ME 8213 - Engineering Analysis - I
Fall 2017
Lecture Summary

Aug. 19- Tuesday
LECTURE 1: Introduction; Course Policies; Examples

Summary:

1. Mathematics is a LANGUAGE commonly used to describe physical systems. As with any language one needs to be familiar with the vocabulary to be able to communicate in this language. Also, as with any language, a single word or statement has many implications beyond the initial observation being communicated. In engineering and science some researchers specialize in writing down their findings using mathematical statements (experimentalist) and others specialize in exploring all the implications of those observations (analysts).
2. This course will explore some basic mathematical tools used in Linear Analysis. This includes finite dimensional (Matrix Algebra) and infinite dimensional (Linear Differential equations) as applied to many problems in engineering. These tools will help you explore/exploit the linear structure that may be present in an observation posed in mathematical language. Thus, the course will introduce/review the vocabulary used in Linear Analysis and will explore the implications of some frequently used statements based on this vocabulary.
3. Linear algebra is much more than Gaussian Elimination and inverting matrices. For instance, linear transforms such as Laplace Transform and Fourier Transforms are linear operations that can be studied using linear algebra (infinite dimensional). Analysis of engineering problems make extensive use of Linear Analysis: Elasticity, Control Systems, Signal Processing, Regression Analysis, Conduction Heat Transfer, Vibrations, etc.
4. Pythagoras Theorem is fundamental in the development of mathematics. It allows for the definition of the length of a vector, it is used in the definition of derivatives (right angle triangle), and it is used to generalized the notion of dot (inner) product to many dimensions.
5. Here is a link for the proof of Pythagoras Theorem as shown in class.
6. Pythagoras Theorem (orthogonality or perpendicularity from the right angle triangle, definition of length) and its role in defining Cartesian coordinates... the birth of Analytical Geometry.
7. Here is a link on a story about how Descartes discovered the Cartesian Coordinates.
8. The importance of notation. What does it mean when we write a column vector?
9. Vectors: finite dimensional, infinite dimensional. Examples,

   Two dimensional vector: $v = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ or $\{v_i\},\ i = 1,2$

   Three dimensional vector: $v = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ or $\{v_i\},\ i = 1,2,3$

Countable infinite dimensional vector: $v = \begin{bmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ \infty \end{bmatrix}$ or $\{v_i\},\ i = 1,2,\cdots,\infty$

10. Countable vs. Uncountable infinities: an infinite dimensional vector can be made of countable elements (as in a sequence) or uncountable elements (as in a function of one variable).

    Uncountable infinite dimensional vector: $f(t) = \sin(2\pi t)\,,\ 0 \le t < 1$

11. Rational numbers. If the decimal part of a number repeats itself indefinitely (in a cyclical manner), then the number is rational. For example N=54.123123123…. is a rational number because 1000N-N=54069, so N=54069/999. Irrational numbers have infinitely long decimal parts that do not repeat in a cyclical fashion.
12. Why are we interested in vectors of more than three dimensions?
13. Here is a link for the short Biography of Georg Cantor (countable/uncountable infinities).

Aug. 22- Tuesday
LECTURE 2: Review: matrix multiplications, dot products.

Summary:

1. Review: how to multiply two matrices.
2. Review of basic trigonometric identities for $sin(A + B)$ and $cos(A + B)$.
3. Read the notes on Dot Product and Cosine Law posted on the "Files and Notes" folder of myCourses website for this course.
4. Notation: the dot product is also written as $(x, y) = x \cdot y$.
5. The dot product can be generalized to higher dimensions by considering the identity $x \cdot y = x^T y = (x, y)$
   and then allowing x and y to be vectors of N dimensions, with N ≥ 0. Common generalizations of the dot product are shown below. The generalized or extended concept of dot product goes by the name of "inner" product.

   a) If f and g are N dimensional vectors, then
   $$(f, g) = f^T g = \sum_{i=0}^{N-1} f_i g_i$$

   b) If f and g are infinite dimensional sequences, then
   $$(f, g) = \sum_{i=0}^{\infty} f_i g_i$$

   c) If $f(x)$ and $g(x)$ are functions of x with $a \le x \le b$, then

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

In cases (b) and (c) it is assumed that the square of the lengths of each vector is finite, i.e., $(f,f) < \infty$, and $(g,g) < \infty$. Infinite dimensional, finite length sequences with the inner product in case (b) are used to create the "little $\ell 2$" vector space. "Square-integrable" functions, i.e., $(f,f) < \infty$, and $(g,g) < \infty$, as would the functions be in case (c), are used to create the "big L2" vector space. The ℓ and L are in honor of Henri Lebesgue who worked with these functions around early 1900's.

6. It is important to understand that $\hat{\imath}$ is not necessarily equal to the vector $[1\ 0\ 0]^T$ and $\hat{\jmath}$ is not necessarily equal to $[0\ 1\ 0]^T$. Instead, the only requirements for $\hat{\imath}$ and $\hat{\jmath}$ are that

$$\hat{\imath} \cdot \hat{\jmath} = 0\ (orthogonal)$$

and

$$\hat{\imath} \cdot \hat{\imath} = \hat{\jmath} \cdot \hat{\jmath} = 1\ \ (unit\ length\ vectors)$$

Other examples of orthogonal, unit vectors $\hat{\imath}$ and $\hat{\jmath}$ are:

$$\hat{\imath} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \frac{1}{\sqrt{3}} \text{ and } \hat{\jmath} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \frac{1}{\sqrt{2}}$$

and

$$\hat{\imath} = \begin{bmatrix} 5 \\ 1 \\ 2 \end{bmatrix} \frac{1}{\sqrt{30}} \text{ and } \hat{\jmath} = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} \frac{1}{\sqrt{5}}$$

Observe how dividing by the length of the vectors normalizes the vectors. (In here, "Normalizing" a vector, means to make the length (magnitude) of the vector equal to one)

When $\hat{\imath}$ is selected as $[1\ 0\ 0]^T$ and $\hat{\jmath}$ is $[0\ 1\ 0]^T$, etc., those coordinate systems are called "natural" coordinates.

7. It is important to understand the meaning of the axes in a coordinate system. For instance, the labels i, j, and k are just that, labels. The i vector can represent a unit direction in a coordinate system. What makes a coordinate system? A coordinate system is used to identify a value of interest. For example the x coordinate may be number of apples and the y coordinate may be number of pears. So the point (5,3) represents 5 apples and 3 pears. The fact that the coordinate systems are perpendicular to each other is used to represent the fact that the choices are independent of each other.
8. How to hide (embed) a 2D circle in five dimensions. A circle in 2D can be drawn using a set of, say one hundred points, where each point has corresponding x and y components and, therefore, can be considered a vector in two dimensions. Then the k-th point of the circle can be expressed as $\bar{p}_k = x_k \cdot \hat{\imath} + y_k \cdot \hat{\jmath}$ with $k = 0,1,\cdots,99$ and where

$$x_k = R \cdot \cos(\theta_k)$$
$$y_k = R \cdot \sin(\theta_k)$$

and

$$\theta_k = k \cdot \left(\frac{2\pi}{100}\right)$$

$$R = 1 \text{ for a unit circle.}$$

So far in this course we have used $\hat{\imath}$ and $\hat{\jmath}$ as unit vectors such that $\hat{\imath} \cdot \hat{\jmath} = 0$ and $\hat{\imath} \cdot \hat{\imath} = \hat{\jmath} \cdot \hat{\jmath} = 1$. For a 2D circle, without much thought, one would take for granted that

$$\hat{\imath} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \hat{\jmath} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

However, $\hat{\imath}$ and $\hat{\jmath}$ can be defined instead to be two unit vectors in 5D. For example, let

$$\hat{\imath} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad \hat{\jmath} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

substituting the above definitions of $\hat{\imath}$ and $\hat{\jmath}$ into $\bar{p}_k = x_k \cdot \hat{\imath} + y_k \cdot \hat{\jmath}$ yields

$$p_k = \frac{x_k}{\sqrt{2}} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} + y_k \cdot \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad or \quad p_k = \begin{bmatrix} x_k/\sqrt{2} \\ 0 \\ 0 \\ y_k \\ x_k/\sqrt{2} \end{bmatrix}$$

So it is seen that p_k is now expressed as a vector in 5D, and for $k = 0,1,\cdots,99$ all the points in the original 2D circle are now embedded in five dimensions.

In practice, the problem is often backwards. Given 100, 5D data points, i.e., each data point consisting of five measurements, find the shape hidden in the data. For the example above, the shape would be a circle. For a more specific example, consider part of a thermodynamic table for water where five physical quantities can be considered: P,T,ν,s,and h. Using 100 data points from this thermodynamic table, it may be of interest to see if there is a shape, i.e., a functional relationship, hidden in this data. Similar problems will be tackled later in the course.

In this example we have learned that $\hat{\imath} \cdot \hat{\jmath} = 0$ and $\hat{\imath} \cdot \hat{\imath} = \hat{\jmath} \cdot \hat{\jmath} = 1$ do not necessarily constrain our problems to two dimensions or three dimensions. Instead, the use of these vectors and the corresponding matrix notation is up to our imagination as long as follow the rules of unit vectors and matrix multiplications. In fact, there are many scientific publications which merit consists on specific applications in several areas of engineering and science where the interpretations and choices of these vectors allow for more convenient solutions to the applications considered.

Exam Questions:

1) Prove that $\bar{u} \cdot \bar{v} = u^T v$, where $\bar{u}$ and u represent the same vector, the first one using the i,j,k,… notation (where i,j,k,… are any set of orthogonal unit vectors, i.e., orthonormal and not necessarily "natural"), and the second one using matrix or array notation. The same relationship applies for $\bar{v}$ and v.
2) Explain how to generalize the concept of dot product for N-dimensional vectors where N is a positive integer, for sequences, and for functions.
3) Use basic trigonometric identities to prove that $\int_0^{2\pi} sin(\theta) \cdot cos(\theta) d\theta = 0$.
4) Use basic trigonometric identities to prove that $\int_0^{2\pi} cos^2(\theta) d\theta = \pi$
5) Explain how to embed, i.e., "hide", a circle in five dimensions.

Aug. 24- Thursday
LECTURE 3: Parallel and Perpendicular components of a vector.

Summary:

1. How to plot a circle in MathCad using parametric notation.
2. How to create a dot product for functions using MathCad.
3. With the understanding of a dot product, the idea of gradient can be readily understood. Below is a summary of the discussion on gradients discussed in class.

   The derivatives of functions of multiple variables result in linear structures, i.e., vector/matrix notation, that can be studied using linear algebra. For example, the total derivative of a temperature distribution $T(x, y)$ at the point (x_o, y_o) can be expressed in terms of the inner (dot) product between the gradient vector and a vector of independent variables (direction):

   $$dT = (\partial T/dx)dx + (\partial T/dy)dy = [\partial T/\partial x \quad \partial T/\partial y] \begin{bmatrix} dx \\ \\ dy \end{bmatrix}$$

   where the partial derivatives are evaluated at the point $p_o = (x_o, y_o) = \begin{bmatrix} x_o \\ y_o \end{bmatrix}$.
   The gradient of $T(x, y)$ is defined as

   $$\nabla T = [\partial T/\,\partial \mathrm{x} \quad \partial T/\partial y]\,.$$

   The vector $\begin{bmatrix} dx \\ dy \end{bmatrix}$ represents an infinitesimal change from the point p_o, that is, $\begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} x - x_o \\ y - y_o \end{bmatrix} = p - p_o$. This vector corresponds to the independent variables and can be expressed as a vector of magnitude dr and direction given by a unit vector $\bar{u}$, i.e.,

   $$\begin{bmatrix} dx \\ dy \end{bmatrix} = \bar{u}\, dr.$$

   With these definitions, dT can be expressed as

$dT = \nabla T \cdot \bar{u}\, dr$ which results in

$dT/dr = \nabla T \cdot \bar{u}$

The last result is the definition of directional derivative: dT/dr is the rate of change of $T(x, y)$ along the direction $\bar{u}$.

The magnitude of the directional derivative will be largest when the dot product $\nabla T \cdot \bar{u}$ is as large as possible. This will happen if the unit vector $\bar{u}$ points in the same direction as the gradient vector ∇T. **Thus, the gradient is the direction of maximum change.**

If the unit vector $\bar{u}$ is perpendicular to the gradient then $\nabla T \cdot \bar{u}$ will be zero resulting in a total derivative equal to zero, i.e., the direction of no temperature change. Thus, the gradient must be perpendicular to the contours of constant temperature, i.e., the constant temperature paths.

4. How to use the inner (dot) product to find components of one vector that are perpendicular and parallel to another vector. Suppose we have two N-dimensional vectors, m and v, and we want the components of vector m that are perpendicular and parallel to vector v. Thus, we can write

$$m = m_{\parallel v} + m_{\perp v}$$

Solving the equation above for $m_{\perp v}$ yields

$$m_{\perp v} = m - m_{\parallel v}$$

Since $m_{\parallel v}$ is a constant multiplied times v, i.e.,

$$m_{\parallel v} = cv$$

then

$$m = m_{\parallel v} + m_{\perp v}$$

can be expressed as

$$m = cv + m_{\perp v}$$

By definition, $v^T m_{\perp v} = 0$ so taking the inner (dot) product with respect to v yields

$$v^T m = cv^T v + \overbrace{v^T m_{\perp v}}^{0}$$

or

$$c = \frac{v^T m}{v^T v}$$

It follows that

$$m_{\parallel v} = cv = \left(\frac{v^T m}{v^T v}\right) v$$

and

$$m_{\perp v} = m - \left(\frac{v^T m}{v^T v}\right) v$$

5. Example: The formula for the average value of a series of n measurements can be obtained using dot products as follows. Consider the problem of finding a constant temperature value $\mathcal{T}$ that is most representative of the n temperature measurements, i.e.,

$$\begin{bmatrix} \mathcal{T} \\ \mathcal{T} \\ \vdots \\ \mathcal{T} \end{bmatrix} \cong \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_n \end{bmatrix} \quad \text{or} \quad \mathcal{T}\bar{1} \cong \bar{T} \quad \text{where } \bar{T} = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_n \end{bmatrix} \text{ and } \bar{1} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Note that $\mathcal{T}$ is the constant to be found. $\bar{T}$ can be expressed in terms of the parallel and perpendicular directions to the vector $\bar{1}$:

$$\bar{T} = \mathcal{T}\bar{1} + \bar{T}_{\perp\bar{1}}$$

Using the concept of Pythagoras theorem or "orthogonality", the smallest (lengthwise) correction to the original measurement vector $\bar{T}$ in order to obtain a vector with constant values, $\mathcal{T}\bar{1}$, is obtained by ignoring or eliminating the perpendicular component, i.e., $\bar{T}_{\perp\bar{1}}$, of $\bar{T}$. To find $\mathcal{T}$ take the dot product with $\bar{1}$ to obtain,

$$\bar{T}\cdot\bar{1} = \mathcal{T}\bar{1}\cdot\bar{1} + \overbrace{\bar{1}\cdot\bar{T}_{\perp\bar{1}}}^{0}$$

Solving for $\mathcal{T}$ yields

$$\mathcal{T} = \left(\frac{\bar{T}\cdot\bar{1}}{\bar{1}\cdot\bar{1}}\right) = \left(\frac{\sum_{k=1}^{n} T_k}{n}\right)$$

so it follows that

$$\mathcal{T} = \frac{\sum_{k=1}^{n} T_k}{n}.$$

This is the formula for the average or "most representative" value of $\bar{T}$ in the "pythagoras theorem" or "least squares" sense.

Exam Questions:

1) Prove that the gradient is the direction of maximum change. Use a scalar function of two variables as an example, e.g, T(x,y).
2) Explain how to find the directional derivative for a function T(x,y) along a direction $\bar{u}$, at the point $p_o = (x_o, y_o)$
3) Given vectors v and w show how to express v in terms of components parallel and perpendicular to w. Clearly explain every step of the process.
4) Provide a proof showing that the average is the "most representative" value of a group of measurements using the least-squares perspective.

Aug. 29- Tuesday

LECTURE 4: Introduction to Gram-Schmidt Orthogonalization.

Summary:

1. Review: The dot product can be generalized to higher dimensions by considering the identity
   $x \cdot y = x^T y = (x, y)$

   and then allowing x and y to be vectors of N dimensions, with N ≥ 0. Common generalizations of the dot product are shown below. The generalized or extended concept of dot product goes by the name of "inner" product.

   a) If f and g are N dimensional vectors, then

   $$(f, g) = f^T g = \sum_{i=0}^{N-1} f_i g_i$$

   b) If f and g are infinite dimensional sequences, then

   $$(f, g) = \sum_{i=0}^{\infty} f_i g_i$$

   c) If $f(x)$ and $g(x)$ are functions of x with $a \leq x \leq b$, then

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

   In cases (b) and (c) it is assumed that the square of the lengths of each vector is finite, i.e., $(f, f) < \infty$, and $(g, g) < \infty$. Infinite dimensional, finite length sequences with the inner product in case (b) are used to create the "little $\ell 2$" vector space. "Square-integrable" functions, i.e., $(f, f) < \infty$, and $(g, g) < \infty$, as would the functions be in case (c), are used to create the "big L2" vector space. The ℓ and L are in honor of Henri Lebesgue who worked with these functions around early 1900's.

2. An example of how to take advantage of vector/matrix notation: Measurement uncertainties in a two phase flow problem. Consider wet steam (liquid + vapor) flowing through a flow splitter. Liquid and vapor mass flowrates are measured at the input and outputs of the flow splitter. All measurements are subject to uncertainties. The problem is to find the best correction to use to obtain corrected flow estimates of the measurements and obey conservation of mass (continuity) as applied to the flow splitter. The problem is first described mathematically using conservation of mass as follows: $(w_1 + v_1) + (w_2 + v_2) = (w_3 + v_3)$, where the subscripts 1 and 2 indicate outlet water and vapor content and the subscript 3 is the inlet water and vapor content. This equation can be conveniently expressed using vector/matrix notation as follows,

$$[w_1 \quad v_1 \quad w_2 \quad v_2 \quad w_3 \quad v_3]\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} = 0$$

$$\text{or} \quad m^T v = m \cdot v = 0$$

where $m^T = [w_1 \quad v_1 \quad w_2 \quad v_2 \quad w_3 \quad v_3]$ and $v = [1 \quad 1 \quad 1 \quad 1 \quad -1 \quad -1]$

The above equation implies that m should be orthogonal to v. In practice this is not the case because of measurement errors. However, one can visualize m in terms of parallel and orthogonal components to v: $m = m_{\parallel v} + m_{\perp v}$. By definition, $m_{\perp v} \cdot v = 0$ implying that $m_{\perp v}$ represents a possible set of corrected measurements. To find $m_{\perp v}$, use the equation

$$m_{\perp v} = m - m_{\parallel v}$$

Since $m_{\parallel v}$ is a constant multiplied times v, i.e.,

$$m_{\parallel v} = cv$$

then

$$m = m_{\parallel v} + m_{\perp v}$$

can be expressed as

$$m = cv + m_{\perp v}$$

Taking the inner (dot) product with respect to v yields

$$v^T m = cv^T v + \overbrace{v^T m_{\perp v}}^{0}$$

or

$$c = \frac{v^T m}{v^T v}$$

It follows that

$$m_{\parallel v} = cv = \left(\frac{v^T m}{v^T v}\right) v$$

and the desired solution is then

$$m_{\perp v} = m - \left(\frac{v^T m}{v^T v}\right) v$$

This solution is known as the "least squares" solution. The "least" part from least-squares, is due to the fact that the shortest correction to m to make it parallel to v is obtained by subtracting $m_{\perp v}$ from m, i.e.,

$$m_{\parallel v} = m - m_{\perp v}$$

The "squares" part is associated with Pythagoras Theorem, in that removing $m_{\perp v}$ from m, is equivalent to removing $|m_{\perp v}|^2$ from

$$|m|^2 = \left|m_{\parallel v}\right|^2 + |m_{\perp v}|^2$$

3. Recall that the following notation is equivalent

$$x \cdot y = (x, y) = x^T y$$

4. Given three vectors, v_1, v_2, and v_3, one can try to find a set of three orthogonal <u>unit</u> (orthonormal) vectors that can be defined as the i, j, k units vectors as follows
First let

$$i = v_1/|v_1|$$

The next vector, v_2, can be expressed as

$$v_2 = a_2\, i + b_2\, j$$

where j is assumed to be another unit vector perpendicular to i. Since i and v_2 are known, and i and j are perpendicular, it is possible to find a_2 by taking the inner (dot) product with i of the left and right side of the above equations

$$v_2 \cdot i = a_2 \overbrace{i \cdot i}^{1} + b_2 \overbrace{j \cdot i}^{0}$$

or

$$(v_2, i) = a_2 \overbrace{(i, i)}^{1} + b_2 \overbrace{(j, i)}^{0}$$

which yields,

$$a_2 = (v_2, i)$$

Now that a_2 is known, and since v_2 is also known (given), the term $b_2\, j$ can be expressed in terms of known values

$$\mathrm{b_2\, j = v_2 - a_2\, i}$$

Evaluating the magnitude of both sides of the above equation yields,

$$\mathrm{b_2 = |v_2 - a_2\, i|}$$

if b_2 is not zero, then

$$j = \frac{[v_2 - a_2\, i]}{b_2}$$

or

$$j = \frac{v_2 - (v_2, i)\, i}{|\, v_2 - (v_2, i)\, i|}$$

The third unit vector can be obtained by expressing v_3 as follows,

$$v_3 = a_3\, i + b_3\, j + c_3\, k$$

where k is another unit vector, perpendicular to the plane spanned by the unit vectors i and j. Following the same idea as before, the coefficients a_3 and b_3 are found using the inner (dot) product to obtain

$$a_3 = (v_3, i) \text{ and } b_3 = (v_3, j)$$

and since

$$c_3\, k = v_3 - a_3\, i - b_3\, j$$

if follows that

$$c_3 = |v_3 - a_3\, i - b_3\, j|$$

if c_3 is not zero, then

$$k = \frac{[v_3 - a_3\, i - b_3\, j\,]}{c_3}$$

or

$$k = \frac{v_3 - (v_3, i)\, i - (v_3, j)\, j}{|\, v_3 - (\, v_3, i)\, i - (v_3, j)\, j|}$$

5. It is important to understand that i is not necessarily equal to the vector $[1\ 0\ 0]^T$ and j is not necessarily equal to $[0\ 1\ 0]^T$. Instead, i is a unit vector in the same direction as v_1 and j is just another unit vector perpendicular to i but in the same plane containing $v_1$and v_2.
6. The process described above shows how to obtain three orthogonal unit vectors provided that a_1, b_2 and c_3 are not zero. This process goes by the name of "Gram-Schmidt Orthogonalization" and is used to find as many orthogonal unit vectors as possible from a given set of vectors.
7. If a coefficient corresponding to a potential unit vector is zero, then this vector must be disregarded and not used. For instance if b_2 is zero, then v_2 should be discarded and the next unit vector is found by using v_3.

Exam Questions:

1. Explain how matrix notation and Pythagoras Theorem can be used to obtain the Least Squares Solution to the following problem. Consider wet steam (liquid + vapor) flowing through a flow splitter. Liquid and vapor mass flowrates are measured at the input and outputs of the flow splitter. All measurements are subject to uncertainties. The problem is to find the best correction to use to obtain corrected flow estimates of the measurements and obey conservation of mass (continuity) as applied to the flow splitter.
2. Given v=v(x) and w=w(x) are “finite energy” functions of x in the interval [-1,1]. show how to express v in terms of components parallel and perpendicular to w. Clearly explain every step of the process.
3. Repeat Question 2, but this time assume v and w are “finite energy” sequences.

Aug. 31- Thursday

LECTURE 5: Review Gram-Schmidt Orthogonalization.

Summary:

1. Review of Lecture 4.

2. Example of Gram-Schmidt applied to three infinite dimensional, uncountable, vectors ($v_1 = 1, v_2 = x$, and $v_3 = x^2$) in the interval from [-1,1]:

$$i = \frac{v_1}{|v_1|} = \frac{1}{\sqrt{\int_{-1}^{1}(1\cdot 1)dx}} = \frac{1}{\sqrt{2}}$$

to find *j* first obtain

$$v_2 - (v_2, i)i = x - \left(\int_{-1}^{1}\left(x\cdot\frac{1}{\sqrt{2}}\right)dx\right)\frac{1}{\sqrt{2}} = x - \frac{1}{2}\overbrace{\int_{-1}^{1} xdx}^{0} = x$$

which implies that 1 and x are already orthogonal in that interval. Normalizing,

$$j = \frac{v_2 - (v_2, i)\, i}{|\, v_2 - (v_2, i)\, i|} = \frac{x}{\sqrt{\int_{-1}^{1}(x\cdot x)dx}} = \frac{x}{\sqrt{2/3}} = \sqrt{\frac{3}{2}}\cdot x$$

Similarly,

$$v_3 - (v_3, i)\, i - (v_3, j)\, j$$

$$= x^2 - \underbrace{\left(\int_{-1}^{1}\left(x^2\cdot\frac{1}{\sqrt{2}}\right)dx\right)\frac{1}{\sqrt{2}}}_{1/3} - \left(\overbrace{\int_{-1}^{1}\left(\overbrace{x^2\cdot\sqrt{\frac{3}{2}}\cdot x}^{\text{this is a cubic}}\right)dx}^{\text{this integral is zero}}\right)\sqrt{\frac{3}{2}}\cdot x$$

simplifying

$$v_3 - (v_3, i)\, i - (v_3, j)\, j = x^2 - \frac{1}{3}$$

and

$$|v_3 - (v_3, i)\, i - (v_3, j)\, j| = \sqrt{\int_{-1}^{1}\left(x^2 - \frac{1}{3}\right)^2 dx} = \frac{2\sqrt{10}}{15}$$

it follows that

$$k = \frac{v_3 - (v_3, i)\, i - (v_3, j)\, j}{|v_3 - (v_3, i)\, i - (v_3, j)\, j|} = \frac{15}{2\sqrt{10}}\left(x^2 - \frac{1}{3}\right)$$

3. A pdf printout of some of the steps shown in MathCad on this lecture has been posted on the Notes section of the course website in MyCourses. The file is called "Example of Gram Schmidt and functions.pdf"

Exam Questions:

1) Explain the process and the logic behind the process of Gram-Schmidt orthogonalization as applied to vectors v1, v2, v3, v4, etc. Do not assume you know whether these vectors are finite dimensional, sequences, or functions.
2) What will happen to the Gram Schmidt procedure if v3 is in the same plane as v1 and v2? Clearly explain your answer.
3) Explain how the answer to question 1 can be made more specific for n-dimensional vectors.
4) Explain how the answer to question 1 can be made more specific for sequences. Can any sequence be used? If not, explain what kind of sequences can be used, i.e., what is the restriction on the sequences?
5) Explain how the answer to question 1 can be made more specific for functions. Can any function be used? If not, explain what is the restriction on the functions?

Sep. 5- Tuesday
LECTURE 6: Lines, Planes, and gradients.

Summary:

1. Review of how to obtain the equation of a parametric line. The "parametric" equation for a line <u>between</u> two points, $p_0 = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$ and $p_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ is given by

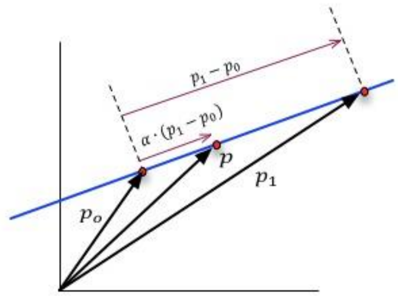
$$p = (1 - \alpha) \cdot p_0 + \alpha \cdot p_1$$

where $0 \le \alpha \le 1$.

Observe that when $\alpha = 0$ then $p = p_0$ and when $\alpha = 1$ then $p = p_1$. Perhaps the equations are better understood by collecting on α to obtain

$$p = \overbrace{p_0}^{\substack{initial \\ point}} + \alpha \cdot \overbrace{(p_1 - p_0)}^{\substack{direction \\ vector \\ (not \\ normalized)}}$$

Geometrically, it is seen that the vector p is the sum of two vectors. The constant vector p_0 and the vector $\alpha \cdot (p_1 - p_0)$, as seen in the picture below. Also observe that if the scalar $\alpha > 1$, then the point p will continue to move past, to the right of p_1, along the same blue line. Similarly, $\alpha < 1$, will lead to p being located to the left of p_1, along the same blue line.



2. A plane can be defined as

   $$n^T(p - p_o) = 0$$

   where n is a vector perpendicular to the plane and p and p_o are any two points on the plane. Although not a requirement, the vector n is typically normalized (in this sense normal would mean an orthogonal **unit** vector). The above equation states the fact that the dot product between two perpendicular vectors, one normal to the plane (n) and one on the plane ($p - p_o$) must equal zero. Thus the equation for the plane can be expressed as

   $$n^T p = n^T p_o$$

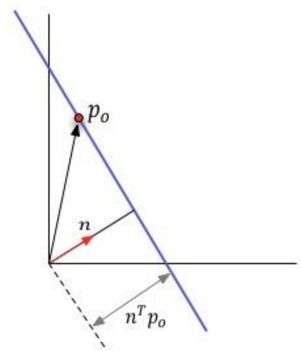
   In three dimensions one can write $p = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ and $n = \begin{bmatrix} n_0 \\ n_1 \\ n_2 \end{bmatrix}$ and the above equation becomes

   $$n_0 x + n_1 y + n_2 z = n^T p_o \quad \text{or}$$

   $$z(x, y) = \left(\frac{n^T p_o}{n_2}\right) - \left(\frac{n_0}{n_2}\right) x - \left(\frac{n_1}{n_2}\right) y$$

3. It is important to observe that the only two pieces of information are needed to define the equation of the plane: 1) the orthogonal vector (which may be normalized later), and 2) a point p_o in the plane.

4. The minimum distance from the plane to the origin is given by taking the dot product between any point p_o in the plane and the normal vector n (this is a unit vector) as seen in the figure below, i.e., $n^T p_o$ is the minimum distance from the plane to the origin.



5. In practice, the equation of a plane is often given as

$$ax + by + cz = d$$

or

$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = d$$

From the above equation, one may want to be able to find the normal to the plane and also the minimum distance from the plane to the origin. To find the normal, the above equation is compared to the previously derived equation of a plane,

$$n_0 x + n_1 y + n_2 z = n^T p_o$$

or

$$\begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} n_0 \\ n_1 \\ n_2 \end{bmatrix} = n^T p_o$$

Observe that the vector $\begin{bmatrix} n_0 \\ n_1 \\ n_2 \end{bmatrix}$ is a normal vector, i.e., a **<u>unit</u>** vector perpendicular to the plane. But, from the original equation $ax + by + cz = d$, the vector of coefficients $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ is not necessarily a unit vector. Comparing the two equations, it is seen that

$$\begin{bmatrix} n_0 \\ n_1 \\ n_2 \end{bmatrix} = \frac{1}{\sqrt{a^2 + b^2 + c^2}} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

That is, the normal vector is obtained by "normalizing" (dividing by its magnitude) the vector of coefficients. To simplify the notation, define the magnitude α as

$$\alpha = \sqrt{a^2 + b^2 + c^2}$$

to obtain

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \alpha \begin{bmatrix} n_0 \\ n_1 \\ n_2 \end{bmatrix}$$

it is possible to write

$$\underbrace{\alpha n_0}_{a} x + \underbrace{\alpha n_1}_{b} y + \underbrace{\alpha n_2}_{c} z = \underbrace{\alpha n^T p_o}_{d}$$

the shortest distance from the plane to the origin is then

$$n^T p_o = \frac{d}{\alpha} = \frac{d}{\sqrt{a^2 + b^2 + c^2}}$$

6. Recall that the equation for the line tangent to the constant temperature line (isotherm) is obtained by setting the total derivative of the temperature function to zero as follows

$$dT = \left(\frac{\partial T}{dx}\right) dx + \left(\frac{\partial T}{dy}\right) dy = \begin{bmatrix} \frac{\partial T}{\partial x} & \frac{dT}{\partial y} \end{bmatrix} \begin{bmatrix} dx \\ dy \end{bmatrix} = 0$$

The equation for the line is an extrapolation of the equation above. The extrapolation is achieved by the following substitution

$$\begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} x - x_o \\ y - y_o \end{bmatrix} = p - p_o$$

Since the gradient is normal (perpendicular) to the constant temperature profile, one can write

$$n^T = \begin{bmatrix} \frac{\partial T}{\partial x} & \frac{dT}{\partial y} \end{bmatrix}$$

Thus, the original equation can be expressed as

$$dT = n^T (p - p_o) = 0$$

This equation is of the same form as the equation of a plane (or line in this case) discussed previously.

7. From the discussions above, it is seen that, in order to find a vector tangent to a contour line, one can choose to obtain an equation for a 3D surface with the specified contour lines. The gradient of the equation for the 3D surface is used to generate the equation of the line tangent to the contour line. To illustrate this, consider a very simple example: finding a vector perpendicular to a circle with radius r and centered at (1,1). The equation for the circle would be $(x-1)^2 + (y-1)^2 = r^2$. To create a 3D surface function with contours equal to the circles just let the constant r^2 be equal to a new (dependent) variable z (think of z as the temperature in the previous discussion). By construction, the equation z(x,y)=constant yields the desired pre-specified contours. Then take the gradient of z with respect to x and y and substitute the numerical values of the desired $p_o = (x_o, y_o)$ location. The equation of the tangent line is given by $n^T(p - p_o)$, where n is the gradient vector.

8. To graph a line in the direction of the gradient (normal to the contour line) just start at the point p_o and head on the direction of the gradient. For the temperature distribution problem, this results in the equation

$$p = p_o + \nabla T]_{p_o} \cdot \alpha$$

or

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x_o \\ y_o \end{bmatrix} + \begin{pmatrix} \left.\frac{\partial T}{\partial x}\right]_{x_o,y_o} \\ \left.\frac{\partial T}{\partial y}\right]_{x_o,y_o} \end{pmatrix} \cdot \alpha$$

Here α is a scalar variable. If the gradient is normalized (converted into a unit vector), then α represents the distance along the gradient direction.

9. In the previous discussion, perpendicular and tangential vectors to a contour line were obtained from a 3D surface used to generate the contour. Suppose now that one needs to find a vector perpendicular to a 3D surface (not to a contour line). Following the same idea as before, a surface in 3D can be thought as a constant valued "contour" surface (better known as a "level-set") of a 4D "hyper-surface". This hyper-surface can be easily obtained by expressing the equation for the original 3D surface in the form as follows $h(x, y, z) = c$ where c is a constant. Thus, $h(x, y, z)$ becomes a "contour" surface (or level set) of the function $w = h(x, y, z)$, where w is just a "dummy" variable name. The gradient $\nabla_{\mathrm{x,y,z}} w$ is then perpendicular to the 3D "contour" surface (or level set) from the 4D hyper-surface defined by w.

10. Recall that for a given function $f(x, y)$, the gradient $\nabla f(x, y)|_{x_o,y_o}$ must be perpendicular to the direction of no change at the point (x_o, y_o). One may ask what else we can learn from the direction of no change. Recalling that a minimum or a maximum of a function occurs when the derivative is zero, then when following a two-dimensional line tangent to a contour (for instance, an isotherm contour in a previous example) there will be a value of z corresponding to each point on the line. However when the line is tangent to a constant value contour, the value of z reaches a minimum (or a maximum) at the point of contact. This is because along the tangent line, on either side of the point of contact, the value of z will be larger (or smaller) than at the point of contact. This also implies that the directional derivative following the line vanishes (is zero) at the point of contact.

11. It is often convenient to rearrange the original 3D surface equation such that $c = 0$.

Exam Questions:

1. Derive the parametric equation of a line using a vector perpendicular to the line and a point on the line. Also derive a parametric equation for the line given two different points on the line.
2. Clearly explain the logic on how to find the equation to a plane (z as a function of x and y) using the normal vector and a point in the plane.

3. Clearly explain the process and the logic behind the method used to find the shortest distance from a plane to the origin and be able to find this distance for an explicit example such as: 3z+2y+x=4. Assume no other information is given to you.
4. Clearly explain the process and the logic behind the process of how to obtain the equation of a line normal to a surface in 3D at a given point p_o. Also, explain how to obtain the equation of a plane tangent to the same surface at the given point p_o.
5. Clearly explain the process and the logic behind the process on how to find the vector normal to a hyper-surface (i.e., in dimensions higher than 3) given by $f(\bar{x}) = constant$ at the point $\bar{x}_o$. Then explain how to find a hyper-plane tangent to the hyper-surface $f(\bar{x}) = constant$ at the point $\bar{x}_o$. Assume $\bar{x}$ is an N dimensional vector, i.e., $f(\bar{x})$ is a function that depends on N scalar variables that have been stacked into vector $\bar{x}$. Then find an explicit formula for the shortest "distance" from this tangential "plane" to the origin.
6. Explain why a minimum or a maximum of a function f(x,y) along a given line in the x-y plane occurs when the line is tangent to a constant value contour of f(x,y).

Sep. 7- Thursday
LECTURE 7: Normal vectors and tangential planes

Summary:

1. Review of Lecture 6 on how to obtain vectors normal to a surface by using the concept of gradient.
2. Review of Lecture 6 on how to obtain the equation of a plane given a point in the plane and a vector normal to the plane.
3. Example: plotting a circle of radius one, and centered at (1,1), then find and plot a vector normal to the circle at a point located on the circle perimeter at an angle of -45 degrees, and finally finding and plotting the line tangent to the circle at the same point.
4. Detailed example on how to use MathCad to create a 3D plot of a sphere with center located at (1,1,1). It was also shown how to use the theory explained in Lecture 6 to obtain vector normal to the sphere at a specific point p_o on the sphere, finally, it was shown how to include a plane tangential to the sphere at p_o.
5. Both examples above have been posted in a file named "Normal and Tangential Example" in the files & notes section of the course website.
6. If your 3D plots appear blank, remember that you may need to set the screen resolution to 16 bits. In Windows 7 this is done as follows:

   1) right click on a blank portion of the desktop and select "Properties"
   2) then select the "Settings" tab,
   3) under "Color quality" change the resolution to 16 bits.

   You may also need to select the borders of the 3D picture to be off: Right click on the 3D graph, and select "borders off"
7. If the point does not appear in the 3D plot, make sure the plot is set to "Data Points" in the General tab of the 3-D Plot Format menu obtained by double-left-clicking on the plot. If the point does not appear, retype the name of the point on the 3-D plot.

Exam Questions:

1. Review the exam questions from Lecture 6.
2. Describe in detail how to obtain the equation of a vector normal to a contour given in 2D at a point p_o located on the perimeter of the contour.
3. Describe in detail how to obtain the equation of the line tangent to the same contour in question 2 above and at the same point p_o.
4. Describe in detail how to obtain the equation of a vector normal to a surface (such as a sphere) given in 3D at a point p_o located on the perimeter of the 3D surface.
5. Describe in detail how to obtain the equation of the plane tangent to the same 3D surface in question 4 above and at the same point p_o.

Sep. 12- Tuesday
LECTURE 8: Vertical slice of a surface in 3D.

Summary:

1. Example of how to create a 3D plot of a decaying cosine wave and then spin the resulting wave around the z axis to obtain a surface plot that resembles a "Mexican hat". The cosine wave, $cos(r)$, can be forced to decay by multiplying the wave against a bell-shaped curve e^{-r^2} to obtain

$$z = cos(r) \cdot e^{-r^2}$$

If the plot decays too fast, then we can force it to decay slower by "stretching" or "widening" the bell-shaped curve as follows:

$$z = cos(r) \cdot e^{-(r/s)^2}$$

where s is the "stretching factor". For instance, a value of $s = 2$ will result in widening the bell-shaped curve by a factor of 2.
Spinning the resulting curve around the z axis is the same as making sure that the same z value results for all points that lie in a circle of radius r. In polar coordinates, one would write

$$z(r, \theta) = cos(r) \cdot e^{-(r/s)^2}$$

That is, the value of z is a function of the radial distance but not of the angular position. Converting back to rectangular coordinates by using $r = \sqrt{x^2 + y^2}$, the following equation is obtained.

$$z(x, y) = cos(\sqrt{x^2 + y^2}) \cdot e^{-\left(\frac{\sqrt{x^2+y^2}}{s}\right)^2}$$

2. How to obtain a 2D plot out of a vertical cut of a 3-D function along any specified direction. A 3D function can be expressed as $z = f(x, y)$ or $z = f(p)$ where p is a vector containing the x and y coordinates. To find the values of z corresponding to a line starting at a point $p_o = (x_o, y_o)$ and following the direction of the vector d one only needs to substitute $p = p_o + \alpha \cdot d$ where α is a scalar value. Thus, z becomes a function of α because the vectors v_o and d are constant, i.e.,

$$z(\alpha) = f\left(p_o + \alpha \frac{d}{|d|}\right)$$

Observe that here, vector d is normalized to be a unit vector so that the magnitude of α represents the length travelled along the direction of vector d.
A 2D plot of z(α) vs. α represents a cut through the 3D function along the direction d and through the point p_o.
The 2D direction vector d can be easily expressed in terms of angles as follows:

$$d = \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}$$

Observe, this latter vector is already normalized.

$$h(r) := \cos(r) \cdot e^{\frac{-r^2}{4}}$$

$$mhat(x, y) := h\left(\sqrt{x^2 + y^2}\right)$$

$$mexhat(p) := mhat\left(p_0, p_1\right)$$

$$xx_0 := 2 \qquad yy_0 := 2 \qquad zz_0 := mhat\left(xx_0, yy_0\right) \qquad p_o := \begin{pmatrix} xx_0 \\ yy_0 \end{pmatrix} \quad point := \begin{pmatrix} xx \\ yy \\ zz \end{pmatrix}$$

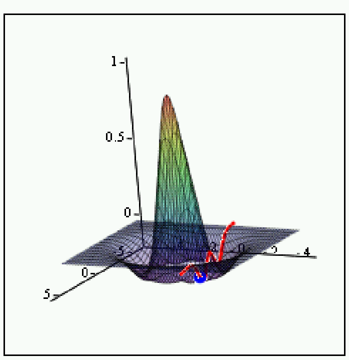
$$d(\theta) := \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$$

$$line2D(\alpha) := p_o + \alpha \cdot d(\theta \cdot deg)$$

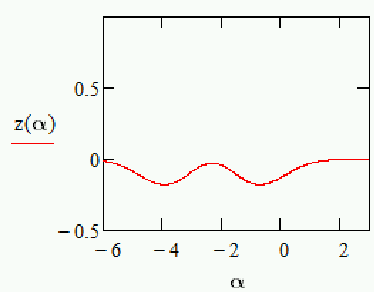
$$z(\alpha) := mexhat(line2D(\alpha))$$

$$z3D(\alpha) := \begin{pmatrix} line2D(\alpha)_0 \\ line2D(\alpha)_1 \\ z(\alpha) \end{pmatrix}$$

$$cutline := CreateSpace(z3D, -8, 4)$$



$$\theta \equiv 10$$



mhat , point , cutline

3. It was explained in class how to obtain the equation of a cone. The process is to first write the equation of a straight line through the origin using $z\ vs.\ r$ coordinates: $z = a \cdot r$. Then the value of r as seen in the $x - y$ coordinates is $r = \sqrt{x^2 + y^2}$. Substituting the equation for r in the equation of the straight line results in the equation for the cone: $z(x, y) = a \cdot \sqrt{x^2 + y^2}$.
4. A pdf file on how to slice a cone in 3d was placed in the Files and Notes section of the website. This information was also discussed in detail during class.
5. A pdf file on how to intersect two cylinders in 3d was placed in the Files and Notes section of the website.
6. A pdf file on how to slice a sphere in 3d was placed in the Files and Notes section of the website. This was not discussed in class.

Exam Questions:

1. Clearly explain the process and the logic used to create a "Mexican hat" function starting from a cosine function.
2. Explain the process and the logic behind the process of how to obtain the equation of a line passing through point p_o and following the direction of vector d. Do not assume that the points are two or three-dimensional.
3. Explain the process and the logic behind the process of how to obtain the profile of the function z(x,y) resulting from a vertical cut passing through point p_o and following the direction of vector d.

Sep. 14- Thursday
LECTURE 9: Intersecting planes and surfaces, Linearity

Summary:

1. A pdf file on how to slice a sphere in 3d was placed in the Files and Notes section of the website. This information was also discussed in detail during class.
2. A pdf file on how to slice a cone in 3d was placed in the Files and Notes section of the website. This information was also discussed in detail during class.
3. A pdf file on how to intersect two cylinders in 3d was placed in the Files and Notes section of the website.
4. A pdf file ("Parametrically Joining Two Solutions.pdf")with an example on how to join two solutions obtained when the plus and minus sign of a square root solution appears by algebraically solving for an intersection of two surfaces (if possible) is given in the Files and Notes section of the website.
5. Review of how to obtain normal vectors to a surface, tangential planes to a surface, and how to find plot points at the intersection of two surfaces.
6. Superposition is commonly used in engineering and is the main idea behind the definition of linearity.
7. Most tractable mathematics used in engineering practice is based on the use of linearity concepts. Thus, the understanding of linearity concepts, i.e., linear algebra, is a central topic in this course.

8. Linearity = superposition, homogeneity & additivity. Homogeneity is described mathematically as $f(cx) = cf(x)$ where c is a scalar and x is a vector. Additivity is described mathematically as follows: $f(x+y) = f(x) + f(y)$.
9. Homogeneity implies that a linear function must pass through the origin: when $c = 0$ then $f(0) = 0$. A common mistake is to think that every straight line is a linear function: a straight line is a linear function only if it passes through the origin.

Exam Questions:

1. Explain the process and the logic behind the process of how generate the equation of cone in 3D.
2. Explain the process and the logic behind the process of how finding the intersection of a plane a 3D surface using a numerical solver. For instance, Given and Find in MathCad.
3. Given a function f(x), explain how to test that this function is linear.
4. Provide a logical explanation on why a straight line that does not pass through the origin cannot be used to represent a linear function.

Sep. 19- Tuesday
LECTURE 10: Linearity – Solving differential equations

Summary:

1) Consider a differential equation $a\,y''' + b\,y'' + c\,y' + d\,y = z(x)$.
In this differential equation the input is $z(x)$ and the resulting output is $y(x)$. If a, b, c, and d are constants or only functions of x, then it is easy to prove that the differential equation is linear. First, the solution process mapping the input $z(x)$ to the resulting output $y(x)$ is labeled $L(\cdot)$, i.e., $y(x) = L(z(x))$. It is not necessary to know the details of how to obtain $L(\cdot)$ to prove linearity. It will be shown later, that knowing the solution to a series of functions that can serve as coordinate vectors for the space of interest is equivalent to knowing $L(\cdot)$ as applied to that space.

The solution process $L(\cdot)$ is linear if given inputs z_1 and z_2 with corresponding outputs $y_1 = L(z_1)$ and $y_2 = L(z_2)$, then $L(\alpha z_1 + \beta z_2) = \alpha y_1 + \beta y_2$ for any two scalars α and β. For example, consider

$$a\,\frac{d^3y_1}{dx^3} + b\,\frac{d^2y_1}{dx^2} + c\,\frac{dy_1}{dx} + d\,y_1 = z_1(x) \quad \text{and}$$

$$a\,\frac{d^3y_2}{dx^3} + b\,\frac{d^2y_2}{dx^2} + c\,\frac{dy_2}{dx} + d\,y_2 = z_2(x)$$

multiplying the first equation by α and the second equation by β and adding the two equations results in

$$a\,\frac{d^3(\alpha y_1+\beta y_2)}{dx^3} + b\,\frac{d^2(\alpha y_1+\beta y_2)}{dx^2} + c\,\frac{d(\alpha y_1+\beta y_2)}{dx} + d\,(\alpha y_1 + \beta y_2) = \alpha z_1(x) + \beta z_2(x)$$

The above, shows that if the input is $(\alpha z_1 + \beta z_2)$ then the output is $(\alpha y_1 + \beta y_2)$, or, equivalently, the output is $(\alpha L(z_1) + \beta L(z_2))$. This shows that the solution process, $L(\cdot)$, to this differential equation is linear, i.e., this is a linear differential equation.

2) Review of undergraduate differential equations: The complete solution of a linear differential equation with constant coefficients can be done in many different ways. Two possible approaches are listed below.

   a) One way is to first solve for the particular and the homogeneous solutions; the final or complete solution is the combination of both solutions. The initial conditions are applied to the complete solution in order to find the values of the constants that arise when solving for the homogeneous solution.

   b) Another way is to first solve for the "free" and the "forced" responses and add both responses to obtain the complete or final solution.

      i) The "free" response is the solution to the differential equation assuming that all the inputs are zero (not the initial conditions). The free response is also known as the "zero input" response.

      ii) The "forced" response, a.k.a. the zero initial conditions response, is the solution to the differential equation assuming zero initial conditions (not the inputs).

3) The advantage of using the "free" and "forced" response method is that this method completely separates the contributions due to the inputs from the natural response of the system due to the initial conditions. In many applications, the input to a system is actually unknown while the output is measured. The solution process to these problems is to first model the system using a differential equation(s), measure the response (the output) of the system starting from zero initial conditions, and then find the input to the differential equation. The methodology for doing this will be shown later in the course.

4) It is important to remember that for a linear differential equation with input $f(t)$ and output $x(t)$, the notation $x(t) = L(f(t))$ should be interpreted as: "the solution of the differential equation to the input $f(t)$ is $x(t)$."

5) Similarly, $x(t) = L\big(af_1(t) + bf_2(t)\big) = aL\big(f_1(t)\big) + bL\big(f_2(t)\big)$ reads: "the solution of the linear ODE to input $af_1(t) + bf_2(t)$ is the weighted sum of two solutions, one obtained by solving the ODE with $f_1(t)$ as the input and the other one obtained by solving the ODE with $f_2(t)$ as the input. The weights are a and b.

6) The method of Laplace Transforms can be used to show by means of partial fraction expansions that, when the input consists of the sum of several functions, then the output will be the sum of the contributions of each individual forced solution, i.e., linearity applies to the "forced" solutions. It is important to remember that Laplace Transforms only apply to linear differential equations with constant coefficients.

7) Linearity can be used to convert a linear differential equation into a simple matrix multiplication. For example, consider the differential equation

$$a\,\frac{d^2y}{dx^2} + c\,\frac{dy}{dx} + c\,y \;=\; f(x)$$

where

$$0 \le x \le \mathcal{L}$$

this equation can be expressed as

$$y(x) = L(f(x))$$

implying that the <u>forced</u> response $y(x)$ is a linear function of the input $f(x)$. A simple Fourier Series (orthogonal) expansion of $f(x)$ is given by

$$f(x) \cong \sum_{i=0}^{N} \big(A_i \sin(\omega_i x) + B_i \cos(\omega_i x)\big)$$

where

$$\omega_i = \frac{2\pi}{\mathcal{L}} i$$

and the coefficients A_i and B_i can be obtained by using inner products and taking advantage of the orthogonality between the sinusoidal functions,

$$A_i = \frac{\big(f(x), \sin(\omega_i x)\big)}{\big(\sin(\omega_i x), \sin(\omega_i x)\big)} \quad and \quad B_i = \frac{\big(f(x), \cos(\omega_i x)\big)}{\big(\cos(\omega_i x), \cos(\omega_i x)\big)}$$

with the inner product defined as $\big(u(x), \mathrm{v}(x)\big) = \int_0^{\mathcal{L}} u(x)\mathrm{v}(x)dx$

It follows that

$$y(x) = L\big(f(x)\big) \cong L\left(\sum_{i=0}^{N} \big(A_i \sin(\omega_i x) + B_i \cos(\omega_i x)\big)\right)$$

Applying linearity,

$$y(x) \cong \sum_{i=0}^{N} \left(A_i \overbrace{L\big(\sin(\omega_i x)\big)}^{LS_i} + B_i \overbrace{L\big(\cos(\omega_i x)\big)}^{LC_i} \right)$$

where LS_i and LC_i represent the forced response of the differential equation to $\sin(\omega_i x)$ and $\cos(\omega_i x)$, respectively. In practice, one could obtain LS_i and LC_i for i=0..N by solving the respective differential equations numerical, e.g., using MathCad, Matlab, Fortran, etc., and applying zero initial conditions.

8) Linearity can be used to convert a linear differential equation into a simple matrix multiplication problem as follows. First assume that the input is a weighted sum (a.k.a. linear combination) of a finite number of functions such as Legendre Polynomials, Chebyshev Polynomials, or Fourier series. Then use linearity to express the solution in terms of the response to each individual polynomial or function used in the weighted

sum representing the input. The following example (not discussed in class) illustrates this concept as applied to the particular solution of a linear differential equation.

The particular solution to a differential equation of the form

$$\dot{x} + x = f(t)$$

can be expressed as a linear function

$$x(t) = L(f(t)).$$

Now, if

$$f(t) = a + bt + ct^2$$

then linearity implies that

$$x(t) = aL(1) + bL(t) + cL(t^2)$$

Define

$$x_1(t) = L(1), \;\; x_2(t) = L(t), \;\; \text{and} \;\; x_3(t) = L(t^2)$$

x_1, x_2, and x_3 can be found as follows:

For $\dot{x}_1 + x_1 = 1$ assume that $x_1 = a$. It follows that $\dot{x}_1 = 0$ and substituting back into the differential equation yields $0 + a = 1$ or $a = 1$. It follows that $x_1 = 1$.

For $\dot{x}_2 + x_2 = t$ assume that $x_2 = bt + a$. It follows that $\dot{x}_2 = b$ and substituting back into the differential equation yields $(b) + (bt + a) = t$. Equating similar polynomial coefficients on both sides of the equations yields $\mathrm{b} = 1$ for the coefficients of t and $a = -b = -1$ for the constant terms. It follows that $x_2 = t - 1$.

For $\dot{x}_3 + x_3 = t^2$ assume that $x_3 = ct^2 + at + b$. It follows that $\dot{x}_3 = 2ct + a$ and substituting back into the differential equation yields $(2ct + a) + (ct^2 + at + b) = t^2$. Equating similar polynomial coefficients on both sides of the equations yields $c = 1$ for the coefficients of t^2,
$a = -2c = -2$ for the coefficients of t, and $b = -a = 2$ for the constant terms. It follows that $x_3 = t^2 - 2t + 2$.

Linearity implies that if $\dot{x} + x = a + bt + ct^2$ then $x = ax_1 + bx_2 + cx_3$ or

$$x = (a)(1) + (b)(t - 1) + (c)(t^2 - 2t + 2)$$

Rearranging,

$$x = (a - b + 2c) + (b - 2c)t + (c)t^2$$

In order to simplify the solution process the coefficients of each of the three solutions (x_1, x_2, and x_3) are expressed in vector form as:

$$x_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad x_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \quad x_3 = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$$

It follows that the particular solution of $\dot{x} + x = a + bt + ct^2$ is given by $x(t) = \alpha + \beta t + \gamma t^2$ where

$$\begin{bmatrix} \gamma \\ \beta \\ \alpha \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} c \\ b \\ a \end{bmatrix}$$

The above is an example of how a linear function between two finite dimensional spaces can be expressed as a matrix multiplication. In this case the solution process $L(\cdot)$ to a differential equation is the linear function, and the three independent input functions ($1, t$, and t^2) correspond to the vectors used to define a three dimensional space. The coordinates of a three dimensional point corresponds to a particular choice of coefficients $c, b,$ and a for quadratic polynomial used as the input. The particular solution can also be represented as a quadratic polynomial with coefficient represented by another three dimensional point with coordinates corresponding to the values of $\gamma, \beta,$ and α.

Exam Questions:

1. Given the following differential equation $a(t)\,\frac{d^2x}{dt^2} + b(t)\,\frac{dx}{dt} + c(t)\,x = f(t)$. Provide a thorough proof showing that this is a linear differential equation.
2. For question 1, explain in plain language what the expression $x(t) = L(f(t))$ means.
3. Given the following orthogonal expansion $y(x) \cong \sum_{i=0}^{N}\left(\alpha_i \sin(\omega_i x) + \beta_i \cos(\omega_i x)\right)$ with $0 \le x \le \mathcal{L}$ and $\omega_i = \frac{2\pi}{\mathcal{L}} i$, provide a thorough explanation on how to obtain α_i and β_i. Merely showing formulas is not enough, you must explain how the formulas are obtained using basic concepts such as inner product.
4. For question 3 above, explain why β_0 is the mean value of $y(x)$ with $0 \le x \le \mathcal{L}$.
5. For the same inner product used in questions 2 and 3, explain why $(1,1) = \mathcal{L}$. Note: the 1 is a constant function.
6. Give the differential equation $a\,\frac{d^2y}{dx^2} + c\,\frac{dy}{dx} + c\,y = f(x)$, provide a thorough explanation on how to obtain an approximate solution using an orthogonal expansion of $f(x)$ and the concept of linearity.
7. Explain how linearity can be used to solve for the particular solution of a differential equation using a Fourier Series expansion of the input function.
8. Consider a linear differential equation with constant coefficients, explain in detail what is the free response and what is the forced response.
9. Consider a linear differential equation with constant coefficients, provide a thorough explanation about the difference between the particular solution and the forced response. Also, explain the difference between the free response and the homogeneous solution.

Sep. 21- Thursday
LECTURE 11: Laplace Transforms Review, linear functions mapping between finite dimensional spaces

- Review of Laplace Transforms and partial fraction expansions:

1. Laplace Transforms are used to transform differential equations into algebraic equations, which can then be easily solved for the output. After solving for the output using algebra, the Laplace Transform is inverted (reversed) to obtain the solution as a function of time. Laplace Transforms replaces the independent variable, t, with a new variable, s. Thus, it is said that the Laplace Transforms changes a variable from the t-domain to the s-domain, i.e., from the time domain to the Laplace domain.
2. Notation: Capital letters are usually used to denote variables in the s-domain, small letters are used for variables in the t-domain (time). Thus a time function x has a Laplace Transform X, i.e., X = L(x). It should be understood that x=x(t) and X=X(s).
3. The Laplace Transform of the derivative term is L(dx/dt) = s·X - x(0).
4. The Laplace Transforms of higher order derivatives can be found by repeated application of the Laplace transform of the first order derivative. Thus, the Laplace Transform of d^2x/dt^2 is obtained by multiplying the Laplace Transform of dx/dt by s and subtracting the new initial condition for dx/dt: L(d^2x/dt^2)= s·L(dx/dt) - x'(0) = s·(sX-x(0)) - x'(0) = s^2·X - s·x(0) - x'(0).
5. With the initial conditions set to zero, multiplying by "s" in the Laplace domain is equivalent to taking the derivative with respect to time in the "t" domain (time domain). It follows that integration in the "s" domain must be achieved by dividing by "s", i.e., multiplying by 1/s.
6. A term of the form e^{-at} is called a "transient" because it eventually goes to zero as time increases.
7. e^{-at} can be written as $e^{-t/\tau}$ where the constant $\tau = 1/a$ is called the "time constant".
8. When t = 4.6τ then $e^{-t/\tau}$ =$e^{-4.6}$ = 0.01. Thus, a transient term "dies out" in a period of time equal to about 4.6 time constants. For this class, a transient term is said to have died-out when it reaches within 1% of 0.01 of its initial value.
9. The Laplace Transform of the Unit step function $\Phi(t)$ also known as the Heaviside Function is L(Φ(t)) = 1/s.
10. A step function of magnitude M is given by M·Φ(t) and its Laplace Transform is M/s.
11. Make sure you **memorize** the Laplace Transforms of the exponential e^{-at} as it is "the most important transform" to be discussed in this course: L($e^{-a\cdot t}$) = 1/(s+a).
12. Examples of several first order differential equations: mass-damper, spring-damper, submerging a solid mass into a constant temperature bath, and a capacitor in parallel with a resistor. The first order differential equations were expressed as $A \cdot \dot{x} + B \cdot x = 0$. Using Laplace Transforms the equation became: $A \cdot (s \cdot X - x(0)) + B \cdot X = 0$. Solving for X, we obtained: $X = \frac{A \cdot x(0)}{A \cdot s + B}$. Dividing numerator and denominator by A resulted in $X = \frac{x(0)}{s + B/A}$ and inverting the Laplace Transform, i.e., using the Laplace Transform table backwards, resulted in $x(t) = x(0) \cdot e^{-a \cdot t}$ where a=B/A and the time constant is τ=1/a= A/B. Thus, the time constant of a first order differential equation can be readily found by dividing the coefficient A by the coefficient B.
13. Partial Fraction expansion is an algebraic technique that reduces a fraction with a high order polynomial in the denominator into a sum of simple fractions with first order polynomials in the denominator and a constant in the numerator. These "simple" fractions in the s-domain become exponentials in the time domain.

14. It is very important to know how to factorize the polynomials in order to solve differential equations through partial fraction expansions.
15. When factorizing a polynomial, it is the NEGATIVE of the roots of the polynomial that is used in each factor.
16. The method of Laplace Transforms can be used to show that, when the input consists of the sum (addition) of several functions, then the output will be the sum of the individual forced solutions to each of the additive terms, i.e., linearity applies to the "forced" solutions. It is important to remember that Laplace Transforms only apply to linear differential equations with constant coefficients.
17. Every linear function between two finite dimensional spaces can be expressed as a matrix multiplication. The proof is as follows. Let $f(\cdot)$ be a linear function mapping points from R^n to points in R^m. This means that if one writes $y = f(x)$, then y is an m-dimensional vector x is an n-dimensional vector. Now let

    $e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{bmatrix}$,..., and $e_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$ be a set of n-dimensional unit vectors.

    The vectors $e_1, e_2, \cdots, e_n$ are known as the "natural" basis vectors.

    Applying the linear function $f(\cdot)$ to the unit vectors above, yield the following set of m-dimensional vectors

    $y_1 = f(e_1), \ldots, y_n = f(e_n)$

    A point $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ in R^n can be expressed as $x = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n$. Using the linearity

    properties of $f(\cdot)$ one can write

    $f(x) = f(x_1 e_1 + \cdots + x_n e_n) = x_1 f(e_1) + \cdots + x_n f(e_n) = x_1 y_1 + \cdots + x_n y_n$

    or

    $f(x) = [y_1 \quad \cdots \quad y_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

    But recall that $y_1, \ldots, y_n$ are each an m-dimensional vector. Therefore

    $M = [y_1 \quad \cdots \quad y_n]$ is an mxn matrix and it follows that

    $f(x) = Mx$ where $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$.

    Thus every linear function between two finite dimensional spaces can be expressed as a matrix multiplication !

Exam Questions:

1. Provide a constructive proof for the fact that every linear function between two finite dimensional spaces can be expressed as a matrix multiplication, i.e., show how to find the matrix.
2. Prove that it takes a time equal to 4.6 time constants to reach within 1% of the initial value of a first order linear differential equation with constant coefficients.