Chapter 1

Basic calculus and algebra review

This section provides a few brief notes on math notation and concepts needed for this text. Not all concepts and formula are presented in a mathematically rigorous way, and you should refer to something like a Math for Engineers text for a more complete treatment. For most of this text, it will be assumed that the reader is familiar with the material treated in this chapter.

1.1 Calculus

1.1.1 Full and partial derivatives

In calculus, we are interested in the *change* or *dependence* of some quantity, *e.g.* u, on small changes in some variable x. If u has value u_0 at x_0 and changes to $u_0 + \delta u$ when x changes to $x_0 + \delta x$, the incremental change can be written as

$$\delta u = \frac{\delta u}{\delta x}(x_0)\delta x. \tag{1.1}$$

The δ (or sometimes written as capital Δ) here means that this is a small, but finite quantity. If we let δx get asymptotically smaller around x_0, we of course arrive at the *partial derivative*, which we denote with ∂ like

$$\lim_{\delta x \to 0} \frac{\delta u}{\delta x}(x_0) = \frac{\partial u}{\partial x}. \tag{1.2}$$

The limit in eq. (1.2) will work as long as u does not do any funny stuff as a function of x, like jump around abruptly. When you think of $u(x)$ as a function (some line on a plot) that depends on x, $\partial u/\partial x$ is the slope of this line that can be obtained by measuring the change δu over some interval δx, and then making the interval progressively smaller.

We call $\frac{\partial u}{\partial x}$ (we also write in shorthand $\partial_x u(x)$ or $u'(x)$; if the variable is time, t, we also use $\dot{u}(t)$ for $\partial u/\partial t$) the partial derivative, because u might also depend on other variables, *e.g.* y and z. If this is the case, the *total derivative* du at some $\{x_0, y_0, z_0\}$ (we will drop (*i.e.*

not write down) the explicit dependence on the variables from now on) is given by the sum of the changes in all variables on which u depends:

$$\mathrm{d}u = \frac{\partial u}{\partial x}\,\mathrm{d}x + \frac{\partial u}{\partial y}\,\mathrm{d}y + \frac{\partial u}{\partial z}\,\mathrm{d}z. \tag{1.3}$$

Here, $\mathrm{d}x$ and similar are placeholders for infinitesimal changes in the variables. This means that eq. (1.3) works as long as $\mathrm{d}x$ is small enough that a linear relationship between δu and δx still holds. In fact, we can perform a Taylor approximation on *any* $u(x)$ around x_0 by

$$u(x) = u(x_0) + \frac{\partial u}{\partial x}(x_0)(x - x_0) + \frac{\partial^2 u}{\partial x^2}(x_0)\frac{(x - x_0)^2}{2!} + \frac{\partial^3 u}{\partial x^3}(x_0)\frac{(x - x_0)^3}{3!}\dots \tag{1.4}$$

Here, $\frac{\partial^2 u}{\partial x^2}$ is the second derivative, the change of the change of u with x. $n!$ denotes the factorial, *i.e.*

$$n! = 1 \times 2 \times 3 \times \dots n. \tag{1.5}$$

So, as long as $\mathrm{d}x = x - x_0$ is small, the derivative will work (for well behaved u). For example, if better approximations are needed, *e.g.*when the strain tensor is not infinitesimal anymore, quadratic and higher terms like the one that goes with the second derivative in the series eq. (1.4) and so on need to be taken into account. Finite difference methods essentially use Taylor approximations to approximate derivatives, as we will see later.

How to compute derivatives Here are some of the most common derivatives of a few functions:

function $f(x)$	derivative $f'(x)$	comment
x^p	px^{p-1}	special case: $f(x) = c = cx^0 \rightarrow f'(x) = 0$ where c, p are constants
$\exp(x) = e^x$	e^x	that's what makes e so special
$\ln(x)$	$1/x$	
$\sin(x)$	$\cos(x)$	
$\cos(x)$	$-\sin(x)$	
$\tan(x)$	$\sec^2(x) = 1/\cos^2(x)$	

If you need to take derivatives of combinations of two or more functions, here called f, g, and h, there are four important rules (with a and b being constants):

Chain rule (inner and outer derivative):

$$\text{If} \quad f(x) = h(g(x)) \tag{1.6}$$

$$f'(x) = h'(g(x))g'(x), \tag{1.7}$$

i.e. derivative of nested functions are given by the outer times the inner derivative.

Sum rule:

$$(af(x)+bg(x))' = af'(x)+bg'(x) \tag{1.8}$$

Product rule:

$$(f(x)g(x))' = f'(x)g(x)+f(x)g'(x) \tag{1.9}$$

Quotient rule:

$$\text{If} \quad f(x) = \frac{g(x)}{h(x)} \tag{1.10}$$

$$f'(x) = \frac{g'(x)h(x)-g(x)h'(x)}{h(x)^2} \tag{1.11}$$

If you need higher order derivatives, those are obtained by successively computing derivatives, *e.g.* the third derivative of $f(x)$ is

$$\frac{\partial^3 f(x)}{\partial x^3} = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}f(x)\right)\right). \tag{1.12}$$

Say, $f(x)=x^3$, then

$$\frac{\partial^3 x^3}{\partial x^3} = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}x^3\right)\right) = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}3x^2\right) = \frac{\partial}{\partial x}6x = 6. \tag{1.13}$$

1.1.2 Divergence and curl

Operators are mathematical constructs that do something with the entity that is written to their right. For example, we had earlier introduced the *gradient operator*, $\underline{\nabla}$ (the del operator is represented by the "Nabla" symbol ∇), which takes derivatives in all directions and, in a Cartesian system, is given by $\underline{\nabla} = \left\{\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right\}$. Note that the operatore $\underline{\nabla}$ is a vector. When applied to scalar field (a distribution of values that depends on spatial location), such as a temperature distribution $T(x,y,z)$ (meaning T is variable with coordinates x, y, and z, assumed implicitly for all properties from now on), the *gradient* operation

$$grad\ T = \underline{\nabla} T = \begin{pmatrix} \frac{\partial T}{\partial x} \\ \frac{\partial T}{\partial y} \\ \frac{\partial T}{\partial z} \end{pmatrix} \tag{1.14}$$

generates a vector from the scalar field which points in the direction of the steepest increase in T.

Consider what $\underline{\nabla}$ can do to a vector field (*i.e.* vectors that vary in space, $\underline{x}$). If

$$\underline{v}(\underline{x}) = \{v_1(\underline{x}), v_2(\underline{x}), v_3(\underline{x})\} \tag{1.15}$$

is a velocity field, then the *divergence* (grad dot product) operation on a vector field

$$div\ \underline{v} = \underline{\nabla} \cdot \underline{v} \tag{1.16}$$

is equivalent to finding the dilatancy (volumetric) strain-rate $\dot{\triangle}$ from the strain-rate tensor components because

$$\dot{\triangle} = \frac{\Delta \dot{V}}{V} = tr(\underline{\underline{\dot{\epsilon}}}) = \sum_i \dot{\epsilon}_{ii} = \dot{\epsilon}_{11} + \dot{\epsilon}_{22} + \dot{\epsilon}_{33} = \frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2} + \frac{\partial v_3}{\partial x_3} = \underline{\nabla} \cdot \underline{v}. \tag{1.17}$$

Here V is volume, and $\Delta \dot{V}$ volume rate-change and, mind you, the strain-rate tensor, $\underline{\underline{\dot{\epsilon}}}$, is defined as

$$\underline{\underline{\dot{\epsilon}}} = \dot{\epsilon}_{ij} = \frac{1}{2}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right). \tag{1.18}$$

In complete analogy, if the vector field are displacements $\underline{u}(\underline{x})$, then $\underline{\nabla} \cdot \underline{u}$ yields the dilatancy, *i.e.* the trace of the strain tensor, $\underline{\underline{\epsilon}}$,

$$\underline{\underline{\epsilon}} = \epsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right). \tag{1.19}$$

Eq. (1.17) illustrates that the divergence has to do with sinks and sources, or volumetric effects. The volume integral over the divergence of a velocity field is equal to the surface integral of the flow normal to the surface. (An electro-magnetics example: For the magnetic field: $div\ \underline{B} = 0$ because there are no magnetic monopoles, but for the electric field: $div\ \underline{E} = q$, with electric charges q being the "source".)

If we take the vector instead of the dot product (see sec. 1.2.2) with the grad operator, we have the *curl* or *rot* operation

$$curl\ \underline{v} = \underline{\nabla} \wedge \underline{v}. \tag{1.20}$$

The curl is a rotation vector just like $\underline{\omega}$. Indeed, if the velocity field is that of a the rigid body rotation, $\underline{v} = \underline{\omega} \wedge \underline{r}$, one can show that $\underline{\nabla} \wedge \underline{v} = \underline{\nabla} \wedge (\underline{\omega} \wedge \underline{r}) = 2\underline{\omega}$.

Second derivatives enter into the *Laplace* operator which appears, *e.g.* in the diffusion equation:

$$\triangle\ T = \nabla^2 T = \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} \tag{1.21}$$

Some rules for second derivatives:

$$curl(grad\ T) = \underline{\nabla} \times (\underline{\nabla} T) \quad = \quad 0 \tag{1.22}$$

$$div(curl\ \underline{v}) = \underline{\nabla} \cdot \underline{\nabla} \times \underline{v} = 0 \tag{1.23}$$

1.1.3 Integrals

Taking an integral

$$F(x) = \int f(x)dx, \tag{1.24}$$

in a general (indefinite) sense, is the inverse of taking the derivative of a function f,

$$F\left(\frac{\partial f(x)}{\partial x}\right) = f(x) + c \tag{1.25}$$

$$\frac{\partial}{\partial x}F\left(\frac{\partial f(x)}{\partial x}\right) = \frac{\partial}{\partial x}(f(x)+c) = f'(x). \tag{1.26}$$

Any general integration of a derivative is thus only determined up to an integration constant, here c, because the derivative, which is the reverse of the integral, of a constant is zero.

Graphically, the definite (with bounds) integral over $f(x)$

$$\int_a^b f(x)dx = F(b) - F(a) \tag{1.27}$$

along x, adding up the value of $f(x)$ over little chunks of dx, from the left $x = a$ to the right $x = b$ corresponds to the area under the curve $f(x)$. This area can be computed by subtracting the analytical form of the integral at b from that at a, $F(b) - F(a)$. If $f(x) = c$ (c a constant), then

$$F(x) = cx + d \tag{1.28}$$

$$F(b) = cb + d \tag{1.29}$$

$$F(a) = ca + d \tag{1.30}$$

$$F(b) - F(a) = c(b-a), \tag{1.31}$$

the area of the box $(b - a) \times c$.

Here are the integrals (anti derivatives) of a few common functions, all only determined up to an integration constant C

function $f(x)$	integral $F(x)$	comment
x^p	$\frac{x^{p+1}}{p+1} + C$	special case: $f(x) = c = cx^0 \rightarrow F(x) = cx + C$
e^x	e^x +C	
$1/x$	$\ln(\|x\|) + C$	
$\sin(x)$	$-\cos(x) + C$	
$\cos(x)$	$\sin(x) + C$	

There are also a few very helpful definite integrals without closed-form anti derivatives, *e.g.*

$$\int_0^\infty e^{-x^2}dx = \frac{\sqrt{\pi}}{2} \tag{1.32}$$

Wolfram Alpha[1], Mathematica, a standard math textbook, table of integrals, or Wikipedia will be of help with more complicated integrals.

A few conventions and rules for integration:

Notation: Everything after the $\int$ sign is usually meant to be integrated over up to the dx, or the next major mathematical operator if the dx is placed next to the $\int$ if the context allows:

$$\int \left(af(x) + bg(x) + \ldots\right) dx = \int af(x) + bg(x) \ldots dx \tag{1.33}$$

$$\int dx\, f(x) = \int f(x) dx \tag{1.34}$$

Linearity:

$$\int_a^b \left(cf(x) + dg(x)\right) dx = c \int_a^b f(x) dx + d \int_a^b g(x) \tag{1.35}$$

Reversal:

$$\int_a^b f(x) dx = -\int_b^a f(x) dx \tag{1.36}$$

Zero length:

$$\int_a^a f(x) dx = 0 \tag{1.37}$$

Additivity:

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx \tag{1.38}$$

Product rules:

$$\int f'(x) f(x) dx = \frac{1}{2} \left(f(x)\right)^2 + C \tag{1.39}$$

$$\int f'(x) g(x) dx = f(x) g(x) - \int f(x) g'(x) dx \tag{1.40}$$

Quotient rule:

$$\int \frac{f'(x)}{f(x)} dx = \ln |f(x)| + C \tag{1.41}$$

[1] https://www.wolframalpha.com/

Gauß theorem The integral over the area Ω of the divergence of a vector field $\underline{f}$ is equivalent to the boundary integral, $\partial\Omega$, over the local normal (to the boundary), $\underline{n}$, dotted with $\underline{f}$:

$$\int_\Omega dA\, \underline{\nabla} \cdot \underline{f} = \int_{\partial\Omega} ds\, \underline{n} \cdot \underline{f}. \tag{1.42}$$

1.2 Linear algebra

TO BE ADDED: matlab conventions for mathematical operations such as dot and cross products.

1.2.1 The dot product

We will make use of the *dot product*, which is defined as

$$c = \underline{a} \cdot \underline{b} = \sum_{i=1}^{n} a_i b_i, \tag{1.43}$$

where $\underline{a}$ and $\underline{b}$ are vectors of dimension n (n-dimensional, geometrical objects with a direction and length, like a velocity) and the outcome of this operation is a scalar (a regular number), c. In eq. (1.43), $\sum_{i=1}^{n}$ means "sum all that follows while increasing the index i from the lower limit, $i = 1$, in steps of of unity, to the upper limit, $i = n$". In the examples below, we will assume a typical, spatial coordinate system with $n = 3$ so that

$$\underline{a} \cdot \underline{b} = a_1 b_1 + a_2 b_2 + a_3 b_3, \tag{1.44}$$

where 1, 2, 3 refer to the vector components along x, y, and z axis, respectively (ADD FIGURE HERE). In the "Einstein summation" convention, we would rewrite $\sum_{i=1}^{n} a_i b_i$ simply as $a_i b_i$, where summation over repeated indices is implied, *i.e.* the $\sum$ is not written.

When we write out the vector components, we put them on top of each other

$$\underline{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} \tag{1.45}$$

or in a list, maybe with curly brackets, like so: $\underline{a} = \{a_1, a_2, a_3\}$. Here, we use $\underline{a}$ to denote a vector, like what is commonly done when writing by hand. You will also see bold font $\mathbf{a}$ to denote vectors as opposed to scalar a, or another common form is $\vec{a}$.

We can write the amplitude (or: length, L_2 norm) of a vector as

$$|\underline{a}| = \sqrt{\sum_{i}^{n} a_i^2} = \sqrt{a_1^2 + a_2^2 + a_3^2} = \sqrt{a_x^2 + a_y^2 + a_z^2}. \tag{1.46}$$

For instance, all of the basis vectors defining the Cartesian coordinate system, $\underline{e}_x$, $\underline{e}_y$, and $\underline{e}_z$ have unity length by definition, $|\underline{e}_i| = 1$. Those $\underline{e}_i$ vectors point along the respective axes of the Cartesian coordinate system so that we can assemble a vector from its components like

$$\underline{a} = \{a_x, a_y, a_z\} = a_x\underline{e}_x + a_y\underline{e}_y + a_z\underline{e}_z. \tag{1.47}$$

For a spherical system, the $\underline{e}_r$, $\underline{e}_\theta$, and $\underline{e}_\phi$ unity vectors can still be used to express vectors but the actual Cartesian components of $\underline{e}_i$ depend on the coordinates at which the vectors are evaluated.

We can restate eq. (1.43) and give another definition of the dot product,

$$\underline{a} \cdot \underline{b} = |\underline{a}||\underline{b}| \cos\theta \tag{1.48}$$

where θ is the angle between vectors $\underline{a}$ and $\underline{b}$. The meaning of this is that if you want to know what component of vector $\underline{a}$ is parallel to $\underline{b}$, you just take the dot product. Say, you have a velocity $\underline{v}$ and want the normal velocity v_n along a vector $\underline{n}$ with $|\underline{n}| = 1$ that is oriented at a 90° angle (perpendicular) to some plate boundary, you can use $v_n = \underline{v} \cdot \underline{n}$.

Also, eq. (1.47) only works because the basis vectors $\underline{e}_i$ of any coordinate system are, by definition, orthogonal (at right angle, perpendicular, at $\theta = 90°$) to each other and $\underline{e}_i \cdot \underline{e}_j = 0$ for all $i \neq j$. Likewise, $\underline{e}_i \cdot \underline{e}_i = 1$ for all i since $\underline{a} \cdot \underline{a} = |\underline{a}|^2$, and basis vectors have unity length by definition. Using the Kronecker δ

$$\delta_{ij} = 1 \quad \text{for} \quad i = j, \quad \text{and} \quad \delta_{ij} = 0 \quad \text{for} \quad i \neq j, \tag{1.49}$$

we can write the conditions for the basis vectors as

$$\underline{e}_i \cdot \underline{e}_j = \delta_{ij}. \tag{1.50}$$

1.2.2 Vector or cross product

This operation is written as $\underline{a} \times \underline{b}$ or $\underline{a} \wedge \underline{b}$ and its result is another vector

$$\underline{c} = \underline{a} \wedge \underline{b} \tag{1.51}$$

that is at a right angle to both $\underline{a}$ and $\underline{b}$ (hence the right-hand-rule, with thumb, index, and middle finger along $\underline{a}$, $\underline{b}$, and $\underline{c}$, respectively). vector $\underline{c}$'s length is given by

$$|\underline{c}| = |\underline{a} \wedge \underline{b}| = |\underline{a}||\underline{b}| \sin\theta, \tag{1.52}$$

that is, $\underline{c}$ is largest when $\underline{a}$ and $\underline{b}$ are orthogonal, and zero if they are parallel. Compare this relationship to eq. (1.48).

In 3-D,

$$\underline{c} = \underline{a} \wedge \underline{b} = \begin{pmatrix} a_y b_z - a_z b_y \\ a_z b_x - a_x b_z \\ a_x b_y - a_y b_x \end{pmatrix} \tag{1.53}$$

(note that there is no i component of $\underline{a}$ or $\underline{b}$ in the i component of $\underline{c}$, this is the aforementioned orthogonality property).

An example for a cross product is the velocity $\underline{v}$ at a point with location $\underline{r}$ in a body spinning with the rotation vector $\underline{\omega}$, $\underline{v} = \underline{\omega} \wedge \underline{r}$. The rotation vector $\underline{\omega}$ is different from, *e.g.*, $\underline{r}$ in that $\underline{\omega}$ has a spin (a sense of rotation) to it (the other right-hand-rule, where your thumb points along the vector and your fingers indicate the counter-clockwise motion). ADD FIGURE

1.2.3 Matrices and tensors

A $n \times m$ matrix is a rectangular table of elements (or entries) with n rows and m columns which are filled with numbers. For example, if $\underline{\underline{A}}$ is 3×3,

$$\underline{\underline{A}} = \begin{pmatrix} a_{xx} & a_{xy} & a_{xz} \\ a_{yx} & a_{yy} & a_{yz} \\ a_{zx} & a_{zy} & a_{zz} \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}. \tag{1.54}$$

You will see matrices also printed like so A, with us here using the handwritten/blackboard version, double underlining like so $\underline{\underline{A}}$. The elements are referred to as a_{ij} where i is the row and j the column. Matrices can be added and or multiplied.

Multiplication of matrix with a scalar

$$f\underline{\underline{A}} = f a_{ij} = f \times \begin{pmatrix} a_{xx} & a_{xy} & a_{xz} \\ a_{yx} & a_{yy} & a_{yz} \\ a_{zx} & a_{zy} & a_{zz} \end{pmatrix} = \begin{pmatrix} fa_{xx} & fa_{xy} & fa_{xz} \\ fa_{yx} & fa_{yy} & fa_{yz} \\ fa_{zx} & fa_{zy} & fa_{zz} \end{pmatrix} \tag{1.55}$$

Multiplication of a matrix with a vector

$$\begin{pmatrix} c_x \\ c_y \\ c_z \end{pmatrix} = \begin{pmatrix} a_{xx} & a_{xy} & a_{xz} \\ a_{yx} & a_{yy} & a_{yz} \\ a_{zx} & a_{zy} & a_{zz} \end{pmatrix} \cdot \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} = \begin{pmatrix} a_{xx}b_x + a_{xy}b_y + a_{xz}b_z \\ a_{yx}b_x + a_{yy}b_y + a_{yz}b_z \\ a_{zx}b_x + a_{zy}b_y + a_{zz}b_z \end{pmatrix} \tag{1.56}$$

or

$$c_i = \sum_j a_{ij} b_j. \tag{1.57}$$

Multiplication of two matrices works like this:

$$\underline{\underline{C}} = \underline{\underline{A}}\,\underline{\underline{B}} \tag{1.58}$$

$$c_{ij} = \sum_k a_{ik} b_{kj}, \tag{1.59}$$

where k goes from 1 to the number of columns in $\underline{\underline{A}}$, which has to be equal to the number of rows in $\underline{\underline{B}}$. Note that, in general, $\underline{\underline{A}}\,\underline{\underline{B}} \neq \underline{\underline{B}}\,\underline{\underline{A}}$!

Special types of matrices and matrix operations

Quadratic matrices Have $n \times n$ rows and columns. All simple physical tensors, such as stress or strain, can be written as quadratic matrices in 3×3.

Identity matrix $\underline{\underline{1}} = \underline{\underline{I}}$, $i_{ij} = \delta_{ij}$, *i.e.* this matrix is unity along the diagonal, and zero for all other elements.

Trace The trace of a $n \times n$ matrix $\underline{\underline{A}}$ is the sum of its diagonal elements

$$tr(\underline{\underline{A}}) = \sum_{i=1}^{n} a_{ii}. \tag{1.60}$$

Determinant The determinant for a 2×2 matrix is computed as

$$\det(\underline{\underline{A}}) = a_{11}a_{22} - a_{12}a_{21} \tag{1.61}$$

and is a measure of area change. For 3×3,

$$\begin{aligned} \det(\underline{\underline{A}}) &= a_{11}\left(a_{22}a_{33} - a_{23}a_{32}\right) \\ &- a_{12}\left(a_{21}a_{33} - a_{23}a_{31}\right) \\ &+ a_{13}\left(a_{21}a_{32} - a_{22}a_{31}\right) \end{aligned} \tag{1.62}$$

(note how the 3×3 determinant is assembled from a pattern of 2×2 determinants; for $n > 3$, a correspondingly more complicated formula applies.

ADD FIGURE

Vector cross product based on the determinant The cross product $\underline{c} = \underline{a} \wedge \underline{b}$ (eq. 1.53) can also be written as the determinant of the matrix

$$\begin{pmatrix} \underline{e}_x & \underline{e}_y & \underline{e}_z \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{pmatrix} \tag{1.63}$$

Invariants The trace

$$I_{\underline{\underline{A}}} = tr(\underline{\underline{A}}) = \sum_i a_{ii} = a_{ii} \tag{1.64}$$

(Einstein summation convention implies summation over all repeated indices), and determinant

$$III_{\underline{\underline{A}}} = \det(\underline{\underline{A}}) \tag{1.65}$$

of a matrix $\underline{\underline{A}}$ are two of the three invariants, *i.e.* properties of a tensor (expressed as a matrix) that are independent of a coordinate system. The third is the "second invariant",

$$II_{\underline{\underline{A}}} = a_{11}a_{22} + a_{11}a_{33} + a_{22}a_{33} - a_{12}^2 - a_{13}^2 - a_{23}^2. \tag{1.66}$$

These expressions arise when finding the eigenvectors and values of a tensor, eq. (1.75).

Transpose of a matrix $(\underline{\underline{A}}^T)_{ij} = a^T_{ij} = a_{ji}$, *i.e.* the transpose has all elements flipped by row and column.

Inverse of $\underline{\underline{A}}$, $\underline{\underline{A}}^{-1}$: The inverse of a matrix is defined via

$$\underline{\underline{A}}^{-1}\underline{\underline{A}} = \underline{\underline{A}}\,\underline{\underline{A}}^{-1} = \underline{\underline{I}}. \tag{1.67}$$

If the inverse exists, then $(\underline{\underline{A}}^{-1})^{-1} = \underline{\underline{A}}$, $(\underline{\underline{A}}^T)^{-1} = (\underline{\underline{A}}^{-1})^T$, and $(\underline{\underline{A}}\,\underline{\underline{B}})^{-1} = \underline{\underline{B}}^{-1}\underline{\underline{A}}^{-1}$. The inverse only exists if $\det(\underline{\underline{A}}) \neq 0$.

For the special case of a 2×2 matrix

$$\underline{\underline{A}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{1.68}$$

the inverse is given by

$$\underline{\underline{A}}^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \tag{1.69}$$

Orthogonal or rotation matrices: For those matrices,

$$\underline{\underline{A}}\,\underline{\underline{A}}^T = \underline{\underline{A}}^T\underline{\underline{A}} = \underline{\underline{1}} \tag{1.70}$$

holds.

If a rotation matrix, $\underline{\underline{R}}$, converts one coordinate system for $\underline{x}$ this vector into $\underline{x}'$, then

$$\underline{y}' = \underline{\underline{R}}\,\underline{y} \tag{1.71}$$

for any vector, and

$$\underline{\underline{\sigma}}' = \underline{\underline{R}}\,\underline{\underline{\sigma}}\,\underline{\underline{R}}^T \tag{1.72}$$

for any matrix, such as the stress tensor.

Eigenvalues and eigen vectors: Any $n \times n$ symmetric matrix $\underline{\underline{A}}$ has n eigen vectors $\underline{v}_i$ that correspond to real eigenvalues λ_i such that

$$\underline{\underline{A}}\underline{v}_i = \lambda_i \underline{v}_i \tag{1.73}$$

An example is the stress matrix which can be written in the principal axes system, where the eigen vectors of the Cartesian representation of the stress matrix are the principal axes.

Eigenvalues can be found using

$$\det(\underline{\underline{A}} - \lambda \underline{\underline{I}}) = 0 \tag{1.74}$$

and eigen vectors subsequently by using the first property, which leads to

$$\det(\underline{\underline{A}} - \lambda \underline{\underline{I}}) = -\lambda^3 + I_{\underline{\underline{A}}}\lambda^2 - II_{\underline{\underline{A}}}\lambda + III_{\underline{\underline{A}}} = 0. \tag{1.75}$$

If a symmetric matrix $\underline{\underline{A}}$ is transformed into the principal axes system, $\underline{\underline{A}}'$, there are no off-diagonal elements

$$\underline{\underline{A}} = \begin{pmatrix} a_{xx} & a_{xy} & a_{xz} \\ a_{yx} & a_{yy} & a_{yz} \\ a_{zx} & a_{zy} & a_{zz} \end{pmatrix} \rightarrow \underline{\underline{A}}' = \begin{pmatrix} a_1 & 0 & 0 \\ 0 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix} \tag{1.76}$$

where the a_1, a_2, and a_3 correspond to the three eigenvalues λ_i. (The coordinate system reference of $\underline{\underline{A}}'$ is then contained in the orientation of the eigen vectors $\underline{v}_i$.) For a matrix in the principal axis system, the invariants are very easily computed:

$$tr(\underline{\underline{A}}') = I_{\underline{\underline{A}}'} \quad = I_{\underline{\underline{A}}} \quad = a_1 + a_2 + a_3 \tag{1.77}$$

$$II_{\underline{\underline{A}}'} \quad = II_{\underline{\underline{A}}} \quad = a_1 a_2 + a_1 a_3 + a_2 a_3 \tag{1.78}$$

$$\det(\underline{\underline{A}}') = III_{\underline{\underline{A}}'} \quad = III_{\underline{\underline{A}}} \quad = a_1 a_2 a_3. \tag{1.79}$$

See also sec. **??** for definitions of invariatns using deviators, such as for the deviatoric stress tensor.

Matrix decomposition Any quadratic tensor $\underline{\underline{A}}$ can be decomposed into a symmetric part $\underline{\underline{A}}^s$ (for which $a^s_{ij} = a^s_{ji}$) and an anti-symmetric part $\underline{\underline{A}}^a$ (for which $a^a_{ij} = -a^a_{ji}$) like $\underline{\underline{A}} = \underline{\underline{A}}^s + \underline{\underline{A}}^a$ (*Cartesian decomposition*). In the case of the deformation matrix $\underline{\underline{F}}$, we call the symmetric part *strain* $\underline{\underline{E}}$ (the infinitesimal strain tensor, $\underline{\underline{\epsilon}}$), and the anti-symmetric part corresponds to a rotation $\underline{\underline{R}}$. The *polar decomposition* is also of interest; we can write $\underline{\underline{F}} = \underline{\underline{R}}\,\underline{\underline{U}} = \underline{\underline{V}}\,\underline{\underline{R}}$ where $\underline{\underline{R}}$ is a rotation matrix and $\underline{\underline{U}}$ and $\underline{\underline{V}}$ are the right- and left-stretch matrices, respectively, and $\underline{\underline{V}} = \left(\underline{\underline{F}}\,\underline{\underline{F}}^T\right)^{1/2}$. The left-stretch matrix describes the deformation in the rotated coordinate system after the rotation $\underline{\underline{R}}$ has been applied to the body.

1.2.4 Tensors

The stress σ and strain ε are examples of second order (rank $r = 2$) tensors which, for $n = 3$, 3-D operations, have 3^r components and can be written as $n \times n$ matrices. You will see tensors also printed as E, we use the handwritten/blackboard version again, double underlining like $\underline{\underline{\varepsilon}}$, making no distinction between tensors and matrices.

Tensors in a Cartesian space are defined by their properties under coordinate transformation. If a quantity $\underline{v}$ remains intact under rotation to a new coordinate system $\underline{v}'$ such that

$$v'_i = L_{ij} v_j = \sum_{j=1}^{3} L_{ij} v_j \tag{1.80}$$

holds, then $\underline{v}$, a vector, is a first order tensor. L_{ij} may be, for example, a rotation matrix. Likewise, a second order tensor $\underline{\underline{T}}$ is defined by remaining intact after rotation into another coordinate system where it is expressed as $\underline{\underline{T}}'$ such that

$$T'_{ij} = L_{ik} T_{kl} L_{jl} = \sum_k L_{ik} \sum_l T_{kl} L_{jl} = \underline{\underline{L}}\,\underline{\underline{T}}\,\underline{\underline{L}}^T \tag{1.81}$$

Bibliography