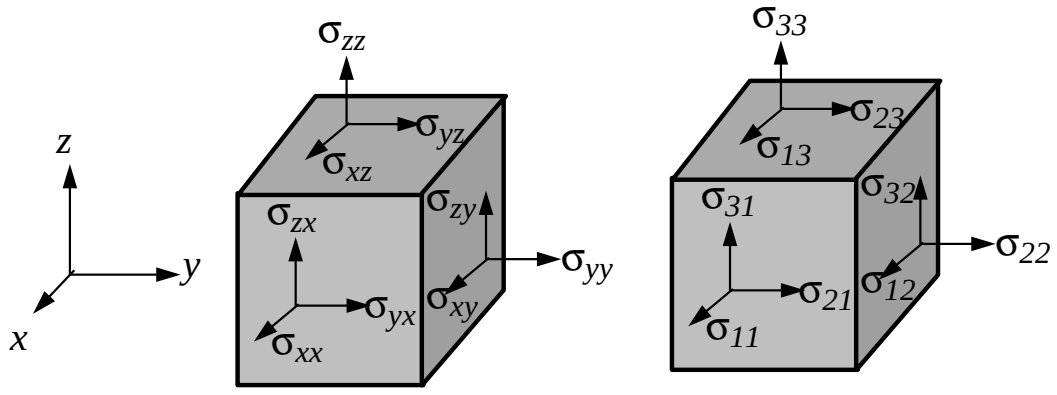

Figure 10-1: Illustration of stress on an oriented volume element.

$$\sigma_{ij} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix} \tag{10-15}$$

There is one special and very simple case of elastic stress, and that is called the hydrostatic stress. It is the case of pure pressure and there are no shear (off-diagonal) stresses (i.e., all $\sigma_{ij} = 0$ for $i \neq j$, and $\sigma_{11} = \sigma_{22} = \sigma_{33}$). An equilibrium system composed of a body in a fluid environment is always in hydrostatic stress:

$$\sigma_{ij} = \begin{bmatrix} -P & 0 & 0 \\ 0 & -P & 0 \\ 0 & 0 & -P \end{bmatrix} \tag{10-16}$$

where the pure hydrostatic pressure is given by P.

Strain is also a rank-2 tensor and it is a physical measure of a how much a material changes its shape.[3]

[3]It is unfortunate that the words of these two related physical quantities, stress and strain, sound so similar. Strain measures the change in geometry of a body and stress measures the forces that squeeze or pull on a body. St**ress** is the pr**ess**; Str**ain** is the g**ain**.

Why should strain be a rank-2 tensor?

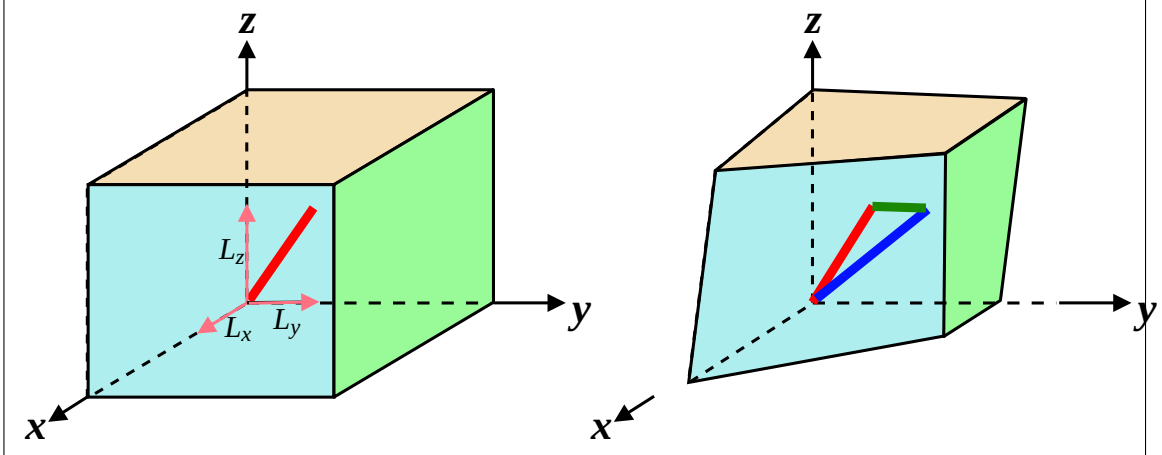

Figure 10-2: Illustration of how strain is defined: imagine a small line-segment that is aligned with a particular direction (one set of indices for the direction of the line-segment); after deformation the end-points of the line segment define a new line-segment in the deformed state. The difference in these two vectors is a vector representing how the line segment has changed from the initial state into the deformed state. The difference vector can be oriented in any direction (the second set of indices)—the strain is a representation of "a difference vectors for all the oriented line-segments" divided by the length of the original line.

Or, using the same idea as that for stress:

$$\epsilon_{ij} = \frac{\Delta L_i}{L_j} \qquad (\text{i.e., } \epsilon_{xz} = \frac{\Delta L_x}{L_z} = \epsilon_{xz} = \frac{\Delta \vec{L} \cdot \hat{i}}{\vec{L} \cdot \hat{k}}) \tag{10-17}$$

If a body that is being stressed hydro-statically is isotropic, then its response is pure dilation (in other words, it expands or shrinks uniformly and without shear):

$$\epsilon_{ij} = \begin{bmatrix} \Delta/3 & 0 & 0 \\ 0 & \Delta/3 & 0 \\ 0 & 0 & \Delta/3 \end{bmatrix} \tag{10-18}$$

$$\Delta = \frac{dV}{V} \tag{10-19}$$

So, for the case of hydrostatic stress, the work term has a particularly simple form:

$$\begin{aligned} V \sum_{i=1}^{3} \sum_{j=1}^{3} \sigma_{ij} d\epsilon_{ij} &= -P dV \\ V \sigma_{ij} d\epsilon_{ij} &= -P dV \qquad (\text{summation convention}) \end{aligned} \tag{10-20}$$

This expression is the same as the rate of work performed on a compressible fluid, such as an ideal gas.

EigenStrains and EigenStresses .....

For any strain matrix, there is a choice of an coordinate system where line-segments that lie along the coordinate axes always deform parallel to themselves (i.e., they only stretch or shrink, they do not twist).

For any stress matrix, there is a choice of an coordinate system where all shear stresses (the off-diagonal terms) vanish and the matrix is diagonal.

These coordinate systems define the eigenstrain and eigenstress. The matrix transformation that takes a coordinate system into its eigenstate is of great interest because it simplifies the mathematical representation of the physical system.

MATHEMATICA® Example: (notebook) Lecture-10

Principal Axes

Diagonalizing the two-dimensional stress matrix

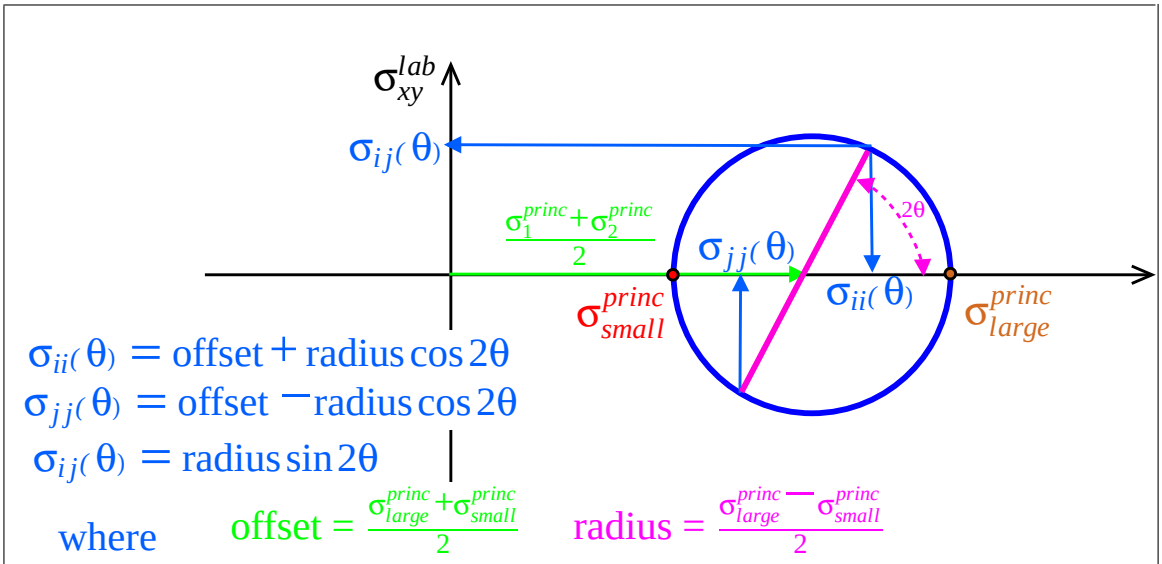

Figure 10-3: Mohr's circle of stress is a way of graphically representing the two-dimensional stresses of identical stress states, but in rotated laboratory frames.
The center of the circle is displaced from the origin by a distance equal to the average of the principal stresses (or average of the eigenvalues of the stress tensor).
The maximum and minimum stresses are the eigenvalues—and they define the diameter in the principal $\theta = 0$ frame.
Any other point on the circle gives the stress tensor in a frame rotated by 2θ from the principal axis using the construction illustrated by the blue lines (and equations).

Quadratic Forms

The example above, where a matrix (rank-2 tensor) represents a material property, can be understood with a useful geometrical interpretation.

For the case of the conductivity tensor $\underline{\sigma}$, the dot product $\vec{E} \cdot \vec{j}$ is a scalar related to the local energy dissipation:

$$e = \vec{E}^T \underline{\sigma} \vec{E} \tag{10-21}$$

The term on the right-hand-side is called a quadratic form, as it can be written as:

$$\begin{aligned} e = &\sigma_{11}x_1^2 + \sigma_{12}x_1x_2 + \sigma_{13}x_1x_3 + \\ &\sigma_{21}x_1x_2 + \sigma_{22}x_2^2 + \sigma_{23}x_2x_3 + \\ &\sigma_{31}x_1x_3 + \sigma_{32}x_2x_3 + \sigma_{33}x_3^2 \end{aligned} \tag{10-22}$$

or, because $\underline{\sigma}$ is symmetric:

$$\begin{aligned} e = &\sigma_{11}x_1^2 + 2\sigma_{12}x_1x_2 + 2\sigma_{13}x_1x_3 + \\ &\sigma_{22}x_2^2 + 2\sigma_{23}x_2x_3 + \\ &\sigma_{33}x_3^2 \end{aligned} \tag{10-23}$$

It is not unusual for such quadratic forms to represent energy quantities. For the case of paramagnetic and diamagnetic materials with magnetic permeability tensor $\underline{\mu}$, the energy per unit volume due to an applied magnetic field $\vec{H}$ is:

$$\frac{E}{V} = \frac{1}{2}\vec{H}^T \underline{\mu} \vec{H} \tag{10-24}$$

for a dielectric (i.e., polarizable) material with electric electric permittivity tensor $\underline{\kappa}$ with an applied electric field $\vec{E}$:

$$\frac{E}{V} = \frac{1}{2}\vec{E}^T \underline{\kappa} \vec{E} \tag{10-25}$$

The geometric interpretation of the quadratic forms is obtained by turning the above equations around and asking—what are the general vectors $\vec{x}$ for which the quadratic form (usually an energy or power density) has a particular value? Picking that particular value as unity, the question becomes what are the directions and magnitudes of $\vec{x}$ for which

$$1 = \vec{x}^T \underline{A} \vec{x} \tag{10-26}$$

This equation expresses a quadratic relationship between one component of $\vec{x}$ and the others. This is a surface—known as the *quadric surface* or *representation quadric*—which is an ellipsoid or hyperboloid sheet on which the quadratic form takes on the particular value 1.

In the principal axes (or, equivalently, the eigenbasis) the quadratic form takes the quadratic form takes the simple form:

$$e = \vec{x_{\text{eb}}}^T \underline{A_{\text{eb}}} \vec{x_{\text{eb}}} = A_{11}x_1^2 + A_{22}x_2^2 + A_{33}x_3^2 \tag{10-27}$$

and the representation quadric

$$A_{11}x_1^2 + A_{22}x_2^2 + A_{33}x_3^2 = 1 \tag{10-28}$$

which is easily characterized by the signs of the coefficients.

In other words, in the principal axis system (or the eigenbasis) the quadratic form has a particularly simple, in fact the most simple, form.

Eigenvector Basis

Among all similar matrices (defined by the similarity transformation defined by Eq. 10-13), the simplest matrix is the diagonal one. In the coordinate system where the similar matrix is diagonal, its diagonal entries are the eigenvalues. The question remains, "what is the coordinate transformation that takes the matrix into its diagonal form?"

The coordinate system is called the eigenbasis or principal axis system, and the transformation that takes it there is particularly simple.

The matrix that transforms from a general (old) coordinate system to a diagonalized matrix (in the new coordinate system) is the matrix of columns of the eigenvectors. The first column corresponds to the first eigenvalue on the diagonal matrix, and the n^{th} column is the eigenvector corresponding the n^{th} eigenvalue.

$$\begin{pmatrix} \text{The} \\ \text{Diagonalized} \\ \text{Matrix} \end{pmatrix} = \begin{pmatrix} \text{Eigenvector} \\ \text{Column} \\ \text{Matrix} \end{pmatrix}^{-1} \begin{pmatrix} \text{The} \\ \text{General} \\ \text{Matrix} \end{pmatrix} \begin{pmatrix} \text{Eigenvector} \\ \text{Column} \\ \text{Matrix} \end{pmatrix} \tag{10-29}$$

This method provides a method for finding the simplest quadratic form.