

Linearity, linear operators, and self adjoint eigenvalue problems

1 Elements of linear algebra

The study of linear partial differential equations utilizes, unsurprisingly, many concepts from linear algebra and linear ordinary differential equations. Here a brief overview of the required concepts is provided.

1.1 Vector spaces and linear combinations

A vector space S is a set of elements - numbers, vectors, functions - together with notions of addition, scalar multiplication, and their associated properties. A *linear combination* of elements $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in S$ is any expression of the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_k\mathbf{v}_k$$

where the coefficients c_j are real or complex scalars. A significant property of vector spaces is that any linear combination of elements in S is also in S . This is easily verified in most cases - for example, \mathbb{R}^n (the set of n -dimensional vectors) and $C^0(\mathbb{R})$ (the set of continuous functions on the real line) are vector spaces.

1.2 Linear transformations and operators

Suppose \mathbf{A} is a $n \times n$ matrix, and \mathbf{v} is a n -dimensional vector. The matrix-vector product $\mathbf{y} = \mathbf{A}\mathbf{v}$ can be regarded as a mapping that takes \mathbf{v} as a input and produces the n -dimensional vector \mathbf{y} as an output. More precisely this mapping is a *linear transformation* or *linear operator*, that takes a vector \mathbf{v} and "transforms" it into \mathbf{y} . Conversely, every linear mapping from $\mathbb{R}^n \rightarrow \mathbb{R}^n$ is represented by a matrix vector product.

The most basic fact about linear transformations and operators is the property of *linearity*. In words, this says that a transformation of a linear combination is the linear combination of the linear transformations. This clearly applies to the case where the transformation is given by a matrix \mathbf{A} , since

$$\mathbf{A}(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1\mathbf{A}\mathbf{v}_1 + c_2\mathbf{A}\mathbf{v}_2. \quad (1)$$

More generally, if f_1, f_2 are elements of a vector space S , then a linear operator \mathcal{L} is a mapping from S to some other vector space (frequently also S) so that

$$\mathcal{L}(c_1f_1 + c_2f_2) = c_1\mathcal{L}f_1 + c_2\mathcal{L}f_2. \quad (2)$$

1.3 Eigenvalues and eigenvectors

Recall that $\mathbf{v} \neq 0, \lambda$ is an eigenvector-eigenvalue pair of \mathbf{A} if

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}.$$

Except in rare cases (for so called "non-normal" matrices), there are exactly n linearly independent eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ corresponding to eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$ (the eigenvalues can sometimes be the same for different eigenvectors, however). Why is this a useful fact? Note that if \mathbf{x} is any n -dimensional vector, it can be written in terms of the basis given by the eigenvectors, i.e.

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n. \quad (3)$$

The c_i solve a linear system $\mathbf{V}\mathbf{c} = \mathbf{x}$ where \mathbf{V} has columns of eigenvectors and \mathbf{c} is the vector $(c_1 c_2 \dots)^T$. When \mathbf{A} is self-adjoint (see below), there is a much easier way to find the coefficients.

Note that action of a linear transformation A on the vector \mathbf{x} can be written simply as

$$\begin{aligned}\mathbf{A}\mathbf{x} &= \mathbf{A}(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n) \\ &= c_1\mathbf{A}\mathbf{v}_1 + c_2\mathbf{A}\mathbf{v}_2 + \dots + c_n\mathbf{A}\mathbf{v}_n \\ &= c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \dots + c_n\lambda_n\mathbf{v}_n.\end{aligned}$$

In other words, *eigenvectors decompose a linear operator into a linear combination*, which is a fact we often exploit.

1.4 Inner products and the adjoint operator

It is frequently helpful to attach geometric ideas to vector spaces. One way of doing this is to specify an inner product, which is a map $S \times S \rightarrow \mathbb{R}$ or $S \times S \rightarrow \mathbb{C}$. The inner product is basically a way of specifying how to measure angles and lengths.

For $\mathbf{v}_1, \mathbf{v}_2 \in S$, we will write an inner product as $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle$. There are many different ways to define inner products, but they must all have the following properties:

1. Symmetry: $\langle v, u \rangle = \langle u, v \rangle$ for every $u, v \in S$. If the inner product is complex valued, this needs to be $\langle v, u \rangle = \langle u, v \rangle$, where \bar{z} denotes the complex conjugate.
2. Linearity in the first variable: for any vector \mathbf{v} and vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ we have

$$\langle c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n, \mathbf{v} \rangle = c_1\langle \mathbf{v}_1, \mathbf{v} \rangle + c_2\langle \mathbf{v}_2, \mathbf{v} \rangle + \dots + c_n\langle \mathbf{v}_n, \mathbf{v} \rangle. \quad (4)$$

3. Positivity: $\langle \mathbf{v}, \mathbf{v} \rangle > 0$ unless $\mathbf{v} = 0$.

Note for complex vector spaces, the first two properties imply “conjugate” linearity in the second variable:

$$\langle \mathbf{v}, c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_n\mathbf{v}_n \rangle = \bar{c}_1\langle \mathbf{v}, \mathbf{v}_1 \rangle + \bar{c}_2\langle \mathbf{v}, \mathbf{v}_2 \rangle + \dots + \bar{c}_n\langle \mathbf{v}, \mathbf{v}_n \rangle. \quad (5)$$

The inner product defines length in the sense that $\langle \mathbf{v}, \mathbf{v} \rangle$ is thought of as the square of the magnitude or *norm* of \mathbf{v} . The inner product also measures how parallel two elements of vector space are. In particular, we define $\mathbf{v}_1, \mathbf{v}_2$ to be *orthogonal* if $\langle \mathbf{v}_1, \mathbf{v}_2 \rangle = 0$.

Once an inner product is defined, then for any linear transformation or operator \mathcal{L} , there is another operator called the *adjoint* of \mathcal{L} , written \mathcal{L}^\dagger . What defines the adjoint is that, for any two vectors $\mathbf{v}_1, \mathbf{v}_2$,

$$\langle \mathcal{L}\mathbf{v}_1, \mathbf{v}_2 \rangle = \langle \mathbf{v}_1, \mathcal{L}^\dagger\mathbf{v}_2 \rangle. \quad (6)$$

This definition is a bit confusing because \mathcal{L}^\dagger is not explicitly constructed. You should think of this as “if I find an operator \mathcal{L}^\dagger that satisfies property (6), it must be the adjoint.”

The dot product for finite dimensional vectors is the best known example of an inner product (there are in fact many ways of defining inner products, even for vectors). As an example, consider the linear operator on n -dimensional (real) vectors given by multiplying by a matrix, i.e. $\mathcal{L}\mathbf{v} = \mathbf{A}\mathbf{v}$. By the usual rules of matrix multiplication

$$(\mathbf{A}\mathbf{v}_1) \cdot \mathbf{v}_2 = \mathbf{v}_1 \cdot (\mathbf{A}^T\mathbf{v}_2),$$

which means that the transpose of \mathbf{A} is same as the adjoint of \mathbf{A} , with respect to the inner product defined by the dot product. Note that if one used a different inner product, the adjoint might be different as well.

1.5 Self-adjointness and expansion in eigenvectors

Sometimes an operator is its own adjoint, in which case it's called *self-adjoint*. Self-adjoint operators have some very nice properties which we will exploit. The most important are

1. The eigenvalues are real.
2. The eigenvectors corresponding to different eigenvalues are orthogonal.

Suppose a matrix \mathbf{A} is symmetric and therefore self-adjoint, and we know its eigenvectors. As in (3), we can try to write any vector \mathbf{x} as a linear combination of the eigenvectors. Taking the inner product of (3) with any particular eigenvector \mathbf{v}_k and using (4), we have

$$\begin{aligned}\langle \mathbf{x}, \mathbf{v}_k \rangle &= \langle c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n, \mathbf{v}_k \rangle \\ &= c_1 \langle \mathbf{v}_1, \mathbf{v}_k \rangle + c_2 \langle \mathbf{v}_2, \mathbf{v}_k \rangle + \dots + c_k \langle \mathbf{v}_k, \mathbf{v}_k \rangle \\ &= c_k \langle \mathbf{v}_k, \mathbf{v}_k \rangle\end{aligned}\tag{7}$$

since \mathbf{v}_k is orthogonal to all eigenvectors except itself. Therefore we have a simple formula for any coefficient c_k :

$$c_k = \frac{\langle \mathbf{x}, \mathbf{v}_k \rangle}{\langle \mathbf{v}_k, \mathbf{v}_k \rangle}.\tag{8}$$

In some cases the eigenvectors are rescaled or *normalized* so that $\langle \mathbf{v}_k, \mathbf{v}_k \rangle = 1$, which means that (8) simplifies to $c_k = \langle \mathbf{x}, \mathbf{v}_k \rangle$.

2 Differential linear operators

We can think of derivatives as linear operators which act on a vector space of functions. Although these spaces are infinite dimensional (recall, for instance, that $1, x, x^2, \dots$ are linearly independent), notions such as linearity, eigenvalues, and adjoints still apply.

As an example, consider the second derivative of a function $d^2 f/dx^2$. This can be thought of as a mapping $f(x)$ to the output $f''(x)$. The basic linearity property (1) is easily verified since

$$\frac{d^2}{dx^2} (c_1 f_1(x) + c_2 f_2(x)) = c_1 \frac{d^2 f_1}{dx^2} + c_2 \frac{d^2 f_2}{dx^2}.$$

There is a technical point to be made here. We sometimes have to worry about what set of functions actually constitutes our vector space. For the example above, a natural vector space for the domain of the linear operator is $C^2(\mathbb{R})$, the set of all twice differentiable functions on the real line. We could have instead chosen the set of infinitely differentiable functions $C^\infty(\mathbb{R})$. Changing the space of functions on which a differential operator acts may affect things like eigenvalues and adjointness properties.

2.1 Linear differential equations

All linear equations involve a linear operator \mathcal{L} . There are two types of linear equations, homogeneous and inhomogeneous, which have the forms

$$\begin{aligned}\mathcal{L}f &= 0, & (\text{homogeneous}), \\ \mathcal{L}f &= g, & (\text{inhomogeneous}).\end{aligned}$$

Here f is the “solution” (the function to be found), \mathcal{L} is some differential linear operator, and g is another given function. As a rule of thumb, identifying a linear equation is just a matter of making sure that each term in the equation is a linear operator acting on the unknown function, or a term which does not involve the unknown.

2.2 The superposition principle

The big advantage of linearity is the *superposition principle*, which permits a divide-and-conquer strategy for solving equations. If we can find enough linearly independent solutions, we can get all solutions simply by forming linear combinations out of these building blocks. More precisely,

Superposition principle for homogeneous equations:

If f_1, f_2, \dots are solutions of $\mathcal{L}f = 0$, then so is any linear combination:

$$\mathcal{L}(c_1f_1 + c_2f_2 + \dots) = c_1\mathcal{L}f_1 + c_2\mathcal{L}f_2 + \dots = 0.$$

This basic idea can be amended for inhomogeneous equations. In this case, one needs to find any single *particular solution* f_p which solves $\mathcal{L}f_p = g$, so that the difference $h = f - f_p$ solves a homogeneous equation $\mathcal{L}h = 0$. The result is

Superposition principle for inhomogeneous equations:

If h_1, h_2, \dots are solutions of $\mathcal{L}h = 0$, then any linear combination of

these plus the particular solution solves the inhomogeneous equation $\mathcal{L}f = g$:

$$\mathcal{L}(f_p + c_1h_1 + c_2h_2 + \dots) = \mathcal{L}f_p + c_1\mathcal{L}h_1 + c_2\mathcal{L}h_2 + \dots = g.$$

2.3 Superposition principle and boundary conditions

One difficulty that arises from using the superposition principle is that, while the equation itself may be satisfied, not every linear combination will satisfy the boundary or initial conditions unless they are of a special form. We call any side condition *linear homogeneous* if it can be written in the form $\mathcal{B}f = 0$, where \mathcal{B} is a linear operator. The most basic types of boundary conditions are of this form: for the Dirichlet condition, \mathcal{B} is the “restriction” operator which simply maps a function to its boundary values. For the Neumann condition, the operator first takes a normal derivative before restriction to the boundary.

We may extend the superposition principle to handle boundary conditions as follows. If f_1, f_2, \dots satisfy linear, homogeneous boundary conditions, then so does an superposition of these $c_1f_1 + c_2f_2 + \dots$. Therefore in most PDEs, the superposition principle is only useful if at least some of the side conditions are linear and homogeneous.

2.4 Sturm-Liouville operators

As an illustration, we will consider a class of differential operators called Sturm-Liouville operators that are fairly easy to work with, and frequently arise in the study of ordinary and partial differential equations. For specified coefficient functions $q(x)$ and $p(x) > 0$, these operators have the form

$$\mathcal{L} = \frac{d}{dx} \left(p(x) \frac{d}{dx} \right) + q(x). \quad (9)$$

(the notation means that to apply \mathcal{L} to a function $f(x)$, we put f on the right hand side and distribute terms so $\mathcal{L}f = (p(x)f')' + q(x)f(x)$) The functions p and q are prescribed, and in the

most basic cases are simply constant. The vector space that \mathcal{L} acts on will be $C_0^\infty[a, b]$, which means that f is infinitely differentiable and $f(a) = f(b) = 0$ (i.e. homogeneous Dirichlet boundary conditions).

2.5 Inner products and self adjointness

As pointed out earlier, there are many ways of defining an inner product. Given two functions f, g in $C_0^\infty[a, b]$, we define the so called L^2 inner product as

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx. \quad (10)$$

We can now show that the Sturm-Liouville operator (9) acting on $C_0^\infty[a, b]$ is self-adjoint with respect to this inner product. How does one compute the adjoint of a differential operator? The answer lies in using integration by parts (or in higher dimensions, Green's formula). For any two functions f, g in $C_0^\infty[a, b]$ we have

$$\begin{aligned} \langle \mathcal{L}f, g \rangle &= \int_a^b \frac{d}{dx} \left(p(x) \frac{df}{dx} \right) g(x) + q(x)f(x)g(x)dx \\ &= \int_a^b -p(x) \frac{df}{dx} \frac{dg}{dx} + q(x)f(x)g(x)dx + p(x)g(x) \frac{df}{dx} \Big|_a^b \\ &= \int_a^b \frac{d}{dx} \left(p(x) \frac{dg}{dx} \right) f(x) + q(x)f(x)g(x)dx - p(x)f(x) \frac{dg}{dx} \Big|_a^b \\ &= \langle f, \mathcal{L}g \rangle. \end{aligned}$$

Integration by parts was used twice to move derivatives off of f and onto g . Therefore (comparing to (6)) \mathcal{L} is its own adjoint. Note that the boundary conditions were essential to make the boundary terms in the integration by parts vanish.

As another example, consider the weighted inner product on $C_0^\infty[0, R]$ defined by

$$\langle f, g \rangle = \int_0^R r f(r)g(r)dr. \quad (11)$$

Then for the linear operator $\mathcal{L}f = r^{-1}(rf)'$, integration by parts twice gives

$$\begin{aligned} \langle \mathcal{L}f, g \rangle &= \int_0^R r \left[r^{-1} \frac{d}{dr} \left(r \frac{df}{dr} \right) \right] g(r)dr \\ &= - \int_0^R r \frac{df}{dr} \frac{dg}{dr} dr, \quad (\text{int. by parts}) \\ &= \int_0^R \frac{d}{dr} \left(r \frac{dg}{dr} \right) f(r) dr, \quad (\text{int. by parts}) \\ &= \int_0^R r \left[r^{-1} \frac{d}{dr} \left(r \frac{dg}{dr} \right) \right] f(r) dr \\ &= \langle f, \mathcal{L}g \rangle. \end{aligned}$$

Note in the final step, a weight factor of r was needed inside the integral to obtain the inner product (11). Thus \mathcal{L} is self adjoint with respect to the weighted inner product. It turns out not to be self adjoint, however, with respect to the unweighted one (10).

What about non-self-adjoint operators? The simplest example is $\mathcal{L} = d/dx$ acting on $C_0^\infty[a, b]$. We again compute

$$\langle \mathcal{L}f, g \rangle = \int_a^b \frac{df}{dx} g(x) = - \int_a^b \frac{dg}{dx} f(x) dx = \langle f, -\mathcal{L}g \rangle.$$

It follows that $\mathcal{L}^\dagger = -\mathcal{L}$. Operators which are their negative adjoints are called *skew symmetric*.

2.6 Eigenvalue problems for differential operators

Analogous to the matrix case, we can consider an *eigenvalue problem* for a linear operator \mathcal{L} , which asks to find non-trivial function-number pairs $(v(x), \lambda)$ which solve

$$\mathcal{L}v(x) = \lambda v(x). \tag{12}$$

It should be noted that sometimes the eigenvalue problem is written instead like $\mathcal{L}v(x) + \lambda v(x) = 0$ (which reverses the sign of λ), but the theory all goes through just the same. The functions $v(x)$ which satisfy this are called *eigenfunctions* and each corresponds to an eigenvalue λ . As with eigenvectors, we can rescale eigenfunctions: if $v(x), \lambda$ is an eigenvector-value pair, so is $(cv(x), \lambda)$ for any scalar $c \neq 0$.

What can we expect to happen in these problems? Since we are working in infinite dimensions, there are often an infinite number of distinct eigenvalues and eigenfunctions. The set of eigenvalues is also called the *discrete spectrum* (there is another part called the essential or continuous spectrum which we will not get into). Often the eigenfunctions form a basis for the underlying vector space, and the set is called *complete*. This property is extremely valuable in PDEs, since it allows for solutions to be expressed as linear combinations of eigenfunctions. The most famous example of this is the Fourier series, which we discuss later.

Self-adjoint operators have some properties equivalent to self-adjoint matrices. In particular, their eigenvalues are real, and their eigenfunctions are orthogonal. Orthogonality is taken with respect to the same inner product which gives self-adjointness.

2.7 Sturm-Liouville eigenvalue problems

Sturm-Liouville operators (9) on $C_0^\infty[a, b]$ have some other nice properties aside from those of any self-adjoint operator. For the eigenvalue problem

$$\frac{d}{dx} \left(p(x) \frac{dv}{dx} \right) + q(x)v(x) + \lambda v(x) = 0,$$

the following properties hold:

1. The real eigenvalues can be ordered $\lambda_1 < \lambda_2 < \lambda_3 \dots$ so that there is a smallest (but not largest) eigenvalue.
2. The eigenfunctions $v_n(x)$ corresponding to each eigenvalue λ_n form a complete set, i.e. for any $f \in C_0^\infty[a, b]$, we can write f as a (infinite) linear combination

$$f = \sum_{n=1}^{\infty} c_n v_n(x). \tag{13}$$

The infinite series in (13) should cause some concern. It should be noted that for functions, there are different ways to define convergence. One way is just to fix each point x and consider the limit of function values, which is called “pointwise” convergence. We actually have something stronger here: the rate at which the sequences at *all* x converge is the same. This is called *uniform convergence*.

Here is the simplest example of a Sturm-Liouville eigenvalue problem. Consider the operator $\mathcal{L} = d^2/dx^2$ on the vector space $C_0^\infty[0, \pi]$. The eigenvalue problem reads

$$\frac{d^2v}{dx^2} + \lambda v = 0, \quad v(0) = 0, \quad v(\pi) = 0. \quad (14)$$

This is just a second order differential equation, and writing down the general solution is easy. Recall that we “guess” solutions of the form $v(x) = \exp(rx)$. Provided $\lambda > 0$, we get $r = \pm i\sqrt{\lambda}$, which means that the general solution has the form

$$v(x) = A \cos(\sqrt{\lambda}x) + B \sin(\sqrt{\lambda}x).$$

Not all of these solutions are valid; we require that $v(0) = 0 = v(\pi)$. Therefore $A \cos(0) + B \sin(0) = 0$, so that $A = 0$. The other boundary condition implies $B \sin(\pi\sqrt{\lambda}) = 0$ which is only true if $\pi\sqrt{\lambda}$ is a multiple of π . Therefore

$$\lambda_n = n^2, \quad n = 1, 2, 3, \dots$$

Corresponding to each of these eigenvalues is the eigenfunction $v_n(x) = \sin(nx)$, $n = 1, 2, 3, \dots$. Recall that we don’t care about the prefactor B since eigenfunctions can always be rescaled.

Properly speaking, we also need to consider the cases $\lambda = 0, \lambda < 0$. For the first case, $v(x) = Ax + B$, and it’s obvious one needs $A = B = 0$ to satisfy the boundary conditions. If $\lambda < 0$, $v(x) = A \exp(\sqrt{|\lambda|x}) + B \exp(-\sqrt{|\lambda|x})$. The boundary conditions imply $A + B = 0$ and $A \exp(\sqrt{|\lambda|\pi}) + B \exp(-\sqrt{|\lambda|\pi}) = 0$, which written in matrix form is

$$\begin{pmatrix} 1 & 1 \\ \exp(\sqrt{|\lambda|\pi}) & \exp(-\sqrt{|\lambda|\pi}) \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Such a system only has nonzero solutions if the matrix is singular, which it is not since its determinant is $\exp(-\sqrt{|\lambda|\pi}) - \exp(\sqrt{|\lambda|\pi}) \neq 0$.

Be careful: there are Sturm-Liouville eigenvalue problems which *do* have non-positive eigenvalues. On the other hand, there is no need to worry about complex eigenvalues because the linear operator is self-adjoint.

2.8 Fourier series

There is a big payoff from fact that the eigenfunctions of the previous example are complete: we can write any function in $C_0^\infty[0, \pi]$ as a linear combination of the eigenfunctions of d^2/dx^2 . In other words we can write any smooth function with zero boundary values as

$$f(x) = \sum_{n=1}^{\infty} B_n \sin(nx). \quad (15)$$

This is the famous Fourier sine series - just one of several types of Fourier series.

Series type	Space of functions	Orthogonal expansion for $f(x)$	Coefficients
Fourier	$f(x) : [-L, L] \rightarrow \mathbb{R}$ $f(-L) = f(L)$ $f'(-L) = f'(L)$	$\frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos(\frac{n\pi x}{L})$ $+ \sum_{n=1}^{\infty} B_n \sin(\frac{n\pi x}{L})$	$A_n = \frac{1}{L} \int_{-L}^L f(x) \cos(\frac{n\pi x}{L}) dx$ $B_n = \frac{1}{L} \int_{-L}^L f(x) \sin(\frac{n\pi x}{L}) dx$
Sine	$f(x) : [0, L] \rightarrow \mathbb{R}$ $f(0) = 0 = f(L)$	$\sum_{n=1}^{\infty} B_n \sin(\frac{n\pi x}{L})$	$B_n = \frac{2}{L} \int_0^L f(x) \sin(\frac{n\pi x}{L}) dx$
Cosine	$f(x) : [0, L] \rightarrow \mathbb{R}$ $f'(0) = 0 = f'(L)$	$\frac{A_0}{2} + \sum_{n=1}^{\infty} A_n \cos(\frac{n\pi x}{L})$	$A_n = \frac{2}{L} \int_0^L f(x) \cos(\frac{n\pi x}{L}) dx$
Complex	$f(x) : [-L, L] \rightarrow \mathbb{C}$ $f(-L) = f(L)$ $f'(-L) = f'(L)$	$\sum_{n=-\infty}^{\infty} c_n \exp(\frac{in\pi x}{L})$	$c_n = \frac{1}{2L} \int_{-L}^L f(x) \exp(\frac{-in\pi x}{L}) dx$

Table 1: Various Fourier series

The question that arises is, how do we actually compute the coefficients B_n ? We have already answered this question in a more general setting. Because the eigenfunctions $\sin(nx)$ are orthogonal, we can use (7) and (8). For the present case, this is equivalent to taking an inner product of (15) with each eigenfunction $\sin(nx)$. This gives the equivalent of (7), namely

$$B_n = \frac{\langle f(x), \sin(nx) \rangle}{\langle \sin(nx), \sin(nx) \rangle} = \frac{\int_0^\pi f(x) \sin(nx) dx}{\int_0^\pi \sin^2(nx) dx}.$$

It should be emphasized that Fourier coefficient formulas like this one don't need to be memorized. They arise quite simply from the basic idea of finding coefficients of a linear combination of orthogonal vectors.

There are other types of Fourier series involving cosines or complex exponentials. These functions are all eigenfunctions of the second derivative operator $\mathcal{L} = d^2/dx^2$, but with different vector spaces of functions on which it acts (remember the technical point earlier: properties of differential operators crucially depend on the vector space in question). Table 2.8 summarizes the standard types of Fourier series. The coefficient formulas all derive from taking inner products with each eigenfunction (note that for the complex series, the inner product is $\langle u, v \rangle = \int_{-L}^L u \bar{v} dx$). We shall see later that other eigenvalue problems give rise to new orthogonal sets of functions: the Bessel functions, and the spherical harmonics.

2.9 Integral operators

Another important type of linear operator involves integrals, such as the "Hilbert-Schmidt" class of integral operators

$$\mathcal{L}u(x) = \int_{\Omega} k(x, y)u(y)dy. \tag{16}$$

(Notice that the output of this operation is a function of y , not x). Here Ω is the domain of u and $k(x, y)$ is called the *kernel*, which is a function from $\Omega \times \Omega$ to \mathbb{R} . Linearity arises simply because of linearity of the integral:

$$\mathcal{L}(c_1u_1 + c_2u_2) = c_1 \int_{\Omega} k(x, y)u_1(y)dy + c_2 \int_{\Omega} k(x, y)u_2(y)dy = c_1\mathcal{L}u_1 + c_2\mathcal{L}u_2.$$

We will not encounter many equations which have integral operators, but some of our solutions will involve integral operators. This is because integral operators are often inverses of differential operators. This means that the inhomogeneous equation $\mathcal{L}u = g$ has a solution $u = \mathcal{L}^{-1}g$ where \mathcal{L}^{-1} is an integral operator.

Integral operators have adjoints. For example, using the L^2 inner product, the operator given in (16) satisfies

$$\begin{aligned}\langle \mathcal{L}u, v \rangle &= \int_{\Omega} v(y) \int_{\Omega} k(x, y) u(x) dx dy \\ &= \int_{\Omega} u(x) \int_{\Omega} k(x, y) v(y) dy dx = \langle u, \mathcal{L}^\dagger v \rangle\end{aligned}$$

where \mathcal{L}^\dagger is another integral operator

$$\mathcal{L}^\dagger v(x) = \int_{\Omega} k(y, x) v(y) dy.$$

This is not the same as \mathcal{L} ; the inputs to the kernel $k(y, x)$ have been reversed.