B

LINEAR VECTOR SPACES, STURM-LIOUVILLE PROBLEMS, AND EIGENFUNCTION EXPANSIONS

The strength of a tree is determined by the strength of the roots

African proverb

This appendix is a brief summary of concepts in abstract linear vector spaces, and application to orthogonal function expansions associated with the Sturm-Liouville equation. For rigorous definitions of terms and justifications of the critical assertions below, refer to any text on mathematical methods in physics and engineering; two readable ones are the books by Friedman [1] and Dettman [2].

B.1 SPACES OF FINITE DIMENSION

B.1.1 Basis vectors and Scalar Products

The reader is probably familiar with N-dimensional Euclidian vector spaces, the three-dimensional Cartesian system being the most obvious example. Any point in such a space can be represented by a set of N real numbers, collectively called a vector $\overline{x}$,

$$\overline{x} = (\xi_1, \xi_2, \ldots, \xi_N) \tag{B.1}$$

where the scalars ξ_i are the "components" of the vector. In a more general (non-Euclidean) space, these numbers could be complex. A *basis set* for an N-dimensional vector space is a set of N *linearly independent* vectors $\overline{x}_1, \overline{x}_2 \ldots \overline{x}_3$ which permit any arbitrary vector in the space to be represented as a linear combination of the basis vectors

$$\overline{A} = a_1\overline{x}_1 + a_2\overline{x}_2 + \ldots a_N\overline{x}_N = \sum_{i=1}^{N} a_i\overline{x}_i \tag{B.2}$$

where the scalars a_i define the "representation" of $\overline{A}$ in this basis set. A basis set is therefore said to *span* the vector space if all vectors in the space can be represented this way. A simple example of a basis set is

$$\overline{x}_1 = (1, 0, \ldots 0), \quad \overline{x}_2 = (0, 1, \ldots 0), \quad \ldots \quad \overline{x}_N = (0, 0, \ldots 1) \tag{B.3}$$

which is analogous to the Cartesian basis in three dimensions.

Linear independence of the basis vectors implies that none can be represented as a linear combination of the others in the set. This in turn implies that the representation of $\overline{A}$ in (B.2) is unique [2]; that is, there is only one possible set of scalars a_i which represents $\overline{A}$ for a given basis set.

There are many possible choices for basis sets in a given vector space. However, we usually try to find basis sets for which the basis vectors are mutually *orthogonal*, for reasons which will be clear shortly. In general vector spaces, orthogonality is defined in terms of a *scalar product*. The scalar product of two vectors, $x = (\xi_1, \xi_2, \ldots \xi_N)$ and $y = (\eta_1, \eta_2, \ldots \eta_N)$, is a single number denoted by $\langle x|y \rangle$, and is usually defined as

$$\langle x|y \rangle = \xi_1\eta_1 + \xi_2\eta_2 + \ldots \xi_N\eta_N = \sum_{i=1}^{N} \xi_i\eta_i \tag{B.4}$$

This is also sometimes called an *inner product*. Note that we have dropped the bar convention to denote vectors; this will be done when it is clear from the context that the quantities involved are vectors. This definition of the scalar product is symmetrical with respect to the two vectors and satisfies

$$\langle x|y \rangle = \langle y|x \rangle \qquad \langle \alpha x|\beta y \rangle = \alpha\beta\langle x|y \rangle \tag{B.5}$$

where α and β are scalars. Two vectors are orthogonal if their scalar product is zero,

$$\langle x|y \rangle = 0 \tag{B.6}$$

which implies that the vectors are, in some geometric sense, perpendicular to each other. The definition of scalar product is also related to the concept of the *length* of a vector, which is taken as

$$|x| = \sqrt{\langle x|x \rangle} \tag{B.7}$$

These definitions are generalizations of the familiar results from real three-dimensional spaces (there the scalar product is written as a "dot" product, $\overline{x} \cdot \overline{y}$). However, for general vector spaces, it is important to recognize that the definition (B.4) is somewhat arbitrary, and does not always lead to a useful definition of "length". When the vector space includes complex numbers, the length as defined by (B.7) could be negative, zero, or complex. In order to preserve the length as a real, nonzero quantity for complex vector spaces (there are usually compelling physical reasons for doing so), the scalar product is re-defined as

$$\langle x|y \rangle = \xi_1^*\eta_1 + \xi_2^*\eta_2 + \ldots \xi_N^*\eta_N = \sum_{i=1}^{N} \xi_i^*\eta_i \tag{B.8}$$

where the asterisk * denotes the complex conjugate. The length defined by (B.7) will then be a real number, often called the vector *norm* by mathematicians. Note that in general the scalar product (B.8) is not symmetrical with respect to interchange of vectors as it is for real spaces, but rather

$$\langle x|y \rangle = \langle y|x \rangle^* \qquad \langle \alpha x|\beta y \rangle = \alpha^*\beta\langle x|y \rangle \tag{B.9}$$

where α and β are complex scalars.

Returning to the representation of an arbitrary vector in terms of the basis set (B.2), we can easily see why orthogonality is a desirable property. If the basis vectors are orthogonal, then $\langle x_i|x_j\rangle = 0$ for $i \neq j$. Forming the scalar product $\langle x_i|A\rangle$ we find that

$$a_i = \frac{\langle x_i|A\rangle}{\langle x_i|x_i\rangle} \tag{B.10}$$

If the basis vectors are *normalized* so that $\langle x_i|x_i\rangle = 1$ for all i, then the basis set is said to be *orthonormal*, and

$$a_i = \langle x_i|A\rangle \tag{B.11}$$

In an orthonormal basis, the scalars a_i can be interpreted as the "projection" of the vector $\overline{A}$ in the direction of the basis vector $\overline{x}_i$. The condition for orthonormality of the basis set is summarized concisely as

$$\langle x_i|x_j\rangle = \delta_{ij} \qquad \text{where} \quad \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & \text{otherwise} \end{cases} \tag{B.12}$$

δ_{ij} is called the *Kronecker delta* function.

B.1.2 Matrices and Eigenvalues

In addition to the scalar product we must also consider linear operations which transform one vector into another vector. A simple example is $y = \alpha x$, where α is a scalar. In this case, each component of y is just the corresponding component of x multiplied by α. This scaling operation thus preserves the original direction of x. A more general operation would involve a rotation as well, where the result y no longer points in the direction of x. This is represented by a matrix equation $y = Ax$, where A is a matrix operator. The reader is assumed to be familiar with basic matrix operations. If x and y are $(N \times 1)$ column vectors, then A is an $(N \times N)$ matrix, and $y = Ax$ represents

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1N} \\ a_{21} & a_{22} & \ldots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \ldots & a_{NN} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{pmatrix} \tag{B.13}$$

or alternatively

$$y_i = \sum_{j=1}^{N} a_{ij} x_j \tag{B.14}$$

where a_{ij} are the components of the matrix A. Note that we can choose to represent any vector as either a row or column vector. Physically, the two representations are the same as along as the components are identical. However, the ordering of matrix operations will dictate a representation; for example, if we were to write $y = xB$, where B is an $N \times N$ matrix, the vector x must be represented as a $1 \times N$ row vector for the operation to make sense.

The matrix equation defines a set of simultaneous linear equations. Typically the matrix A and the vector y are known, and we wish to solve for the vector x. This requires finding the inverse matrix A^{-1}, which is defined by

$$AA^{-1} = A^{-1}A = I \tag{B.15}$$

where I is the *identity matrix* or *identity operator*, given by

$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix} \qquad AI = IA = A \tag{B.16}$$

The vector x is then given by

$$x = A^{-1}y \tag{B.17}$$

The solution for x exists if the matrix has an inverse. The condition that for the matrix to have an inverse is that the determinant of the matrix be non-zero, $|A| \neq 0$. A matrix with a zero determinant, $|A| = 0$, is called a *singular* matrix. Some other matrix definitions are

- *Matrix transpose*: The transpose of a matrix A is denoted by A^T and obtained by interchanging rows and columns. The elements of the transpose matrix are $a^T_{ij} = a_{ji}$. The transpose of the product of two vectors is $(AB)^T = B^T A^T$.
- *Symmetrical matrix*: When $A = A^T$, or $a_{ij} = a_{ji}$, the matrix is symmetrical.
- *Hermitian matrix*: When $A = (A^*)^T$, or $a_{ij} = a^*_{ji}$, the matrix is Hermitian. If all the elements are real, a Hermitian matrix is symmetrical.

In general, matrix operations are not symmetrical with respect to the ordering of the operations; that is, Ax and xA do not produce the same vector in general, $Ax \neq xA$. This is *not* because Ax produces a column vector and xA produces a row vector; physically, as we noted earlier, there is no difference between a vector and its transpose. We can express this equivalence as $x \Leftrightarrow x^T$ or $Ax \Leftrightarrow (Ax)^T$. According to the definitions above, if Ax and $(Ax)^T$ have the same components, $Ax \Leftrightarrow xA^T$. Therefore, Ax and xA are the same vector only when the matrix is symmetrical, $A = A^T$.

The matrix *eigenvalue* equation is especially important in physics and is defined by

$$Ax = \lambda x \tag{B.18}$$

where λ is a scalar. Values of λ for which this equation is satisfied are called the eigenvalues of the matrix A. The vectors x corresponding to each eigenvalue are called the eigenvectors. The eigenvalue equation can be rewritten as

$$(A - \lambda I)x = 0 \tag{B.19}$$

using the identity matrix. If the matrix $A - \lambda I$ is not singular, then from (B.17) we see that the only solution for this equation is the trivial one $x = 0$. A non-trivial solution will only exist if the determinant is zero; that is, $|A - \lambda I| = 0$, or

$$\begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} - \lambda & \dots & a_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & \dots & a_{NN} - \lambda \end{vmatrix} = 0 \tag{B.20}$$

This defines the *characteristic equation* for the eigenvalues. The set of eigenvalues for a matrix A are called the *spectrum* of A. The set of eigenvectors for which $\lambda = 0$ define the *null space* of A. Numerical techniques for rapidly evaluating eigenvalues and eigenvectors are readily available in most mathematics packages like Mathematica or MATLAB; source code for such routines can be found in the super book *Numerical Recipes* by Press *et al.* [3].

B.1.3 Symmetric and Hermitian Matrices

Earlier the advantage of using an orthonormal basis set for representing vectors was demonstrated, but no mention was made of how to actually construct such an orthonormal basis. For problems involving symmetric or Hermitian matrix operators, it can be shown that the eigenvectors of the matrix form a very useful orthogonal basis, and that the eigenvalues of such matrices are all real. Before showing these properties, we introduce the matrix *adjoint* operator. The adjoint of a matrix A will be denoted by $A^\dagger$, and is defined by the relation

$$\langle y|Ax\rangle = \langle A^\dagger y|x\rangle \tag{B.21}$$

A matrix operator is said to be *self-adjoint* when $A = A^\dagger$. It is important to recognize that the adjoint operator depends on the definition of the scalar product. For example, using (B.8) we find

$$\langle y|Ax\rangle = \sum_k y_k^* \sum_j a_{kj} x_j = \sum_j \left(\sum_k a_{kj}^* y_k \right)^* x_j = \langle (A^*)^T y|x\rangle \tag{B.22}$$

so for this scalar product, the adjoint is given by the complex conjugate transpose of the matrix, $A^\dagger = (A^*)^T$. For real vector spaces using the scalar product (B.4), a similar analysis shows $A^\dagger = A^T$. Therefore, from the definitions given earlier, we see that Hermitian matrix operators are self-adjoint operators in a complex scalar-product space, whereas symmetric matrix operators are self-adjoint using the real vector product (B.4).

We are now in a position to motivate the assertions mentioned earlier, which are restated as follows:

- A self-adjoint operator in either a real or complex vector space has purely real eigenvalues
- The eigenvectors of a self-adjoint operator in either a real or complex vector space provide an orthogonal basis for the space.

The first property will be shown for real vector spaces using (B.5). It is assumed that in a real vector space the matrix operator is real (all components are real numbers), but the eigenvalues and eigenvectors may not necessarily be real. From (B.18) we can therefore write $Ax = \lambda x$ and $Ax^* = \lambda^* x^*$. Using the former we find $\langle x^*|Ax\rangle = \lambda |x|^2$. But from the self-adjoint property, $\langle x*|Ax\rangle = \langle Ax^*|x\rangle = \lambda^* |x|^2$. Therefore, $\lambda = \lambda^*$; the eigenvalues are real. The proof for Hermitian operators in complex spaces is similar.

For two different eigenvectors and eigenvalues, labeled by subscripts i and j, we write $Ax_i = \lambda_i x_i$ and $Ax_j = \lambda_j x_j$. From the self-adjoint property and the properties of the real scalar product (B.5),

$$\langle x_j | Ax_i \rangle - \langle Ax_j | x_i \rangle = (\lambda_i - \lambda_j)\langle x_i | x_j \rangle = 0 \tag{B.23}$$

so as long as $\lambda_i \neq \lambda_j$, the eigenvectors are orthogonal. A similar proof holds for Hermitian matrices in complex vector spaces. It can be shown [2] that eigenvectors belonging to the same eigenvalue (a *degenerate* eigenvalue) can be orthogonalized using a process called the *Gram-Schmidt orthogonalization procedure* whereby new orthogonal vectors are constructed from linear combinations of the degenerate eigenvectors.

All of the operators we will deal with will be self-adjoint (or we will contrive to make them so), and therefore the eigenvectors can always be used to represent arbitrary vectors in the space. This choice of basis is often referred to as the one which *diagonalizes* the matrix, for the following reason. Consider the matrix equation $y = Ax$. Assuming we have an orthonormal basis set x_i where $\langle x_i | x_j \rangle = \delta_{ij}$ we can represent the vectors x and y as

$$x = \sum_j \alpha_j x_j \qquad\qquad y = \sum_k \beta_k x_k \tag{B.24}$$

Since $y = Ax$ we have

$$\sum_k \beta_k x_k = \sum_j \alpha_j A x_j \tag{B.25}$$

Forming the scalar product of both sides with x_i and using the orthonormal property gives

$$\beta_i = \sum_j \langle x_i | Ax_j \rangle \alpha_j \tag{B.26}$$

This has the form of a matrix equation $\beta = L\alpha$, where the vectors α and β are representations of the original vectors x and y in the chosen basis, and L is a representation of the original matrix operator with elements $L_{ij} = \langle x_i | Ax_j \rangle$. Clearly if the basis set corresponds to the eigenvectors of A, the matrix L will be diagonal, with the diagonal elements corresponding to the eigenvalues of A.

B.2 SPACES OF INFINITE DIMENSION; HILBERT SPACES

The preceding vector concepts can be generalized to infinite sequences of numbers in such a way that arbitrary functions can be viewed as vectors. For example, the function $\sin(x)$ defined over the range $0 \leq x \leq \pi$ may represent a vector where each component of the vector is the value of the function at some point in the interval. The motivation for this viewpoint stems from the ability to represent arbitrary vectors in terms of a suitably chosen basis set, which in turn simplifies the solution of certain linear equations. The theory of Fourier series, as well as other so-called "special functions" of mathematical physics—Bessel functions, spherical harmonics, Legendre and Hermite polynomials, etc.—are all treated in a unified way using concepts of abstract linear vector spaces.

A linear operator in abstract vector spaces is a generalization of the matrix concept. A linear operator $\mathcal{L}$ has the property that it maps a vector $|x\rangle$, defined in some space S, into another vector $\mathcal{L}|x\rangle$ in the same space. The set of vectors $|x\rangle$ for which the mapping is defined is called the *domain* of $\mathcal{L}$. The set of vectors $\mathcal{L}|x\rangle$ is called the *range* of $\mathcal{L}$. For example, consider the differential operator $\mathcal{L} = d/dx$ defined over the interval $0 \le x \le 1$, and take the space to be all real numbers. The domain of $\mathcal{L}$ in this case is the set of all real functions of x that have well-defined first derivatives on this interval, and the range of $\mathcal{L}$ is the set of values corresponding to these first derivatives.

The *adjoint* of a linear operator $\mathcal{L}$ is denoted as $\mathcal{L}^*$, and is defined by the following relation

$$\langle y|\mathcal{L}\, x\rangle = \langle \mathcal{L}^* y|x\rangle \tag{B.27}$$

which is an obvious generalization of (B.21). An operator is *self-adjoint* if $\mathcal{L} = \mathcal{L}^*$. Clearly the definition of the scalar product will influence whether or not an operator is self-adjoint. The scalar product in *real* vector spaces is commonly taken as

$$\langle f|g\rangle = \int_a^b f(t)g(t)\,dt \tag{B.28}$$

where the limits of integration are taken over the domain of the functions f and g. This is a straightforward generalization of the earlier N-dimensional scalar products, where the discrete summation becomes an integral. Naturally, this is not the only possible choice of scalar product; when the functions are complex,

$$\langle f|g\rangle = \int_a^b f^*(t)g(t)\,dt \tag{B.29}$$

is a better choice. Note that (B.28) and (B.29) satisfy the original properties (B.5) and (B.9), respectively. In both cases, a vector or function can only have a finite "length" $\sqrt{\langle f|f\rangle}$ if the functions are "square integrable", which is sometimes referred to as *Lebesque integrable* by mathematicians. The space of Lebesque integrable functions with a complex scalar product is called a *Hilbert space*, and is fundamentally important in most branches of physics. Note that Hilbert space contains the set of real functions that are square-integrable with a real scalar product as a special case.

B.2.1 Orthogonal Function Expansions and Completeness

Like the finite-dimensional spaces, we define an orthonormal basis for an abstract space as a sequence of functions or vectors x_i which satisfy

$$\langle x_i|x_j\rangle = \delta_{ij} \tag{B.30}$$

Assuming that we can find such a basis, the next question is: under what conditions can an arbitrary vector or function in the abstract space be represented by a linear combination of the vectors x_i? In N-dimensional space, an arbitrary vector could be represented *exactly* in terms of N different, linearly independent vectors. We might guess that in an infinite-dimensional space, an infinite number of basis functions are required. Consider approximating the function $f(t)$ over the range $a \le t \le b$ by a sequence of the orthonormal functions,

$$f(t) = \sum_{i=1}^{N} \alpha_i x_i(t) \tag{B.31}$$

We would like to choose the expansion coefficients so that the sum approximates the function as closely as possible. The "best" choice is usually defined in a least mean-square-error sense, where the α_i are chosen so as to minimize the positive quantity

$$\int_a^b \left| f(t) - \sum_{i=1}^{N} \alpha_i x_i(t) \right|^2 dt \ge 0 \tag{B.32}$$

Expanding the left hand side and using the orthogonality property (B.30) gives

$$\int_a^b |f|^2 dt - \sum_{i=1}^{N} \alpha_i \langle x_i|f\rangle - \sum_{i=1}^{N} \alpha_i^* \langle x_i|f\rangle^* + \sum_{i=1}^{N} |\alpha_i|^2 \ge 0 \tag{B.33}$$

Adding and subtracting $\sum_i \langle x_i|f\rangle^2$ gives

$$\int_a^b |f|^2 dt + \sum_{i=1}^{N} |\alpha_i - \langle x_i|f\rangle|^2 \; dt - \sum_{i=1}^{N} \langle x_i|f\rangle^2 \ge 0 \tag{B.34}$$

The mean-square error is therefore minimized by choosing

$$\alpha_i = \langle x_i|f\rangle \tag{B.35}$$

using the scalar product (B.29). This is a generalization of the finite dimensional result (B.11). Using this result, the mean-square error is

$$\int_a^b |f|^2 \, dt - \sum_{i=1}^{N} |\alpha_i|^2 \ge 0 \tag{B.36}$$

A basis is said to be complete when this error is identically zero in the limit of $N \to \infty$. That is, an orthonormal basis $x_i(t)$ is *complete* when the coefficients $\alpha_i = \langle x_i|f\rangle$ can be used to represent any square-integrable function $f(t)$ in the sense of

$$\lim_{N\to\infty} \int_a^b \left| f(t) - \sum_{i=1}^{N} \alpha_i x_i(t) \right| dt = 0 \tag{B.37}$$

Note that this does not imply that the series $\sum_i \alpha_i x_i$ will represent $f(t)$ *exactly* at every point on the interval $a \le t \le b$. When $f(t)$ is discontinuous, the series will converge to the *mean* value of $f(t)$ at the discontinuity.

It has been shown [4] that all the orthonormal basis functions corresponding to physical operators that are typically encountered in physics are complete in the sense of (B.37). When the set is complete, (B.36) gives *Parseval's theorem*

$$\int_a^b |f(t)|^2 \, dt = \sum_{i=1}^{\infty} |\alpha_i|^2 \tag{B.38}$$

Inserting (B.35) into (B.31) and using the scalar product (B.29), we find

$$f(t) = \sum_{i=1}^{\infty} x_i(t) \int_a^b x_i^*(t') f(t')\, dt' = \int_a^b \left[\sum_{i=1}^{\infty} x_i(t) x_i^*(t') \right] f(t')\, dt' \tag{B.39}$$

The quantity in brackets can be recognized as the Dirac delta function

$$\sum_{i=1}^{\infty} x_i(t) x_i^*(t') = \delta(t - t') \tag{B.40}$$

This is called the *completeness relation*. Note that the arguments of the two functions in the sum are different.

B.2.2 Self-Adjoint Operators

When the operator is self-adjoint, the eigenvectors form an orthonormal set, which is the same result that we found for finite-dimensional spaces. The proof is correspondingly similar. The eigenvectors x_i and eigenvalues λ_i are defined by

$$\mathcal{L}\, x_i = \lambda_i x_i \tag{B.41}$$

Using the definition of the adjoint gives

$$\langle x_i | \mathcal{L}\, x_j \rangle - \langle \mathcal{L}\, x_i | x_j \rangle = 0 = (\lambda_j - \lambda_i) \langle x_i | x_j \rangle \tag{B.42}$$

where the real scalar product (B.28) was assumed. As long as $\lambda_i \neq \lambda_j$, eigenvectors are mutually orthogonal, $\langle x_i | x_j \rangle = 0$. When the eigenvalues belonging to different eigenvectors are the same (degenerate), then the eigenvectors may still be orthogonal, but if not, they can be made so using the Gram-Schmidt procedure as in the finite-dimensional case. Therefore we can construct an orthonormal basis from the eigenvectors so that

$$\langle x_i | x_j \rangle = \delta_{ij} \tag{B.43}$$

As mentioned in the previous section, this basis will always be complete for the operators we will encounter.

We will be interested in solving equations of the form $\mathcal{L}\, f = a$, where we are given a vector a, and wish to solve for the unknown vector f. Using the eigenvectors of $\mathcal{L}$ as a basis, we can expand f as

$$f = \sum_i \alpha_i x_i \qquad \alpha_i = \langle x_i | f \rangle \tag{B.44}$$

Using the self-adjoint property of $\mathcal{L}$ we write

$$\mathcal{L}\, f = a \quad \Rightarrow \quad \langle x_i | \mathcal{L}\, f \rangle = \langle \mathcal{L}\, x_i | f \rangle = \lambda_i \langle x_i | f \rangle = \langle x_i | a \rangle \tag{B.45}$$

where we have assumed the scalar product (B.28). Therefore, combining the last two equations, the unknown function is given by

$$f = \sum_i \frac{\langle x_i | a \rangle}{\lambda_i} x_i \tag{B.46}$$

which is a useful result.

Any operator can be represented by a matrix in the same sense as (B.26). Alternatively, we can write

$$\mathcal{L} f = \sum_n c_n x_n \tag{B.47}$$

Using the self-adjoint property and (B.41),

$$c_n = \langle x_n | \mathcal{L} f \rangle = \langle \mathcal{L}\, x_n | f \rangle = \lambda_n \langle x_n | f \rangle = \lambda_n \alpha_n \tag{B.48}$$

so that

$$\mathcal{L} f = \sum_n \lambda_n \alpha_n x_n \tag{B.49}$$

This is called the *spectral* representation for the operator, and is related to the diagonalization process described earlier in connection with (B.26).

B.2.3 Continuous Spectrum

So far we have only considered the case when there is a discrete set of eigenvalues and eigenfunctions, each labeled by a subscript. There are many problems which give a continuous spectrum of eigenvalues and corresponding eigenfunctions. For simplicity, we will label an eigenfunction corresponding to a certain eigenvalue λ as x_λ. In the case of a continuous spectrum the eigenfunctions of a self-adjoint operator are orthogonal in the sense of

$$\langle x_\lambda | x_{\lambda'} \rangle = \delta(\lambda - \lambda') \tag{B.50}$$

where $\delta(\lambda - \lambda')$ is the Dirac delta function. An arbitrary square-integrable function $f(t)$ is expanded in terms of continuous eigenfunctions as an integral instead of a discrete sum, where the integral is taken over all the allowed eigenvalues

$$f(t) = \int_{\lambda_a}^{\lambda_b} \alpha(\lambda) x_\lambda(t)\, d\lambda \tag{B.51}$$

where the function $\alpha(\lambda)$ is analogous to the expansion coefficients of the earlier discrete sums. Using the orthogonality property we can easily see that $\alpha(\lambda)$ is again given by

$$\alpha(\lambda) = \langle x_\lambda | f \rangle \tag{B.52}$$

Substituting this result back into the expansion of $f(t)$ we find the completeness relation for continuous eigenfunctions, analogous to (B.40),

$$\int_{\lambda_a}^{\lambda_b} x_\lambda(t) x_\lambda(t')\, d\lambda = \delta(t - t') \tag{B.53}$$

From the definition of completeness we can also find Parseval's theorem for a continuous spectrum,

$$\int_a^b |f(t)|^2\; dt = \int_{\lambda_a}^{\lambda_b} |\alpha(\lambda)|^2\; d\lambda \tag{B.54}$$

B.3 STURM-LIOUVILLE PROBLEMS

Many important differential equations in physics are second order and can be cast in the general form

$$\frac{1}{w(x)}\frac{d}{dx}\left(p(x)\frac{d\psi(x)}{dx}\right) + q(x)\psi(x) - \lambda\psi(x) = 0 \tag{B.55}$$

where $p(x)$, $q(x)$, and $w(x)$ describe both the physical situation and the choice of coordinate system, and λ is a constant. All of the partial differential equations equations of electromagnetics can be written in this form after separation of variables (assuming a separable coordinate system), in which case λ is a separation constant. The equation is written in operator form as

$$\mathcal{L}\psi - \lambda\psi = 0 \qquad \text{where} \quad \mathcal{L} \equiv \frac{1}{w(x)}\frac{d}{dx}p(x)\frac{d}{dx} + q(x) \tag{B.56}$$

From previous sections we know that the eigenvectors of the operator can be used to construct an orthonormal set, provided we define a scalar product such that the operator is self-adjoint. The most common choice in electromagnetics is a real scalar product

$$\langle\psi_i|\psi_j\rangle \equiv \int_a^b w(x)\psi_i(x)\psi_j(x)\,dx \tag{B.57}$$

Using this scalar product we find

$$\langle\psi_i|\mathcal{L}\psi_j\rangle - \langle\mathcal{L}\psi_i|\psi_j\rangle = p(x)\left[\psi_i\frac{d\psi_j}{dx} - \psi_j\frac{d\psi_i}{dx}\right]_a^b \tag{B.58}$$

which follows after a simple integration by parts. This shows that the operator can be made self-adjoint (symmetric) if all the functions ψ satisfy certain types of boundary conditions at $x = a$ and $x = b$:

1. $\psi = 0$ at either boundary (Dirichlet conditions)
2. $d\psi/dx = 0$ at either boundary (Neumann conditions)
3. $d\psi/dx + \alpha\psi = 0$, at either boundary, where α is a constant
4. $\psi(a) = \psi(b)$ and $p(a)\dfrac{d\psi(a)}{dx} = p(b)\dfrac{d\psi(b)}{dx}$ (periodic boundary conditions).

When these boundary conditions are used, $\langle\psi_i|\mathcal{L}\psi_j\rangle = \langle\mathcal{L}\psi_i|\psi_j\rangle$, and the Sturm-Liouville operator is self-adjoint using the scalar product (B.57). Note that the boundary conditions at either end need not be the same; for example, we may have a combination of #1 and #2, such as $\psi(a) = 0$ and $d\psi(b)/dx = 0$. This is a case of *mixed* boundary conditions.

When the operator is self-adjoint, the eigenfunctions belonging to different eigenvectors are then orthogonal as in (B.42), and can be used to construct an orthonormal basis so that

$$\int_a^b w(x)\psi_i(x)\psi_j(x)\,dx = \delta_{ij} \tag{B.59}$$

if the eigenvalue spectrum is discrete, or

$$\int_a^b w(x)\psi_i(x)\psi_j(x)\,dx = \delta(\lambda_i - \lambda_j) \tag{B.60}$$

if the eigenvalue spectrum is continuous. The function $w(x)$ is called a “weighting” function, and is always defined to be positive.

Another choice for the scalar product is useful for complex eigenfunctions and is given by

$$\langle\psi_i|\psi_j\rangle \equiv \int_a^b w(x)\psi_i^*(x)\psi_j(x)\,dx \tag{B.61}$$

which the reader may recognize from quantum mechanics. In this case we find that for the Sturm-Liouville operator to be self-adjoint (Hermitian), we must stipulate that p, q, and w be *real* functions, which gives

$$\langle\psi_i|\mathcal{L}\,\psi_j\rangle - \langle\mathcal{L}\,\psi_i|\psi_j\rangle = p(x)\left[\psi_i^*\frac{d\psi_j}{dx} - \psi_j\frac{d\psi_i^*}{dx}\right]_a^b \tag{B.62}$$

which in turn implies basically the same restrictions on boundary conditions as listed above, with the exception that the constant α in boundary condition #3 must be real. For many problems in electromagnetics, the stipulations for the Sturm-Liouville operator to be self-adjoint are satisfied for both real and complex scalar products, and so they can (and often are) used interchangeably. However, the different requirements on the functions p and q should be kept in mind, since these may be important in some problems, such as wave propagation in lossy media.

It can be shown [4] that the eigenfunctions of the Sturm-Liouville operator are also complete in the sense of (B.37). In summary, then, we have the following properties of the Sturm-Liouville equation when the boundary conditions listed above are used over the range $a \leq x \leq b$:

- There are an infinite number of eigenvalues
- All eigenvalues are real
- Eigenfunctions corresponding to different eigenvalues are orthogonal with respect to the weighting function using the scalar product (B.57)
- The eigenfunctions form a complete set in the range $a \leq x \leq b$

Completeness relations similar to (B.40) can be easily developed for the scalar product (B.61). For the discrete case,

$$\sum_i \frac{\psi_i^*(x)\psi_i(x')}{\langle\psi_i|\psi_i\rangle} = \frac{\delta(x - x')}{w(x)} \tag{B.63}$$

where the term $\langle\psi_i|\psi_i\rangle$ has been left in the summation for the frequent situations involving eigenfunctions that are not normalized.

B.3.1 Summary of Spectral Expansion Technique

The general problem we will encounter has the form

$$(\mathcal{L} - \lambda) f = a \tag{B.64}$$

where a is a given "source" function, λ is a constant (not necessarily one of the eigenvalues of $\mathcal{L}$), and f is the unknown to be solved for. Using the above properties of the operator and the eigenvalues and eigenfunctions defined by

$$\mathcal{L}\,\psi_i = \lambda_i \psi_i \tag{B.65}$$

we can expand f as

$$f = \sum_i \alpha_i \psi_i \qquad \alpha_i = \langle \psi_i | f \rangle \tag{B.66}$$

Using the self-adjoint property of $\mathcal{L}$ we write

$$\langle \psi_i | (\mathcal{L} - \lambda) f \rangle = (\lambda_i - \lambda) \langle \psi_i | f \rangle = \langle \psi_i | a \rangle \tag{B.67}$$

Therefore, as long as λ does not coincide with one of the eigenvalues, the unknown function is given by

$$f = \sum_i \frac{\langle \psi_i | a \rangle}{(\lambda_i - \lambda)} \psi_i \tag{B.68}$$

which is a useful result. When $\lambda = 0$ this reduces to the earlier result (B.46). When λ does coincide with one of the eigenvalues, say $\lambda = \lambda_k$, then it can be shown [2] that the solution is

$$f = C\psi_k + \sum_{i \neq k} \frac{\langle \psi_i | a \rangle}{(\lambda_i - \lambda_k)} \psi_i \tag{B.69}$$

where C is an arbitrary constant.

REFERENCES

[1] B. Friedman, *Principles and Techniques of Applied Mathematics*, Dover: New York, 1990. First published by Wiley: New York, 1956.

[2] J.W. Dettman, *Mathematical Methods in Physics and Engineering*, Dover: New York, 1988. First published by McGraw-Hill: New York, 1962.

[3] Press et. al., *Numerical Recipes*, Cambridge University Press, New York, 1989

[4] F.W. Byron and R.W. Fuller, *Mathematics of Classical and Quantum Physics*, Dover: New York,1992. First published in two volumes by Addison-Wesley, 1969.