MATRIX ANALYSIS II

FURTHER INTRODUCTION AND SOME APPLICATIONS TO PHYSICAL PROBLEMS

1952

H. V. McINTOSH

Table of Contents for Part II

FINITE GROUPS

By defining an abstract group one attempts to capture the essence of multiplication theory. For this purpose we define:

An abstract group $\mathcal{G}$ is a collection of elements $\{a, b, \cdots\}$ for which there is defined a rule of composition, known as group multiplication, satisfying the following four postulates:

1. The group is closed under multiplication
2. Group multiplication is associative
3. The group contains a left identity
4. Each element of the group has a left inverse

An example is the matrix group consisting of the following six matrices:

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad D = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad F = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Writing their multiplication table we have

$\cdot$	I	A	B	C	D	F
I	I	A	B	C	D	F
A	A	B	I	D	F	C
B	B	I	A	F	C	D
C	C	F	D	I	B	A
D	D	C	F	A	I	B
F	F	D	C	B	A	I

since the product of any two of these matrices is again one of them, this set of matrices is closed. The products are associative because this is already true of matrix products. $\mathbb{I}$ serves as the identity, and the fact that $\mathbb{I}$ occurs once in every column shows that each of these matrices has an inverse which is also in the group. Thus the group postulates are all satisfied, and these matrices actually form a group.

From those four postulates many consequences may be deduced, some of which have already been shown independently, for matrix multiplication. For instance, the postulates required only left identities and left inverses. There are also right identities and inverses as we now see.

Theorem 1: $ea = a \implies ae = a$

Proof:

$ea = e$ postulate 3

$(a^{-1}a)e = a^{-1}a$ postulate 4: $e = a^{-1}a$

($\implies$ means "implies that")

$a^{-1}(ae) = a^{-1}a$ associativity

$(a^{-1})^{-1}a^{-1}(ae) = (a^{-1})^{-1}a^{-1}a$ There is an $(a^{-1})^{-1}$

$\left((a^{-1})^{-1}a^{-1}\right)(ae) = \left((a^{-1})^{-1}a^{-1}\right)a$ associactivity

$eae = ea$

$ae = a$

Q.E.D.

As may be seen from this proof, in which each step has been written down explicitly, many of the steps which one might have taken for granted, such as cancelling on both sides of an equation, depend upon the group postulates being true in order that they may be legitimate.

Theorem ii: The identity is unique

Proof: Suppose not. Then

$e a^{-1} = a^{-1}$ hyp

$e' a^{-1} = a^{-1}$

$\therefore e a^{-1} = e' a^{-1}$

$(e a^{-1}) a = (e' a^{-1}) a$ a has an inverse a^{-1}

$e(a^{-1}a) = e'(a^{-1}a)$ association

$e = e'$ by Theorem i

Q. E. D.

Theorem iii: $a^{-1}a = e \Rightarrow a a^{-1} = e$; left inverses are right inverses.

Proof:

$ae = a$ Theorem i

$a(a^{-1}a) = a$ postulate 4

$(a a^{-1}) a = a$

$\therefore a a^{-1} = e$ Theorem iii

Q.E.D.

Theorem iv: The inverse in unique:

$\left.\begin{aligned} ba &= e \\ ea &= e \end{aligned}\right\}$ hyp

$ba = ca$

$(ba) a^{-1} = (ca) a^{-1}$ postulate 4

$b(aa^{-1}) = c(aa^{-1})$ postulate 2

$be = ce$ Theorem iii

$b = c$ Theorem i

Q.E.D.

this shows that $(a^{-1})^{-1} = a$, since $aa^{-1} = e$ and $(a^{-1})^{-1} = e$.

As with matrices, the inverse of a product of several elements is the product of the inverses in the opposite order

$$(abc---)^{-1} = ---c^{-1}b^{-1}a^{-1}$$

since

$$---c^{-1}b^{-1}a^{-1}abc--- = e$$

A group whose elements, in addition to satisfying the group postulates, also obey the commutative rule is called an abelian or a commutative group. Thus

$$ab = ba$$

Such a situation will be reflected in the symmetry of the group table.

The order of a group is the number of elements in the group. Thus the example was in a group of order 6. A group is called finite if its order is finite.

Now, the powers of an element of a finite group cannot all be different. Thus consider

$a^0 = e$

$a = a$

$a^2 = a \cdot a$

$a^3 = a(aa) = (aa)a$ etc.

a^4

...

At the very worst all these powers would be different, so that upon coming to a^h, where h is the order of the group, an element must finally repeat. This repetition may very well occur before all the elements of the group have been used up. Let the first power which repeats be u.

$$a^u = a^k \qquad k < u;\ \text{suppose } k \neq 0$$

then $$a^{u-1} = a^{k-1}$$

contrary to the hypothesis, and $k = 0$, $a^u = e$, $a^{u-1} = a^{-1}$. Thus the first repitition must be with the identity. This number u is called the order of the element. The collection of elements

$$\left\{ e, a, a^2, \ldots, a^{u-1} \right\}$$

is called the period of a.

Now the period of an element satisfies all the group postulates, so that it is again a group. Such a collection of elements of a group which themselves form a group is called a subgroup. Furthermore, a group consisting of the powers of a certain element alone is called a cyclic group; so that the period of an element is a cyclic subgroup of the original group. A cyclic group is necessarily abelian, since the associative rule may be invoked to show

$$a^l a^m = a^m a^l = a^{l+m}$$

as with integers, matrix powers, etc.

It was seen upon inspection of the group table of the example that each element occured just once in each row and column. That this must necessarily be true generally may be seen quite easily. The proof given here for a row holds for a column by changing the orders of multiplication in the proof.

Suppose that an element occurred twice in a certain row. This would mean that both

$$\ell s = f$$
$$\ell t = f$$

where f is the repeated element in the ℓ-row. If this were true, multiply by ℓ^{-1} and see that

$$s = t$$

which is impossible since s and t are supposedly distinct group elements. On the other hand the element must occur once, for suppose f did not occur in the ℓ-row at all; but $\ell^{-1}f$ is an element of the group so that the product $\ell(\ell^{-1}f) = f$ must occur in the ℓ-row.

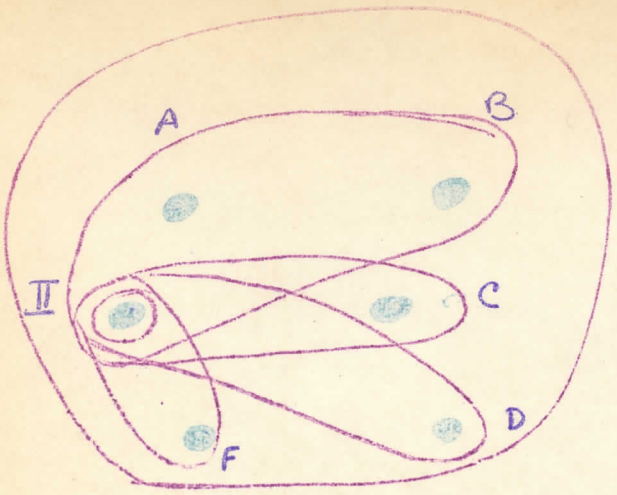
This has an important consequence that, if one has written a sum as

$$\sum_{a \in \mathcal{G}} f_a$$

where a certain set of quantities has been indexed, not by integers, as is common, but by group elements, and one summand taken for each group element, the sum may be rewritten in one of two ways.

$$\sum_{a \in \mathcal{G}} f_a = \sum_{a \in \mathcal{G}} f_{ab} = \sum_{a \in \mathcal{G}} f_{ba}$$

One method of partitioning a group into smaller units consists in dividing it up into subgroups. Each element sits in at least one subgroup, since its period is a subgroup. However, an element may sit in several subgroups, and $+ a \varepsilon \mathcal{G}_j$ means "a is an element of $\mathcal{G}_j$"



each subgroup certainly contains the identity.

The subgroups of the example we have been using are diagrammed at the right, as may be verified by checking their multiplication table.

A sometimes more useful partitioning of a group is the partitioning into equivalence classes. Two elements are "equivalent" when there stands between them a relation R with the properties

1) aRa (reflexive law)

2) $aRb \Rightarrow bRa$ (symmetric law)

3) aRb and $bRc \Rightarrow aRc$ (transitive law)

Ordinary equality is such a relation, and in fact, equivalence may be considered a sort of generalized equality. The advantage of an equivalence relation is that it partitions the group into non-overlapping subdivisions, which are called equivalence classes. All elements equivalent to a certain element sit in its equivalence class.

Each element sits in some equivalence class, for by the reflective property it sits in its own equivalence class. Furthermore if two classes have an element in common they coincide. For let $a \in K_b$, the class determined by b, and $a \in K_c$, the class determined by c. Then

$$aRb$$
$$aRc$$

by the symmetric law, cRa, and by the transitive law, since cRa, aRb, cRb and c sits in the equivalence class of b. Then by the transitive law, all the elemtns of K_c are elements of K_b.

For example, a group may be partitioned into cosets of a subgroup. If $\mathcal{H}$ is a subgroup of $\mathcal{G}$, containing the elements $h_0 = e, h_1, \cdots, h_{l-1}$, then the l elements $\{ph_1 = p, ph_1, \cdots, ph_{l-1}\}$ are called a left coset of $\mathcal{H}$. It is symbolized by $p\mathcal{H}$.

$$p\mathcal{H} = \{ph_1, ph_2, \cdots, ph_{l-1}, p\}$$

If p multiplies the elements of $\mathcal{H}$ on the right, the coset is a right coset called $\mathcal{H}p$.

$$\mathcal{H}p = \{p, h_1p, h_2p, \cdots, h_{l-1}p\}$$

Now, the left cosets, say, form equivalence classes in the groups, as do the right cosets. To show this for left cosets, let aRb when a sits in the coset defined by b, i.e., $a = bh_i \quad h_i \in \mathcal{H}$

$$a = bh_i$$

$$\therefore \quad b = ah_i^{-1} = ah_j \qquad \therefore aRb \Rightarrow bRa$$

$$a = ae = ah_1 \qquad \therefore aRa$$

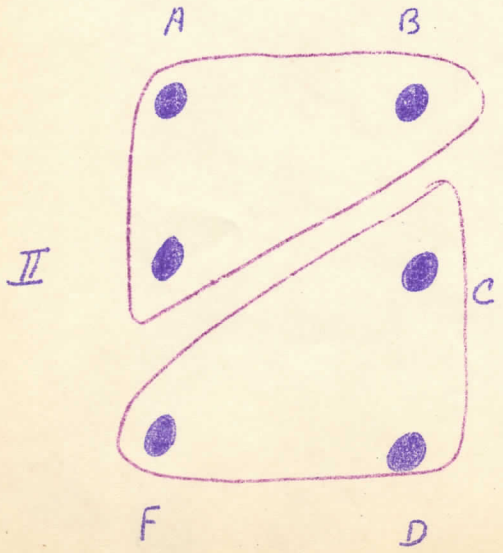
$$b = ch_j$$

$$a = bh_i$$

$$\therefore \quad a = ch_jh_i = ch_{i_e} \qquad \therefore akb \cdot bRc \Rightarrow aRc$$

Q.E.D.

Considering again the example, let $\mathcal{H} = \{\mathrm{I}, A, B\}$



$\mathrm{I}\mathcal{H} = \{\mathrm{I}, A, B\} = \mathcal{H}$

$A\mathcal{H} = \{A, B, \mathrm{I}\} = \mathcal{H}$

$B\mathcal{H} = \{B, \mathrm{I}, A\} = \mathcal{H}$

$C\mathcal{H} = \{C, \sigma, D\} = C\mathcal{H}$

$D\mathcal{H} = \{D, C, F\} = C\mathcal{H}$

$F\mathcal{H} = \{F, D, C\} = C\mathcal{H}$

Each coset of $\mathcal{H}$ contains the same number of elements, namely the number of elements in the group $\mathcal{H}$. If $\mathcal{G}$ has g elements and $\mathcal{H}$, h, the cosets exhaust $\mathcal{G}$ by layers so that if there are n cosets

$$nh = g$$

and, for finite groups the order of the subgroup must be a divisor of the order of the group. The number n is called the index of the subgroup. It also follows that since the period of an element is a subgroup, its index must divide the order of the group, and Hence the order of an element must divide the order of the group.

In particular,

$a^q = e$ (where q divides g the order of the group)

If one has a matrix group under consideration,

$$A^g = \mathbb{I}$$

and eigenvalves of A must satisfy

$$\lambda^g = 1$$

$$\lambda = e^{2\pi i \frac{k}{g}}$$

$$|\lambda| = 1$$

the elements of a finite group are then, in a rather general sense, roots of unity.

Obviously a group of prime order has only the trivial subgroups, consisting of the identity alone, and the entire group. Hence it must be cyclic, consisting of the powers of one of the elements other than e.

Another equivalence relation which may be formed in a group consists in letting aRb mean $a = p^{-1}bp \quad p \in \mathcal{G}$. Such an element, a, is called "conjugate" to b. The proof that this is an equivalence relation is as follows:

aRa $\qquad a = e^{-1} a e$

$aRb \Rightarrow bRa$ $\qquad a = p^{-1} b p$

$b = p a p^{-1} = (p^{-1})^{-1} a p^{-1}$

$aRb \cdot bRc \Rightarrow aRc$ $\qquad a = p^{-1} b p$

$b = q^{-1} c q$

$\therefore \; a = p^{-1} q^{-1} c q p$

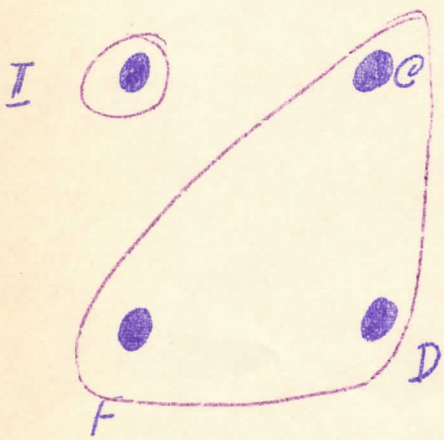
$a = (qp)^{-1} c q p$

Q.E.D.

The equivalence classes in this case are just called classes.

The group identity sits in a class by itself. If the group is abelian, each element is in a class by itself, since in the first case $p^{-1} e p = e$: in the second $p^{-1} a p = a p^{-1} p = a$. No class, other than that containing e, may be a subgroup, since they all lack e.

The classes of the example are shown at the left.



$R^{-1} \mathrm{I} R = \mathrm{I} \longrightarrow \{\mathrm{I}\}$

$\mathrm{I}^{-1} A \mathrm{I} = A$

$A^{-1} A A = A$

$B^{-1} A B = A$

$C^{-1} A C = B$ $\longrightarrow$ $\{A, B\}$

$D^{-1} A D = B$

$F^{-1} A F = B$

$\mathrm{I}^{-1} C \mathrm{I} = C$ $\qquad C^{-1} C C = C$

$A^{-1} C A = D$ $\qquad D^{-1} C D = F$ $\longrightarrow \{C, D, F\}$

$B^{-1} C B = F$ $\qquad F^{-1} C F = D$

In the case of matrix group, the classes are all distinguished by having the same trace, since

$$\chi(P^{-1}AP) = \chi(A)$$

Also the elements of a class all have the same order. If

$$a^n = e \qquad (p^{-1}ap)^n = p^{-1}a^n p = e$$

A collection of elements from a group, without especial reference to any property which they may possess, is called a complex. There is a sort of complex alegbra, which forms a convenient notation for some of the manipulations which one may perform upon a complex.

Let $\mathcal{C}$ be a complex $\{f, g, h, \ldots p\}$. By $s\mathcal{C}$ is meant the complex $\{sf, sg, sh, \ldots, sp\}$, and correspondingly $\mathcal{C}s = \{fs, gs, \ldots, ps\}$. From this it is apparent that $t(s\mathcal{C}) = (ts)\mathcal{C}$, since in either event the elements of $\mathcal{C}$ are multiplied by ts.

If $\mathcal{C}$ is a subgroup, $s\mathcal{C}$ is a left coset, as before. Furthermore, if

$h \in \mathcal{C}$, $h\mathcal{C} = \mathcal{C}$, which has already been shown.

By the product of two complexes $\mathcal{C}\mathcal{D}$ one means the complex containing all the products of some element from $\mathcal{C}$ as a prefactor and some element from as a present factor. Thus if

$$\mathcal{C} = \{f, g, h, \ldots p\} \qquad \mathcal{D} = \{q, r, s, \ldots v\}$$

$$\mathcal{C}\mathcal{D} = \{fq, fr, \ldots, fv; gq, gr, \ldots gv; \ldots; pq, pr, \ldots pv\}$$

As was shown before, if $\mathcal{C}$ is a group, $\mathcal{C}\mathcal{C} = \mathcal{C}$.

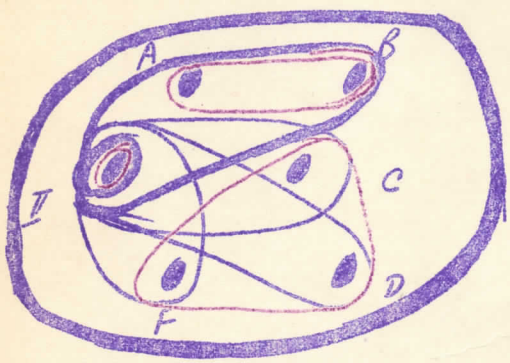
A subgroup whose left cosets coincide with its right cosets is called a normal divisor. Thus, for each $p \in \mathcal{G}$

$$np = pn$$

In this case

$$p^{-1}np = p^{-1}pn = n$$

Now, $p^{-1}\eta p$ separates η into classes, so that a normal divisor must contain an integral number of classes. This is borne out by the example we have been considering. The subgroups are enclosed in the black line; the classes by a red line. From this it appears that $\{\Pi, A, B\}$ is a normal divisor of the group, a conclusion borne but by calculation. In addition there are two other trivial normal divisors, the subgroup containing only the identity, and the whole group itself.



A group having no other normal divisor than the trivial ones is called simple. Apparently all groups of prime order are simple.

The reason that the normal divisors of a group are important is that their cosets form a group under complex multiplication, and that this group, called the factor group, and designed by G/η has certain important properties.

That the cosets form a group may be readily seen. The product of two such cosets is again a coset, since

$$\begin{aligned}(p\eta)(q\eta) &= (\eta p)(q\eta)\\ &= \eta(pq\eta)\\ &= \eta\eta pq\\ &= \eta(pq) = (pq)\eta\end{aligned}$$

and the coset multiplication is closed. It is associative since the multiplication of the elements themselves has this property. The coset with e is the identity.

$$(e\eta)(q\eta) = q\eta$$

while the inverse of g $\mathcal{H}$ is the coset formed by $g^{-1}\mathcal{H}$, since

$$(g^{-1}\mathcal{H})(g\mathcal{H}) = \mathcal{H}g^{-1}g\mathcal{H} = \mathcal{H}$$

and all the group postulates have been verified.

It was claimed in the beginning that a group was a concept so defined as to caputre the essence of multiplication, without specific reference to the nature of the quantities being discussed other than their multiplicative properties. Thus it may happen that many concrete objects behave as do the elements of a certain group. This causes one to attempt to identify the various elements with one another. Such an identification is usually called a mapping. The mapping is to be made in such a way that it preserves multiplication. Thus if one has a mapping $a \to a^1$, $b \to b^1$, etc., it is required that

$$(ab)^1 = a^1 b^1$$

such a mapping is called an homorphism, and implies, in a sense, that one really is calling the same thing by a different name, in exhibiting two homomorphic groups.

It is readily seen that the image of the identity under an homorphism is a new identity since

$$a^1 = (ea)^1 = e^1 a^1$$

and that inverses go over into inverses, since

$$e^1 = (a^{-1}a)^1 = a^{-1}a^1 = (a^1)^{-1}a^1$$

the associative law is also preserved, since multiplication,is, as well as closure. Hence the homorphic image of a group is also a group, which is not surprising since homorphism was specifically designed to give such a relationship.

The mapping may or may not be one to one. Thus to each group element

there may be associated one or more elements of the new group. In the special case that there is but one image the mapping is reversible. Then one speaks of isomorphism.

The importance of the factor group lies in the fact that if a group is homorphic to another, the other must be ismorphic to one of its factor groups. If they are isomorphic as it stands, the factor group must be the factor group of the subgroup containing only the identity. Consider the elements of the group $\mathcal{G}$ which go over into the identity of the new group $\bar{\mathcal{G}}$. The unit surely has this property as well as certain other elements, say $e_1, e_2, \cdots, e_n$

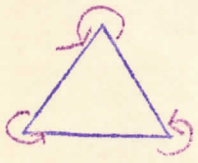
$$e, e_1, e_2, \cdots, e_n \to \bar{e}$$

call this complex $\mathcal{E}$. $\mathcal{E}$ is surely a group, since $\overline{e_i e_j} = \bar{e}_i \bar{e}_j = \bar{e}$ the product of two elements sent into the unit is also sent into the unit, and $\mathcal{E}$ is closed. The elements of $\mathcal{E}$ inherit the associative law from $\mathcal{G}$: Since $e \to \bar{e}$ it contains a unit and since $e^{-1} \to \bar{e}^{-1} = \bar{e}$, it contains inverses. Therefore $\mathcal{E}$ is a group, a subgroup of $\mathcal{G}$. It is also a normal divisor, since $p^{-1} \mathcal{E} p$ contains elements $p^{-1} e_i p$ which are imaged into $\bar{p}^{-1} \bar{e}_i \bar{p} = \bar{p}^{-1} \bar{e} \bar{p} = \bar{e}$ so that $p^{-1} \mathcal{E} p = \mathcal{E}$

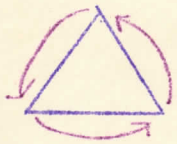
Now if $g \to \bar{g}$, so does $g \mathcal{E}$ since $g e_i \to \bar{g} \bar{e}_i = \bar{g}$. Thus each element of the cosets of $\mathcal{E}$ is mapped into a single image element. But there cosets are just the elements of the factor group, so that the factor group $\mathcal{G}/\mathcal{E}$ is isomorphic to $\bar{\mathcal{G}}$ as may be verified by writing down the mapping. This may be diagrammed:

$$\underbrace{e\ g_2 \cdots g_n}_{\bar{e}}\ ;\ \underbrace{g_{n+1} \cdots g_{2n}}_{\bar{g}_2}\ ;\ \cdots\ ;\ \underbrace{g_{(n-1)n} \cdots g_{kn}}_{\bar{g}_n}$$

To illustrate these things consider again the example we have been using. Although it was introduced as a pure matrix group, it is isomorphic to the group of symmetries of an equilateral triahgle.

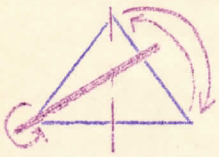


identity



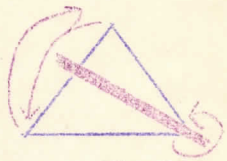
120° rotation

-120° rotation



mirror reflection

mirror reflection



mirror reflection

This may be seen by drawing a triangle in space and observing that the

matrices, operating upon $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ affect it as in the diagram above.

The subgroups are the whole group; the identity; the three rotations of 0°, 120°, 240°; and the various reflections in the symmetry axes. The reflections are of order two, since the square of a reflection, i.e., a reflection applied twice, gives the identity, while the rotations are of order 3, since 3 120° rotations in either direction are required to bring the figure back to the original position.

The classes however represent symmetries which may be transformed into one another operations entirely within the group itself, it being recalled that it takes two matrices to transform another, by taking $P^{-1}AP$. The unit of course is transformed into itself in this way, since the unit matrix is the same in each coordinate system. The various reflective symmetries

are gotten from each other by 120° rotations or by the identity transformation, since one merely rotates the symmetry plane. This gives clearly the class $\{C, D, F\}$. On the other hand, the two non zero rotations go over into each other, even under reflections in the various planes, so that $\{A, B\}$ is another class.

The normal divisor is the subgroup of the rotations.

If one takes the determinant of each of the matrices, he obtains an holomorphic image of the factor group

$$|\mathbb{I}| = 1 \qquad |C| = -1$$
$$|A| = 1 \qquad |D| = -1$$
$$|B| = 1 \qquad |F| = -1$$

so that $\mathcal{G}/\mathcal{H} \cong \{1, -1\}$. This is apparently the separation of the symmetry group into the rotations and the reflections with the multiplication table.

	rotation	reflection
rotation	rotation	reflection
reflection	reflection	rotation

The other two factor groups are the group itself $\mathcal{G}/\{\mathbb{I}\}$ and $\{\mathbb{I}\} = \mathcal{G}/\mathcal{G}$ It seems that the name "factor" group comes from some idea of having "divided out" the normal group from the rest of the group, leaving the cosets behind.

GROUP REPRESENTATIONS

The purpose of introducing some of the theory of groups has been that if a set of matrices is isomorphic to some group, it will be possible to deduce further information concerning the matrices by applying some of the concepts of group theory.

A set of matrices isomorphic to a certain group $\mathcal{G}$ is called a matrix representation of the group. If the isomorphism is actually a holomorphism,

and to each group element there belongs a separate matrix, the representation is called faithful. If the isomorphism is multiple valved, then the representation must be a faithful representation of a factor group of $\mathcal{G}$.

Now, by the group postulates, the matrices in question must be square and non-singular, that closure and the presence of inverses hold. Likewise the unit matrix must be present in the representation. Since the matrices are square it is possible to speak of the dimensionality of the representation with respect to the order of the matrices in question.

Two representations of a group which differ only by a similarity transformation are called equivalent.

A representation is called irreducible if it is impossible to partially diagonalize all the matrices of the representation simultaneously. Otherwise the representation is called reducible.

By the previous consideration of isomorphisms, appears that the matrix representing the identy must be the unit matrix, and that the representations of inverse elements are inverse matrices.

Let the matrices in question be indexed by their group elements. Then

$$\mathbb{D}(a)\mathbb{D}(a^{-1}) = \mathbb{D}(a a^{-1}) = \mathbb{D}(a) = \mathbb{I}$$

$$\therefore \mathbb{D}(a^{-1}) = \mathbb{D}(a)^{-1}$$

Each matrix representation is equivalent to a representation with unitary matrices. For let

$$H = \sum_{a \in \mathcal{G}} \mathbb{D}(a)\, \overline{\mathbb{D}(a)}^{*}$$

since H is hermitean, diagonalize it by S, a unitary matrix.

$$S^{-1} H S = \Lambda$$

Now, H is positive definite, so that all its eigenvalues are positive. Thus one can form

$\Lambda^{1/2}$

by taking the positive square roots of the eigenvalues. Then

$$\Pi = \Lambda^{-1/2} \sum_{a \in g} S^{-1} \mathcal{D}(a) S \overline{S^{-1} \mathcal{D}(a) S}^{*} \Lambda^{-1/2}$$

then $\Lambda^{-1/2} S^{-1} \mathcal{D}(a) S \Lambda^{1/2}$ will be the desired unitary matrix.

This is true since

$$\Lambda^{-1/2} S^{-1} \mathcal{D}(a) S \Lambda^{1/2} \quad \overline{\Lambda^{-1/2} S^{-1} \mathcal{D}(a) S \Lambda^{1/2}}^{*}$$

$$= \Lambda^{-1/2} S^{-1} \mathcal{D}(a) S \Lambda^{1/2} \Lambda^{1/2} S^{-1} \overline{\mathcal{D}(a)}^{*} S \Lambda^{-1/2}$$

$$= \Lambda^{-1/2} S^{-1} \mathcal{D}(a) \{ \sum_{b \in g} \mathcal{D}(b) \overline{\mathcal{D}(b)}^{*} \} \overline{\mathcal{D}(a)}^{*} S \Lambda^{-1/2}$$

$$= \Lambda^{-1/2} S^{-1} \{ \sum_{b \in g} \mathcal{D}(a) \mathcal{D}(b) \overline{\mathcal{D}(b)}^{*} \overline{\mathcal{D}(a)}^{*} \} S \Lambda^{-1/2}$$

$$= \Lambda^{-1/2} S^{-1} \{ \sum_{b \in g} \mathcal{D}(ab) \overline{\mathcal{D}(ab)}^{*} \} S \Lambda^{-1/2}$$

$$= \Lambda^{-1/2} S^{-1} H S \Lambda^{-1/2}$$

$$= \Pi \qquad \text{Q. E. D.}$$

Thus $S \Lambda^{1/2}$ transforms the representation. The proof of this theorem is certainly not obvious intuitively, although it is not contradicted by the remark previously made that the absolute values of a matrices which are elements of a finite group are 1, a fact also true of unitary matrices.

It was shown that two matrices commute only when they have a set of coincident eigenvectors, in which case they could be simultaneously diagonalized. Now, in a group of irreducible matrices not only can the matrices not all be diagonalized, they cannot even be partially diagonalized. This means that their eigenvectors must certainly present a formidable array, and that only the unit matrix or its scalar multiples may commute with all of them. The proof follows the lines of the commutability proof. Assume that the representation

is irreducible and has been brought to the unitary form. Suppose there were a matrix M such that

$$\mathbb{D}(a)\, M = M \mathbb{D}(a)$$

then also $$\bar{M}^{*}\, \overline{\mathbb{D}(a)}^{*} = \overline{\mathbb{D}(a)}^{*}\, \bar{M}^{*}$$

or, since $\mathbb{D}(a)$ is unitary

$$\mathbb{D}(a)\, \bar{M}^{*} = \bar{M}^{*}\, \mathbb{D}(a)$$

and the hermitean conjugate of M commutes with all the $\mathbb{D}$'s. Then so does $\frac{M+\bar{M}^{*}}{2}$ and $\frac{M-\bar{M}^{*}}{2i}$, the hermitean and antihermitean parts of M, respectively. Thus it suffices to prove the proposition for a hermitean matrix N. The advantage of choosing a hermitean matrix is that it may be diagonalized, say by T.

$$\Lambda = T^{-1} N T$$

Consider now the matrices

$$\mathbb{L}(a) = T^{-1}\, \mathbb{D}(a)\, T$$

which must commute with Λ, since if

$$\mathbb{D}(a) N = N \mathbb{D}(a$$

$$T^{-1}\, \mathbb{D}(a) T T^{-1} N T = T^{-1} N T T^{-1}\, \mathbb{D}(a)\, T$$

$$\mathbb{L}(a)\, \Lambda = \Lambda\, \mathbb{L}(a)$$

writing this out elementwise

$$\begin{bmatrix} \lambda_1 l_{11}(a) & \lambda_2 l_{12}(a) & \cdots & \lambda_n l_{1n}(a) \\ \lambda_1 l_{21}(a) & \lambda_2 l_{23}(a) & \cdots & \lambda_n l_{2n}(a) \\ \cdots & & & \\ \lambda_1 l_{n1}(a) & \lambda_2 l_{n2}(a) & \cdots & \lambda_n l_{nn}(a) \end{bmatrix} = \begin{bmatrix} \lambda_1 l_{11}(a) & \lambda_1 l_{12}(a) & \cdots & \lambda_1 l_{1n}(a) \\ \lambda_2 l_{21}(a) & \lambda_2 l_{22}(a) & \cdots & \lambda_2 l_{2n}(a) \\ \cdots & & & \\ \lambda_n l_{n1}(a) & \lambda_n l_{nn}(a) & \cdots & \lambda_n l_{nn}(a) \end{bmatrix}$$

now

$$\lambda_j\, l_{ij}(a) = \lambda_i\, l_{ij}(a)$$

if

$$\lambda_j \neq \lambda_i \qquad l_{ij}(a) = 0 \qquad a \in g$$

also, $\ell_{ji}(a) = 0$

so that if one eigenvalue, say λi , were to differ from the rest, the row and column would be zero save for $\ell_{ii}(a)$, for all $a \in G$ and the matrices will be partially diagonalized, hence reducible, contrary to hypothesis. Thus all the λ's are equal, N is a multiple of the unit matrix.

Q.E.D.

The same proposition is true for representations of different dimensionality; namely,

Schur's Lemma: Let there be two representations $\{\mathbb{D}^{(1)}(a)\}$ and $\{\mathbb{D}^{(2)}(a)\}$ of the same group which are irreducible and let there be such a matrix M , that

$$M\mathbb{D}^{(1)}(a) = \mathbb{D}^{(2)}(a)M \qquad a \in G$$

with ℓ = dimension of $\mathbb{D}^{(1)}$ and ℓ_2 = dimension of $\mathbb{D}^{(2)}$. Then for $\ell_1 \neq \ell$, $M = 0$. For $\ell_1 = \ell_2$, either $M = 0$ or is a matrix with non-vanishing determinant and the two representations are equivalent. Again, prepare the representations by bringing them to unitary form.

$$R^{-1}MSS^{-1}\mathbb{D}^{(1)}(a)S = R^{-1}\mathbb{D}^{(2)}(a)RR^{-1}MS,$$

and consider $R^{-1}MS$ in place of M , together with the two representations in unitary form. Thus it will do no harm to take them as being already in that form.

$$M\mathbb{D}^{(1)}(a) = \mathbb{D}^{(2)}(a)M \qquad 1)$$

$$\overline{\mathbb{D}^{(1)}(a)}^{*}\,\overline{M}^{*} = \overline{M}^{*}\,\overline{\mathbb{D}^{(2)}(a)}^{*}$$

$$\mathbb{D}^{(1)}(a^{-1})\,\overline{M}^{*} = \overline{M}^{*}\,\mathbb{D}^{(2)}(a^{-1}) \qquad 2)$$

premultiplying 1) by $\overline{M}^{*}$. postmultiplying 2) by M

$$\overline{M}^{*}M\,\mathbb{D}^{(1)}(a) = \overline{M}^{*}\,\mathbb{D}^{(2)}(a)M$$

$$\mathbb{D}^{(1)}(a^{-1})\,\overline{M}^{*}M = \overline{M}^{*}\,\mathbb{D}^{(2)}(a^{-1})M$$

noticing that the name of the dummy variable is unimportant.

$$\bar{M}^* M \mathbb{D}^{(1)}(a) = \mathbb{D}^{(1)}(a) \bar{M}^* M$$

so that $\bar{M}^* M$ commutes with all the matrices of the representation (1) It thus satisfies all the conditions of the last proposition, and

$$\bar{M}^* M = \lambda \mathbb{I}$$

Now, when $l_1 \neq l_2$, either $\lambda = 0$, $\bar{M}^* M = 0$, whence $M = 0$ the condition in question being that the hermitean length of each of the columns of M is zero, a condition which is fulfilled only if each column is identically zero. Otherwise $\lambda \neq 0$, $|M| \neq 0$, and M possesses an inverse, whereupon

$$M^{-1} \mathbb{D}^{(2)}(a) M = \mathbb{D}^{(1)}(a)$$

and the two representations are equivalent.

For $l_1 \neq l_2$, M is a non-square matrix with

$$\bar{M}^* M = \lambda \mathbb{I}_1$$

and

$$M \bar{M}^* = \mu \mathbb{I}$$

as may be seen by eliminating $\mathbb{D}^{(1)}(a)$ rather than $\mathbb{D}^{(2)}(a)$ from 1) and 2) of the last page. Now, one of the pair, $\bar{M}^* M$ is a column of short rows if M is not square. Say it is M , for the sake of argument. It was proven that $|\bar{M} M|$ gave the square of the volume enclosed by the rows of M . Since there are more rows than the dimensionality of the space this volume must be zero, since the figure is by necessity flat. This was not shown of $|\bar{M}^* M|$, but the proof can be readily extended to include a sort of "hermitean volume". Thus $|\bar{M}^* M| = 0 = \lambda^{l_1}$, say; hence $\lambda = 0$, and by a previous argument $M = 0$

Q.E.D.

With the aid of these propositions certain orthogonality relations may

be deduced for the elements of the matrices of an irreducible representation. With this object in view we notice that

$$M = \sum_{a \in g} \mathfrak{D}^{(2)}(a)\, Q\, \mathfrak{D}^{(1)}(a)^{-1}$$

with Q an arbitrary matrix conformable with $\mathfrak{D}^{(1)}(a)$ and $\mathfrak{D}^{(1)}(a)$ is such a matrix as contemplated in Schur's lemma.

$$\mathfrak{D}^{(2)}(p) M = \sum_{a \in g} \mathfrak{D}^{(2)}(p)\, \mathfrak{D}^{(2)}(a)\, Q\, \mathfrak{D}^{(1)}(a)^{-1}$$

$$= \sum_{a \in g} \mathfrak{D}^{(2)}(pa)\, Q\, \mathfrak{D}^{(1)}(pa)^{-1} \mathfrak{D}^{(1)}(p)$$

$$= M \mathfrak{D}^{(1)}(p)$$

this means that either the dimensionality of $\mathfrak{D}^{(1)}$ and $\mathfrak{D}^{(2)}$ are the same, and $M = \lambda \mathbb{I}$, or the dimensionalities are different, and $M = 0$

$$[M]_{il} = 0 = \sum_{a \in g} \sum_{j} \sum_{k} [\mathfrak{D}^{(2)}(a)]_{ij} [Q]_{jk} [\mathfrak{D}^{(1)}(a)^{-1}]_{kl}$$

This is true for each Q of the right order, so take $[Q]_{ij} = 0$ except $[Q]_{ik}$

$$0 = \sum_{a \in g} [\mathfrak{D}^{(2)}(a)]_{io} [\mathfrak{D}^{(1)}(a)^{-1}]_{xj}$$

If the $\mathfrak{D}$'s have been brought to the unitary form, this reads

$$0 = \sum_{a \in g} [\mathfrak{D}^{(2)}(a)]_{ir} [\mathfrak{D}^{(1)}(a)]^{*}_{jt}$$

In either case if one imagines a vector space in which the coordinates have been named after group elements, those relations require certain vectors in "group space" to be perpendicular. These vectors are corresponding elements of the matrices of representation belonging to different group elements.

From the other proposition; that the matrix which commutes with each element of an irreducible representation is a multiple of the unit matrix

follows another orthogonality relation. Namely, such a matrix M is

$$\sum_{a \in \mathcal{G}} \mathbb{D}^{(1)}(a)\, Q\, \mathbb{D}^{(1)}(a)^{-1}$$

since
$$\mathbb{D}^{(1)}(p)\, M = \sum_{a \in \mathcal{G}} \mathbb{D}^{(-1)}(p)\mathbb{D}^{(1)}(a)\, Q\, \mathbb{D}^{(-1)}(a^{-1})$$
$$= \sum_{a \in \mathcal{G}} \mathbb{D}^{(1)}(pa)\, Q\, \mathbb{D}^{(1)}(pa)^{-1}\mathbb{D}^{(1)}(p)$$
$$= M\, \mathbb{D}^{(1)}(p)$$

thus
$$M = \lambda \mathbb{I}$$
$$\lambda\, \delta_{\alpha\beta} = \sum_{a \in \mathcal{G}} \sum_{\mu} \sum_{\nu} [\mathbb{D}^{(1)}(a)]_{\alpha\mu}\, [Q]_{\mu\nu}\, [\mathbb{D}^{(1)}(a^{-1})]_{\nu\beta}$$

where again we may set
$$[Q]_{\mu\nu} = \delta_{\gamma\mu}\, \delta_{\eta\nu}$$

$$\lambda_{\gamma\eta}\, \delta_{\alpha\beta} = \sum_{a \in \mathcal{G}} [\mathbb{D}^{(1)}(a)]_{\alpha\gamma}\, [\mathbb{D}^{(1)}(a^{-1})]_{\eta\beta}$$

where $\lambda_{\gamma\eta}$ has been so indexed since the constant in question may very well depend upon which $[Q]_{\mu\nu}$ we set equal to 1,

to evaluate λ, set $\alpha = \beta$, sum,

$$\sum_{\alpha=1}^{\ell} \lambda_{\gamma\eta}\, \delta_{\alpha\alpha} = \sum_{a \in \mathcal{G}} \sum_{\alpha} [\mathbb{D}^{(1)}(a)]_{\alpha\gamma}\, [\mathbb{D}^{(1)}(a^{-1})]_{\eta\alpha}$$
$$= \sum_{a \in \mathcal{G}} [\mathbb{D}^{(1)}(e)]_{\gamma\eta} = \sum_{a \in \mathcal{G}} \delta_{\gamma\eta}$$

If ℓ is the dimension of the representation, g is the order of the group, this gives

$$\lambda_{\gamma\eta}\, \ell = g\, \delta_{\gamma\eta}$$
$$\lambda_{\gamma\eta} = \frac{g}{\ell}\, \delta_{\gamma\eta}$$

This gives

$$\frac{g}{\ell}\,\delta_{\gamma\eta}\,\delta_{\alpha\beta} = \sum_{a\in g} [\mathbb{D}^{(i)}(a)]_{\alpha\gamma}\,[\mathbb{D}^{(i)}(a^{-1})]_{\eta\beta}$$

which, if the representation is unitary, may be written

$$\frac{g}{\ell}\,\delta_{\gamma\eta}\,\delta_{\alpha\beta} = \sum_{a\in g} [\mathbb{D}^{(i)}(a)]_{\alpha\gamma}\,[\mathbb{D}^{(i)}(a)]^{*}_{\beta\eta}$$

This and the preceding formula may be combined. Suppose that there are several irreducible representations of a group, $\mathbb{D}^{(1)}$ of order ℓ_1 ; $\mathbb{D}^{(2)}$ of order ℓ_2 , --- , $\mathbb{D}^{(n)}$ of order ℓ_n . Suppose furthermore that they have been brought to the unitary form. Then

$$\sum_{a\in g} [\mathbb{D}^{(j)}(a)]_{\mu\nu}\sqrt{\frac{\ell_j}{g}}\cdot[\mathbb{D}^{(j')}(a)]^{*}_{\mu'\nu'}\sqrt{\frac{\ell_{j'}}{g}} = \delta_{jj'}\,\delta_{\mu\mu'}\,\delta_{\nu\nu'}$$

There are just $\ell_1^2 + \ell_2^2 + \cdots + \ell_n^2$ different matrix elements with which to form such vectors, but no more than g vectors may stand mutually perpendicular in group space. Thus $\ell_1^2 + \cdots + \ell_n^2$ can be no greater than g , thereby limiting the possible dimensionalities of irreducible representations.

The traces of the matrices of a representation gives the quantities known as group characters. The set of numbers $\{\chi^{(j)}(e), \chi^{(j)}(a), \cdots, \chi^{(j)}(g_{g})\}$ is called the character of the representation $[\mathbb{D}^{(j)}(a)]$, where

$$\chi^{(j)}(a) = \sum_{\mu=1}^{\ell_j} [\mathbb{D}^{(j)}(a)]_{\mu\mu}$$

$$= \mathcal{L}\,\mathbb{D}^{(j)}(a)$$

As we have already seen, the trace of each matrix of a class is the same, so that the character of a class contains just the one common trace. Furthermore, since traces are invarient under similarity transformation, the characters

of equivalent representations are the same.

There is an orthogonality relations for the traces just as there is for the matrix elements. Namely, since

$$\sum_{a\in g} [D^{(j)}(a)]_{\alpha\alpha} [D^{(j')}(a)]^*_{\beta\beta} = \frac{g}{\ell_j} \delta_{\alpha\beta}\delta_{jj'}$$

$$\sum_{a\in g} \chi^{(j)}(a)\chi^{(j')}(a)^* = \sum_{\alpha}\sum_{\beta} \frac{g}{\ell_j}\delta_{\alpha\beta}\delta_{jj'}$$

$$= \delta_{jj'}\frac{g}{\ell_j}\sum_{\alpha} 1$$

$$\sum_{a\in g} \chi^{(j)}(a)\chi^{(j')}(a)^* = g\,\delta_{jj'}$$

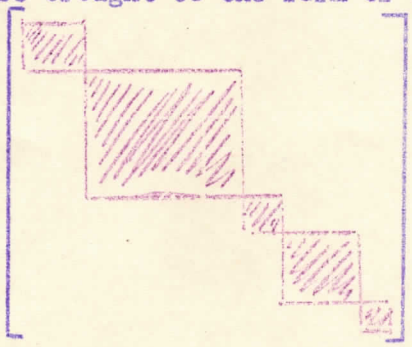
Suppose one now sums over classes, weighting each trace by the number of group elements in that class,

$$\sum_{c\,c\in g} \chi_c^{(j)}(a)\cdot\chi_c^{(j')}(a)^* \rho_c = g\,\delta_{jj'}$$

however, now there are only k summands, if there are k classes in g, so that there are now at most k different sets of class characters possible. Since all equivalent representations have the same character, it follows that there are no more than k different irreducible representations.

An interesting consequence of this is that, with an abelian group of order g there are g classes. Thus there are g different irreducible representations; but from the last orthogonality relation, $\ell_1^2, \ell_2^2, \ldots \ell_g^2 = g$ hence $\ell_i = 1$, and all the irreducible representations are one dimensional. But this corresponds to the fact that all the matrices of the representation commute, hence can be simultaneously diagonalized, with 1×1 matrices or numbers in their diagonals.

These orthogonality relations for the character of a group finally enables one to partially diagonalize a given representation as far as possible. This provides a sort of canonical form for representations, just as the diagonal form is a canonical form for individual matrices. Thus the matrices of the representation may be brought to the form of setep matrices.



Now, the trace of such a matrix is surely the sum of the traces of the submatrices, and is invariant under a similarity transformation, so that this relation holds even when $\mathbb{D}(a)$ is not partially diagonalized. Thus

$$\chi(a) = \sum_{j=1}^{n} \chi^{(j)}(a)\, c_j$$

where $\chi^{(j)}$ is the trace of the j^{th} irreducible submatrix. However

$$\sum_{a \in g} \chi(a)\, \chi^{(k)}(a)^* = \sum_{a \in g} \sum_{j} c_j \chi^{(j)}(a)\, \chi^{(k)}(a)^*$$

$$= g\, c_k$$

so that the number of times which a given irreducible representation occurs is given by

$$c_k = \frac{1}{g} \sum \chi(a)\, \chi^{(k)}(a)^*$$

LINEAR DIFFERENTIAL EQUATIONS

Let there be an n^{th} order linear homogeneous differential equation, such as

$$a_n(t)\frac{d^n x(t)}{dt^n} + a_{n-1}(t)\frac{d^{n-1}x(t)}{dt^{n-1}} + \cdots + a_0(t)x(t) = 0$$

Dividing through by $a_n(t) \neq 0$, setting

$$\alpha_1(t) = -\frac{a_{n-1}(t)}{a_n(t)}; \quad \alpha_2(t) = -\frac{\alpha_{n-2}(t)}{a_n(t)}; \quad \cdots$$

this becomes

$$\frac{d^n x}{dt^n} = \alpha_1\frac{d^{n-1}x}{dt^{n-1}} + \alpha_2\frac{d^{n-2}x}{dt^{n-2}} + \cdots + \alpha_n x$$

which may be rewritten

$$\frac{d}{dt}\left(\frac{d^{n-1}x}{dt^{n-1}}\right) = \alpha_1\frac{d^{n-1}x}{dt^{n-1}} + \cdots + \alpha_n x$$

while the following equations are trivally true

$$\frac{d}{dt}\left(\frac{d^{n-2}x}{dt^{n-2}}\right) = \frac{d^{n-1}x}{dt^{n-1}}$$

$$\frac{d}{dt}\left(\frac{d^{n-3}x}{dt^{n-3}}\right) = \frac{d^{n-2}x}{dt^{n-2}}$$

$$\cdots$$

$$\frac{d}{dt}\left(\frac{d^{n-n}x}{dt^{n-n}}\right) = \frac{dx}{dt}$$

These equations may all be gathered into a single matrix equation

$$\frac{d}{dt}\begin{bmatrix} \frac{d^{n-1}x}{dt^{n-1}} \\ \frac{d^{n-2}x}{dt^{n-2}} \\ \vdots \\ \frac{d^{0}x}{dt^{0}} \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{n-1} & \alpha_n \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{d^{n-1}x}{dt^{n-1}} \\ \frac{d^{n-2}x}{dt^{n-2}} \\ \vdots \\ x \end{bmatrix}$$

which is a matrix equation of a type that is readily solvable. Write it in the shorter form:

$$\frac{d\mathcal{X}}{dt} = M\mathcal{X}$$

If the matrix M commutes with itself for all values of t, such as will be the case when the α's are constants, the solution is

$$\mathcal{X}(t) = e^{\int_{t_0}^{t} M(\sigma)\,d\sigma}\,\mathcal{X}(t_0)$$

Supposing for the moment that the equation in question has constant coefficients, the solution reduces to

$$\mathcal{X}(t) = e^{M\cdot(t-t_0)}\,\mathcal{X}(t_0)$$

The matrix exponential may be calculated with the aid of Sylvester's theorem. The eigenvalues of M satisfy

$$\begin{bmatrix} \alpha_1 - \lambda & \alpha_2 & \cdots & \alpha_{n-1} & \alpha_n \\ 1 & -\lambda & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & & \\ 0 & 0 & \cdots & 1 & -\lambda \end{bmatrix} = 0$$

Expansion of this determinant according to the first row shows the characteristic equation to be

$$(-\lambda)^n + \alpha_1 (-\lambda)^{n-1} - \alpha_2 (-\lambda)^{n-2} + \cdots + \alpha_n = 0$$

or

$$\lambda^n - \{\alpha_1 \lambda^{n-1} + \alpha_2 \lambda^{n-2} + \cdots + \alpha_n\} = 0$$

This equation has certain roots, λ_i. It will be noted that this equation is structurally quite similar to the original differential equation; in fact, $d^k x/dt^k$ has been replaced by λ^k. This is called the characteristic equation of the differential equiation, and its roots are the eigenvalues of the matrix M. Thus the solution of the differential equiation is

$$\xi(t) = \sum_{(i)}^{n} e^{\lambda_i (t-t_0)} |i\rangle\langle i| \xi(t_0)$$

and

$$x(t) = \sum_{(i)}^{n} e^{\lambda_i (t-t_0)} \langle n; i\rangle\langle i| \xi(t_0)\rangle$$

$$= \sum_{(i)}^{n} c_i e^{\lambda_i (t-t_0)}$$

where the c's are determined from the initial conditions $\xi(t_0)$ and the eigenvectors; but may usually be found more readily if the c's are treated as arbitrary constants, whose values are to be found by substitution. However, the form of the solution in principle gives the solution in terms of the initial value of the function and its derivatives throughh the $(n-1)$ st. Likewise in the last form in which it was written, with arbitrary constants, it is apparent that there must be specified n constants to determine the solution.

The rule for determining the characteristic exponents is likewise the usual rule which is given; namely one substitutes an unknown for the deriva-

tive operator, solves the resulting algebraic equation, obtaining the roots of the equation as the numerical factors in the exponents.

This process is successful when all the eigenvalues of M are distinct, for then M has a complete system of eigenvectors and Sylvester's theorem applies. The situation which occurs when an eigenvalues is repeated is illustrated by the following example of a second order differential equation:

$$\frac{d^2x}{dt^2} + 2\frac{dx}{dt} + 1 = 0$$

The matrix M is then

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

with $\lambda = -1$ as a double eigenvalue. It has the sole eigenvector

$$|1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

so that Sylvester's theorem will fail. However the degeneracy is clearly due to the particular values of the elements of the matrix, so that some insight into the failure of Sylvester's expansion may be gained by changing the elements slightly, and examining the result.

In this case we change the coefficients of the differential equation slightly. Write a new matrix

$$M' = \begin{bmatrix} -2 & -1+\varepsilon^2 \\ 1 & 0 \end{bmatrix}$$

corresponding to a differential equation

$$\frac{d^2x}{dt^2} + 2\frac{dx}{dt} + (1-\varepsilon^2) = 0$$

The eigenvalues of M' are the solutions of

$$\begin{bmatrix} -2-\lambda & -1+\varepsilon^2 \\ 1 & -\lambda \end{bmatrix} = 0$$

which gives $\lambda = -1 \pm \varepsilon$

with the eigenvectors

$$[1 \quad 1 \pm \varepsilon] \text{ and } \begin{bmatrix} 1 \\ -1 \\ \hline 1 \pm \varepsilon \end{bmatrix}$$

Thus it is clear that as $\varepsilon \to 0$, the two eigenvectors coalesce, to give but one distinct eigenvector for the matrix M . However, if M' is expanded by Sylvester's theorem, there results

$$M' = \frac{-1+\varepsilon}{1-\frac{1+\varepsilon}{1-\varepsilon}} \begin{bmatrix} 1 & 1+\varepsilon \\ -\frac{1}{1-\varepsilon} & -\frac{1+\varepsilon}{1-\varepsilon} \end{bmatrix} + \frac{-1-\varepsilon}{1-\frac{1-\varepsilon}{1+\varepsilon}} \begin{bmatrix} 1 & 1-\varepsilon \\ -\frac{1}{1+\varepsilon} & -\frac{1-\varepsilon}{1+\varepsilon} \end{bmatrix} .$$

where the terms $\left\{1-\frac{1+\varepsilon}{1-\varepsilon}\right\}$ and $\left\{1-\frac{1-\varepsilon}{1+\varepsilon}\right\}$ arise from normalizing the projective operators. Clearing fractions this gives

$$\frac{(1-\varepsilon)^2}{2\varepsilon} \begin{bmatrix} 1 & 1+\varepsilon \\ -\frac{1}{1-\varepsilon} & -\frac{1+\varepsilon}{1-\varepsilon} \end{bmatrix} + \frac{-(1+\varepsilon)^2}{2\varepsilon} \begin{bmatrix} 1 & 1+\varepsilon \\ -\frac{1}{1+\varepsilon} & -\frac{1-\varepsilon}{1+\varepsilon} \end{bmatrix}$$

which adds up correctly to give M' . If we now calculate $e^{M'(t-t_0)}$, we find

$$\frac{e^{(-1+\varepsilon)(t-t_0)}}{1-\frac{1+\varepsilon}{1-\varepsilon}}\begin{bmatrix} 1 & 1+\varepsilon \\ \frac{-1}{1-\varepsilon} & -\frac{1+\varepsilon}{1-\varepsilon} \end{bmatrix} + \frac{e^{(-1-\varepsilon)(t-t_0)}}{1-\frac{1-\varepsilon}{1+\varepsilon}}\begin{bmatrix} 1 & 1-\varepsilon \\ -\frac{1}{1+\varepsilon} & -\frac{1-\varepsilon}{1+\varepsilon} \end{bmatrix}$$

and multiplying this by $\begin{bmatrix} x(t_0) \\ \dot{x}(t_0) \end{bmatrix}$ to find the solution of the differential equation,

$$x(t) = \frac{e^{(-1+\varepsilon)(t-t_0)}}{2\varepsilon}\left\{\dot{x}(t_0)+(1+\varepsilon)x(t_0)\right\}$$

$$+\frac{e^{(-1-\varepsilon)(t-t_0)}}{-2\varepsilon}\left\{\dot{x}(t_0)+(1-\varepsilon)x(t_0)\right\}$$

$$=\frac{1}{2}x(t_0)\left\{e^{(-1+\varepsilon)(t-t_0)}+e^{(-1-\varepsilon)(t-t_0)}\right\}+\left\{x(t_0)+\dot{x}(t_0)\right\}e^{-1\cdot(t-t_0)}\left\{\frac{e^{\varepsilon(t-t_0)}-e^{-\varepsilon(t-t_0)}}{2\varepsilon}\right\}$$

Now, taking $\lim_{\varepsilon \to 0}$, the solution becomes

$$x(t) = x(t_0)e^{-1\cdot(t-t_0)} + x\left\{t_0+\dot{x}(t_0)\right\}e^{-1\cdot(t-t_0)}\left\{t-t_0\right\}$$

$$x(t) = e^{-1\cdot(t-t_0)}\left\{c_1+c_2 t\right\}$$

which is the familiar solution from the ordinary theory of linear differential equations with constant coefficients. The origin of the powers of as well as exponentials lies in the fact that two solutions coalesce, leaving an indeterminate form, which is evaluated by a process which amounts to differentiating the exponentials with respect to the eigenvalues, which brings down one or more powers of $(t-t_0)$. The example shown here may be extended to cases of higher degeneracy to show this. Since we are striving only for a verification of the ordinary theory here, we leave the details of

of such a proof to the reader.

For a linear equation without constant coefficients it will be necessary to use the series form for the solution to the matrix differential equation rather than the exponential form, due to commutation difficulties, a discussion which we also forsake as irrelevant here.

The same technique which converts a differential equation into a matrix equation may be used to solve recursion formulae, and is, of course, most useful for recursion formulae with constant coefficient. Thus suppose that there is given a recursion formula among unknown quantities $D_n \quad n = 1, 2, \cdots$

$$D_n = \alpha_1 D_{n-1} + \alpha_2 D_{n-2} + \cdots + \alpha_K D_{n-K}$$

to this may be added the trivial formulae

$$D_{n-1} = D_{n-1}$$

$$D_{n-2} = D_{n-2}$$

$$D_{n-K+1} = D_{n-K+1}$$

which has the matrix formula

$$\begin{bmatrix} D_n \\ D_{n-1} \\ \cdots \\ D_{n-K+1} \end{bmatrix} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_K \\ 1 & 0 & \cdots & 0 \\ \cdots & & & \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} D_{n-1} \\ D_{n-2} \\ \cdots \\ D_{n-K+1} \end{bmatrix}$$

$$\begin{aligned} \mathcal{D}_n &= M \mathcal{D}_{n-1} \\ &= M^2 \mathcal{D}_{n-2} \\ &= \cdots \\ &= M^n \mathcal{D}_0 \end{aligned}$$

diagonalizing M, this can be written

$$\mathcal{D}_n = \sum_{i=1}^{K} \lambda_i^n |i\rangle\langle i| \mathcal{D}_0$$

where λ_i is an eigenvalue of M. Since $D_n = \langle 1 | M^n | 1 \rangle$

$$D_n = \sum_{i=1}^{k} \lambda_i^{\,n} \langle 1 | i \rangle \langle i | 1 \rangle$$

$$= \sum_{i=1}^{k} c_i \lambda_i^{\,n}$$

Where $c_i = \langle 1 | i \rangle \langle i | 1 \rangle$, a constant which depends upon the boundary conditions. Thus it seems that, like a differential equation, this type of equation must be provided with k boundary values, usually values of D_n for lower n, and that there occurs in a canonical solution just k arbitrary constants which must be fitted to them. Just as with the linear differential equation with constant coefficients, the λ's are obtained from solving a characteristic equation

$$\lambda^n - a_1 \lambda^{k-1} - a_2 \lambda^{n-2} + \cdots - a_k \lambda^{n-k} = 0$$

which results from substituting λ^p for D_p in the recursion formula.

As an example of the matrix solution of recursion formulae consider a class of $n \times n$ determinants which will occur in the later text.

$$D_n = \begin{vmatrix} \alpha & 1 & 0 & \cdots & 0 \\ 1 & \alpha & 1 & \cdots & 0 \\ 0 & 1 & \alpha & \cdots & 0 \\ \cdots & \cdots & \cdots & & \\ 0 & 0 & 0 & \cdots & \alpha \end{vmatrix}$$

Expanding according to the first row, and recognizing similar determinants of lesser order, the following recursion relationsis obtained.

$$D_n = \alpha D_{n-1} - D_{n-2}$$

Now, $D_1 = \alpha$, $D_2 = \alpha^2 - 1$, from which one can deduce the simpler boundary conditions $D_0 = 1$, $D_{-1} = 0$.

The equations

$$D_{n+1} = \alpha D_n - D_{n-1}$$

$$D_n = D_n$$

form the matrix equation

$$\begin{bmatrix} D_{n+1} \\ D_n \end{bmatrix} = \begin{bmatrix} \alpha & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} D_n \\ D_{n-1} \end{bmatrix}$$

which gives

$$\begin{bmatrix} D_{n+1} \\ D_n \end{bmatrix} = \begin{bmatrix} \alpha & -1 \\ 1 & 0 \end{bmatrix}^{n+1} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

which will yield a solution as soon as the matrix power has been calculated.

The eigenvalues of the matrix are the solution of the equation

$$\begin{bmatrix} \alpha - \lambda & -1 \\ 1 & -1 \end{bmatrix} = 0$$

and are

$$\lambda = \frac{\alpha}{2} \pm \sqrt{\frac{\alpha^2}{4} - 1}$$

making the substitution

$$\frac{\alpha}{2} = \cosh \varphi$$

the eigenvalues become

$$\lambda_1 = e^{\varphi} \qquad \lambda_2 = e^{-\varphi}$$

The eigenvectors are readily determined to be

$$|1\rangle \sim \begin{bmatrix} 1 \\ e^{-\varphi} \end{bmatrix} \qquad |2\rangle \sim \begin{bmatrix} 1 \\ e^{\varphi} \end{bmatrix}$$

$$\langle 1| \sim [1, -e^{-\varphi}]$$

$$\langle 2| \sim [1, -e^{\varphi}]$$

so that Sylvester's theorem reads

$$\begin{bmatrix} \alpha & -1 \\ 1 & 0 \end{bmatrix}^{n+1} = \frac{e^{(n+1)\varphi}}{1-e^{-2\varphi}} \begin{bmatrix} 1 & -e^{-\varphi} \\ e^{-\varphi} & -e^{-2\varphi} \end{bmatrix} + \frac{e^{-(n+1)\varphi}}{1-e^{2\varphi}} \begin{bmatrix} 1 & -e^{+\varphi} \\ e^{\varphi} & -e^{2\varphi} \end{bmatrix}$$

so that the $(n+1)$ st power is

$$\frac{1}{\sinh \varphi} \begin{bmatrix} \sinh(n+2)\varphi & -\sinh(n+1)\varphi \\ \sinh(n+1)\varphi & -\sinh n\varphi \end{bmatrix}$$

and D_n is

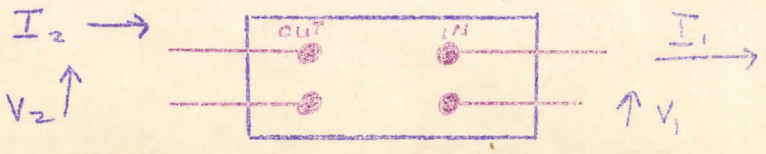
$$D_n = \frac{\sinh(n+1)\varphi}{\sinh \varphi} \qquad \varphi = \operatorname{arccosh} \frac{\alpha}{2}$$

THE TRANSMISSION LINE

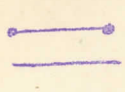
The solution of the electrical equations for the steady state for a transmission line illustrates the use which can be made of the concept of a "transfer" matrix, which, in the present case is used to calculate the voltage and current in one part of the transmission line in terms of the voltagee and current elsewhere. The procedure resembles that used in solving recursion formulae where one solves for certain of the unknowns in terms of others which are given.

The elemental calculations are done with certain simple four terminal networks. The entire transmission line is then synthesized from these elements. Four-terminal networks are networks with two "in" and two "out" terminals, and one strives to express the current and voltage at the output terminals in terms of the current and voltage of the input terminals.

Current and voltage are reckoned by the following scheme:

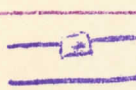


and the equations for typical elements are

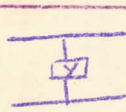


$V_2 = V_1$, $I_1 = I_1$ $\quad \begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$ — closed circuit

$V_2 = 0$, $I_2 = 0$ $\quad \begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$ — open circuit



$V_2 = V_1 + Z J_1$, $I_2 = J_1$ $\quad \begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} 1 & Z \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_1 \\ J_1 \end{bmatrix}$ — series impedence



$V_2 = V_1$, $J_2 = Y V_1 + I_1$ $\quad \begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ Y & 1 \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$ — parallel admittance

More complicated networks are formed by connecting the simple ones shown above in series, or in "cascade" as it is sometimes called. For instance, if two impedances are connected in series

$$\begin{bmatrix} V_3 \\ I_3 \end{bmatrix} = \begin{bmatrix} 1 & Z_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_2 \\ I_2 \end{bmatrix} \qquad 1)$$

$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} 1 & Z_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix} \qquad 2)$$

substituting 2) in 1)

$$\begin{bmatrix} V_3 \\ I_3 \end{bmatrix} = \begin{bmatrix} 1 & Z_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & Z_1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} 1 & Z_1 + Z_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$$

The impedances add, as would be expected, to write the matrices for a cascade network the matrices for the elements must be written down in order, from left to right, just as the elements appear in the networks, as is apparent in the following example

$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ Y & 1 \end{bmatrix} \begin{bmatrix} 1 & Z \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$$

the transfer matrix is

$$\begin{bmatrix} 1 & z \\ y & 1+yz \end{bmatrix}$$

while

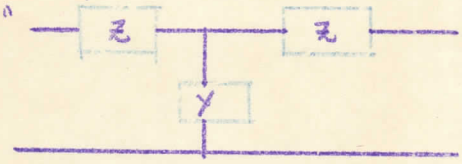
$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ y & 1 \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$$

gives a transfer matrix $\begin{bmatrix} 1+yz & z \\ y & 1 \end{bmatrix}$

and the electrical assymetry of the two networks is reflected in the non-commutativity of the matrices representing their elements.

Two common elements for transmission lines are the "T" and "π" sections, so named from their shape. Their transfer matrices are:

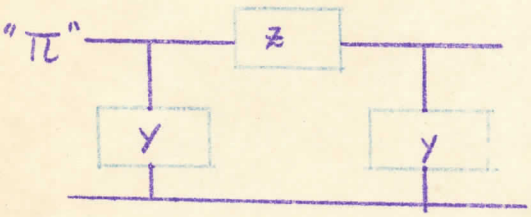
"T"



$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ y & 1 \end{bmatrix} \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$$

$$= \begin{bmatrix} 1+yz & z(2+yz) \\ y & 1+yz \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$$

"π"



$$\begin{bmatrix} V_2 \\ I_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ y & 1 \end{bmatrix} \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ y & 1 \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$$

$$= \begin{bmatrix} 1+yz & z \\ y(2+yz) & 1+yz \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$$

Considering the matrix for a π-section, its eigenvalues can be calculated to be

$$\lambda = (1+yz \pm \sqrt{(1+yz)^2 - 1}$$

Making the hyperbolic substitution

$$\cosh\varphi = 1 + \gamma z$$

they become

$$\lambda_1 = e^{\varphi} \qquad \lambda_2 = e^{-\varphi}$$

Making the additional substitution

$$z_0 = \sqrt{\frac{z}{\gamma}}\,\frac{1}{\sqrt{2+\gamma z}}$$

The eigenvectors are readily shown to be

$$|1\rangle = \frac{1}{\sqrt{2z_0}}\begin{bmatrix} z_0 \\ 1 \end{bmatrix} \qquad \langle 1| = \frac{1}{\sqrt{2z_0}}\begin{bmatrix} 1 & z_0 \end{bmatrix}$$

$$|2\rangle = \frac{1}{\sqrt{2z_0}}\begin{bmatrix} z_0 \\ -1 \end{bmatrix} \qquad \langle 2| = \frac{1}{\sqrt{2z_0}}\begin{bmatrix} 1 & -z_0 \end{bmatrix}$$

So that the transfer matrix is, by Sylvester's theorem

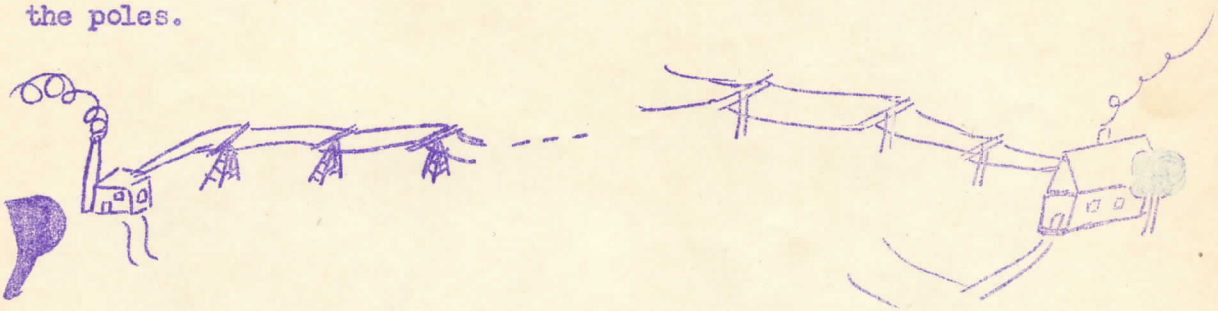
$$\begin{bmatrix} 1+\gamma z & z \\ \gamma(2+\gamma z) & 1+\gamma z \end{bmatrix} = \frac{e^{\varphi}}{2}\begin{bmatrix} 1 & z_0 \\ \frac{1}{z_0} & 1 \end{bmatrix} + \frac{e^{-\varphi}}{2}\begin{bmatrix} 1 & -z_0 \\ -\frac{1}{z_0} & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \cosh\varphi & z_0 \sinh\varphi \\ \frac{1}{z_0}\sinh\varphi & \cosh\varphi \end{bmatrix}$$

Either by Sylvester's theorem, or by direct calculation it is apparent that when the transfer matrix is written in this form it can be raised to powers quite easily since

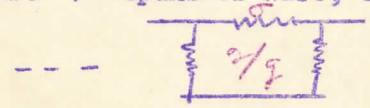
$$\begin{bmatrix} \cosh\varphi & z_0 \sinh\varphi \\ \frac{1}{z_0}\sinh\varphi & \cosh\varphi \end{bmatrix}^n = \begin{bmatrix} \cosh n\varphi & z_0 \sinh n\varphi \\ \frac{1}{z_0}\sinh n\varphi & \cosh n\varphi \end{bmatrix}$$

which is the transfer matrix for a transmission line composed of n Π-sections. As an example, consider a power line, in which the resistance of the spans of wire has been taken into account as well as the leakage of the insulators on the poles.



Suppose the resistance of a span of wire is r ohms, the resistance of the insulation to ground $1/g$ ohms. Schematically the line is

where the end leakages have been taken as $2/g$ ohms to simplify the problem. If there are n spans of wire, one has



with a transfer matrix

$$\begin{bmatrix} 1+\frac{rg}{2} & r \\ \frac{g}{2}\left(2+\frac{rg}{2}\right) & 1+\frac{rg}{2} \end{bmatrix}^n$$

set

$$\varphi = \operatorname{arccosh}\left(1+\frac{rg}{2}\right)$$

$$Z_0 = \sqrt{\frac{r}{g}}\,\frac{1}{\sqrt{1+rg/4}}$$

Suppose the problem which is given is the determination of the effective impedance of the line as seen from the powerhouse.

$$Z_{eff} = \frac{V_{PH}}{I_{PH}}$$

$$\begin{bmatrix} V_{PH} \\ I_{PH} \end{bmatrix} = \begin{bmatrix} \cosh n\varphi & Z_0 \sinh n\varphi \\ \frac{1}{Z_0}\sinh n\varphi & \cosh n\varphi \end{bmatrix} \begin{bmatrix} V_1 \\ I_1 \end{bmatrix}$$

$$Z_{eff} = \frac{V_1 \cosh n\varphi + Z_0 I_1 \sinh n\varphi}{\frac{V_1}{Z_0} \sinh n\varphi + I_1 \cosh n\varphi}$$

$$\lim_{n\to\infty} Z_{eff} = \frac{V_1 e^{n\varphi} + Z_0 I_1 e^{n\varphi}}{\frac{V_1}{Z_0} e^{n\varphi} + I_1 e^{n\varphi}} \qquad \left(e^{-n\varphi} \doteq 0\right)$$

$$= Z_0 = \sqrt{\frac{r}{g}} \frac{1}{\sqrt{1 + \frac{r}{g}\frac{1}{\varphi}}}$$

If, as is the case in practice, both r and g are small, The second radical is approximately 1, giving

$$Z_{eff} \doteq \sqrt{\frac{r}{g}}$$

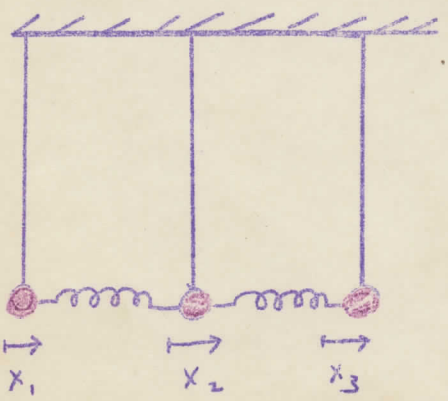
which is the geometric mean between the resistance of the spans and the resistance to ground. To choose some numbers, if $r = 0.1$ ohm, $g = 10^{-6}$ ohm, the resistance of the long line would be about 300 ohms.

THE COUPLED HARMONIC OSCILLATOR

The expansion of the elastic constant for a vibrational problem according to Sylvester's theorem has its interpretation in terms of normal modes. If a certain displacement is an eigenvector for the elastic constant, the restoring force will always be opposite in direction to the displacement, and proportional to it. The resulting motion is a particularly simple type of motion since then all the particles will vibrate with simple harmonic motion, and the "shape" of the motion, so to speak, will remain constant and only its amplitude will vary with time. Such motion is called a normal mode. The situation is frequently described by saying the "time is separable," meaning that the solution to the problem may be written as a product of two factors, one

depending only upon the time, and the other depending only upon the coordinates of the particles.

The displacements corresponding to eigenvectors can then be discovered by noticing which displacements of the particles give rise to forces always pushing back directly along the line of displacement. As an example consider three pendula whose hobs are all of the mase mass, m, and the lengths, ℓ, of whose suspensions are all the same; but which are, however, interconnected by springs of elastic constant k, as illustrated in the diagram. Considering only plane motion with small enough amplitude to justify the approximation $\theta \doteq \sin\theta$ the forces are:



$$f_1 = -kx_1 + kx_2 - \frac{mg}{\ell} x_1$$

$$f_2 = kx_1 - 2kx_2 + kx_3 - \frac{mg}{\ell} x_2$$

$$f_3 = kx_2 - kx_3 - \frac{mg}{\ell} x_3$$

Equating the restoring faces to the inertia forces, and writing the equations in matrix form:

$$\begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{bmatrix} = - \begin{bmatrix} \frac{k}{m} + \frac{g}{\ell} & -\frac{k}{m} & 0 \\ -\frac{k}{m} & \frac{2k}{m} + \frac{g}{\ell} & -\frac{k}{m} \\ 0 & -\frac{k}{m} & \frac{k}{m} + \frac{g}{\ell} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

which has the form

$$\ddot{X} = -K X \qquad 1)$$

The eigenvalues of K are the solutions of the equation

$$\begin{vmatrix} \frac{k}{m} + \frac{g}{\ell} - \lambda & -\frac{k}{m} & 0 \\ -\frac{k}{m} & \frac{2k}{m} + \frac{g}{\ell} - \lambda & -\frac{k}{m} \\ 0 & -\frac{k}{m} & \frac{k}{m} + \frac{g}{\ell} - \lambda \end{vmatrix} = \left(\frac{k}{m} + \frac{g}{\ell} - \lambda\right)^2 \left(\frac{2k}{m} + \frac{g}{\ell} - \lambda\right) - \frac{2k^2}{m^2}\left(\frac{k}{m} + \frac{g}{\ell} - \lambda\right) = 0$$

and are

$$\lambda_1 = \frac{g}{\ell}$$

$$\lambda_2 = \frac{g}{\ell} + \frac{k}{m}$$

$$\lambda_3 = \frac{g}{\ell} + \frac{3k}{m}$$

The eigenvectors can be calculated from the formula:

$$|i\rangle\langle i| = \frac{\prod_{j \neq i} (K - \lambda_j \mathbb{I})}{\prod_{j \neq i} (\lambda_i - \lambda_j)}$$

$$|1\rangle\langle 1| = \frac{(K - \lambda_2 \mathbb{I})(K - \lambda_3 \mathbb{I})}{(\lambda_1 - \lambda_2)(\lambda_1 - \lambda_3)}$$

$$= \frac{1}{\left(-\frac{k}{m}\right)\left(-\frac{3k}{m}\right)} \begin{bmatrix} 0 & -\frac{k}{m} & 0 \\ -\frac{k}{m} & \frac{k}{m} & -\frac{k}{m} \\ 0 & -\frac{k}{m} & 0 \end{bmatrix} \begin{bmatrix} -\frac{2k}{m} & -\frac{k}{m} & 0 \\ -\frac{k}{m} & -\frac{k}{m} & -\frac{k}{m} \\ 0 & -\frac{k}{m} & -\frac{2k}{m} \end{bmatrix}$$

$$= \frac{1}{3} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \therefore \quad \boxed{|1\rangle = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}$$

Likewise

$$|2\rangle\langle 2| = \frac{(K - \lambda_1 \mathbb{I})(K - \lambda_3 \mathbb{I})}{(\lambda_2 - \lambda_1)(\lambda_2 - \lambda_3)}$$

$$= \frac{1}{\frac{k}{m}\left(-\frac{2k}{m}\right)} \begin{bmatrix} -\frac{k}{m} & -\frac{k}{m} & 0 \\ -\frac{k}{m} & \frac{2k}{m} & -\frac{k}{m} \\ 0 & -\frac{k}{m} & \frac{k}{m} \end{bmatrix} \begin{bmatrix} -\frac{2k}{m} & -\frac{k}{m} & 0 \\ -\frac{k}{m} & -\frac{k}{m} & -\frac{k}{m} \\ 0 & -\frac{k}{m} & -\frac{2k}{m} \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\therefore \quad \boxed{|2\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}}$$

and

$$|3\rangle\langle 3| = \frac{(K-\lambda_1 \mathbb{I})(K-\lambda_2 \mathbb{I})}{(\lambda_3-\lambda_1)(\lambda_3-\lambda_2)}$$

$$= \frac{1}{\frac{3k}{m}\,\frac{2d}{m}} \begin{bmatrix} \frac{3k}{m} & -\frac{3k}{m} & 0 \\ -\frac{3k}{m} & \frac{3k}{m} & -\frac{3k}{m} \\ 0 & -\frac{3k}{m} & \frac{3k}{m} \end{bmatrix} \begin{bmatrix} 0 & -\frac{3k}{m} & 0 \\ -\frac{3k}{m} & \frac{3k}{m} & -\frac{3k}{m} \\ 0 & -\frac{3k}{m} & 0 \end{bmatrix}$$

$$= \frac{1}{6} \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

$$\therefore \quad |3\rangle = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

Therefore the matrix

$$U = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix}$$

diagonalizes K and

$$U^{-1}KU = \begin{bmatrix} g/\ell & 0 & 0 \\ 0 & g/\ell + k/m & 0 \\ 0 & 0 & g/\ell + 3k/m \end{bmatrix}$$

Under the substitution $U^{-1}X = Y$, equation 1) reads

$$\ddot{Y} = -U^{-1}KUY$$

which gives the following set of differential equations when the corresponding elements are equated.

$$\ddot{y}_1 = -\frac{g}{\ell}\, y_1$$

$$\ddot{y}_2 = -\left(\frac{g}{\ell} + \frac{k}{m}\right) y_2$$

$$\ddot{y}_3 = -\left(\frac{g}{\ell} + \frac{3k}{m}\right) y_3$$

which have as their solutions

$$y_1 = a_1 \sin(\sqrt{g/\ell}\, t + \delta_1)$$

$$y_2 = -a_2 \sin(\sqrt{g/\ell + g/m}\, t + \delta_2)$$

$$y_3 = a_3 \sin(\sqrt{g/\ell + 3g/m}\, t + \delta_3)$$

So that by making the unsubstitution $\mathcal{X} = U\mathcal{Y}$

$$\mathcal{X} = |1\rangle y_1 + |2\rangle y_2 + |3\rangle y_3$$

and one obtains the positions of the individual pendula as a function of time according to

$$x_i = \langle i|1\rangle y_1 + \langle i|2\rangle y_2 + \langle i|3\rangle y_3$$

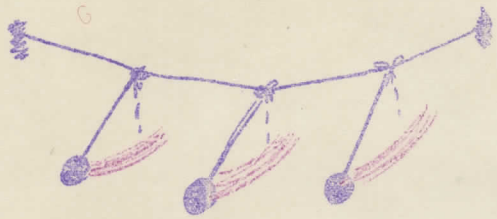
The quantities $y_1 = \langle 1|x\rangle$ are called normal coordinates, while motion described by just one of them is a normal mode. There is, of course, a geometrical interpretation of the normal coordinate as projections of the eigenvectors on the coordinate axes, but since it is difficult to see the connection of such a space directly with the motion of the particle, this interpretation does not have a direct physical meaning.

A slightly modified problem, which however has the same equation of motion as the one discussed consists in three identical pendula hung from a string which has been rather loosely stretched between two supports. The coupling between the pendula which swing in a plane perpendicular to the supporting string is provided through this support rather than by the connecting springs contemplated in the problem.

The normal mode of lowest frequency has the eigenvector

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

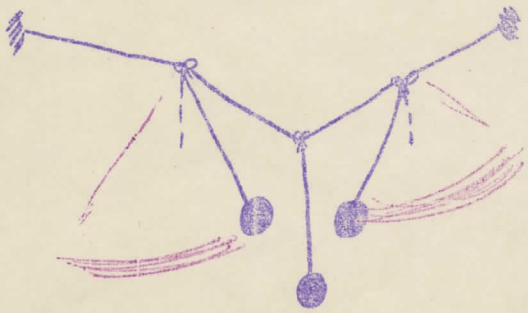
so that when it is excited all the pendula have the same displacement of any given time. Thus there is no force on the supports between pendula, and all three swing as a unit. Since there is no additional force they swing with the frequency that they would have when uncoupled; namely $\nu = \frac{1}{2\pi}\sqrt{g/\ell}$



The second eigenvector is

$$\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

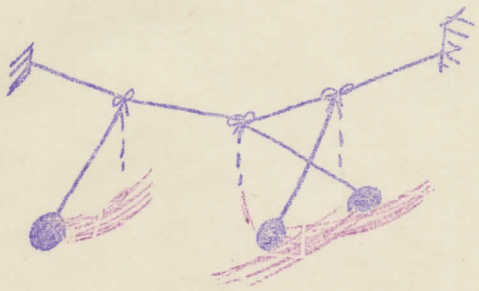
and describes motion in which the center hob is at rest and the other two are displaced oppositely, but by the same amount. There is no resultant force on the center hob, since the influences of the side hobs cancel, being equal, but oppositely directed. Nevertheless the end pendula manage to exert a force on one another such that one acts to retard the other. The action is symmetric between the end hobs, and altogether the condition for an eigenvector is seen to be fulfilled. The increased restoring force on the end hobs raises their frequency from the uncoupled $\nu = \frac{1}{2\pi}\sqrt{g/\ell}$ to $\frac{1}{2\pi}\sqrt{g/\ell + k/m}$.



The third eigenvector is

$$\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

so that in this case the two end hobs move together but with the center hob always displaced twice as far in the opposite direction. The frequency of this motion is still higher than that of the other normal modes due to the greater retaining faces exerted due to the coupling, and becomes $\frac{1}{2\pi}\sqrt{\frac{k}{m} + \frac{3k}{m}}$. Again inspection shows that the forces are always opposite and proportional to the displacements, so that the condition for eigenvectors of the elastic constant is again fulfilled.



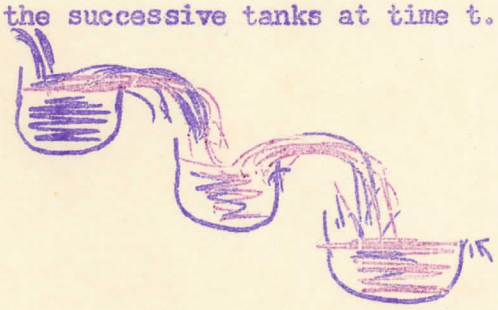
The fact that the motion is described in terms of eigenvectors of the elastic constant shows that motion originally started according to one normal mode can nveer proceed according to another. A rthaer striking example of this independence is afforded by considering the motion resulting after the center hob has been displaced and then released, the end hobs starting from rest. This is readily seen to be a combination of motion of the types 1 and 3, 2 being absent. Type 2 provides the motion in which the end hobs move in opposite direction, so it follows that from these initial conditions one will never expect to see the two end hobs going in opposite direction, a conclusion borne out by calculation and experiment. Actually the center hob will

move with gradually increasing amplitude, finally coming to rest when the outer hobs have taken up the motion, which in their turn begin to lose amplitude, setting the center hob to swinging once more, and so on.

A ROW OF TANKS

Suppose the first tank in a row contains a mixture of G-g gallons of liquid A and G gallons of liquid B. From time $t = 0$, liquid A is pumped into the first tank at the rate of r gallons per minute and a mixture of A and B is thus forced from each tank to the next at the same rate. Assuming that the tanks are perfectly stirred, calculate the amount of liquid B in any tank as a function of the time.

Such a problem as this helps to illustrate the nature of a diffusion process. Suppose that $x_n(t)$ $n = 0, 1, \ldots$ denotes the amount of liquid B in the successive tanks at time t.



In a time Δt sec,

$$r \frac{x_{k-1}}{G} \Delta t$$

gallons of B flow into tank k, while

$$r \frac{x_k}{G} \Delta t$$

gallons flow out. Then

$$\Delta x_k = \frac{r \Delta t}{G} (x_{k-1} - x_k)$$

gives the change of liquid B in the k th tank in Δt sec. Thus the rate of flow of B is

$$\frac{dx_k}{dt} = \frac{r}{G} (x_{k-1} - x_k)$$

and setting $x_{-1} = 0$ will give correctly $\frac{dx_0}{dt}$. The equation above is and element of a matric equation.

$$\begin{bmatrix} \frac{dx_0}{dt} \\ \frac{dx_1}{dt} \\ \cdots \end{bmatrix} = \begin{bmatrix} -1 & 0 & & \\ 1 & -1 & 0 & \\ 0 & 1 & -1 & \\ & & \cdots & \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \cdots \end{bmatrix}$$

$$\frac{d\mathcal{X}}{dt} = \frac{r}{G}\left\{\mathcal{S}^{-1} - \mathcal{I}\right\}\mathcal{X}$$

Where

$$\mathcal{S}^{-1} = \begin{bmatrix} 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

The solution of such an equation is

$$\mathcal{X}(t) = e^{\frac{r}{G}\left\{\mathcal{S}^{-1} - \mathcal{I}\right\}t}\,\mathcal{X}(0)$$

$$= e^{\frac{r}{G}\mathcal{S}^{-1}t}\, e^{-\frac{r}{G}\mathcal{I}t}\,\mathcal{X}(0)$$

$$= \left\{\mathcal{I} + \frac{r}{G}\mathcal{S}^{-1}t + \frac{1}{2!}\left(\frac{r}{G}\mathcal{S}^{-1}t\right)^2 + \cdots\right\} e^{-\frac{r}{G}t}\,\mathcal{I}\,\mathcal{X}(0)$$

or written out in full

$$\mathcal{X}(t) = e^{-\frac{r}{G}t} \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{r}{G}t & 1 & 0 & 0 \\ \frac{1}{2!}\left(\frac{r}{G}t\right)^2 & \frac{r}{G}t & 1 & 0 \\ & & \text{etc.} & \end{bmatrix} \begin{bmatrix} x_0(0) \\ x_1(0) \\ \cdots \end{bmatrix}$$

but

$$\mathcal{X}(0) = \begin{bmatrix} g \\ 0 \\ 0 \\ \cdots \end{bmatrix} \quad \therefore \quad \mathcal{X}(t) = e^{-\frac{r}{G}t} \begin{bmatrix} g \\ g\frac{r}{G}t \\ g\left(\frac{r}{g}\right)^2 t^2 \end{bmatrix}$$

or, generally,

$$\chi_n(t) = \frac{q}{n!}\left(\frac{r}{G}\right)^n t^n e^{-\frac{r}{G}t}$$

This represents a considerable simplification over the series solution in terms of iterated integrals

$$\mathcal{X}(t) = \left\{\mathbb{I} + \int_0^t \Phi(\sigma)\,d\sigma + \int_0^t\int_0^{\sigma_1} \Phi(\sigma_1)\,\Phi(\sigma_2)\,d\sigma_2\,d\sigma_1 + \cdots\right\}\mathcal{X}(0)$$

but this latter solution provides a detailed physical picture of the diffusion process. We have

$$\Phi = \frac{r}{G}\left\{\mathcal{S}^{-1} - \mathbb{I}\right\}$$

and the differential equation reads

$$d\mathcal{X} = \Phi(\sigma_k)\,\mathcal{X}(\sigma_{k-1})\,d\sigma_k$$

so that this is the expression for the amount of fluid flowing out of the tank in $d\sigma_k$ minutes at the time σ_k. By inspecing Φ we find that it is an operator consisting of two parts; $\mathcal{S}^{-1}$ shifts chemical into one tank from the preceding one, while $-\mathbb{I}$ discharges fluid from the tank in question.

However $\mathcal{X}$ has not only the interpretation of being the amount of liquid in a tank, but also the probability of finding liquid B in a unit volume of liquid. Thus $d\mathcal{X}$ is the probability for liquid B to flow out of the various tanks, written as a sum of the probability of flowing in from the previous tank or out the present tank. Thus in the general term of the series expansion,

$$\int_0^t\int_0^{\sigma_1}\cdots\int_0^{\sigma_{k-1}} \Phi(\sigma_1)\,\Phi(\sigma_2)\cdots\Phi(\sigma_k)\,d\sigma_k\cdots d\sigma_2\,d\sigma_1\,\mathcal{X}_0$$

$$= \int_0^t\int_0^{\sigma_1}\cdots\int_0^{\sigma_{k-1}} \Phi(\sigma_1)\,\Phi(\sigma_2)\cdots\left\{\Phi(\sigma_k)\,\mathcal{X}_0\,d\sigma_k\right\}\cdots d\sigma_2\,d\sigma_1$$

The term $\overline{\Phi}(\sigma_k)\chi_0 d\sigma_k$ is the fractional probability of the liquid flowing out from a given tank in $d\sigma_k$ minutes; the first integral is then the probability that it flows out in σ_{k-1} minutes, since the probabilities that it flows out in any time $d\sigma_k$ are independent. The next integration gives the probability that if it flows from the given tank and also from the next tank; and so on until the whole term gives the probability of flowing past k tanks, i.e., of going from one to the next, and from that to the next and to the next and so on. The reason for integrating from 0 to σ_{p-1} is the p^{th} integration is clear, for the liquid cannot flow into the $p+1^{st}$ tank until it has arrived in the p^{th}.

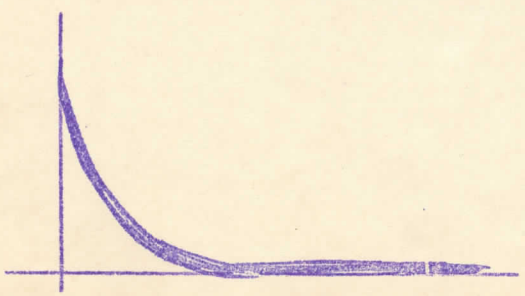
$\chi(t)$ gives the probability of finding liquid in the various tanks in terms of its probability of having come from one tank up the line, two tanks, ——, or having stayed where it started.

Since the order in which the shifts occur seem to be unimportant as long as they are all performed, the formula

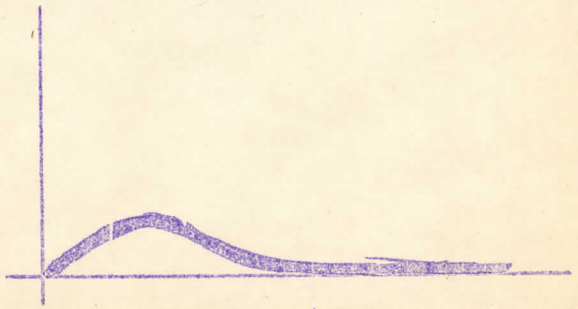
$$\chi_n(t) = \frac{q}{n!}\left(\frac{r}{c}\right)^n \left\{^n e^{-\frac{r}{c}t}\right.$$

is obtained at once; q is the probability of having liquid to start with; is the probability that the liquid will go to the next tank, hence $\left(\frac{r}{c}t\right)^n$ is the probability for having shifted down to the n^{th} tank, while $n!$ is the number of ways it could have been left in a tank enroute, while $e^{-\frac{r}{c}t}$ gives the probability of remaining in a tank once arrived there.

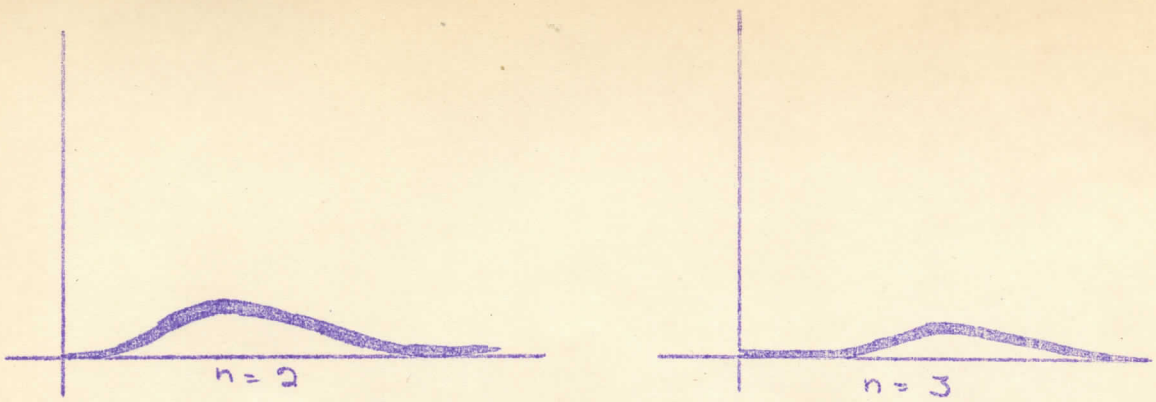
The amount of liquid as a function of time is plotted below.



$n = 0$



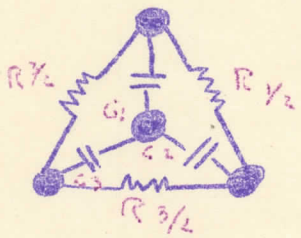
$n = 1$



The curve starts off as t^n, rerlecting the small chance of a particle soon arriving far down the line, builds up to a maximum, and falls off when the exponential dominates, reflecting the small chance of keeping the liquid long in a tank once accumulated there.

AN ELECTRICAL PROBLEM

The present problem introduces no new concepts, but affords a means of comparing the solution by eigenvectors and by series of the last two problems.



Suppose that three condensers, all having the same capacitance C, and three resistors, all having a common resistance R are connected according to the diagram above. The condensers are given certain initial charges, and problem is to determine their subsequent charges as a function of time.

Writing Kirchoff's equations, and letting i range through 1, 2, 3 cyclically, one has

$$R I_{i+\frac{1}{2}} + \frac{Q_{i+1}}{C} - \frac{Q_i}{C} = 0$$

$$I_{i-\frac{1}{2}} + I_i - I_{i+\frac{1}{2}} = 0$$

so that

$$R I_i - \frac{1}{C}\left\{Q_{i+1} - 2Q_i + Q_{i-1}\right\} \qquad i = 1,2,3$$

These equations can be written in matrix form as

$$\begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \end{bmatrix} = \frac{1}{RC}\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}\begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \end{bmatrix}$$

The solution, in terms of series, is

$$Q(t) = e^{\int_{t_0}^{t} \frac{1}{RC}\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} d\sigma} \; Q(0)$$

The matrix exponential is just

$$\frac{t-t_0}{RC}\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}$$

Writing $M = S' - 2\Pi + S^{-1}$

where

$$M = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \qquad S' = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Since these matrices all commute,

$$e^{M\frac{t-t_0}{RC}} = e^{S'\frac{t-t_0}{RC}} \; e^{S^{-1}\frac{t-t_0}{RC}} \; e^{\Pi\,-\frac{2(t-t_0)}{RC}}$$

obviously

$$e^{-2\Pi\frac{(t-t_0)}{RC}} = e^{-\frac{2(t-t_0)}{RC}\Pi}$$

But, since S' is orthogonal, $S^{-1} = S'$, and $e^{\frac{t}{RC}S'} = e^{\frac{t}{RC}S^{-1}}$, so that it will suffice to evaluate just one of these exponentials.

$$e^{S'\frac{t-t_0}{RC}} = \Pi + S'\frac{(t-t_0)}{RC} + \frac{1}{2!}\left(\frac{t-t_0}{RC}\right)^2 + \cdots$$

$$= \begin{bmatrix} 1 + \frac{1}{3!}\left(\frac{t-t_0}{RC}\right)^3 + \frac{1}{6!}\left(\frac{t-t_0}{RC}\right)^6 + \cdots & \frac{t-t_0}{RC} + \frac{1}{4!}\left(\frac{t-t_0}{RC}\right)^4 + \frac{1}{7!}\left(\frac{t-t_0}{RC}\right)^7 + \cdots & S_3 \\ \frac{1}{2!}\left(\frac{t-t_0}{RC}\right)^2 + \frac{1}{5!}\left(\frac{t-t_0}{RC}\right)^5 + \frac{1}{8!}\left(\frac{t-t_0}{RC}\right)^8 + \cdots & S_1 & S_2 \\ S_2 & S_3 & S_1 \end{bmatrix}$$

These series $S_1 = 1 + \left(\frac{t-t_0}{RC}\right)^3 \frac{1}{3!} + \cdots$; $S_2 = \left(\frac{t-t_0}{RC}\right) + \frac{1}{4!}\left(\frac{t-t_0}{RC}\right)^4 + \cdots$; $S_3 = \frac{1}{2!}\left(\frac{t-t_0}{RC}\right)^2 + \cdots$

can be evaluated from noticing that they all satisfy the differential equation

$$\frac{d^3 S}{dt^3} = S$$

with the boundary conditions $S_1(0)=1$, $S_1'(0)=0$, $S_1''(0)=0$; $S_2(0)=0$, $S_2'(0)=1$, $S_2''(0)=0$; $S_3(0)=0$, $S_3'(0)=0$, $S_3''(0)=1$

and are readily seen to be

$$S_1 = \frac{1}{3}\left(e^{\frac{t-t_0}{RC}} + e^{w\frac{t-t_0}{RC}} + e^{w^2\frac{t-t_0}{RC}}\right) \qquad w = e^{\frac{2\pi i}{3}}$$

$$S_2 = S_1' = \frac{1}{3}\left(e^{\frac{t-t_0}{RC}} + we^{w\frac{t-t_0}{RC}} + w^2 e^{w^2\frac{t-t_0}{RC}}\right)$$

$$S_3 = S_1'' = \frac{1}{3}\left(e^{\frac{t-t_0}{RC}} + w^2 e^{w\frac{t-t_0}{RC}} + we^{w^2\frac{t-t_0}{RC}}\right)$$

and the matrix $e^{S'\frac{t-t_0}{RC}}$ is

$$e^{S'\frac{t-t_0}{RC}} = \begin{bmatrix} S_1 & S_2 & S_3 \\ S_3 & S_1 & S_2 \\ S_2 & S_3 & S_1 \end{bmatrix}$$

and the solution to the problem is

$$\begin{bmatrix} Q_1(t) \\ Q_2(t) \\ Q_3(t) \end{bmatrix} = \begin{bmatrix} S_1 & S_2 & S_3 \\ S_3 & S_1 & S_2 \\ S_2 & S_3 & S_1 \end{bmatrix} \begin{bmatrix} S_1 & S_3 & S_2 \\ S_2 & S_1 & S_3 \\ S_3 & S_2 & S_1 \end{bmatrix} \begin{bmatrix} e^{\frac{-2(t-t_0)}{RC}} & 0 & 0 \\ 0 & e^{\frac{-2(t-t_0)}{RC}} & 0 \\ 0 & 0 & e^{\frac{-2(t-t_0)}{RC}} \end{bmatrix} \begin{bmatrix} Q_1(t_0) \\ Q_2(t_0) \\ Q_3(t_0) \end{bmatrix}$$

$$= \frac{1}{3} \begin{bmatrix} 1+2e^{-\frac{3}{RC}(t-t_0)} & 1-e^{-\frac{3}{RC}(t-t_0)} & 1-e^{-\frac{3}{RC}(t-t_0)} \\ 1-e^{-\frac{3}{RC}(t-t_0)} & 1+2e^{-\frac{3}{RC}(t-t_0)} & 1-e^{-\frac{3}{RC}(t-t_0)} \\ 1-e^{-\frac{3}{RC}(t-t_0)} & 1-e^{-\frac{3}{RC}(t-t_0)} & 1+2e^{-\frac{3}{RC}(t-t_0)} \end{bmatrix} \begin{bmatrix} Q_1(t_0) \\ Q_2(t_0) \\ Q_3(t_0) \end{bmatrix}$$

so that

$$Q_i(t) = \tfrac{1}{3}\{Q_1(t_0)+Q_2(t_0)+Q_3(t_0)\} - \tfrac{1}{3} e^{-\frac{3}{RC}(t-t_0)} \{Q_{i-1}(t_0) - 2Q_i(t_0) + Q_{i+1}(t_0)\}$$

Setting $\frac{1}{3}\{Q_1(t_0)+Q_2(t_0)+Q_3(t_0)\} = \langle Q(t_0)\rangle$ the average charge at time t_0 is

$$Q_i(t) = \langle Q(t_0)\rangle - \tfrac{1}{3} e^{-\frac{3}{RC}(t-t_0)} \{Q_{i-1}(t_0) - 2Q_i(t_0) + Q_{i+1}(t_0)\}$$

thus the matrix, $e^{\int_{t_0}^{t} M \frac{1}{RC} d\sigma}$, is an operator which expresses the charge at the time t in terms of the charge at the time t. It was of course obtained from the product

$$Q(t) = \left\{\mathbb{I} - \tfrac{\Delta t}{RC} M\right\}\left\{\mathbb{I} - \tfrac{\Delta t}{RC} M\right\} \cdots \left\{\mathbb{I} - \tfrac{\Delta t}{RC} M\right\} Q(t_0)$$

which carries the solution forward an infinitesimal amount at a time. However it is factorable into three operators

$$e^{-\frac{2t}{RC}\mathbb{I}} \qquad e^{S^{-1}\frac{t}{RC}} \qquad e^{S'\frac{t}{RC}}$$

The first of these produces a discharge of the condensers through an effective resistance $R/2$, which would be just the resistance offered by the two resistors if they were grounded to the other plate of the condensers to which they connect.

The other two operators carry the charge to the left and to the right cyclically, and they can be interpreted in terms of a diffusion process for the charge if they are expaned in Taylor's series, such as was done in the last problem.

Although the problem is now solved, it is interesting to solve it by another method, for the purpose of comparing the two solutions. Thus we obtain a solution by finding the eigenvectors of the matrix

$$\frac{1}{RC}\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix}$$

for when this matrix is diagonalized the solution in terms of normal coordinates, discussed in a previous problem, is possible. The task of finding the eigenvector is simplified by noticing that this matrix can be written as

$$\frac{1}{RC}\left\{S' - 2\mathbb{I} + S'^{-1}\right\} = \frac{1}{RC} M$$

These matrices having been previously defined. Since they all commute, the eigenvectors of S' are those of S'^{-1}. But the eigenvectors and eigenvalues are readily seen to be

$$|1\rangle = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix} \qquad |2\rangle = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\\omega\\\omega^2\end{bmatrix} \qquad |3\rangle = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\\omega^2\\\omega\end{bmatrix}$$

$$\lambda_1 = 1 \qquad \lambda_2 = \omega \qquad \lambda_3 = \omega^2$$

where $\omega = e^{\frac{2\pi i}{3}}$. These are the eigenvectors of M; the eigenvalues of M are the sum of the eigenvalues of S', $-2\mathbb{I}$, and S'^{-1}. Thus they are

$$\lambda_1^{(M)} : 1 - 2 + 1 = 0$$

$$\lambda_2^{(M)} \quad \omega - 2 + \omega = -3$$

$$\lambda_3^{(M)} \cdot \omega^2 - 2 + \omega^2 = -3$$

which gives a double eigenvalue. Since this is true it is possible to take linear combinations of $|2\rangle$ and $|3\rangle$ for the purpose of obtaining real eigenvectors. Such a choice is

$$|1\rangle = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix} \qquad |2\rangle = \frac{1}{\sqrt{6}}\begin{bmatrix}1\\-2\\1\end{bmatrix} \qquad |3\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\-1\end{bmatrix}$$

$$\lambda_1 = 0 \qquad \lambda_2 = 3 \qquad \lambda_3 = 3$$

Using Sylvester's theorem to write the solution of the differential equation:

$$Q(t) = \left\{ e^{0\,\frac{t-t_0}{RC}} |1\rangle\langle 1| + e^{-\frac{3t-t_0}{RC}} |2\rangle\langle 2| + e^{-3\,\frac{t-t_0}{RC}} |3\rangle\langle 3| \right\} Q(t_0)$$

The normal coordinates then have this significance; charges which have been distributed on the condensers as indicated by the eigenvectors will cause their condensers all to discharge at the same rate, for the eigenvalues of $\frac{1}{RC} M$ are the reciprocals time constants of the normal modes. Thus corresponds to a situation in which all the condensers have been equally charged; no charge flows, corresponding to the time constant ∞.

The second normal mode results when a unit charge is placed on the first and third condensers, and -2 charges on the second condenser. These then discharge at the same rate, no charge flowing between 1 and 3, and the relation between 2 and 3 and 1 and 3 being symmetrical.

Finally the third normal mode represents a unit charge on the first condenser, no charge on the second, -1 charge on the third. The second condenser remains with charge 0, while one and three discharge into one another. Clearly this relation is symmetrical giving a uniform rate of discharge for all the condensers. These latter two charge distributions could have been rotated 120° or 240° and produced the same results, since all the condensers are equivalent. This is the reason for the double eigenvalues $-3/RC$, for then any linear combination of these latter normal modes is again a normal mode. For instance, $\frac{1}{3}\{|2\rangle + |3\rangle\}$ gives a normal mode of the third type, but with zero charge on the third rather than the second condenser.

The solution

$$Q_i(t) = \langle Q(t_0)\rangle - \tfrac{1}{3} e^{-\frac{3}{RC}(t-t_0)} \left\{ Q_{i-1}(t_0) - 2\,Q_i(t_0) + Q_{i+1}(t_0) \right\}$$

then has two parts: $\langle Q(t_0)\rangle$ is the charge represented by the eigenvectors $|1\rangle$, and remains on the condensers throughout the experiment. That this charge should be the average charge is fairly obvious, since there is no flow of current when they all have the same charge, so that after a long time one should expect the fluctuations from average to have been evened out. This is done by the term $\frac{1}{3}e^{-\frac{3}{RC}(t-t_0)}\{Q_{i-1}(t_0) - 2Q_i(t) + Q_{i+1}(t_0)\}$ which is given by the sum of the last two normal modes.

THE SMALL VIBRATIONS OF THE OZONE MOLECULE

The ozone molecule problem illustrates the use which can be made of the fact that if two matrices commute they must have a common set of eigenvectors. The elastic constant for the ozone molecule is fairly complicated; however, one can find some relatively simple matrices which commute with it. My knowing their eigenvectors, the problem of finding those of the elastic constant is simplified, for the matrix which diagonalizes the commuting matrix will partially diagonalize the elastic constant.

Such commuting matrices can be discovered from the symmetry of the molecule, for if it has a certain symmetry, such that when two particles are interchanged the matrix of the elastic constant is unchanged, this permutation matrix will then commute with the elastic constant, for one will have

$$P^{-1}KP = K$$

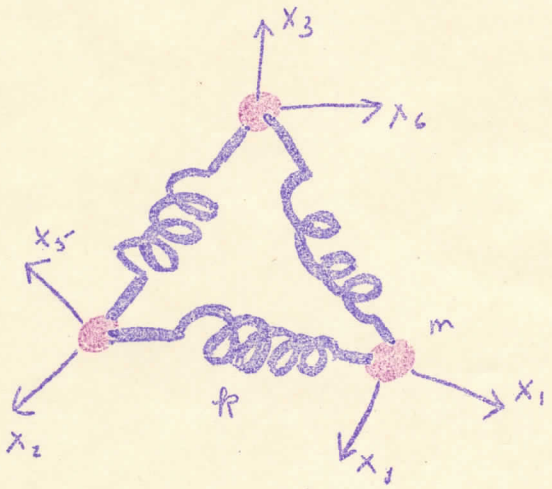
which can be rearranged to read

$$KP = PK$$

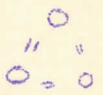
which is the desired result.

In this problem but one symmetry is considered, but for more complicated problems one might be forced to take advantage of more symmetries. This requires only the repeated application of the technique described here however.

Only the plane motion of the molecule is considered, and the particle coordinates are chosen according to the diagram below. The two coordinates of a particular particle are chosen to be orthogonal so as to make the mass matrix diagonal; otherwise accelerations in the two directions would not be independent. Springs have been inserted in place of valence bonds under the assumption that the force required for elongation along a bond is much greater than that required to change the angle between bonds.



O_3 :



The equations of motion can be obtained by calculating the force in a particular direction when the particles are displaced in tern along the successive coordinate axes, and adding the results. We use the abbreviations

$s = \sin 30°$ $\quad c = \cos 30°$ $\quad k$ = elastic constant of spring

$$f_1 = -2kc^2x_1 - kc^2x_2 - kc^2x_3 + 0x_4 - ksc\,x_5 + ksc\,x_6$$

$$f_2 = -kc^2x_1 - 2kc^2x_2 - kc^2x_3 + ksc\,x_4 + 0x_5 - ksc\,x_6$$

$$f_3 = -kc^2x_1 - kc^2x_2 - 2kc^2x_3 - ksc\,x_4 + ksc\,x_5 + 0x_6$$

$$f_4 = 0x_1 + ksc\,x_2 - ksc\,x_3 - 2ks^2x_4 + ks^2x_5 + ks^2x_6$$

$$f_5 = -ksc\,x_1 + 0x_2 + ksc\,x_3 + ks^2x_4 - 2ks^2x_5 + ks^2x_6$$

$$f_6 = ksc\,x_1 - ksc\,x_2 + 0x_3 + ks^2x_4 + ks^2x_5 - 2ks^2x_6$$

in matrix form, with m as the mass of an atom, and inertia forces substituted for restoring forces:

$$m \begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \\ \ddot{x}_4 \\ \ddot{x}_5 \\ \ddot{x}_6 \end{bmatrix} = -k \begin{bmatrix} 2c^2 & c^2 & c^2 & 0 & sc & -sc \\ c^2 & 2c^2 & c^2 & -sc & 0 & sc \\ c^2 & c^2 & 2c^2 & sc & -sc & 0 \\ 0 & -sc & sc & 2s^2 & -s^2 & -s^2 \\ sc & 0 & -sc & -s^2 & 2s^2 & -s^2 \\ -sc & sc & 0 & -s^2 & -s^2 & 2s^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}$$

or $$\ddot{\mathfrak{X}} = -\frac{k}{m} K \mathfrak{X}$$

The matrix K is to be diagonalized, to discover the normal modes.

The permutation of the particles corresponding to a 120° rotation of the molecule about its center of gravity is a physical symmetry of the molecule. The coordinate exchange resulting from this permutation is

$$x_1 \to x_3 \qquad x_4 \to x_6$$
$$x_2 \to x_1 \qquad x_5 \to x_4$$
$$x_3 \to x_2 \qquad x_6 \to x_5$$

whence P, the orthogonal permutator in question, is easily verified to commute with K and is

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} S' & 0 \\ 0 & S' \end{bmatrix}$$

where S' is the matrix which performs 120° rotation about the axes (1, 1, 1) in three dimensions. Its properties have already been studied, and its eigenvectors found to be:

$|1\rangle$ $\frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}$ $\lambda_1 = 1$ $\quad$ $|2\rangle$ $\frac{1}{\sqrt{3}}\begin{bmatrix}1\\\omega\\\omega^2\end{bmatrix}$ $\lambda_2 = \omega$ $\quad$ $|3\rangle$ $\frac{1}{\sqrt{3}}\begin{bmatrix}1\\\omega^2\\\omega\end{bmatrix}$ $\lambda_3 = \omega^2$

if $\mathcal{S}'\omega = \lambda\omega$, , also $\mathcal{S}'\mathcal{O} = \lambda\mathcal{O}$ so that the eigenequations for P is satisfied by:

$$\begin{bmatrix}\mathcal{S}' & \mathcal{O}\\ \mathcal{O} & \mathcal{S}'\end{bmatrix}\begin{bmatrix}\omega\\ \mathcal{O}\end{bmatrix} = \lambda\begin{bmatrix}\omega\\ \mathcal{O}\end{bmatrix}\ ;\ \begin{bmatrix}\mathcal{S}' & \mathcal{O}\\ \mathcal{O} & \mathcal{S}'\end{bmatrix}\begin{bmatrix}\mathcal{O}\\ \omega\end{bmatrix} = \lambda\begin{bmatrix}\mathcal{O}\\ \omega\end{bmatrix}$$

which gives, as double degenerate eigenvectors for P :

$\sqrt{3}\,|1\rangle$:

$$\underbrace{\begin{bmatrix}1\\1\\0\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\0\\1\\1\end{bmatrix}}_{1} \qquad \underbrace{\begin{bmatrix}1\\\omega\\\omega^2\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\0\\1\\\omega^2\end{bmatrix}}_{\omega} \qquad \underbrace{\begin{bmatrix}1\\\omega^2\\\omega\\0\\0\end{bmatrix}\begin{bmatrix}0\\0\\0\\1\\\omega\end{bmatrix}}_{\omega^2}$$

λ_i :

with each eigenvalue double. This means that every vector in the plane spanned by two eigenvectors belonging to one of these eigenvalues also belongs to it. This means that the eigenvectors of K cannot be completely determined, but rather that they are restricted to lie in three absolutely perpendicular planes. However, the matrix which diagonalizes P will partially diagonalize K, leaving 2×2 submatrices which can easily be diagonalized separately.

A matrix U which will diagonalize P is then:

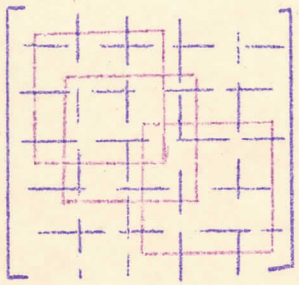
$$U = \begin{bmatrix}
\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & 0 & 0 & 0\\
\frac{1}{\sqrt{3}} & \frac{\omega}{\sqrt{3}} & \frac{\omega^2}{\sqrt{3}} & 0 & 0 & 0\\
\frac{1}{\sqrt{3}} & \frac{\omega^2}{\sqrt{3}} & \frac{\omega}{\sqrt{3}} & 0 & 0 & 0\\
0 & 0 & 0 & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}}\\
0 & 0 & 0 & \frac{1}{\sqrt{3}} & \frac{\omega}{\sqrt{3}} & \frac{\omega^2}{\sqrt{3}}\\
0 & 0 & 0 & \frac{1}{\sqrt{3}} & \frac{\omega^2}{\sqrt{3}} & \frac{\omega}{\sqrt{3}}
\end{bmatrix} = \begin{bmatrix}\mathcal{U} & \mathcal{O}\\ \mathcal{O} & \mathcal{U}\end{bmatrix}$$

where Ω is suitably defined.

Writing K in terms of submatrices also as $K = \begin{bmatrix} \alpha & \beta \\ -\beta & \gamma \end{bmatrix}$

$$U^{-1}KU = \begin{bmatrix} \Omega^{-1}\alpha\Omega & \Omega^{-1}\beta\Omega \\ -\Omega^{-1}\beta\Omega & \Omega^{-1}\gamma\Omega \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3/4 & 0 & 0 & 3\sqrt{1}/4 & 0 \\ 0 & 0 & 3/4 & 0 & 0 & -3\sqrt{1}/4 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -3\sqrt{1}/4 & 0 & 0 & 3/4 & 0 \\ 0 & 0 & 3\sqrt{1}/4 & 0 & 0 & 3/4 \end{bmatrix}$$

This matrix does not present quite the expected block form, the reason for this being that in the matrix U the normalized eigenvectors belonging to the same eigenvalue were not arranged side by side. This was done so that KU could be partitioned into convenient submatrices for multiplying the product above out. However, the submatrices may be recognized for the following diagram, where elements of a submatrix are connected by heavy lines:



A matrix L

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

may be used to arrange the vectors of U in a more desirable order, and thus to place $U^{-1}KU$ in the form $L^{-1}U^{-1}KUL$ which is

$$\begin{bmatrix} \Lambda V & \mathcal{O} & \mathcal{O} \\ \mathcal{O} & \mathit{II} & \mathcal{O} \\ \mathcal{O} & \mathcal{O} & \mathit{II} \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3/4 & -3i/4 & 0 & 0 \\ 0 & 0 & 3i/4 & 3/4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3/4 & 3i/4 \\ 0 & 0 & 0 & 0 & -3i/4 & 3/4 \end{bmatrix}$$

The eigenvalues and eigenvectors of these 2x2 submatrices are

Matrix	Eigenvalue	Eigenvector
ΛV	3	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$
	0	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$
II	$3/2$	$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \end{bmatrix}$
	0	$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix}$
$\Lambda\Lambda$	0	$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix}$
		$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \end{bmatrix}$

and matrices Q, Q', Q'' which diagonalize each of these blocks separately, are given, in their respective orders, by

ΛV $\quad Q = \begin{bmatrix} \mathit{II} & \mathcal{O} & \mathcal{O} \\ \mathcal{O} & \mathit{II} & \mathcal{O} \\ \mathcal{O} & \mathcal{O} & \mathit{II} \end{bmatrix} \quad -\mathit{II}$

II $\quad Q' = \begin{bmatrix} \mathit{II} & \mathcal{O} & \mathcal{O} \\ \mathcal{O} & \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ i/\sqrt{2} & -i/\sqrt{2} \end{bmatrix} & \mathcal{O} \\ \mathcal{O} & \mathcal{O} & \mathit{II} \end{bmatrix}$

$\Lambda\Lambda$ $\quad Q'' = \begin{bmatrix} \mathit{II} & \mathcal{O} & \mathcal{O} \\ \mathcal{O} & \mathit{II} & \mathcal{O} \\ \mathcal{O} & \mathcal{O} & \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -i/\sqrt{2} & i/\sqrt{2} \end{bmatrix} \end{bmatrix}$

so that finally $Q^{-1}Q'^{-1}Q''^{-1}\Lambda^{-1}U^{-1}KUL^{-}Q'Q'Q$ is diagonal.

It is:

$$\begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

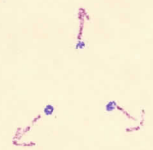
and K' eigenvalues and eigenvectors are the columns of $(uLQ''Q'Q)$.

$$\underbrace{\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \\ 0 \\ 0 \\ 0 \end{bmatrix}}_{3} \quad \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} \\ \omega^2/\sqrt{6} \\ \omega/\sqrt{6} \\ -i/\sqrt{6} \\ -i\omega^2/\sqrt{6} \\ -i\omega/\sqrt{6} \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} \\ -\omega/\sqrt{6} \\ \omega^2/\sqrt{6} \\ i/\sqrt{6} \\ i\omega/\sqrt{6} \\ i\omega^2/\sqrt{6} \end{bmatrix}}_{0} \quad \underbrace{\begin{bmatrix} 1/\sqrt{6} \\ \omega^2/\sqrt{6} \\ -\omega/\sqrt{6} \\ i/\sqrt{6} \\ i\omega^2/\sqrt{6} \\ i\omega/\sqrt{6} \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} \\ \omega/\sqrt{6} \\ \omega^2/\sqrt{6} \\ -i/\sqrt{6} \\ -i\omega/\sqrt{6} \\ -i\omega^2/\sqrt{6} \end{bmatrix}}_{3/2}$$

It is of interest to note that Q is the unit matrix, and that the upper left hand block of $L^{-1}L^{-1}KUL$ which it would have diagonalized, was already diagonal, even though it had two distinct eigenvalues and hence was not a multiple of the unit matrix. This was purely accidental, and was due to a fortuitous choice of eigenvectors for U. Thus Q could have been omitted from the discussion and advantage taken of this coincidence. Furthermore, Q, Q^1, Q^{11} could have been combined into one matrix.

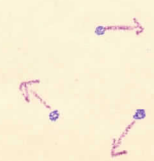
It is now possible to sketch the normal mode patterns. Taking the proper linear combinations to obtain real eigenvectors, there results

$|1\rangle$: $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ frequency: $\sqrt{\frac{3k}{m}}$



uniform dilution

$|2\rangle$ $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ " 0



rotation about c, g

$|3\rangle$ $\begin{bmatrix} 1 \\ -s \\ -s \\ 0 \\ c \\ c \end{bmatrix}$ " 0 translation

$|4\rangle$ $\begin{bmatrix} 0 \\ c \\ -c \\ -s \\ -s \\ -s \end{bmatrix}$ " 0 translation

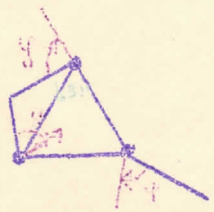
$|3\rangle \cos\varphi - |4\rangle \sin\varphi$ gives translation in any desired direction by proper choice of φ .

$|5\rangle$ $\begin{bmatrix} 1 \\ -s \\ -s \\ 0 \\ c \\ -c \end{bmatrix}$ " $\sqrt{\frac{3K}{2m}}$

$|6\rangle$ $\begin{bmatrix} 0 \\ -c \\ c \\ 1 \\ -s \\ -s \end{bmatrix}$ " $\sqrt{\frac{3K}{2m}}$

By taking the combination $|5\rangle \cos\varphi + |6\rangle \sin\varphi$

there results, with frequency $\sqrt{\frac{3K}{2m}}$;



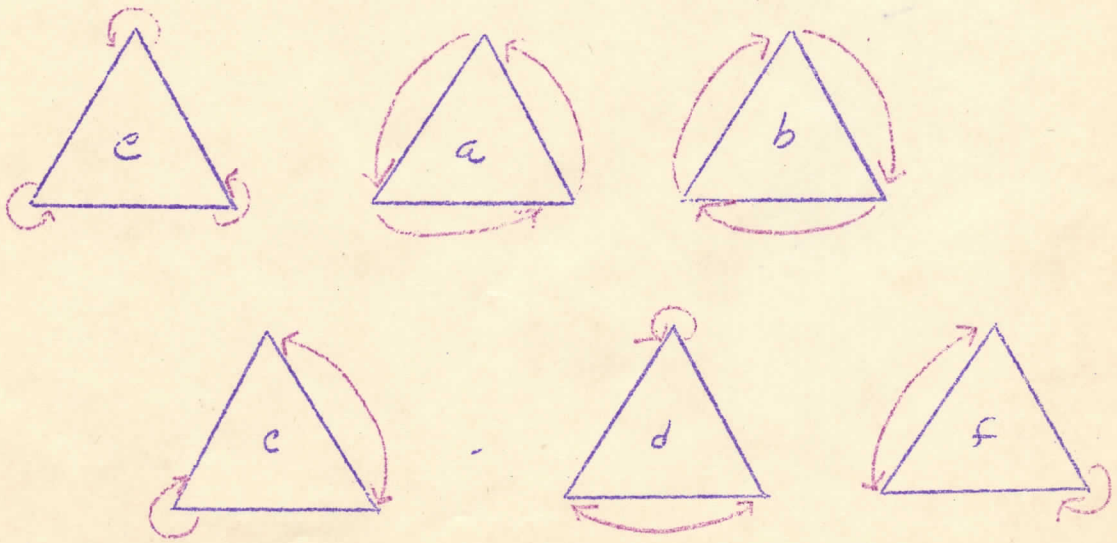
Many modes which should be expected are recognized, namely the two translational modes, with zero frequency, corresponding to the fact that the molecule can move in any direction at a constant velocity and experience no restoring force. The rotational mode was also to be expected, however the motion illustrated will eventually stretch the molecule out and produce a restoring force, despite the zero restoring force indicated by the zero eigenvalue. The solution to this paradox lies in the fact that one discusses "small" motions along the arc of a circle may be replaced by motion along a tangent; as is the case here. The elastic constant $-\frac{15}{m}K$ as originally written is only correct to first orders in the displacements.

The uniform dilation obviously gives restoring forces in the direction opposite to the displacements, and proportional to them, which is a property of motion according to a normal mode. The remaining two modes are degenerates and they are not obvious.

The motion of the molecule from arbitrary initial conditions may be synthesized from these modes by methods already discussed.

It will be recalled that the submatrix containing the rotation and dilation modes was already diagonalized by the matrix U, hence they should be invariant under 120° rotations, which inspection shows to actually be the case.

Among other questions which one might ask about this problem is the following: It was noticed that there were two degenerate modes, one with eigenvalue $3k/2m$ which was doubly degenerate, and one with eigenvalue 0 which was triply degenerate. The question is: "Are any of these degeneracies imposed by the symmetry of the molecule?" Reflection shows that this is the case, but one may make use of group theory to see precisely how and to what extent. The ozone molecule is symmetric under the permutation group on three letters. Its symmetries are $G = \{e, a, b, c, d, f\}$



With the multiplication table

	e	a	b	c	d	f
e	e	a	b	c	d	f
a	a	b	e	d	f	c
b	b	ee	a	f	c	d
c	c	f	d	e	b	a
d	d	c	f	a	e	b
f	f	d	e	b	a	ee

This group has three classes

$$\{e\} \ , \ \{a, \ b\} \ , \ \{c, d, f\}$$

and one non-trivial normal divisor

$$\{e, a\ b\}$$

The group is of order 6, and since it has three classes there are three irreducible representations, whose dimensionalities satisfy

$$l_1^2 , l_2^2 + l_3^2 = 6 \ \therefore \ l_1 = 2, \ l_2 = l_3 = 1$$

The irreducible representations are easily discovered. The representation in which each element is 1 is irreducible and of the dimension 1. Call it

α:

$e \to 1$ $\quad$ $c \to 1$

$a \to 1$ $\quad$ $d \to 1$ $\quad$ $\mathcal{G}/\mathcal{G}$

$b \to 1$ $\quad$ $f \to 1$

The representation β

β:

$e \to 1$ $\quad$ $c \to -1$

$a \to 1$ $\quad$ $d \to -1$ $\quad$ $\mathcal{G}/\{e, a, b\}$

$b \to 1$ $\quad$ $f \to -1$

is also irreducible, and is the factor group of $\mathcal{G}$ by the non-trivial normal divisor, and separates the rotational symmetries from the reflective symmetries.

The third representation must be holomorphic to $\mathcal{G}/\{e\}$, and is two dimensional. The plane 120° rotation and reflections are elements of this representation, γ'.

$$\gamma': \quad e \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad a \to \begin{bmatrix} \cos 120^\circ & \sin 120^\circ \\ -\sin 120^\circ & \cos 120^\circ \end{bmatrix}$$

$$b \to \begin{bmatrix} \cos 120^\circ & -\sin 120^\circ \\ \sin 120^\circ & \cos 120^\circ \end{bmatrix}$$

$$c \to \begin{bmatrix} -\cos 120^\circ & -\sin 120^\circ \\ -\sin 120^\circ & \cos 120^\circ \end{bmatrix} \qquad d \to \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$f \to \begin{bmatrix} -\cos 120^\circ & \sin 120^\circ \\ \sin 120^\circ & \cos 120^\circ \end{bmatrix}$$

To find the characters, we write a table

element / representation	e	a	b	c	d	f
	1	1	1	1	1	1
	1	1	1	-1	-1	-1
	2	-1	-1	0	0	0

Traces of Elements of Representations

so that the rows of the table are the group characters. Note that the orthogonality condition for group characters hold.

It is now necessary to determine which representation of $\mathcal{G}$ is to be used to commute with K. It must be six-dimensional, but reducible to two equal three dimensional representations. If the three-dimensional representation is

$$S:\quad e\to\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}=E \qquad a\to\begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}=A \qquad b\to\begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix}=B$$

$$c\to\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}=C \qquad d\to\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}=D \qquad f\to\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}=F$$

The desired representation will be

$$\mathcal{E}:\ e\to\begin{bmatrix}E&0\\0&E\end{bmatrix}\quad a\to\begin{bmatrix}A&0\\0&A\end{bmatrix}\quad b\to\begin{bmatrix}B&0\\0&B\end{bmatrix}\quad c\to\begin{bmatrix}C&0\\0&-C\end{bmatrix}\quad D\to\begin{bmatrix}D&0\\0&-D\end{bmatrix}\quad f\to\begin{bmatrix}F&0\\0&-F\end{bmatrix}$$

The character of this representation will then be

T	e	a	b	c	d	f
$\chi(T)$	6	0	0	0	0	0

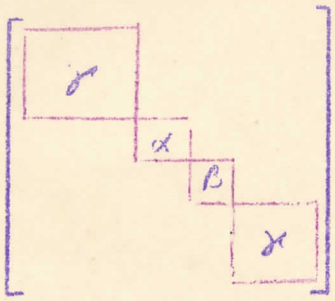
To find out how often the irreducible representations α, β, γ are contained in the reducible representation $\mathcal{E}$, one applies the formula

$$n(K)=\frac{1}{g}\sum_{G\in\mathfrak{G}}\chi(a)\,\chi^{(K)*}(a)$$

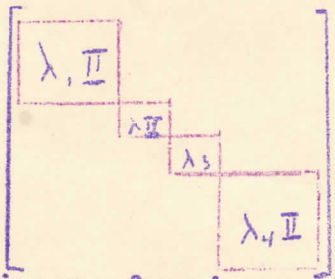
where g is the number of elements in $\mathfrak{G}$, $n(K)$ is the number of times the K^{th} irreducible representation occurs

$$n(\alpha)=\tfrac{1}{6}\{6\cdot1+0\cdot1+0\cdot1+0\cdot1+0\cdot1+0\cdot1\}$$
$$=1$$
$$n(\beta)=\tfrac{1}{6}\{6\cdot1+0\cdot1+0\cdot1+0\cdot1+0\cdot1+0\cdot1\}$$
$$=1$$
$$n(\gamma)=\tfrac{1}{6}\{6\cdot2+0\cdot1+0\cdot1+0\cdot1+0\cdot1+0\cdot1\}$$
$$=2$$

This means that the representation $\mathcal{E}$ can be brought to the partially diagonal form.



and $\mathcal{K}$, since it commutes with each matrix of must, when diagonalized, also be of the form



Thus two double degeneracies are forced upon $\mathcal{K}$. One is clearly the double degeneracy with eigenvalue ω_0. The other is the pair of translation modes with eigenvalue zero. That the rotational mode had zero eigenvalue also was purely accidental, and was not forced by the symmetry. That this is true may be seen by binding the center of gravity to the origin by an isotropic linear force. This will change the frequency of the translational motion without changing the zero frequency for rotation, although the symmetry of the problem has been preserved.

This discussion also shows that the eigenvectors are free to be chosen in specific subspaces, but to obtain much more detailed information the method originally propounded in the problem must be resorted to.

THE ASSYMETRICALLY COUPLED OSCILLATOR

Sometimes the solution to a particular problem is difficult because of difficulties arising in the finding of eigenvectors for the matrices which may arise in its solution. However, if the matrix differs but slightly from

one whose eigenvectors are already known, there is an approximation method available. Suppose that it is desired to find the eigenvalues and eigenvectors of the matrix

$$H + \varepsilon H' \qquad 1)$$

and that the eigenvalues and eigenvectors of H are already known. If one can find a matrix M such that

$$[M, H] = H'$$

he will make a small error in adding to 1) enough terms to complete the series

$$H + \varepsilon [M, H] + \frac{\varepsilon^2}{2!} [M, [M, H]] + \frac{\varepsilon^3}{3!} [M, [M, [M, H]]] + \cdots$$

which may be recognized as

$$e^{\varepsilon M} H e^{-\varepsilon M}$$

Since now $H + \varepsilon H' = e^{\varepsilon M} H e^{-\varepsilon M}$

$$H = e^{-\varepsilon M} \{ H + \varepsilon H' \} e^{\varepsilon M}$$

Now, if U diagonalizes H, $U e^{\varepsilon M}$ will diagonalize $H + \varepsilon H'$ from above.

Since the matrix whose columns are eigenvectors is the one which diagonalizes matrices, it is apparent that the new eigenvectors are the columns of

$$U e^{\varepsilon M}$$

But, in the eigensystem of H, these are just the columns of $e^{\varepsilon M}$. To first order in ε, $e^{\varepsilon M}$ is $\mathbb{I} + \varepsilon M$ so that the corrections to the eigenvectors are in fact just the columns of M.

The new eigenvalues are given by

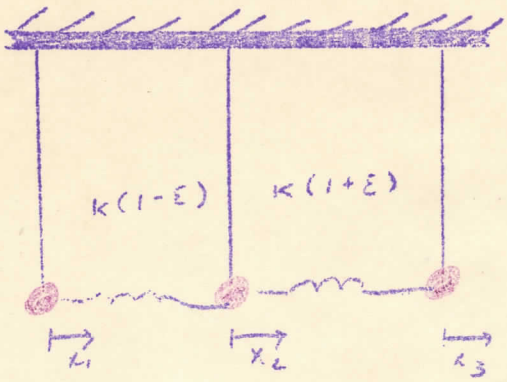
$$\langle i + \varepsilon \delta i | H + \varepsilon H' | i + \varepsilon \delta i \rangle$$

$$= \lambda_i \langle i | M | i \rangle + \varepsilon \{ \langle \delta i | M | i \rangle + \langle i | M | \delta i \rangle + \varepsilon \langle i | H' | i \rangle + O(\varepsilon^2) \}$$

which gives, in view of the extremal properties of the eigenvectors

$$\lambda'_i = \lambda_i + \varepsilon \langle i | H' | i \rangle$$

For an example consider the three coupled pendula considered in a previous problem, where the alteration of using a slightly stronger spring on the right and a weaker one on the left has been made.



The quation of motion was

$$\begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{bmatrix} = - \begin{bmatrix} k/m + g/\ell & -k/m & 0 \\ -k/m & 2k/m + g/\ell & -k/m \\ 0 & -k/m & k/m + g/\ell \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\ddot{\vec{x}} = -H\vec{x}$$

while the perturbation to the problem consists in the matrix

$$\varepsilon H' = \begin{bmatrix} \varepsilon k/m & -\varepsilon k/m & 0 \\ -\varepsilon k/m & 0 & \varepsilon k/m \\ 0 & \varepsilon k/m & -\varepsilon k/m \end{bmatrix}$$

which represents the force due to the modified springs. The solutions to the original equation are already known; the eigenvalues and eigenvectors of the elastic constant were $\lambda_1 = g/\ell$ $\lambda_2 = g/\ell + k/m$ $\lambda_3 = g/\ell + 3k/m$

$$|1\rangle = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad |2\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \qquad |3\rangle = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

In the eigensystem of H,

$$H = \begin{bmatrix} g/\ell & 0 & 0 \\ 0 & g/\ell + \hbar/m & 0 \\ 0 & 0 & g/\ell + 2\hbar/m \end{bmatrix} \qquad \varepsilon H' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{\varepsilon K\sqrt{3}}{m} \\ 0 & \frac{\varepsilon K\sqrt{3}}{m} & 0 \end{bmatrix}$$

The matrix M which solves the commutator equation

$$[M, H] = H'$$

is given by

$$\langle i|M|j\rangle = \frac{\langle i|H'|j\rangle}{\lambda_j - \lambda_i}$$

and is

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{\varepsilon\sqrt{3}}{2} \\ 0 & -\frac{\varepsilon\sqrt{3}}{2} & 0 \end{bmatrix}$$

so that the new, approximate eigenvectors, in the eigensystem are

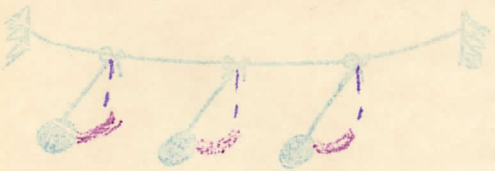
$$|1\rangle' : \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \qquad |2\rangle' = \begin{bmatrix} 0 \\ 1 \\ -\frac{\varepsilon\sqrt{3}}{2} \end{bmatrix} \qquad |3\rangle' = \begin{bmatrix} 0 \\ \frac{\varepsilon\sqrt{3}}{2} \\ 1 \end{bmatrix}$$

Note that although they contain terms of order ε, they are still normalized, to first order in ε.

These approximate eigenvectors were given in terms of the eigenvectors as a basis; transforming them back to the laboratory system, they are

$$|1\rangle' = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad |2\rangle' = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 - \varepsilon/2 \\ \varepsilon \\ -1 - \varepsilon/2 \end{bmatrix} \qquad |3\rangle' = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 + 3\varepsilon/2 \\ -2 \\ 1 - \frac{3\varepsilon}{2} \end{bmatrix}$$

While the new, approximate eigenvalues are the same as the old ones. The



first normal mode is unchanged from the unperturbed case since there is no influence due to springs in this mode. In the second mode the center bob moves slightly, the end bob lags by half the displacement of the center. This compensates the un-

equal coupling. In the third mode one of the end bobs leads, on the left, to

compensate for the weak coupling while the one on the right lags, to compensate for the slightly stronger coupling to the center. Net result on center is the same as before hence no frequency change.

This process yields an approximate solution, but the series cannot be extended as it is to terms of order ϵ^2, since this error was already neglected at the start of the problem. One can however expand $H+\epsilon H'$ by Sylvester's theorem with the approximate eigenquantities, expecting it to be correct to order ϵ, and obtaining a small correction of order ϵ^2. The matrix could then be expanded in a commutator series according to powers of ϵ^2, and the process continued until a recognizable series, or the desired accuracy whichever might be preferable was attained.

Actually, in practice, the whole series in powers of ϵ is assumed, and the various terms calculated by evaluating the coefficient of ϵ by suitable

trickery.

Since the matrix $\varepsilon H'$ is regarded as a small perturbation on the matrix H, the process just described is called a perturbation calculation. Depending upon the power of ε to which it is carried, one speakes of n^{th} order perturbations.

The matrix M remains arbitrary to the extent that any matrix which commutes with H may be added to it, as may be seen from the process by which one solves out a commutator bracket. But such a matrix would be diagonal in the eigensystem of H, and if M is to represent a rotation, must be anti-hermitian, so that these diagonal elements must be purely imaginary. This means zero when dealing with a real matrix.

This process certainly fails in the case that two eigenvalues are equal, i.e. when H is degenerate. This situation is discussed in the next problem.

OZONE ~~A~~ WEAK SPRING

The approximation formula for the eigenvectors in the last problem fails when two of the eigenvalues of the unperturbed matrix happen to be equal. The reason for this is that the eigenvectors of the matrix $H+\varepsilon H'$ may be quite well defined, while some of the eigenvectors of H may be taken quite arbitrarily in some subspace. Since then two sets of eigenvectors need not even approximately coincide, and, which is more important, may not even coincide in the limit as $\varepsilon \to 0$, it is impossible to transform between them by an infinitesimal rotation. A similar situation occurs when two eigenvalues have say nearly the same value, for then the reciprocal of their difference will be very large, corresponding to a large change in the direction of the eigenvectors, relative to the small pertubation, ε . Such a situation gives slow convergence when one attempts to carry the perturbation calculation further. On the other hand,

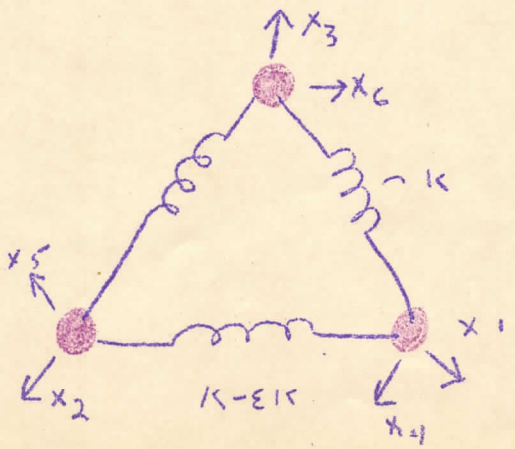
if the numerator $\langle i|H'|j\rangle$ of such a term is also small, it will cancel the small denominator, indicating that the eigenvectors already coincide approximately, or, in the degenerate case, if the element $\langle i|H'|j\rangle$ is zero, it will indicate that the eigenvectors already agree in the limit as $\varepsilon \to 0$, and the calculation will proceed without further complication. For, as may be seen by examining the derivation of the formula for solving for an unknown inside a commutator braclet, this will satisfy that formula.

However, if the elements $\langle i|H'|j\rangle$ zero, the degenerate part of H must commute with the corresponding part of H', which is readily seen to be the case, since this degenerate part is merely a multiple of the unit matrix. Since any vector is the space spanned by the degenerate part may be chosen as an eigenvector, it is a simple matter to choose those which are eigenvectors of H' rather than those which were originally chosen. This then gives the correct eigenvector from which to start the perturbation calculation.

This amounts to diagonalizing a submatrix.

The process is called a zero order perturbation, since one is mere preparing a convenient set of vectors for the first order perturbation calculation.

The ozone molecule considered previously, since it contains degenerate eigenvectors, will serve as an example. Let the perturbation be provided by a spring whose elastic constant differs slightly from that of the others.



The unperturbed elastic constant is already known to be:

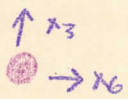
$$H = \frac{k}{m}\begin{bmatrix} 2c^2 & c^2 & c^2 & 0 & sc & -sc \\ c^2 & 2c^2 & c^2 & -sc & 0 & sc \\ c^2 & c^2 & 2c^2 & sc & -sc & 0 \\ 0 & -sc & sc & 2c^2 & -s^2 & -s^2 \\ sc & 0 & -sc & -s^2 & 2s^2 & -s^2 \\ -sc & sc & 0 & -s^2 & -s^2 & 2s^2 \end{bmatrix}$$

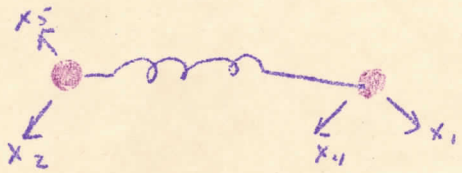
with eigenvectors and eigenvalues

$$\frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\\0\\0\\0\end{bmatrix} \qquad \frac{1}{\sqrt{3}}\begin{bmatrix}0\\0\\0\\1\\1\\1\end{bmatrix} \quad \frac{1}{\sqrt{3}}\begin{bmatrix}1\\-s\\-s\\0\\-c\\c\end{bmatrix} \quad \frac{1}{\sqrt{3}}\begin{bmatrix}0\\c\\-c\\1\\-s\\s\end{bmatrix} \qquad \frac{1}{\sqrt{3}}\begin{bmatrix}1\\-s\\-s\\0\\c\\-c\end{bmatrix} \quad \frac{1}{\sqrt{3}}\begin{bmatrix}0\\-c\\c\\1\\-s\\s\end{bmatrix}$$

$3k/m$ $\qquad$ 0 $\qquad$ $3k/2m$

The perturbing matrix is found by investigating the influence of a small spring connected thusly





elastic constant

and is

$$\varepsilon H' = \varepsilon\frac{k}{m}\begin{bmatrix} -c^2 & -c^2 & 0 & sc & -sc & 0 \\ -c^2 & -c^2 & 0 & sc & -sc & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ sc & sc & 0 & -s^2 & s^2 & 0 \\ -sc & -sc & 0 & s^2 & -s^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

diagonalizing H, and observing the effect of this transformation on H'

$U^{-1}HU =$ $\frac{k}{m}$ $\begin{bmatrix} 3 & & & & \\ & 0 & & & \\ & & 0 & & \\ & & & 0 & \\ & & & & 3/2 \\ & & & & & 3/2 \end{bmatrix}$ $U^{-1}\varepsilon H'U =$ $\frac{\varepsilon k}{m}$ $\begin{bmatrix} -4c^2 & 0 & 0 & 0 & -2c^2 & 3c \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -2c^2 & 0 & 0 & 0 & -c^2 & \frac{3c}{2} \\ 3c & 0 & 0 & 0 & \frac{3c}{2} & -3c^2 \end{bmatrix}$

Since H is degenerate, the corresponding blocks of H' must be diagonalized. They are:

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -c^2 & 3c/2 \\ 3c/2 & -3c^2 \end{bmatrix}$$

The first is already diagonal. The eigenvalues and eigenvectors of the second are

$$\lambda_1 = 0 \quad |1\rangle = \begin{bmatrix} c \\ s \end{bmatrix} \qquad \lambda_2 = -3 \quad |2\rangle = \begin{bmatrix} s \\ -c \end{bmatrix}$$

Let V be

$$V = \begin{bmatrix} \Pi & 0 \\ 0 & \begin{bmatrix} c & s \\ s & -c \end{bmatrix} \end{bmatrix}$$

Then the matrix

$$V^{-1}U^{-1}H'UV = \frac{\varepsilon k}{m} \begin{bmatrix} -3 & 0 & 0 & 0 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -3 & 0 & 0 & 0 & 0 & -3 \end{bmatrix}$$

will then be the correct zero order approximation, to which the perturbation calculation may be applied. The matrix M which yields the corrections to the eigenvectors is:

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 2\varepsilon \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -2\varepsilon & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

so that upon writing the first order approximation to the eigenvectors in the laboratory system, we have

$$\frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1+2\varepsilon C \\ 1-2\varepsilon C \\ a \\ 0 \\ 0 \end{bmatrix} \qquad 3\frac{k}{m}(1-\varepsilon)$$

$$\frac{1}{\sqrt{3}}\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ -S \\ -S \\ 0 \\ -C \\ C \end{bmatrix} \quad \frac{1}{\sqrt{3}}\begin{bmatrix} 0 \\ C \\ -C \\ 1 \\ -S \\ -S \end{bmatrix} \qquad 0$$

$$\frac{1}{\sqrt{3}}\begin{bmatrix} C \\ -C \\ 0 \\ S \\ S \\ -1 \end{bmatrix} \qquad \frac{3k}{2m}$$

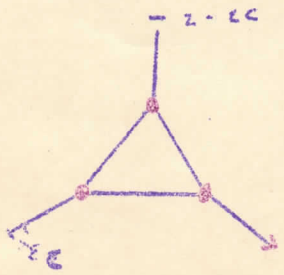
$$\frac{1}{\sqrt{3}}\begin{bmatrix} S+2\varepsilon \\ S+2\varepsilon \\ -1+2\varepsilon \\ -C \\ C \\ 0 \end{bmatrix} \qquad \frac{3k}{2m}(1-2\varepsilon$$

with their corresponding corrected eigenvalues.

It is interesting to note that the degeneracy has been removed for the modes of eigenvalue $3k/2m$, but not for the other degenerate mode, with eigenvalue 0. If one were to apply the second order perturbation calculations to the problem, it might be necessary to repeat this zero-order process again before continuing. In the present example it is obvious that the degeneracy would not be removed until an anisotropic force acting upon the center of mass of the molecule, i.e., an external force, was provided

To sketch the eigenvectors, we have

$$\text{II} \sim \begin{bmatrix} 1 \\ 1+2\varepsilon \\ 1-2\varepsilon \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \lambda_1 = \frac{3k}{m}(1-\varepsilon)$$



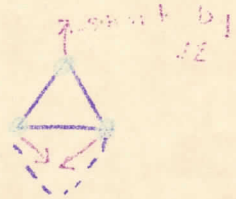
dilation

$\|2\rangle$	rotation	unperturbed
$\|3\rangle$	translation	unperturbed
$\|4\rangle$	translation	unperturbed

$|5\rangle$ $\begin{bmatrix} c \\ -c \\ s \\ s \\ -s \\ -s \end{bmatrix}$ $\lambda_5 = \frac{3k}{2m}$

(unperturbed to first order)

$|6\rangle$ $\begin{bmatrix} -(s+2\varepsilon) \\ -(s+2\varepsilon) \\ (1-2\varepsilon) \\ c \\ -c \\ 0 \end{bmatrix}$ $\lambda_6 = \frac{3k}{2m}(1-2\varepsilon)$



The adjustment of the molecule to the new situation is apparent, especially in the removal of the degeneracy between the modes 5 and 6, since the perturbation has clearly destroyed the previously existing symmetry.

PERTURBATION FORMULAE

Having discussed a technique for obtaining approximate eigenvectors and eigenvalues, we now attempt to obtain a series form for the perturbation calculations. Suppose that a matrix is given in the form of an infinite series:

$$H = H^{(0)} + \varepsilon H^{(1)} + \varepsilon^2 H^{(2)} + \cdots$$

where ε is a small parameter. Now assume the series

$$\lambda_i = \lambda_i^{(0)} + \varepsilon \lambda_i^{(1)} + \varepsilon^2 \lambda_i^{(2)} + \cdots$$

$$|i\rangle = |i\rangle^{(0)} + \varepsilon |i\rangle^{(1)} + \cdots$$

$$\langle i| = \langle i|^{(0)} + \varepsilon \langle i|^{(1)} + \cdots$$

which are power series expansions for the eigenvalues and eigenvectors of H, The corresponding eigenvalue equation may be written in the form

$$\{H - \lambda_i \Pi\} |i\rangle = 0$$

substituting the above series into this expression:

$$\left\{\sum_n \varepsilon^n (H^{(n)} - \lambda_i^{(n)} \Pi\right\} \left\{\sum_n \varepsilon^n |i\rangle^{(n)}\right\} = 0$$

multiplying this expression out, and equating the coefficients of the various powers of ε to zero, there results a system of equations:

$$\{H^{(0)} - \lambda_i^{(0)} \mathbb{I}\} |i\rangle^{(0)} = 0$$

$$\{H^{(0)} - \lambda_i^{(0)} \mathbb{I}\} |i\rangle^{(1)} = -\{H^{(1)} - \lambda_i^{(1)} \mathbb{I}\} |i\rangle^{(0)}$$

$$\{H^{(0)} - \lambda_i^{(0)} \mathbb{I}\} |i\rangle^{(n)} = -\sum_{m=1}^{n} \{H^{(m)} - \lambda_i^{(m)} \mathbb{I}\} |i\rangle^{(n-m)}$$

with a similar set of equations for the left eigenvectors. The first equations of these series are just the characteristic equations for the unperturbed matrix $H^{(0)}$.

The terms of the series for the eigenvalue are obtained by premultiplication of each of these equations by $\langle i|^{(0)}$. The left hand side of the equation vanishes since

$$\langle i|^{(0)} \{H^{(0)} - \lambda_i^{(0)} \mathbb{I}\} = 0.$$

leaving

$$\sum_{m=1}^{n} \langle i|^{(0)} \{H^{(m)} - \lambda_i^{(m)} \mathbb{I}\} |i\rangle^{(n-m)} = 0$$

for $n = 1$, this gives

$$\lambda_i^{(1)} = \langle i|^{(0)} H^{(1)} |i\rangle^{(0)}.$$

for $n = 2$,

$$\lambda_i^{(2)} = -\lambda_i^{(1)} \langle i|^{(0)} |i\rangle^{(1)} + \langle i|^{(0)} H^{(1)} |i\rangle^{(1)} + \langle i|^{(0)} H^{(2)} |i\rangle^{(0)}$$

and so on. Now, $\lambda_i^{(n)}$ can be calculated from quantities bearing a lesser superscript, so that it is possible to solve

$$\{H^{(0)} - \lambda_i^{(0)} \mathbb{I}\} |i\rangle^{(n)} = -\{H^{(n)} - \lambda_i^{(n)} \mathbb{I}\} |i\rangle^{(0)}$$

for $|i\rangle^{(n)}$, since the equation involves nothing else but already known quantities. If there were an inverse matrix $\{H^{(0)} - \lambda_i^{(0)} \mathbb{I}\}^{-1}$, this process would be straightforward. But there is no such matrix. This means that $|i\rangle^{(n)}$ will be indeterminate, and may have any component at all in the direction of $|i\rangle^{(0)}$. It is nevertheless possible to determine the remaining

components of $|i\rangle^{(1)}$. Suppose that $H^{(0)}$ has a complete set of eigenvectors and that λ_i is not a multiple eigenvalue. Then

$$\sum_{j=1}^{n} |j\rangle^{(0)}\langle j|^{(0)} = \mathbb{I}$$

$$\sum_{j=1}^{n} \{H^{(0)} - \lambda_i^{(0)}\mathbb{I}\}|j\rangle^{(0)}\langle j|^{(0)}|i\rangle^{(1)} = -\{H^{(1)} - \lambda_i^{(1)}\mathbb{I}\}|i\rangle^{(0)}$$

$$\langle \ell|^{(0)} \sum_{j\neq i}^{n} \{\lambda_j^{(0)} - \lambda_i^{(0)}\}|j\rangle^{(0)}\langle j|^{(0)}|i\rangle^{(1)} = -\langle \ell|^{(0)}\{H^{(1)} - \lambda_i^{(1)}\mathbb{I}\}|i\rangle^{(0)}$$

$$\{\lambda_\ell^{(0)} - \lambda_i^{(0)}\}\langle \ell|^{(0)}|i\rangle^{(1)} = -\langle \ell|^{(0)}\{H^{(1)} - \lambda_i^{(1)}\mathbb{I}\}|i\rangle^{(0)}$$

$$\langle \ell|^{(0)}|i\rangle^{(1)} = -\frac{\langle \ell|^{(0)}\{H^{(1)} - \lambda_i^{(1)}\mathbb{I}\}|i\rangle^{(0)}}{\lambda_\ell^{(0)} - \lambda_i^{(0)}}$$

The components of $|i\rangle^{(1)}$ in the eigensystem of $H^{(0)}$ have not been discovered; but the zero denominator when $i = \ell$ forestalls any attempt to find the i^{th} component of $|i\rangle^{(1)}$, since the formula then gives the indeterminate form o/o.

Inspection of the formula for $\lambda_i^{(2)}$ shows that it would be convenient to choose $|i\rangle^{(1)}$ so that $\langle i|^{(0)}|i\rangle^{(1)}$ vanishes, which is always possible when one component of $|i\rangle^{(1)}$ is arbitrary. This enables the formula for $|i\rangle^{(1)}$ to be put in a neater form. Observing that when $\langle i|^{(0)}|i\rangle^{(1)} = 0$

$$|i\rangle^{(1)} = \sum_{\ell=1}^{n} |\ell\rangle^{(0)}\langle \ell|^{(0)}|i\rangle^{(1)} = \sum_{\ell\neq i} |\ell\rangle^{(0)}\langle \ell|^{(0)}|i\rangle^{(1)}$$

since a zero term has been omitted from the summation;

$$|i\rangle^{(1)} = -\sum_{\ell\neq i} \frac{|\ell\rangle^{(0)}\langle \ell|^{(0)}\{H^{(1)} - \lambda_i^{(1)}\mathbb{I}\}|i\rangle^{(0)}}{\lambda_\ell^{(0)} - \lambda_i^{(0)}}$$

$$= -\sum_{\ell\neq i} \frac{\langle \ell|^{(0)}H^{(1)}|i\rangle^{(0)}}{\lambda_\ell^{(0)} - \lambda_i^{(0)}}|\ell\rangle^{(0)}$$

These formula, for $|i\rangle^{(1)}$ and $\lambda_i^{(1)}$ are clearly the same results as already obtained, $\langle i|^{(1)}$ may be found by a method similar to that here sketched for $|i\rangle^{(1)}$. But we are in a position to obtain the second order quantities without much additional labor. With the choice of $\langle i|^{(0)}|i\rangle^{(1)} = 0$, we have

$$\lambda_i^{(2)} = \langle i|^{(0)} H^{(1)} |i\rangle^{(1)} + \langle i|^{(0)} H^{(2)} |i\rangle^{(0)}$$

Thus the second order correction to the eigenvalue depends not only on the diagonal elements of the second order correction to the matrix, but also the second order terms remaining from $H^{(1)}$. It will be recalled that those terms did not appear in the first order because of an extremal property of the eigenvectors.

The equation for $|i\rangle^{(2)}$ is

$$\left\{H^{(0)} - \lambda_i^{(0)} \pm\right\}|i\rangle^{(2)} = -\left\{H^{(1)} - \lambda_i^{(1)} \mathbb{I}\right\}|i\rangle^{(1)} - \left\{H^{(2)} - \lambda_i^{(2)}\right\}|i\rangle^{(0)}$$

The same observations made for $|i\rangle^{(1)}$ apply here, as does the method for solving for $|i\rangle^{(2)}$. As before, set $\langle i|^{(0)}|i\rangle^{(2)} = 0$. Upon performing the necessary algebra, the result can be shown to be

$$|i\rangle^{(2)} = \sum_{\ell \neq i} \frac{1}{\lambda_\ell^{(0)} - \lambda_i^{(0)}} \left\{ \sum_{\substack{m \neq i \\ \ell \neq i}} \frac{\langle i|^{(0)} H^{(1)} |m\rangle^{0} \langle m|^{(0)} H^{(1)} |i\rangle^{(0)}}{\lambda_m^{(0)} - \lambda_i^{(0)}} - \langle \ell|^{(0)} H^{(2)} |i\rangle^{(0)} \right\} |\ell\rangle$$

It is clear that the series may be extended, term by term, to any desired order.

These formula may be summarized; recalling that $\langle i|^{(0)} A |j\rangle^{(0)} = [A]_{ij}$ in the eigensystem, we have

$$\lambda_i^{(1)} = H_{ii}^{(1)}$$

$$|i\rangle^{(1)} = -\sum_{\ell \neq i} \frac{H_{\ell i}^{(1)}}{\lambda_\ell^{(0)} - \lambda_i^{(0)}} |\ell\rangle^{(0)}$$

$$\lambda_i^{(2)} = -\sum_{\ell \neq i} \frac{H_{i\ell}^{(1)} H_{\ell i}^{(1)}}{\lambda_\ell^{(0)} - \lambda_i^{(0)}} + H_{ii}^{(2)}$$

$$|i\rangle^{(2)} = \sum_{i \neq j} \frac{1}{\lambda_i^{(0)} - \lambda_j^{(0)}} \left\{ \sum_{\substack{m \neq i \\ \ell \neq i}} \frac{H_{jm}^{(1)} H_{mi}^{(1)}}{\lambda_m^{(0)} - \lambda_i^{(0)}} - H_{ji}^{(2)} \right\} |\ell\rangle^{(0)}$$

etc.

The case of repeated eigenvalues must be treated by zero$^{\text{th}}$ order perturbation theory in which a correct set of eigenvectors of $H^{(o)}$ is prepared so that the eigenvectors of H approach them in the limit. This process has already been discussed.

ENERGY CONSIDERATIONS

The vibrational problems presented have all been discussed from the point of view of the forces operating between the particles, and the energy considerations have been lacking altogether. Suppose that the potential energy of a system of particles is given by $V(x_1, x_2, \cdots, x_n)$, where the x's are the various particle coordinates. Then, if $x_1 = 0, x_2 = 0, \cdots, x_n = 0$ is the particle configuration when the system is in equilibrium, it is possible to develop the potential in a Taylor's series about the origin.

$$V(x_1, x_2, \cdots, x_n) = V(0, 0, \cdots, 0) + \frac{\partial V}{\partial x_1} x_1 + \frac{\partial U}{\partial x_2} x_2 + \cdots + \frac{\partial U}{\partial x_n} x_n + \cdots$$

$$= V(0) + \sum \frac{\partial U}{\partial x_i}\Big|_{x_i = 0} x_i + \frac{1}{2!} \sum \sum x_i \frac{\partial^2 V}{\partial x_i \partial x_j}\Big|_{x=0} x_j + \cdots$$

The potential is indeterminate by an additive constant, so that it is possible to set $V(0) = 0$. Since the particles are presumed in equilibrium at $x_i^\circ = 0$ $\frac{\partial V}{\partial x_i}\Big|_{x=0} = 0$; there are no forces acting at the origin. This leaves the second term as an approximation to the potential.

$$V(x_i) \doteq \frac{1}{2} \sum_i \sum_j x_i \frac{\partial^2 V}{\partial x_i \partial x_j}\Big|_{\substack{x_i = 0 \\ x_j = 0}} x_j$$

which may be a good approximation, provided that it itself does not vanish, and that the x_i's and d are sufficiently small that the series converges. But the terms

$$\frac{\partial^2 V}{\partial x_i \partial x_j}\Big|_{\substack{x_i = 0 \\ x_j = 0}}$$

are just the elements of the matrix of elastic constants for an equivalent linear problem, and this first approximation is one-half the quadratic form of the elastic constant matrix:

$$V(x_i) = \frac{1}{2} \langle x_i | K | x_j \rangle$$

A result which could also have been obtained from the expression

$$V(x_i) = \int \langle \text{force} | d x_i \rangle$$

$$= \int \langle x_i | K | d x_j \rangle$$

But, the eigenvectors provide extremals of the quadratic form of their matrix, so that the motion according to normal modes is motion which makes the potential energy of the system an extremal. This is a circumstance which may sometimes be used to advantage either in determining normal modes by inspection or in calculating them.

The fact that the normal coordinates are orthogonal for a symmetric matrix shows that the total energy of the system can be written as a sum of the energies in the normal modes and that the normal modes do no work on one another. Newton's third law requires the matrix of elastic constants to be symmetric, by requiring that the k^{th} force due to a displacement along the j^{th} coordinate of a particle be equal to the force when the particles are interchanged, i.e., the j^{th} force on the i^{th} particle.

THE TRIPLE PENDULUM

The problem of discovering the eigenvalues and eigenvectors of a matrix is of considerable importance, in veiw of the convenience afforded by using Sylvester's theorem in the solution of problems. Although they can be found for small matrices rather easily, as the site of a matrix increases, so does

the tediousness of the arithemetic involved in carrying out the formulae for eigenvectors and eigenvalues.

In addition to the techniques already presented, which, with the exception of the perturbation theories, have been exact in nature, there are a few, primarily numerical in nature, which are often useful. We show two of them here.

The first consists in noticing the approximate form of the high powers of a matrix. If it has a complete set of eigenvectors, we see that:

$$M^n = \sum_i \lambda_i^n |i\rangle\langle i|$$

The largest eigenvalue dominates this expression, and one commits a small error in writing:

$$M^n \doteq \lambda_{max}^n |k\rangle\langle k|$$

for sufficiently large n.

Now the trace of such a large power is

$$\mathrm{Tr}\, M^n = \sum \lambda_i^n$$

which becomes in the limit, as we also take n^{th} roots.

$$\lambda_{max} \doteq \sqrt[n]{\lambda_1^n + \lambda_2^n + \cdots + \lambda_R^n}$$

also:

$$\lim_{n\to\infty} \left\{ \frac{1}{\lambda_{max}^n} M^n \right\}$$

gives

$$\lim_{n\to\infty} \frac{1}{\lambda_{max}^n} \sum \lambda_i^n |i\rangle\langle i| = \lim_{n\to\infty} \sum \left(\frac{\lambda_i}{\lambda_{max}}\right)^n |i\rangle\langle i| = |k\rangle\langle k|$$

if λ_K is the greatest eigenvalue. If the eigenvalue is degenerate one obtains instead, the projective operator into its subspace which will not be a linear subspace, but of a dimensionality corresponding to the multiplicity of the eigenvalue.

The next largest eigenvalue and eigenvector may obviously be found by applying the same process to the matrix

$$\{M - \lambda_{max} |max\rangle\langle max|\}$$

and so on. This method has as disadvantages the facts that it finds only the largest eigenvalue and corresponding eigenvectors each time, and that the extraction of succeeding eigenvalues depends upon the accuracy with which the preceding ones have been discovered. Furthermore, the process converges rapidly only if the eigenvalues are well separated in value, the smaller the ration

$$\frac{\lambda \text{ next largest}}{\lambda \text{ largest}}$$

the faster.

The second process claims that it is an reasonable guess for the value of some eigenvalue and $|i\rangle$ is an approximation to the corresponding eigenvector then

$$\{M - \lambda \mathbb{I}\}^{-1} |i\rangle$$

will be an even better one. It will not be normalized. To see the inner working of this process, write the matrix in its eigensystem. Then one has

$$\begin{bmatrix} \frac{1}{\lambda_1 - \lambda} & & & \\ & \frac{1}{\lambda_2 - \lambda} & & \\ & & \cdots & \\ & & & \frac{1}{\lambda_n - \lambda} \end{bmatrix}$$

and since $\lambda - \lambda_i$ is small, its reciprocal will be large, and hence the matrix product

$$\begin{bmatrix} \frac{1}{\lambda - \lambda} & & & \\ & \frac{1}{\lambda_2 - \lambda} & & \\ & & \cdots & \\ & & & \frac{1}{\lambda_n - \lambda} \end{bmatrix} \quad \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_4 \end{bmatrix}$$

will emphasize the i^{th} component of the vector, and if the new approximation is normalized each time, one has upon repeating the process several times:

$$|i\rangle_n = \begin{bmatrix} \left(\frac{\lambda_i - \lambda}{\lambda_1 - \lambda}\right)^n & & \\ & \left(\frac{\lambda_i - \lambda}{\lambda_2 - \lambda}\right)^n & \\ & & \ddots \end{bmatrix} |i\rangle_0 .$$

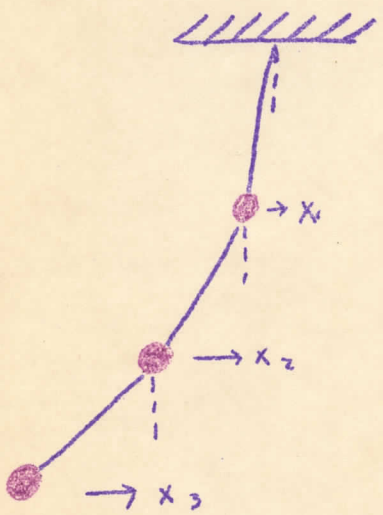
which becomes in the limit

$$|i\rangle_\infty = \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & 1 & \\ & & & 0 \end{bmatrix} |i\rangle_0 = |i\rangle \text{ actual}$$

since in this system $|i\rangle \text{ actual} = |i\rangle$.

This method depends for its success upon an initial good guess for λ_i and upon the other eigenvalues differing considerably from the one under consideration.

The first process may be illustrated by the problem of the triple pendulum.



Three pendula, whose hobs have equal masses and whose suspensions each have equal lengths are hung one from another, given certain initial displacements, and left to swing.

Rather than displace the particles, and calculate the forces arising, it is more convenient in this problem to apply given forces and see how far the particles are displaced. Thus applying a force to the first particle horizontally and considering small unit forces we have all the particles displaced equally by an amount

$$a = \frac{l}{3mg}$$

If the unit force is applied to the second hob, the first will be displaced a

distance a, but the second and third will be displaced further, by a distance a + b, where

$$b = \frac{l}{2mg}$$

Finally if a unit horizonal force is applied to the third hob, displace the first by a, the second by a+ b, and the third by (a+ b + c) with

$$c = \frac{l}{mg}$$

This gives a set of equations

$$\gamma_1 = a f_1 + a f_2 + a f_3$$
$$\gamma_2 = a f_1 + (a+b) f_2 + (a+b) f_3$$
$$\gamma_3 = a f_1 + (a+b) f_2 + (a+b+c) f_3$$

or, equating the γ's to the inertia forces, writing a matrix equation

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{l}{6g} \begin{bmatrix} 2 & 2 & 2 \\ 2 & 5 & 5 \\ 2 & 5 & 11 \end{bmatrix} \begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{bmatrix}$$

To determine the normal modes and frequencies of their vibrations one must find the eigenvalue and eigenvectors of

$$M \cdot \begin{bmatrix} 2 & 2 & 2 \\ 2 & 5 & 5 \\ 2 & 5 & 11 \end{bmatrix}$$

If λ_i are the eigenvalues of this matrix, the frequencies of the motion are given by

$$\omega_i = \sqrt{\frac{6g}{\lambda_i l}}$$

We obtain by calculation

$$M = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 5 & 5 \\ 2 & 5 & 11 \end{bmatrix} \qquad \Sigma M = 18$$

$$M^2 = \begin{bmatrix} 12 & 24 & 36 \\ 24 & 54 & 84 \\ 36 & 84 & 150 \end{bmatrix} \qquad \sqrt{\operatorname{tr} M^2} = 14.70$$

$$M^3 = \begin{bmatrix} 144 & 324 & 540 \\ 324 & 738 & 1242 \\ 540 & 1242 & 2142 \end{bmatrix} \qquad \sqrt[3]{\operatorname{tr} M^3} = 14.45$$

It then appears that M^3 will be a reasonable stopping point for two-figure accuracy; dividing M^3 by its trace, we have

$$|1\rangle\langle 1| \doteq \frac{M^3}{\operatorname{tr} M^3} \qquad \begin{bmatrix} .0475 & .107 & .178 \\ .107 & .244 & .410 \\ .178 & .410 & .707 \end{bmatrix}$$

$$|1\rangle \doteq \begin{bmatrix} .218 \\ .493 \\ .841 \end{bmatrix}$$

These results may be compared with "Pipes" page 200, who does a more extensive approximation

$$\lambda_1 \doteq 14.4309$$

$$|1\rangle \doteq \begin{bmatrix} .254885 \\ .584225 \\ 1.000000 \end{bmatrix}$$

If $\lambda_1 |1\rangle\langle 1|$ is subtracted from M we have to subtract

$$\begin{bmatrix} .68 & 1.55 & 2.57 \\ 1.55 & 3.52 & 5.80 \\ 2.57 & 5.80 & 10.2 \end{bmatrix}$$

to give

$$M' = \begin{bmatrix} 1.32 & 0.45 & -0.57 \\ 0.45 & 1.48 & -0.80 \\ -0.57 & -0.80 & 0.80 \end{bmatrix} \qquad \operatorname{Tr} M' = 3.60$$

$$M'^2 = \begin{bmatrix} 2.26 & 1.72 & -1.56 \\ 1.72 & 3.03 & -2.08 \\ -1.56 & -2.08 & 1.60 \end{bmatrix} \qquad \sqrt{\operatorname{Tr} M'^2} = 2.61$$

$$M'\binom{4}{3} = \begin{bmatrix} 10.50(4) \\ 4.65(3) & & \sim \\ & 16.48(4) \\ & 6.46(3) & \\ & & (4)9.34 \\ \sim & & (3)3.83 \end{bmatrix} \qquad \sqrt[3]{\operatorname{Tr} M'^3} = 2.46$$

$$\sqrt[4]{\operatorname{Tr} M'^4} = 2.48$$

$$\frac{M'^2}{\operatorname{Tr} M'^2} \doteq |2\rangle\langle 2| \begin{bmatrix} .33 & .25 & -.23 \\ .25 & .44 & -.30 \\ -.23 & -.30 & .23 \end{bmatrix}$$

$$|2\rangle \doteq \begin{bmatrix} .57 \\ .67 \\ -.48 \end{bmatrix}$$

$$\lambda_2 |2\rangle\langle 2| = \begin{bmatrix} .86 & .65 & -.60 \\ .65 & 1.15 & -.78 \\ -.60 & -.78 & .60 \end{bmatrix}$$

This process continues with

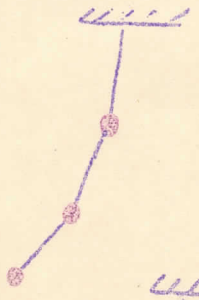
$$M'' = \begin{bmatrix} .46 & -.20 & -.03 \\ -.20 & .33 & -.20 \\ .03 & -.02 & .20 \end{bmatrix} \qquad \operatorname{Tr} M'' = .99$$

$$|3\rangle \doteq \begin{bmatrix} .68 \\ -.57 \\ .45 \end{bmatrix}$$

When these results are checked by multiplication, it is found that (1) is determined fairly well, as is λ_1, λ_2, and λ_3. However, $|3\rangle$ is quite badly off, due presumably to an error in the subtraction 11-10.8. This illustrates the care with which the first eigenvectors should be calculated; even though the accuracy is not needed for then, it shows up later in the calculation.

Sketching the normal modes we have

(1)



lowest frequency

(2)

(3)

highest frequency

Note that in such a problem as this, the end of the chain has a tendency to whip around, while the end of the chain nearer the supports, having a much greater restoring force moves but slightly.

QUASI - CONTINUUA

THE VIBRATING STRING

The vibrating string illustrates the use of a discrete model as a basis from which to obtain a differential equation, through a limiting process by which a supposedly continuous medium is more or less accurately approximated.

Suppose that there are n balls, each of mass m, suspended between two supports by springs of elastic constant k, and that gravity can be neglected. The following diagram illustrates the scheme:

The equations of motion can be obtained by considering the forces on a certain ball in terms of the displacements of the neighboring balls.

$$m\ddot{u}_i = k\{u_{i-1} - 2u_i + u_{i+1}\} \qquad i \neq 0,\ i+1 \neq n+1$$

$$m\ddot{u}_1 = k\{-2u_1 + u_2\}$$

$$m\ddot{u}_n = k\{u_{n-1} - 2u_n\}$$

Written in matrix form this is

$$\begin{bmatrix} \ddot{u}_1 \\ \ddot{u}_2 \\ \cdots \\ \ddot{u}_n \end{bmatrix} = \frac{k}{m} \begin{bmatrix} -2 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & \cdots & 0 \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \cdots \\ u_n \end{bmatrix}$$

This matrix is inconvenient to use, primarily because it cannot be written as a sum of simpler matrices which commute in any obvious fashion. The difficulty is clearly due to the supports at the end of the string. One may attempt to remove it by adding a $n+1$ $\underline{st}$ ball, and then continuing the string periodically, later imposing the boundary condition that $u_{n+1} = 0$. This scheme fails, however, since the boundary condition implies a force of constraint which is not contained in the equation of motion. But if such a force is

already provided in the equation of motion there can be no objection to such a procedure, and this will be the case if an $(n+1)^{st}$ ball is added, and then the string doubled in length according to the condition that

$$u_{n-1+1} = u_{n+1+1}$$

or

$$u_i = u_{2n+i+2}$$

and then continued periodically. This gives a matrix equation, when the periodicity is borne in mind

$$\begin{bmatrix} \ddot{u}_1 \\ \ddot{u}_2 \\ \cdots \\ \ddot{u}_{2n+1} \\ \ddot{u}_{2n+2} \end{bmatrix} = \frac{K}{m} \begin{bmatrix} -2 & 1 & 0 & \cdots & 1 \\ 1 & -2 & 1 & \cdots & 0 \\ \cdots & & & & \\ 1 & 0 & 0 & \cdots & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \cdots \\ u_{2n+1} \\ u_{2n+2} \end{bmatrix}$$

where the elastic constant may now be written as the form of three matrices:

$$\frac{K}{m} \{ S' - 2\Pi + S^{-1} \}$$

where S' is the matrix

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

These matrices all obviously commute. The eigenvalues of S' are the solution of

$$\begin{bmatrix} -\lambda & 1 & 0 & \cdots & 0 \\ 0 & -\lambda & 1 & \cdots & 0 \\ \cdots & & & & \\ 1 & 0 & 0 & \cdots & -\lambda \end{bmatrix} = 0$$

The development of this determinant by Laplace's method gives

$$\lambda_K = e^{2\pi i \frac{K}{2n+2}}$$

so that the eigenvalues are the distinct $2n+2^{nd}$ roots of unity. The eigenvalue equation

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & & & \\ 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \cdots \\ a_{2n+2} \end{bmatrix} = \lambda_K \begin{bmatrix} a_1 \\ a_2 \\ \\ a_{2n+2} \end{bmatrix}$$

quickly yields the eigenvectors

$$|1\rangle = \frac{1}{\sqrt{2n+2}} \begin{bmatrix} e^{2\pi i \frac{2n+1}{2n+2}} \\ \cdots \\ e^{2\pi i \frac{1}{2n+2}} \end{bmatrix} \quad |2\rangle = \frac{1}{\sqrt{2n+2}} \begin{bmatrix} e^{2\pi i \, 2\cdot\frac{2n+1}{2n+2}} \\ \cdots \\ e^{2\pi i \, 2\frac{1}{2n+2}} \\ 1 \end{bmatrix} \quad |\ell\rangle = \frac{1}{\sqrt{2n+2}} \begin{bmatrix} e^{2\pi i \frac{2n+1}{2n+2}\ell} \\ \cdots \\ e^{2\pi i \frac{1}{2n+2}\ell} \\ 1 \end{bmatrix}$$

Thus $S' - 2\frac{1}{\pi} + S^{-1}$ has the eigenvalues

$$e^{2\pi i \frac{k}{2n+2}} - 2 + e^{-2\pi i \frac{k}{2n+2}}$$

or

$$-4 \sin^2 \frac{k\pi}{2n+2}$$

Since these are degenerate, one can pick linear combinations of them to give

$$|\ell\rangle^{(1)} = \frac{1}{2}\{|\ell\rangle + |2n+2-\ell\rangle\} = \frac{1}{2}\| e^{2\pi i \frac{2n+2-s}{2n+2}\ell} + e^{2\pi i \frac{(2n+2-s)(2n+2-\ell)}{2n+2}} \|$$

$$= \frac{1}{2}\| e^{-2\pi i \frac{s\ell}{2n+2}} + e^{2\pi i \frac{s\ell}{2n+2}} \|$$

$$= \| \cos \frac{2\pi s\ell}{2n+2} \|$$

$$|\ell\rangle^2 = \frac{i}{2}\{|\ell\rangle - |2n+2-\ell\rangle\} = \| \sin \frac{2\pi s\ell}{2n+2} \|$$

So that we may choose, as eigenvectors of K

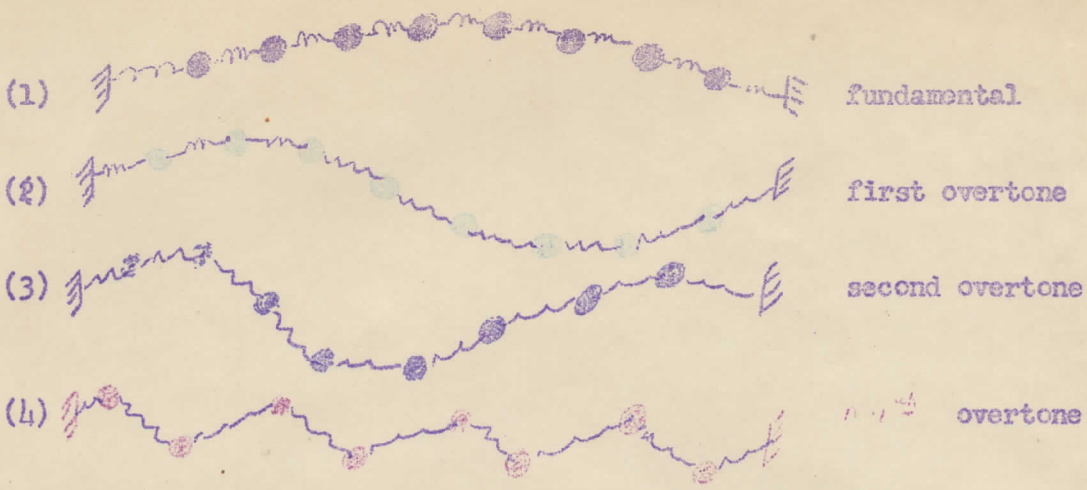
$$|1\rangle = \begin{bmatrix} \sin \frac{\pi}{2n+2} \\ \sin \frac{2\pi+2}{2n+2} \\ \cdots \\ \sin \frac{2\pi n}{2n+2} \end{bmatrix} \quad |2\rangle = \begin{bmatrix} \sin 2\frac{2\pi}{2n+2} \\ \sin 2\frac{2\pi+2}{2n+2} \\ \cdots \\ \sin 2\frac{2\pi n}{2n+2} \end{bmatrix} \quad \text{etc.}$$

and drop the cosine eigenvectors, by imposing the boundary condition that $U_{n+1} = 0$, which leaves just n eigenvectors, and the degeneracy is removed.

We have

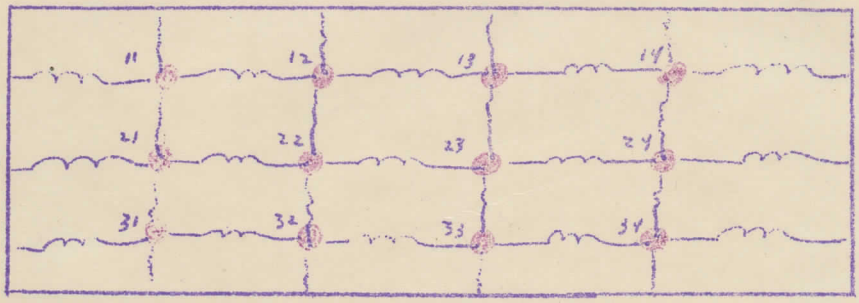
$$\sqrt{\lambda_\sigma} = 2 \sin \sqrt{\frac{k}{m}} \frac{\sigma\pi}{2n+2} \qquad \sigma = 1, \cdots, n$$

The normal modes may now be sketched:



THE VIBRATING BEDSPRING

It is possible to solve two dimensional problems with the same methods used for one dimensional problem. As an example, consider the small vibrations of the following collection of springs and masses.



Except for the edges the equations of motion read:

$$\ddot{U}_{i,j} = \frac{k}{m}\{-4U_{i,j} + U_{i+1,j} + U_{i-1,j} + U_{i,j+1} + U_{i,j-1}\}$$

and it is formally correct there if one puts $U = 0$ for any coordinates that lie beyond the rim of the figure. In matrix form this becomes: $\ddot{U} = -KU$

$$K = \frac{k}{m} \cdot \left[\begin{array}{cccc|cccc|cccc} -4 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -4 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -4 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -4 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 & -4 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & -4 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & -4 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & -4 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -4 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -4 \end{array}\right]$$

or, writing the appropriate submatrices:

$$\mathbb{K} = \begin{bmatrix} \mathbb{N} & \mathbb{I} & \mathbb{O} \\ \mathbb{I} & \mathbb{N} & \mathbb{I} \\ \mathbb{O} & \mathbb{I} & \mathbb{N} \end{bmatrix}$$

Since each element of $\mathbb{U}$, the coordinate vector, has not one but two indices, and since the elements of the elastic constant should have not two, but four indices, some convention other than the usual one must be used for ordering the elements. The one adopted here is to list them in the library order, that is the first indices are ordered first, among the vectors, then the second indices with the same first indices are ordered, With matrices, the first pair of indices determines the row, the second pair the column in which the element is located, by the same convention.

To find the eigenvectors and eigenvalues of $\mathbb{K}$, one may first diagonalize the supermatrix. Since all its elements commute with one another, it is apparent that this may be done as though the elements were Scalars. The determinant which must be evaluated to find the eigenvalues is the one treated in the discussion on recursion formulae. The eigenvalues are the matrices

$$\lambda_{k'} = \mathbb{N} - 2\cos\frac{k'\pi}{4}\mathbb{I}$$

and the eigenvalues of these eigenvalues are

$$\lambda_{k'k} = -4\left\{\sin^2\frac{k'\pi}{2\cdot 4} + \sin^2\frac{k\pi}{2\cdot 5}\right\}$$

and are in their turn the eigenvalues of $\mathbb{K}$. The eigenvectors of the supermatrix are the same as those of the matrix treated in the vibrating string problem, which are:

$$|1i\rangle = \begin{bmatrix} (\sin\frac{\pi}{3+1}\cdot i)\,\underline{\underline{I}} \\ (\sin\frac{2\pi}{3+1}\cdot i)\,\underline{\underline{I}} \\ \sin\frac{3\pi}{3+1}\cdot i\;\underline{\underline{I}} \end{bmatrix}, \text{ i.e. } |1\rangle = \frac{1}{2}\begin{bmatrix} \underline{\underline{I}} \\ \sqrt{2}\,\underline{\underline{I}} \\ \underline{\underline{I}} \end{bmatrix} \quad |2\rangle = \frac{1}{\sqrt{2}}\begin{bmatrix} \underline{\underline{I}} \\ \underline{\underline{0}} \\ -\underline{\underline{I}} \end{bmatrix} \quad |3\rangle = \frac{1}{2}\begin{bmatrix} \underline{\underline{I}} \\ -\sqrt{2}\,\underline{\underline{I}} \\ \underline{\underline{I}} \end{bmatrix}$$

and the eigenvectors of the eigenvalues of the supermatrix are gotten from the same formula, their matrices being similar in structure. Thus the matrix which diagonalizes the supermatrix is

$$\mathcal{H} = \begin{bmatrix} \frac{1}{2}\underline{\underline{I}} & \frac{1}{\sqrt{2}}\underline{\underline{I}} & \frac{1}{2}\underline{\underline{I}} \\ \frac{\sqrt{2}}{2}\underline{\underline{I}} & \underline{\underline{0}} & -\frac{\sqrt{2}}{2}\underline{\underline{I}} \\ \frac{1}{2}\underline{\underline{I}} & -\frac{1}{\sqrt{2}}\underline{\underline{I}} & \frac{1}{2}\underline{\underline{I}} \end{bmatrix}$$

The matrix which diagonalizes the diagonal elements of the diagonalized supermatrix is

$$\mathcal{H} = \begin{bmatrix} \sin\frac{\pi}{5} & \sin\frac{2\pi}{5} & \sin\frac{3\pi}{5} & \sin\frac{4\pi}{5} \\ \sin\frac{2\pi}{5} & \sin\frac{4\pi}{5} & \sin\frac{6\pi}{5} & \sin\frac{8\pi}{5} \\ \sin\frac{3\pi}{5} & \sin\frac{6\pi}{5} & \sin\frac{9\pi}{5} & \sin\frac{12\pi}{5} \\ \sin\frac{4\pi}{5} & \sin\frac{8\pi}{5} & \sin\frac{12\pi}{5} & \sin\frac{16\pi}{5} \end{bmatrix}$$

So that the matrix:

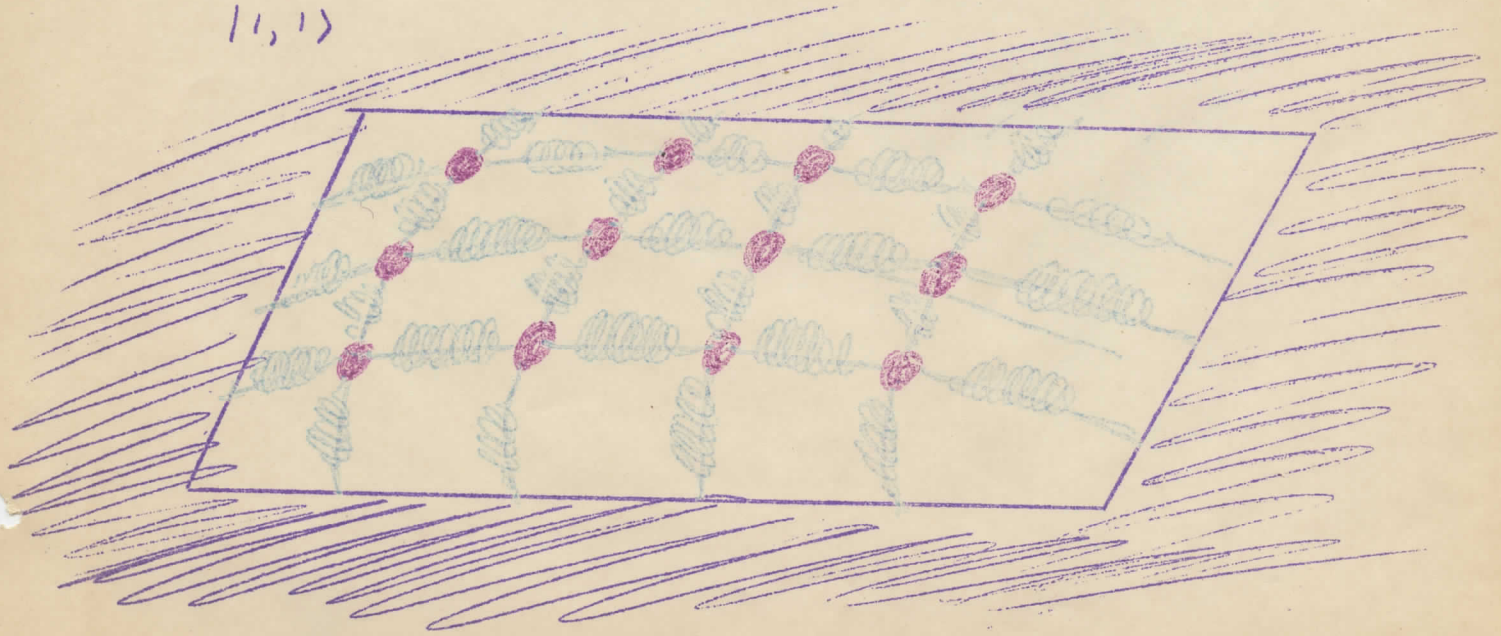
$$\begin{bmatrix} \mathcal{H} & 0 & 0 \\ 0 & \mathcal{H} & 0 \\ 0 & 0 & \mathcal{H} \end{bmatrix} = \mathcal{H}\,\underline{\underline{I}}$$

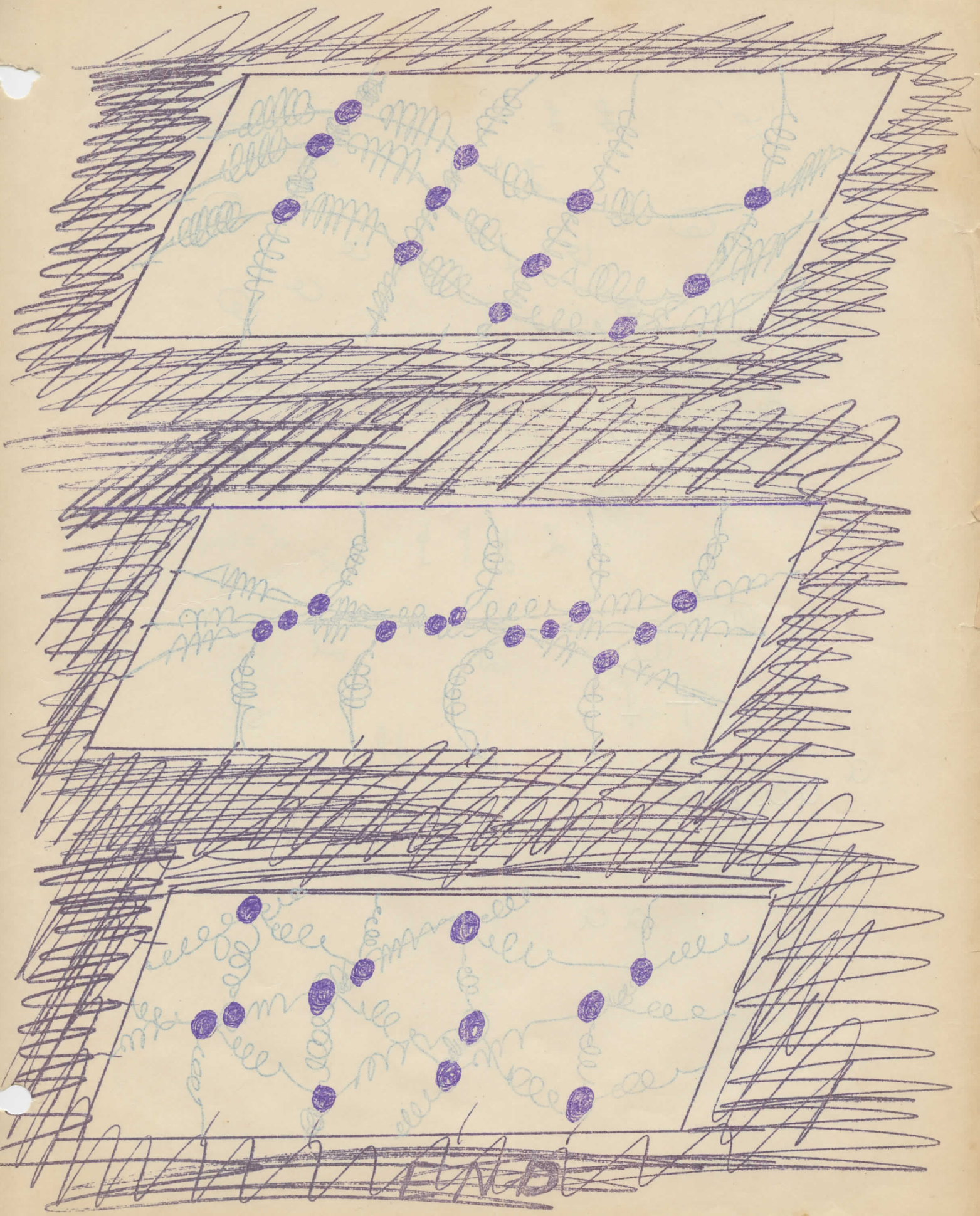
completes the diagonalization by diagonalizing the diagonal elements of the diagonalized supermatrix. The final eigenvectors are then

$$|K';K\rangle = \begin{bmatrix} \sin\frac{K\pi}{5} & \sin\frac{K'\pi}{4} \\ \sin\frac{2K\pi}{5} & \sin\frac{K'\pi}{4} \\ - & - \\ \sin\frac{\pi K}{5} & \sin\frac{2K'\pi}{4} \\ \sin\frac{2\pi K}{5} & \sin\frac{2K'\pi}{4} \\ - & - \\ \sin\frac{\pi K}{5} & \sin\frac{3K'\pi}{4} \\ \sin\frac{2\pi K}{5} & \sin\frac{3K'\pi}{4} \\ - & - \\ \sin\frac{4\pi K}{5} & \sin\frac{3\pi K'}{4} \end{bmatrix} = |\sin\frac{K\pi a}{5} \sin\frac{K'\pi b}{4}\rangle$$

Note that each element of the eigenvectors have been written as a product of two functions, one of y alone and one of y alone, for as y varies the first index varies and one moves in the y direction in the network, while as the second index K varies the y position of the ball under consideration changes. This has again separated the variables, a consequence of finding the normal modes. Sketching some at the normal modes we have:

$|1,1\rangle$




THE END