5.80 Small-Molecule Spectroscopy and Dynamics
Fall 2008

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Chemistry 5.76
Revised February, 1982

**NOTES ON MATRIX METHODS**

# 1. Matrix Algebra

Margenau and Murphy, The Mathematics of Physics and Chemistry, Chapter 10, give almost all the background relevant for physical applications. Schiff, Quantum Mechanics, Chapter VI, and many other texts, give useful brief summaries. Here we shall note only a few of the most important definitions and theorems.

The *matrix* is a square or rectangular array of numbers that can be added to or multiplied into another matrix according to certain rules. For a matrix **A**, the elements are denoted $\mathbf{A}_{ij}$, where indices refer to rows and columns respectively.

*Addition*. If two matrices **A** and **B** have the same *rank* (same number of rows and the same number of columns) they may be added element by element. If the sum matrix is called $\mathbf{C} = \mathbf{A} + \mathbf{B}$ then

$$C_{ij} = A_{ij} + B_{ij}. \tag{1.1}$$

The addition is commutative:

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}. \tag{1.2}$$

*Multiplication*. Two matrices **A** and **B** can be multiplied together to give **C**= **AB** by the following rule

$$C_{ij} = \Sigma_k A_{ik} B_{kj}. \tag{1.3}$$

If **A** has $m$ rows and $n$ columns, **B** must have $n$ rows and $m$ columns if **C** is a square matrix with $m$ rows and $m$ columns. Each element $C_{ij}$ of the product is a sum of products along the $i^{\text{th}}$ row of **A** and the $j^{\text{th}}$ column of **B**. For example, for $2 \times 2$ matrices:

$$\begin{pmatrix} A_{11} A_{12} \\ A_{21} A_{22} \end{pmatrix} \begin{pmatrix} B_{11} B_{12} \\ B_{21} B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}$$

Note that in general matrix multiplication is not commutative:

$$\mathbf{AB} \neq \mathbf{BA}. \tag{1.4}$$

From the definitions it follows at once that the distributive law of multiplication still holds for matrices:

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}. \tag{1.5}$$

Likewise, the associative law holds:

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}. \tag{1.6}$$

*Inverse of a matrix*. Matrix division is not defined. A matrix may or may not possess an inverse $\mathbf{A}^{-1}$, which is defined by

$$\mathbf{AA}^{-1} = \mathbf{E} \text{ and } \mathbf{A}^{-1}\mathbf{A} = \mathbf{E} \tag{1.7}$$

where $\mathbf{E}$ is the *unit matrix* (consisting of 1's along the diagonal and 0's elsewhere). The Matrix $\mathbf{A}$ is said to be *nonsingular* if it possesses an inverse and singular if it does not. If $\mathbf{A}$ is nonsingular and of finite rank, it can be shown to be square and the $ij$ element of its inverse is just the cofactor of $A_{ji}$ divided by the determinant of $\mathbf{A}$. Thus $\mathbf{A}$ is singular if its determinant vanishes. It is readily verified from (1.3) and (1.7) that for nonsingular matrices

$$(\mathbf{ABC})^{-1} = \mathbf{C}^{-1}\mathbf{B}^{-1}\mathbf{A}^{-1}. \tag{1.8}$$

The *determinant* of a square matrix is found from the usual rule for the computation of the determinant of a square array of numbers, e.g.

$$\begin{vmatrix} A_{11}A_{12}A_{13} \\ A_{21}A_{22}A_{23} \\ A_{31}A_{32}A_{33} \end{vmatrix} = A_{11}A_{22}A_{33} + A_{12}A_{23}A_{31} + A_{13}A_{32}A_{21} - A_{13}A_{22}A_{31} - A_{12}A_{21}A_{33} - A_{11}A_{32}A_{23}.$$

The complementary minor of an element $A_{ij}$ within a determinant is the smaller determinant formed by crossing out row $i$ and column $j$; thus, for the above example the minor of $A_{21}$ is

$$\begin{vmatrix} A_{12}A_{13} \\ A_{32}A_{33} \end{vmatrix}$$

Another important quantity derived from a matrix is the *trace* (also called the *spur* or *diagonal sum*), which is the sum of the diagonal elements,

$$\mathrm{Tr}\mathbf{A} = \Sigma_i A_{ii}.$$

*Transpose and Adjoint Matrices*. The transpose $\mathbf{A}'$ of a square matrix is formed by interchanging rows and columns, i.e. $A'_{ij} = A_{ji}$. The adjoint $\mathbf{A}^\dagger$ is obtained by taking the complex conjugate of each element of the transpose, i.e. $A^\dagger_{ij} = A^\star_{ji}$. From the definitions it is readily verified that the adjoint of a product of matrices is the product of their adjoints in the reverse order,

$$(\mathbf{ABC})^\dagger = \mathbf{C}^\dagger\mathbf{B}^\dagger\mathbf{A}^\dagger. \tag{1.9}$$

*Special Matrices*. A matrix is *real* if all the elements are real. It is *symmetric* if it is the same as its transpose; *Hermitian* if it is the same as its adjoint; *orthogonal* if it is the same as its inverse; *unitary* if its adjoint is the same as its inverse. In summary:

| Type of Matrix | Definition Relation |
|---|---|
| Real | $\mathbf{A}^\star = \mathbf{A}$ |
| Symmetric | $\mathbf{A}' = \mathbf{A}$ |
| Hermitian | $\mathbf{A}'^\dagger = \mathbf{A}$ |
| Orthogonal | $\mathbf{A}' = \mathbf{A}^{-1}$ |
| Unitary | $\mathbf{A}^\dagger = \mathbf{A}^{-1}$ |

Also, a matrix is *diagonal* if all its non-diagonal elements are zero and *null* if all of its elements are zero. Be forewarned that various authors use different symbols to denote some of these special matrices.

*Transformations of Matrices*. Two square matrices **A** and **B** are said to be related by a similarity transformation if

$$\mathbf{A} = \mathbf{T}^{-1}\mathbf{B}\mathbf{T} \tag{1.10}$$

where **T** is a nonsingular square matrix. Clearly, the inverse transformation gives

$$\mathbf{B} = \left(\mathbf{T}^{-1}\right)^{-1}\mathbf{A}\left(\mathbf{T}^{-1}\right) = \mathbf{T}\mathbf{A}\mathbf{T}^{-1}. \tag{1.11}$$

The form of any matrix equation is unaffected by transformation, thus the equation

$$\mathbf{AB} + \mathbf{CDE} = \mathbf{F}$$

may be transformed into

$$\mathbf{T}^{-1}\mathbf{ABT} + \mathbf{T}^{-1}\mathbf{CDET} = \mathbf{T}^{-1}\mathbf{FT}$$

which is equivalent to

$$\left(\mathbf{T}^{-1}\mathbf{AT}\right)\left(\mathbf{T}^{-1}\mathbf{BT}\right) + \left(\mathbf{T}^{-1}\mathbf{CT}\right)\left(\mathbf{T}^{-1}\mathbf{DT}\right)\left(\mathbf{T}^{-1}\mathbf{ET}\right) = \left(\mathbf{T}^{-1}\mathbf{FT}\right)$$

where the parentheses indicate the transforms of the individual matrices. This invariance of matrix equations makes it possible to work with any convenient transformation of a set of matrices without affecting the validity of any results obtained.

A matrix **M** is said to be *diagonalized* by the transformation **T** if the resulting matrix $\mathbf{T}^{-1}\mathbf{MT}$ is diagonal. We denote this diagonal matrix by $\mathbf{\Lambda}$, with elements $\lambda_i\delta_{ij}$; the elements $\lambda_i$ are called the *eigenvalues* of the matrix **M**. To find $\mathbf{\Lambda}$ explicitly, we have

$$\mathbf{T}^{-1}\mathbf{MT} = \mathbf{\Lambda}$$

or

$$\mathbf{MT} = \mathbf{T}\mathbf{\Lambda}. \tag{1.12}$$

If we write out the $ij$ element of this equation, we obtain

$$\Sigma_k M_{ik} T_{kj} = \Sigma_k T_{ik} \lambda_k \delta_{kj} = T_{ij} \lambda_j \qquad \text{for } i = 1, N \tag{1.13}$$

where $\lambda_j$ is a particular eigenvalue and the index $k$ is summed over from 1 to the rank $N$ of the matrix **M**. (1.13) defines a set of $N$ homogeneous algebraic equations for the unknown transformation matrix element $T_{ij}$, where $j$ is fixed. The necessary and sufficient condition that these equations have a solution is that the determinant of their coefficients vanish, or that the determinant of the square matrix $\left(M_{ik} - \lambda_j \delta_{ik}\right)$ be zero. This provides a single algebraic equation, called the *secular equation*, which is of order $N$ and has $N$ roots $\lambda_j$.

The matrices which occur in quantum mechanics are primarily Hermitian matrices and unitary matrices. There is a fundamental theorem which states that any Hermitian matrix can be diagonalized by a unitary transformation. This has several important corollaries, including:

1. The eigenvalues of a Hermitian matrix are unique, except perhaps for the order in which they are arranged.
2. The eigenvalues of a Hermitian matrix are real.
3. A matrix that can be diagonalized by a unitary transformation and has real eigenvalues is Hermitian.
4. The necessary and sufficient condition that two Hermitian matrices can be diagonalized by the same unitary transformation is that they commute.

## 2. The Two-State Problem

We consider the problem of diagonalizing a $2 \times 2$ matrix **M**. Suppose first that **M** is a real symmetric matrix.

**Method I**. Solution of the secular equation is a direct but cumbersome procedure to find the eigenvalues ($\lambda_1$ and $\lambda_2$) and the transformation coefficients ($T_{ij}$). The condition

$$\mathbf{T}^{-1}\mathbf{M}\mathbf{T} = \mathbf{\Lambda}$$

is equivalent to

$$\mathbf{MT} = \mathbf{T}\mathbb{1} \tag{2.1}$$

$$\begin{pmatrix} M_{11} M_{12} \\ M_{12} M_{22} \end{pmatrix} \begin{pmatrix} T_{11} T_{12} \\ T_{21} T_{22} \end{pmatrix} = \begin{pmatrix} T_{11} T_{12} \\ T_{21} T_{22} \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

This gives 2 sets of 2 equations, one for each value of $\lambda$:

$$M_{11}T_{11} + M_{12}T_{21} = T_{11}\lambda_1 \qquad (2..2a)$$
$$M_{12}T_{11} + M_{22}T_{21} = T_{21}\lambda_1$$

and

$$M_{11}T_{12} + M_{12}T_{22} = T_{12}\lambda_2 \qquad (2..2b)$$
$$M_{12}T_{12} + M_{22}T_{22} = T_{22}\lambda_2.$$

If we drop the second subscript on $T$ and that on $\lambda$, these two sets are the same, namely

$$(M_{11} - \lambda)T_1 + M_{12}T_2 = 0 \qquad (2..3)$$
$$M_{12}T_1 + (M_{22} - \lambda)T_2 = 0.$$

In order that there exist nontrivial solutions (other than $T_1 = T_2 = 0$) the determinant of the coefficients must vanish:

$$\begin{vmatrix} M_{11} - \lambda & M_{12} \\ M_{12} & M_{22} - \lambda \end{vmatrix} = 0. \qquad (2..4)$$

This is the secular equation. In expanded form, it is

$$(M_{11} - \lambda)(M_{22} - \lambda) - M_{12}^2 = 0$$

or

$$\lambda^2 - (M_{11} + M_{22})\lambda + M_{11}M_{22} - M_{12}^2 = 0. \qquad (2..5)$$

The roots are given by

$$\lambda = \frac{1}{2}\left[(M_{11} + M_{22}) \pm (M_{11} + M_{22})^2 - 4\left(M_{11}M_{22} - M_{12}^2\right)\right]^{1/2} \qquad (2..6)$$

so

$$\lambda_1 = \frac{1}{2}(M_{11} + M_{22}) + \frac{1}{2}\left[(M_{11} - M_{22})^2 + 4M_{12}^2\right]^{1/2} \qquad (2..7a)$$
$$\lambda_2 = \frac{1}{2}(M_{11} + M_{22}) - \frac{1}{2}\left[(M_{11} - M_{22})^2 + 4M_{12}\right]^{1/2} \qquad (2..7b)$$

With these values of the roots, the customary prescription calls for solving (2.2) for the ratios

$$\frac{T_{11}}{T_{21}} = \frac{M_{11}}{\lambda_1 - M_{11}} = \frac{\lambda_1 - M_{22}}{M_{12}} \qquad (2..8a)$$
$$\frac{T_{12}}{T_{22}} = \frac{M_{12}}{\lambda_2 - M_{11}} = \frac{\lambda_2 - M_{22}}{M_{12}}. \qquad (2..8b)$$

The calculation of the coefficients is completed by normalizing them according to

$$T_{11}^2 + T_{21}^2 = 1 \tag{2..9a}$$

$$T_{12}^2 + T_{22}^2 = 1. \tag{2..9b}$$

This is a rather tedious procedure even for a $2 \times 2$ problem! The reader may readily confirm that the solution has the form

$$T_{11} = \cos\theta, \qquad T_{12} = \mp \sin\theta$$

$$T_{21} = \pm \sin\theta, \qquad T_{22} = \cos\theta.$$

As indicated, there is an ambiguity in sign (which corresponds to rotation via $+\theta$ or $-\theta$). The upper sign will be used. Also, Eqs. (2.8) give

$$\tan\theta = \frac{2M_{12}}{(\lambda_1 - \lambda_2) + (M_{11} - M_{22})}. \tag{2..10}$$

Note also that, as with any quadratic equation, (2.5) can also be written in factored form as

$$(\lambda - \lambda_1)(\lambda - \lambda_2) = 0.$$

Thus, we may compare (2.5) with

$$\lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2 = 0$$

and thus establish the sum and product rules:

$$\lambda_1 + \lambda_2 = M_{11} + M_{22}$$

$$\lambda_1\lambda_2 = \begin{vmatrix} M_{11} M_{12} \\ M_{12} M_{22} \end{vmatrix}.$$

**Method II**. A real symmetric matrix can be diagonalized by an orthogonal transformation

$$\mathbf{T} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{2..11}$$

in which the rotation angle $\theta$ is chosen to eliminate the off-diagonal elements in the transformed matrix. Thus we want

$$\mathbf{T}^{-1}\mathbf{M}\mathbf{T} = \mathbf{\Lambda}$$

where $\mathbf{\Lambda}$ is a diagonal matrix with elements $\lambda_1$ and $\lambda_2$. The transformation gives

$$\begin{aligned}\mathbf{T}^{-1}\mathbf{MT} &= \begin{pmatrix} C & S \\ -S & C \end{pmatrix}\begin{pmatrix} M_{11} & M_{12} \\ M_{12} & M_{22} \end{pmatrix}\begin{pmatrix} C & -S \\ S & C \end{pmatrix} \\ &= \begin{pmatrix} CM_{11} - SM_{12} & CM_{12} + SM_{22} \\ -SM_{11} + CM_{12} & -SM_{12} + CM_{22} \end{pmatrix}\begin{pmatrix} S & -S \\ S & C \end{pmatrix} \\ &= \begin{pmatrix} C^2M_{11} + 2CS\,M_{12} + S^2M_{22} & (C^2 - S^2)M_{12} - SC(M_{11} - M_{22}) \\ (C^2 - S^2)M_{12} - SC(M_{11} - M_{22}) & S^2M_{11} - 2CS\,M_{12} + C^2M_{22} \end{pmatrix}\end{aligned}$$

where $C = \cos\theta$ and $S = \sin\theta$. The off-diagonal elements are eliminated if

$$(C^2 - S^2)M_{12} - SC(M_{11} - M_{22}) = 0, \tag{2..12}$$

that is, if $\theta$ is chosen such that

$$\tan 2\theta = \frac{2CS}{C^2 - S^2} = \frac{2M_{12}}{M_{11} - M_{22}}. \tag{2..13}$$

With this choice of $\theta$ the diagonal elements are given by

$$\lambda_1 = C^2M_{11} + 2CS\,M_{12} + S^2M_{22} \tag{2..14a}$$

$$\lambda_2 = S^2M_{11} - 2CS\,M_{12} + C^2M_{22}. \tag{2..14b}$$

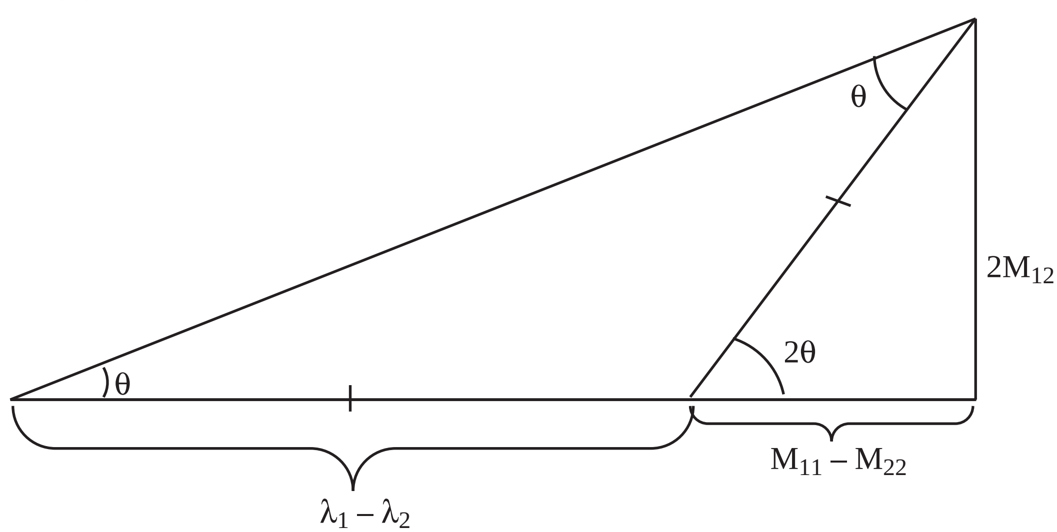
Note again that the trace of the matrix (sum of the diagonal elements),

$$\lambda_1 + \lambda_2 = M_{11} + M_{22} \tag{2..15a}$$

and the determinant of the matrix

$$\lambda_1\lambda_2 = M_{11}M_{22} - M_{12}^2 \tag{2..15b}$$

are both unchanged by the transformation. The equivalence of **Methods I** and **II** is apparent from the following geometrical construction:



**Treatment of a Hermitian Matrix**. If **M** is a Hermitian matrix, its off-diagonal elements will in general be complex. However, the diagonalization problem is easily reduced to that for a real symmetric matrix. A Hermitian matrix may be written in the form

$$\begin{pmatrix} M_{11} M_{12} \\ M_{21} M_{22} \end{pmatrix} = \begin{pmatrix} |M_{11}| & |M_{12}|e^{ia} \\ |M_{12}|e^{-ia} & |M_{22}| \end{pmatrix} \tag{2..16}$$

since $M_{12} = M_{21}^{\star}$. Also, the diagonal elements are real. The unitary transformation

$$\mathbf{A} = \begin{pmatrix} e^{ia} & 0 \\ 0 & 1 \end{pmatrix} \tag{2..17}$$

will convert **M** into a real symmetric matrix, since

$$\begin{aligned} \mathbf{A}^{-1}\mathbf{M}\mathbf{A} &= \begin{pmatrix} e^{-ia} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} |M_{11}| & |M_{12}|e^{ia} \\ |M_{12}|e^{-ia} & |M_{22}| \end{pmatrix} \begin{pmatrix} e^{ia} & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} |M_{11}|e^{-ia} & |M_{12}| \\ |M_{12}|e^{ia} & |M_{22}| \end{pmatrix} \begin{pmatrix} e^{-ia} & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} |M_{11}| & |M_{12}| \\ |M_{12}| & |M_{22}| \end{pmatrix}. \end{aligned}$$

The diagonalization is then completed as before. The complete transformation matrix is

$$\begin{aligned} \mathbf{AT} &= \begin{pmatrix} e^{ia} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} C & -S \\ S & C \end{pmatrix} \\ &= \begin{pmatrix} e^{ia}C & -e^{ia}S \\ S & C \end{pmatrix}. \end{aligned} \tag{2..18}$$

# 3. Numerical Methods for Larger Matrices

We now consider an $N \times N$ symmetric matrix.

**Method I**. The secular equation gives, when all values of $i$ and $j$ are used, $N$ sets of $N$ equations for the $T_{ij}$'s and $\lambda_j$, one set for each $\lambda_j$. If the subscript $j$ is dropped, these sets are all the same, and have the form

$$\begin{aligned} M_{11}T_1 + M_{12}T_2 + \cdots + M_{1N}T_N &= T_1\lambda \\ M_{21}T_1 + M_{22}T_2 + \cdots + M_{2N}T_N &= T_2\lambda. \end{aligned} \tag{3..1}$$

This single set of $N$ homogeneous equations in $N$ unknown $T_i$'s can be solved for the ratios of the $T_i$ elements if the determinant of the coefficients vanishes,

$$\begin{vmatrix} M_{11} - \lambda & M_{12} \ldots & M_{1N} \\ M_{21} & M_{22} - \lambda \ldots & M_{2N} \end{vmatrix} = 0. \tag{3..2}$$

This is an N[th] degree equation in $\lambda$ and the roots give the $N$ eigenvalues $\lambda_j$. For each allowed value of $\lambda_j$ obtained from (3.2), we may substitute into (3.1) and solve for the $N-1$ ratios $T_{ij}\big/T_{1j}$, and then complete the determination of the $T_{ij}$ by one of the normalization requirements,

$$\sum_i T_{ij}^2 = 1$$

or

$$T_{ij} = \left(\sum_{i=2}^{N}\left[1+\left(T_{ij}/T_{1j}\right)^2\right]\right)^{-1/2} . \qquad (3..3)$$

This procedure is clumsy but useful for small matrices.

**Method II**. Other methods are more effective with large matrices and electronic computers. Consider the effect of the following transformation:

$$T = \begin{array}{c} \\ \\ \\ \text{row } m \\ \\ \\ \\ \text{row } n \\ \\ \\ \\ \end{array} \left(\begin{array}{ccccccccccc} 1 & & & \text{column } m & & & & \text{column } n & & & \\ & 1 & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & C & & & & -S & & & \\ & & & & 1 & & & & & & \\ & & & & & 1 & & & & & \\ & & & & & & 1 & & & & \\ & & & S & & & & C & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & 1 & \\ & & & & & & & & & & 1 \end{array}\right) \qquad (3..4)$$

in which $T_{ij} = \delta_{ij}$ if neither $i$ nor $j = n$ or $m$ and $T_{mm} = T_{nn} - \cos\theta$, $T_{mn} = -T_{nm} = \sin\theta$. The elements of the transformed matrix,

$$\mathbf{M}' = \mathbf{T}'\mathbf{M}\mathbf{T}$$

are given by

$$M'_{ij} = \Sigma_{k1} T_{ki} M_{k1} T_{1j}. \qquad (3..5)$$

(a) If both $i$ and $j$ differ from $m$ and $n$,

$$M'_{ij} = \Sigma_{k1}\delta_{ki} M_{k1}\delta_{1j} = M_{ij} \qquad (3..6)$$

and these elements are unchanged by the transformation.

(b) If $i = m$ and $j \neq m$ or $n$,

$$M'_{mj} = \Sigma_{k1} T_{km} M_{k1} \delta_{1j} = \Sigma_k T_{km} M_{kj}.$$

The index $k$ runs over all values but $T_{km} = 0$ except for $k = m$ and $k = n$, so

$$M'_{mj} = \cos\theta M_{mj} + \sin\theta M_{nj}.$$

Similarly, for $i = n$ and $j \neq m$ or $n$,

$$M'_{nj} = -\sin\theta M_{mj} + \cos\theta M_{nj}.$$

Since the transformed matrix is symmetric, we have

$$M'_{jm} = M'_{mj} = CM_{mj} + S\,M_{nj} \tag{3..7a}$$

$$M'_{jn} = M'_{nj} = -S\,M_{mj} + CM_{nj}. \tag{3..7b}$$

(c) The elements $M_{mm}$, $M_{nn}$, and $M_{mn}$ transform exactly as they did for a $2 \times 2$ matrix, and

$$M'_{mm} = C^2 M_{mm} + 2CS\,M_{mn} + S^2 M_{nn} \tag{3..8a}$$

$$M'_{nn} = S^2 M_{mm} - 2CS\,M_{mn} + C^2 M_{nn} \tag{3..8b}$$

$$M'_{mn} = (C^2 - S^2) M_{mn} - CS\,(M_{mm} - M_{nn}) \tag{3.9}$$

These simple relations are the basis of two methods which diagonalize a large matrix by use of successive $2 \times 2$ rotations.

**Method IIA**. In one procedure, we select the largest nondiagonal element $M_{ij}$ of the parent matrix, set $i = m$, $j = n$, and use (3.9) to make $M'_{mn} = 0$, with

$$\tan 2\theta = \frac{2M_{mn}}{M_{mm} - M_{nn}}. \tag{3..10}$$

The remainder of the parent matrix is then transformed using (3.6), (3.7), and (3.8). Again the largest nondiagonal element is selected, and the procedure repeated. The first $M_{ij}$ which was originally eliminated might not become different from zero, but if so it will be much smaller than before. The process is continued until all nondiagonal elements become close enough to zero to be neglected.

This method gives rapid convergence and provides quite accurate results for the $T_{ij}$ elements. The final result for the **T**–matrix is obtained by multiplying together the series of $2 \times 2$ rotations, i.e.,

$$\mathbf{T} = \mathbf{T}_1 \mathbf{T}_2 \mathbf{T}_3 \ldots$$

**Method IIB**. Another procedure involves making the $M'_{m-1,n}$ element vanish by use of (3.7b), which requires

$$M'_{m-1,n} = 0 = CM_{n,m-1}S\,M_{m,m-1}$$

or

$$\tan\theta = \frac{M_{m-1,n}}{M_{m-1,m}}. \tag{3..11}$$

In practice, for an $N \times N$ matrix, we would proceed by eliminating the $M'_{1N}$ element (by a rotation with $m = 2$ and $n = N$) and then the $M'_{1,N-1}$ element (by a rotation with $m = 2$ and $n = N - 1$), etc. With this method, once an element is made zero it is never changed by a later rotation, and $m$ and $n$ are systematically varied until all possible elements are eliminated. Usually the first row is treated and then the second row, etc. However, it is impossible to eliminate the elements adjacent to the diagonal, such as $M_{12}$, etc., as this would require $m$ and $n$ to be the same (e.g. both = 2 for the $M_{12}$ case), which would be no rotation. Hence so far this method reduces the matrix to the form

$$\begin{pmatrix} M_{11} & M_{12} & & \mathbf{0} \\ M_{12} & M_{22} & M_{23} & \\ \mathbf{0} & M_{23} & M_{33} & \\ & & & \ddots \end{pmatrix}$$

with zero everywhere except the diagonal and adjacent to it. This is called a *tridiagonal* matrix.

The secular equation for a tridiagonal matrix is relatively easy to solve. Several efficient methods are available. One of the best is due to Givens [see *J. Assoc. Computing Machinery* **4**, 298 (1957)]. His method gives very accurate eigenvalues and is very fast, but unfortunately does not give good transformation matrices. Continued fraction methods are very convenient, if there are no difficulties from near degeneracies. For example, for a $5 \times 5$ tridiagonal matrix, we have

$$(M_{11} - \lambda)T_1 + M_{12}T_2 = 0$$

$$M_{12}T_1 + (M_{22} - \lambda)T_2 + M_{23}T_3 = 0$$

$$M_{23}T_2 + (M_{33} - \lambda)T_3 + M_{34}T_4 = 0$$

$$M_{34}T_3 + (M_{44} - \lambda)T_4 + M_{45}T_5 = 0$$

$$M_{45}T_4 + (M_{55} - \lambda)T_5 = 0.$$

The first equation may be rewritten as

$$T_1/T_2 = M_{12}/(M_{11} - \lambda),$$

the second as

$$M_{12}(T_1/T_2) + M_{22} - \lambda = -M_{23}(T_3/T_2).$$

Combining these we have

$$\frac{T_2}{T_3} = \frac{-M_{23}}{M_{22} - \lambda + M_{12}(T_1/T_2)} = \frac{-M_{23}}{M_{22} - \lambda - \frac{M_{12}^2}{M_{11}-\lambda}}.$$

Similarly, the third equation gives

$$M_{23}(T_2/T_3) + M_{33} - \lambda = -M_{34}(T_4/T_3)$$

or

$$\frac{T_3}{T_4} = \frac{-M_{34}}{M_{33} - \lambda - \frac{M_{23}^2}{M_{22}-\lambda-\frac{M_{12}^2}{M_{11}-\lambda}}}.$$

The fourth equation gives an analogous expansion for $T_4/T_5$ and the fifth gives

$$\lambda = M_{55} + M_{45}\,(T_4/T_5)\,.$$

Thus, finally we have

$$\lambda = M_{55} - \frac{M_{45}^2}{M_{44} - \lambda - \frac{M_{34}^2}{M_{33}-\lambda-\frac{M_{23}^2}{M_{22}-\lambda-\frac{M_{12}^2}{M_{11}-\lambda}}}}.$$

The trial $\lambda$ is selected and substituted in the right hand side, generating a new value of $\lambda$ on the left. This new value is used as a second approximation on the right and the process repeated until the trial $\lambda$ and the final $\lambda$ agree to within the desired accuracy. In the absence of degeneracies this procedure converges very rapidly. It also provides the ratios $T_i/T_{i+1}$ at each stage. For detailed discussion of this continued fraction method, see Swalen and Pierce, *J. Math. Phys.* **2**, 736 (1961) and Pierce, *ibid.*, 740 (1961). See J. H. Wilkinson, The Algebraic Eigenvalue Problem (Oxford, 1965).