

Topics In Linear Algebra and Its Applications

James Emery

6/16/2010

Contents

0.1	Introduction	2
0.2	Linear Transformations and Linear Operators	6
0.3	Determinants	7
0.4	Permutations	9
0.5	Kernel and Range	9
0.6	Quotient Spaces	10
0.7	Direct Sums	10
0.8	The Inner Product	10
0.9	Inner Product Spaces	10
0.10	Orthonormal Vectors	11
0.11	Inequalities	11
0.11.1	The Cauchy-Schwarz Inequality	11
0.11.2	Bessel's Inequality	12
0.11.3	The Triangle Inequality	14
0.12	Gram-Schmidt Orthogonalization	15
0.13	The Parallelogram Law	15
0.14	Quadratic Forms	16
0.15	Canonical Forms	17
0.16	Upper Triangular Form	17
0.17	Isometries, Rotations, Orthogonal Matrices	17
0.18	Rotation Matrices	18
0.19	Exponentials of Matrices and Operators in a Banach Algebra	18
0.20	Eigenvalues	19
0.21	The Characteristic Polynomial and the Cayley-Hamilton Theorem	20
0.22	Unitary Transformations	21
0.23	Transpose, Trace, Self-Adjoint Operators	21

0.24	The Spectrum	22
0.25	The Spectral Theorem	22
0.26	Tensors	23
0.27	Application of Linear Algebra to Vibration Theory, Normal Coordinates	23
	0.27.1 A Simple Spring Example	23
	0.27.2 Decoupling Equations	24
	0.27.3 Example: Coupled Oscillators	27
	0.27.4 The General Problem of Linear Vibration	37
0.28	Projection Operators	38
0.29	Functional Analysis	38
0.30	Hamel Basis	39
0.31	Numerical Linear Algebra	39
0.32	Quantum Mechanics	39
0.33	The Schrodinger Wave Equation	40
0.34	The Postulates of Quantum Mechanics	41
0.35	The Bra and Ket Notation of Dirac	42
0.36	Example: The Hydrogen Atom	42
0.37	The Relation Between Linear Algebra, Functional Analysis, and Abstract Solutions to Problems	43
0.38	Bibliography	43

0.1 Introduction

Linear algebra is the study of finite dimensional vector spaces and linear transformations. A vector space is a quadruple $(V, F, +, *)$. V is a set of vectors, F is a field of scalars, $+$ is the operation of vector addition, and $*$ is the operation of scalar multiplication. We usually do not write the multiplication operator, that is, we write $\alpha * v$ as αv . Let $\alpha, \beta \in F$ and $u, v, w \in V$. The following axioms are satisfied:

1. $u + v = v + u$
2. $u + (v + w) = (u + v) + w$
3. There is a zero element so that $0 \in V$ so that $u + 0 = u$.
4. For each $u \in V$, there is an inverse element $-u$ so that $u + (-u) = 0$

5. $\alpha(u + v) = \alpha u + \alpha v$.
6. $(\alpha + \beta)u = \alpha u + \beta u$.
7. $(\alpha\beta)u = \alpha(\beta u)$.
8. $1u = u$

A finite set of vectors v_1, v_2, \dots, v_n is linearly independent if

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0$$

implies that each α_i is zero. Otherwise the set is called linearly dependent. A subset S of V is a subspace of V if the sum of any two elements in S is in S and the scalar product of any element in F with any element in S is in S . That is S is closed under addition and scalar multiplication. The subspace spanned by vectors

$$v_1, v_2, \dots, v_n$$

is

$$S = \{\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n : \alpha_i \in F\}$$

Theorem The nonzero vectors v_1, v_2, \dots, v_n are linearly dependent if and only if some one of them is a linear combination of the preceding ones.

Proof. Suppose v_k can be written as a linear combination of v_1, \dots, v_{k-1} . Then we have a linear combination of v_1, \dots, v_k set equal to zero with $\alpha_k = -1$, so that these vectors are linearly dependent. Conversely, suppose v_1, \dots, v_n are dependent. Then we can find a set of α_i so that

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0,$$

and at least one of them is not zero. Let k be the largest index for which α_i is not zero, then dividing by α_k we find that v_k is a linear combination of the preceding vectors.

Corollary. Any finite set of vectors contains a linear independent subset that spans the same space.

Theorem. Let vectors v_1, v_2, \dots, v_n span V . Suppose the vectors u_1, u_2, \dots, u_k are linearly independent. Then $n \geq k$.

Proof. The set $u_1, v_1, v_2, \dots, v_n$ is linearly dependent and spans V , so some v_j is dependent on its predecessors. Then the set $u_1, v_1, v_2, \dots, v_{j-1}, v_{j+1}, \dots, v_n$

spans V , and is dependent. We may continue this, adding a u_i while removing a v_j and still having a set that spans V and is dependent. This can be continued until the u_i are exhausted, otherwise the v_j would be exhausted first and some subset of the u_1, u_2, \dots, u_k would then be dependent, which is not possible. Therefore there are more v_j than u_i , which forces

$$n \geq k.$$

Definition A *basis* of a vector space V is a set of linearly independent vectors that spans V . A vector space is finite dimensional if it has a finite basis.

Theorem Suppose a vector space has a finite basis

$$A = \{v_1, v_2, \dots, v_n\}.$$

Then any other basis also has n elements.

Proof Let $B = \{u_1, u_2, u_3, \dots\}$, which is possibly infinite, be a second basis of V . By the previous theorem $n \geq k$ for any subset

$$u_1, u_2, \dots, u_k$$

of B . It follows that

$$B = \{u_1, u_2, u_3, \dots, u_m\},$$

for some m , and that $n \geq m$. Reversing the role of A and B , we apply the previous theorem again to get $m \geq n$, which proves the corollary. We conclude that the dimension of a finite dimensional vector space can be well defined as the number of elements in any basis. Any vector $v \in V$ can be represented as a linear combination of the basis elements:

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \dots + \alpha_n v_n.$$

This representation is unique, because if we subtract two different representations we would get a representation of the zero vector as a linear combination of the basis vectors with at least one nonzero coefficient, which contradicts that the vectors are linearly independent.

The scalar coefficients are called the coordinates of v , and form an element in the cartesian n -product of the scalar field F . These n -tuples of scalars themselves form an n dimensional vector space and are isomorphic to the original vector space. Let U and V be two vector spaces. A function

$$T : U \rightarrow V$$

is called a linear transformation if

1. For $u_1, u_2 \in U, T(u_1 + u_2) = T(u_1) + T(u_2)$.
2. For $\alpha \in F, u \in U, T(\alpha u) = \alpha T(u)$.

A linear transformation from U to itself, is called a linear operator. Associated with every linear transformation is a matrix. Let

$$\{u_1, u_2, u_3, \dots, u_n\}$$

be a bases of U and let

$$\{v_1, v_2, \dots, v_m\}$$

a bases of V . For each $u_j, T(u_j)$ is in V , so that it may be written as a linear combination of the basis vectors, we have

$$T(u_j) = \sum_{i=1}^m a_{ij}v_i.$$

Now let $u \in U$ and let its coordinates be x_1, \dots, x_n . Vector u is represented by the coordinate vector

$$\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

We have

$$\begin{aligned} T(u) &= T\left(\sum_{j=1}^n x_j u_j\right) = \sum_{j=1}^n x_j T(u_j) \\ &= \sum_{j=1}^n x_j \sum_{i=1}^m a_{ij} v_i \\ &= \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j\right) v_i \\ &= \sum_{i=1}^m y_i v_i, \end{aligned}$$

where the y_1, y_2, \dots, y_m are the components of the vector $T(u)$ in vector space V . We have shown that the coordinate vector x of u is mapped to the coordinate vector y of $T(u)$ by matrix multiplication,

$$\begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

Now suppose we have two linear transformations

$$T : U \rightarrow V,$$

and

$$S : V \rightarrow W.$$

The composite transformation is

$$ST : U \rightarrow W,$$

defined by $ST(u) = S(T(u))$.

Theorem. If A is the matrix of linear transformation S , and B is the matrix of linear transformation T , then the matrix multiplication product AB is the matrix of ST .

0.2 Linear Transformations and Linear Operators

Examples:

Finite dimensional linear transformation

$$\begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

Integral linear transformation

$$f \rightarrow g,$$

where

$$g(y) = \int k(x, y)f(x)dx$$

Differential operator

$$f \rightarrow g,$$
$$g(x) = \left(5\frac{d^2}{dx^2} + 2\frac{d}{dx} - 3\right)f(x)$$

0.3 Determinants

A determinant is a multilinear functional defined on a square matrix. A linear functional is a mapping from a vector space to a field of scalars. A multilinear functional is a function defined on a cartesian product of the vector space. The column vectors of a matrix may be considered to be vectors of the vector space V . The set of n column vectors constitute a point in the cartesian product. The functional is linear in the sense that

$$f(v_1, v_2, \dots, v_k + v'_k, \dots, v_n) = f(v_1, v_2, \dots, v_k, \dots, v_n) + f(v_1, v_2, \dots, v'_k, \dots, v_n),$$

and

$$f(v_1, v_2, \dots, \alpha v_k, \dots, v_n) = \alpha f(v_1, v_2, \dots, v_k, \dots, v_n).$$

A multilinear functional is alternating if interchanging a pair of variables changes the sign of the function.

Definition. A determinant $D(A)$ of an n dimensional square matrix A is the unique alternating multilinear functional defined on the n column vectors of A , which takes value 1 on the identity matrix.

Properties.

1. If two column vectors of a matrix are identical, then the determinant is zero. This follows because interchanging the columns changes the sign of the determinant, but the new matrix has not changed, so the value of the determinant is the same. The determinant must be zero.
2. Adding a multiple of one column to a second does not change the value of the determinant. This is clear from

$$\begin{aligned} & D(v_1, \dots, v_i, \dots, v_j + \alpha v_i, \dots, v_n) \\ &= D(v_1, \dots, v_i, \dots, v_j, \dots, v_n) + \alpha D(v_1, \dots, v_i, \dots, v_i, \dots, v_n) \\ &= D(v_1, \dots, v_i, \dots, v_j, \dots, v_n) + 0. \end{aligned}$$

Example To compute the determinant of

$$\begin{bmatrix} 2 & 2 \\ 3 & 4 \end{bmatrix}$$

subtract the first column from the second, then three times the second from the first, getting

$$\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

So $D(A) = 2D(I) = 2$, where I is the identity matrix. Once we have a matrix in diagonal form, we see from its definition as a multilinear functional, that the determinant is equal to the product of each multiplier of each column times the determinant of the identity. That is, the value equals the product of the diagonal elements.

Example *Cramers's Rule* Suppose we have a system of n equations in n unknowns x_1, x_2, \dots, x_n written in the form

$$x_1v_1 + x_2v_2 + \dots, x_nv_n = v.$$

We have

$$\begin{aligned} D(v, v_2, v_3, \dots, v_n) &= D(x_1v_1 + x_2v_2 + \dots + x_nv_n, v_2, \dots, v_n) \\ &= x_1D(v_1, v_2, v_3, \dots, v_n). \end{aligned}$$

So that unknown x_1 is given by

$$x_1 = \frac{D(v, v_2, v_3, \dots, v_n)}{D(v_1, v_2, v_3, \dots, v_n)}.$$

There is clearly a similar expression for each of x_2, \dots, x_n .

To compute a determinant we can perform permutations on the columns and add scalar multiples of columns to other columns, to put the matrix into diagonal form. Once in diagonal form, (or triangular form) the determinant equals the product of the diagonal elements.

There is an alternate definition of the determinant involving permutations. Let a be a n by n matrix. Consider the sum

$$\sum_{\sigma} s(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\dots a_{n\sigma(n)}$$

where σ is a permutation of the integers $1, 2, 3, 4, \dots, n$ and $s(\sigma)$ is the sign of the permutation. The sign of the identity permutation is one, and interchanging a pair of elements, a transposition, changes the sign of the permutation. Notice that this is a multilinear functional, and is alternating, and further equals one on the identity matrix. Therefore it is the unique such functional, and so is equal to the determinant.

Properties.

1. $D(A^T) = D(A)$.
2. $D(AB) = D(A)D(B)$.
3. Expansion by minors about row i . Let A_{ij} be the matrix obtained from A by deleting row i and column j . Then

$$D(A) = \sum_{j=1}^n (-1)^{i+j} a_{i,j} D(A_{ij}).$$

0.4 Permutations

See the paper **Group Theory** by James Emery (Document groups.tex).

0.5 Kernel and Range

Let U be a n dimensional vector space. Let T be a linear transformation from U to a vector space V . Let $K(T)$ be the kernel of T and $R(T)$ the range of T . Then

$$\dim(K(T)) + \dim(R(T)) = n$$

Let u_1, u_2, \dots, u_p be a basis of the kernel of T . It can be extended to be a full basis of U . Suppose V is also n dimensional. Let A be the matrix of T with respect to this basis. Then the first column of A must be all zeroes, so that $D(A)$ is zero. Let U have a second basis and let S be the linear transformation mapping the first bases to the second. Let B be the matrix of S . Then B has an inverse B^{-1} and

$$D(B)D(B^{-1}) = D(BB^{-1}) = D(I) = 1.$$

Then $D(B)$ is not zero. The matrix of T with respect to the second basis is $D(AB) = D(A)D(B)$, so that in general the determinant of the matrix of T with respect to bases of U and V is zero if and only if the kernel of T is not zero, and this is true if and only if T has an inverse.

0.6 Quotient Spaces

0.7 Direct Sums

0.8 The Inner Product

The inner product of two vectors x and y in a complex or real vector space is written (x, y) . The dot product of vector analysis is an inner product. The inner product has the following properties:

1. $(x, y) = \overline{(y, x)}$,
2. $(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1(x_1, y) + \alpha_2(x_2, y)$,
3. $(x, x) \geq 0$; $(x, x) = 0$, if and only if $x = 0$.

From these properties we get

$$(x, \alpha y) = \overline{(\alpha y, x)} = \overline{\alpha(y, x)} = \overline{\alpha} \overline{(y, x)} = \overline{\alpha}(x, y)$$

An example of an inner product on a vector space of complex valued continuous functions of a real variable, say defined on an interval $[a, b]$ is

$$(f, g) = \int_a^b f(x)\bar{g}(x)dx.$$

0.9 Inner Product Spaces

Let V be a vector space with an inner product (u, v) . A basis can be used to construct an orthonormal basis (Gram-Schmidt Orthogonalization). An inner product space is also a normed linear space with the norm

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

The Cauchy-Schwartz inequality is

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

The triangle inequality is

$$\|u + v\| \leq \|u\|^2 + \|v\|^2.$$

0.10 Orthonormal Vectors

A set of vectors $S = \{x_i\}$ is an orthonormal set if each pair is orthogonal

$$(x_i, x_j) = 0,$$

and each vector has unit norm

$$(x_i, x_i) = \|x_i\| = 1.$$

If x is an arbitrary vector and $S = \{x_i\}$ is an orthonormal set, then the coefficients

$$c_i = (x, x_i),$$

are called the Fourier coefficients of x with respect to the the set S . If x can be expressed as a linear combination of the orthonormal vectors S , then there is a Fourier expansion of x

$$x = \sum_{i=1}^n c_i x_i.$$

Examples of orthonormal sets include the trigonometric functions and the various orthogonal polynomials used widely in applied mathematics.

0.11 Inequalities

0.11.1 The Cauchy-Schwarz Inequality

Given two vectors x and y , we have

$$0 \leq \|x - \lambda y\|$$

$$\begin{aligned}
&= (x - \lambda y, x - \lambda y) \\
&= (x, x) + (x, -\lambda y) + (-\lambda y, x) + (-\lambda y, -\lambda y) \\
&= \|x\|^2 - \bar{\lambda}(x, y) - \lambda(y, x) + |\lambda|^2 \|y\|^2.
\end{aligned}$$

Let

$$\lambda = \frac{(x, y)}{(y, y)}$$

Then

$$\bar{\lambda} = \frac{(y, x)}{(y, y)}.$$

The inequality becomes

$$0 \leq \|x\|^2 - \frac{|(x, y)|^2}{\|y\|^2}.$$

So

$$|(x, y)|^2 \leq \|x\|^2 \|y\|^2,$$

or

$$|(x, y)| \leq \|x\| \|y\|.$$

This is the Cauchy-Schwarz inequality. If x is a multiple of y , say

$$x = \alpha y,$$

then

$$\|(x, y)\| = \|\alpha\| \|y\|^2 = \|\alpha\| \|y\| \|y\| = \|x\| \|y\|.$$

So we have equality. Conversely, if we have equality, x is a multiple of y .

$$|(x, y)| \leq \|x\| \|y\|.$$

0.11.2 Bessel's Inequality

Proposition If $S = \{x_i, i = 1..n\}$ is an orthonormal set, then

$$\sum_{i=1}^n |(x, x_i)|^2 \leq \|x\|^2.$$

Proof. We have

$$\begin{aligned}
0 &\leq \left\| x - \sum_{i=1}^n (x, x_i) x_i \right\|^2 \\
&= \left(x - \sum_{i=1}^n (x, x_i) x_i, x - \sum_{j=1}^n (x, x_j) x_j \right) \\
&= \|x\|^2 - \sum_{j=1}^n (x, x_j)(x, x_j) - \sum_{i=1}^n (x, x_i)(x_i, x) + \sum_{i=1}^n \sum_{j=1}^n ((x, x_i)(x, x_j)(x_i, x_j)) \\
&= \|x\|^2 - \sum_{j=1}^n (x, x_j)(x, x_j) - \sum_{i=1}^n (x, x_i)(x_i, x) + \sum_{j=1}^n ((x, x_j)(x, x_j)) \\
&= \|x\|^2 - \sum_{i=1}^n (x, x_i)(x_i, x) \\
&= \|x\|^2 - \sum_{i=1}^n |(x, x_i)|^2
\end{aligned}$$

That is

$$\sum_{i=1}^n |(x, x_i)|^2 \leq \|x\|^2.$$

The coefficients

$$(x, x_i)$$

are called the Fourier coefficients of x with respect to the orthonormal set $\{x_i, i = 1..n\}$. If x is a linear combination of the set S ,

$$x = \sum_{i=1}^n \alpha_i x_i,$$

then

$$(x, x_i) = \alpha_i (x_i, x_i) = \alpha_i.$$

So Bessel's inequality is an equality if and only if x is a linear combination of elements of S .

The Cauchy-Schwarz inequality can be proved from the Bessel inequality. Indeed, given x and y , let

$$x_1 = \frac{y}{\|y\|}.$$

Then the set containing just this vector is an orthonormal set, so we have by the Bessel inequality

$$\|x\|^2 \geq |(x, x_i)|^2 = |(x, y/\|y\|)|^2 = |(x, y)|^2/\|y\|^2.$$

Thus

$$|(x, y)| \leq \|x\|\|y\|.$$

0.11.3 The Triangle Inequality

$$\begin{aligned} \|x + y\|^2 &= (x + y, x + y) \\ &= \|x\|^2 + (x, y) + (y, x) + \|y\|^2 \\ &= \|x\|^2 + 2RE(x, y) + \|y\|^2 \\ &\leq \|x\|^2 + 2|(x, y)| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2. \end{aligned}$$

So

$$\|x + y\| \leq \|x\| + \|y\|$$

So $\sqrt{(x, x)}$ is a norm. Also

$$d(x, y) = \|x - y\|$$

is a metric, distance function. Indeed we have

$$\|x - y\| = \|x - z + z - y\| \leq \|x - z\| + \|z - y\|.$$

0.12 Gram-Schmidt Orthogonalization

Given a vector v and a vector u , the projection of v onto u has length

$$\|v\| \cos(\theta),$$

where θ is the angle between v and u . This can be written as

$$\frac{(v, u)}{\|u\|}$$

The projection of v in the direction of u , or the component of v in the direction u is

$$\frac{(v, u)}{\|u\|} \frac{u}{\|u\|} = \frac{(v, u)}{\|u\|^2} u.$$

So suppose we have an orthogonal set $W = w_1, w_2, \dots, w_k$ and a vector v , not in W . Let V be the space spanned by W and v . We can find a replacement for v , w_{k+1} , so that $w_1, w_2, \dots, w_k, w_{k+1}$ is an orthogonal basis for V . We subtract from v its projections in the direction of each of the w_1, w_2, \dots, w_k . So

$$w_{k+1} = v - \frac{(v, w_1)}{\|w_1\|^2} w_1 - \frac{(v, w_2)}{\|w_2\|^2} w_2 - \frac{(v, w_k)}{\|w_k\|^2} w_k$$

Using this process, given a basis of a space, we can compute from this basis an orthogonal basis. See **Numerical Methods**, Dahlquist and Bjorck, 1974, Prentice-Hall, for the modified Gram-Schmidt process, which is an equivalent process with better numerical stability.

0.13 The Parallelogram Law

Consider a parallelogram in the plane. Let s_1 and s_2 be vectors on the sides of the parallelogram. Let vectors d_1 and d_2 be the diagonals. Let x and y be the half diagonals. Then

$$s_1 = x + y,$$

and

$$s_2 = x - y$$

Draw a figure to see this. We find in general for any x and y in an inner product space that

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

, which follows by expanding the inner products. In the case of our parallelogram, we have

$$\|s\|^2 + \|s_2\|^2 = \frac{\|d_1\|^2 + \|d_2\|^2}{2}.$$

That is *The sum of the squares of the sides of a parallelogram is equal to the average of the squares of the two diagonals.* Thus the general identity

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

is called the parallelogram law.

0.14 Quadratic Forms

A symmetric matrix defines a quadratic form, a homogeneous second degree algebraic function. So if matrix M is symmetric and x is a column vector, then

$$q(x) = x^T M x,$$

is a quadratic form. For example,

$$(x_1 + 5x_2 + 2x_3)^2 = x_1^2 + 10x_1x_2 + 4x_1x_3 + 25x_2^2 + 20x_2x_3 + 4x_3^2,$$

is a quadratic form. Each term in the expansion is of the second degree. As a matrix equation this is

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} 1 & 5 & 2 \\ 5 & 25 & 10 \\ 2 & 10 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Every quadratic form can be diagonalized. That is there is a change of basis so that the matrix M becomes a diagonal matrix and the quadratic form becomes

$$q(x) = \sum_{i=1}^n a_i x_i^2.$$

Quadratic forms arise in various practical applications such as in vibration problems and in expressions for energy. Diagonalizing corresponds to the introduction of the so called normal coordinates. The elements on the diagonal of the diagonal matrix are the eigenvalues.

0.15 Canonical Forms

When matrices can not be diagonalized, there exists various canonical forms, which characterize the matrices and its properties. Among these are: Jordon Normal form, Rational Normal form, Triangular form, and LU decomposition.

0.16 Upper Triangular Form

For every matrix M , there is a matrix T so that

$$T^{-1}MT$$

is upper triangular. One may make an eigenvector of M the first column of T , to get a partitioned matrix where the first column has zeroes below the first element of the first column. Then one may apply induction to the remaining lower dimensional submatrix. See Richard Bellman **The Stability Theory of Differential Equations** Dover, for a simple proof.

0.17 Isometries, Rotations, Orthogonal Matrices

An isometry is a transformation that preserves length.

$$\|T(v)\| = \|v\|.$$

Let T be a linear transformation that is an isometry. Let T have a matrix representation M in some orthogonal basis. Then we shall show that M is an orthogonal matrix. This means that the norm of any column is one, and that any two columns are orthogonal.

We have

$$Mv \cdot Mv = v \cdot v$$

That is, if v is a column vector, then

$$(Mv)^T(Mv) = v^T M^T M v = v^T v.$$

Let v_i be the column vector whose i th element is one, and all other elements are zero. Then letting $P = M^T M$, we see that $P_{ii} = 1$. Letting $v = v_i + v_j$, for i not j , we find that $P_{ij} = -P_{ji}$. But P is symmetric. So $P_{ij} = 0$. Thus P is the identity matrix. So we have shown that the inverse of an orthogonal matrix is its transpose. This also implies that its row vectors are orthonormal. In two space or three space, a rotation matrix, representing a rotation about a fixed axis, is clearly an isometry. Hence it is orthogonal. Because

$$\det(M) = \det(M^T),$$

the determinant must be 1, or -1 . Clearly any real eigenvalue of an isometry must be 1 or -1 . This follows from the value of the determinant. An orthogonal matrix whose determinant is 1, called proper. A proper orthogonal transformation represents a rotation. A proper orthogonal matrix defines three Euler angles, which explicitly shows that the matrix represents a rotation. See Nobel, **Applied Linear Algebra**. Also for the determination of the rotation axis of the matrix, which is an eigenvector, see the discussion in the paper on rotation matrices contained in an issue of the IEEE Journal on Computational Geometry and Graphics. Also see the computer codes in the Emery Fortran and C++ libraries.

0.18 Rotation Matrices

See the paper **Rotations** by James Emery (Document rotation.tex).

0.19 Exponentials of Matrices and Operators in a Banach Algebra

See See the paper **Rotations** by James Emery (Document rotation.tex).

0.20 Eigenvalues

If a is a square matrix, and

$$av = \lambda v$$

then λ is called an eigenvalue and v the corresponding eigenvector. For example suppose

$$a = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$$

The characteristic equation is:

$$\begin{vmatrix} 1 - \lambda & 2 \\ 0 & 3 - \lambda \end{vmatrix} = 0 \\ = (1 - \lambda)(3 - \lambda).$$

Eigenvalues are the roots of the characteristic equation.

Eigenvalues $\lambda_1 = 1, \lambda_2 = 3$. Corresponding eigenvectors:

$$v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$v_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Matlab:

```
a =
```

```
    1    2  
    0    3
```

```
lambda=eig(a)
```

```
lambda =
```

```
    1  
    3
```

```
[v lambda]=eig(a)
```

```
v =
```

```
    1.0000    0.7071  
         0    0.7071
```

```
lambda =
```

```
    1    0  
    0    3
```

A symmetric matrix can always be diagonalized using the eigenvalues and the eigenvectors. The eigenvalues are the elements of the diagonal. This is the simplest example of what is called the spectral theorem. In vibration theory, this is equivalent to using normal coordinates. In quantum mechanics eigenanalysis plays a large role.

0.21 The Characteristic Polynomial and the Cayley-Hamilton Theorem

The characteristic polynomial of a n by n matrix M , $f(\lambda)$ is the polynomial

$$\det(M - \lambda I),$$

where I is the identity matrix. The roots of the characteristic polynomial are the eigenvalues of M .

The Cayley-Hamilton theorem says that a matrix is a root of its characteristic polynomial. That is suppose

$$a_3\lambda^3 + a_2\lambda^2 + a_1\lambda + a_0$$

is the characteristic polynomial of a matrix M . Then

$$a_3M^3 + a_2M^2 + a_1M + a_0I = 0$$

Example

0.22 Unitary Transformations

In a Hilbert space, a unitary transformation U has the property that

$$(Ux, Uy) = (x, y).$$

A complex n by n matrix M is unitary if its inverse is the conjugate transpose of M .

A real orthogonal matrix M is a matrix so that the its inverse is the transpose.

A unitary transformation is

0.23 Transpose, Trace, Self-Adjoint Operators

The trace of a matrix is the sum of the diagonal elements. The trace has the following property. If A and B are square matrices then

$$\text{trace}(AB) = \text{trace}(BA).$$

This follows because

$$\begin{aligned} \text{trace}(AB) &= \sum_{i=1}^n \sum_{j=1}^n a_{ij}b_{ji} \\ &= \sum_{j=1}^n \sum_{i=1}^n a_{ij}b_{ji} \\ &= \sum_{j=1}^n \sum_{i=1}^n b_{ji}a_{ij} \\ &= \text{trace}(BA). \end{aligned}$$

If

$$M' = PMP^{-1},$$

then

$$\text{trace}(M') = \text{trace}(PMP^{-1}) = \text{trace}(P^{-1}PM) = \text{trace}(M).$$

0.24 The Spectrum

The root word of spectrum is a word meaning appearance or image. One might speculate that spectrum was first used for the image of lines obtained from a spectroscope. I suppose the use of spectrum in mathematical analysis must have come after the invention of quantum mechanics.

A bounded linear operator T is an operator such that there exists some number M so that

$$\|Tx\| \leq M\|x\|$$

for all x .

The spectrum of a bounded linear operator T is the set λ in which the operator

$$T - \lambda I$$

has an inverse. In general the spectrum may consist of several parts, the continuous spectrum, the residual spectrum, and the point spectrum (the eigenvalues). A finite linear transformation has only a point spectrum.

0.25 The Spectral Theorem

Various versions of the spectral theorem in linear algebra show how a linear transformation may be characterized by its eigenvalues and eigenvectors. In the most simple form of a symmetric matrix it is shown that the matrix may be diagonalized by introducing a basis of eigenvectors. The eigenvectors are orthogonal to each other and the eigenvalues are real. A very good survey of the spectral theorem, in both finite dimensional and infinite dimensional spaces is

Lorch Edward R, **The Spectral Theorem**, In *Studies in Modern Analysis* Volume I, R. C. Buck editor, MAA Studies in Mathematics, The Mathematical Association of America, 1962.

0.26 Tensors

A tensor is a multilinear functional. That is, it is a function defined on a cartesian product of vector spaces. It is a linear transformation in each individual vector space v_i . Thus

$$f(v_1, v_2, \dots, v_i + u_i, \dots, v_n) = f(v_1, v_2, \dots, v_i, \dots, v_n) + f(v_1, v_2, \dots, u_i, \dots, v_n)$$

and so on.

0.27 Application of Linear Algebra to Vibration Theory, Normal Coordinates

See the document Vibration, (vibra.tex).

0.27.1 A Simple Spring Example

Spring force:

$$F = kx.$$

Spring Potential Energy:

$$V = k \frac{x^2}{2}.$$

Kinetic Energy:

$$T = m \frac{\dot{x}^2}{2}$$

Lagrangian:

$$L = T - V.$$

Lagrange's Equation:

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0.$$

We have

$$\begin{aligned}\frac{\partial L}{\partial \dot{x}} &= m\dot{x}. \\ \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} &= m\ddot{x}. \\ \frac{\partial L}{\partial x} &= -kx\end{aligned}$$

So Lagrange's equation is

$$m\ddot{x} + kx = 0.$$

Trying $e^{\alpha t}$ as a solution, we find

$$(m\alpha^2 + k)e^{\alpha t} = 0$$

or

$$\alpha = i\sqrt{k/m} = i\omega,$$

where

$$\omega^2 = \frac{k}{m}.$$

So

$$x = e^{i\omega t}$$

The differential equation can be written as

$$\frac{\partial^2}{\partial t^2} x = -\omega^2 x.$$

So $x = e^{i\omega t}$ and $x = e^{-i\omega t}$ are eigenfunctions, and $-\omega^2$ is an eigenvalue of the linear operator

$$\frac{\partial^2}{\partial t^2}.$$

0.27.2 Decoupling Equations

Consider the equation

$$M\ddot{X} + KX = F,$$

where M is a square n -dimensional symmetric mass matrix, K is a square n -dimensional symmetric stiffness matrix, and F is a n -dimensional force. Both M and K must be positive definite because they represent the kinetic

and strain quadratic forms. If either of them were not positive definite, there would be a nodal displacement or a nodal displacement velocity giving a negative energy. We assume a solution of the form

$$X = X_0 \exp(i\omega t).$$

Consider the homogeneous problem with $F = 0$. We have

$$-M\omega^2 X_0 + K X_0 = 0.$$

Let us write X for X_0 . Then we have

$$(K - \omega^2 M)X = 0$$

This is a generalized eigenvalue problem. This homogeneous problem has a nonzero solution if and only if the determinant is zero:

$$|K - \omega^2 M| = 0.$$

This determinant is called the characteristic function, which is an n th degree polynomial in the variable ω^2 . The roots are called eigenvalues. In the case of symmetric matrices the eigenvalues are real. To show this, let the inner product be

$$(u, v) = \sum_{i=1}^n u_i \bar{v}_i,$$

where the subscripted values are the vector components, and where the bar indicates the complex conjugate. Then if λ is an eigenvalue, with eigenvector v , because M and K are self-adjoint (real symmetric), we have

$$\begin{aligned} \lambda(Mv, v) &= (\lambda Mv, v) = (Kv, v) = (v, \bar{K}^T v) \\ &= (v, Kv) = (v, \lambda Mv) = \bar{\lambda}(v, Mv) = \bar{\lambda}(\bar{M}^T v, v) = \bar{\lambda}(Mv, v). \end{aligned}$$

Therefore

$$\lambda = \bar{\lambda},$$

and so λ is real.

In this vibration problem the eigenvalues are not only real, but positive, because the matrices are positive definite. Under certain conditions, (1) If

they are distinct, or (2) If one of the matrices is positive definite, then the vibration problem can be decoupled. Suppose that the eigenvalues are distinct. That is, the eigenfrequencies $\omega_1, \dots, \omega_n$ are distinct. Since the determinant of the matrix

$$K - \omega_i^2 M$$

is zero for each i , its column vectors are linearly dependent, so that there exists a nonzero vector X_i such that

$$(K - \omega_i^2 M)X_i = 0.$$

The vectors X_1, \dots, X_n are called eigenvectors or modal vectors. Let P be a matrix which has as columns the eigenvectors

$$P = [X_1 | X_2 | \dots | X_n]$$

By the definition of the eigenvectors, we have

$$\begin{aligned} K[X_1 | X_2 | \dots | X_n] &= [\omega_1^2 M X_1 | \omega_2^2 M X_2 | \dots | \omega_n^2 M X_n] \\ &= M[X_1 | X_2 | \dots | X_n] \Lambda = MP\Lambda, \end{aligned}$$

where Λ is the diagonal matrix of eigenvalues. Then

$$P^T K P = P^T M P \Lambda.$$

We will show that the eigenvectors have certain orthogonality properties. We will show that if ω_i^2 is not equal to ω_j^2 , then vector X_i is perpendicular to X_j . Let A be a matrix, then we write A^T for the transpose. Because K and M are symmetric

$$K^T = K$$

and

$$M^T = M$$

We have

$$X_i^T K X_j = X_i^T \omega_j^2 M X_j = \omega_j^2 X_i^T M X_j$$

On the other hand

$$X_i^T K X_j = (K^T X_i)^T X_j = (K X_i)^T X_j = (\omega_i^2 M X_i)^T X_j$$

$$= \omega_i^2 X_i^T M^T X_j = \omega_i^2 X_i^T M X_j.$$

So

$$\omega_j^2 X_i^T M X_j = \omega_i^2 X_i^T M X_j.$$

But ω_j and ω_i are not equal, so that we must have

$$X_i^T M X_j = 0,$$

and then

$$X_i^T K X_j = 0.$$

Now let

$$X_i^T M X_i = \mu_i^2.$$

Any multiple of an eigenvector is still an eigenvector, so we scale the eigenvectors so that each

$$\mu_i^2 = 1$$

Then

$$P^T M P = I.$$

Then we have

$$P^T K P = P^T M P \Lambda = I \Lambda = \Lambda.$$

Both matrices have been diagonalized.

0.27.3 Example: Coupled Oscillators

Consider two masses m_1 and m_2 . They can each move in a horizontal direction. Let the masses be connected with springs, k_1, k_2 , and k_3 . Spring k_1 is connected on the left with a fixed point, and on the right with mass m_1 . Spring k_2 is connected on the left with m_1 , and on the right with mass m_2 . Spring k_3 is connected on the left with m_2 , and on the right with a fixed point. Let u_1 and u_2 be the displacement from equilibrium of masses m_1 and m_2 respectively. The kinetic energy of the system is

$$T = \frac{1}{2} m_1 \dot{u}_1^2 + \frac{1}{2} m_2 \dot{u}_2^2$$

and the potential energy is

$$V = \frac{1}{2} k_1 u_1^2 + \frac{1}{2} k_2 (u_2 - u_1)^2 + \frac{1}{2} k_3 u_2^2.$$

The Lagrangian is

$$L = T - V.$$

The equations of motion are given by Lagrange's equations

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{u}_1} - \frac{\partial L}{\partial u_1} = 0$$

and

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{u}_2} - \frac{\partial L}{\partial u_2} = 0.$$

We have

$$\frac{\partial L}{\partial \dot{u}_1} = \frac{\partial T}{\partial \dot{u}_1} = m_1 \dot{u}_1$$

and

$$\frac{\partial L}{\partial \dot{u}_2} = \frac{\partial T}{\partial \dot{u}_2} = m_2 \dot{u}_2.$$

Also

$$-\frac{\partial L}{\partial u_1} = \frac{\partial V}{\partial u_1} = k_1 u_1 - k_2(u_2 - u_1)$$

and

$$-\frac{\partial L}{\partial u_2} = \frac{\partial V}{\partial u_2} = k_2(u_2 - u_1) + k_3 u_2.$$

Thus Lagrange's equations are

$$m_1 \ddot{u}_1 + k_1 u_1 - k_2(u_2 - u_1) = 0$$

and

$$m_2 \ddot{u}_2 + k_2(u_2 - u_1) + k_3 u_2 = 0.$$

In matrix form the equations become

$$M \begin{bmatrix} \ddot{u}_1 \\ \ddot{u}_2 \end{bmatrix} + K \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 0$$

where

$$M = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix},$$

and

$$K = \begin{bmatrix} (k_1 + k_2) & -k_2 \\ -k_2 & (k_3 + k_2) \end{bmatrix}.$$

Let

$$u_1 = U_1 e^{i\omega t}$$

and

$$u_2 = U_2 e^{i\omega t}.$$

Then

$$(K - \omega^2 M) \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = 0$$

This equation has a nonzero solution if and only if

$$\det(K - \omega^2 M) = 0.$$

Let $\lambda = \omega^2$. The determinant is

$$(k_1 + k_2 - \lambda m_1)(k_3 + k_2 - \lambda m_2) - k_2^2,$$

which we equate to zero to find the eigenvalues λ . Now let us specialize the problem and set the spring constants k_1 and k_3 , to a common value k . And also set the two masses m_1 and m_3 , to a common value m . Let the coupling spring be a weak spring.

That is, let spring constant k_2 be equal to a fraction of k . The equation for the eigenvalues is the quadratic equation

$$(k + k_2 - \lambda m)(k + k_2 - \lambda m) - k_2^2 = 0.$$

That is, it is

$$A\lambda^2 + B\lambda + C = 0,$$

with

$$A = m^2$$

$$B = -2(k + k_2)m$$

and

$$C = (k + k_2)^2 - k_2^2 = k^2 + 2kk_2.$$

Let $a = (k + k_2 - \lambda m)$ and $b = -k_2$. Then since the determinant is 0, we must have

$$a^2 = b^2$$

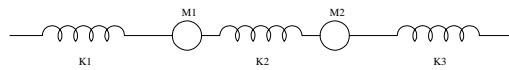


Figure 1: Coupled springs, With masses m_1 and m_1 . The spring constants are k_1, k_2 and k_3 . k_2 is the constant for the coupling spring and is much smaller than the other two constants.

So either

$$a = b,$$

or

$$a = -b.$$

The equations for the eigenvectors are

$$aU_1 + bU_2 = 0,$$

and

$$bU_1 + aU_2 = 0.$$

If $a = b$, then the eigenvector is

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

On the other hand, if $a = -b$, then the eigenvector is

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The eigenvalues are λ_1 and λ_2 . Let a_i be the value of a for eigenvalue λ_i , and let b_i be the value of b for eigenvalue λ_i . Now let us compute with values assigned to the constants. Let $m = 1$, $k = 1$, and $k_2 = 1/10$. The computation is done with a Matlab script (Octave script). Call it **coupled.m**:

```
m=1.;
k=1.;
k_2=1/10;
A= m*m;
B= -2*(k+k_2)*m ;
C= k^2 + 2 * k * k_2;
lambda_1=(-B + sqrt(B*B - 4*A*C))/(2*A)
omega_1=sqrt(lambda_1)
a_1 = (k + k_2 - lambda_1 * m)
b_1 = -k_2
lambda_2=(-B - sqrt(B*B - 4*A*C))/(2*A)
omega_2=sqrt(lambda_2)
```

```

a_2 = (k + k_2 - lambda_2 * m)
b_2 = -k_2
omega=(omega_1+omega_2)/2
t1=2*pi/omega
phi=(omega_1 -omega_2)/2
t2=2*pi/phi

```

The output is:

```

octave-3.0.0.exe:5> coupled
lambda_1 = 1.2000
omega_1 = 1.0954
a_1 = -0.10000
b_1 = -0.10000
lambda_2 = 1.00000
omega_2 = 1.00000
a_2 = 0.10000
b_2 = -0.10000
omega = 1.0477
t1 = 5.9970
phi = 0.047723
t2 = 131.66
octave-3.0.0.exe:6> exit

```

The first eigenvalue is

$$\lambda_1 = 1.2$$

a_1 and b_1 are equal, so the eigenvector is

$$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

The corresponding frequencies are

$$\omega_1 = \pm\sqrt{\lambda_1} = \pm 1.0954$$

Hence two linearly independent solutions are

$$e^{i\omega_1 t} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and

$$e^{-i\omega_1 t} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

So any linear combination of these solutions is also a solution. That is, the space spanned by these two solutions is a solution. But both $\cos(\omega_1 t)$ and $\sin(\omega_1 t)$ can be written as a linear combination of $e^{i\omega_1 t}$ and $e^{-i\omega_1 t}$. Hence this solution space is the same as the space spanned by

$$\cos(\omega_1 t) \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and

$$\sin(\omega_1 t) \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

The second eigenvalue is

$$\lambda_2 = 1.0$$

a_2 and b_2 are not equal, so the eigenvector is

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The corresponding frequencies are

$$\omega_1 = \pm\sqrt{\lambda_2} = \pm 1.0$$

Hence two linearly independent solutions are

$$e^{i\omega_2 t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and

$$e^{-i\omega_2 t} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Again the solution space is spanned by

$$\cos(\omega_2 t) \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and

$$\sin(\omega_2 t) \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Suppose the initial conditions are

$$\begin{aligned} u_1(0) &= A \\ \frac{du_1(0)}{dt} &= 0 \\ u_2(0) &= 0 \\ \frac{du_2(0)}{dt} &= 0 \end{aligned}$$

Let the solution be

$$u = c_{11} \cos(\omega_1 t) \begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_{12} \sin(\omega_1 t) \begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_{21} \cos(\omega_2 t) \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_{22} \sin(\omega_2 t) \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

where the constants are to be determined by the initial conditions.

From the first initial condition we have

$$c_{11} + c_{21} = A$$

From the third initial condition we have

$$-c_{11} + c_{21} = 0$$

Thus

$$c_{21} = A/2$$

and

$$c_{11} = A/2.$$

Differentiating our solution, we have

$$\frac{du}{dt} = -c_{11}\omega_1 \sin(\omega_1 t) \begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_{12}\omega_1 \cos(\omega_1 t) \begin{bmatrix} 1 \\ -1 \end{bmatrix} - c_{21}\omega_2 \sin(\omega_2 t) \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_{22}\omega_2 \cos(\omega_2 t) \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

From the second initial condition we have

$$c_{12}\omega_1 + c_{22}\omega_2 = 0.$$

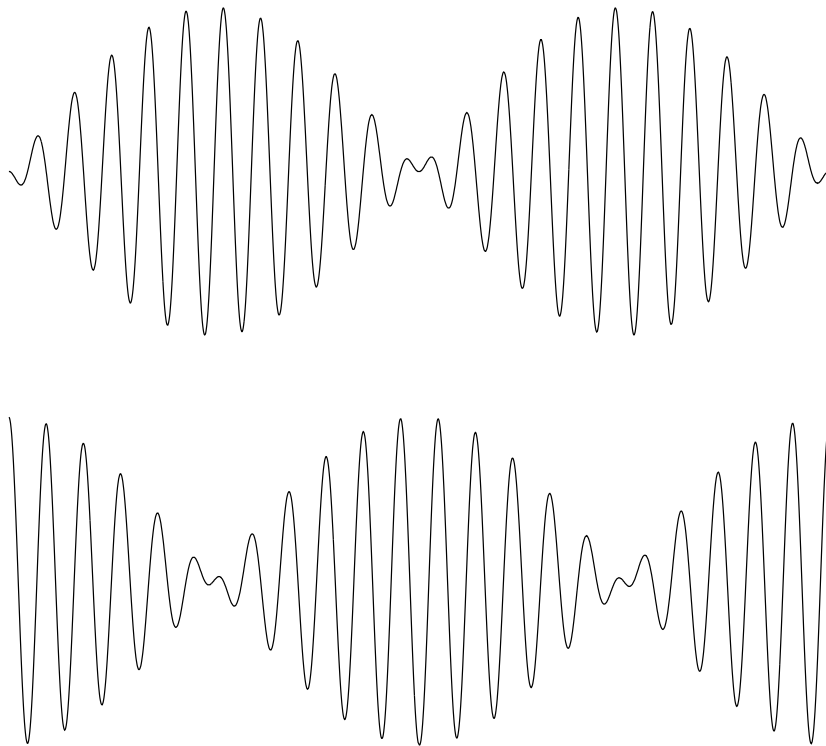


Figure 2: Coupled oscillations. The lower curve is the oscillation of mass m_1 . The upper curve is the oscillation of mass m_2 .

From the forth initial condition we have

$$-c_{12}\omega_1 + c_{22}\omega_2 = 0.$$

Therefore $c_{22} = 0$ and $c_{12} = 0$. Therefore our solution is

$$u_1 = \frac{A}{2} \cos(\omega_1 t) + \frac{A}{2} \cos(\omega_2 t).$$

$$u_2 = -\frac{A}{2} \cos(\omega_1 t) + \frac{A}{2} \cos(\omega_2 t).$$

Now if $\omega = (\omega_1 + \omega_2)/2$ and $\phi = (\omega_1 - \omega_2)/2$, then

$$\cos(\omega_1 t) + \cos(\omega_2 t) = 2 \cos(\omega t) \cos(\phi t)$$

So

$$u_1 = A \cos(\omega t) \cos(\phi t).$$

Similarly

$$u_2 = -A \sin(\omega t) \sin(\phi t).$$

For the numbers we have used here the oscillating frequency is

$$\omega = 1.0477,$$

and the corresponding period is

$$T_1 = 5.9970.$$

The modulating frequency is

$$\phi = 0.047723,$$

and the modulating period is

$$t_2 = 131.66.$$

The oscillators operate as follows. The first oscillator starts at maximum amplitude A , while the second oscillator is stopped at zero amplitude. The amplitude of the first oscillator begins to decrease, according to the modulating factor $\cos(\phi t)$. At $t = t_2/4$, the amplitude of the first oscillator has decreased to zero, whereas the amplitude of the second oscillator has increased to its maximum amplitude A , according to the modulating factor $\sin(\phi t)$. The energy of the first oscillator has been transferred to the second. This behavior repeats, as energy flows back and forth between the two oscillators.

0.27.4 The General Problem of Linear Vibration

The reference for this section is:

Bradbury T C, **Theoretical Mechanics**, John Wiley, 1968, p352.

The kinetic and potential energies are given as quadratic forms

$$T = \frac{1}{2}g_{ij}\dot{q}_i\dot{q}_j,$$

and

$$V = \frac{1}{2}k_{ij}q_iq_j.$$

The Lagrange equations of motion become

$$g_{ij}\ddot{q}_j + k_{ij}q_j = 0.$$

Write this as a matrix equation

$$G\ddot{Q} + KQ = 0,$$

where G and K are symmetric n by n matrices. This equation has a solution

$$Q = X \exp(\omega t).$$

Then

$$(K - \lambda G)X = 0,$$

where

$$\lambda = \omega^2.$$

This is a generalized eigenvalue problem. The quadratic forms are positive definite. The eigenvalues are real, and the eigenvectors (normal modes) are orthogonal. Let the rows of matrix S be constructed out of eigenvectors, which are scaled to be of unit length with metric G (S , or S^T is called the modal matrix in Thomson "Theory of Vibration."). That is,

$$X^T G X = \delta_{ij}.$$

Then

$$S G S^T = I.$$

Let Λ be a diagonal matrix of the eigenvalues. Then

$$KS^T = \Lambda GS.$$

Define normal coordinates Q' by

$$Q = Q'S^T.$$

Then the equations of motion become

$$GS^T\ddot{Q}' + KS^TQ' = 0.$$

Multiplying by S and simplifying, we get

$$\ddot{Q}' + \Lambda Q' = 0.$$

The problem could also be solved by diagonalizing the original energy quadratic forms. Any two positive definite quadratic forms can be simultaneously diagonalized. See for example, D. E. Littlewood **A University Algebra**, 2nd ed. 1958 Dover, p53.

0.28 Projection Operators

A projection operator is a linear operator P that has the property that

$$P(P(v)) = P(v).$$

So for example in the simple geometric case of projecting a 3-d vector to a 2-d subspace, that is to a plane, the projection then lies in the plane, hence projecting it again to the plane does not alter it because it is already in the plane.

0.29 Functional Analysis

When our vector spaces are infinite dimensional, for example consider the vector space of all continuous functions on an interval $[a, b]$, we enter the realm of functional analysis. This area of mathematics arose out of the problems of solving differential and integral equations, and in such areas as Fourier Analysis. So one studies Hilbert Spaces, which are complete inner product spaces, and Banach Spaces, which are complete normed linear spaces. In a further increase in abstraction one studies Topological Vector spaces in which there may be no inner product or norm.

0.30 Hamel Basis

Hamel constructed a basis for the vector space of real numbers over the rational field. He did this using Zorn's Lemma. So for example the vectors $\sqrt{2}$ and $\sqrt{3}$ are linearly independent. Otherwise

$$\sqrt{\frac{3}{2}}$$

would be a rational number.

For if we assumed that

$$\sqrt{\frac{3}{2}}$$

were rational we could write

$$\sqrt{\frac{3}{2}} = \frac{n}{m},$$

where n and m have no common factor. Then

$$\begin{aligned}\frac{3}{2} &= \frac{n^2}{m^2}, \\ 3m^2 &= 2n^2,\end{aligned}$$

so n^2 and thus n must have a factor of 3. Hence m^2 and hence m must have a factor of 3. This contradicts our assumption.

See Angus Taylor **Introduction to Functional Analysis** for a treatment of Hamel basis.

0.31 Numerical Linear Algebra

See Golub and Van Loan, and books on Matlab.

0.32 Quantum Mechanics

Quantum Mechanics was treated by John Von Neumann in terms of Hilbert Spaces and Linear Operators in Hilbert Space, and involving eigenvalues and eigenvectors of such spaces of functions. One runs into things like delta functions, which are approximations to functions, but not actually functions at

all. Thus we use mathematical objects called distributions. These distributions are functionals and not to be confused with probability distributions. Von Neumann made great contributions to the study of Banach Algebras.

0.33 The Schrodinger Wave Equation

From the photoelectric effect, the energy of a photon is

$$E = h\nu,$$

where h is Planck's constant, and ν is the frequency of the photon. From relativity theory, the energy of a particle is given by the equation

$$E^2 = p^2v^2 + m^2c^4,$$

where v is the particle velocity and m is its mass. For a photon the velocity is the velocity of light c , and the mass is zero. Thus for a photon

$$E^2 = p^2c^2.$$

This implies that the wavelength is

$$\lambda = c/\nu = \frac{E/p}{E/h} = h/p.$$

Let \mathbf{k} be the wave number vector. By definition

$$k\lambda = 2\pi.$$

We have

$$\mathbf{p} = \frac{\mathbf{k}h}{k\lambda} = \hbar\mathbf{k},$$

where

$$\hbar = \frac{h}{2\pi}.$$

A plane wave takes the form

$$e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} = e^{(i/\hbar)(\mathbf{p}\cdot\mathbf{r}-Et)}$$

A wave packet is represented as a sum of such plane waves, written as the Fourier transform

$$\psi(\mathbf{r}, t) = \frac{1}{\sqrt{2\pi\hbar}} \int a(\mathbf{p}) e^{(i/\hbar)(\mathbf{p}\cdot\mathbf{r} - Et)} d\mathbf{p}.$$

This wave function satisfies the equation

$$(\hbar^2/2m)\nabla^2\psi - (\hbar/i)\partial\psi/\partial t = 0,$$

because $p^2/(2m) - E = 0$. It is called the Schrodinger wave equation. The group velocity of the wave packet is

$$v_g = \frac{d\omega}{dk} = \frac{dE}{dp}.$$

Using

$$E = \frac{mc^2}{\sqrt{1 - v^2/c^2}},$$

and

$$p = \frac{mv}{\sqrt{1 - v^2/c^2}}.$$

We find that

$$\frac{dE}{dp} = \frac{pc^2}{E} = v.$$

Thus the group velocity of the wave packet corresponds to the particle velocity. $|\psi|^2$ is proportional to the probability density function for the location of the particle. This is the Born postulate.

0.34 The Postulates of Quantum Mechanics

The dynamical variables of a mechanical system are interpreted as operators, operating on wave functions which are elements of a Hilbert space. The identification with classical mechanics is done somewhat in the manner of the mathematical theory of distributions. The expected values of a variable (i.e. the measured values) form a discrete set which is the set of eigenvalues of the operator.

0.35 The Bra and Ket Notation of Dirac

A ket $|v\rangle$ vector is an element of a Hilbert space (a complete complex inner product space). A bra vector $\langle u|$ is an element of the dual space, that is, it is a linear operator. Thus the value of u on v is

$$u(v) = \langle u|v\rangle,$$

which also may be interpreted as an inner product.

In another way of looking at it, a bra vector $\langle v|$ is a row matrix of components with respect to some basis, of a vector v , that is if

$$v = \sum_{i=1}^n v_i b_i$$

then

$$v = [v_1 \quad v_2 \quad \dots \quad v_n]$$

and a ket vector $|v\rangle$ is the conjugate transpose, and thus a column vector of components.

From a mathematical point of view, this bra and ket business seems rather complicated, confused, and is not clearly necessary.

Chapter VII in Messiah outlines some linear algebra in this bra ket notation.

Section 2.3 of Shankar gives another explanation of this notation.

0.36 Example: The Hydrogen Atom

Schrodinger's equation is a partial differential equation and can be solved by the method of separation of variable. That is a solution can be written as a product, where each product function uses independent variables. Then usually the partial differential equation becomes a set of ordinary differential equations and eigenvalue problems. These eigenvalues than characterize quantum energy levels. In the case of the hydrogen atom,

0.37 The Relation Between Linear Algebra, Functional Analysis, and Abstract Solutions to Problems

Ultimately the solution of applied problems must involve a finite set of numbers rather than an abstract infinite theoretical solution. But the abstract solution can define properties of the solution and indeed existence of a solution and uniqueness. This is important, for finite approximations of a solution that does not exist will fail. So finally we typically will use the results of linear algebra to find arbitrary close approximations to applied problems. Specifically we require results from numerical linear algebra. That is we need algorithms that are robust, accurate, and render roundoff and truncation error small.

0.38 Bibliography

The number of books written on linear algebra and functional analysis is huge. So I list here some favorites of mine.

- [1] Lange Serge, **Linear Algebra** Addison Wesley, 1968
- [2] Lawson and Hanson, **Solving Linear Least Squares Problems**, (JEL), 1974, (least squares, singular value decomposition).
- [3] Stoker J J, **Nonlinear Vibrations**, (JEL).
- [4] Akhiezer N I, Glazman I M, **Theory of Linear Operators in Hilbert Space**, V I 1961, V II 1963, (two volumes Dover 1993).
- [5] Halmos Paul R, **Finite-Dimensional Vector Spaces**, Springer-Verlag, 1974.
- [6] Halmos Paul R, **Introduction to Hilbert Space and the Theory of Spectral Multiplicity**, 2nd ed., Chelsea, 1951.
- [7] Herstein I N, **Topics in Algebra**, Blaisdell, 1964., (Frobenius Theorem on Division Algebras, Quaternions).
- [8] Strang Gilbert, **Linear Algebra and Its Applications**, Academic Press, 1980, 2nd Edition..
- [9] Fausett Laurene V, **Applied Numerical Analysis Using Matlab**, Prentice-Hall, 1999..
- [10] Golub Gene H, Van Loan Charles F, **Matrix Computations**, John Hop-

kins University Press, 1989.

[11]Lorch Edward R, **The Spectral Theorem**, In *Studies in Modern Analysis* Volume I, R. C. Buck editor, MAA Studies in Mathematics, The Mathematical Association of America, 1962.

[12]Taylor Angus E, **Introduction to Functional Analysis**, John Wiley, 1967.

[13]Courant Richard, Hilbert David, **Methods of Mathematical Physics**, Volumes I and 2, Wiley, Springer , 1937.

[14]Bamberg Paul, Sternberg Shlomo, **A Course in Mathematics for Students of Physics**,, volumes 1 and 2 , Cambridge University Press, 1988..

[15]Yosida Kosaku, **Functional Analysis**, 2nd Edition, 1968, Springer-Verlag.

[16]Reed Michael, Simon Barry, 1972,, **Methods of Modern Mathematical Physics, Part 1. Functional Analysis**, 1972, Academic Press.

[17]Banach Stefan, **Theorie Des Operations Lineares**, Chelsea, 1932.

[18]Banach Stefan, **Theory of Linear Operations**, English translation by F. Jellet, 1987, Elsevier..

[19]Noble Ben, **Applied Linear Algebra**, Prentice-Hall 1969.

[20]Nering Evar D, **Linear Algebra and Matrix Theory**, John Wiley, 1963.

[21]Powell John L, Crasemann Bernd **Quantum Mechanics** Addison-Wesley, 1965.