The equation \((A-\lambda B)x = 0\): We consider this generalised eigenvalue problem when \(A\) and \(B\) are symmetric matrices and in addition \(B\) is positive definite. (The Hermitian case is identical.) Throughout \(<,>\) will denote the usual inner product on \(\mathbb{R}^n\) or \(\mathbb{C}^n\; the context will make clear which is being used. We define a new inner product on \(\mathbb{R}^n\) (or \(\mathbb{C}^n\)) by \((X,Y) := <BX,Y>\).

Lemma 2: The eigenvalues of \((A-\lambda B)x = 0\) are all real.

Proof: If the eigenvector \(X \in \mathbb{R}^n\) has eigenvalue \(\lambda \in \mathbb{R}\), then 
\[
\lambda <X,Y> = <AX,X> = <X,AX> = <X,\lambda BX> = \lambda <X,X>
\]
Thus \(\lambda = \frac{1}{\bar{\lambda}}\) so that \(\lambda \in \mathbb{R}\) and in particular \(X \in \mathbb{R}^n\).

Theorem 3: If \(A\) and \(B\) are symmetric \(n \times n\) matrices with \(B\) positive definite, then there exists a basis of eigenvectors \(X_1, \ldots, X_n\) of the equation \((A-\lambda B)x = 0\) which are orthonormal with respect to \(B\) (i.e. \(<BX_i, X_j> = \delta_{ij}\))

Proof: We remark that since \(B\) is positive definite \(B^{-1}\) exists, so we can define \(A^{-1} = B^{-1}A\) and observe

(i) \(X_1\) is an eigenvector of \((A-\lambda B)x = 0\) with eigenvalue \(\lambda_1\) if and only if it is an eigenvector of \((A - \lambda I)x = 0\) with eigenvalue \(\lambda_1\).

(ii) \(A_1\) is symmetric with respect to \((, )\) because 
\[
<A_1X, Y> = <BA_1X, Y> = <AX, Y> = <X, AY> = <X, B_1^tAY> = <BX, A_1^tY> = <X, A^tY>
\]

(iii) By lemma 2 there exists an eigenvector \(X_1\) of \(A^t\) which we may assume to have unit length with respect to \((, )\). If \([X_1]\) denotes the subspace spanned by \(X_1\), then \([X_1]\) is invariant under \(A^t\) and therefore (by lemma 1) \([X_1]\) is also invariant under \(A^t\).

(iv) The argument is now completed by induction with \(A^t\) restricted to \([X_1]\).
**Corollary 4:** If A and B are as in theorem 3 then there exists a matrix Q such that

\[
Q^tAQ = \begin{pmatrix}
\lambda_1 & & \\
& \lambda_2 & \\
& & \lambda_n
\end{pmatrix}
\]

and \(Q^tBQ = I\).

**Proof:** Let \(Q = [X_1, X_2, \ldots, X_n]\) the matrix whose \(i\)th column is the \(i\)th eigenvector (as in theorem 3) with eigenvalue \(\lambda_i\).

**Remark:**

1. If \(B = I\), then corollary 4 is the usual statement that every symmetric matrix can be orthogonally diagonalized. That symmetry is necessary here is obvious since \(Q^tAQ = 0\) where 0 is diagonal and \(Q^tQ = I\) imply \(A = QDQ^t = (QDQ^t)^t = A^t\).

2. While symmetry is used to show that the eigenvalues are real it is not the key point. Indeed, \(A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\) has real distinct eigenvalues but cannot be orthogonally diagonalized. Of course we see that \(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\) is an eigenvector whose orthogonal complement, i.e. the line \(\{t[0]\} : t \in \mathbb{R}\) is not invariant under \(A\).

**Theorem 5:** If \(A\) is an \(n \times n\) skew-symmetric matrix, then there exists an orthonormal basis for \(\mathbb{R}^n\) with respect to which \(A\) is tridiagonal. That is there exists \(Q\) satisfying \(Q^tQ = I\) and

\[
Q^tAQ = \begin{pmatrix}
0 & -b_3 & \\
-1 & 0 & -b_4 \\
& -1 & 0 \\
& & -1 & -b_5 \\
& & & 0 & -b_6 \\
& & & & -1 & 0 \\
& & & & & & 0
\end{pmatrix}
\]

**Proof:** By the argument of lemma 2 one sees that all the eigenvalues of \(A\) are pure imaginary. If \(m(t)\) denotes the minimal polynomial for \(A\) (over \(\mathbb{R}\)) and if \(ib, b \neq 0\), is an eigenvalue of \(A\) with eigenvector \(X\), then \(0 = m(A)X = m(ib)X\) implies \(ib\) is a root of \(m(t) = 0\). Accordingly \(m(t) = q(t)(t^2 + b^2)\). Therefore, since \(m(t)\) is minimal, there exists a unit vector \(X_1 \in \mathbb{R}^n\) such that \((A^2 + b^2)X_1 = 0\). If we define \(X_2 = (AX_1)/b\) then

(i) \(AX_1 = bX_2\) and \(AX_2 = (A^2X_1)/b = bX_1\).

(ii) \(<X_2, X_2> = <(AX_1)/b, (AX_1)/b> = <X_1, X_1> = b^2 = <X_1, X_1> = 1\).

(iii) \(<X_1, X_2> = <X_1, (AX_1)/b> = -<(AX_1)/b, X_1> = -<X_2, X_2> = 0\).

(iv) The subspace \([X_1, X_2]\) spanned by \(X_1\) and \(X_2\) is invariant under \(A\) and therefore (by lemma 1) is also its orthogonal complement \([X_1, X_2]^\perp\). We now continue by induction on \(A\) restricted to \([X_1, X_2]^\perp\).

(v) The case of zero eigenvalues is easily taken care of and it is clear that the basis produced is the required one with \(Q\) being the change of basis matrix.

**School of Mathematical Sciences, N.I.M.C., Dublin.**

From 2-Manifold, No. 1