Geological and Geophysical Sciences 539
Mechanics of the Earth
Fall Term 1971

Chapter 1: Review of Linear Algebra

Definition: A vector space is a set V of elements called vectors satisfying the following axioms

(1) addition
$$x + y = y + x$$
$$x + (y + z) = (x + y) + z$$
$$x + 0 = x$$
$$x + (-x) = 0$$

(2) scalar multiplication
$$a(bx) = (ab)x$$
$$1x = x$$

(3) scalar multiplication is distributive
$$a(x+y) = ax + ay$$
$$(a+b)x = ax + bx$$

For our purposes the scalars a, b may be either real $a, b \in R$ or complex $a, b \in C$.

Examples

1. let $a, b \in R$, let x, y denote ordinary vectors in Euclidean 3-space.

2. Let $\mathcal{P}$ be the set of all polynomials with complex coefficients, in a variable t. Let the scalars $a, b \in \mathbb{C}$. Then $\mathcal{P}$ is a vector space

3. <u>arrays</u> of a given order and dimension

examples of arrays

$A = (A_1, \cdots, A_n)$, an array of order 1 and dimension n

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & & & \vdots \\ \vdots & & & \vdots \\ A_{m1} & \cdots & \cdots & A_{mn} \end{bmatrix},$$

an array of order 2 and dimension $m \times n$ (m by n)

It has mn entries

$$A = \left\{ \begin{bmatrix} A_{111} & \cdots & A_{11n} \\ \vdots & & \vdots \\ A_{1m1} & \cdots & A_{1mn} \end{bmatrix}, \cdots\cdots, \begin{bmatrix} A_{\ell 11} & \cdots & A_{\ell 1n} \\ \vdots & & \vdots \\ A_{\ell m1} & \cdots & A_{\ell mn} \end{bmatrix} \right\},$$

an array of order 3 and dimension $\ell \times m \times n$

It is convenient to use the Einstein <u>index convention</u> for arrays

(1) $A_{ijk} = B_{ijk}$ means that A and B have the same order (here 3) and the same dimension, and that each entry of A = the corresponding entry of B, i.e., the equality holds for all possible values of i, j, k.

(2) $A_{ij}B_{jkl}$ stands for $\sum_{j=1}^{q} A_{ij}B_{jkl}$, the sum being over all possible values of j. To write $A_{ij}B_{jkl}$ implies that the possible values of j are the same for A and B. Thus if A is $l \times m$, then B must be $m \times p \times q$.

Definition: if A and B are arrays of the same order and dimension, and a and b are real or complex numbers, then $aA + bB$ is defined in the obvious way. For example for arrays of order 3.

$$(aA+bB)_{ijk} = aA_{ijk} + bB_{ijk}$$

Using the above definition it is easy to show that the arrays of a given order order and dimension with real (complex) entries constitute a vector space over the field of real (complex) numbers.

Array nomenclature

(1) an array of order 1 and dimension n is called an n-tuple. The vector space of real n-tuples is denoted R^n. The vector space of complex n-tuples is denoted C^n.

(2) an array of order 2 and dimension

$m \times n$ is called an <u>$m \times n$ matrix</u>.

A special matrix is the <u>identity matrix</u> I_n.
For an n, I_n is defined as

$$(I_n)_{ij} = \left\{ \begin{array}{ll} 1 & i=j \\ 0 & i \neq j \end{array} \right. , \quad i,j = 1, \cdots, n \Bigg\}$$

$$I_n = \underbrace{\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}}_{\leftarrow n \rightarrow} \updownarrow n$$

<u>The Kronecker delta δ</u>: This is an array of order 2 and dimension $n \times n$. It is in fact just I_n
That is $\delta_{ij} = 1$ if $i=j$ and $\delta_{ij} = 0$ if $i \neq j$
The notation is ambiguous, the same symbol δ is used, regardless of n. Often $n = 3$, of course.
<u>Properties of δ</u>

(1) $\delta_{ij} = \delta_{ji}$
(2) $\delta_{ii} = n$
(3) $\delta_{ij} B_{jkl} = B_{ikl}$

<u>The 3-dimensional alternating symbol ϵ</u>: This is an array of order 3 and dimension $3 \times 3 \times 3$. There are thus 27 entries. Its definition is

$$\left.\begin{array}{llll} \epsilon_{123} = & \epsilon_{231} = & \epsilon_{312} & = 1 \\ \epsilon_{213} = & \epsilon_{132} = & \epsilon_{321} & = -1 \end{array}\right\} \text{ six non-zero entries}$$

$\epsilon_{ijk} = 0$ otherwise (i.e., if there is a repeated index — 21 entries zero)

<u>Properties of ϵ:</u>

(1) $\epsilon_{ijk} = \epsilon_{jki} = \epsilon_{kij} = -\epsilon_{jik} = -\epsilon_{ikj} = -\epsilon_{kji}$

(2) if $A_{ij} = A_{ji}$ then $A_{ij}\,\epsilon_{ijk} = 0$

(3)
$$\epsilon_{ijk}\,\epsilon_{lmn} = \delta_{il}\delta_{jm}\delta_{kn} + \delta_{im}\delta_{jn}\delta_{kl} + \delta_{in}\delta_{jl}\delta_{km} - \delta_{im}\delta_{jl}\delta_{kn} - \delta_{il}\delta_{jn}\delta_{km} - \delta_{in}\delta_{jm}\delta_{kl}$$

(4) $\epsilon_{ijk}\,\epsilon_{imn} = \delta_{jm}\delta_{kn} - \delta_{jn}\delta_{km}$

(5) $\epsilon_{ijk}\,\epsilon_{ijl} = 2\delta_{kl}$

(6) if A is a 3×3 matrix, $\det A = \epsilon_{ijk}\,A_{1i}A_{2j}A_{3k}$

(7) suppose $\hat{x}_1, \hat{x}_2, \hat{x}_3$ are 3 mutually orthogonal unit vectors and $\vec{A} = A_i\hat{x}_i$ and $\vec{B} = B_i\hat{x}_i$. Then

$$\vec{A}\times\vec{B} = (\epsilon_{ijk}\,A_j B_k)\,\hat{x}_i \quad \text{or}$$
$$(\vec{A}\times\vec{B})_i = \epsilon_{ijk}\,A_j B_k$$

Problem 1:

(1) prove (1)–(7) above

(2) suppose $\hat{x}_1, \hat{x}_2, \hat{x}_3$ are mutually orthogonal unit vectors and $\vec{A} = A_i \hat{x}_i$, $\vec{B} = B_i \hat{x}_i$, $\vec{C} = C_i \hat{x}_i$ Write $\vec{A} \cdot \vec{B} \times \vec{C}$ in terms of A_i, B_i, C_i, and ϵ_{ijk}

(3) use (4) to show that $\vec{A} \times (\vec{B} \times \vec{C}) = (\vec{A} \cdot \vec{C}) \vec{B} - (\vec{A} \cdot \vec{B}) \vec{C}$

Enough about arrays. Consider again a general vector space V.

Vector space terminology

1. linear combination: if $u_1, \dots u_n \in V$ are vectors and $a_1, \dots a_n$ are scalars, the vector $v = a_i v_i = a_1 v_1 + \cdots a_n v_n$ is called a linear combination of $v_1, \dots v_n$.

2. linear dependence and independence: the vectors $u_1, \dots u_n$ are said to be linearly dependent if there are scalars $a_1, \dots a_n$, at least one of which is not zero, such that $a_i u_i = 0$. Conversely if $a_i u_i = 0$ implies $a_i = 0$, $i = 1, \dots, n$, then $u_1, \dots u_n$ are linearly independent.

Theorem: if $u_1 \neq 0$, then $u_1, \dots, u_n$ are linearly dependent if and only if some u_i is a linear combination of $u_1, \dots, u_{i-1}$.

7

<u>Example</u>: in 3-space

u_1 is linearly dependent $\Leftrightarrow u_1 = 0$

u_1, u_2 are " " $\Leftrightarrow$ they are colinear

u_1, u_2, u_3 are " " $\Leftrightarrow$ they are coplanar

<u>any</u> four vectors are linearly dependent

3. <u>subspace</u>: a subset M of a vector space V is called a subspace if $x, y \in M$ implies that every linear combination $ax + by \in M$.

<u>Example</u>: the subspaces of 3-space are $\{0\}$, all the straight lines thru 0, all the planes containing 0, and 3-space itself

4. space <u>spanned</u> or <u>generated</u> by $u_1, \dots u_n$: for given vectors $u_1, \dots u_n$, the set of all their linear combinations is a subspace denoted by $[\![u_1, \dots u_n]\!]$. We say that $u_1, \dots u_n$ span or generate the space $[\![u_1, \dots, u_n]\!]$.

5. <u>finite-dimensional vector space</u>: a space spanned by finitely many of its elements

<u>Theorem</u>: if $u_1, \dots u_m$ are linearly independent vectors in V and $v_1, \dots v_n$ span V, then $m \leq n$

6. <u>basis</u> : a basis for a vector space V is a linearly independent set of vectors which span V.

<u>Theorem</u> : Any finite-dimensional vector space has a basis and any two of its bases have the same number of elements. (This number is called the <u>dimension</u> of the space.

<u>Theorem</u> : If $v_1, \dots v_m$ are linearly independent vectors in an n-dimensional vector space V, then $m \leq n$ and there is a basis for V containing $v_1, \dots, v_m$.

<u>Theorem</u>: If $v_1, \dots, v_m$ span an n-dimensional vector space V, then $m \geq n$ and from among $v_1, \dots, v_m$, a basis for V may be chosen

<u>Theorem</u> If $v_1, \dots v_n$ are a basis for V and if $v \in V$, then there exist unique numbers $a_1, \dots, a_n$ such that

$v = a_1 v_1 + \dots + a_n v_n = a_i v_i$

These numbers are called the "coordinates" or "expansion coefficients" of v relative to the basis $v_1, \dots v_n$.

Proof of existence: since $v_1, \dots v_n$ span V, every vector in V is a linear combination of $v_1, \dots v_n$

Proof of uniqueness: if $v = a_i v_i = b_i v_i$, then $(a_i - b_i) v_i = 0$. But $v_1, \dots v_n$ are linearly independent so $a_i - b_i = 0$ or $a_i = b_i$.

<u>Definition</u>: If two subspaces U and V in a vector space W are disjoint ($U \cap V = 0$, where $\cap$ denotes set-theoretic intersection, i.e., U and V have only the zero vector in common) and U and V together span W, then W is said to be the <u>direct sum</u> of U and V, written $W = U \oplus V$.

<u>Theorem</u>: $W = U \oplus V$, iff every vector $z \in W$ can be written in the form $z = x + y$, $x \in U$, $y \in V$, in one and only one way.

<u>Theorem</u>: If $W = U \oplus V$ and $\dim U = m$, $\dim V = n$, then $\dim W = m + n$

<u>Theorem</u>: If W is an $(m+n)$-dimensional vector space and if U is an m-dimensional subspace of W, then there exists an n-dimensional subspace V in W such that $W = U \oplus V$. V is said to be the <u>complement</u> of U and U and V are called <u>complementary subspaces</u>.

Inner Product Spaces, Lengths, and Angles

Definition: Let V be a real (complex) vector space. An *inner product* in V is a rule which assigns to each ordered pair of vectors $u, v \in V$ a real (complex) number denoted by (u, v) in such a way that if u, v, w are any vectors $\in V$ and a and b are any real (complex) numbers, then

(i) $(u, v) = (v, u)^*$ ($*$ denotes complex conjugate)

(ii) $(u, av + bw) = a(u, v) + b(u, w)$

(iii) if $u \neq 0$, then $(u, u) > 0$, furthermore $(u, u) = 0$ iff $u = 0$.

It is easy to see that

(i) $(u, 0) = (0, u) = 0$ for any u

(ii) (u, u) is real for any u

(iii) $(au + bv, w) = a^*(u, w) + b^*(v, w)$

Examples:

1. V ordinary real 3-space. Let $(u, v) = \vec{u} \cdot \vec{v}$, the ordinary dot product.

2. V the space C^n of complex n-tuples $z = (z_1, \dots, z_n)$

   Let $(z, w) = z_1^* w_1 + \cdots z_n^* w_n = z_i^* w_i$

3. Let V be all complex-valued functions defined and continuous on the real interval $[a, b]$. If $f, g \in V$, define

$$(f, g) = \int_a^b f^*(x)\, g(x)\, dx$$

<u>Definition</u>: An inner product on V enables us to define the <u>length</u> $\|v\|$ of any vector $v \in V$ as

$$\|v\| = (v, v)^{1/2}$$

Secondly if V is a <u>real</u> vector space we can define the <u>angle</u> θ between any two non-zero vectors $u, v \in V$ as that angle $0 \le \theta \le \pi$ such that

$$\cos\theta = \frac{(u, v)}{\|u\|\,\|v\|}$$

How do we know that θ has the usual properties of an angle, i.e., how do we know that $\cos\theta$ lies between -1 and $+1$?
That it does is a fundamental property of inner products which can be stated even for <u>complex</u> inner product spaces. This property is known as Schwarz's inequality.

<u>Schwarz's inequality</u> $\quad |(u,v)| \leq \|u\|\,\|v\| \quad \star$

Proof: If u or v is zero $\star$ is trivial. Assume $\|u\| > 0$, $\|v\| > 0$. For fixed u, v and any two scalars a, b we have

$$(au + bv,\, au + bv) \geq 0$$

$$a^*a\,(u,u) + a^*b\,(u,v) + ab^*(v,u) + b^*b\,(v,v) \geq 0$$

but since $(u,v) = (v,u)^*$

$$|a|^2\|u\|^2 + |b|^2\|v\|^2 + a^*b\,(u,v) + (a^*b(u,v))^* \geq 0$$

$$-2\,\mathrm{Re}\,[a^*b\,(u,v)] \leq |a|^2\|u\|^2 + |b|^2\|v\|^2$$

This inequality true for any a, b

Suppose we write (u,v) in polar form

$$(u,v) = |(u,v)|\,e^{i\alpha}$$

If V is a real inner product space $\alpha = 0$ or π

Now set

$$a = \|u\|^{-1} e^{i\alpha} \quad , \quad b = -\|v\|^{-1}$$

$$2\,\mathrm{Re}\left[\frac{e^{-i\alpha}\,|(u,v)|\,e^{i\alpha}}{\|u\|\;\|v\|}\right] \leq 1 + 1$$

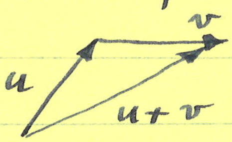
or

$$|(u,v)| \leq \|u\|\;\|v\|$$

We can also show that the definition of length behaves like a length. For one thing we have $\|av\| = |a|\,\|v\|$.

We also have the triangle inequality

__Triangle__ __inequality__: $\quad \| u + v \| \le \|u\| + \|v\|$



Proof: $\quad \| u+v \|^2 = (u+v, u+v) = (u,u)$

$$+ (u,v) + (v,u) + (v,v)$$
$$= \|u\|^2 + (u,v) + (u,v)^* + \|v\|^2$$
$$= \|u\|^2 + 2\,\mathrm{Re}\,(u,v) + \|v\|^2$$
$$\le \|u\|^2 + 2\,|(u,v)| + \|v\|^2$$
$$\le \|u\|^2 + 2\,\|u\|\,\|v\| + \|v\|^2$$
$$= (\|u\| + \|v\|)^2$$

Thus $\quad \|u+v\| \le \|u\| + \|v\|$

The most important relation among the vectors of an inner product space is orthogonality

__Definition__: Two vectors $u, v \in V$ are orth__ogonal__ if $\quad (u,v) = 0$

Note that if $(u,v) = 0$, then $(v,u) = 0$ also

Note also that in a real vector space, two vectors are orthogonal iff the angle between them is $\pi/2$.

__Definition__: A [redacted] set of vectors in a finite-dimensional vector space V is called an __orthonormal set__ $\{u_1, \cdots, u_n\}$ if $(u_i, u_j) = \delta_{ij}$. An orthonormal set of vectors are all mutually orthogonal and all have unit length.

It is an important result that in any finite dimensional inner product space V it is always possible to find a basis which consists of an an orthonormal set of vectors. Given a basis $v_1, \dots v_n$, there is a standard process, the <u>Gram - Schmidt Orthogonalization process</u>, for constructing another basis $u_1, \dots u_n$ which is orthonormal. Define

$$u_1 = v_1 / \| v_1 \|$$
$$w_2 = v_2 - (u_1, v_2) u_1$$
$$u_2 = w_2 / \| w_2 \|$$
$$w_3 = v_3 - (u_1 v_3) u_1 - (u_2, v_3) u_2$$
$$u_3 = w_3 / \| w_3 \|$$
$$w_4 = v_4 - (u_1 v_4) u_1 - (u_2, v_4) u_2 - (u_3, v_4) u_3$$
$$u_4 = w_4 / \| w_4 \|$$

etc.

In this construction w_i is always arranged to be orthogonal to $u_{i-1}, \dots u_1$ and w_i is a linear combination of $v_i, \dots v_1$ in which the coefficient of v_i is non-zero. Hence $\| w_i \| \neq 0$ and u_i is a well-defined unit vector orthogonal to $u_{i-1}, \dots u_1$.

Orthonormal bases are extremely useful for performing calculations. For example.

1. If $u_1, \cdots u_n$ is an orthonormal basis of an n-dimensional vector space V and $u = a_i u_i$, then the expansion coefficients a_i are given by

$$a_i = (u_i, u)$$

The quantities a_i are also called the <u>components</u> of u w.r.t. u_i

2. If $u = a_i u_i$ and $v = b_i u_i$, then the inner product (u, v) is given by

$$(u, v) = \overset{*}{a}_i b_i \quad \text{(Parseval's identity)}$$

3. If $u = a_i u_i$, then the length $\|u\|$ is given by

$$\|u\| = (\overset{*}{a}_i a_i)^{1/2} = \left(\sum_i |a_i|^2\right)^{1/2}$$

The relation given above in 3. is called the <u>completeness relation</u>

$$\|u\|^2 = \sum_i |(u_i, u)|^2$$

An orthonormal basis $v_1, \ldots, v_n$ of a vector space V is also called a <u>complete orthonormal set</u> of vectors in V.

<u>Problem</u> 2 (Projection theorem). Let V be an n-dimensional vector space with an inner product. Let W be a subspace of V, $\dim W = m$. Define $W_\perp$ (read W perp, the <u>orthogonal</u> <u>complement</u> of W) to be the set of all vectors in V which are orthogonal to <u>every</u> member of W.

(1) Prove that $W_\perp$ is a subspace of V.

(2) show that $V = W \oplus W_\perp$, and hence $\dim W_\perp = n-m$

Change of orthonormal basis

Suppose that V is an inner product space, and that $v_1, \dots v_n$ and $v_1', \dots v_n'$ are two different orthonormal bases for V.
The vectors in one basis may of course be expressed as a linear combination of the vectors in the other basis

$$v_j' = (v_i, v_j')\, v_i$$
$$v_j = (v_i', v_j)\, v_i'$$

Now interchange i and j in the second of the above two eqns and substitute the first eqn into the second

$$\begin{aligned} v_i &= (v_j', v_i)\, v_j' \\ &= (v_j', v_i)(v_k, v_j')\, v_k \\ &= (v_k', v_i)(v_j, v_k')\, v_j \\ &= (v_j, v_k')(v_k', v_i)\, v_j \end{aligned}$$

Thus

$$(v_j, v_k')(v_k', v_i) = \delta_{ij}$$ also
$$(v_i, v_k')(v_k', v_j) = \delta_{ij}$$

Now denote $Q_{ij} = (v_i', v_j)$
Q is an $n \times n$ matrix
Now for any $n \times n$ matrix, one defines the conjugate transpose or adjoint matrix by the relation

$$(A^+)_{ij} = A_{ji}^* \qquad (* \text{ means complex conjugate})$$

The adjoint of the matrix Q is

$$Q^+_{ij} = Q_{ji}^* = (v_j', v_i)^*$$
$$= (v_i, v_j')$$

Now note that the relation
$$(v_i, v_k')(v_k', v_j) = \delta_{ij}$$
may be written in terms of Q_{ij} and Q^+_{ij}

$$Q^+_{ik} Q_{kj} = \delta_{ij}$$
or in matrix notation
$$Q^+ Q = I$$
One may also show that $QQ^+ = I$
Hence the matrix Q and its adjoint Q^+ are inverses of one another. Such a matrix is said to be unitary.

The above discussion may thus be summarized by the statement

If $\{v_i\}$ and $\{v_i'\}$ are any two orthonormal bases of an n dimensional vector space V, then they are related to each other by means of the unitary matrix Q where $QQ^+ = Q^+Q = I$ and

$$Q_{ij} = (v_i', v_j)$$
$$v_j' = Q^+_{ij} v_i$$
$$v_j = Q_{ij} v_i'$$

In matrix notation $v' = vQ^+$

$$v = v'Q$$

Now consider an arbitrary vector $v \in V$. It may be written in terms of $\{v_i\}$ and $\{v_i'\}$

$$v = a_i v_i = a_i' v_i'$$

where the components a_i and a_i' are given by

$$a_i = (v_i, v)$$
$$a_i' = (v_i', v)$$

How are the components a_i' relative to the orthonormal basis $\{v_i\}$ related to the components a_i relative to the basis $\{v_i'\}$? The answer is by means of the unitary matrix Q.

$$a_i' = (v_i', v) = (v_i', a_j v_j) = (v_i', v_j)\, a_j = Q_{ij} a_j$$

likewise

$$a_i = (v_i, v) = (v_i, a_j' v_j') = (v_i, v_j') a_j' = Q_{ij}^+ a_j'$$

In summary

$$a_i' = Q_{ij} a_j$$
$$a_i = Q_{ij}^+ a_j'$$

or in matrix notation

$$a' = Qa$$
$$a = Q^+ a'$$

<u>Operations on vector spaces.</u>

Definition: If f is a rule which assigns to each member of a set U a unique member of the set V, then we say f maps U <u>into</u> V, written $f: U \to V$. The member of V assigned to $u \in U$ is written $f(u)$.

Definition: If $f: U \to V$ and $g: V \to W$ then gf is that rule which assigns to any $u \in U$ the member $g[f(u)] \in W$. Thus $gf: U \to W$

Definition: Let $f: U \to V$ and $g: U \to V$. Suppose V is a real (complex) vector space. Let a, b be real (complex) numbers and define $af + bg$ to be that rule which assigns to $u \in U$ the vector $af(u) + bg(u)$ $\in$ the vector space V. Note that with this definition the set of all mappings of U into a vector space V is itself a vector space.

Definition: Let V and W be real (or complex) vector spaces. A mapping $f: V \to W$ is linear if for any real (or complex) scalars a, b any any $u, v \in V$ we have

$$f(au + bv) = af(u) + bf(v)$$

If $V = W$, f is a linear operator on V.

Examples:

1. consider the mapping $f: V \to V$ defined by $f(v) = v$. This operator is called the identity operator for obvious reasons. It is often denoted by I. $I(v) = v$.
2. consider $f: R^3 \to R^3$ (real 3-space) $f(v) = \vec{\Omega} \times \vec{v}$ where $\vec{\Omega}$ is a fixed vector.
3. $f(v) = av$ for any v, a a fixed scalar
4. Consider $f: V \to V$, V a finite-dimensional vector space. Let $v_1, \dots, v_n$ be an orthonormal basis for V and let A_{ij} be any $n \times n$ matrix of fixed real

or complex numbers. For any vector $v \in V$ we can write $v = a_i v_i$. Now define
$A(v) = v_i A_{ij} a_j$. Clearly $A(v)$ is a vector $\in V$ and depends linearly on v. Hence A is a linear operator.
$A(v)$ is commonly written Av
Note that if the vector v is taken to be one of the basis vectors, say v_j, then one has
$A v_j = v_i A_{ij}$

In fact all linear operators on a finite-dimensional vector space V are of the form of the operator A in example 4 above. This can be seen by doing

Problem 3: Let V be an n-dimensional vector space and let $A: V \to V$ be a linear operator on V. Let $v_1, \dots v_n$ be an orthonormal basis. Define $y_1, \dots y_n$ to be the vectors in V obtained by the application of A to $v_1, \dots v_n$.

$$A v_i = y_i$$

Now $y_i \in V$ and hence may be written as a linear combination of the basis elements v_i

$$y_i = A v_i = A_{ij} v_i .$$

The $n \times n$ matrix A is called the matrix of A relative to the basis $\{v_i\}$. Thus with the aid of a fixed basis $\{v_i\}$ we have made correspond an $n\times n$ matrix A_{ij} with any linear operator $A: V \to V$.

1. Prove that knowledge of the matrix A_{ij} completely determines the operator A.

2. If $\{v_1', \dots v_n'\}$ is another orthonormal basis for V and $\tilde{A}_{ij}$ is the matrix of A relative to $\{v_i'\}$, how are the matrices A and $\tilde{A}$ related. Express your answer in terms of the unitary matrix Q which defines the change of orthonormal basis.

Chapter 2: Spherical Harmonics

1. Preliminaries

Let Ω denote the surface of the unit sphere centered on $\vec{0}$ in ordinary 3-space.

Definition: The vector space $\mathcal{L}_2(\Omega)$ consists of all complex valued functions $f: \Omega \to C$ such that $\int_\Omega f^* f \, dA < \infty$

Such functions are said to be square integrable.

On this vector space an inner product is defined

$$(f, g) = \int_\Omega f^* g \, dA \qquad (*)$$

Thus $\mathcal{L}_2(\Omega)$ is a complex inner product space.

We are particularly interested in examining two subspaces of $\mathcal{L}_2(\Omega)$ which we will now define.

Definition: A polynomial in 3-space with complex coefficients $P_\ell(\vec{r})$ is called a homogeneous polynomial if it involves only terms of total degree ℓ, i.e.

$$P_\ell(\vec{r}) = \sum C_{\alpha\beta\gamma} x^\alpha y^\beta z^\gamma \text{ where } \alpha + \beta + \gamma = \ell$$

Note that if a homogeneous polynomial is known for all unit vectors $\hat{r}$ (i.e. for all $\vec{r} \in \Omega$, on the unit sphere) then it is in fact known for all $\vec{r}$ since

$$P_\ell(\vec{r}) = P_\ell(r\hat{r}) = r^\ell P_\ell(\hat{r})$$

It thus suffices to consider the restriction of homogeneous polynomials to the surface of the unit sphere Ω

Definition: $\mathcal{P}_\ell$ = vector space of all polynomials in 3-space homogeneous of degree ℓ and having complex coefficients.

$\mathcal{P}_\ell(\Omega)$ = space of all functions $P_\ell: \Omega \to C$ defined by restricting the elements of $\mathcal{P}_\ell$ to Ω.

Each element $P_\ell \in \mathcal{P}_\ell$ determines a unique element $R_\ell \in \mathcal{P}_\ell(\Omega)$ and vice-versa.

Problem 4: Show that $\mathcal{P}_\ell(\Omega)$ is a subspace of $\mathcal{L}_2(\Omega)$ and that $\mathcal{P}_\ell$ is a complex inner product space with inner product (*). Show that the dimension of $\mathcal{P}_\ell$ is $\dfrac{(\ell+1)(\ell+2)}{2}$.

Definition: A scalar function $\phi(\vec{r})$ in three space is harmonic in an open set V of R^3 if $\nabla^2\phi = 0$ in V.

Example: if $\phi(\vec{r})$ is the gravitational potential of a mass distribution lying outside a volume V, then ϕ is harmonic in V.

Definition: $\mathcal{H}_\ell$ = the vector space of all harmonic members of $\mathcal{P}_\ell$.
(If a polynomial is harmonic in any open set, it is harmonic everywhere.)

Any member of $\mathcal{H}_\ell$ is called a solid spherical harmonic of degree ℓ

$\mathcal{H}_\ell(\Omega)$ = the space of all functions $Y_\ell : \Omega \to C$ defined by restricting the elements of $\mathcal{H}_\ell$ to Ω.

The element $Y_\ell \in \mathcal{H}_\ell(\Omega)$ determined by $H_\ell \in \mathcal{H}_\ell$ is called the surface spherical harmonic of degree ℓ determined by H_ℓ.

$\mathcal{H}_\ell$ is an inner product space with inner product (*) and is a subspace of $\mathcal{P}_\ell$.

2. The operator $\vec{\Lambda}$ (the infinitesimal rotation operator)

Let $f(\vec{r})$ be a scalar field defined in an open set V of R^3. The field $f(\vec{r})$ is said to be differentiable at the point $\vec{r} \in V$ if there exists a vector $\vec{F}$, depending on $\vec{r}$, such that

$$f(\vec{r}+\delta\vec{r}) = f(\vec{r}) + \delta\vec{r}\cdot\vec{F} + O(|\delta\vec{r}|^2)$$

The vector $\vec{F}$ is generally written $\vec{\nabla}f(\vec{r})$ and is called the <u>gradient</u> of the scalar field f at $\vec{r}$

The gradient $\vec{\nabla}$ may be considered to be a linear operator which maps a differentiable scalar field $f(\vec{r})$ into a vector field $\vec{\nabla}f(\vec{r})$.

The gradient is of course a very familiar linear operator. We wish to consider a related operator $\vec{\Lambda}$, which also maps a differentiable scalar field $f(\vec{r})$ into a vector field $\vec{\Lambda}f(\vec{r})$.

Consider

$$f(\vec{r}+\delta\vec{\omega}\times\vec{r}) = f(\vec{r}) + (\delta\vec{\omega}\times\vec{r})\cdot\vec{\nabla}f(\vec{r}) + O(|\delta\omega|^2)$$

$$= f(\vec{r}) + \delta\vec{\omega}\cdot\vec{r}\times\vec{\nabla}f(\vec{r}) + O(|\delta\omega|^2)$$

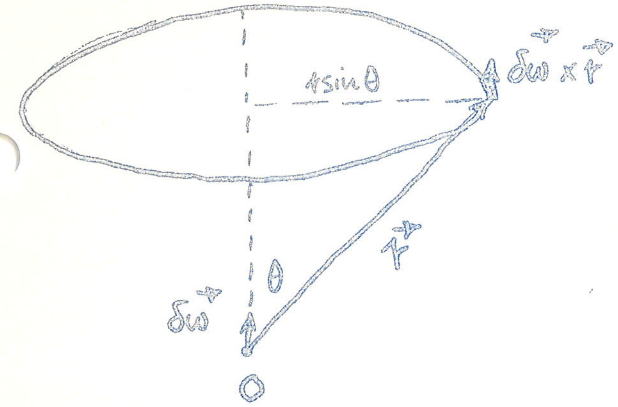
$$\boxed{\text{Now define } \vec{\Lambda} = \vec{r}\times\vec{\nabla}}$$

$$\boxed{f(\vec{r}+\delta\vec{\omega}\times\vec{r}) = f(\vec{r}) + \delta\vec{\omega}\cdot\vec{\Lambda}f(\vec{r}) + O(|\delta\vec{\omega}|^2)}$$

$\vec{\Lambda}$ is called the <u>infinitesimal</u> <u>rotation operator</u>.

Why the name?

The vector $\delta\vec{\omega}\times\vec{r}$ is perpendicular to both $\delta\vec{\omega}$ and $\vec{r}$ (right hand rule) and has magnitude $|\delta\vec{\omega}||\vec{r}|\sin\theta$.



From the figure it is clear that the vector $\vec{r} + \vec{\delta\omega} \times \vec{r}$ is obtained from $\vec{r}$ by rotating the whole space rigidly thru the small angle $|\vec{\delta\omega}|$ about the axis $\vec{\delta\omega}$. (This is not quite true, there being an error of order $|\vec{\delta\omega}|^2$). The linear vector operator $\vec{\Lambda}$ allows one to compute the value of the scalar field f at the infinitesimally rotated vector $\vec{r} + \vec{\delta\omega} \times \vec{r}$, via the boxed equation above. By the same token the gradient operator $\vec{\nabla}$ might also be called the infinitesimal translation operator since it allows one to compute the value of f at the infinitesimally translated vector $\vec{r} + \delta\vec{r}$.

Properties of $\vec{\Lambda}$ easy in Cartesian coordinates

Let $\hat{x}, \hat{y}, \hat{z}$ be a right-handed Cartesian axis system. Define the three scalar linear operators $\Lambda_x, \Lambda_y, \Lambda_z$ by

$$\Lambda_x f = \hat{x} \cdot \vec{\Lambda} f$$
$$\Lambda_y f = \hat{y} \cdot \vec{\Lambda} f$$
$$\Lambda_z f = \hat{z} \cdot \vec{\Lambda} f$$

Let $\hat{x} = \hat{x}_1, \hat{y} = \hat{x}_2, \hat{z} = \hat{x}_3$

$$\vec{r} = r_i \hat{x}_i, \quad \partial_i = \frac{\partial}{\partial r_i}$$

Then $\vec{\Lambda} = \vec{r} \times \vec{\nabla} = \Lambda_i \hat{x}_i$ where

$$\Lambda_i = \epsilon_{ijk} r_j \partial_k$$

or $\vec{\Lambda} = \Lambda_x \hat{x} + \Lambda_y \hat{y} + \Lambda_z \hat{z}$ where

$$\Lambda_x = y\partial_z - z\partial_y, \quad \Lambda_y = z\partial_x - x\partial_z, \quad \Lambda_z = x\partial_y - y\partial_x$$

1. $\boxed{\Lambda^2 = r^2\nabla^2 - r_k\partial_k(r_j\partial_j + 1)}$

$\boxed{\text{Note: } \Lambda^2 \text{ is defined by } \Lambda^2 f = \hat{r}\times\nabla\cdot(\hat{r}\times\nabla)f}$

Proof: $\Lambda^2 = \Lambda_i\Lambda_i = (\epsilon_{ijk} r_j \partial_k)(\epsilon_{ilm} r_l \partial_m)$

$= \epsilon_{ijk}\epsilon_{ilm} r_j\partial_k r_l \partial_m$

$= (\delta_{jl}\delta_{km} - \delta_{jm}\delta_{kl}) r_j \partial_k r_l \partial_m$

$= r_j\partial_k r_j \partial_k - r_j\partial_k r_k \partial_j$

$= r_j r_j \partial_k\partial_k + r_j\delta_{kj}\partial_k - r_j(\partial_k r_k)\partial_j - r_j r_k \partial_k\partial_j$

$= r^2\nabla^2 + r_j\partial_j - 3r_j\partial_j - r_j r_k\partial_k\partial_j$

$= r^2\nabla^2 - 2r_j\partial_j - r_k\partial_k(r_j\partial_j) + r_j\partial_j$

$= r^2\nabla^2 - r_k\partial_k(r_j\partial_j) - r_k\partial_k$

2. $\boxed{\vec{\Lambda}\times\vec{\Lambda} = -\vec{\Lambda}}$

Proof: $(\vec{\Lambda}\times\vec{\Lambda})_i = \epsilon_{ijk}(\epsilon_{jlm} r_l\partial_m)(\epsilon_{kpq} r_p \partial_q)$

$= (\epsilon_{ijk}\epsilon_{jlm}\epsilon_{kpq}) r_l \partial_m r_p \partial_q$

$= (\delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp})\epsilon_{jlm}(r_l r_p \partial_m\partial_q + r_l\delta_{mp}\partial_q)$

$= (\delta_{ip}\epsilon_{qlm} - \delta_{iq}\epsilon_{plm})(r_l r_p\partial_m\partial_q + r_l\delta_{mp}\partial_q)$

$= \delta_{ip}\epsilon_{qlm} r_l \delta_{mp}\partial_q = \epsilon_{qli} r_l \partial_q = -\epsilon_{ilq} r_l\partial_q = -\Lambda_i$

Note this means that $\Lambda_x\Lambda_y - \Lambda_y\Lambda_x = -\Lambda_z$, etc.

Now let us define the <u>commutator</u> $[A,B]$ of any two linear operators A and B acting on the same vector space V by the relation

$[A,B] = AB - BA$ Note that $[A,B] = -[B,A]$

Then we have shown that $[\Lambda_x, \Lambda_y] = -\Lambda_z$

$$[\Lambda_y, \Lambda_z] = -\Lambda_x$$

$$[\Lambda_z, \Lambda_x] = -\Lambda_y$$

or rewritten in a single equation

$$\boxed{[\Lambda_i, \Lambda_j] = -\epsilon_{ijk}\Lambda_k}$$

3. $\Lambda^2 = (\Lambda_x + i\Lambda_y)(\Lambda_x - i\Lambda_y) + \Lambda_z^2 - i\Lambda_z$

$$= (\Lambda_x - i\Lambda_y)(\Lambda_x + i\Lambda_y) + \Lambda_z^2 + i\Lambda_z$$

Proof:

$$(\Lambda_x + i\Lambda_y)(\Lambda_x - i\Lambda_y) = \Lambda_x^2 + \Lambda_y^2 + i(\Lambda_y\Lambda_x - \Lambda_x\Lambda_y)$$

$$= \Lambda_x^2 + \Lambda_y^2 + i\Lambda_z$$

Thus

$$(\Lambda_x + i\Lambda_y)(\Lambda_x - i\Lambda_y) - i\Lambda_z + \Lambda_z^2 = \Lambda_x^2 + \Lambda_y^2 + \Lambda_z^2 = \Lambda^2$$

If we define two operators

$$\boxed{\begin{aligned} \Lambda_+ &= \Lambda_x + i\Lambda_y \\ \Lambda_- &= \Lambda_x - i\Lambda_y \end{aligned}}$$

Then the above two relations may be written

$$\boxed{\Lambda^2 = \Lambda_+\Lambda_- + \Lambda_z^2 - i\Lambda_z = \Lambda_-\Lambda_+ + \Lambda_z^2 + i\Lambda_z}$$

It is simple to obtain the commutation relations for Λ_+ and Λ_-

$$\boxed{\begin{aligned} [\Lambda_+, \Lambda_-] &= 2i\Lambda_z \\ [\Lambda_+, \Lambda_z] &= -i\Lambda_+ \\ [\Lambda_-, \Lambda_z] &= i\Lambda_- \end{aligned}}$$

4. $\nabla^2 \vec{\Lambda} = \vec{\Lambda} \nabla^2$

In other words ∇^2 and $\vec{\Lambda}$ <u>commute</u> $\boxed{[\nabla^2, \vec{\Lambda}] = 0}$

Proof:
$$\partial_\ell \partial_\ell (\epsilon_{ijk} r_j \partial_k) = \epsilon_{ijk} \partial_\ell (\delta_{j\ell} \partial_k + r_j \partial_\ell \partial_k)$$
$$= \epsilon_{ijk} [\delta_{j\ell} \partial_\ell \partial_k + \delta_{j\ell} \partial_\ell \partial_k + r_j \partial_\ell \partial_\ell \partial_k]$$
$$= 2\epsilon_{i\ell k} \partial_\ell \partial_k + \epsilon_{ijk} r_j \partial_k \partial_\ell \partial_\ell$$
$$= (\epsilon_{ijk} r_j \partial_k) \partial_\ell \partial_\ell$$

Properties of $\vec{\Lambda}$ easy in spherical coordinates

1. $\vec{\Lambda} = \vec{r} \times \vec{\nabla} = \vec{r} \times \left(\hat{r} \partial_r + \hat{\theta} \frac{1}{r} \partial_\theta + \hat{\phi} \frac{1}{r\sin\theta} \partial_\phi \right)$

$$= -\hat{\theta} \frac{1}{\sin\theta} \partial_\phi + \hat{\phi} \partial_\theta$$

Thus $\vec{\Lambda}$ commutes with ∂_r and with any function of r alone

$$\boxed{\begin{aligned} [\partial_r, \vec{\Lambda}] &= 0 \\ [f(r), \vec{\Lambda}] &= 0 \end{aligned}}$$

$\vec{\Lambda} f(\vec{r}) = \vec{\Lambda} f(r, \theta, \phi)$ depends only on the values of f on Ω_r, the sphere of radius r.

2. $\hat{\theta} = \hat{x}\cos\theta\cos\phi + \hat{y}\cos\theta\sin\phi - \hat{z}\sin\theta$

$\hat{\phi} = -\hat{x}\sin\phi + \hat{y}\cos\phi$

$$\vec{\Lambda} = \hat{\theta}\left(-\frac{1}{\sin\theta}\partial_\phi\right) + \hat{\phi}\,\partial_\theta$$

$$= \hat{x}\left(-\sin\phi\,\partial_\theta - \cot\theta\cos\phi\,\partial_\phi\right) + \hat{y}\left(\cos\phi\,\partial_\theta - \cot\theta\sin\phi\,\partial_\phi\right) + \hat{z}\,\partial_\phi$$

Thus

$$\Lambda_x = -\sin\phi\,\partial_\theta - \cot\theta\cos\phi\,\partial_\phi$$

$$\Lambda_y = \cos\phi\,\partial_\theta - \cot\theta\sin\phi\,\partial_\phi$$

$$\boxed{\Lambda_z = \partial_\phi}$$

3. $$\boxed{\begin{aligned} \Lambda_+ &= ie^{i\phi}\left[\partial_\theta + i\cot\theta\,\partial_\phi\right] \\ \Lambda_- &= ie^{-i\phi}\left[-\partial_\theta + i\cot\theta\,\partial_\phi\right] \end{aligned}}$$

4. Now $\Lambda^2 = \Lambda_+\Lambda_- + \Lambda_z^2 - i\Lambda_z$

$$= e^{i\phi}\left(\partial_\theta + i\cot\theta\,\partial_\phi\right)e^{-i\phi}\left(\partial_\theta - i\cot\theta\,\partial_\phi\right) + \partial_\phi^2 - i\partial_\phi$$

$$= \partial_\theta^2 + i\left(\csc^2\theta - \cot^2\theta\right)\partial_\phi + \cot\theta\,\partial_\theta + \left(\cot^2\theta + 1\right)\partial_\phi^2 - i\partial_\phi$$

$$= \partial_\theta^2 + \cot\theta\,\partial_\theta + \frac{1}{\sin^2\theta}\partial_\phi^2$$

$$\boxed{\Lambda^2 = \partial_\theta^2 + \cot\theta\,\partial_\theta + \frac{1}{\sin^2\theta}\partial_\phi^2 = \frac{1}{\sin\theta}\partial_\theta\left(\sin\theta\,\partial_\theta\right) + \frac{1}{\sin^2\theta}\partial_\phi^2}$$

5. $\boxed{[\vec{\Lambda}, \Lambda^2] = 0}$

Proof: since the Cartesian axis system $\hat{x}, \hat{y}, \hat{z}$ is arbitrary, it suffices to show it for Λ_z.
But $\Lambda_z = \partial_\phi$ and $\Lambda^2 = \frac{1}{\sin\theta}\partial_\theta(\sin\theta\,\partial_\theta) + \frac{1}{\sin^2\theta}\partial_\phi^2$, and these clearly commute.

6. $\boxed{\nabla^2 = \frac{1}{r}\partial_r^2 r + \frac{1}{r^2}\Lambda^2}$

This is the formula for ∇^2 in spherical coordinates written in terms of the operator Λ^2.

Proof:
$$r_i \partial_i = \vec{r}\cdot\vec{\nabla} = \vec{r}\cdot\left(\hat{r}\,\partial_r + \frac{1}{r}\hat{\theta}\,\partial_\theta + \frac{1}{r\sin\theta}\hat{\phi}\,\partial_\phi\right)$$
$$= r\partial_r$$

Now from above
$$\Lambda^2 = r^2\nabla^2 - r_k\partial_k(r_j\partial_j + 1)$$
$$= r^2\nabla^2 - r\partial_r(r\partial_r + 1)$$
$$= r^2\nabla^2 - r\partial_r^2 r$$

Thus $$\nabla^2 = \frac{1}{r}\partial_r^2 r + \frac{1}{r^2}\Lambda^2$$

7. We can now offer a more straightforward proof of the relation

$\boxed{[\nabla^2, \vec{\Lambda}] = 0}$

Proof: $\nabla^2 = \frac{1}{r}\partial_r^2 r + \frac{1}{r^2}\Lambda^2$, and $[\vec{\Lambda}, \Lambda^2] = [\vec{\Lambda}, \partial_r] = [\vec{\Lambda}, f(r)]$
$= 0$

8. one last thing. $\boxed{\text{If } H_\ell \in \mathcal{H}_\ell \text{, then } \Lambda^2 H_\ell = -\ell(\ell+1) H_\ell}$

Proof: $\nabla^2 H_\ell = 0$ so

$$\frac{1}{r}\partial_r^2 r\left[r^\ell H_\ell(\hat{r})\right] + \frac{1}{r^2}\Lambda^2\left[r^\ell H_\ell(\hat{r})\right] = 0$$

$$(\ell+1)\ell r^{\ell-2} H_\ell(\hat{r}) + \frac{1}{r^2}\Lambda^2 r^\ell H_\ell(\hat{r}) = 0$$

dividing by $r^{\ell-2}$

$$\Lambda^2 H_\ell(\hat{r}) = -\ell(\ell+1) H_\ell(\hat{r})$$

3. <u>The operator $\vec{L}$</u>, <u>the angular momentum operator</u>

<u>Definition</u>: If V is finite dimensional complex inner product space and $B: V \to V$ is a linear operator on V, then the <u>adjoint</u> of B, written B^*, is that unique linear operator such that for any $f, g \in V$

$$(f, Bg) = (B^* f, g)$$

Problem 4:

(1) prove that B^* is linear and that $(B^*)^* = B$

(2) prove that if a is any complex number $(aB)^* = a^* B^*$

(3) prove that if A and B are two linear operators $(AB)^* = B^* A^*$

(4) Let $\{v_1, \dots v_n\}$ be an orthonormal basis for V and let $[B]$ be the <u>matrix</u> of the linear operator B relative to $\{v_i\}$. show that $[B^*]$ is $B^\dagger$, the adjoint matrix

(5) Let $\{v_i'\}$ be another orthonormal basis obtained from $\{v_i\}$ by means of an orthogonal matrix Q. Use the result of problem 3 to show that $[B^*] = B^\dagger$ in every orthonormal basis system.

<u>Definition</u>: An operator $B: V \to V$ is <u>self-adjoint</u> or <u>Hermitian</u> if $B^* = B$. If $B^* = -B$, it is anti Hermitian

Problem 5: Show that <u>any</u> linear operator may be written as the sum of a Hermitian linear operator and an anti-Hermitian linear operator

Now define $\vec{L} = \frac{1}{i}\vec{\Lambda}$	$\hbar\vec{L}$ is the quantum mechanical angular momentum operator

Theorem: $L_j : \mathcal{P}_\ell \to \mathcal{P}_\ell$ and $L_j : \mathcal{H}_\ell \to \mathcal{H}_\ell$, $j = 1, 2, 3$ and L_j are Hermitian (thus Λ_j are anti-Hermitian)

Proof: (i) $L_z : \mathcal{P}_\ell \to \mathcal{P}_\ell$ since $L_z = \frac{1}{i}(x\partial_y - y\partial_x)$ and if $P_\ell \in \mathcal{P}_\ell$ so is $L_z P_\ell$. Same for L_x, L_y

(ii) $L_z : \mathcal{H}_\ell \to \mathcal{H}_\ell$ since if $H_\ell \in \mathcal{H}_\ell$ then $L_z H_\ell \in \mathcal{P}_\ell$ and we must show that $\nabla^2 L_z H_\ell = 0$, but $\nabla^2 L_z H_\ell = \frac{1}{i}\nabla^2 \Lambda_z H_\ell$ $= \frac{1}{i}\Lambda_z \nabla^2 H_\ell = \frac{1}{i}\Lambda_z 0 = 0$. Same for L_x, L_y.

(iii) Now we want to prove that $L_z : \mathcal{P}_\ell \to \mathcal{P}_\ell$ is Hermitian, $L_z^* = L_z$. That is $(L_z f, g) = (f, L_z g)$ for any $f, g \in \mathcal{P}_\ell$.

But $(f, L_z g) = \int_\Omega f^* L_z g \, dA = \frac{1}{i}\int_\Omega f^* \partial_\phi g \, dA$

$$= \frac{1}{i}\int_0^\pi d\theta \sin\theta \int_0^{2\pi} d\phi \, f^* \partial_\phi g \qquad \text{integrate by parts}$$

$$= -\frac{1}{i}\int_0^\pi \sin\theta \, d\theta \int_0^{2\pi} d\phi \, \partial_\phi f^* g = -\frac{1}{i}\int_\Omega (\partial_\phi f^*) g \, dA$$

$$= \int_\Omega \left(\frac{1}{i}\partial_\phi f\right)^* g \, dA = \int_\Omega (L_z f)^* g \, dA = (L_z f, g)$$

Since $\hat{x}, \hat{y}, \hat{z}$ is arbitrary, it is clear that L_x, L_y are Hermitian also, but just in case you don't believe it

Problem 6: Given a Cartesian axis system $\hat{x}, \hat{y}, \hat{z}$, show that L_y, and L_x are Hermitian.

It is clear that if $L_i : \mathcal{P}_\ell \to \mathcal{P}_\ell$ is Hermitian, then so is $L_i : \mathcal{H}_\ell \to \mathcal{H}_\ell$

Corollaries:

1. $L^2 = -\Lambda^2$ is a self adjoint linear operator on $\mathcal{P}_\ell$ and $\mathcal{H}_\ell$

2. $L_+ = \frac{1}{i}\Lambda_+$ and $L_- = \frac{1}{i}\Lambda_-$ are Hermitian conjugate operators on $\mathcal{P}_\ell$ and $\mathcal{H}_\ell$, that is
$L_+^* = L_-$ and $L_-^* = L_+$

Proof: 1. $(L^2)^* = (L_xL_x + L_yL_y + L_zL_z)^* = L_x^*L_x^* + L_y^*L_y^* + L_z^*L_z^*$
$= L^2$

2. $L_+^* = (L_x + iL_y)^* = L_x^* - iL_y^* = L_x - iL_y = L_-$

Useful $\vec{L}$ identities easily obtainable from $\vec{\Lambda}$ identities are

$$\vec{L} \times \vec{L} = i\vec{L}$$

$$L^2 = L_+L_- + L_z^2 - L_z = L_-L_+ + L_z^2 + L_z$$

Commutation relations

$$[L^2, \vec{L}] = [L^2, L_z] = [L^2, L_\pm] = 0$$

$$[L_z, L_+] = L_+$$

$$[L_z, L_-] = -L_-$$

$$[L_+, L_-] = 2L_z$$

4. The vector space $\mathcal{P}_\ell$

Definition: Two subspaces U and V of an inner product space W are orthogonal if for every $u \in U$ and $v \in V$, the inner product $(u, v) = 0$. We write $U \perp V$

If $l \neq l'$, then $\mathcal{H}_l \perp \mathcal{H}_{l'}$

Proof: Take $H_l \in \mathcal{H}_l$ and $H_{l'} \in \mathcal{H}_{l'}$

$$(H_l, \Lambda^2 H_{l'}) = (\Lambda^2 H_l, H_{l'})$$

$$-l'(l'+1)(H_l, H_{l'}) = -l(l+1)(H_l, H_{l'})$$

Thus $[l'(l'+1) - l(l+1)](H_l, H_{l'}) = 0$ or

$$(l'-l)(l'+l+1)(H_l, H_{l'}) = 0$$

hence if $l' \neq l$, $(H_l, H_{l'}) = 0$

Now define $r^\nu \mathcal{H}_l$ to be the space of all functions of the form $r^\nu H_l$ where $H_l \in \mathcal{H}_l$. Then we have the following

<u>Theorem</u>: The space $\mathcal{P}_l$ has a unique direct sum decomposition of the form

$$\boxed{\mathcal{P}_l = \mathcal{H}_l \oplus r^2 \mathcal{H}_{l-2} \oplus \cdots \oplus r^{l-\tau(l)} \mathcal{H}_{\tau(l)}}$$

where $\tau(l) \equiv l \bmod 2 = 0$ if l even, 1 if l odd.

The assertion is that for any $P_l \in \mathcal{P}_l$ there exist unique harmonic polynomials $H_l \in \mathcal{H}_l$, $H_{l-2} \in \mathcal{H}_{l-2}, \dots, H_{\tau(l)} \in \mathcal{H}_{\tau(l)}$ such that

$$\boxed{P_l = H_l + r^2 H_{l-2} + \cdots + r^{l-\tau(l)} H_{\tau(l)}}$$

Proof:

(i) uniqueness: say $\tilde{P}_\ell$ and $\tilde{\tilde{P}}_\ell$ both have harmonic polynomial expansions of the above form. Then their difference $P_\ell = \tilde{P}_\ell - \tilde{\tilde{P}}_\ell$ does also. Look on Ω. There if $P_\ell = 0$, then $(P_\ell, H_{\ell-2\nu}) = 0$, but because $\mathcal{H}_\ell \perp \mathcal{H}_{\ell'}$, $(P_\ell, H_{\ell-2\nu}) = (H_{\ell-2\nu}, H_{\ell-2\nu}) = 0$, hence $H_{\ell-2\nu} = 0$ for any ν. Hence $\tilde{H}_{\ell-2\nu} = \tilde{\tilde{H}}_{\ell-2\nu}$.

(ii) To prove existence we first need a lemma

$$\boxed{\nabla^2 r^{n+2} H_\ell = (n+2)(2\ell+n+3) r^n H_\ell}$$

Proof:

$$\nabla^2 r^{n+2} H_\ell(\vec{r}) = \left[\frac{1}{r}\partial_r^2 r + \frac{1}{r^2}\Lambda^2\right] r^{\ell+n+2} H_\ell(\hat{r})$$

$$= \frac{1}{r}\partial_r^2 r^{\ell+n+3} H_\ell(\hat{r}) + r^{\ell+n}\Lambda^2 H_\ell(\hat{r})$$

$$= (\ell+n+3)(\ell+n+2) r^{\ell+n} H_\ell(\hat{r}) - \ell(\ell+1) r^{\ell+n} H_\ell(\hat{r})$$

$$= \left[(\ell+n+3)(\ell+n+2) - \ell(\ell+1)\right] r^n H_\ell(\vec{r})$$

$$= (n+2)(2\ell+n+3) r^n H_\ell(\vec{r}).$$

Now we prove existence by induction

(a) true for $\ell = 0$ $\quad \mathcal{P}_0 = \mathcal{H}_0$, $\quad 1$ a basis

$\mathcal{P}_1 = \mathcal{H}_1$, $\quad (x, y, z)$ a basis

(b) suppose true for ℓ. We want to prove for $\ell+2$.

Consider $P_{\ell+2} \in \mathcal{P}_{\ell+2}$. Now $\nabla^2 P_{\ell+2} \in \mathcal{P}_\ell$ so by assumption $\exists$ harmonic polynomials $H_\ell, H_{\ell-2}, \cdots H_{\tau(\ell)}$ such that

$$\nabla^2 P_{\ell+2} = H_\ell + r^2 H_{\ell-2} + \cdots + r^{\ell-\tau(\ell)} H_{\tau(\ell)}$$

But by the lemma, this is

$$\nabla^2 P_{\ell+2} = \nabla^2 \left[\frac{r^2 H_\ell}{2(2\ell+3)} + \frac{r^4 H_{\ell-2}}{4(2\ell+1)} + \frac{r^6 H_{\ell-4}}{6(2\ell-1)} + \cdots + \frac{r^{\ell+2-\tau(\ell)} H_{\tau(\ell)}}{(\ell+2-\tau(\ell))(\ell+3-\tau(\ell))} \right]$$

Thus $P_{\ell+2} = F_{\ell+2} + r^2 H_\ell' + r^4 H_{\ell-2}' + r^{\ell+2-\tau(\ell)} H_{\tau(\ell)}'$

where $F_{\ell+2}$ is such that $\nabla^2 F_{\ell+2} = 0$ and $F_{\ell+2} \in \mathcal{P}_{\ell+2}$.

Hence $F_{\ell+2}$ is harmonic or $F_{\ell+2} = H_{\ell+2}' \in \mathcal{H}_{\ell+2}$.

Also $\tau(\ell+2) = \tau(\ell)$. Hence

$$P_{\ell+2} = H_{\ell+2}' + r^2 H_\ell' + \cdots + r^{\ell+2-\tau(\ell+2)} H_{\tau(\ell+2)}$$

and we are done,

The theorem has an immediate corollary.

Any polynomial $F(\vec{r}) = F(x,y,z)$ can be written as a finite sum of harmonic homogeneous polynomials, i.e., as a sum of solid spherical harmonics

$$F(\vec{r}) = \sum_{\ell=0}^{L} \sum_{n=0}^{N} r^n H_\ell^{(n)}(\vec{r}) \qquad \text{where } H_\ell^{(n)}(\vec{r}) \in \mathcal{H}_\ell$$

Another immediate corollary is that if $F(\hat{r})$ is any polynomial mapping $F : \Omega \to C$, then F may be written as a finite sum of surface spherical harmonics $Y_\ell \in \mathcal{H}_\ell(\Omega)$.

$$\boxed{F(\hat{r}) = \sum_{\ell=0}^{L} Y_\ell(\hat{r}) \ , \ Y_\ell \in \mathcal{H}_\ell(\Omega)}$$

It is now a simple matter to prove the following completeness theorem.

<u>Theorem</u>: If $f : \Omega \to C$ is a <u>continuous</u> function, then for any $\epsilon > 0$ there is an L and a sequence of surface spherical harmonics $Y_\ell \in \mathcal{H}_\ell(\Omega)$, $\ell = 0, 1, \ldots L$ such that for every $\hat{r}$ on Ω

$$\left| f(\hat{r}) - \sum_{\ell=0}^{L} Y_\ell(\hat{r}) \right| < \epsilon .$$

This theorem may be stated simply by saying that any <u>continuous</u> function $f : \Omega \to C$ may be approximated <u>uniformly</u> by a series of surface spherical harmonics.

Proof: We know that <u>any</u> <u>polynomial</u> function $P : \Omega \to C$ may be written as a finite sum of surface spherical harmonics. Therefore if we can find a polynomial function $P : \Omega \to C$ such that

$$\left| f(\hat{r}) - P(\hat{r}) \right| < \epsilon$$

then we have proved the theorem.
But the existence of such a polynomial function is exactly the assertion of the well-known

<u>Weierstrass</u> <u>Approximation Theorem</u> : Suppose V is a bounded, closed subset of R^3 and $f : V \to C$ is continuous. Then for any $\epsilon > 0$ $\exists$ a polynomial (not necessarily homogeneous) $P(\vec{r})$ such that

$$\left| P(\vec{r}) - f(\vec{r}) \right| < \epsilon \quad \text{for every } \vec{r} \in V.$$

Proof: Take a sphere of radius b centered at zero and containing V in its interior. Extend f to be continuous in this sphere $B(\vec{0}, b)$ (this is why V must be closed). Define $f=0$ for $|r| \geq b$. Then f is continuous everywhere and vanishes outside $B(\vec{0}, b)$.

Now define the function

$$K_n(\vec{r}, \vec{\rho}) = C_n(b) \left[1 - \frac{|\vec{r} - \vec{\rho}|^2}{(2b)^2} \right]^n$$

where $C_n(b)$ is chosen so that

$$\int_{B(\vec{0}, 2b)} K_n(\vec{0}, \vec{\rho}) \, d^3\vec{\rho} = 1$$

Thus $C_n(b) \, 4\pi \int_0^{2b} \left(1 - \frac{r^2}{4b^2}\right)^n r^2 \, dr = 1$ or

$$4\pi C_n(b) (2b)^3 \int_0^1 (1-u^2)^n u^2 \, du = 1$$

$$C_n(b) = \frac{1}{4\pi (2b)^3 \dfrac{\Gamma(n+1)\Gamma(3/2)}{2\Gamma(n+5/2)}}$$

Now let

$$P_n(\vec{r}) = \int_{R^3} K_n(\vec{r}, \vec{\rho}) f(\vec{\rho}) \, d^3\vec{\rho}$$

It is clear that $P_n(\vec{r})$ is a polynomial of degree $2n$

We now assert that $\lim_{n \to \infty} P_n(\vec{r}) = f(\vec{r})$ uniformly in $B(\vec{0}, b)$

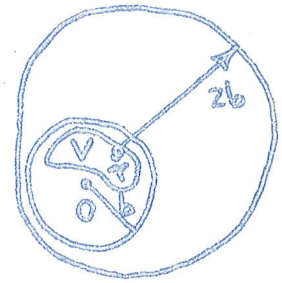
To see this

$$P_n(\vec{r}) = \int_{R^3} C_n(b)\left[1 - \frac{|\vec{r}-\vec{\rho}|^2}{(2b)^2}\right]^n f(\vec{\rho})\, d^3\vec{\rho}$$

substitute $\vec{x} = \vec{\rho} - \vec{r}$

$$P_n(\vec{r}) = \int_{R^3} C_n(b)\left[1 - \frac{|\vec{x}|^2}{(2b)^2}\right]^n f(\vec{r}+\vec{x})\, d^3\vec{x}$$

$$= \int_{B(\vec{0},2b)} C_n(b)\left[1 - \frac{|\vec{x}|^2}{(2b)^2}\right] f(\vec{r}+\vec{x})\, d^3\vec{x}$$



$$= \int_{B(\vec{0},2b)} C_n(b)\left[1 - \frac{|\vec{x}|^2}{(2b)^2}\right]^n \left[f(\vec{r}) + f(\vec{r}+\vec{x}) - f(\vec{r})\right] d^3\vec{x}$$

$$= f(\vec{r}) \int_{B(\vec{0},2b)} C_n(b)\left[1 - \frac{|\vec{x}|^2}{(2b)^2}\right]^n d^3\vec{x} + \int_{B(\vec{0},2b)} C_n(b)\left[1 - \frac{|\vec{x}|^2}{(2b)^2}\right]^n \left[f(\vec{r}+\vec{x}) - f(\vec{r})\right] d^3\vec{x}$$

$$= f(\vec{r}) + \int_{B(\vec{0},2b)} C_n(b)\left[1 - \frac{|\vec{x}|^2}{(2b)^2}\right]^n \left[f(\vec{r}+\vec{x}) - f(\vec{r})\right] d^3\vec{x}$$

Now since f is continuous in V, given $\epsilon > 0$, $\exists\ \delta(\epsilon) > 0$ such that if $|\vec{x}| < \delta$, then $|f(\vec{r}+\vec{x}) - f(\vec{r})| < \epsilon$

Let $B_\epsilon(\vec{0},2b)$ denote $B(\vec{0},2b) - B(\vec{0},\epsilon)$ and let $M = \max_V [f(\vec{r}+\vec{x}) - f(\vec{r})]$

Then $$|P_n(\vec{r}) - f(\vec{r})| < \epsilon \int_{B(0,2\epsilon)} C_n(b)\left[1 - \frac{|\vec{x}|^2}{(2b)^2}\right]^n d^3\vec{x} + M \int_{B_\epsilon(\vec{0},2b)} C_n(b)\left[1 - \frac{|\vec{x}|^2}{(2b)^2}\right]^n d^3\vec{x}$$

$$< \epsilon \int_{B(0,2b)} \text{same} + M \int_{B_\epsilon(0,2b)} \text{same}$$

$$|P_n(\vec{r}) - f(\vec{r})| < \epsilon + M \int_{B_\epsilon(0,2b)} C_n(b)\left[1 - \frac{|\vec{x}|^2}{(2b)^2}\right]^n d^3\vec{x}$$

It remains to show that for any $\epsilon > 0$, the integral

$$\int_{B_\epsilon(\vec{0},2b)} C_n(b)\left[1-\frac{|\vec{x}|^2}{(2b)^2}\right]^n d^3\vec{x} \to 0 \quad \text{as} \quad n \to \infty$$

Problem 7: Show it. Hint, if you need one. The above proof is an extension to $\mathbb{R}^3$ of the proof in Courant and Hilbert, vol. 1, pp. 65-68.

5. The space $\mathcal{H}_\ell$

We must now investigate the structure of the space $\mathcal{H}_\ell$.

<u>Dimension</u>:

$$\mathcal{P}_\ell = \mathcal{H}_\ell \oplus r^2 \mathcal{H}_{\ell-2} \oplus r^4 \mathcal{H}_{\ell-4} \oplus \cdots\cdots$$

$$= \mathcal{H}_\ell \oplus r^2 \left[\mathcal{H}_{\ell-2} \oplus r^2 \mathcal{H}_{\ell-4} \oplus \cdots\right]$$

$$= \mathcal{H}_\ell \oplus r^2 \mathcal{P}_{\ell-2}$$

Thus $\dim \mathcal{P}_\ell = \dim \mathcal{H}_\ell + \dim \mathcal{P}_{\ell-2}$

$$\frac{(\ell+1)(\ell+2)}{2} = \dim \mathcal{H}_\ell + \frac{\ell(\ell-1)}{2}$$

$$\boxed{\dim \mathcal{H}_\ell = 2\ell+1}$$

Examples:

ℓ	$2\ell+1$	basis for $\mathcal{H}_\ell$
0	1	1
1	3	x, y, z
2	5	$xy, xz, yz, x^2-y^2, 2z^2-x^2-y^2$

Problem 8: Use the Gram-Schmidt orthonormalization process to generate orthonormal bases for $\mathcal{H}_0, \mathcal{H}_1, \mathcal{H}_2$. Remember the inner product is $(f,g) = \int_\Omega f^* g \, dA$.

Construction of an orthonormal basis in $\mathcal{H}_\ell$ for each Cartesian axis system $\hat{x}, \hat{y}, \hat{z}$ in R^3

First a *definition*: Let V be a real (complex) vector space and let T be a linear operator on V. Let λ be a real (complex) number. Then $v \in V$ is said to be an *eigenvector* of T corresponding to the eigenvalue λ if

(i) $v \neq 0$, and

(ii) $T(v) = \lambda v$

Note that if v is an eigenvector of T belonging to eigenvalue λ and a is any real (complex) number $\neq 0$, then av is also an eigenvector belonging to eigenvalue λ. Thus if λ is an eigenvalue of T, there is a vector $\hat{n} \in V$ of unit length $\|\hat{n}\| = 1$ such that $T(\hat{n}) = \lambda \hat{n}$.

Example:

$(x+iy)^\ell$ and $(x-iy)^\ell$ are both in $\mathcal{H}_\ell$

Now $(x+iy)^\ell = (r\sin\theta e^{i\phi})^\ell = r^\ell \sin^\ell\theta e^{i\ell\phi}$

$(x-iy)^\ell = (r\sin\theta e^{-i\phi})^\ell = r^\ell \sin^\ell\theta e^{-i\ell\phi}$

Now $L_z = \frac{1}{i}\Lambda_z = \frac{1}{i}\partial_\phi$, so

$$\boxed{\begin{aligned} L_z(x+iy)^\ell &= \ell(x+iy)^\ell \\ L_z(x-iy)^\ell &= -\ell(x-iy)^\ell \end{aligned}}$$

Thus $(x+iy)^\ell$ and $(x-iy)^\ell \in \mathcal{H}_\ell$ are both eigenvectors (or eigenfunctions of the operator L_z. The associated eigenvalues are, respectively ℓ and $-\ell$

We are going to make $H_\ell^{-\ell} = (x-iy)^\ell = r^\ell \sin^\ell\theta\, e^{-i\ell\phi}$ the first element of our orthonormal basis for $\mathcal{H}_\ell$.
Thus we must normalize it.

$$(H_\ell^{-\ell}, H_\ell^{-\ell}) = \int_\Omega |H_\ell^{-\ell}|^2\, dA = \int_0^{2\pi} d\phi \int_0^\pi \sin\theta\, d\theta\, \sin^{2\ell}\theta$$

$$= 2\pi \int_0^\pi \sin^{2\ell+1}\theta\, d\theta = 4\pi \frac{2^{2\ell}(\ell!)^2}{(2\ell+1)!}$$

Now define

$$\boxed{Y_\ell^{-\ell} = \sqrt{\frac{2\ell+1}{4\pi}} \left[\frac{(2\ell)!}{0!}\right]^{1/2} \frac{1}{2^\ell \ell!} \sin^\ell\theta\, e^{-i\ell\phi}}$$

Then $(Y_\ell^{-\ell}, Y_\ell^{-\ell}) = \| Y_\ell^{-\ell}\|^2 = 1$

We are now going to construct an orthonormal basis for $\mathcal{H}_\ell$ by operating on $Y_\ell^{-\ell}$ with the L_+ operator.

Suppose that H_ℓ^m is an eigenfunction of $L_z \in \mathcal{H}_\ell$ with eigenvalue m, i.e. suppose that

$L_z H_\ell^m = m H_\ell^m$ and $H_\ell^m \in \mathcal{H}_\ell$

Then $L_+ H_\ell^m = H_\ell^{m+1}$ is also in $\mathcal{H}_\ell$ and $L_z H_\ell^{m+1} = (m+1) H_\ell^{m+1}$

Proof: $L_z H_\ell^{m+1} = L_z L_+ H_\ell^m = L_+ (L_z + 1) H_\ell^m = L_+ (m+1) H_\ell^m$

$= (m+1) L_+ H_\ell^m = (m+1) H_\ell^{m+1}$!

Note that we have not shown that H_ℓ^{m+1} is an eigenfunction of L_z with eigenvalue $m+1$ since we have not investigated whether or not $H_\ell^{m+1} \neq 0$.

Is H_ℓ^{m+1} an eigenvector of L_z, i.e., is $H_\ell^{m+1} \neq 0$?

$$(H_\ell^{m+1}, H_\ell^{m+1}) = (L_+ H_\ell^m, L_+ H_\ell^m) = (H_\ell^m, L_- L_+ H_\ell^m)$$

$$= (H_\ell^m, (L^2 - L_z^2 - L_z) H_\ell^m) = [\ell(\ell+1) - m^2 - m] (H_\ell^m, H_\ell^m)$$

$$= (\ell - m)(\ell + m + 1)(H_\ell^m, H_\ell^m)$$

Thus $H_\ell^{m+1} \neq 0$ if $H_\ell^m \neq 0$ and $m \neq -\ell-1$ and $m \neq \ell$

Also $H_\ell^m = 0$ if $m > \ell$

<u>Theorem</u>: Define $Y_\ell^{-\ell}$ as above and define $Y_\ell^m = 0$ if $|m| > \ell$

Now define

$$\boxed{Y_\ell^m = \left[\frac{(\ell-m)!}{(\ell+m)!}\right]^{\frac{1}{2}} \frac{1}{\sqrt{(2\ell)!}} (L_+)^{\ell+m} Y_\ell^{-\ell} \qquad m = -\ell, -\ell+1, \cdots, \ell-1, \ell}$$

Then $Y_\ell^{-\ell}, Y_\ell^{-\ell+1}, \cdots Y_\ell^{\ell-1}, Y_\ell^\ell$ are an orthonormal basis for H_ℓ, and we have

(1). $L_+ Y_\ell^m = \sqrt{(\ell-m)(\ell+m+1)}\, Y_\ell^{m+1}$

(2). $L_- Y_\ell^m = \sqrt{(\ell+m)(\ell-m+1)}\, Y_\ell^{m-1}$

(3). $L_z Y_\ell^m = m Y_\ell^m$

These are true for all m

Proof:

(1) for $m \leq -\ell-2$ trivial

for $m = -\ell-1$ verify

for $-\ell \leq m \leq \ell-1$

$$L_+ Y_\ell^m = \left[\frac{(\ell-m)!}{(\ell+m)!}\right]^{1/2} \frac{1}{\sqrt{(2\ell)!}} (L^+)^{\ell+m+1} Y_\ell^{-\ell} = \sqrt{(\ell-m)(\ell+m+1)}\, Y_\ell^{m+1}$$

for $m=l$

$$(L_+ Y_l^l, L_+ Y_l^l) = (Y_l^l, L_- L_+ Y_l^l) =$$

$$(Y_l^l, (L^2 - L_z^2 - L_z) Y_l^l) = [l(l+1) - l^2 - l](Y_l^l, Y_l^l) = 0$$

for $m > l$ trivial

(2) apply L_- to (1), get

$$(L^2 - L_z^2 - L_z) Y_l^m = \sqrt{(l-m)(l+m+1)}\, L_- Y_l^{m+1}$$

$$(l-m)(l+m+1) Y_l^m = \sqrt{(l-m)(l+m+1)}\, L_- Y_l^{m+1}$$

$$L_- Y_l^{m+1} = \sqrt{(l-m)(l+m+1)}\, Y_l^m$$

change m to $m-1$

$$L_- Y_l^m = \sqrt{(l+m)(l-m+1)}\, Y_l^{m-1}$$

(3) prove by induction on m. True for $m=-l$. Now assume true for m, then true for $m+1$ because

$$L_z Y_l^{m+1} = L_z \frac{L_+ Y_l^m}{\sqrt{(l-m)(l+m+1)}} = \frac{L_+ (L_z + 1) Y_l^m}{\sqrt{(l-m)(l+m+1)}}$$

$$= \frac{(m+1) L_+ Y_l^m}{\sqrt{(l-m)(l+m+1)}} = (m+1) Y_l^{m+1}$$

Thus true for $m = -l, \dots 0, \dots, l$

Trivial for $|m| > l$.

Now we must still prove that Y_l^m are orthonormal and linearly independent. That $\|Y_l^m\| = 1$ is also proved by induction on m (for $|m| \le l$)

It is true for $m=-l$. Suppose $\|Y_l^m\| = 1$ and $-l \le m \le l-1$

Then

$$(Y_\ell^{m+1}, Y_\ell^{m+1}) = \frac{1}{(\ell-m)(\ell+m+1)} (L_+ Y_\ell^m, L_+ Y_\ell^m)$$

$$= \frac{(Y_\ell^m, L_- L_+ Y_\ell^m)}{(\ell-m)(\ell+m+1)} = \frac{(Y_\ell^m, (L^2 - L_z^2 - L_z) Y_\ell^m)}{(\ell-m)(\ell+m+1)} = (Y_\ell^m, Y_\ell^m) = 1$$

The induction fails if $m = \ell$.

Now Y_ℓ^m is an eigenvector of the operator L_z with eigenvalue m when $|m| \le \ell$

Problem 9: Prove that the eigenvectors of an Hermitian linear operator on a finite dimensional vector space V are mutually orthogonal

Since $L_z : \mathcal{H}_\ell \to \mathcal{H}_\ell$ is Hermitian, $Y_\ell^{-\ell}, \dots Y_\ell^{\ell}$ are mutually orthogonal. Therefore they are linearly independent. Since there are $2\ell+1$ of them, they are therefore a basis for $\mathcal{H}_\ell$. Therefore they are an orthonormal basis for $\mathcal{H}_\ell$. Q.E.D.

Corollary:

$$\boxed{Y_\ell^m = \left[\frac{(\ell+m)!}{(\ell-m)!}\right]^{1/2} \frac{1}{\sqrt{(2\ell)!}} (L_-)^{\ell-m} Y_\ell^\ell \qquad |m| \le \ell}$$

proof: true if $m = \ell$. suppose true for m. Then

$$Y_\ell^{m-1} = \frac{1}{\sqrt{(\ell+m)(\ell-m+1)}} L_- Y_\ell^m = \left[\frac{(\ell+m-1)!}{(\ell-m+1)!}\right]^{1/2} \frac{1}{\sqrt{(2\ell)!}} (L_-)^{\ell-m+1} Y_\ell^\ell$$

now replace $m-1$ by m

We have thus shown that an orthogonal basis in $\mathcal{H}_\ell$ may be constructed from $2\ell+1$ eigenvectors of the operator $L_z : \mathcal{H}_\ell \to \mathcal{H}_\ell$. If $\psi_\ell \in \mathcal{H}_\ell$ and ψ_ℓ is an eigenvector of L_z with eigenvalue m, then $\psi_\ell = \alpha Y_\ell^m$ for some complex constant α. The operators L_+ and L_- are called <u>ladder operators</u>, because they convert an eigenvector of L_z with associated eigenvalue m to another eigenvector with associated eigenvalue $m \pm 1$.

<u>Explicit expressions for Y_ℓ^m:</u>

$$Y_\ell^m = \sqrt{\frac{2\ell+1}{4\pi}}\,\frac{1}{2^\ell \ell!}\left[\frac{(\ell-m)!}{(\ell+m)!}\right]^{\frac{1}{2}} (L_+)^{\ell+m} \sin^\ell\theta\, e^{-i\ell\phi}$$

$$L^+ = e^{i\phi}\left[\partial_\theta + i\cot\theta\,\partial_\phi\right]$$

Now for any function $f(\theta)$ alone

$$L_+ e^{iq\phi} f(\theta) = e^{i(q+1)\phi}\left(\partial_\theta - q\cot\theta\right) f(\theta) = e^{i(q+1)\phi} \sin^q\theta\, \partial_\theta \sin^{-q}\theta\, f(\theta)$$

$$(L_+)^2 e^{iq\phi} f(\theta) = e^{i(q+2)\phi} \sin^{q+1}\theta\, \partial_\theta\left(\frac{1}{\sin\theta}\right)\partial_\theta \sin^{-q}\theta\, f(\theta)$$

$$= e^{i(q+2)\phi} \sin^{q+2}\theta \left(\frac{1}{\sin\theta}\partial_\theta\right)^2 \sin^{-q}\theta\, f(\theta)$$

in general

$$(L_+)^\nu e^{iq\theta} f(\theta) = e^{i(q+\nu)\phi} \sin^{(q+\nu)}\theta \left(\frac{1}{\sin\theta}\partial_\theta\right)^\nu \sin^{-q}\theta\, f(\theta)$$

Now set $\nu = \ell+m$, $q = -\ell$, and $f(\theta) = \sin^\ell\theta$

$$\boxed{Y_\ell^m = \sqrt{\frac{2\ell+1}{4\pi}}\,\frac{1}{2^\ell \ell!}\left[\frac{(\ell-m)!}{(\ell+m)!}\right]^{1/2} e^{im\phi} \sin^m\theta \left(\frac{1}{\sin\theta}\partial_\theta\right)^{\ell+m} \sin^{2\ell}\theta\,.}$$

Now let $\mu = \cos\theta$

The <u>associated Legendre function of degree l and order m</u> is defined by

$$\boxed{P_l^m(\mu) = \frac{(1-\mu^2)^{m/2}}{2^l\, l!}\left(\frac{d}{d\mu}\right)^{l+m} (\mu^2-1)^l}$$

Thus we can write an explicit expression for Y_l^m in terms of P_l^m

$$\boxed{Y_l^m(\theta,\phi) = (-1)^m \sqrt{\frac{2l+1}{4\pi}} \left[\frac{(l-m)!}{(l+m)!}\right]^{1/2} e^{im\phi}\, P_l^m(\cos\theta)}$$

In particular $$\boxed{P_l^0(\mu) = P_l(\mu) = \frac{1}{2^l l!}\left(\frac{d}{d\mu}\right)^l (\mu^2-1)^l}$$

is <u>Rodrigues' ~~[illegible]~~ formula</u> for the <u>Legendre polynomial</u> $P_l(\mu)$

Also $P_l^l(\mu) = \frac{(1-\mu^2)^{l/2}}{2^l l!}(2l)!$ so

~~[illegible]~~

$$\boxed{Y_l^l = (-1)^l \sqrt{\frac{2l+1}{4\pi}}\, \frac{1}{2^l l!}\left[\frac{(2l)!}{(0!)}\right]^{1/2} e^{il\phi} \sin^l\theta}$$

Now $Y_l^m = (-1)^l \sqrt{\frac{2l+1}{4\pi}}\, \frac{1}{2^l l!}\left[\frac{(l+m)!}{(l-m)!}\right]^{1/2} (L_-)^{l-m}\, e^{il\phi} \sin^l\theta$

$L_- = e^{-i\phi}[-\partial_\theta + i\cot\theta\, \partial_\phi]$

$$L_- \left[e^{iq\phi} f(\theta)\right] = -e^{i(q-1)\phi} \left(\partial_\theta + q\cot\theta\right) f(\theta) = -e^{i(q+1)\phi} \frac{1}{\sin^q\theta} \partial_\theta \sin^q\theta\, f(\theta)$$

$$(L_-)^2 \left[e^{iq\phi} f(\theta)\right] = (-1)^2 e^{i(q-2)\phi} \frac{1}{\sin^{q-1}\theta} \partial_\theta \frac{1}{\sin\theta} \partial_\theta \sin^q\theta\, f(\theta)$$

$$= e^{i(q-2)\phi} \frac{1}{\sin^{q-2}\theta} \left(-\frac{1}{\sin\theta}\partial_\theta\right)^2 \sin^q\theta\, f(\theta).$$

$$(L_-)^\nu \left[e^{iq\phi} f(\theta)\right] = e^{i(q-\nu)\phi} \frac{1}{\sin^{q-\nu}\theta} \left(-\frac{1}{\sin\theta}\partial_\theta\right)^\nu \sin^q\theta\, f(\theta)$$

Now set $\nu = l-m$, $q=l$, $f(\theta) = \sin^l\theta$

$$\boxed{Y_l^m = (-1)^l \sqrt{\frac{2l+1}{4\pi}} \frac{1}{2^l l!} \left[\frac{(l+m)!}{(l-m)!}\right]^{1/2} e^{im\phi} \frac{1}{\sin^m\theta} \left(-\frac{1}{\sin\theta}\partial_\theta\right)^{l-m} \sin^{2l}\theta \qquad |m| \le l}$$

Using this we can get another expression for $P_l^m(\mu)$

$$P_l^m(\mu) = \frac{(-1)^m}{2^l l!} \left[\frac{(l+m)!}{(l-m)!}\right] \frac{1}{(1-\mu^2)^{m/2}} \left(\frac{d}{d\mu}\right)^{l-m} (\mu^2-1)^l$$

Also we can see that

$$\boxed{Y_l^{-m} = (-1)^m Y_l^{m*}}$$

comparing the above with (*)

Now lets consider some particular values of Y_l^m

1. $Y_\ell^m(-\hat{r}) = (-1)^\ell Y_\ell^m(\hat{r})$ since $Y_\ell^m \in P_\ell$

$P_\ell^m(-\mu) = (-1)^{\ell+m} P_\ell^m(\mu)$

2. $Y_\ell^m(\theta,\phi) = \text{const.}\; e^{im\phi} \sin^{|m|}\theta \,[1 + O(\sin\theta)]$

$\therefore \quad Y_\ell^m(\hat{z}) = Y_\ell^m(-\hat{z}) = 0$ if $m \neq 0$

$\therefore \quad P_\ell^m(\pm 1) = 0$ if $m \neq 0$

3. $P_\ell^0(\mu) = \frac{1}{2^\ell \ell!}\left(\frac{d}{d\mu}\right)^\ell (\mu^2-1)^\ell$

let $x = \mu - 1$

$$P_\ell^0(1+x) = \frac{1}{2^\ell \ell!}\left(\frac{d}{dx}\right)^\ell x^\ell (x+2)^\ell$$

$$= \frac{1}{\ell!}\left(\frac{d}{dx}\right)^\ell x^\ell \left(1+\frac{x}{2}\right)^\ell$$

Thus $P_\ell^0(1) = 1$

Likewise $P_\ell^0(-1) = (-1)^\ell$

Thus

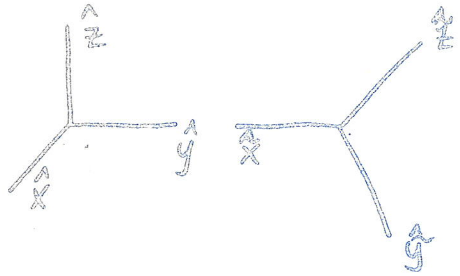
$$\boxed{\begin{aligned} &Y_\ell^0(\hat{z}) = Y_\ell^0(-\hat{z}) = \sqrt{\frac{2\ell+1}{4\pi}} \\ &Y_\ell^m(\hat{z}) = Y_\ell^m(-\hat{z}) = 0 \quad m \neq 0 \end{aligned}}$$

4. Problem 10. Show that $\int_{-1}^{1} P_\ell^0(\mu) P_{\ell'}^0(\mu)\, d\mu = \frac{\delta_{\ell\ell'}}{\ell+\frac{1}{2}}$

(orthogonality relation for Legendre polynomials)

Show that $\int_{-1}^{1} P_\ell^m(\mu) P_{\ell'}^{m'}(\mu)\, d\mu = \frac{\delta_{mm'}\delta_{\ell\ell'}}{\ell+\frac{1}{2}} \frac{(m+\ell)!}{(\ell-m)!}$

Change of basis in H_ℓ induced by a change of basis in $\mathbb{R}^3$



Y_ℓ^m defined in terms of $\langle \hat{x}, \hat{y}, \hat{z} \rangle$

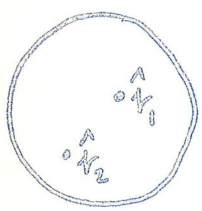
$\tilde{Y}_\ell^{\tilde{m}}$ defined in terms of $\langle \hat{\tilde{x}}, \hat{\tilde{y}}, \hat{\tilde{z}} \rangle$

$\{ Y_\ell^{-\ell}, \dots Y_\ell^{\ell} \}$ and $\{ \tilde{Y}_\ell^{-\tilde{\ell}}, \dots \tilde{Y}_\ell^{\tilde{\ell}} \}$ are both orthonormal bases for H_ℓ. Therefore there is a $2\ell+1 \times 2\ell+1$ unitary matrix $Q^{m\tilde{m}}$ such that

$$Y_\ell^m = \sum_{\tilde{m}=-\ell}^{\ell} Q_{m\tilde{m}} \tilde{Y}_\ell^{\tilde{m}} \qquad \text{where}$$

$$\sum_{m=-\ell}^{\ell} Q_{m\tilde{m}} (Q_{m\tilde{m}'})^* = \delta_{\tilde{m}\tilde{m}'}$$

Consider this basis transformation at two arbitrary points $\hat{r}_1, \hat{r}_2 \in \Omega$



$$Y_\ell^m(\hat{r}_1) = \sum_{\tilde{m}=-\ell}^{\ell} Q_{m\tilde{m}} \tilde{Y}_\ell^{\tilde{m}}(\hat{r}_1)$$

$$Y_\ell^m(\hat{r}_2) = \sum_{\tilde{m}=-\ell}^{\ell} Q_{m\tilde{m}} \tilde{Y}_\ell^{\tilde{m}}(\hat{r}_2)$$

In matrix notation (if we denote) $Y(\hat{r}_1) = \begin{bmatrix} Y_\ell^{-\ell}(\hat{r}_1) \\ \vdots \\ Y_\ell^{\ell}(\hat{r}_1) \end{bmatrix}$

$Y(\hat{r}_1) = Q\tilde{Y}(\hat{r}_1)$

$Y(\hat{r}_2) = Q\tilde{Y}(\hat{r}_2)$

Now $Y^+(\hat{r}_1)Y(\hat{r}_2) = (Q\tilde{Y}(\hat{r}_1))^+ Q\tilde{Y}(\hat{r}_2) = \tilde{Y}^+(\hat{r}_1) Q^+ Q \tilde{Y}(\hat{r}_2)$

$= \tilde{Y}^+(\hat{r}_1)\tilde{Y}(\hat{r}_2)$

$$\boxed{\sum_{m=-\ell}^{\ell} Y_\ell^m(\hat{r}_1)^* Y_\ell^m(\hat{r}_2) = \sum_{\tilde{m}=-\ell}^{\ell} \tilde{Y}_\ell^{\tilde{m}}(\hat{r}_1)^* \tilde{Y}_\ell^{\tilde{m}}(\hat{r}_2) \quad \text{true for any } \hat{r}_1, \hat{r}_2 \;\; \langle \hat{x}, \hat{y}, \hat{z} \rangle \; \langle \hat{\tilde{x}}, \hat{\tilde{y}}, \hat{\tilde{z}} \rangle}$$

Now take $\tilde{\hat{x}}, \tilde{\hat{y}}, \tilde{\hat{z}}$ so that $\tilde{\hat{z}} = \hat{r}_1$. Then

$$\tilde{Y}_\ell^{\tilde{m}}(\hat{r}_1) = \tilde{Y}_\ell^{\tilde{m}}(\tilde{\hat{z}}) = 0 \quad \text{if } \tilde{m} \neq 0 \quad \text{and}$$

$$\tilde{Y}_\ell^0(\hat{r}_1) = \sqrt{\frac{2\ell+1}{4\pi}} \quad \text{thus}$$

$$\sqrt{\frac{2\ell+1}{4\pi}}\, \tilde{Y}_\ell^0(\hat{r}_2) = \sum_{m=-\ell}^{\ell} Y_\ell^m(\hat{r}_1)^* Y_\ell^m(\hat{r}_2)$$

and $\tilde{Y}_\ell^0(\hat{r}_2) = \sqrt{\frac{2\ell+1}{4\pi}}\, P_\ell^0(\cos\Theta)$ where $\cos\Theta = \hat{r}_2 \cdot \tilde{\hat{z}} = \hat{r}_2 \cdot \hat{r}_1$

Thus

$$\boxed{\frac{2\ell+1}{4\pi} P_\ell^0(\hat{r}_1 \cdot \hat{r}_2) = \sum_{m=-\ell}^{\ell} Y_\ell^m(\hat{r}_1)^* Y_\ell^m(\hat{r}_2) \quad \text{for any } \langle \hat{x}, \hat{y}, \hat{z} \rangle}$$

The addition theorem for spherical harmonics

In particular, taking $\hat{r}_1 = \hat{r}_2$

$$\frac{2\ell+1}{4\pi} = \sum_{m=-\ell}^{\ell} Y_\ell^m(\hat{r}_1)^* Y_\ell^m(\hat{r}_1) = \sum_{m=-\ell}^{\ell} |Y_\ell^m(\hat{r}_1)|^2 \quad \text{for any } \hat{r}_1$$

Thus

$$\boxed{|Y_\ell^m(\hat{r})| \le \sqrt{\frac{2\ell+1}{4\pi}} \quad \text{for any } \ell, m,\ \hat{r} \in \Omega}$$

Schur's lemma

Every change of Cartesian axis system $\langle \hat{x}, \hat{y}, \hat{z} \rangle \to \langle \tilde{\hat{x}}, \tilde{\hat{y}}, \tilde{\hat{z}} \rangle$ gives rise to a change of orthonormal basis for H_ℓ of the form $Y_\ell^m(\hat{r}) \to \tilde{Y}_\ell^m(\hat{r})$ via a unitary matrix $Q_{mm'}$. Thus, associated with every change of Cartesian axis system (i.e., associated with every rigid rotation R of the original Cartesian axis system $\langle \hat{x}, \hat{y}, \hat{z} \rangle$) there is associated a $(2\ell+1)$-dimensional unitary matrix $Q_{(\ell)}$ (for every $\ell = 0, 1, 2, \ldots$). The set of all such $(2\ell+1)$-dimensional unitary matrices $Q_{(\ell)}$ is said to be a $(2\ell+1)$-dimensional <u>representation</u> of the rotation group $O^+(3)$ (i.e. the set of all rigid rotations R). When we come to discussing the response of the Earth to the tidal potential generated by the moon and sun, we will have need for the following theorem concerning this $(2\ell+1)$-dimensional representation of the rotation group.

Theorem (Special case of Schur's lemma)

Suppose that D is a $(2\ell+1)$-dimensional matrix which commutes with every member of the $(2\ell+1)$-dimensional representation of the rotation group $O^+(3)$. That is, suppose that for <u>every</u> $(2\ell+1)$-dimensional unitary matrix $Q_{(\ell)}$ of the form

$$\tilde{Y}_\ell(\hat{r}) = Q_{(\ell)} Y_\ell(\hat{r})$$

we have

$$D Q_{(\ell)} = Q_{(\ell)} D$$

Then D is a constant multiple of the $(2\ell+1)$-dimensional identity matrix I, i.e.

$$D = d\, I$$

where d is a scalar.

Proof: Let $\langle \hat{x}, \hat{y}, \hat{z} \rangle$ be a Cartesian axis system and let $Y_\ell^m(\hat{r}) = N_\ell^m P_\ell^m(\cos\theta) e^{im\phi}$ be the associated spherical harmonics $\left(N_\ell^m = (-1)^m \sqrt{\frac{2\ell+1}{4\pi}} \left[\frac{(\ell-m)!}{(\ell+m)!} \right] \right)$

Consider first a change of Cartesian axis system $\langle \hat{x}, \hat{y}, \hat{z} \rangle \to \langle \hat{\tilde{x}}, \hat{\tilde{y}}, \hat{z} \rangle$ such that the $\hat{z}$-axis remains fixed. Any such change involves merely a rotation through some angle α about the $\hat{z}$-axis, the angle θ remains the same, and the new surface spherical harmonics are merely

$$\tilde{Y}_\ell^m(\hat{r}) = N_\ell^m P_\ell^m(\cos\theta) e^{im(\phi+\alpha)}$$

Thus in this case the matrix $Q_{(\ell)}$ takes the form

$$Q_{(\ell)}^{mm'}(0,0,\alpha) = e^{im\alpha} \delta_{mm'} \quad , \text{or}$$

$$Q_\ell(0,0,\alpha) = \begin{pmatrix} e^{-i\ell\alpha} & & 0 \\ & \ddots & \\ 0 & & e^{i\ell\alpha} \end{pmatrix}$$

Clearly any matrix D which commutes with all such matrices $Q_{(\ell)}(0,0,\alpha)$ is necessarily a diagonal matrix

Now consider a change of Cartesian axis system of the form $\langle \hat{x}, \hat{y}, \hat{z} \rangle \to \langle \tilde{\hat{x}}, \hat{y}, \tilde{\hat{z}} \rangle$, a rotation about some angle β about the $\hat{y}$ axis. The angle ϕ remains unchanged but θ is increased to $\theta+\beta$. The new surface spherical harmonics are

$$\tilde{Y}_\ell^m(\hat{r}) = N_\ell^m P_\ell^m(\cos(\theta+\beta))\, e^{im\phi}$$

and we have

$$\tilde{Y}_\ell(\hat{r}) = Q_{(\ell)}(0,\beta,0) Y_\ell(\hat{r})$$

Now consider the point $\hat{r} = (0,0)$ or $\theta=0$, $\phi=0$. Now $\tilde{Y}_\ell^m(0,0) = N_\ell^m P_\ell^m(\cos\beta)$ is in general non-zero, but $Y_\ell^m(0,0) = 0$ unless $m=0$. Thus $Q_{0m}^{(\ell)}$ is non-vanishing for all m, i.e. no zeroes occur in the $m=0$ row of the transformation matrix $Q_{(\ell)}(0,\beta,0)$. Now this fact can be used to show that all elements of the diagonal matrix D are in fact equal. To see this, assume that the diagonal matrix D with diagonal elements d_m commutes with $Q_{(\ell)}(0,\beta,0)$. Then the elements in the $m=0$ row of the products $DQ_{(\ell)}$ and $Q_{(\ell)}D$ are

$$\left[d_0 Q_\ell(0,\beta,0)\right]_{0m} = \left[Q_\ell(0,\beta,0) d_m\right]_{0m},$$

and since $\left[Q_\ell(0,\beta,0)\right]_{0m} \neq 0$ for all m, this implies that $d_0 = d_m$ for all m. Thus $D = d_0 I$. Q.E.D.

Expansion in spherical harmonics

Let $f: \Omega \to \mathbb{C}$ be in $\mathcal{L}_2(\Omega)$ and let $(\hat{x}, \hat{y}, \hat{z})$ be a right handed Cartesian axis system which generates a set $\{Y_\ell^m,\ \ell = 0, 1, 2, \ldots,\ -\ell \le m \le \ell\}$ of canonical surface spherical harmonics. Consider the problem of making a least squares approximation to f by a finite linear combination of surface spherical harmonics Y_ℓ^m.

$$\boxed{\text{Define} \quad f_\ell^m = (Y_\ell^m, f) = \int_\Omega Y_\ell^{m*}(\hat{r})\, f(\hat{r})\, dA}$$

We wish to determine coefficients g_ℓ^m such that

$$\left\| f - \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} g_\ell^m Y_\ell^m \right\| = \int_\Omega \left|\left(f - \sum_\ell^L \sum_m g_\ell^m Y_\ell^m \right)\right|^2 dA = \text{minimum}$$

But

$$\int_\Omega \left| f - \sum_\ell^L \sum_m g_\ell^m Y_\ell^m \right|^2 dA = \int_\Omega \left(f - \sum_\ell^L \sum_m g_\ell^m Y_\ell^m \right)^* \left(f - \sum_\ell^L \sum_m g_\ell^m Y_\ell^m \right) dA$$

$$= \int_\Omega |f|^2 dA + \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} \left|(g_\ell^m - f_\ell^m)\right|^2 - \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} \left| f_\ell^m \right|^2$$

Hence the best approximation is clearly given by $g_\ell^m = f_\ell^m$

Call this best approximation $f_L(\hat{r})$

$$\boxed{f_L(\hat{r}) = \sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} f_\ell^m Y_\ell^m(\hat{r}) \quad \text{where} \quad f_\ell^m = (Y_\ell^m, f)}$$

<u>Theorem</u>: $\lim_{L\to\infty} \| f - f_L \| = 0$, i.e.

$$\lim_{L\to\infty} \int_\Omega | f(\hat{r}) - f_L(\hat{r}) |^2 dA = 0$$

The proof of this theorem is omitted.
The content of the theorem is often stated by saying that the set $\{ Y_\ell^m \}$ of surface spherical harmonics is an orthonormal basis for $\mathcal{L}_2(\Omega)$. Every function $f \in \mathcal{L}_2(\Omega)$ may be written in the form $\sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_\ell^m Y_\ell^m(\hat{r})$ in a unique manner and the series is said to <u>converge to f in the mean</u>.

<u>Further results</u>: The above result concerns mean square convergence. What can be said about pointwise convergence, i.e. when can we say that

$$\lim_{L\to\infty} f_L(\hat{r}) = f(\hat{r})$$

at any or all $\hat{r} \in \Omega$? Answer: It can be shown that if $f : \Omega \to C$ is <u>continuously differentiable</u> on Ω, then $f_L(\hat{r}) \to f(\hat{r})$ <u>uniformly</u>.

If however $f : \Omega \to C$ is merely continuous on Ω but not necessarily differentiable, then it is possible that $\lim_{L\to\infty} f_L(\hat{r})$ can <u>fail to exist</u> <u>almost everywhere</u>, i.e., everywhere except on a set of total area zero.

Thus mere continuity of $f: \Omega \to C$ does not by any means assure pointwise convergence of the series of partial sums $f_L(\hat{r})$ to $f(\hat{r})$. We know however that any continuous $f: \Omega \to C$ can be uniformly approximated by a series

$$\sum_{\ell=0}^{L} \sum_{m=-\ell}^{\ell} \gamma_\ell^m Y_\ell^m(\hat{r})$$

(see theorem on page 37).

Can such a series be exhibited? The answer is yes. It can be shown that ~~[illegible]~~ if one defines the so-called Césaro mean as the arithmetic mean of the partial sums $f_L(\hat{r})$

$$\sigma_L(\hat{r}) = \frac{f_0(\hat{r}) + \dots \; f_{L-1}(\hat{r})}{L}, \quad L = 1, 2, \dots$$

then

$$\sigma_L(\hat{r}) \to f(\hat{r}) \text{ at every } \hat{r} \text{ where } f(\hat{r}) \text{ is continuous}$$

Internal and External Dirichlet Problems

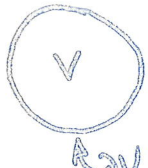
Theorem: Suppose $f \in \mathcal{L}_2(\Omega)$ and $f_\ell^m = (Y_\ell^m, f)$. Suppose $\nabla^2 \phi = 0$ in $0 \leq r < a$ and $\int_\Omega |\phi(a\hat{r}) - f(\hat{r})|^2 d\Omega = 0$. Then

$$\phi(r\hat{r}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} f_\ell^m \left(\frac{r}{a}\right)^\ell Y_\ell^m(\hat{r}) \text{ in } r \leq a \text{ and if } 0 \leq \tau < 1$$

then the sum and all its derivatives are absolutely and uniformly convergent for $0 \leq r \leq \tau a$.

Proof:

1. ϕ is uniquely determined by $\nabla^2\phi = 0$ in $0 \le r < a$ and $\int_\Omega |\phi(a\hat{r}) - f(\hat{r})|^2 dA = 0$, for if ϕ and $\tilde{\phi}$ both satisfy these conditions and $\psi = \phi - \tilde{\phi}$, then $\nabla^2\psi = 0$ in V and $\psi = 0$ on ∂V. Then using Green's identity



$$\int_V |\nabla\psi|^2 dV = \int_V [\nabla\cdot(\psi\nabla\psi) - \psi\nabla^2\psi]\, dV = \int_{\partial V} \hat{n}\cdot\psi\nabla\psi\, dA$$

$= 0$, hence $\nabla\psi = \text{constant} = 0$, hence $\psi = \text{constant}$ but $\psi = 0$ on ∂V, hence $\psi = 0$.

2. if $\phi(r\hat{r})$ is the series given above, then ϕ satisfies $\nabla^2\phi = 0$ for $r < a$, $\phi = f$ on $r = a$.

(a) $|f_\ell^m| \le \|f\|$ by Schwarz inequality

$|Y_\ell^m| \le \sqrt{\frac{2\ell+1}{4\pi}}$

thus the ℓm term in the series is bounded by

$\|f\| \sqrt{\frac{2\ell+1}{4\pi}}\, \tau^\ell$

Thus by the Weierstrass M-test, the series converges absolutely and uniformly for $0 \le r \le \tau a$.

Therefore it is differentiable term by term

(b) $\nabla^2\phi = \sum_\ell \sum_m f_\ell^m \nabla^2 \left(\frac{r}{a}\right)^\ell Y_\ell^m(\hat{r}) = 0$ and

$\phi(a\hat{r}) = \sum_\ell \sum_m f_\ell^m Y_\ell^m(\hat{r}) = f(\hat{r})$

Corollary: (The mean value theorem for harmonic functions)

If $\nabla^2\phi = 0$ in $B(\vec{r}, a)$, then

$$\phi(\vec{r}) = \frac{1}{4\pi a^2} \int_{\partial B(\vec{r}, a)} \phi(\vec{\rho})\, dA$$

i.e., a function ϕ harmonic in a sphere of radius a is equal to its mean taken over the surface of the sphere

Proof: suffices to prove for $\vec{r}=0$, just shift origin for any other $\vec{r}$

But $\phi(0) = f_0^0 Y_0^0 = Y_0^0 \int_\Omega Y_0^{0*} f = \frac{1}{4\pi}\int_\Omega f(\vec{\rho})\, dA$

$$= \frac{1}{4\pi a^2}\int_{\Omega_a} f(\vec{\rho})\, dA = \frac{1}{4\pi a^2}\int_{\Omega_a} \phi(\vec{\rho})\, dA$$

Corollary: If $\nabla^2\phi = 0$ in an open set V and $\vec{r}_0 \in V$ and $\phi(\vec{r}_0) \geq \phi(\vec{r})$ for all $\vec{r}\in V$, then $\phi(\vec{r}) = \phi(\vec{r}_0)$ for all $\vec{r}\in V$.

(A harmonic function with an interior maximum is a constant). Same true of minimum.

<u>Theorem</u>: Suppose $f \in \mathcal{L}_2(\Omega)$ and $f_\ell^m = (Y_\ell^m, f)$. Suppose

1. $\nabla^2\phi = 0$ for $r > a$
2. $\lim_{r\to\infty} \phi(r\hat{r}) = 0$ uniformly for all $\hat{r}$ on Ω
3. $\int_\Omega |\phi(r\hat{r}) - f(\hat{r})|^2 dA = 0$

Then $\phi(r\hat{r}) = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell} f_\ell^m \left(\frac{a}{r}\right)^{\ell+1} Y_\ell^m(\hat{r})$ for $a \leq r < \infty$

and if $\tau > 1$, then the series and all its derivatives converge absolutely and uniformly for $r \geq \frac{a}{\tau}$.

<u>Proof</u>: If ϕ and $\tilde{\phi}$ satisfy 1.2.3. above then $\psi = \phi - \tilde{\phi}$ has $\nabla^2\psi = 0$, $\psi = 0$ on ∂V, $\psi = 0$ at ∞. If $\psi \not\equiv 0$, then ψ has either an interior maximum or minimum, but this is impossible, hence $\psi \equiv 0$

The series does satisfy 1.2.3. above, since

$$\nabla^2 \frac{1}{r^{\ell+1}} Y_\ell^m(\hat{r}) = \left(\frac{1}{r^2}\partial_r^2 r + \frac{\Lambda^2}{r^2}\right) r^{-\ell-1} Y_\ell^m(\hat{r})$$

$$= \left[-\ell(-\ell-1) - \ell(\ell+1)\right] r^{-\ell-3} Y_\ell^m(\hat{r}) = 0$$

V. Operators on vector spaces

1. <u>mapping</u> : f <u>maps</u> a set U <u>into</u> a set V if f is a rule which assigns to each $u \in U$ a unique member of V
$f: U \to V$ (notation). The $v \in V$ assigned to u is written $f(u)$

2. <u>sum of two mappings</u> : $f: U \to V$ and $g: U \to V$
define $af + bg : U \to V$ to be that rule which assigns to $u \in U$, the vector $af(u) + bg(u)$

3. composition or <u>product</u> : $f: U \to V$ and $g: V \to W$
product gf assigns to u $g[f(u)] \in W$
$gf: U \to W$

4. <u>linear mapping</u> : Let V, W be <u>vector spaces</u>
$f: V \to W$ is linear if $\forall$ $u, v \in V$
$$f(au + bv) = af(u) + bf(v)$$

4.5. <u>linear operator on a vector space</u> V : if $V = W$ above $f: V \to V$

Examples

1. $f(v) = v$ $f: V \to V$
identity operator, denoted by $I(v) = v$

2. $V \equiv R_3$ $\quad f: R^3 \to R^3$

$$f(v) = \underline{\Omega} \times \underline{v}, \quad \underline{\Omega} \text{ fixed}$$

3. $f(v) = av$, a a fixed scalar

4. $V \equiv C^\infty[R^3]$ all ∞- cont. differ scalar fcns on R^3.
   $f(v) = \nabla V$, gradient operator
   $f: V \to V$

5. V finite dimensional, let $\{v_1, \dots v_n\}$ be an orthonormal basis, let A_{ij} be any $n \times n$ matrix of fixed numbers. $\forall v$ we have $v = a_i v_i$. Now write $A(v) = v_i A_{ij} a_j$. Clearly $A(v) \in V$ and depends linearly on V. Hence $A: V \to V$ is a linear operator on V. $A(v)$ often Av
   Note that

$$A v_j = v_i A_{ij}$$

In fact <u>all</u> linear operators on a finite dim vector space V are of the form of example 5 above.

5. <u>Matrix of a linear operator</u>
   V finite dim $\{v_1, \dots v_n\}$ an orthonormal basis. <u>Define</u> y_j by

$$y_j = A v_j \quad \text{, now } y_j \in V \text{ hence}$$
$$y_j = A v_j = A_{ij} v_i$$

$$\star \quad \boxed{A v_j = A_{ij} v_i} \qquad \leftarrow \text{typo in notes}$$

The matrix A defined by $\star$ is said to be the <u>matrix of the lin. op. w.r.t.</u> $\{v_i\}$.
Given $A: V \to V$ $\exists$ a matrix A w.r.t. every orthonormal basis. Knowledge of any such matrix completely determines $A: V \to V$, since for any $v = a_j v_j \in V$

$$A v = A a_j v_j = a_j A v_j = a_j y_j$$
$$= a_j A_{ij} v_i$$

i.e. if $z = b_i v_i = Av$, then

$$b_i = A_{ij} a_j \quad , \text{or}$$
$$b = Aa$$

6. Change of orthonormal basis

Say we know the matrix A of a lin. op $A: V \to V$ w.r.t. one orthonormal basis $\{v_i\}$
What is the matrix A' w.r.t. another $\{v_i'\}$?

$$A v_j = A_{ij} v_i$$

now $v_j = Q_{mj} v_m'$

$$v_i = Q_{ki} v_k'$$

$$A(Q_{mj} v_m') = Q_{mj} A(v_m') =$$
$$A_{ij} Q_{ki} v_k'$$

mult by Q_{jl}^{-1}

$$A(v_m') Q_{mj} Q_{jl}^{-1} = Q_{ki} A_{ij} Q_{jl}^{-1} v_k'$$

$$A(v_m') \delta_{ml} = A(v_l') = Q_{ki} A_{ij} Q_{jl}^{-1} v_k'$$

$A(v_l') = A_{kl}' v_k'$ since unitary $= Q_{ki} A_{ij} Q^{\dagger}_{jl} v_k'$

Thus $\boxed{\begin{aligned} A'_{kl} &= Q_{ki} A_{ij} Q^{+}_{jl} \\ A' &= Q A Q^{+} = Q A Q^{-1} \end{aligned}}$

or

Concepts 1. linear operator on V

2. matrix of a linear operator w.r.t. a basis

$A v_j = A_{ij} v_i$

if $z = b_i v_i = A(a_i v_i)$, then

$b_i = A_{ij} a_i$ or $b = Ab$

3. change of matrix under change of linear operator.

End of Chapter 1.

Lecture 16. Homework due Thursday

Strongly recommend doing tensor problems.

Chapter 3: Tensors

Now restrict attn to vectors in R_3. Dot product $\underline{a}\cdot\underline{b}$. $\underline{a}\cdot\underline{b} = |\underline{a}||\underline{b}|\cos\theta$

orthonormal bases are CAS $\hat{x}_1, \hat{x}_2, \hat{x}_3$

unitary-orthog.

Section 1: Linear functionals

Definition: Let V be a vector space and let f be a function which assigns to each vector $\vec{v} \in V$ a scalar $f(\vec{v})$. Then f is a functional on V. If, for any scalars a and b and any vectors $\vec{u}$ and $\vec{v}$, we have

$$f(a\vec{u} + b\vec{v}) = af(\underline{u}) + bf(\underline{v})$$

then f is a linear functional on V.

Corollary: if f is a linear functional, then $f(\vec{0}) = 0$ since $f(\vec{0}) = f(0\vec{0}) = 0 f(\vec{0}) = 0$

Definition: if f and g are functionals and a and b are scalars, we define $af + bg$ to be the functional which assigns to any vector $\vec{v} \in V$ the scalar $af(\vec{v}) + bg(\vec{v})$

$$(af + bg)(\vec{v}) = af(\vec{v}) + bg(\vec{v}) \qquad (*)$$

Corollary — if f and g are linear functionals on V, then so is $af + bg$

Corollary — the set of all linear functionals on V is itself a vector space (proof left to reader)

It is possible to examine linear functionals on an arbitrary vector space V. We shall restrict ourselves to the study of linear functionals and tensors defined on R^3 (real 3-space), since it is for the most part tensors in real 3-space which we shall have a need for. You should recognize that everything we say can be readily extended to the general case. If one wishes to consider a <u>complex</u> vector space V, one discovers that it is a fact of life that complex conjugates $*$'s appear in many of the formulae in this chapter. For example, because of the way we have defined the inner product (page 10) one finds that in order to make things come out right, one has to <u>redefine</u> eqn ($\star$) on page (57) to read

$$(af+bg)(\vec{v}) = a^* f(\vec{v}) + b^* g(\vec{v})$$

You might find it instructive to see what changes are necessary in this chapter in order that everything is valid for an arbitrary vector space V.

From now on $V \equiv R^3$
Any Cartesian axis system $\{\vec{0};\hat{x}_1,\hat{x}_2,\hat{x}_3\}$ is an orthonormal basis for R^3

<u>Theorem</u>: Let f be a linear functional on R^3. Then
(i) $\exists$ exactly one vector $\vec{f} \in R^3$ such that for every $\vec{v} \in R^3$ we have

$$\boxed{f(\vec{v}) = (\vec{f},\vec{v}) = \vec{f}\cdot\vec{v}}$$

We say that $\vec{f}$ represents or is generated by f, and vice-versa.

(ii) If f and g are linear functionals represented by $\vec{f}$ and $\vec{g}$, then $a\vec{f} + b\vec{g}$ represents $af + bg$.

Proof: (i) existence — Let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ be an arbitrary Cartesian axis system and let

$$\boxed{\vec{f} = [f(\hat{x}_i)]\hat{x}_i}$$

(Note: if V is complex, we must take $\vec{f} = [f(\vec{v}_i)]^* \vec{v}_i$)

Now if $\vec{v} \in R^3$, then $\vec{v} = (\hat{x}_i \cdot \vec{v})\hat{x}_i$, so $f(\vec{v}) =$

$$f[(\hat{x}_i \cdot \vec{v})\hat{x}_i] = (\hat{x}_i \cdot \vec{v}) f(\hat{x}_i) = (f(\hat{x}_i)\hat{x}_i) \cdot \vec{v} = \vec{f} \cdot \vec{v}$$

(i) uniqueness — Suppose $\exists$ two vectors $\vec{f}, \vec{g}$ such that $f(\vec{v}) = \vec{f} \cdot \vec{v} = \vec{g} \cdot \vec{v}$ for all $\vec{v}$. Then in particular it must be true for $\vec{v} = \vec{f} - \vec{g}$, so $(\vec{f} - \vec{g}) \cdot (\vec{f} - \vec{g}) = 0$ or $\vec{f} - \vec{g} = 0$ or $\vec{f} = \vec{g}$

(ii) $(af + bg)(\vec{v}) = af(\vec{v}) + bg(\vec{v}) = a\vec{f} \cdot \vec{v} + b\vec{g} \cdot \vec{v}$

$= (a\vec{f} + b\vec{g}) \cdot \vec{v}$ We have not shown that all f are so generated, but this obvious

The above theorem shows that for all practical purposes a vector $\vec{v} \in R^3$ may be thought of a identical with the linear functional which represents it. In other words we have another way of thinking about vectors besides the usual geometric picture that a vector has direction and length. A vector is also a linear functional.

This new geometrical picture of a vector does not appear at first sight to be a very useful one, but in fact it is an extremely useful correspondence since it is possible to make a simple extension of the concept of a linear functional in order to give a completely geometrical (i.e., one which does not in any way depend on the choice of a basis) definition of a tensor.

A tensor of order q in 3-space is defined in most physics books as a rule which assigns to every basis $\hat{x}_1, \hat{x}_2, \hat{x}_3$ in 3-space an array of order q of dimension $\underbrace{3 \times 3 \times \cdots \times 3}_{q}$, in such a way that the arrays assigned to different bases are related in a certain way. This definition leads to the same objects which we are about to define in a purely geometric manner. The objection to the older definition is that it leads to long-winded algebraic calculations to prove things which are geometrically obvious, and its notation stifles rather than promotes the use of geometric intuition in visualizing tensors.

Section 2: Multilinear functionals

Definition: A multilinear functional of order q on a vector space V is a rule which assigns a real (or complex) number $M(\vec{u}_1, \dots \vec{u}_q)$ to any ordered q-tuple of vectors $\vec{u}_1, \dots \vec{u}_q \in V$, in such a way that M is linear separately in each of its arguments. That is, if $\vec{u}_1, \dots \vec{u}_{i-1}, \vec{u}_{i+1} \dots \vec{u}_q$ are held fixed then for any a and b, and any $\vec{u}_i, \vec{v}_i \in V$ we have

$$M(\vec{u}_1, \dots, a\vec{u}_i + b\vec{v}_i, \dots \vec{u}_q) = aM(\vec{u}_1 \dots \vec{u}_i, \dots \vec{u}_q) + bM(\vec{u}_1, \dots \vec{v}_i, \dots \vec{u}_q)$$

It is important to note that the order of the vectors in the q-tuple is important. It is not assumed that

$$M(\vec{u}_1, \vec{u}_2, \dots \vec{u}_q) = M(\vec{u}_2, \vec{u}_1, \dots \vec{u}_q)$$

Example: Suppose f and g are linear functionals. Define $M(\vec{u},\vec{v}) = f(\vec{u})\, g(\vec{v})$. Then M is a multilinear functional of order two. Multilinear functionals of order two are called bilinear functionals.

Example: Suppose $f_1, \dots f_n$ are linear functionals. Define $M(\vec{v}_1, \dots \vec{v}_n) = f_1(\vec{v}_1) \dots f_n(\vec{v}_n)$. Then M is a multilinear functional of order n.

Example: Let V be an inner product space and let $M(\vec{u},\vec{v}) = (\vec{u},\vec{v}) = \vec{u}\cdot\vec{v}$. Then M is a bilinear functional

Example: Let V be R^3 (real 3-space) and let $M(\vec{u},\vec{v},\vec{w}) = \vec{u}\cdot(\vec{v}\times\vec{w})$. Then M is a multilinear functional of order 3 (trilinear functional) in 3-space.

Section 3 : Tensors

Definition: A tensor of order q is a multilinear functional of order q defined on an inner product space V (in our case R^3)

The two terms "tensor" and multilinear functional are synonymous Note that a tensor of order 1 is just a vector. Both are linear functionals on V. To avoid exceptional cases in theorems it is convenient also to define tensors of order zero.

Definition: A tensor of order zero is simply any scalar

Example $1, \pi, -19/11$, etc.

Four operations

scalar mult $\left.\right\}$ vector space $\overset{9}{\otimes} R^3$ 61

sum

tensor product

trace

Section 4: Tensor Arithmetic

Definition (sum and product by scalars)

Suppose a and b are scalars and M and N are tensors of the same order q over a vector space V. Then we denote by $aM + bN$ that tensor of order q which assigns to any ordered q-tuple $\vec{u}_1, \ldots \vec{u}_q$ the scalar

$$aM(\vec{u}_1, \ldots \vec{u}_q) + bN(\vec{u}_1, \ldots \vec{u}_q)$$

Note the appearance of $*$'s in this equation if V is complex

Definition (Tensor Product) If L is a tensor of order p and M a tensor of order q over vector space V, we define LM to be that function which assigns to any ordered $p+q$-tuple $\vec{u}_1, \ldots \vec{u}_p, \vec{v}_1, \ldots \vec{v}_q$ the scalar

$$LM(\vec{u}_1, \ldots \vec{u}_p, \vec{v}_1, \ldots \vec{v}_q) = L(\vec{u}_1, \ldots \vec{u}_p)\, M(\vec{v}_1, \ldots \vec{v}_q)$$

The function LM is called the "tensor product" or "outer product" of L and M and is sometimes written $L \otimes M$

Note that we most certainly do not have $LM = ML$.

The first thing to notice about $aM + bN$ and $LM = L \otimes M$ is that they are tensors, $aM + bN$ is a tensor of order q and LM is a tensor of order $p+q$.

The proofs that $aM + bN$ and LM are tensors are straightforward. One merely has to demonstrate that $aM + bN$ depends linearly on each of $\vec{u}_1, \ldots \vec{u}_q$ and that LM depends linearly on each of $\vec{u}_1, \ldots \vec{u}_p, \vec{v}_1, \ldots \vec{v}_q$. It might be instructive to go thru the proofs

Problem 11: Go thru them.

The rules of tensor arithmetic are easily deduced from the definitions. In fact the rules are exactly the same as the rules for array arithmetic.

Addition: if A, B, C have the same order q then

$A + B = B + A$

$A + (B+C) = (A+B)+C$

Scalar multiplication $(ab)A = a(bA)$

Scalar multiplication is distributive

$a(A+B) = aA + aB$

$(a+b)A = aA + bA$

$0A = 0$, where the 0 tensor of order q is defined to be that multilinear which assigns to every ordered q-tuple the scalar 0.

<u>Remark</u>: We have just seen that if M and N are tensors of order q over a vector space V, then so is $aM + bN$. Furthermore the above arithmetical rules are exactly those possessed by a vector space. It follow that

> The set of all tensors of order q over an inner product space is another vector space, usually denoted $\underbrace{V \otimes V \otimes \cdots \otimes V}_{q} = \otimes^q V$

The other arithmetic rules governing the tensor product are these

For any L, M, N

$(LM)N = L(MN)$ associative rule

For L of order p, M and N of order q

$$\left.\begin{aligned} L(aM + bN) &= a(LM) + b(LN) \\ (aM+bN)L &= a(ML) + b(NL) \end{aligned}\right\} \text{distributive rules}$$

Note: in general $LM \neq ML$, but if L is of order zero (a scalar), then $LM = ML$.

Note also that the associative rule for tensor products means that we can omit parentheses in expressions like LMN.

Definition: If $\vec{f}$ and $\vec{g}$ are two vectors $\in R^3$ and f and g are the associated linear functionals, then the bilinear functional (tensor product of f and g) defined by

$$\vec{f}\vec{g}(\vec{u},\vec{v}) = f(\vec{u})\,g(\vec{v}) = (\vec{f}\cdot\vec{u})(\vec{g}\cdot\vec{v})$$
$$= f_i g_j u_i v_j$$

is called a dyad.

Likewise the tensor product of q vectors is called a polyad of order q (one speaks of triads, etc.)

$$\vec{f}_1 \ldots \vec{f}_q(\vec{u}_1, \ldots, \vec{u}_q) = (\vec{f}_1\cdot\vec{u}_1)\ldots(\vec{f}_q\cdot\vec{u}_q)$$

Problem 12: If $\vec{f}$ and $\vec{g}$ are two vectors in R^3, give a geometrical relation between $\vec{f}$ and $\vec{g}$ such that $\vec{f}\vec{g} = \vec{g}\vec{f}$

Section 5 — Two Special Tensors

1. Let $\vec{u}, \vec{v} \in R^3$, and define the identity tensor to be the second order tensor I defined by $I(\vec{u},\vec{v}) = \vec{u}\cdot\vec{v}$. The reason for the name will be apparent later

2. The alternating tensor in R^3. For every $\vec{u},\vec{v},\vec{w} \in R^3$ define call this $A(\vec{u},\vec{v},\vec{w}) = \vec{u}\cdot(\vec{v}\times\vec{w})$

It might be thought at this stage that every tensor of order q is in fact a polyad of order q. This is not true if $q>1$ and $n \equiv \dim V > 1$. For example

Problem 12: Let $V \equiv \mathbb{R}^3$. Show that the identity tensor is not a dyad. Now find out whether the alternating tensor is a triad. Prove your answer.

Section 6. Trace or Contraction

Let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ and $\{\hat{x}_1', \hat{x}_2', \hat{x}_3'\}$ be two different Cartesian axis systems in $\mathbb{R}^3$. Now

$$\hat{x}_j' = (\hat{x}_i \cdot \hat{x}_j')\hat{x}_i$$

Now dot $\hat{x}_k'$ into the above equation

$$(x_j' \cdot x_k') = (\hat{x}_i \cdot \hat{x}_j')(\hat{x}_i \cdot \hat{x}_k') \quad \text{but } \hat{x}_j' \cdot x_k' = \delta_{jk}$$

so

$$\boxed{(\hat{x}_i \cdot \hat{x}_j')(\hat{x}_i \cdot \hat{x}_k') = \delta_{jk}}$$

Now consider the bilinear functional T and consider the sum $T(\hat{x}_i, \hat{x}_i)$ where $\{x_i\}$ is an orthonormal basis. We assert that the value of this sum depends only on T and not on the particular Cartesian axis system used for its calculation. To see this, we write

$$T(\hat{x}_i, \hat{x}_i) = T[(\hat{x}_j' \cdot \hat{x}_i)x_j', (\hat{x}_k' \cdot \hat{x}_i)\hat{x}_k']$$

$$= (\hat{x}_j' \cdot \hat{x}_i)(\hat{x}_k' \cdot \hat{x}_i)\, T(\hat{x}_j', \hat{x}_k'), \text{ but}$$

$(\hat{x}_j' \cdot \hat{x}_i)(\hat{x}_k' \cdot \hat{x}_i) = \delta_{jk}$, so

$$\boxed{T(\hat{x}_i, \hat{x}_i) = T(\hat{x}_j', \hat{x}_j')}$$

Thus we may adopt the following

Definition: (trace of a second order tensor)

If T is a second order tensor, then the trace or contraction of T, written $\operatorname{tr} T$ is defined as $T(\hat{x}_i, \hat{x}_i)$ where $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ is an arbitrary Cartesian axis system. The value of the trace does not depend on which Cartesian axis system is used to compute it. The trace of a second order tensor is of course a scalar.

More generally we can define the trace of a tensor of any order q. If T is a tensor of order q, and $\vec{u}_1, \dots \vec{u}_{q-2}$ are any $q-2$ vectors $\in \mathbb{R}^3$, we can always compute the sum

$$T(\vec{u}_1, \dots \vec{u}_{r-1}, \underset{\substack{\uparrow \\ r\text{th place}}}{\hat{x}_i}, \vec{u}_{r+1}, \dots \vec{u}_{s-1}, \underset{\substack{\uparrow \\ s\text{th place}}}{\hat{x}_i}, \vec{u}_{s+1}, \dots, \vec{u}_{q-2})$$

Exactly the same argument as above shows that the above sum of three terms is independent of the axis system $\hat{x}_1, \hat{x}_2, \hat{x}_3$ used to compute it. The sum depends only on T and on $\vec{u}_1, \dots \vec{u}_{q-2}$. Moreover it depends linearly on $\vec{u}_1, \dots \vec{u}_{q-2}$. Thus the sum defines a tensor of order $q-2$.

$$\operatorname{tr} T(\vec{u}_1, \dots \vec{u}_{r-1}, \vec{u}_{r+1}, \dots \vec{u}_{s-1}, \vec{u}_{s+1}, \dots \vec{u}_{q-2})$$

<u>Definition</u> Let T be a tensor of order q and let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ be any Cartesian axis system. For any vectors $\vec{u}_1, \dots \vec{u}_{q-2}$, define the $q-2$ order tensor $tr_{rs}T$, called the <u>contraction or trace on the r, s indices</u>, by

$$\boxed{tr_{rs}T(\vec{u}_1, \dots \vec{u}_{q-2}) = T(\vec{u}_1, \dots \vec{u}_{r-1}, \hat{x}_i, \vec{u}_{r+1}, \dots \vec{u}_{s-1}, \hat{x}_i, \vec{u}_{s+1}, \dots \vec{u}_{q-2})}$$

The resulting tensor of order $q-2$ does not depend on $\{\hat{x}_i\}$.

Note that our definition of trace refers to a basis, but it doesn't matter which basis we use. The trace $tr_{rs}T$ is thus a geometrical object. It should thus be possible to define $tr_{rs}T$ without ever even introducing a basis, and in fact it is, but it is somewhat complicated, and not very enlightening

Rules for operating with traces are simple to prove

$$tr_{rs}(R+S) = tr_{rs}R + tr_{rs}S$$

$$tr_{rs}(aR) = a\, tr_{rs}R$$

<u>Problem 13</u>: Find the traces of the following tensors M

1. $M = \vec{f}\vec{g}$
2. $M = I$
3. $M = A$ (alternating tensor)
4. $M = \vec{f}\vec{g}\vec{h}$ (find $tr_{12}M$, $tr_{13}M$, $tr_{23}M$)
5. $\vec{\Omega}$ a fixed vector, $M(\vec{u}, \vec{v}) = \vec{u} \cdot \vec{\Omega} \times \vec{v}$
6. $M(\vec{u}, \vec{v}, \vec{w}, \vec{x}) = (\vec{u} \cdot \vec{v})(\vec{w} \cdot \vec{x})$, find $tr_{12}M$, $tr_{23}M$, $tr_{34}M$

Define transpose $T^T(\vec{u}, \vec{v}) = T(v, u)$ and $T_{rs}T$ more generally

Lecture 17: 2. Review defn of tensor, 3. four operations of tensor arithmetic, stress geometric nature of defn. 1. Isomorphism: linear fcnal and vectors $\in R^3$.

Section 7: Cartesian Components of Tensors

The components of any vector $\vec{f} \in R_3$ are $f_i = (\hat{x}_i, \vec{f})$, $\vec{f} = f_i \hat{x}_i$. Now if f is the linear functional generated by $\vec{f}$, we can write the f_i in terms of f, since

$$\boxed{f_i = \vec{f} \cdot \hat{x}_i = f(\hat{x}_i)}$$

The above eqn makes it reasonable to call $f(\hat{x}_i)$ the i^{th} component of the linear functional f with respect to the basis $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$. This suggests a definition which applies to tensors of order two or more.

Definition – Let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ be a Cartesian axis system. Let T be a tensor of order $q \geq 1$. Then the components of T relative to $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ are defined by

$$\boxed{T_{i_1 \dots i_q} = T(\hat{x}_{i_1}, \dots \hat{x}_{i_q})}$$

Example: a tensor of order 2 has nine components

Example: components of I

Section 8: The vector space $\otimes^q R^3$

If we know the tensor T, we can calculate its Cartesian components w.r.t. any Cartesian axis system $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$. The tensor T completely determines its components. Now we know that if the tensor is of order 1 (a vector) and we know only its components f_i w.r.t. an arbitrary $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$, then in fact the vector itself is completely determined since

$$\vec{f} = f_i \hat{x}_i$$

Is a tensor T of any order $\underline{q}$ uniquely determined by its components? This question is answered affirmatively by the following theorem.

<u>Theorem</u>: Let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ be an orthonormal basis for $\mathbb{R}^3$. Then the 3^q polyads of order $\underline{q}$, $\hat{x}_{i_1} \cdots \hat{x}_{i_q}$ constitute a basis for the space $\otimes^q \mathbb{R}^3$ of all tensors of order $\underline{q}$. In fact any tensor T of order $\underline{q}$ $\in \otimes^q \mathbb{R}^3$ may be written as

$$\boxed{T = T_{i_1 \cdots i_q} \hat{x}_{i_1} \cdots \hat{x}_{i_q} = T(\hat{x}_{i_1}, \cdots \hat{x}_{i_q})\, \hat{x}_{i_1} \cdots \hat{x}_{i_q}}$$

<u>Proof</u>: To show that $T = T(\hat{x}_{i_1}, \cdots \hat{x}_{i_q}) \hat{x}_{i_1} \cdots \hat{x}_{i_q}$, we must show that these two tensors assign the same number to every ordered $\underline{q}$-tuple $\vec{u}^{(1)}, \cdots \vec{u}^{(q)}$ of vectors.

$$T(\vec{u}^{(1)}, \cdots \vec{u}^{(q)}) = T\left[(\hat{x}_{i_1} \cdot \vec{u}^{(1)})\hat{x}_{i_1}, \cdots \quad (\hat{x}_{i_q} \cdot \vec{u}^{(q)})\hat{x}_{i_q}\right]$$

$$= (\hat{x}_{i_1} \cdot \vec{u}^{(1)}) \cdots (\hat{x}_{i_q} \cdot \vec{u}^{(q)})\, T(\hat{x}_{i_1}, \cdots \hat{x}_{i_q})$$

$$= T(\hat{x}_{i_1}, \cdots \hat{x}_{i_q}) \left[\hat{x}_{i_1} \cdots \hat{x}_{i_q} (\vec{u}^{(1)}, \cdots \vec{u}^{(q)})\right]$$

recall defn $a_i u_i = 0 \Rightarrow a_i = 0$

$$= \left[T(\hat{x}_{i_1}, \cdots \hat{x}_{i_q}) \hat{x}_{i_1} \cdots \hat{x}_{i_q}\right] (\vec{u}_{(1)}, \cdots \vec{u}_{(q)})$$

Hence the 3^q polyads $\hat{x}_{i_1} \cdots \hat{x}_{i_q}$ span $\otimes^q \mathbb{R}^3$.

To prove they are a basis we must prove that they are linearly independent. Suppose $\exists$ scalars $a_{i_1 \cdots i_q}$ such that

$a_{i_1 \cdots i_q} \hat{x}_{i_1} \cdots \hat{x}_{i_q} = 0$, 0 here meaning the multilinear functional which assigns the number 0 to every ordered $\underline{q}$-tuple

Then in particular, it must assign 0 to the q-tuple $\hat{x}_{j_1}, \ldots \hat{x}_{j_q}$ for some fixed $j_1, \ldots j_q$.

$$[a_{i_1 \ldots i_q} \hat{x}_{i_1} \ldots \hat{x}_{i_q}](\hat{x}_{j_1}, \ldots, \hat{x}_{j_q})$$

$$= a_{i_1 \ldots i_q} [(\hat{x}_{i_1} \cdot \hat{x}_{j_1}) \ldots (\hat{x}_{i_q} \cdot \hat{x}_{j_q})]$$

$$= a_{i_1 \ldots i_q} \delta_{i_1 j_1} \ldots \delta_{i_q j_q}$$

$$= a_{j_1 \ldots j_q} = 0 \text{ by hypothesis}$$

But since $j_1, \ldots, j_q$ is arbitrary, all the $a_{i_1 \ldots i_q}$ are zero. Thus the 3^q polyads are linearly independent. Thus they form a basis.

Corollary — the dimension of $\otimes^q R^3$ is 3^q

A tensor completely det. by its components w.r.t. any basis

Corollary — If $T_{i_1 \ldots i_q}$ are the components of the q^{th} order tensor T relative to $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ and if $\vec{u}^{(1)} = u_i^{(1)} \hat{x}_i, \ldots$ $\vec{u}^{(q)} = u_i^{(q)} \hat{x}_i$, then

$$T(\vec{u}^{(1)}, \ldots \vec{u}^{(q)}) = T_{i_1 \ldots i_q} u_{i_1}^{(1)} \ldots u_{i_q}^{(n)}$$

Section 9: Tensor arithmetic in terms of components

If $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ is an orthonormal basis for R^3, then not only does every tensor of order q define a unique array of order q and dimension 3^q, but also every such array defines a unique q^{th} order tensor over R^3. If $T_{i_1 \ldots i_q}$ are the elements of such an array then T is a tensor of order q,

$$T = T_{i_1 \ldots i_q} \hat{x}_{i_1} \ldots \hat{x}_{i_q}, \text{ and}$$

$$T(\hat{x}_{i_1}, \ldots, \hat{x}_{i_q}) = T_{i_1 \ldots i_q}$$

This one-to-one correspondence between tensors and arrays of order q depends on the basis $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$. Such a correspondence exists for every basis, but it is a different correspondence for every basis.

A number of useful arithmetic properties of tensor components follow directly from the definitions:

1. Let R and T be tensors of order q, and let all components be relative to the same basis $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$. Then

$$(aR + bT)_{i_1 \cdots i_q} = a R_{i_1 \cdots i_q} + b T_{i_1 \cdots i_q}$$

Because of this the vector space $\otimes^q \mathbb{R}^3$ is essentially identical with the vector space of q^{th} order arrays of dimension 3^q. Each basis of V is said to generate an isomorphism between these two spaces. ~~Example: tensors of order 2 $\leftrightarrow$ matrices (3×3) All rules of arithmetic are the same.~~

Continuing relations among components, we have

2. Let S be a tensor of order q and T a tensor of order p. Then

$$(ST)_{i_1 \cdots i_q j_1 \cdots j_p} = S_{i_1 \cdots i_q} T_{j_1 \cdots j_p}$$

3. $$(\mathrm{tr}_{rs} T)_{i_1 \cdots i_{q-2}} = T_{i_1 \cdots i_{r-1} \underset{\substack{\uparrow \\ r^{th}\text{ place}}}{j} i_r \cdots i_{s-2} \underset{\substack{\uparrow \\ s^{th}\text{ place}}}{j} i_{s-1} \cdots i_{q-2}}$$

Section 10 — Examples of Tensor Components

Example 1: Any second order tensor over R^3 has nine components

$$T_{ij} = T(\hat{x}_i, \hat{x}_j)$$

Example 2: I, the identity tensor has components

$$I_{ij} = I(\hat{x}_i, \hat{x}_j) = \hat{x}_i \cdot \hat{x}_j = \delta_{ij}$$

The components of I are the same in every Cartesian axis system (the elements of the 3x3 identity matrix). Such a tensor is called an isotropic tensor.

Example 3: The components of a qth order polyad may be expressed in terms of the components of the q vectors which comprise it

$$(\vec{f}\vec{g}\vec{h})_{ijk} = f_i g_j h_k \text{ , for example}$$

Example 4: The alternating tensor A has components

$A_{ijk} = \hat{x}_i \cdot (\hat{x}_j \times \hat{x}_k)$. Now if $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ is right handed $A_{ijk} = \epsilon_{ijk}$, but if it is left handed, $A_{ijk} = -\epsilon_{ijk}$. The components of A in any right handed Cartesian axis system are the same, but they change sign in a left-handed system. Such a tensor is called an improper isotropic tensor.

Section 11: Relations between tensor components relative to different bases

If T is a qth order tensor whose components $T_{i_1 \dots i_q}$ relative to $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ are known, then T is determined and hence its components $T'_{j_1 \dots j_q}$ relative to some other orthonormal basis $\{\hat{x}_1', \hat{x}_2', \hat{x}_3'\}$ are determined, and ought to be calculable.

There are several ways to do the calculation

1. $T'_{j_1 \cdots j_q} = T(\hat{x}'_{j_1}, \ldots \hat{x}'_{j_q})$

$$= T[(\hat{x}_{i_1} \cdot \hat{x}'_{j_1})\hat{x}_{i_1} \cdots (\hat{x}_{i_q} \cdot \hat{x}'_{j_q})\hat{x}_{i_q}]$$

$$= (\hat{x}_{i_1} \cdot \hat{x}'_{j_1}) \cdots (\hat{x}_{i_q} \cdot \hat{x}'_{j_q}) T(\hat{x}_{i_1}, \ldots \hat{x}_{i_q})$$

$$= (\hat{x}_{i_1} \cdot \hat{x}'_{j_1}) \cdots (\hat{x}_{i_q} \cdot \hat{x}'_{j_q}) T_{i_1 \cdots i_q}$$

or 2. $T = T_{i_1 \cdots i_q} \hat{x}_{i_1} \cdots \hat{x}_{i_q} = T'_{j_1 \cdots j_q} \hat{x}'_{j_1} \cdots \hat{x}'_{j_q}$

Now evaluate the above equation at the $\underline{q}$-tuple $\hat{x}'_{j_1}, \ldots \hat{x}'_{j_q}$ to obtain

$$T'_{j_1 \cdots j_q} = (\hat{x}'_{j_1} \cdot \hat{x}_{i_1}) \cdots (\hat{x}'_{j_q} \cdot \hat{x}_{i_q}) T_{i_1 \cdots i_q}$$

Both these methods lead of course to the same result which we state as a

<u>Theorem</u>: If $T_{i_1 \cdots i_q}$ and $T'_{j_1 \cdots j_q}$ are the components of a tensor T w.r.t. two orthonormal bases $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ and $\{\hat{x}'_1, \hat{x}'_2, \hat{x}'_3\}$, then

$$\boxed{T'_{j_1 \cdots j_q} = (\hat{x}'_{j_1} \cdot \hat{x}_{i_1}) \cdots (\hat{x}'_{j_q} \cdot \hat{x}_{i_q}) T_{i_1 \cdots i_q}} \quad (*)$$

The above formula is the starting point of the usual older formulation of tensor algebra.

Section 12: Tensors of order two

Let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ and $\{\hat{x}'_1, \hat{x}'_2, \hat{x}'_3\}$ be two orthonormal bases and define the <u>orthogonal</u> matrix Q by

$$\boxed{\begin{aligned} &Q_{ij} = (\hat{x}_i' \cdot \hat{x}_j) \\ &(\hat{x}_i, \hat{x}_k')(\hat{x}_k', \hat{x}_j) = \delta_{ij}, \text{ or} \\ &Q^T Q = I \end{aligned}}$$

See page 17 of these notes. An orthogonal matrix is the special case of a unitary matrix for <u>real</u> matrices.

Equation (*) can be written as a matrix equation for tensors of order two, and this observation provides us with a reasonably orderly method for carrying out calculations

$$\begin{aligned} T_{ij}' &= (\hat{x}_i' \cdot \hat{x}_k)(\hat{x}_j' \cdot \hat{x}_l) T_{kl} \\ &= (\hat{x}_i' \cdot \hat{x}_k) T_{kl} (\hat{x}_l \cdot \hat{x}_j') \\ &= Q_{ik} T_{kl} Q_{lj}^T \end{aligned}$$

$$\boxed{\begin{aligned} &T_{ij}' = Q_{ik} T_{kl} Q_{lj}^T, \text{ or in matrix notation} \\ &T' = Q T Q^T \end{aligned}}$$

<u>Problem 14</u>: All tensors of order zero may be thought of as isotropic in the sense that the value of a tensor of order zero may be thought of as its component.

Prove that the only isotropic tensor of order 1 is the linear functional generated by the zero vector

Prove that every isotropic tensor of order 2 is a scalar multiple of the identity tensor.

Can be generalized: non-zero isotropic (strongly) tensors always have even order. Non-zero skew-isotropic tensors always have odd order ≥ 3.

Chapter 4 Second Order Tensors and Linear Operators

Tensors of orders zero and one are easy to visualize (scalars and vectors). One of the aims of this chapter is to give a description of second order tensors that will make them almost as easy to visualize. It will be shown that a second order tensor on R^3 can be thought of as a linear operator on R^3, then we have to learn how to visualize linear operators.

Recall the <u>definition</u> of a linear operator on R^3, a function $\phi: R_3 \to R_3$ which assigns to each $\vec{v} \in R^3$ another <u>vector</u> $\phi(\vec{v}) \in R^3$ in such a way that

$$\phi(a\vec{u} + b\vec{v}) = a\phi(\vec{u}) + b\phi(\vec{v})$$

Throughout this chapter <u>all</u> linear operators will be of the type $\phi: R^3 \to R^3$

<u>Examples</u>:

1. $\phi(\vec{v}) = \vec{v}$, called the identity operator and usually denoted by I, $I(\vec{v}) = \vec{v}$
2. $\vec{\Omega}$ a fixed vector, let $\phi(\vec{v}) = \vec{\Omega} \times \vec{v}$
3. $\phi(\underline{v}) = a\underline{v}$

See page 20 for other examples ($\phi: R^3 \to R^3$)

Now let ϕ be any linear operator on R^3. We claim that ϕ generates a unique second order tensor T_ϕ on R^3. For every ordered pair of vectors $\vec{u}, \vec{v}$, we <u>define</u>

$$\boxed{T_\phi(\vec{u}, \vec{v}) = \vec{u} \cdot \phi(\vec{v})} \quad (*)$$

(matrix of linear operator, see p. 20 or any text)

4. given $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$
$\underline{v} = v_i \hat{x}_i$
define $\phi(\underline{v}) = \hat{x}_i A_{ij} v_j$

Lecture 18 review: defn tensor of order q, four arith. operations, tensor components w.r.t. arb. Cart. axis system. Second order tensors and linear operators.

Answer G. Brown's question

Clearly $T(\vec{u},\vec{v})$ is a real number which depends linearly on $\vec{u}$ if $\vec{v}$ is fixed. But if $\vec{u}$ is fixed, $T(\vec{u},\vec{v})$ also depends linearly on $\vec{v}$ since $\phi(\vec{v})$ is a linear operator. **Defn such that ϕ generates a T_ϕ.**

$$T(\vec{u}, a\vec{v}+b\vec{w}) = \vec{u}\cdot\phi(a\vec{v}+b\vec{w}) = \vec{u}\cdot[a\phi(\vec{v})+b\phi(\vec{w})]$$

$$= a[\vec{u}\cdot\phi(\vec{v})] + b[\vec{u}\cdot\phi(\vec{w})] = aT(\vec{u},\vec{v}) + bT(\vec{u},\vec{w})$$

Thus T, defined by $(*)$ is a bilinear functional, i.e., a second order tensor. Denote it by T_ϕ.

made it to here

Every ϕ generates a unique T_ϕ, that is if $\phi=\psi$, then $T_\phi = T_\psi$. It is also easy to show that if $T_\phi = T_\psi$, then $\phi=\psi$. If $T_\phi = T_\psi$ for every ordered pair $\vec{u},\vec{v}$, then $\vec{u}\cdot\phi(\vec{v}) = \vec{u}\cdot\psi(\vec{v})$ for every $\vec{u},\vec{v}$. If we fix $\vec{v}$, this equation is true for every $\vec{u}$, hence $\phi(\vec{v}) = \psi(\vec{v})$ for every $\vec{v}$, hence $\phi=\psi$.

must show mapping between $\mathcal{L}$ and L is 1-1

Thus different linear operators ϕ and ψ generate different tensors T_ϕ and T_ψ. The obvious question is whether every second order tensor is generated by some linear operator. Given an arbitrary second order tensor T, does there exist a ϕ such that $(*)$ holds for all $\vec{u},\vec{v}$. We already know there is no more than one such ϕ. Is there one?

must show mapping is onto.

The answer is yes.

Let $\vec{v}_0$ be a fixed vector. Then $T(\vec{u},\vec{v}_0)$ depends linearly on $\vec{u}$ and is thus a linear functional. Thus $\exists$ a unique vector $\vec{f}_0 \in \mathbb{R}^3$ such that $T(\vec{u},\vec{v}_0) = \vec{u}\cdot\vec{f}_0$

Similarly if $\vec{v}_1$ is another fixed vector, $\exists$ a vector $\vec{f}_1$ such that $T(\vec{u}, \vec{v}_1) = \vec{u} \cdot f_1$. In general for any fixed vector $\vec{v}$, $\exists$ a unique vector $\vec{f}$ such that

$$T(\vec{u}, \vec{v}) = \vec{u} \cdot \vec{f}$$

Since the vector $\vec{f}$ is determined uniquely by $\vec{v}$, we can write $\vec{f} = \phi(\vec{v})$. Thus $\exists$ a function ϕ such that $\forall \vec{u}, \vec{v}$

$$T(\vec{u}, \vec{v}) = \vec{u} \cdot \phi(\vec{v})$$

Is ϕ linear? Yes. $T(\vec{u}, a\vec{v} + b\vec{w}) = aT(\vec{u}, \vec{v}) + bT(\vec{v}, \vec{w})$

But by definition of ϕ

$$T(\vec{u}, a\vec{v} + b\vec{w}) = \vec{u} \cdot \phi(a\vec{v} + b\vec{w})$$
$$T(\vec{u}, \vec{v}) = \vec{u} \cdot \phi(\vec{v})$$
$$T(\vec{u}, \vec{w}) = \vec{u} \cdot \phi(\vec{w})$$

Hence $\vec{u} \cdot \phi(a\vec{v} + b\vec{w}) = \vec{u} \cdot [a\phi(\vec{v}) + b\phi(\vec{w})]$

But this equation is valid for any $\vec{u}$. Hence ϕ is linear

Now we summarize the above results in the following

<u>Theorem</u> — Every linear operator ϕ $(\phi : \mathbb{R}^3 \to \mathbb{R}^3)$ generates a unique second order tensor T_ϕ defined by $(*)$. Two different linear operators generate different tensors. Every second order tensor is generated by exactly one linear operator. In other words $(*)$ defines a one-to-one, onto mapping from linear operators to tensors of order-two.

Is this mapping an isomorphism? In other words if ϕ generates T_ϕ and ψ generates T_ψ does $a\phi + b\psi$ generate $aT_\phi + bT_\psi$? We can't really ask this yet since we haven't yet even defined the linear operator $a\phi + b\psi$.

Definition: The symbol $a\phi + b\psi$ will denote the function which assigns to any vector $\vec{v}$ the vector $a\phi(\vec{v}) + b\psi(\vec{v})$

$$(a\phi + b\psi)(\vec{v}) = a\phi(\vec{v}) + b\psi(\vec{v})$$

Theorem — If a and b are real numbers and ϕ and ψ are linear operators, then $a\phi + b\psi$ is a linear operator and the second order tensor generated by $a\phi + b\psi$ is

$$T_{a\phi + b\psi} = aT_\phi + bT_\psi$$

The proof is the usual trivial exercise and is left to the reader.

The two theorems permit us to confuse linear operators with second order tensors via equation (*). Thus if we wish to be able to visualize a second order tensor T, all we have to do is learn how to visualize linear operators

Section 2 The transpose

Given a linear operator ϕ, we could have defined a tensor S not by the eqn (*) but by

$$S(\vec{u}, \vec{v}) = \phi(\vec{u}) \cdot \vec{v}$$

It turns out that the tensor S is immediately expressible in terms of T. If $T(\vec{u}, \vec{v}) = \vec{u} \cdot \phi(\vec{v})$ mistake

and $S(\vec{u}, \vec{v}) = \phi(\vec{u}) \cdot \vec{v}$, then $S(\vec{v}, \vec{u}) = \phi(\vec{v}) \cdot \vec{u} = T(\vec{v}, \vec{u})$ (u, v)

The tensor S is said to be the transpose of the tensor T (Defn:) usually written T^T

<u>Definition</u> If T is a second order tensor, the <u>transpose</u> of T is that tensor T^T such that for any vectors $\vec{u}, \vec{v} \in R^3$,

$$\boxed{T^T(\vec{u}, \vec{v}) = T(\vec{v}, \vec{u})}$$

It follows immediately that relative to any Cartesian axis system $\boxed{T^T_{ij} = T_{ji}}$

example: if use old defn of tensor, this defn of transpose, must then verify for every coordinate system.

Moreover if S and T are second order tensor such that relative to <u>some</u> Cartesian axis system $S_{ij} = T_{ji}$, then $S = T^T$! One more fact worth noting is that $(T^T)^T = T$

Now returning to argument on the last page, if we define a second order tensor S by

$$S(\vec{u}, \vec{v}) = \phi(\vec{u}) \cdot \vec{v}$$

then we know that $S(\vec{u}, \vec{v})$ generates a linear operator ψ. i.e.

Theorem : if ϕ is any linear operator, then $\exists$ a unique linear operator ψ such that for any $\vec{u}, \vec{v}$

$$\phi(\vec{u}) \cdot \vec{v} = \vec{u} \cdot \psi(\vec{v})$$

This theorem allows us to make the following definition

<u>Definition</u> (transpose of a linear operator) If ϕ is a linear operator, its <u>transpose</u> ϕ^T is that unique linear operator

such that $\boxed{\vec{u}\cdot\phi(\vec{v}) = \phi^T(\vec{u})\cdot\vec{v}}$ (*)

It is clear that $(\phi^T)^T = \phi$

note that $T_{\phi^T} = (T_\phi)^T$

In a complex vector space the analog of the transpose of a linear operator is the adjoint, defined exactly by (*) $(u, \phi(v)) = (\phi^+(u), v)$ and denoted by ϕ^+, ϕ^*, ϕ^A, or ϕ^T. Even in a real vector space ϕ^T is sometimes called the adjoint operator. Now read page 78.5.

If $\phi = \phi^T$, called self-adjoint or Hermitian. Will be recognized by quantum mechanics.

be sure to do these

Problem 15 Find the linear operator corresponding to the following tensors

(i) $T(\vec{u},\vec{v}) = \vec{u}\cdot\vec{v}$ do this one in lecture.

(ii) $T(\vec{u},\vec{v}) = (\vec{u}\times\vec{\Omega})\cdot\vec{v}$, $\vec{\Omega}$ fixed

(iii) $T = \vec{f}\vec{g}$

(iv) $T = T_{ij}\hat{x}_i\hat{x}_j$ $\qquad T(\vec{v}) = \hat{x}_i(T_{ij}v_j)$

Problem 16 Find the transposes or adjoints of the linear operators in problem 15.

Section 3 Operator Products

Definition The product of two not necessarily linear operators is defined by

$$\boxed{\phi\psi(\vec{v}) = \phi[\psi(\vec{v})]}$$

The reader can verify that

Only linear operators can have a transpose.

An afterthought on the transpose.

Suppose $\phi: V \to V$ is an arbitrary (not necessarily linear) operator on a vector space V and suppose that ϕ has an adjoint, i.e. suppose $\exists$ an operator $\phi^{+}: V \to V$ such that

$$(u, \phi(v)) = (\phi^{+}(u), v) \quad (*)$$

Then it can easily be shown that ϕ must be a linear operator, since $(*)$ implies that $(u, \phi(v))$ depends linearly on v as well as on u. Hence $\exists$ a second order tensor T_ϕ such that $(u, \phi(v)) = T_\phi(u, v)$ for all u and v. Thus for all v, $\phi(v)$ is a linear operator on V.

We have thus proved the

Theorem: Suppose $\phi: R^3 \to R^3$ has a transpose. Then ϕ is linear.

(i) if ϕ and ψ are linear, so is $\phi\psi$

(ii) $\phi(\psi\omega) = (\phi\psi)\omega$ (associative law)

$(\psi+\omega)\phi = \psi\phi + \omega\phi$

$(a\phi)\psi = a(\phi\psi)$ for any real a

ϕ, ψ, ω not necessarily linear

(iii) if in addition ϕ is linear, then

$$\phi(\psi+\omega) = \phi\psi + \phi\omega$$

$$\phi(a\psi) = a(\phi\psi)$$

(iv) if ϕ and ψ are linear $(\phi\psi)^T = \psi^T\phi^T$ (note reversed order)

Question: What is the tensor generated by $\phi\psi$? What is $T_{\phi\psi}$? It is not $T_\phi T_\psi$ since that is a fourth order tensor

$$T_{\phi\psi}(\vec{u}, \vec{v}) = \vec{u} \cdot \phi\psi(\vec{v}) = \vec{u} \cdot \phi[\psi(\vec{v})]$$

$$= \phi^T(\vec{u}) \cdot \psi(\vec{v}) = [\phi^T(\vec{u}) \cdot \hat{x}_i][\psi(\vec{v}) \cdot \hat{x}_i]$$ (introduce an arbitrary orthonormal basis)

$$= [\vec{u} \cdot \phi(\hat{x}_i)][\hat{x}_i \cdot \psi(\vec{v})]$$

$$= T_\phi(\vec{u}, \hat{x}_i) T_\psi(\hat{x}_i, \vec{v}) = T_\phi T_\psi(\vec{u}, \hat{x}_i, \hat{x}_i, \vec{v})$$

$$= tr_{23} T_\phi T_\psi(\vec{u}, \vec{v})$$

Thus we conclude that

$$\boxed{T_{\phi\psi} = tr_{23} T_\phi T_\psi}$$

Section 4: Gibbs notation for second order tensors

Equations involving vectors and second order tensors may be written in a compact and easily remembered form by utilizing the notation conventions introduced by J. Willard Gibbs, perhaps America's greatest Yale man. **See his book, <u>Vector Analysis</u> (Dover).**

We denote a vector by a single arrow $\vec{f}$ or by an underline $\underline{f}$.

We denote a second order tensor by a double arrow $\overleftrightarrow{T}$ or by a double underline $\underline{\underline{T}}$. The second order tensor generated by a linear operator T we write as $\overleftrightarrow{T}$ or $\underline{\underline{T}}$.

tensor sans serif **T**
vectors bold

Thus where before we wrote $T_{a\phi+b\psi} = aT_\phi + bT_\psi$, now we merely write

$$\overleftrightarrow{a\phi+b\psi} = a\overleftrightarrow{\phi} + b\overleftrightarrow{\psi}$$

We also write $\overleftrightarrow{\phi^T} = \overleftrightarrow{\phi}^T$

or: tensors cap. bold, vectors l.c. bold

Henceforth we shall write $\overleftrightarrow{T}(\vec{u},\vec{v})$ in the form $\vec{u}\cdot\overleftrightarrow{T}\cdot\vec{v}$, and we shall write $\phi(\vec{v}) = \overleftrightarrow{\phi}\cdot\vec{v}$ and $\phi^T(\vec{u})$ as $\overleftrightarrow{\phi}^T\cdot\vec{u}$ or $\vec{u}\cdot\overleftrightarrow{\phi}$. We thus rewrite $T_\phi(\vec{u},\vec{v}) = \vec{u}\cdot\phi(\vec{v}) = \phi^T(\vec{u})\cdot\vec{v}$ in the form

$$\vec{u}\cdot\overleftrightarrow{\phi}\cdot\vec{v} = \vec{u}\cdot(\overleftrightarrow{\phi}\cdot\vec{v}) = (\vec{u}\cdot\overleftrightarrow{\phi})\cdot\vec{v}$$

and thus in any such expression we may omit the parentheses.

For dyads $\vec{f}\vec{g}$ (this counts as a double arrow) we have

$$\vec{u}\cdot\vec{f}\vec{g}\cdot\vec{v} = (\vec{u}\cdot\vec{f})(\vec{g}\cdot\vec{v})$$

The linear operator which generates $\overrightarrow{fg}$ assigns to vector

$\vec{v}$ the vector $(\vec{f}\vec{g})\cdot\vec{v}$. The notation suggests that this is $\vec{f}(\vec{g}\cdot\vec{v})$. See problem 17.

One more piece of notation, we define

$$\overleftrightarrow{S}\cdot\overleftrightarrow{T} = tr_{23}(\overleftrightarrow{S}\,\overleftrightarrow{T}) \qquad (\underline{\underline{S}}\cdot\underline{\underline{T}})_{ij} = S_{ik}T_{kj}$$

Thus the eqn $T_{\phi\psi} = tr_{23}\, T_\phi T_\psi$ becomes $\overleftrightarrow{\phi\psi} = \overleftrightarrow{\phi}\cdot\overleftrightarrow{\psi}$

↑ not really — this is a fourth order tensor in Gibbs notation

Problem 17 Verify the following expressions

(i) $(\vec{f}\vec{g})\cdot\vec{v} = \vec{f}(\vec{g}\cdot\vec{v})$

(ii) $\vec{u}\cdot(\vec{f}\vec{g}) = (\vec{u}\cdot\vec{f})\vec{g}$

(iii) $(\underline{\underline{S}}\cdot\underline{\underline{T}})\cdot\underline{u} = \underline{\underline{S}}\cdot(\underline{\underline{T}}\cdot\underline{u})$

(iv) $\underline{u}\cdot(\underline{\underline{S}}\cdot\underline{\underline{T}}) = (\underline{u}\cdot\underline{\underline{S}})\cdot\underline{\underline{T}}$

(v) $\underline{\underline{R}}\cdot(\underline{\underline{S}}\cdot\underline{\underline{T}}) = (\underline{\underline{R}}\cdot\underline{\underline{S}})\cdot\underline{\underline{T}}$

(vi) $(\vec{f}\vec{g})\cdot(\vec{h}\vec{k}) = \vec{f}(\vec{g}\cdot\vec{h})\vec{k}$

(vii) $(\underline{\underline{R}}\cdot\underline{\underline{S}})^T = (\underline{\underline{S}}^T\cdot\underline{\underline{R}}^T)$

(viii) $\underline{\underline{R}}\cdot(a\underline{\underline{S}} + b\underline{\underline{T}}) = a\underline{\underline{R}}\cdot\underline{\underline{S}} + b\underline{\underline{R}}\cdot\underline{\underline{T}}$

(ix) $(a\underline{\underline{S}} + b\underline{\underline{T}})\cdot\underline{\underline{R}} = a\underline{\underline{S}}\cdot\underline{\underline{R}} + b\underline{\underline{T}}\cdot\underline{\underline{R}}$

(x) $(a\underline{\underline{R}} + b\underline{\underline{S}})\cdot\underline{v} = a\underline{\underline{R}}\cdot\underline{v} + b\underline{\underline{S}}\cdot\underline{v}$

(xi) $\underline{v}\cdot(a\underline{\underline{R}} + b\underline{\underline{S}}) = a\underline{v}\cdot\underline{\underline{R}} + b\underline{v}\cdot\underline{\underline{S}}$

(xii) $\underline{\underline{R}}\cdot(a\underline{u} + b\underline{v}) = a\underline{\underline{R}}\cdot\underline{u} + b\underline{\underline{R}}\cdot\underline{v}$

(xiii) $(a\underline{u} + b\underline{v})\cdot\underline{\underline{R}} = a\underline{u}\cdot\underline{\underline{R}} + b\underline{v}\cdot\underline{\underline{R}}$

The $\cdot$ product is called the Gibbs dot product. It is defined $A\cdot B$ when A, B are either vectors or second order tensors. If both A and B are vectors $\vec{A}, \vec{B}$, then $\vec{A}\cdot\vec{B}$ is the scalar product or inner product. The Gibbs dot product is completely associative and distributive with vector and tensor addition. It is not commutative

$\underline{\underline{A}}\cdot\underline{\underline{B}} \neq \underline{\underline{B}}\cdot\underline{\underline{A}}$, in general.

One more notation is useful: double dot

examples back of page 83

Section 5: Components of Linear Operators

$$\underline{\underline{A}} : \underline{\underline{B}} = tr(\underline{\underline{A}}\cdot\underline{\underline{B}}^T) = tr(\underline{\underline{A}}^T\cdot\underline{\underline{B}}) = tr_{13}tr_{24}\underline{\underline{A}}\,\underline{\underline{B}}$$

Definition The components of a linear operator T w.r.t. a Cartesian axis system $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ are defined to be the components of the tensor $\overleftrightarrow{T}$ (generated by T) relative to that system. We denote them T_{ij}.

Malvern also defines

$$\underline{\underline{A}} \cdot\cdot\, \underline{\underline{B}} = tr(\underline{\underline{A}}\cdot\underline{\underline{B}}) = tr_{14}tr_{23}\underline{\underline{A}}\,\underline{\underline{B}}$$

$$T_{ij} = \overleftrightarrow{T}(\hat{x}_i, \hat{x}_j) = \hat{x}_i \cdot \overleftrightarrow{T} \cdot \hat{x}_j$$

Now $\overleftrightarrow{T} = T_{ij}\hat{x}_i\hat{x}_j$ while $T(\hat{x}_k) = \overleftrightarrow{T}\cdot\hat{x}_k =$ $(T_{ij}\hat{x}_i\hat{x}_j)\cdot\hat{x}_k = T_{ij}\hat{x}_i(\hat{x}_j\cdot\hat{x}_k) = T_{ij}\hat{x}_i\delta_{jk} = \hat{x}_i T_{ik}$, so

$$\boxed{T(\hat{x}_j) = \hat{x}_i T_{ij}}$$

$$\underline{\underline{A}} : \underline{\underline{B}} = A_{ij}B_{ij}$$

$$\underline{\underline{A}} \cdot\cdot\, \underline{\underline{B}} = A_{ij}B_{ji}$$

Definition: Suppose T is a linear operator with components T_{ij} relative $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$. The 3x3 matrix

$$\begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix}$$ is called the matrix of T relative to $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$

and will be denoted $\mathcal{M}(\hat{x}_1, \hat{x}_2, \hat{x}_3; T)$ or sometimes when there is no chance of confusion, merely by the same symbol T.

Every Cartesian axis system thus produces or generates a one-to-one correspondence between all 3x3 matrices and all linear operators. We see later that this correspondence preserves all the rules of arithmetic so it is an isomorphism.

Suppose we know T_{ij} for some linear operator T. Then we can easily compute $T(\vec{v})$ for any $\vec{v}$.

$$T(\vec{v}) = T(v_j \hat{x}_j) = v_j T(\hat{x}_j) = \hat{x}_i v_j T_{ij}$$

If $\vec{y} = T(\vec{v})$, then $y_i = T_{ij} v_j$, or

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$, or in matrix notation $y = Tv$

The above is true as long as all matrices are computed relative to the same axis system

Section 6: The correspondence between matrix algebra and operator algebra.

Write $M(\hat{x}_1, \hat{x}_2, \hat{x}_3; T)$ as $M(T)$ for short. Then $[M(T)]_{ij} = T_{ij}$ if T_{ij} are the components of the tensor $\overleftrightarrow{T}$ generated by the linear operator T. It is straightforward to show that the rules of operator algebra are the same as the rules of matrix algebra.

1. $M(aS + bT) = aM(S) + bM(T)$
2. $M(ST) = M(S)M(T) \leftarrow$ (matrix multiplication)
3. $M(I) = I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, the identity matrix
4. $M(S^T) = [M(S)]^T$, the matrix transpose of $M(S)$

<u>Problem 18</u>. Prove 1. thru 4. above. The easiest way to prove these is to note that if $\tilde{T}$ is any matrix and T is a linear operator such that $T(\hat{x}_j) = \tilde{T}_{ij}\hat{x}_i$, then $\tilde{T} = \mathcal{M}(T)$.

Lecture 19 Review here
1-1 → 2 ord. tensors, lin op.
each (x,y,z) → 1-1 corr. lin op matrices
Give physical examples here, many are linear phenomenological laws

<u>Section 7: Inverse Operators</u>

<u>Definition</u>: An operator $S: R^3 \to R^3$ is called <u>invertible</u> if $\forall\, \vec{v} \in R^3\ \exists$ exactly one vector $\vec{w}$ such that $S(\vec{w}) = \vec{v}$. If to each $\vec{v}$ we assign the unique $\vec{w}$ such that $S(\vec{w}) = \vec{v}$, then we have a rule for assigning vectors $\vec{w}$ to vectors $\vec{v}$. This rule is called the <u>inverse</u> mapping of S and is written S^{-1}. Must be both 1-1 and onto

If S is <u>invertible</u> then $S^{-1}(\vec{v}) = \vec{w}$ if and only if $S(\vec{w}) = \vec{v}$

1-1 iff rt. inverse
onto iff left inverse

1. if S is invertible, $SS^{-1} = S^{-1}S = I$
For $SS^{-1}(\vec{v}) = S[S^{-1}(\vec{v})] = S(\vec{w}) = \vec{v} = I(\vec{v})$,
and $S^{-1}S(\vec{w}) = S^{-1}(\vec{v}) = \vec{w} = I(\vec{w})$

2. If S is an operator and $\exists$ an operator T such that $TS = ST = I$, then S is invertible and $T = S^{-1}$.
Let $\vec{w} = T(\vec{v})$. Then $S(\vec{w}) = ST(\vec{v}) = I(\vec{v}) = \vec{v}$, so
onto → there is at least one $\vec{w}$ such that $S(\vec{w}) = \vec{v}$. But there is only one since if $\vec{v} = S(\vec{w}_1) = S(\vec{w}_2)$, then $TS(\vec{w}_1) = TS(\vec{w}_2)$, so $I(\vec{w}_1) = I(\vec{w}_2)$, so $\vec{w}_1 = \vec{w}_2$.

3. If T and S are linear operators, only one of $TS = I$, $ST = I$ is needed. Consider

$TS + T - I$

4. If S is invertible, so is S^{-1} and $(S^{-1})^{-1} = S$

5. If S and T are invertible, so is ST and $(ST)^{-1} = T^{-1}S^{-1}$

$$(T^{-1}S^{-1})(ST) = T^{-1}(S^{-1}S)T = T^{-1}IT = T^{-1}T = I$$

6. If S is invertible and linear, S^{-1} is linear

$$S[aS^{-1}(\vec{u}) + bS^{-1}(\vec{v})] = aS[S^{-1}(\vec{u})] + bS[S^{-1}(\vec{v})]$$

$= a\vec{u} + b\vec{v}$, hence $S^{-1}(a\vec{u} + b\vec{v}) = aS^{-1}(\vec{u}) + bS^{-1}(\vec{v})$

7. transposes of inverses. Suppose $T = S^{-1}$. Then $TS = ST = I$, Clearly $I^T = I$ so $S^T T^T = T^T S^T = I$

so $B^T = (A^T)^{-1}$ or $\qquad T^T = (S^T)^{-1}$ or

$$(A^{-1})^T = (A^T)^{-1} \qquad (S^{-1})^T = (S^T)^{-1}$$

8. S^{-1} exists if and only if $[M(S)]^{-1}$ exists, and $M(S^{-1}) = [M(S)]^{-1}$

(Recall that a 3x3 matrix M is <u>non-singular</u> (or invertible) if $\exists$ a matrix N such that $MN = NM = I$, and that the matrix N is called the inverse matrix to M, written M^{-1})

Thus a linear operator S is <u>invertible</u> iff its matrix $M(S)$ w.r.t. <u>any</u> $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ is invertible.

Maybe give some physical examples here: stress tensor, inertia tensor, thermal cond. tensor, dielectric tensor

Section 8 : Determinants

1. Determinants of Matrices ($n \times n$), without proofs.

Definition The *determinant* of an $n \times n$ matrix M may be defined by

$$\det M = \epsilon^{i_1 \ldots i_n} M_{i_1 1} \ldots M_{i_n n}$$

where $\epsilon^{i_1 \ldots i_n}$ is the alternating symbol in n dimensions, defined by

$\epsilon_{i_1 \ldots i_n} = 1$ if $(i_1 \ldots i_n)$ is an even permutation of $(1 \ldots n)$

$\epsilon_{i_1 \ldots i_n} = -1$ if $(i_1 \ldots i_n)$ is an odd permutation of $(1 \ldots n)$

$\epsilon_{i_1 \ldots i_n} = 0$ if $i_1, \ldots, i_n$ are not all different

Recall that: $\det M^T = \det M$

$$\det(MN) = (\det M)(\det N)$$

since $\det I = 1$, $\det M^{-1} = \dfrac{1}{\det M}$

Finally recall that M is non-singular iff $\det M \neq 0$. If $\det M \neq 0$, then

$$(M^{-1})_{ij} = \frac{1}{\det M}(\Gamma_{ji})$$

where the *cofactor* Γ_{ij} is

$\Gamma_{ij} = (-1)^{i+j} \det M_{(ij)}$ where $M_{(ij)}$ is the $(n-1)\times(n-1)$ matrix obtained by deleting the i^{th} row and j^{th} column of M

2. Determinants of second order tensors:

Let $\overleftrightarrow{M}$ be a second order tensor over R^3 and let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ and $\{\hat{x}_1', \hat{x}_2', \hat{x}_3'\}$ be two Cartesian axis systems. Relative to the two systems, the components of M are

$M_{ij} = \hat{x}_i \cdot \overleftrightarrow{M} \cdot \hat{x}_j$ and $M_{ij}' = \hat{x}_i' \cdot \overleftrightarrow{M} \cdot \hat{x}_j$. Now

$$M_{ij}' = (\hat{x}_i' \cdot \hat{x}_k)\hat{x}_k \cdot \overleftrightarrow{M} \cdot \hat{x}_\ell (\hat{x}_\ell \cdot x_j')$$

$$= (x_i' \cdot x_k) M_{k\ell} (\hat{x}_\ell \cdot x_j')$$

Define the matrix U by $U_{ij} = (\hat{x}_i' \cdot \hat{x}_j)$. Then U is an <u>orthogonal</u> matrix, $U^T = U^{-1}$, or $(U^{-1})_{ij} = \hat{x}_i \cdot \hat{x}_j'$, and in matrix notation, we have

$M' = UMU^T$, or $M' = UMU^{-1}$. Thus

$$\det M' = \det(UMU^{-1}) = (\det U)(\det M)(\det U^{-1})$$

$$= (\det U)(\det M)\left(\frac{1}{\det U}\right).$$ Thus

$$\boxed{\det M' = \det M}$$

The determinant of the component matrix M_{ij} of a second order tensor $\overleftrightarrow{M}$ is independent of the axis system used to calculate. It thus makes sense to

<u>Define</u>: The determinant of a second order tensor $\overleftrightarrow{M}$ is the determinant of its component matrix in <u>any</u> coordinate system.

We can also show that

$$\det(\overleftrightarrow{M}\cdot\overleftrightarrow{N}) = \det\overleftrightarrow{M}\,\det\overleftrightarrow{N}$$

since the respective component matrices satisfy

$$(\overleftrightarrow{M}\cdot\overleftrightarrow{N})_{ij} = \overleftrightarrow{M}_{ik}\overleftrightarrow{N}_{kj}$$

3. Determinants of linear operators

Definition: $\boxed{\det M = \det\overleftrightarrow{M}}$

The above equation is equivalent to $\det MN = \det M \det N$

Also $\det I = 1$ if I is the identity operator, so if M^{-1} exists, then $\det M^{-1} = \det M$

We also have $\det M^{T} = \det M$

Problem 18: A second order tensor is called antisymmetric if $A = -A^{T}$. Show that the determinant of an antisymmetric second order tensor over R^3 is zero. In fact prove that the determinant of an antisymmetric tensor over any n-dimensional vector space is zero if n is odd.

Section 9 : Eigenvalues and Eigenvectors

1. Eigenvalues and columns for matrices

Let A be an $n \times n$ matrix, and v be an $n \times 1$ column matrix of complex numbers, and λ be a complex number. Then v is an <u>eigencolumn</u> of A corresponding to <u>eigenvalue</u> λ if

(1) $v \neq 0$ i.e. $v \neq \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$

(2) $Av = \lambda v$, i.e. $\begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \\ A_{n1} & \cdots & A_{nn} \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \lambda \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$

<u>Theorem (prove it yourself)</u>

The eigenvalues of λ are the roots of the polynomial equation $\det(A - \lambda I) = 0$.

The polynomial $\det(A - \lambda I)$ is a poly of degree n in λ, the highest degree term is $(-1)^n \lambda^n$

<u>Definition</u>. The polynomial $\det(A - \lambda I)$ is called the <u>characteristic polynomial</u> of the $n \times n$ matrix A.

If A is a 3×3 matrix, the characteristic polynomial takes the form

$$\det(A - \lambda I) = -\lambda^3 + (A_{11} + A_{22} + A_{33})\lambda^2$$

$$- \left(\begin{vmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{vmatrix} + \begin{vmatrix} A_{11} & A_{13} \\ A_{31} & A_{33} \end{vmatrix} + \begin{vmatrix} A_{22} & A_{23} \\ A_{32} & A_{33} \end{vmatrix} \right) + \det A$$

Problem 19 Find the characteristic polynomial, eigenvalues, and eigencolumns of $\begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$

Show thereby that not all real matrices have even one real eigenvalue.

Problem 20 Show that if v_1 and v_2 are eigencolumns corresponding to the same eigenvalue, then so is $av_1 + bv_2$.

Note that in problem 19, $n=2$. It is clear that if n is odd, then any $n \times n$ real matrix must have one (at least) real eigenvalue.

The so-called fundamental theorem of algebra states that every real or complex polynomial in λ has at least one complex root. Hence any real or complex $n \times n$ matrix has at least one complex eigenvalue, whatever the value of n.
Has n if count repeated roots

Definition: A matrix is Hermitian if $A_{ij} = A_{ji}^*$ or rewritten in matrix notation, if $A = (A^T)^*$.

A real Hermitian matrix is symmetric

Theorem: Every eigenvalue of a Hermitian matrix A is real.

Problem 21 Prove the above theorem

2. Eigenvalues and Eigenvectors for Linear Operators

Definition: A vector $\vec{v}$ is an eigenvector of a linear operator T belonging to eigenvalue λ if

1. $\vec{v} \neq \vec{0}$, and
2. $T \cdot \vec{v} = \lambda \vec{v}$ or $T(\vec{v}) = \lambda \vec{v}$

Note that if $\vec{v}$ is an eigenvector, so is $a\vec{v}$. Hence if λ is eigenvalue $\exists$ a vector $\hat{n}$ of <u>unit length</u> such that $T(\hat{n}) = \lambda\hat{n}$.

Now let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ be a Cartesian axis system, let

$v = v_i\hat{x}_i$

Now $T(\vec{v}) = \hat{x}_i T_{ij} v_j$, or $[T(\vec{v})]_i = T_{ij} v_j$

Thus the condition that $\vec{v}$ be an eigenvector may be written

$T_{ij} v_j = \lambda v_i$ or $(T_{ij} - \lambda\delta_{ij}) v_j = 0$

Thus we have the

<u>Theorem</u>: Let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ be a Cartesian axis system. Then $\vec{v} = v_i\hat{x}_i$ is an eigenvector of T with eigenvalue λ if and only if $\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$ is an eigencolumn of the matrix $M(\hat{x}_1, \hat{x}_2, \hat{x}_3; T)$ with eigenvalue λ

In practice finding eigenvalues and eigenvectors of linear operators is carried out by computing eigenvalues and eigenvectors of their matrices

<u>Example</u>: Let $T(\vec{v}) = \vec{\Omega} \times \vec{v}$, $\vec{\Omega}$ a fixed vector

Method 1: If $\vec{\Omega} \times \vec{u} = \lambda\vec{u}$, then $\lambda\vec{u}\cdot\vec{u} = 0$, so if $\vec{u} \neq 0$, $\lambda = 0$. Thus $\lambda = 0$ is the only eigenvalue. If $\vec{\Omega} \times \vec{u} = 0$, then $\vec{u} = a\vec{\Omega}$. The eigenvectors of T are the scalar multiples of $\vec{\Omega}$.

Method 2: Any Cartesian axis system can be used for the computation. Choose one in which $\vec{\Omega} = \Omega\hat{x}_3$. Then $T_{ij} = \hat{x}_i \cdot (\Omega\hat{x}_3 \times \hat{x}_j)$, so $T_{12} = -\Omega$, $T_{21} = \Omega$, $T_{ij} = 0$ otherwise

$$\det \begin{vmatrix} -\lambda & -\Omega & 0 \\ \Omega & -\lambda & 0 \\ 0 & 0 & -\lambda \end{vmatrix} = -\lambda^3 - \lambda\Omega^2 = -\lambda(\lambda^2+\Omega^2).$$ The eigenvalues are $\lambda = 0$, $\lambda = \pm i\Omega$. The only real eigenvalue is $\lambda = 0$. The corresponding eigencolumns are

$$\begin{bmatrix} 0 & -\Omega & 0 \\ \Omega & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 0 \quad \text{or} \quad \begin{matrix} -\Omega v_1 = 0 \\ \Omega v_2 = 0 \\ 0 v_3 = 0 \end{matrix}$$, i.e. the eigenvectors are

$a\hat{x}_3$, i.e., the non-zero multiples of $\vec{\Omega}$.

3. The characteristic polynomial of a linear operator

Let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ and $\{\hat{x}_1', \hat{x}_2', \hat{x}_3'\}$ be two Cartesian axis systems, and let T and T' be the matrices of a linear operator T relative to these two Cartesian axis systems.

Define a matrix U by $U_{ij} = (\hat{x}_i' \cdot \hat{x}_j)$. Then

$T_{ij}' = (\hat{x}_i' \cdot \hat{x}_k)(\hat{x}_j' \cdot x_l) T_{kl}$, or $T' = UTU^T$, where $U^{-1} = U^T$ or $UU^T = I$. Thus $T' = UTU^{-1}$. Thus

$T' - \lambda I = U(T - \lambda I)U^{-1}$, so

$$\det(T' - \lambda I) = (\det U)(\det(T - \lambda I)) \det U^{-1}$$
$$= \det(T - \lambda I)$$

The characteristic polynomial of the matrices T and T' are the same.

Therefore we can define

<u>Definition</u>: The characteristic polynomial of a linear operator T is the characteristic polynomial of the matrix M(T) relative to any Cartesian axis system.

The characteristic polynomial is often written

$$-\lambda^3 + T_{I}\lambda^2 + T_{II}\lambda + T_{III} \quad \text{where}$$

$$T_{I} = T_{11} + T_{22} + T_{33} = \operatorname{tr} \overleftrightarrow{T}$$

note

$$-T_{II} = \begin{vmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{vmatrix} + \begin{vmatrix} T_{11} & T_{13} \\ T_{31} & T_{33} \end{vmatrix} + \begin{vmatrix} T_{22} & T_{23} \\ T_{32} & T_{33} \end{vmatrix}$$

$T_{III} = \det T$ These are the three <u>principal</u> invariants.

We already knew that $\operatorname{tr}\overleftrightarrow{T}$ and $\det T$ were invariant under a change of orthonormal basis. The coefficients T_{I}, T_{II}, T_{III} are called the rotational invariants of the linear operator T.

Knowledge of the rotational invariants of an operator does <u>not</u> completely determine the operator, e.g.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

note

$\operatorname{tr}T, \operatorname{tr}T^2, \operatorname{tr}T^3$ also can serve as invariants. The above three can be written in terms of these or vice-versa.

Lecture 20 review here

Chapter 5 — Some Particular Kinds of Operators

1. Preliminary — the wedge operator $\wedge$

The wedge operator $\wedge : R^3 \otimes R^3 \to R^3$ is a linear operator which acts on a second order tensor and converts it into a vector in R^3.

<u>Definition</u>: Let A be the third order alternating tensor and for any second order tensor $\overleftrightarrow{T} \in R^3 \otimes R^3$, define

$$\wedge \overleftrightarrow{T} = tr_{23}\left[tr_{35}\left(\overleftrightarrow{A} \otimes \overleftrightarrow{M} \right) \right]$$

Effect on a dyad — let $\vec{f}$ and $\vec{g}$ be any two vectors. What is $\wedge \vec{f}\vec{g}$? To answer this, we view $\wedge \vec{f}\vec{g}$ as a linear functional

$$\wedge \vec{f}\vec{g}\,(\vec{u}) = A(\vec{u}, \hat{x}_i, \hat{x}_j)\left[\vec{f}\vec{g}\,(\hat{x}_i, \hat{x}_j) \right]$$

$$= A(\vec{u}, \hat{x}_i, \hat{x}_j)(\vec{f}\cdot\hat{x}_i)(\vec{g}\cdot\hat{x}_j) = A\left(\vec{u}, \hat{x}_i(\vec{f}\cdot\hat{x}_i), \hat{x}_j(\vec{g}\cdot\hat{x}_j)\right)$$

$$= A(\vec{u}, \vec{f}, \vec{g}) = \vec{u}\cdot(\vec{f}\times\vec{g})$$

$$\boxed{\text{Hence } \wedge \vec{f}\vec{g} = \vec{f}\times\vec{g}, \text{ the cross product of } \vec{f} \text{ and } \vec{g}}$$

Components of $\wedge \overleftrightarrow{T}$ in an arbitrary axis system $\{\hat{x}_i, \hat{x}_j, \hat{x}_k\}$

$$\boxed{(\wedge \overleftrightarrow{T})_i = A_{ijk} T_{jk}}$$

The wedge operation can be considered a generalization of the cross product (sometimes called the wedge product and written $\vec{f} \wedge \vec{g}$).

Armed with the wedge operator, we are now prepared to discuss the first of three characteristic types of operators, namely antisymmetric operators

<u>Definition</u>: A linear operator A is called <u>antisymmetric</u> if $A = -A^T$. A linear operator S is called <u>symmetric</u> if $S = S^T$.

<u>Problem 20</u>: Show that any linear operator $M: \mathbb{R}^3 \to \mathbb{R}^3$ may be written as the sum of a symmetric linear operator and an antisymmetric linear operator, and show furthermore that this decomposition is unique.

2. Antisymmetric operators and antisymmetric tensors

Suppose Ω is an antisymmetric operator $\Omega: \mathbb{R}^3 \to \mathbb{R}^3$ and $\overleftrightarrow{\Omega}$ is the corresponding antisymmetric tensor.

Denote the vector $-\frac{1}{2} \wedge \overleftrightarrow{\Omega}$ by $\vec{\Omega}$

$$\vec{\Omega} = -\frac{1}{2} \wedge \overleftrightarrow{\Omega} \quad \text{, thus}$$

$$\Omega_i = -\frac{1}{2} A_{ijk} \Omega_{jk} = \mp \frac{1}{2} \epsilon_{ijk} \Omega_{jk}$$

$-$ if $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ is right handed, $+$ if it is left handed

Now relative to a right handed Cartesian axis system $A_{ijk} = \epsilon_{ijk}$, and so

$$\Omega_i \epsilon_{ilm} = \epsilon_{ilm} \left(-\frac{1}{2} \epsilon_{ijk} \Omega_{jk}\right) = -\frac{1}{2} \epsilon_{ilm} \epsilon_{ijk} \Omega_{jk}$$

$$= -\frac{1}{2} \left(\delta_{lj} \delta_{mk} - \delta_{lk} \delta_{mj}\right) \Omega_{jk} = -\frac{1}{2} \left(\Omega_{lm} - \Omega_{ml}\right)$$

$$= -\Omega_{lm}, \quad \text{since } \Omega_{lm} = -\Omega_{ml}$$

Thus if $\overleftrightarrow{\Omega}$ is an antisymmetric tensor and $\vec{\Omega}$ is defined as above, then

$$\boxed{\overleftrightarrow{\Omega} = -\vec{\Omega}\cdot\overleftrightarrow{A} \text{ , or in components} \quad \Omega_{jk} = -\Omega_i A_{ijk}} \quad (*)$$

The converse is also easily seen to be true. If $\vec{\Omega}$ is any vector, then $(*)$ above defines an antisymmetric second order tensor. This is called (Hodge) duality

It follows that there is a one-to-one correspondence between antisymmetric tensors and vectors in $\mathbb{R}^3$. Obviously this correspondence is linear.

Therefore antisymmetric tensors are relatively easy to visualize. They can be thought of as vectors.

What is $\overleftrightarrow{\Omega}\cdot\vec{v}$ (or $\Omega(\vec{v})$) in terms of the vector $\vec{\Omega}$.

From $(*)$

$$\overleftrightarrow{\Omega}\cdot\vec{v} = -\vec{\Omega}\cdot\overleftrightarrow{A}\cdot\vec{v} \text{ , so}$$

$$(\overleftrightarrow{\Omega}\cdot\vec{v})_j = -\Omega_i A_{ijk} v_k = -A_{ijk}\Omega_i v_k = +A_{jik}\Omega_i v_k$$
$$= (\vec{\Omega}\times\vec{v})_j$$

$$\boxed{\overleftrightarrow{\Omega}\cdot\vec{v} = \Omega(\vec{v}) = \vec{\Omega}\times\vec{v}} \quad (**)$$

This equation shows explicitly how an arbitrary antisymmetric operator behaves.

Note also that

$$\vec{v}\cdot\overleftrightarrow{\Omega} = \overleftrightarrow{\Omega}^T\cdot\vec{v} = -\overleftrightarrow{\Omega}\cdot\vec{v} = -\vec{\Omega}\times\vec{v}$$

$$\boxed{\vec{v}\cdot\overleftrightarrow{\Omega} = -\vec{\Omega}\times\vec{v} = \vec{v}\times\vec{\Omega}}$$

What is the matrix $M(\hat{x}_1, \hat{x}_2, \hat{x}_3; \Omega)$ corresponding to the linear operator Ω? From (**)

$$\begin{bmatrix} \Omega_{11} & \Omega_{12} & \Omega_{13} \\ \Omega_{21} & \Omega_{22} & \Omega_{23} \\ \Omega_{31} & \Omega_{32} & \Omega_{33} \end{bmatrix} = \pm \begin{bmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{bmatrix}$$

where the + sign is taken if $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ is right handed, and the − sign if it is left-handed

Thus

$$\begin{bmatrix} \Omega(\vec{v})_1 \\ \Omega(\vec{v})_2 \\ \Omega(\vec{v})_3 \end{bmatrix} = \pm \begin{bmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \pm \begin{bmatrix} \Omega_2 v_3 - \Omega_3 v \\ \Omega_3 v_1 - \Omega_1 v_3 \\ \Omega_1 v_2 - \Omega_2 v_1 \end{bmatrix}$$

We now proceed to examine the next most easily visualized type of second order tensor or linear operator.

3. Orthogonal operators

<u>Definition</u>: An orthogonal linear operator is a linear operator U such that $U^T U = U U^T = I$, i.e. $U^{-1} = U^T$.

If U is orthogonal $UU^T = I$, so $\det(UU^T) = \det I = 1$, thus $(\det U)(\det U^T) = 1$, but $\det U^T = \det U$. So we have

<u>If U is orthogonal</u>, <u>$\det U = \pm 1$</u>.

Define: A $\left\{\begin{matrix}\text{proper}\\ \text{improper}\end{matrix}\right\}$ orthogonal operator is one for which

$\det U = \pm 1. \left(\left\{\begin{matrix}+1\\ -1\end{matrix}\right\}\right)$. A matrix U is called a

~~[illegible]~~ $\left\{\begin{matrix}\text{proper}\\ \text{improper}\end{matrix}\right\}$ orthogonal matrix if it is the matrix of a

$\left\{\begin{matrix}\text{proper}\\ \text{improper}\end{matrix}\right\}$ orthogonal operator.

To verify that an operator is orthogonal, one verifies that its matrix relative to any Cartesian axis system is orthogonal. A matrix U is orthogonal if and only if $UU^T = U^TU = I$. Actually it is only necessary to verify one of these conditions, because if $U^TU = I$, then $UU^T = I$ for any square matrix whatsoever.

To see this: If $\exists$ A, B such that $AB = I$, then A is invertible and $A^{-1}B$, i.e. $BA = I$ also.

If $AB = I$, then $(\det A)(\det B) = 1$, so $\det A \neq 0$. Thus A^{-1} exists, and $AA^{-1} = A^{-1}A = I$. But if $AB = I$, then $A^{-1} = A^{-1}I = A^{-1}(AB) = (A^{-1}A)B = B$. Hence not only $AB = I$, but also $BA = I$.

Thus either of the two equations $UU^T = U^TU = I$ implies the other. In terms of components, the conditions for orthogonality are

$$\sum_{j=1}^{3} U_{ij}U_{kj} = \sum_{j=1}^{3} U_{ji}U_{jk} = \delta_{ik}.$$

An equivalent specification is this. If $U = \begin{pmatrix} U_{11} & U_{12} & U_{13} \\ U_{21} & U_{22} & U_{23} \\ U_{31} & U_{32} & U_{33} \end{pmatrix}$

is a matrix, then U is orthogonal if its rows are "orthogonal" and of "unit length" or if its "columns are "orthogonal" and of unit length.

<u>Problem 21:</u> Is $\begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{6}} \end{pmatrix}$ a proper orthogonal matrix.

Verify your answer in four ways, $U^TU = I$, $UU^T = I$, rows orthogonal & of unit length, same for columns.

<u>Problem 22</u>: Let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$, $\{\hat{x}_1', \hat{x}_2', \hat{x}_3'\}$ be Cartesian axis systems, and let $U_{ij} = \hat{x}_i' \cdot \hat{x}_j$. Show that U is orthogonal.

4. Rigid Rotations

Lecture 21 review defn of orthogonal operator

<u>Definition</u>: An operator R (not necessarily linear) is a <u>rigid rotation about $\vec{0}$</u> if $\forall \vec{u}$ and $\vec{v}$, the length $|R(\vec{u}) - R(\vec{v})|$ is the same as the length $|\vec{u} - \vec{v}|$, and if $R(\vec{0}) = \vec{0}$;

$$\boxed{\begin{aligned} &|R(\vec{u}) - R(\vec{v})| = |\vec{u} - \vec{v}| \quad \forall \vec{u}, \vec{v} \in R^3 \\ &R(\vec{0}) = \vec{0} \end{aligned}}$$

The reason for the name is obvious.

1. An operator R is a rigid rotation about $\vec{0}$ if and only if $R(\vec{u}) \cdot R(\vec{v}) = \vec{u} \cdot \vec{v} \quad \forall \vec{u}, \vec{v}$.

Proof: If $R(\vec{u}) \cdot R(\vec{v}) = \vec{u} \cdot \vec{v}$, then $R(\vec{0}) = \vec{0}$, and $[R(\vec{u}) - R(\vec{v})] \cdot [R(\vec{u}) - R(\vec{v})] = R(\vec{u}) \cdot R(\vec{u}) - 2R(\vec{u}) \cdot R(\vec{v}) + R(\vec{v}) \cdot R(\vec{v}) = \vec{u} \cdot \vec{u} - 2\vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{v} = (\vec{u} - \vec{v}) \cdot (\vec{u} - \vec{v})$

If, on the other hand, R is a rigid rotation about then $R(\vec{u}) = R(\vec{u}) - \vec{0} = R(\vec{u}) - R(\vec{0})$, and

$|R(\vec{u}) - R(\vec{0})| = |\vec{u} - \vec{0}| = |\vec{u}|$. Thus $R(\vec{u}) \cdot R(\vec{u}) = \vec{u} \cdot \vec{u}$,

Now $[R(\vec{u}) - R(\vec{v})] \cdot [R(\vec{u}) - R(\vec{v})] = (\vec{u} - \vec{v}) \cdot (\vec{u} - \vec{v})$, so

$R(\vec{u}) \cdot R(\vec{u}) - 2R(\vec{u}) \cdot R(\vec{v}) + R(\vec{v}) \cdot R(\vec{v}) = \vec{u} \cdot \vec{u} - 2\vec{u} \cdot \vec{v} + \vec{v} \cdot \vec{v}$.

Then since $R(\vec{u}) \cdot R(\vec{u}) = \vec{u} \cdot \vec{u}$, and $R(\vec{v}) \cdot R(\vec{v}) = \vec{v} \cdot \vec{v}$, we have $R(\vec{u}) \cdot R(\vec{v}) = \vec{u} \cdot \vec{v}$.

2. If $R : \mathbb{R}^3 \to \mathbb{R}^3$ is a rigid rotation about $\vec{0}$, then R is linear.

Proof: Let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ be a Cartesian axis system. Let $\hat{x}_i' = R(\hat{x}_i)$. Then $\hat{x}_i' \cdot \hat{x}_j' = R(\hat{x}_i) \cdot R(\hat{x}_j) = \hat{x}_i \cdot \hat{x}_j = \delta_{ij}$.

Thus $\{\hat{x}_1', \hat{x}_2', \hat{x}_3'\}$ is another orthonormal basis. Now consider

$\hat{x}_i' \cdot R(a\vec{u} + b\vec{v}) = R(\hat{x}_i) \cdot R(a\vec{u} + b\vec{v}) \overset{\text{using 1.}}{=} \hat{x}_i \cdot (a\vec{u} + b\vec{v})$

$= a(\hat{x}_i \cdot \vec{u}) + b(\hat{x}_i \cdot \vec{v}) = a[R(\hat{x}_i) \cdot R(\vec{u})] + b[R(\hat{x}_i) \cdot R(\vec{v})]$

$= \hat{x}_i' \cdot [aR(\vec{u}) + bR(\vec{v})]$. Thus relative to the basis $\{\hat{x}_i'\}$ the two vectors $R(a\vec{u} + b\vec{v})$ and $aR(\vec{u}) + bR(\vec{v})$ have the same components. Hence they are the same vector.

3. $R : \mathbb{R}^3 \to \mathbb{R}^3$ is a rigid rotation about zero if and only if $R^T = R^{-1}$! before 2. we didn't even know R^T existed

First suppose R is a rigid rotation. Then R is linear and $\vec{u} \cdot \vec{v} = \vec{u} \cdot I(\vec{v}) = R(\vec{u}) \cdot R(\vec{v}) = \vec{u} \cdot R^T R(\vec{v})$

so $\overleftrightarrow{R^T R} = \overleftrightarrow{I}$, so $R^T R = I$. If R is linear and $R^T = R^{-1}$, then $R(\vec{u}) \cdot R(\vec{v}) = \vec{u} \cdot R^T R(\vec{v}) = \vec{u} \cdot I(\vec{v}) = \vec{u} \cdot \vec{v}$, so R is a rigid rotation about zero.

4. R is a rigid rotation about zero if and only if R is linear and maps every (or at least one) Cartesian axis system into another Cartesian axis system. Proof: In the course of proving 2., we showed that if R is a rigid rotation and $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ an orthonormal basis, then so is $\{R(\hat{x}_1), R(\hat{x}_2), R(\hat{x}_3)\}$. Now suppose $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ is a Cartesian axis system and $\{\hat{x}_1^P, \hat{x}_2^P, \hat{x}_3^P\}$ another and R is the unique linear operator such that $R(\hat{x}_i) = \hat{x}_i^P$. Then $R(\vec{u}) = R(u_i \hat{x}_i) = u_i R(\hat{x}_i) = u_i x_i^P$, and likewise $R(\vec{v}) = v_i \hat{x}_i^P$, so $R(\vec{u}) \cdot R(\vec{v}) = u_i v_i = \vec{u} \cdot \vec{v}$.

Theorem (summarizing the above results): The operator $U: \mathbb{R}^3 \to \mathbb{R}^3$ is orthogonal iff it has any one of the following properties, which are then equivalent

1. U is a rigid rotation $|U(\vec{u}) - U(\vec{v})| = |\vec{u} - \vec{v}|$, $U(\vec{0}) = \vec{0}$
2. $U(\vec{u}) \cdot U(\vec{v}) = \vec{u} \cdot \vec{v}$ $\quad \forall \vec{u}, \vec{v}$
3. U^T and U^{-1} exist and are equal
4. $U^T U = I$
5. $U U^T = I$
6. U is linear and maps every Cartesian axis system into another
7. U " " " at least one " " " " .

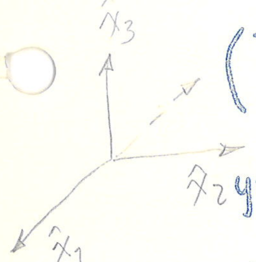
Corollary: If U is orthogonal, so are U^T and U^{-1}

Corollary: If U and V are orthogonal, so is UV, since $(UV)^T = V^T U^T = V^{-1} U^{-1} = (VU)^{-1}$.

A special improper orthog. operator

Now if $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ is a Cartesian axis system; then so is $\{-\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ and if $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ is $\left\{\begin{matrix}\text{right-handed}\\ \text{left-}\ "\end{matrix}\right\}$ then $\{-\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ is $\left\{\begin{matrix}\text{left-handed}\\ \text{right-}\ "\end{matrix}\right\}$. The linear operator $\tilde{U}$

$\tilde{U}(\hat{x}_1) = -\hat{x}_1$, $\tilde{U}(\hat{x}_2) = \hat{x}_2$, $\tilde{U}(\hat{x}_3) = \hat{x}_3$ is orthogonal and is called a <u>simple reflection</u>. Its matrix relative to $\{\hat{x}_i\}$ is



$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, and $\det \tilde{U} = -1$. In this the reflection is in the 23 or yz-plane. The operator $\tilde{U}$ is its own inverse, $\tilde{U}^2 = I$.

Now if U is <u>any</u> improper orthogonal operator, then $\tilde{U}U$ and $U\tilde{U}$ are orthogonal operators, and $\det(\tilde{U}U) = \det(U\tilde{U}) = \det U \det \tilde{U} = (-1)(-1) = 1$, so $\tilde{U}U = P_1$ and $U\tilde{U} = P_2$ are <u>proper</u> orthogonal operators. Then $\tilde{U}(\tilde{U}U) = \tilde{U}P_1 = \tilde{U}^2 U = IU = U$ and $(U\tilde{U})\tilde{U} = P_2\tilde{U} = U\tilde{U}^2 = UI = U$, so we have

<u>Corollary</u> <u>Any</u> improper orthogonal operator U is the product of a proper orthogonal operator and a simple reflection.

$$U = \tilde{U}P_1 = P_2\tilde{U}$$

The orthogonal operators form a group, the group of orthogonal operators on R^n is denoted $O(n)$. The proper orthogonal operators also form a group, denoted $O^+(n)$. The improper orthogonal operators do <u>not</u>, since $\det(UV) = \det U \det V = (-1)(-1) = +1$.

5. Rotations in a Plane : Before trying to visualize orthogonal operators on $\mathbb{R}^3$, we shall examine $\mathbb{R}^2$.

Let U be an orthogonal operator on the plane $U: \mathbb{R}^2 \to \mathbb{R}^2$.

Let $\begin{pmatrix} U_{11} & U_{12} \\ U_{21} & U_{22} \end{pmatrix}$ be the matrix relative to $\{\hat{x}_1, \hat{x}_2\}$. The matrix U is orthogonal, so $U_{11}^2 + U_{21}^2 = U_{12}^2 + U_{22}^2 = 1$, and $U_{11}U_{12} + U_{21}U_{22} = 0$. Now $\exists$ angles α, β between 0 and 2π such that $U_{ij} = \hat{x}_i \cdot U(\hat{x}_j)$ or $U_{11} = \cos\alpha$, $U_{21} = \sin\alpha$, $U_{12} = \sin\beta$, $U_{22} = -\cos\beta$. Moreover $\cos\alpha\sin\beta - \sin\alpha\cos\beta = 0$, so $\sin(\alpha - \beta) = 0$ and $\alpha - \beta = 0$ or $\pm\pi$.

Case 1: $\alpha = \beta$. Then $U = \begin{pmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{pmatrix}$. In this case

do case 2 first

$\det U = -1$, U is improper. The characteristic polynomial is $\lambda^2 - 1 = 0$, so U has two real eigenvalues $\lambda = +1$, $\lambda = -1$. The corresponding eigencolumns are $\begin{pmatrix} 1+\cos\alpha \\ \sin\alpha \end{pmatrix}$ and $\begin{pmatrix} -\sin\alpha \\ 1+\cos\alpha \end{pmatrix}$

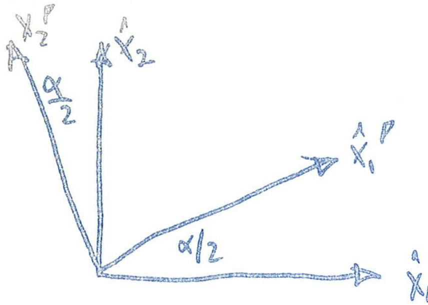
Now $1+\cos\alpha = 2\cos^2\frac{\alpha}{2}$, and $\sin\alpha = 2\sin\frac{\alpha}{2}\cos\frac{\alpha}{2}$, so the unit or normalized eigencolumns are

$\begin{pmatrix} \cos\frac{\alpha}{2} \\ \sin\frac{\alpha}{2} \end{pmatrix}$ for $\lambda = +1$, and $\begin{pmatrix} -\sin\frac{\alpha}{2} \\ \cos\frac{\alpha}{2} \end{pmatrix}$ for $\lambda = -1$. Thus $\exists$ two unit eigenvectors for the operator U, namely

$\hat{x}_1' = \hat{x}_1 \cos\frac{\alpha}{2} + \hat{x}_2 \sin\frac{\alpha}{2}$, $\hat{x}_2' = -\hat{x}_1 \sin\frac{\alpha}{2} + \hat{x}_2 \cos\frac{\alpha}{2}$.

Now $U(\hat{x}_1') = \hat{x}_1'$, and $U(\hat{x}_2') = -\hat{x}_2'$, while clearly $\hat{x}_1' \cdot \hat{x}_2' = 0$. The matrix of U relative to $\{\hat{x}_1', \hat{x}_2'\}$ is $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, and U is simply a reflection in the $\hat{x}_1'$ axis.

There is a picture on the next page.



$$\hat{x}_1' = \hat{x}_1 \cos\frac{\alpha}{2} + \hat{x}_2 \sin\frac{\alpha}{2}$$
$$\hat{x}_2' = -\hat{x}_1 \sin\frac{\alpha}{2} + \hat{x}_2 \cos\alpha/2$$
$$U(\hat{x}_1') = \hat{x}_1' \; ; \; U(\hat{x}_2') = -\hat{x}_2'$$
$$U' = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

<u>Case 2</u>: $\alpha - \beta = \pm\pi$. Then $\sin\beta = -\sin\alpha$, $\cos\beta = -\cos\alpha$

$U = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$: $\det U = +1$, U is proper. The characteristic polynomial is $\lambda^2 - 2\lambda\cos\alpha + 1 = 0$ $\Rightarrow$ <u>no</u> real eigenvalues unless $\alpha = 0$ or π ($U = I$ or $-I$)

$$U(\hat{x}_1) = \hat{x}_1 \cos\alpha + \hat{x}_2 \sin\alpha$$
$$U(\hat{x}_2) = -\hat{x}_1 \sin\alpha + \hat{x}_2 \cos\alpha$$

Thus U is simply a rotation of the whole plane through the angle α about $\vec{0}$. note $(U)_{\text{case }2} = (U)_{\text{case }1} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$

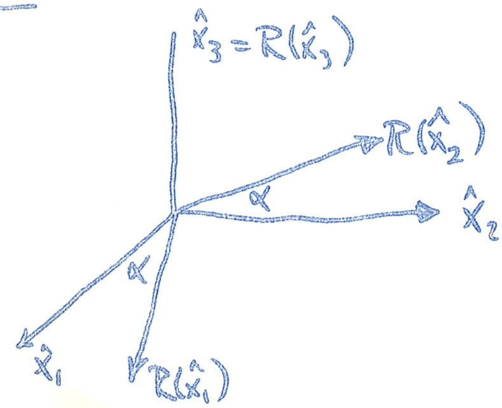
Summarizing: Proper orthogonal operators are rigid rotations about $\vec{0}$ thru some angle α; improper orthogonal operators are reflections of the plane about some line. Proper orthogonal operators (on $\mathbb{R}^2$) have no real eigenvalues (only $\pm I$ are exceptions), improper orthogonal operators always have one eigenvector with eigenvalue $+1$, and another perpendicular to the first with eigenvalue -1.

6. Rotations in Three-dimensional space $\mathbb{R}^3$.

If R is a linear operator $R: \mathbb{R}^3 \to \mathbb{R}^3$ for which a Cartesian axis system can be found such that $M(R)$ has the form

$$M(R) = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

then we call R a <u>rotation thru the angle α about the $\hat{x}_3$ axis</u>. The reason for the name is clear



Any such operator is clearly a proper orthogonal operator. We propose now to prove the converse, which is called

> <u>Eulers Theorem</u>: Every proper orthogonal operator on $\mathbb{R}^3$ is a rotation operator through some angle α between 0 and π about some direction $\hat{n}$. Unless $\alpha = 0$ or π, α and $\hat{n}$ are uniquely determined by the operator

To prove Euler's theorem, we must first prove a lemma.

<u>Lemma</u>: If U is a proper orthogonal operator, $\lambda = 1$ is an eigenvalue of U.

Proof of lemma: U has at least one real eigenvalue. If $U(\vec{v}) = \lambda\vec{v}$, then $U(\vec{v})\cdot U(\vec{v}) = \lambda^2 \vec{v}\cdot\vec{v}$. Since $\vec{v} \neq \vec{0}$, $\lambda^2 = 1$, and $\lambda = \pm 1$. Thus <u>any</u> orthogonal operator $U: R_3 \to R_3$ has either an eigenvalue $+1$ or -1. Consider the second case; i.e. suppose $\lambda = -1$ is an eigenvalue. Let $\hat{n}$ be a unit eigenvector, so $U(\hat{n}) = -\hat{n}$. Now if $\vec{v}\cdot\hat{n} = 0$, then $U(\vec{v})\cdot U(\vec{n}) = -U(\vec{v})\cdot\hat{n} = 0$. Thus if $\vec{v}$ is in the plane perpendicular to $\hat{n}$ ($\vec{v} \in$ orthogonal complement to $\hat{n}$), then so is $U(\vec{v})$. The plane (orthogonal complement) is said to be <u>invariant</u> under U. Thus U generates a two dimensional operator $\tilde{U}$ in that plane. If $\vec{u}$ and $\vec{v}$ are in that plane, then $\tilde{U}(\vec{u})\cdot\tilde{U}(\vec{v}) = \vec{u}\cdot\vec{v}$, so $\tilde{U}$ is a two-dimensional orthogonal operator.

Let $\hat{x}_3 = \hat{n}$, and let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ be a Cartesian axis system with $\hat{x}_1, \hat{x}_2$ in the plane $\perp \hat{n}$. Let $\tilde{U} = \begin{pmatrix} \tilde{U}_{11} & \tilde{U}_{12} \\ \tilde{U}_{21} & \tilde{U}_{22} \end{pmatrix}$ relative to $\{\hat{x}_1, \hat{x}_2\}$. Then $U(\hat{x}_1) = \tilde{U}_{11}\hat{x}_1 + \tilde{U}_{12}\hat{x}_2$, $U(\hat{x}_2) = \tilde{U}_{21}\hat{x}_1 + \tilde{U}_{22}\hat{x}_2$ and $U(\hat{x}_3) = -\hat{x}_3$ (since -1 is an eigenvalue). Thus relative to $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$

$$U = \begin{pmatrix} \tilde{U}_{11} & \tilde{U}_{12} & 0 \\ \tilde{U}_{21} & \tilde{U}_{22} & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

and $\det U = -\det\tilde{U}$.

Thus, since $\det U = 1$, $\det\tilde{U} = -1$, and $\tilde{U}$ is an improper orthogonal operator on $\mathbb{R}^2$, an object all about which we already know. In particular $\exists$ a non-zero vector $\vec{u}$ in the plane normal to $\hat{n}$ such that $\tilde{U}(\vec{u}) = \vec{u}$. But then $U(\vec{u}) = \vec{u}$, so $+1$ is an eiv

Now that the lemma is proved, the proof of Euler's theorem is easy, and in fact similar to the proof of the lemma. This time let $\hat{x}_3$ be a unit vector such that $U(\hat{x}_3) = \hat{x}_3$. The existence of such a vector is assured by the lemma.

Now as in the proof of the lemma, U generates a two dimensional orthogonal operator $\tilde{U}$ on the plane $\perp$ to $\hat{x}_3$; take $\hat{x}_1, \hat{x}_2$ in the plane such that $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ is a Cartesian axis system. Then relative to this system

$$U = \begin{pmatrix} \tilde{U}_{11} & \tilde{U}_{12} & 0 \\ \tilde{U}_{21} & \tilde{U}_{22} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Thus $\det U = \det \tilde{U}$, so $\det \tilde{U} = +1$, so $\tilde{U}$ is proper. Thus

$$U = \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and $\tilde{U}$ is a rotation about the angle α in the plane $\perp$ to the $\hat{x}_3$-axis. Thus U is a rotation operator about the $\hat{x}_3$ axis. If $\pi < \alpha \leq 2\pi$, then we choose a new axis system $\hat{x}_1' = \hat{x}_1$, $\hat{x}_2' = \hat{x}_2$, $\hat{x}_3' = -\hat{x}_3$, and the matrix of U relative to $\{\hat{x}_1', \hat{x}_2', \hat{x}_3'\}$ will be a rotation about $\hat{x}_3' = -\hat{x}_3$ through the angle $\alpha' = 2\pi - \alpha$.

The uniqueness is obvious, so Eulers theorem is proved.

<u>Corollary</u>: (an easy way to find the angle α)

$$\boxed{\operatorname{tr} U = 1 + 2\cos\alpha}$$

<u>Corollary</u>: If U is an improper orthogonal operator on $\mathbb{R}^3$, then U is a rigid rotation about some axis $\hat{n}$ followed (or preceded) by a simple reflection.

Problem 23: Relative to a right handed $\{\hat{x}, \hat{y}, \hat{z}\}$

$$U = \frac{1}{125}\begin{bmatrix} 100 & 60 & -45 \\ -60 & 109 & 12 \\ 45 & 12 & 116 \end{bmatrix}$$

Using a desk calculator, verify that U is orthogonal. Find its rotational axis and the angle of the rotation it represents.

Note that Euler's theorem does provide a 1-1 mapping between orthogonal operators and vectors, but the mapping is <u>not</u> linear, i.e. vector addition is not preserved.

We now come to the study of

Lecture 22 begin here.

7. <u>Symmetric operators</u>

Definition: An operator S is symmetric if $S^T = S$. Relative to some Cartesian axis system $S_{ij} = S_{ji}$.

You should already have seen (problem 20) the relation

$$Q = \underbrace{\tfrac{1}{2}(Q + Q^T)}_{\text{symmetric}} + \underbrace{\tfrac{1}{2}(Q - Q^T)}_{\text{anti-symmetric}}$$

for <u>any</u> operator Q.

<u>Definition</u>: $\frac{1}{2}(Q + Q^T)$ is called the symmetric part and $\frac{1}{2}(Q - Q^T)$ the antisymmetric part of any linear operator Q.

<u>Theorem</u>: <u>Every</u> eigenvalue λ of a symmetric linear operator $S: \mathbb{R}^3 \to \mathbb{R}^3$ is real.

<u>Proof</u>: Let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ be an orthonormal basis and let S_{ij} be the matrix of S relative to this basis. There is at least one complex eigenvalue. Thus $\exists$ a non-zero $\vec{v}$ such that

$$S_{ij} v_j = \lambda v_i$$

Taking the complex conjugate, since S_{ij} are real

$$S_{ij} v_j^* = \lambda^* v_i^*$$

From the first equation $v_i^* S_{ij} v_j = \lambda v_i^* v_i$

From the second $v_i S_{ij} v_j^* = \lambda^* v_i^* v_i$, or rearranging indices $v_i^* S_{ij} v_j = \lambda^* v_i^* v_i$, since $S_{ji} = S_{ij}$. Hence

$$v_i^* v_i (\lambda^* - \lambda) = 0$$

Now $\vec{v} \neq 0$, so $\lambda^* = \lambda$.

But then λ is real.

<u>Corollary</u>: S has an eigenvector in $\mathbb{R}^3$ (a real eigenvector)

Note: in a complex vector space, the corresponding result is that every eigenvalue of a <u>Hermitian</u> linear operator is real.

<u>Problem 24</u>: Prove this.

State theorem first — proof very similar to that of Euler's theorem

8. <u>Eigenvalues and Eigenvectors of Symmetric Operators</u>

Let S be a symmetric linear operator $S: \mathbb{R}^3 \to \mathbb{R}^3$. Let λ_1 be a real eigenvalue for S and let $\hat{a}_1$ be the corresponding unit eigenvector. Now suppose

$\vec{v} \cdot \hat{n}_1 = 0$. Then $S(\vec{v}) \cdot \hat{n}_1 = \vec{v} \cdot S^T(\hat{n}_1)$ $= \vec{v} \cdot S(\hat{n}_1) = \lambda_1 \vec{v} \cdot \hat{n}_1 = 0$. Thus if $\vec{v}$ is in the plane $p \perp$ to $\hat{n}_1$, so is $S(\vec{v})$, thus S generates a linear operator $\tilde{S}$ in the plane p. If $\vec{u}, \vec{v} \in p$, $\vec{u} \cdot \tilde{S}(\vec{v}) = \vec{u} \cdot S(\vec{v}) = S(\vec{u}) \cdot \vec{v} = \tilde{S}(\vec{u}) \cdot \vec{v}$, hence $\tilde{S}$ is symmetric. The linear operator $\tilde{S}$ is called the <u>restriction</u> of S to the plane p, and the plane p is said to be <u>invariant</u> under S. Now let λ_2 be a real eigenvalue for $\tilde{S}$, and let $\hat{n}_2$ be the corresponding unit eigenvector. Then $\tilde{S}(\hat{n}_2) = \lambda_2 \hat{n}_2$, so $S(\hat{n}_2) = \lambda_2 \hat{n}_2$. Also $\hat{n}_2 \cdot \hat{n}_1 = 0$. Now let $\hat{n}_3$ be a unit vector $\perp$ to both $\hat{n}_2$ and $\hat{n}_1$. Then $S(\hat{n}_3) \cdot \hat{n}_2 = \hat{n}_3 \cdot S^T(\hat{n}_2) = \hat{n}_3 \cdot S(\hat{n}_2) = \lambda_2 \hat{n}_3 \cdot \hat{n}_2 = 0$. Similarly $S(\hat{n}_3) \cdot \hat{n}_1 = 0$. Thus $S(\hat{n}_3)$ must be some constant multiple of $\hat{n}_3$, since $S(\hat{n}_3) \perp \hat{n}_2$ and $S(\hat{n}_3) \perp \hat{n}_1$. Thus $S(\hat{n}_3) = \lambda_3 \hat{n}_3$. We have just proved the

Theorem: If $S: \mathbb{R}^3 \to \mathbb{R}^3$ is symmetric, there is a Cartesian axis system $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}$ such that $\hat{n}_1, \hat{n}_2, \hat{n}_3$ are eigenvectors (possibly with different eigenvalues) of S

<u>Corollary</u>: If $S: \mathbb{R}^3 \to \mathbb{R}^3$ is symmetric, there is a Cartesian axis system $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}$ such that the matrix of S relative to that system is diagonal

$$M(\hat{n}_1, \hat{n}_2, \hat{n}_3; S) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

The diagonal elements are the eigenvalues of S.

Proof: Relative to $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}$, $S_{ij} = \hat{n}_i \cdot S(\hat{n}_j)$
$= \lambda_j \hat{n}_i \cdot \hat{n}_j$ (no sum on j) $= \lambda_j \delta_{ij}$ (no sum on j)

<u>Corollary</u>: If $\overleftrightarrow{S}$ is a real symmetric tensor, there is a Cartesian axis system $\{\hat{n}_1, \hat{n}_2, \hat{n}_3\}$ such that

$$\overleftrightarrow{S} = \sum_{i=1}^{3} \lambda_i \hat{n}_i \hat{n}_i$$

where λ_i are the eigenvalues of the linear operator generated by S.

<u>Corollary</u>: If S is a symmetric matrix, there is an orthogonal matrix U such that USU^T is a diagonal matrix, the diagonal elements being the eigenvalues of λ.

Proof: Pick a Cartesian axis system $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$, let S be the linear operator whose matrix relative to this axis system is S. Let $\{\hat{x}_1', \hat{x}_2', \hat{x}_3'\}$ be an axis system consisting of eigenvectors of S, and let S' be the matrix of S relative to this system. Then

$$S_{ij}' = (\hat{x}_i' \cdot \hat{x}_k)(\hat{x}_j' \cdot \hat{x}_l) S_{kl}$$

<u>rows</u> of U (<u>columns</u> of U^T) are the eigenvectors in the unprimed system

Define $U_{ij} = \hat{x}_i' \cdot \hat{x}_j$

Then U is orthogonal, and $S' = USU^T$, and S' is orthogonal with diagonal entries the eigenvalues.

It should be noted that the eigenvectors $\hat{n}_1, \hat{n}_2, \hat{n}_3$ are not necessarily unique, if for example $\lambda_2 = \lambda_3$, then any linear combination $a\hat{n}_2 + b\hat{n}_3$ is an eigenvector with eigenvalue $\lambda_2 = \lambda_3$.

Problem 25: Find all the eigenvalues and eigenvectors of linear operator S whose matrix relative to some axis system is

$$S = \begin{pmatrix} 3 & -\sqrt{6} & \sqrt{3} \\ -\sqrt{6} & 4 & \sqrt{2} \\ \sqrt{3} & \sqrt{2} & 5 \end{pmatrix}$$

Exhibit two different Cartesian axis systems of eigenvectors.

9. Eigenspaces, multiplicity, and commutation

If $\vec{v}_1$ and $\vec{v}_2$ are eigenvectors belonging to the same eigenvalue, then so are $a\vec{v}_1 + b\vec{v}_2 \neq 0$.

Definition: The set of all eigenvectors $\vec{v}$ of a linear operator Q which belong to one particular eigenvalue is called the eigenspace of Q belonging to eigenvalue λ. and may be denoted $\{Q, \lambda\}$. $\{Q, \lambda\}$ is a subspace.

In R^3, the eigenspace belonging to a particular λ can be

1. a straight line thru the origin $(\dim\{Q,\lambda\} = 1)$
2. a plane thru the origin $(\dim\{Q,\lambda\} = 2)$
3. all of R^3 $(\dim\{Q,\lambda\} = 3)$.

The dimension of the eigenspace belonging to a particular eigenvalue λ_1 is called the degeneracy of λ_1 or the geometric multiplicity of λ_1.

If the characteristic polynomial of Q has the form $(\lambda_1 - \lambda)^{n_1}(\lambda_2 - \lambda)^{n_2} \cdots (\lambda_r - \lambda)^{n_r}$, then n_1 is called the algebraic multiplicity of the eigenvalue λ_1.

It is easy to see that for an arbitrary (not necessarily symmetric) linear operator $Q: \mathbb{R}^3 \to \mathbb{R}^3$, the geometric multiplicity of an eigenvalue λ_1 is always $\leq$ the algebraic multiplicity of that eigenvalue. Let Q_1 be the restriction of Q to $\{Q, \lambda_1\}$. Then $\det(Q_1 - \lambda I)$ is a factor of $\det(Q - \lambda I)$. If $m_1 =$ geometric multiplicity of $\lambda_1 = \dim \{Q, \lambda_1\}$, then $\det(Q_1 - \lambda I) = (\lambda_1 - \lambda)^{m_1}$. Thus $m_1 \leq$ algebraic multiplicity of λ_1. An example when they are unequal is given on page 104 of Halmos.

For a symmetric linear operator $S: \mathbb{R}^3 \to \mathbb{R}^3$, the geometric and algebraic multiplicities are the same. Let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ be a Cartesian axis system of eigenvectors

$$M(\hat{x}_1, \hat{x}_2, \hat{x}_3; S) = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

If $\lambda_1, \lambda_2, \lambda_3$ are all different, each has algebraic & geometric multiplicity 1. If $\lambda_1 \neq \lambda_2 = \lambda_3$, then the plane containing $\hat{x}_2$ & $\hat{x}_3$ is the eigenspace of λ_2 while the characteristic poly is $(\lambda_1 - \lambda)(\lambda_2 - \lambda)^2$; thus λ_2 has geometric & algebraic multiplicity 2. If $\lambda_1 = \lambda_2 = \lambda_3$, the geometric and algebraic multiplicities of λ_1 are both 3.

If S is symmetric clearly any two eigenspaces belonging to different eigenvectors are orthogonal.

Finally, Any symmetric linear operator is completely determined by its eigenvalues and eigenspaces.

What if S and T are symmetric linear operators with the same eigenvalues but not the same eigenspaces.

<u>Theorem</u>: Suppose S and T are symmetric with the same characteristic polynomial. Then there is a proper orthogonal operator U such that $T = USU^{-1} = USU^T$. Conversely if $T = USU^{-1}$, T and S have the same characteristic polynomial.

Proof: There exists a Cartesian axis system $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ such that $S(\hat{x}_i) = \lambda_i \hat{x}_i$ and another $\{\hat{y}_1, \hat{y}_2, \hat{y}_3\}$ such that $T(\hat{y}_i) = \lambda_i \hat{y}_i$. Let U be defined by $U(\hat{x}_i) = \hat{y}_i$. Then $U^T(\hat{y}_i) = \hat{x}_i$, and $SU^T(\hat{y}_i) = S(\hat{x}_i) = \lambda_i \hat{x}_i$, so $USU^T(\hat{y}_i) = \lambda_i U(\hat{x}_i) = \lambda_i \hat{y}_i = T(\hat{y}_i)$

Hence $T = USU^T$. If $\det U = -1$, replace $\hat{y}_1$ by $-\hat{y}_1$ in the above argument.

The converse is trivial; if $T = USU^{-1}$, then $U(S - \lambda I)U^{-1} = T - \lambda I$ so $\det(T - \lambda I) = \det(S - \lambda I)$

$T = USU^T$ may be thought of as a rotated version of S.

Now what if the eigenspaces of S and T are the same but not their eigenvalues?

This question is answered by the following theorem, the infinite-dimensional version of which is at the ♡ of quantum mechanics.

<u>Theorem</u> : If $S: R^3 \to R^3$ and $T: R^3 \to R^3$ are symmetric, there is a Cartesian axis system whose unit vectors are eigenvectors for both S and T iff $ST = TS$ (S and T commute).

Proof : first if there is such an axis system $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ then $M(\hat{x}_1, \hat{x}_2, \hat{x}_3 ; S) = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \lambda_3 \end{pmatrix} = \mathcal{S}$, and

$$M(\hat{x}_1, \hat{x}_2, \hat{x}_3 ; T) = \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_3 \end{pmatrix} = \mathcal{T}$$

Clearly $\mathcal{S}\mathcal{T} = \mathcal{T}\mathcal{S}$, thus $ST = TS$.

Now if $ST = TS$, Let $\vec{v} \in \{T, \lambda\}$ so $T\vec{v} = \lambda \vec{v}$
Then $T(S(\vec{v})) = TS(\vec{v}) = ST(\vec{v}) = S(\lambda \vec{v}) = \lambda S(\vec{v})$
so $S(\vec{v}) \in \{T, \lambda\}$. Thus S may be restricted to $\{T, \lambda\}$ on the restriction is symmetric, so $\{T, \lambda\}$ contains a set of mutually perpendicular unit vectors which are eigenvectors of S and form a Cartesian axis system in $\{T, \lambda\}$. Being in $\{T, \lambda\}$ they are also eigenvectors of T. This construction can be repeated for each eigenspace of T, the result is a Cartesian axis system consisting of eigenvectors of both S and T.

10. Positive definite symmetric operators

If Q is an arbitrary linear operator $Q: R^3 \to R^3$, then Q generates a function on the surface of the unit sphere $|\hat{n}| = 1$.

Assign to every $\hat{n} \in \Omega$ the real number

$$\boxed{q(\hat{n}) = \hat{n} \cdot Q(\hat{n}) = \hat{n} \cdot \overleftrightarrow{Q} \cdot \hat{n}}$$

If A is antisymmetric $\hat{n} \cdot A(\hat{n}) = \hat{n} \cdot \overleftrightarrow{A} \cdot \hat{n}$ $= \hat{n} \cdot (\vec{A} \times \hat{n}) = \vec{A} \cdot (\hat{n} \times \hat{n}) = 0$, so all antisymmetric operators generate the same function on the sphere.

Thus any operator Q and its symmetric part generate the same function on the Ω.

However different symmetric operators generate different functions on the unit sphere. Thus if $\hat{n} \cdot S(\hat{n})$ is known for all $\hat{n}$, then S is determined if S is symmetric.

To see this, write $\vec{v} = v\hat{v}$, where $\hat{v} \in \Omega$.

Then $\vec{v} \cdot S(\vec{v}) = v^2 \hat{v} \cdot S(\hat{v})$, so if $\hat{n} \cdot S(\hat{n})$ is known for all $\hat{n} \in \Omega$, then $\vec{v} \cdot S(\vec{v})$ is known for all $\vec{v}$.

Now $(\vec{u} + \vec{v}) \cdot S(\vec{u} + \vec{v}) = \vec{u} \cdot S(\vec{u}) + \vec{u} \cdot S(\vec{v}) + \vec{v} \cdot S(\vec{u}) + \vec{v} \cdot S(\vec{v})$ and $\vec{v} \cdot S(\vec{u}) = S^T(\vec{v}) \cdot \vec{u} = S(\vec{v}) \cdot \vec{u}$, so

$$2\vec{u} \cdot S(\vec{v}) = \underbrace{(\vec{u} + \vec{v}) \cdot S(\vec{u} + \vec{v})}_{\text{known}} - \underbrace{\vec{u} \cdot S(\vec{u})}_{\text{known}} - \underbrace{\vec{v} \cdot S(\vec{v})}_{\text{known}}$$

Thus $\vec{u} \cdot S(\vec{v})$ is known for all $\vec{u}, \vec{v}$; thus S is known.

<u>Definition</u>: The symmetric linear operator S is positive definite (positive semidefinite) if $\forall \vec{v}$, $\vec{v} \cdot S(\vec{v}) > 0$, $(\vec{v} \cdot S(\vec{v}) \geq 0)$.

Theorem: A symmetric linear operator S is positive definite (semidefinite) iff all its eigenvalues are positive (non-negative).

Proof: Let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ be an axis system of eigenvectors

Then $\vec{v} \cdot S(\vec{v}) = v^i v^j \hat{x}_i \cdot S(\hat{x}_j) = \lambda_1 (v_1)^2 + \lambda_2 (v_2)^2 + \lambda_3 (v_3)^2$

Thus if $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 > 0$, $\vec{v} \cdot S(\vec{v}) > 0$

Now say $\vec{v} \cdot S(\vec{v}) > 0$ but some λ, say $\lambda_3 < 0$.

Then take $v_3 = 1$, $v_1, v_2 = 0$. Then $S(\hat{x}_3) \cdot \hat{x}_3 < 0 \Rightarrow$ contradiction

---

Positive definite symmetric operators may be thought of as ellipsoids. In fact the set of positions $\vec{v}$ such that $\vec{v} \cdot S(\vec{v}) = 1$ is an ellipsoidal surface and if this surface is determined then S is determined.

Once again, let $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ be a Cartesian axis system of eigenvectors $S(\hat{x}_i) = \lambda_i \hat{x}_i$.

$$\vec{v} \cdot S(\vec{v}) = \lambda_1 (v_1)^2 + \lambda_2 (v_2)^2 + \lambda_3 (v_3)^2$$

$$\boxed{\begin{array}{l} \vec{v} \cdot S(\vec{v}) = 1 \quad \text{becomes} \\ \lambda_1 (v_1)^2 + \lambda_2 (v_2)^2 + \lambda_3 (v_3)^2 = 1 \end{array}}$$

This is the eqn of an ellipsoid whose principal semi-axes have directions $\hat{x}_1, \hat{x}_2, \hat{x}_3$ and magnitudes $\frac{1}{\sqrt{\lambda_1}}, \frac{1}{\sqrt{\lambda_2}}, \frac{1}{\sqrt{\lambda_3}}$. If an ellipsoid with principal semi-axes $\hat{x}_1, \hat{x}_2, \hat{x}_3$ of length a, b, c is given, then this ellipsoid is generated by the linear operator S whose tensor $\overleftrightarrow{S}$ is

$$\boxed{\overleftrightarrow{S} = \lambda_1 \hat{x}_1 \hat{x}_1 + \lambda_2 \hat{x}_2 \hat{x}_2 + \lambda_3 \hat{x}_3 \hat{x}_3}$$

$\lambda_1 = 1/a^2$

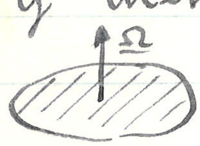
$\lambda_2 = 1/b^2$

$\lambda_3 = 1/c^2$

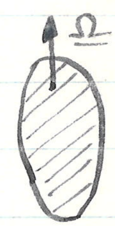
The ellipsoid generated by the stress tensor is the <u>stress ellipsoid</u>, that by the inertia tensor is the <u>inertia ellipsoid</u>. The latter is certainly positive definite but not necessarily the former.
One still defines the "stress quadric"

$$\lambda_1 (u_1)^2 + \lambda_2 (u_2)^2 + \lambda_3 (u_3)^2 = 1$$

but since one of λ_i maybe $-$, can be hyperboloid instead.
Note if inertia ellipsoid looks like



then body looks like



<u>Geometry of</u> symmetric operators
Another ellipsoid is useful also
What does S do to the unit sphere
Say eivs $\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$

$$S(v_i \hat{x}_i) = \lambda_1 v_1 \hat{x}_1 + \lambda_2 v_2 \hat{x}_2 + \lambda_3 v_3 \hat{x}_3$$

Suppose $v \in \Omega$ or $\underline{v} \cdot \underline{v} = 1$

$$S(v) = w_i \hat{x}_i \qquad w_1 = \lambda_1 v_1 \quad w_2 = \lambda_2 v_2 \quad w_3 = \lambda_3 v_3$$

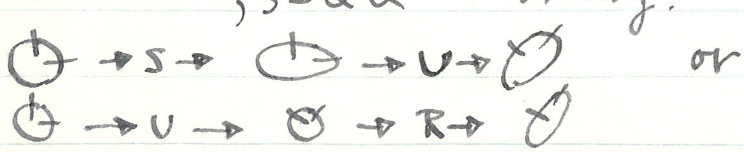
$$v_1^2 + v_2^2 + v_3^2 = 1$$

$$\frac{w_1^2}{\lambda_1^2} + \frac{w_2^2}{\lambda_2^2} + \frac{w_3^2}{\lambda_3^2} = 1$$

$S(\underline{v})$ lies on an ellipsoid whose principal semi-axes are $\lambda_1 \hat{x}_1, \lambda_2 \hat{x}_2, \lambda_3 \hat{x}_3$

note knowledge of this ellipsoid does not determine S unless pos. def.
On the other hand, this ellipsoid is always an ellipsoid.

Polar decomp. theorem: we won't prove
asserts
arbitrary $Q = RU = US$ where
$R = USU^T$, $S = U^T R U$
proof by constr. R, S pos. def. symm.
$\to R^2 = QQ^T; S^2 = Q^T Q$ U orthog.
R is "square root" of QQ^T



S, R are rotated versions of each other

Lecture 22 Review: symm. $\{\hat{x}_i\}$ of eifs – very useful
e.g. use to prove

Cayley–Hamilton theorem for symmetric lin. operators

Theorem: S symmetric with char poly

$$\lambda^3 - c_1\lambda^2 - c_2\lambda - c_3 = 0$$

Then $S^3 - c_1 S^2 - c_2 S - c_3 I = 0$, (i.e. S satisfies its own char. eqn.)

Proof: let $Q = S^3 - c_1 S^2 - c_2 S - c_3 I$

consider $S\hat{x}_i = \lambda_i \hat{x}_i$ in eigenvector $\{\hat{x}_i\}$
then

$$S^2\hat{x}_i = S(S(\hat{x}_i)) = S(\lambda_i \hat{x}_i)$$
$$= \lambda_i S(\hat{x}_i) = \lambda_i^2 \hat{x}_i$$

$$S^3\hat{x}_i = \lambda_i^3 x_i$$

then $Q\hat{x}_i = (\lambda_i^3 - c_1\lambda_i^2 - c_2\lambda_i - c_3)\hat{x}_i$

but λ_i is a root so $Q\hat{x}_i = 0$
similarly $Q\hat{x}_i = 0$, $i = 1,2,3$
but then

$$Qv = Q(v_i\hat{x}_i) = v_i Q\hat{x}_i = 0$$

so Q is the zero operator

Result in fact true for all operators. Proved differently. (use Jordan form)

11. The norm of a linear operator

Definition: If $M: R^3 \to R^3$ is a linear operator, then one defines the norm of M by

$$\|M\| = \sup\left\{ \frac{|M(v)|}{|v|} : v \in R^3 \text{ and } v \neq 0 \right\}$$

or equivalently

$$\|M\| = \sup\left\{ |M(v)| : v \in R^3 \text{ and } |v| = 1 \right\}$$

Remarks:

1. for any $v \in R^3$, $|M(v)| \leq \|M\|\,|v|$ and $\|M\|$ is the smallest constant with this property

2. if $\exists$ a constant k such that $|M(v)| \leq k|v|$ for all $v \in R^3$, then $\|M\| \leq k$

3. For any $u, v \in R^3$ $|(u, Mv)| \leq \|M\|\,|u|\,|v|$ and $\|M\|$ is the smallest constant with this property

Proof: $|(u, Mv)| \leq |u|\,|Mv|$ by Schwarz's inequality so $|(u, Mv)| \leq \|M\|\,|u|\,|v|$. Now suppose $|(u, Mv)| \leq k|u|\,|v|$ for any u, v. Then let $u = Mv$ and we have $|Mv|^2 \leq k|v|\,|Mv|$ so $|Mv| \leq k|v|$ so $\|M\| \leq k$.

4. $\|M^T\| = \|M\|$

5. $\|aM\| = |a|\,\|M\|$ if a is any real number

$\|K\| \ll 1$ implies that components K_{ij} in any Cart. axis system are all $\ll 1$

$$\|K\| = \max \{ |K(v)| : v \in R_3 \text{ and } |v| = 1 \}$$

$$= \max \{ (\underset{\substack{\uparrow \\ \text{matrix } K \text{ in } \{x_1, x_2, x_3\}}}{K})(v) : |v| = 1 \}$$

now consider $K\hat{x}_i = K_{ji}\hat{x}_j$

so

$$|K\hat{x}_i|^2 = \sum_{j=1}^{3} (K_{ji})^2$$

so $\displaystyle\sum_{j=1}^{3} (K_{ji})^2 \leq \|K\|^2$ since $|\hat{x}_i| = 1$

then $\displaystyle\sum_{i=1}^{3}\sum_{j=1}^{3} (K_{ij})^2 \leq 3\|K\|^2$

thus

$$\|K\| \ll 1 \quad \Rightarrow$$

$$\sum_{i,j=1}^{3} (K_{ij})^2 = K_{11}^2 + \dots K_{33}^2 \ll 1$$

thus $K_{ij} \ll 1$ all i, j

Now if $K_{ij} \ll 1$, then $K_{ij}' \ll 1$ any other $\{x_i'\}$ since

$$K_{ij}' = \underbrace{(x_i' \cdot x_k)}_{\leq 1}\underbrace{(x_j' \cdot x_\ell)}_{\leq 1} K_{k\ell}$$

Moreover if $K_{ij} \ll 1$ any $\{\hat{x}_i\}$, then
$\| K \| \ll 1$

$$\forall \underline{v} \quad |Kv|^2 = |(K)(v)|^2$$

$$= \sum_{j=1}^{3} (K_{ij} v_j)^2 \le \sum_{j=1}^{3} \left(\sum_{i=1}^{3} (K_{ij})^2 \right) \left(\sum_{i=1}^{3} v_i^2 \right)$$

$$\le \left(\sum_{i,j=1}^{3} K_{ij}^2 \right) |v|^2$$

let $k^2 = \left(\sum_{i,j}^{3} K_{ij}^2 \right)$

$|Kv|^2 \le k^2 |v|^2$ or

$|Kv| \le k |v|$ thus

$\| K \| \le k$ or $\| K \| \ll K_{11}^2 + \cdots K_{33}^2$

Thus $K_{ij} \ll 1 \Rightarrow \| K \| \ll 1$.

So small operators have small components.

6. $\| M+N \| \leq \| M \| + \| N \|$

proof: for any $v \in \mathbb{R}^3$, $|(M+N)v| = |Mv + Nv|$

$\leq |Mv| + |Nv| \leq \|M\| |v| + \|N\| |v|$. Now use remark 2.

7. $\| MN \| \leq \|M\| \|N\|$

proof: for any $v \in \mathbb{R}^3$, $|MNv| \leq \|M\| |Nv| \leq \|M\| \|N\| |v|$

now use remark 2.

12. Small linear operators and linear operators near the identity.

Definition: We say that K is a small operator if $\| K \| \ll 1$.

Show that small operators have small components

Definition: We say that M is an operator near the identity if $M = I + K$ where K is small.

Operators near the identity have comp. near $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Properties of operators near the identity.

1. if M is near the identity, so is M^T.

2. if M_1 and M_2 are near the identity, so is $M_1 M_2$

proof: $M_1 = I + K_1$, $M_2 = I + K_2$

$M_1 M_2 = (I+K_1)(I+K_2) = I + K_1 + K_2 + K_1 K_2$

but $\| K_1 + K_2 + K_1 K_2 \| \leq \|K_1\| + \|K_2\| + \|K_1\| \|K_2\| \ll 1$

if $\|K_1\| \ll 1$ and $\|K_2\| \ll 1$.

3. If M_1 and M_2 are near the identity say
$M_1 = I + K_1$, $M_2 = I + K_2$, then
$M_1 M_2 = I + K_1 + K_2 +$ error of order $\|K_1\| \|K_2\|$
Hence $M_1 M_2 = M_2 M_1 +$ error of order $\|K_1\| \|K_2\|$
(always commute)

13. Orthogonal linear operators near the identity.
Suppose Q is an <u>orthogonal</u> linear operator near the identity.

1. (Theorem): Then $Q = I + K$ where K is <u>antisymmetric</u>
$K = -K^T +$ error of order $\|K\|^2$.

proof:
$QQ^T = I$ so $I + K + K^T + KK^T = I$. Hence
$KK^T + K + K^T = 0$. But $\|KK^T\| \leq \|K\| \|K^T\| = \|K\|^2$.
$K \approx -K^T$

2. (Theorem): Then $\det Q = 1$

proof: consider $\det \begin{bmatrix} 1+K_{11} & K_{12} & \cdots \\ & 1+K_{22} & \\ & & \ddots \; 1+K_{33} \end{bmatrix}$

only the diagonal term is $O(1)$. All other terms contain at least one K_{ij} and are of order $\|K\|$ or less.
Thus $\det Q \simeq 1$, but $\det Q$ is either $+1$ or -1
hence $\det Q = 1$.

3. If Q is an orthogonal operator near the identity, then relative to any Cartesian axis system,

$$Q_{ij} = \begin{pmatrix} 1 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 1 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 1 \end{pmatrix} + O(\Omega^2)$$

We can write this as

$$\begin{pmatrix} 1 & -\Omega_3 & 0 \\ \Omega_3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & \Omega_2 \\ 0 & 1 & 0 \\ -\Omega_2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -\Omega_1 \\ 0 & \Omega_1 & 1 \end{pmatrix} + O(\Omega^2)$$

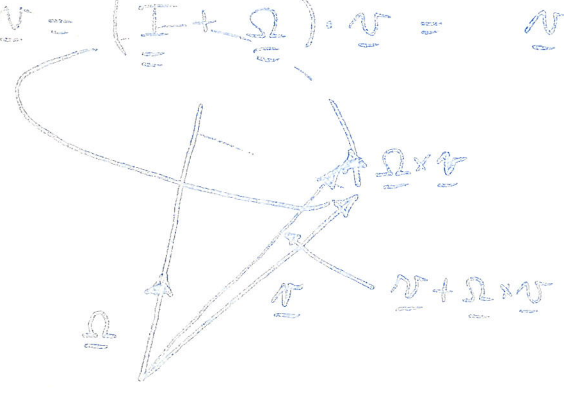
or as

$$\begin{pmatrix} \cos\Omega_3 & -\sin\Omega_3 & 0 \\ \sin\Omega_3 & \cos\Omega_3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\Omega_2 & 0 & \sin\Omega_2 \\ 0 & 1 & 0 \\ -\sin\Omega_2 & 0 & \cos\Omega_2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\Omega_1 & -\sin\Omega_1 \\ 0 & \sin\Omega_1 & \cos\Omega_1 \end{pmatrix} + O(\Omega^2)$$

Thus Q is a rotation thru angle Ω_1 about $\hat{x}_1$ followed by Ω_2 about $\hat{x}_2$ followed by Ω_3 about $\hat{x}_3$.

If we choose our axis system so that $\Omega_2 = \Omega_3 = 0$, then Q is a rotation about $\vec{\Omega}$ thru the angle $|\vec{\Omega}|$

$$\underline{\underline{Q}} \cdot \underline{v} = (\underline{\underline{I}} + \underline{\underline{\Omega}}) \cdot \underline{v} = \underline{v} + \underline{\Omega} \times \underline{v}$$



(see picture)

now whereas the rotation vector associated with finite rotations does not satisfy the law of vector addition, this law is satisfied for infinitesimal rotations

since
$$(I+K_1)(I+K_2) \cong I+K_1+K_2$$
$$= I+K_2+K_1$$
$$\cong (I+K_2)(I+K_1)$$

This leads me to define the instantaneous angular rotation of a body.
Will talk about this next time.

Symmetric linear operators near the identity. In the $\{\hat{x}_i\}$ of eigenvectors

$$\begin{pmatrix} 1+\rho_1 & & \\ & 1+\rho_2 & \\ & & 1+\rho_3 \end{pmatrix} \approx I$$

thus the eivs $1+\rho_1$, $1+\rho_2$, $1+\rho_3$ are all ~ 1
~~En~~ hence positive definite.

14. Arbitrary linear operators near the identity.

Say $M = I + K$ is near the identity. Then define the symmetric and anti-symmetric parts of K by

$$S = (K + K^T)/2$$
$$A = (K - K^T)/2$$

so that $K = S + A$

Now note that

$$M = I + K = I + S + A = (I+S)(I+A) + \text{error of order } \|K\|^2$$

But $I+S$ is a <u>symmetric</u> linear operator and $I+A$ is an <u>orthogonal</u> linear operator.

Hence to within an error of order $\|K\|^2$, an arbitrary linear operator near the identity may be written as the product of a symmetric linear operator $I+S$ and an orthogonal operator $I+A$, and in a unique way.

$$I + K \cong (I+S)(I+A) \approx (I+A)(I+S) \text{ to within an error of order } \|K\|^2$$

This is a special case of the so-called <u>polar decomposition</u> theorem. The result may be shown to be valid for arbitrary linear operators. Any linear operator M may be written in a unique way in the form

$M = US = RU$, where $R = USU^T$, U orthogonal, R, S symmetric