PHYS/GEOL/APS 6610
Earth and Planetary Physics I

Eigenfunction Expansions, Sturm-Liouville Problems, and Green's Functions

1. Eigenfunction Expansions

In seismology in general, as in the simple oscillating systems discussed earlier, we are concerned with finding the solutions of homogeneous differential equations with the ultimate object of treating inhomogeneous equations. In one dimension, the homogeneous equations are of the form $\mathcal{L}(y) = 0$ and the inhomogeneous equations are

$$\mathcal{L}(y) = f(x), \tag{1}$$

where $f(x)$ is a prescribed or general function and the boundary conditions to be satisfied by the solution at the end points 0 and L are given. The expression $\mathcal{L}$ represents here a general linear differential operator and not the Laplace Transform! For example, for the damped SHO $\mathcal{L} = m\frac{d^2}{dx^2} + 2b\frac{d}{dx} + \omega_0^2$. In general, unless $f(x)$ is particularly simple, one cannot simply integrate the inhomogeneous equation to solve for $y(x)$. The idea has been in our discussions above and is, in fact, generally true, that one seeks to exploit the linearity of the operator $\mathcal{L}$ by building up the required solution as a superposition of, generally, an infinite number of terms. This method is particularly efficient if we can find suitable functions which, when acted upon by $\mathcal{L}$, somehow eliminate the derivatives. The investigation of this choice and its consequences is the subject of this subsection.

Suppose that we can find a set of functions $\{y_n(x)\}$ $(n = 0, 1, 2, \ldots)$ such that

$$\text{Property One:} \qquad \mathcal{L}(y_n) = -\lambda_n y_n, \tag{2}$$

so that derivatives would be eliminated. Then, as a possible solution to (1) try the superposition

$$y(x) = \sum_n a_n y_n(x), \tag{3}$$

which when substituted into equation (1) yields:

$$f(x) = \mathcal{L}(y) = \mathcal{L}\left(\sum_n a_n y_n\right) = \sum_n a_n \mathcal{L}(y_n) = -\sum_n a_n \lambda_n y_n. \tag{4}$$

This has resulted in a purely algebraic equation, but at the price of introducing the set of unknowns $\{a_n\}$. This can be put right if, in addition, the set $\{y_n(x)\}$ is in some sense mutually orthogonal:

$$\text{Property Two:} \qquad \int_0^L y_m^*(z) y_n(z)\,dz = 0, \qquad \text{m} \neq \text{n}, \tag{5}$$

where $*$ denotes complex conjugate. The orthogonality expressed by this equation is called Hermitian orthogonality due to the complex conjugate. Multiplying both sides of equation (4) by y_m^* and integrating yields:

$$a_m = -\frac{1}{\lambda_m}\frac{\int_0^L y_m^*(z) f(z)\,dz}{\int_0^L y_m^*(z) y_m(z)\,dz}. \tag{6}$$

Equations (3) and (6) form a complete solution as long as the set of functions $\{y_n(x)\}$ exist that satisfy the Properties One and Two given by equations (2) and (5). Functions which satisfy equation (2) are called *eigenfunctions* of the operator $\mathcal{L}$ and, hence, equation (3) is known as an *eigenfunction expansion.* The quantities λ_n are the corresponding *eigenvalues.* The general idea of expansion in terms of a set of orthogonal eigenfunctions is the basis of Fourier Series solutions to differential equations, with which you are familiar. You should be able to see why Fourier Series work so well, sines and cosines are solutions to the SHO equation and are, therefore, eigenfunctions of the SHO differential operator. They are also orthogonal. Thus, the Fourier coefficients can be computed. However, the ideas presented here are much more broadly based than that as we will now see.

The question remains, however, whether for a given operator $\mathcal{L}$ a suitable set of functions can be found. We cannot deal with this problem in general but it is worth pointing out that at least for linear operators of a particular form, such suitable sets of functions can be found and that fairly broad types of boundary conditions can be accommodated.

2. Sturm-Liouville Theory

Confine attention to second-order linear differential equations that are so common in wave propagation problems in which $\mathcal{L}$ has the form:

$$\mathcal{L}(y) \equiv p(x)y^{''} + r(x)y^{'} - q(x)y, \qquad \text{with } r(x) = p^{'}(x), \tag{7}$$

where p, q, and r are real functions of x. The class of differential equations of the form:

$$\mathcal{L}(y) = -\lambda\rho(x)y, \tag{8}$$

were first studied intensively in the 1830s by Sturm and Liouville. Writing equations (7) and (8) together yields

$$(py')' - qy + \lambda\rho y = 0 \tag{9}$$

This is known as the *Sturm-Liouville (S-L) equation* and linear differential operators of the forms given by equation (8) clearly satisfy Property One (eqn (2)), although equation (8) has been slightly generalized to include a weighting function $\rho(x)$. The only conditions on the weighting function are that it is real valued and does not change sign. This latter requirement means that we can assume that it is everywhere positive without loss of generality. Its introduction also requires a generalized definition of orthogonality and the expansion coefficients:

$$\int_0^L \rho(z)y_m^*(z)y_n(z)dz = 0, \qquad \text{m} \neq \text{n}, \tag{10}$$

$$a_m = -\frac{1}{\lambda_m}\frac{\int_0^L \rho(z)y_m^*(z)f(z)dz}{\int_0^L \rho(z)y_m^*(z)y_m(z)dz}. \tag{11}$$

The operation in equation (10) defines the inner product in a function space which is actually an infinite vector space. An infinite vector space with an inner product is called a a Hilbert space in which the eigenfunctions reside. We will say very little more about this.

The satisfaction of Condition One is one of the reasons why S-L equations have been studied so intensively. Another reason is that, although the form looks very restrictive, many of the most important equations in mathematical physics are S-L equations. For example,

$$
\begin{aligned}
(1-x^2)y'' - 2xy' + l(l+1)y &= 0 && \text{Legendre's equation,} && (12)\\
((1-x^2)y')' + \left[l(l+1) - \frac{m^2}{1-x^2}\right] y &= 0 && \text{Associated Legendre equation,} && (13)\\
y'' - 2xy' + 2\alpha y &= 0 && \text{Hermite's equation,} && (14)\\
xy'' + (1-x)y' + \alpha y &= 0 && \text{Laguerre's equation,} && (15)\\
(1-x^2)y'' - xy' + n^2 y &= 0 && \text{Chebyshev equation,} && (16)\\
y'' + 2by' + \omega_0^2 y &= 0 && \text{Simple Harmonic Oscillator equation} && (17)
\end{aligned}
$$

Bessel's equation ($x^2y'' + xy' + (x^2 - n^2)y = 0$) is also an S-L equation with an appropriate change of variables ($\xi = x/a$). It should be noted that any second-order linear differential equation

$$p(x)y'' + r(x)y' + q(x)y + \lambda\rho(x)y = 0, \tag{18}$$

can be converted to the required type by multiplying through by the factor

$$F(x) = \exp\left[\int^x \frac{r(z) - p'(z)}{p(z)} dz\right], \tag{19}$$

provided that the indefinite integral is defined. It then takes on the S-L form

$$\left(F(x)p(x)y'\right)' - (-F(x)q(x)y) + \lambda F(x)\rho(x)y = 0, \tag{20}$$

but clearly with a different, but still non-negative, weighting function ($F(x)\rho(x)$).

Second order linear differential operators, $\mathcal{L}$, for which $\mathcal{L}(y)$ can be written in the form

$$\mathcal{L}(y) = (py')' - qy, \tag{21}$$

where p and q are real functions of x are known as *self-adjoint* operators. This is a bit of a simplification but will suffice here. Therefore, the study of the eigenfunctions and eigenvalues of self-adjoint operators is synonymous with the study of S-L equations. Another useful definition is the following: $\mathcal{L}$ is said to be *Hermitian* if

$$\int_0^L y_m^*(x)\mathcal{L}y_n(x)dx = \left(\int_0^L y_n^*(x)\mathcal{L}y_m(x)dx\right)^* \tag{22}$$

where y_m and y_n are arbitrary functions satisfying the boundary conditions. The quantity on the left side of this equation is called the (m,n) matrix element of $\mathcal{L}$, or $\mathcal{L}_{mn}$, or $< y_m|\mathcal{L}y_n >$. The final notation, of course, is the bra-ket notation of Dirac, but is common in normal mode seismology. In this notation, Hermiticity is stated as $< y_m|\mathcal{L}y_n > = < \mathcal{L}y_m|y_n >$.

S-L equations satisfy Condition One since we seek solutions to the eigenvalue problem given by equation (8). It remains to show that such equations also satisfy Condition Two, that is that the eigenfunctions of self-adjoint operators are orthogonal in the generalized sense of equation (10). It will be left as an exercise to show that

$$(\lambda_m^* - \lambda_n) \int_0^L y_m^* \rho y_n dx = 0, \tag{23}$$

from which the reality of the eigenvalues and the orthogonality of the eigenfunctions follow almost immediately. The derivation of equation (23) requires the specification of boundary conditions at both ends of the range of the free variable (i.e., at both 0 and L). The boundary condition required is the following:

$$[y_m^* p y_n']_{x=a} = [y_m^* p y_n']_{x=b}, \qquad \text{for all m,n} , \tag{24}$$

where $y_m(x)$ and $y_n(x)$ are any two solutions of the S-L equations. Again, this appears to be pretty restrictive, but is actually a pretty mild assumption that is met by many commonly ocurring cases, e.g., $y(0) = y(L) = 0, y(0) = y'(L) = 0, p(0) = p(L) = 0$, and many more.

A last consideration is the normalization of the eigenfunctions. Equation (10) only places a constraint on the eigenfunctions when $m \neq n$. When $m = n$, because of the linearity of $\mathcal{L}$, the normalization is arbitrary. We will assume for definiteness that they are normalized so that $\int y_n^2 \rho dx = 1$. In this case, equations (10) and (11) can be rewritten as

$$< y_m|\rho y_n > \; = \; \int_0^L \rho(z) y_m^*(z) y_n(z) dz = \delta_{mn}, \tag{25}$$

$$a_m = -\lambda_m^{-1} < y_m|\rho f > \; = \; -\lambda_m^{-1} \int_0^L \rho(z) y_m^*(z) f(z) dz. \tag{26}$$

Eigenfunctions corresponding to equal eigenvalues are said to be *degenerate*.

Summarizing then, equations of the form (1) admit the following solution

$$y(x) = -\sum_m \lambda_m^{-1} y_m(x) \int_0^L \rho(z) y_m^*(z) f(z) dz = -\sum_m y_m(x) \lambda_m^{-1} < y_m|\rho f >, \tag{27}$$

if the differential operator is linear and self-adjoint (i.e., results in a differential equation that is a Sturm-Liouville equation); that is, if equations (8), (25), and (24) hold. In addition, the eigenvalues λ_m are real and the eigenfunctions y_n are orthonormal. In addtion, they form a complete set. By a *complete set* we mean that any function satisfying the boundary conditions can be represented as a (potentially infinite) sum of the eigenfunctions. We will not attempt to prove this here. Resulting from this is the so-called *completeness relation* or *closure property* of the eigenfunctions:

$$\rho(x) \sum_n y_n^*(x) y_n(z) = \delta(z - x), \tag{28}$$

which you are asked to prove as an exercise.

3. Green's (or Green) Functions

Starting with equation (27), assume that we can interchange the order of summation and integration:

$$y(x) \; = \; \int_0^L \left\{ \rho(z) \sum_m [-\lambda_m^{-1} y_m(x) y_m^*(z)] \right\} f(z) dz, \tag{29}$$

$$= \; \int_0^L G(x,z) f(z) dz. \tag{30}$$

In this form, the solution to S-L problems has clearer properties, it is an integral of two factors, of which (1) the first is determined entirely by the boundary conditions and the eigenfunctions y_m, and hence by $\mathcal{L}$ itself, and (2) the second, $f(z)$, depends purely on the right-hand side of equation (1). Thus, there is the possibility of finding, once and for all, for any given function $\mathcal{L}$, a function $G(x,z)$ which will enable us to solve equation (1) for any right-hand side; that is any forcing function. The solution will be in the form of an integral which, at worst, can be evaluated numerically. This function, $G(x,z)$, is called the *Green's function* for the operator $\mathcal{L}$. This approach is somewhat similar to the use of Laplace Transforms in that we have reduced the problem to "quadrature" as the British say, but once $G(x,z)$ is found the remaining work to produce a solution is remarkably simple. In addition, Green's functions lend themselves to generalization to multiple dimensions and direct application to partial differential equations. For these reasons, Green's function methods are of greater practical significance that Laplace Transform methods which are mostly used theoretically.

One expression for Green's functions has already been given and can be seen by comparing equations (29) and (30),

$$G(x,z) = -\sum_m \lambda_m^{-1} \rho(z) y_m(x) y_m^*(z). \tag{31}$$

Alternately, we note that equation (30) is, by construction, a solution to equation (1). Hence

$$\mathcal{L}(y) = \int_0^L \mathcal{L}[G(x,z)] f(z) dz = f(x). \tag{32}$$

Now, recall that a delta function is defined as a function that satisfies the following properties:

$$\begin{aligned}
\delta(x - x_0) &= 0 \qquad \text{if } x \neq x_0, & (33)\\
\delta(x - x_0) &= \infty \qquad \text{if } x = x_0, & (34)\\
\int_0^L \delta(x - x_0) dx &= 1 \qquad \text{if } 0 \leq x_0 \leq L, & (35)\\
\int_0^L \delta(x - x_0) f(x) dx &= f(x_0) \quad \text{if } 0 \leq x_0 \leq L. & (36)
\end{aligned}$$

A delta function can be *represented* as limits of a number of different functions; e.g., as an infinitessimally thick Gaussian, a sinc function with infinite frequency, the inverse Fourier Transform of an exponential, and a number of others. The only essential requirements are knowledge of the area under such a curve and that undefined operations such as differentiation are not carried out on the delta-function while some nonexplicit representation is being employed.

Using the property of the delta-function given by equation (36), equation (32) can be rewritten as

$$\mathcal{L}(y) = \int_0^L \left\{ \mathcal{L}[G(x,z)] - \delta(z - x) \right\} f(z) dz = 0. \tag{37}$$

For this to hold for any function f, it must be the case that

$$\mathcal{L}[G(x,z)] = \delta(z - x). \tag{38}$$

Note that in this equation, z is only a parameter and all the differential operations implicit in $\mathcal{L}$ act on the variable x.

Putting equation (38) in words, *the Green's function G is the solution of the differential equation obtained by replacing the right-hand side of equation (1) by a delta function. Thus, the solution to equation (1) given by equation (29) is the superposition of the effects of isolated 'impulses' of size $f(z)dz$ occurring at positions $x = z$.* Since each impulse has effects at locations other than where it acts, the total result at any position x must by obtained by integrating over all z:

$$f(x) = \int_0^\infty f(z)\delta(z-x)dz. \tag{39}$$

That is, $f(x)$ is a limiting case of a whole set of impulses.

1.10 Applications of Superposition and Green's Functions

Superposition

Consider, as an example, the spatial part of the string problem where we have included a forcing term

$$y'' + k^2 y = f(x) \tag{40}$$

to be solved on the interval $[0, L]$ with initial conditions

$$y(0) = y(L) = 0. \tag{41}$$

Temporarily, let the string length $L = \pi$ to simply the calculations, we won't have to carry around a π/L term. We'll translate back to a string of length L by simply replacing every π with an L and every n with an $n\pi/L$. The weight function ρ is unity. In this case, $\mathcal{L} = d^2/dx^2 + k^2$. We seek eigenfunctions satisfying the S-L equation

$$y'' + k^2 y + \lambda y = 0, \tag{42}$$

$$\mathcal{L}y = -\lambda y. \tag{43}$$

These are obviously $y_n = A_n \sin nx + B_n \cos nx$ corresponding to eigenvalues λ_n given by $n^2 = \lambda_n + k^2$. The boundary conditions require that n be a positive integer and that $B_n = 0$. Thus, the eigenfunctions are $y_n = A_n \sin nx$ and the normalization condition, equation (25), requires that $A_n = \sqrt{2/\pi}$. Using equation (26),

$$a_n = -(n^2 - k^2)^{-1} \int_0^\pi \left(\frac{2}{\pi}\right)^{1/2} \sin nz dz, \tag{44}$$

and finally that the solution in terms of the given function $f(x)$ is

$$y(x) = \sum_n a_n y_n = \sum_n a_n A_n \sin nx = -\frac{2}{\pi}\sum_{n=1}^\infty \frac{\sin nx}{n^2 - k^2}\int_0^\pi f(z)\sin nz dz. \tag{45}$$

Upon transforming back to a string of length L we get:

$$y(x) = -\frac{2}{L}\sum_{n=1}^\infty \frac{\sin(n\pi x/L)}{(n\pi/L)^2 - k^2}\int_0^\pi f(z)\sin(n\pi z/L)dz. \tag{46}$$

Note that this solution is also the Fourier Series form because of the particular form of the linear operator involved. However the above method is a general model for all equations involving S-L-like operators.

Green's Functions

From (46), the Green's function for the undamped SHO is immediately apparent:

$$G(x,z) = -\frac{2}{L}\sum_{n=1}^{\infty}\frac{\sin(n\pi x/L)}{(n\pi/L)^2 - k^2}\sin(n\pi z/L). \tag{47}$$

We call this the *discrete form of the Green's Function* and illustrates a symmetry relation for Green's functions also apparent in equation (31), that $G(x,z) = (G(z,x))^*$.

This form is not very convenient for computation since it involves an infinite sum. A more useful form of this Green's function can be obtained by using the fact that the Green's function is the solution of the differential equation (40) in which the forcing function has been replaced by the delta function, $\delta(z-x)$:

$$\frac{d^2G(x,z)}{dx^2} + k^2 G(x,z) = \delta(z-x) \tag{48}$$

The Green's function will depend on the initial conditions. The solution of equation (48) subject to the initial conditions that $y(0) = y'(0) = 0$ will be left as an exercise, with the result:

$$G(x,z) = \frac{1}{k}\sin k(x-z), \tag{49}$$

if $0 < z < x$ and 0 otherwise. Using this approach produces Green's functions that act as integral kernels and are called the *continuous form of the Green's Function*. Using this expression for the Green's function the solution to equation (40) subject to the different initial conditions listed directly above can be written:

$$y(x) = \int_0^x \frac{1}{k}\sin k(x-z) f(z)dz. \tag{50}$$

Finding the Green's function for the boundary conditions given in equation (41) is somewhat more complicated, but it is instructive, so let's do so but keep the string length equal to L (rather than π) here. $G(x,z)$ still satisfies equation (48). For x equal to anything but z we have

$$\begin{aligned} G(x,z) &= A\sin kx \qquad (0 < x < z), & (51)\\ &= B\sin k(x-L) \qquad (z < x < L). & (52)\end{aligned}$$

To determine the constants A and B we need to apply conditions on G and its first derivative at $x = z$. To find the appropriate conditions, we integrate equation (48) from $x = x - \epsilon$ to $x = z + \epsilon$ and then let $\epsilon \to 0$. Since $\int d^2G/dx^2 = dG/dx$, we find that

$$\frac{dG}{dx}\Big|_{z-\epsilon}^{z+\epsilon} + \int_{z-\epsilon}^{z+\epsilon} G(x,z)dx = \int_{z-\epsilon}^{z+\epsilon}\delta(x-z)dx = 1, \tag{53}$$

so that letting $\epsilon \to 0$ the second term on the right-hand-side goes to zero and the change in slope at $x = z$ is 1. Integrating again gives

$$G|_{z-\epsilon}^{z+\epsilon} = 0, \tag{54}$$

which implies that G is continuous at $x = z$. This yields the pair of simultaneous equations

$$A \sin kz = B \sin k(z - L) \tag{55}$$

$$kA \cos kz + 1 = kB \cos k(z - L), \tag{56}$$

which upon solution give

$$A = \frac{\sin k(z - L)}{k \sin kL} \qquad B = \frac{\sin kz}{k \sin kL} \tag{57}$$

and the Green's function is

$$G(x, z) = (k \sin kL)^{-1} \sin kx \sin k(z - L) \qquad 0 < x < z \tag{58}$$

$$= (k \sin kL)^{-1} \sin kz \sin k(x - L) \qquad z < x < L \tag{59}$$

from which it is immediately apparent that

$$y(x) = \frac{\sin kx}{k \sin kL} \int_0^x f(z) \sin k(z - L) dz + \frac{\sin k(x - L)}{k \sin kL} \int_x^L f(z) \sin(kz) dz. \tag{60}$$

Following the same method it is possible to show that a solution to the differential equation

$$y'' + p(x)y' + q(x)y = f(x) \tag{61}$$

with $y(0) = y(L) = 0$ is given by

$$y(x) = y_2(x) \int_0^x \frac{y_1(x')f(x')}{W(x')} dx' + y_1(x) \int_x^L \frac{y_2(x')f(x')}{W(x')} dx', \tag{62}$$

where $y_1(x)$ and $y_2(x)$ are solutions of the homogeneous equation with $y_1(0) = y_2(0) = 0$, and W is the Wronskian of $y_1(x)$ and $y_2(x)$: $W(x) = y_1(x)y_2'(x) - y_2(x)y_1'(x)$. Recall that if $W \neq 0$, y_1 and y_2 are linearly independent. Also as in the above, we can find that a particular solution y_p of equation (61) is

$$y_p(x) = y_2(x) \int \frac{y_1(x)f(x)}{W(x)} dx + y_1(x) \int \frac{y_2(x)f(x)}{W(x)} dx. \tag{63}$$

This particular solution is exactly the same as that obtained by variation of parameters, but may seem somewhat less arbitrary.