Math 421 - Lecture notes

TABLE OF CONTENTS

Chapter 1

Linear algebra

1.1 Vector Spaces and Matrix Algebra

1.1.1 Introduction

The Laplace transform we studied in the last chapter is an example of a linear operator. The natural domain of a linear operator is a linear space, roughly a collection of objects that is closed under addition and scalar multiplication. For example, for any integer n, the collection of n-tuples $\mathbb{R}^n$ is an example of a vector space. In the first part of this section we study the basic properties of linear space.

The natural linear operator that acts on $\mathbb{R}^n$ is represented by a $m \times n$ matrix. The second part of this section deals with matrix algebra: the interaction between linear operators on $\mathbb{R}^n$.

1.1.2 Vector spaces

Definition 1.1.1. *A vector space is essentially a collection that is closed under addition and scalar multiplication. That is let V be a vector space. Then for any $x, y \in V$ and scalars a, b,*

$$ax + by \in V.$$

The usual properties of addition and multiplication apply, such as commutatitivity, associativity and distributivity. The (unique) identity for the addition operation is called the 0 vector:

$$x + 0 = 0 + x = x.$$

Finally every element x has its additive inverse $-x$:

$$x + (-x) = -x + x = 0.$$

Example 1.1.2. *The typical examples of a vector space include*

a. $\mathbb{R}^n$, the space of n-tuples.

b. $C[0,\infty)$*, the space of continuous functions on* $[0,\infty)$*.*

c. $C^n[0,\infty)$*, the space of functions that are continously differentiable up to the nth order.*

In general, one just check that the collection is closed under addition and scalar multiplication. The 0 element and the existence of additive inverse are usually obvious.

1.1.3 Subspace

A subset W of a vector space V is a subspace if W itself is a vector space. For example, $C[0,\infty)$ is a subspace of $C^1[0,\infty)$. Another example is the collection of 3-tuples with zero first entry being a subspace of $\mathbb{R}^3$.

1.1.4 Linear independence and basis

Definition 1.1.3. *A subset* W *of a vector space* V *is linearly independent if the following holds for any* $w_1,\cdots,w_n \in W$*:*

$$\sum_{i=1}^{n} c_i w_i = 0$$

if and only if

$$c_1 = c_2 = \cdots = c_n = 0.$$

We say W *is linearly dependent if the above fails.*

Example 1.1.4. *The collection* $\{[1,0,0],[0,1,0],[0,0,1]\}$ *is a linearly independent subset of* $\mathbb{R}^3$*. The collection* $\{[1,0,0],[0,1,0],[1,2,0]\}$ *is a* **linearly dependent** *subset of* $\mathbb{R}^3$*.*

Definition 1.1.5. *A subset* W *of a vector space* V *is a basis of* V *if the following holds:*

a. W *is linearly-independent.*

b. For any v *in* V*, there exist* $w_1,\cdots,w_n$ *and scalars* $c_1,\cdots,c_n$ *such that*

$$\sum_{i=1}^{n} c_i w_i = v.$$

Remark: It is a fact that any vector space has a basis. Moreover, all bases of a vector space have the same size. The size of a basis of a vector space is referred to as its dimension.

Example 1.1.6. *The collection* $\{[1,0,0],[0,1,0],[0,0,1]\}$ *is a basis of* $\mathbb{R}^3$*. Thus (obviously) the dimension of* $\mathbb{R}^3$ *is 3.*

1.1.5 Matrix algebra

Definition 1.1.7. *A matrix is a rectangular array of elements*

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

The above matrix has m rows and n columns. We also say it is a $m \times n$ matrix.

Let A, B be $m \times n$ matrices. We say $A = B$ if $a_{ij} = b_{ij}$ for any i, j. We define $A + B$ to be a $m \times n$ matrix such that

$$(A + B)_{ij} = a_{ij} + b_{ij}.$$

Similarly, for a scalar c, we define cA to be a $m \times n$ matrix such that

$$(cA)_{ij} = ca_{ij}.$$

Now let A be $m \times n$ and B be $n \times k$ matrices. We define AB to be a $m \times k$ matrix such that

$$(AB)_{ij} = \sum_{l=1}^{n} a_{il} b_{lj}.$$

Matrix addition and multiplication enjoy the usual properties of addition and multiplication such as commutatitivity, associativity and distributivity. We won't go into the details.

Finally, let A be a $m \times n$ matrix. The transpose of A, denoted as A^T, is a $n \times m$ matrix such that

$$A^T_{ij} = A_{ji}.$$

The following is a frequently used result about the transpose of a product of matrices:

$$(AB)^T = B^T A^T.$$

1.1.6 Special matrices

A $n \times n$ matrix A is upper triangular if all the entries below the diganoal are 0: $A_{ij} = 0$ for all $i > j$. We say $A_{n \times n}$ is lower triangular if all the entries above the diagonal are 0: $A_{ij} = 0$ for all $j > i$.

If only the diagonal entries of A is non-zero then we say A is a diagonal matrix.

A diagonal matrix is a scalar matrix if all the diagonal entries are equal.

A scalar matrix is an identity matrix, denoted as I, if all the diagonal entries are 1s.

A matrix A is symmetric if $A = A^T$.

1.1.7 System of linear equations

A system of m linear equations and n unknowns are of the form $Ax = b$ where A is a $m \times n$ matrix and b is a vector in $\mathbb{R}^m$. The standard way to solve the system is to reduce the augmented matrix $[A|b]$ to row echelon form or reduced row echelon form. Roughly speaking, a *row echelon form* transforms A into a matrix that is as close to upper triangular as possible while a *reduced row echelon form* transforms A into a matrix that is as close to the identity matrix as possible.

The operations that one can perform on a matrix to reduce it to echelon forms are: swapping rows, multiplying a row with a scalar, adding two rows and replace the sum with one of the rows.

We say a system is *inconsistent* if it has no solution and is *consistent* otherwise. If $m < n$ the system is *under-determined* and if $m > n$ it is *over-determined*. An under-determined system always has at least one solution (and usually infinitely many solutions) while an over-determined system usually is inconsistent. If $m = n$ it is of interest to see if there is only one solution (solution is unique). Also note that the system $Ax = 0$ is always consistent (no matter the size of A) since $x = 0$ is always a trivial solution.

1.2 Rank of a matrix and determinants

1.2.1 Introduction

In studying a linear system of equations $Ax = b$, it is clear that the properties of A play a crucial role in the existence of uniqueness of the solution. We have mentioned two in the last section: over-determined ($m > n$) and under-determined ($m < n$). These tell us whether we have more (or less) equations than the unknowns. Note, however that not all equations tell us something “new". For example, the following are essentially 1 equation:

$$\begin{aligned} x + y &= 1 \\ 2x + 2y &= 2. \end{aligned}$$

If we look at the matrix $\begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$ superficially, we may conclude that this system is neither over- nor under-determined since $m = n = 2$. Yet this system is under-determined because only the second equation is redundant. We will see that in stead of the size of the rows, a better way to capture how many equations are “relevant" is the rank of a matrix. By comparing the rank with the number of the unknowns, we will be able to tell whether we can expect the existence of a solution and whether it is unique.

An alternative way to investigate the existence and uniqueness of the solution is via the determinant of a matrix, which is only defined if a matrix is square.

1.2.2 Rank of a matrix

Definition 1.2.1. *The rank of a $m \times n$ matrix A is the maximum number of linearly independent row vectorsi n A.*

Remark: Since elementary row operations do not change the linear structure among the rows, the rank of a matrix is equal to the number of non-zero rows in its row echelon form or reduced row echelon form. This is how we find the rank of a matrix in practice.

Example 1.2.2. *After row reduction, the matrix*

$$A = \begin{bmatrix} 1 & 1 & -1 & 3 \\ 2 & -2 & 6 & 8 \\ 3 & 5 & -7 & 8 \end{bmatrix}$$

is reduced to

$$\tilde{A} = \begin{bmatrix} 1 & 1 & -1 & 3 \\ 0 & 1 & -2 & -1/2 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Thus rank(A) = 2.

In the previous example, suppose we want to solve the system $Ax = b$. After reduction, the augmented matrix $[A|b]$ is reduced to $[\tilde{A}|\tilde{b}]$ where $\tilde{A}$ is given above. It is clear that if the the system is consistent if and only if the last entry of $\tilde{b}$ is zero. We capture this fact by saying a system $Ax = b$ is consistent if and only if $rank(A) = rank([A|b])$.

On the other hand if the system $Ax = b$ is consistent then in the above example, it is clear that only the first two equations are relevant. There are 4 unknowns in the system. Thus we have the freedom to choose 2 of them to be arbitrary values (usually the last two unknowns by convention). If we let r denote the rank of the matrix then $n-r$ is the number of free unknowns we can assign arbitrary values to.

1.2.3 Determinants

Roughly speaking, determinant is a mapping from the space of $n \times n$ matrices to the real line. To give its definition we need to introduce some notation. For a $n \times n$ matrix A, we denote A^{ij} to be the $n-1 \times n-1$ matrix obtained by deleting the ith row and jth column. The determinant of A is defined inductively as followed:

a. $det(A) = a_{11}a_{22} - a_{12}a_{21}$ for $A_{2\times 2}$.

b. Generally, for a fixed j,

$$det(A) = \sum_{i=1}^{m} (-1)^{i+j} det(A^{ij}).$$

This is expansion along the j-th column. Also for a fixed i,

$$det(A) = \sum_{j=1}^{n} (-1)^{i+j} det(A^{ij}).$$

This is expansion along the i-th row. It is a remarkable fact that we obtain a single number, called the determinant of A, for expansion along any row or any column.

Example 1.2.3. *Expanding along the first row:*

$$det(\begin{bmatrix} 2 & 4 & 7 \\ 6 & 0 & 3 \\ 1 & 5 & 3 \end{bmatrix}) = 2(0 \times 3 - 3 \times 5) - 4(6 \times 3 - 3 \times 1) + 7(6 \times 5 - 0 \times 1) = 120.$$

1.2.4 Properties of determinants

We list here some basic properties of determinants:

a. $det(A) = det(A^T)$.

This is easy to see for 2×2 matrices. For 3×3 and higher it is easy to see that expanding along a row of A is the same as expanding along a column of A^T and the resulting matrices after deleting specific row and column of A and A^T are in fact transposes of each other. Thus the result is true by induction.

b. $det(AB) = det(A)det(B)$.

c. Let A' be the resulting matrix after swapping two rows (columns) of A. Then $det(A') = -det(A)$.

d. Let A' be the result of multiplying a row (column) of A with a constant c. Then $det(A') = cdet(A)$.

e. Let A' be the result of multiplying a row of A wit a constant, add it to another row and replace that row. Then $det(A') = det(A)$.

Remarks: c, d, e can be derived from b by represnting the elementary row operations with action by a matrix. That is we can show $A' = EA$ for some matrix E. It is easy to show that the matrix E has the expected determinant (e.g. $det(E) = -1$ in statement c).

f. If A has a zero row (column) then $det(A) = 0$. This is easy to see by expanding along the zero row or column.

g. If A have two identical rows or columns then $det(A) = 0$. This follows from e and f.

h. If A is triangular then $\det(A) = \prod_{i=1}^{n} a_{ii}$. This can be seen by expanding along the first column (if A is upper triangular) or the first row (if A is lower triangular).

1.3 Inverse of a Matrix and Cramer's Rule

1.3.1 Inverse of a matrix

We have seen that a $m \times n$ matrix A is a *linear* mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$. When $m = n$ the linear map A might be one-to-one and onto. That is it might have an inverse, denoted as A^{-1}. From its definition as an inverse, A^{-1} satisfies

$$AA^{-1} = A^{-1}A = I.$$

Conversely, if there exists a matrix B such that $AB = BA = I$ then B is the inverse of A. In fact, we only need either $AB = I$ or $BA = I$ to conclude $B = A^{-1}$ (this fact needs some explaination which we'll skip).

Example 1.3.1.

$$\begin{aligned} A &= \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \\ A^{-1} &= \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}. \end{aligned}$$

If A^{-1} exists we say A is *non-singular*. Otherwise we sa A is *singular.*

1.3.2 Some properties of inverse

a. The inverse is unique. That is if there are B, C such that $BA = CA = I$ then $B = C$.

Reason: The above equation implies $(B - C)A = 0$. Since A^{-1} exists (from either the existence of B or C), multipliying both sides with A^{-1} on the right gives $B - C = 0$ or $B = C$.

b. $(AB)^{-1} = B^{-1}A^{-1}$.

Reason: Clearly $B^{-1}A^{-1}AB = I$.

c. $(A^{-1})^{-1} = A$.

Reason: From definition of the inverse.

d. $(A^T)^{-1} = (A^{-1})^T$.

Reason: $(A^{-1})^T A^T = (AA^{-1})^T = I^T = I$.

1.3.3 Finding the inverse

Let A be a $n \times n$ matrix. Let $C_{ij} = (-1)^{i+j} M_{ij}$ where M_{ij} is the determinant of the $(n-1) \times (n-1)$ submatrix obtained by deleting the ith row and jth column of A. Denote

$$adjA = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & & & \vdots \\ C_{n1} & C_{n2} & \cdots & C_{nn} \end{bmatrix}^T = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{1n} \\ C_{12} & C_{22} & \cdots & C_{2n} \\ \vdots & & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}.$$

Suppose $det(A) \neq 0$ then

$$A^{-1} = \frac{1}{det(A)} adj A.$$

Note that this result also implies that A^{-1} exits if and only if $det(A) \neq 0$. To show that it is true, we need to show

$$A \frac{1}{det(A)} adj A = I,$$

or equivalently

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{1n} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{1n} \\ C_{12} & C_{22} & \cdots & C_{2n} \\ \vdots & & & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix} = \begin{bmatrix} det(A) & 0 & \cdots & 0 \\ 0 & det(A) & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & det(A) \end{bmatrix}.$$

That is we need to show

$$\sum_{j=1}^{n} a_{ij} C_{ij} = det(A), \forall i$$

and

$$\sum_{j=1}^{n} a_{ij} C_{kj} = 0, \forall i \neq k.$$

The first equation is exactly the definition of $det(A)$, expanding along the ith row. The second equation can be explained as followed. For simplicity let's say we want to show

$$\sum_{j=1}^{n} a_{1j} C_{2j} = 0.$$

Imagine that we replace the second row of A with the first row and call the resulting matrix $\tilde{A}$. If we calculate $\tilde{C}_{2j}$ of $\tilde{A}$ it is the same as C_{2j} of $A, \forall j$ since C_{2j} is obtained exactly by deleting the second row of A and the jth column (so the second row's information is irrelevant). Now since $\tilde{A}$ has two identical rows, its determinant is 0. We thus have

$$\sum_{j=1}^{n} a_{1j} C_{2j} = 0 = \sum_{j=1}^{n} \tilde{a}_{2j} \tilde{C}_{2j} = det(\tilde{A}) = 0.$$

In practice, an easier way to find the inverse if to look at the $n \times 2n$ augmented matrix $[A|I]$ and perform Gaussian elimination until we obtain $[I|B]$. The matrix B is exactly A^{-1}. The reason is because Gaussian elimination is equivalent to being multiplied on the left by a matrix. That is there must be a matrix B such that

$$B[A|I] = [I|B].$$

But then $BA = I$ or $B = A^{-1}$.

1.3.4 Solving system of equations with inverse

Consider the $n \times n$ system $Ax = b$. If A^{-1} exists then the system has a unique solution: $x = A^{-1}b$. In particular, the homogenous system $Ax = 0$ has a unique and trivial solution $x = 0$ if and only if A is invertible. Conversely, $Ax = 0$ has a non-zero solution if and only if A is non-invertible. Also note the aside fact that finding the inverse to solve for a system is much more costly in terms of computation than using Gaussian elimination.

1.3.5 Cramer's rule

Consider the $\nu times n$ system $Ax = b$ and suppose that A is invertible. Then we can represent $x = A^{-1}b$. Note however that we do not have a representation for x_1 (or $x_2, \cdots, x_n$) individually using the inverse approach. If for some reason our interest is to only find a particular x_i in the whole vector x, Cramer's rule gives us the way to do so.

Theorem 1.3.2. *Cramer's rule Let $A^i, i = 1, \cdots, n$ be the matrix that resulted from replacing the ith column with the b column. Then*

$$x_i = \frac{det(A^i)}{det(A)}.$$

Example 1.3.3. *Consider the system*

$$\begin{bmatrix} 3 & 2 & 1 \\ 1 & -1 & 3 \\ 5 & 4 & -2 \end{bmatrix} = \begin{bmatrix} 7 \\ 3 \,. \\ 1 \end{bmatrix}.$$

Then

$$A^1 = \begin{bmatrix} 7 & 2 & 1 \\ 3 & -1 & 3 \\ 1 & 4 & -2 \end{bmatrix}$$

and

$$x_!1 = \frac{-39}{13} = -3.$$

1.4 Eigenvalues and Eigenvectors

1.4.1 Introduction

Let A be a $n \times n$ matrix. We have mentioned that A is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^n$. It is of interest and importance to look at the vectors that A "fixed". That is to find $\mathbf{v} \in \mathbb{R}^n, \mathbf{v} \neq 0$ so that

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

for some constant λ. In the above equation,the action of A on v essentially does not change v, aside from a constant factor. Intuitively, one can believe that $\mathbf{v}$ captures some essential characteristic of A because A does not change $\mathbf{v}$. We say λ is an eigenvalue and $\mathbf{v}$ is an eigenvector associated with the eigenvalue λ. (Note: there can be more than one eigenvectors associated with one eigenvalue). It turns out that under certain conditions, we can re-construct A from its complete set of eigenvalues and eigenvectors. This is the diagonalization result and it shows how eigenvalues and eigenvectors give information about the original matrix.

1.4.2 Finding eigenvalues

The system $A\mathbf{v} = \lambda\mathbf{v}$ is equivalent to $(A - \lambda I)\mathbf{v} = 0$. The requirement that it has a non-trivial solution $\mathbf{v}$ is equivalent to requiring that $A - \lambda I$ is singular or

$$det(A - \lambda I) = 0.$$

This is the equation for the eigenvalues of A. The LHS is a polynomial in λ, usually of degree n. Since eigenvalues are the roots of a polynomial, it can be repeated, real or complex-valued.

Example 1.4.1. *An example of real, repeated eigenvalues:*

$$\begin{aligned} A &= \begin{bmatrix} 3 & 4 \\ -1 & 7 \end{bmatrix} \\ det(A - \lambda I) &= (\lambda - 5)^2 = 0 \text{ iff } \lambda = 5. \end{aligned}$$

Example 1.4.2. *An example of repeated complex eigenvalues:*

$$\begin{aligned} A &= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \\ det(A - \lambda I) &= \lambda^2 + 1 = 0 \text{ iff } \lambda = \pm i. \end{aligned}$$

1.4.3 Finding eigenvectors

Once we find a specific value of λ, we can proceed to sovle the system $(A - \lambda I)\mathbf{v} = 0$. Since $det(A - \lambda I) = 0$ this system has infinitely many solutions by construction. As a consequence eigenvector is *not unique*. That is if $\mathbf{v}$ is an eigenvector then $c\mathbf{v}$ is also an eigenvector for any $c \neq 0$ becaus

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

implies

$$A(c\mathbf{v}) = \lambda(c\mathbf{v}).$$

What we present as eigenvectors are the basis vectors of the solution space. Note that there are $n - r$ basis vectors of the solution space (where r is the rank of the matrix $A - \lambda I$).

Example 1.4.3.

$$A = \begin{bmatrix} 3 & 4 \\ -1 & 7 \end{bmatrix}.$$

We have showed $\lambda = 5$ is the eigenvalue.

$$A - 5I = \begin{bmatrix} -2 & 4 \\ -1 & 2 \end{bmatrix}.$$

And thus an eigenvector would be $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Example 1.4.4.

$$A = \begin{bmatrix} 9 & 1 & 1 \\ 1 & 9 & 1 \\ 1 & 1 & 9 \end{bmatrix}.$$

You can easily check that $\lambda = 8$ is an eigenvalue.

$$A - 8I = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Here $A - 8I$ is of rank 1 so we would expect two (linearly independent) eigenvectors associated with eigenvalue 8. Indeed if $\mathbf{v} = [v_1, v_2, v_3]^T$ then the only equation $\mathbf{v}$ satisfies is $v_1 + v_2 + v_3 = 0$. Thus $\mathbf{v}_1 = [-1, 0, 1]^T$ and $\mathbf{v}_2 = [-1, 1, 0]^T$ would be 2 eigenvectors associated with eigenvalue 8.

1.4.4 Some properties

Here we list some properties about eigenvalues and eigenvectors that can be useful in their computation.

a. If $a + ib$ is an eigenvalue of a (real entry) matrix A then $a - ib$ is also an eigenvalue of A. If $\mathbf{v}$ is an eigenvector associated with $a + ib$ then $\bar{\mathbf{v}}$ (the complex conjugate of $\mathbf{v}$ taken entriwise) is an eigenvector associated with $a - ib$.

Reason: This comes from an algebra fact that complex roots of a real entry polynomial come in conjugate pairs. Since A has real entry,

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

implies

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

b. A matrix A is singular iff it has an eigenvalue 0.

Reason: An eigenvalue equalling 0 is equivalent to $det(A) = 0$ iff A is singular.

c. Conversely, a matrix A is non-singular (invertible) iff all of its eigenvalues are non-zero. In this case, if λ is an eigenvalue of A with eigenvector $\mathbf{v}$ then $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} with the same eigenvector $\mathbf{v}$.

Reason: $A\mathbf{v} = \lambda\mathbf{v}$ implies $\mathbf{v} = \lambda A^{-1}\mathbf{v}$ or $\frac{1}{\lambda}\mathbf{v} = A^{-1}\mathbf{v}$.

d. If A is triangular then its eigenvalues are exactly its diagonal entries.

Reason: If A is triangular then $A-\lambda I$ is also triangular and $det(A-\lambda I) = \prod_{i=1}^{n}(a_{ii} - \lambda)$. Clearly the roots of this polynomial are $\lambda_{ii}, i = 1, \cdots, n$.

1.5 Orthogonal Matrices and Diagonalization of Matrices

1.5.1 Introduction

The notion of orthogonality is of fundamental importance in many areas. The motivation is quite natural: two vectors $\mathbf{v}_1, \mathbf{v}_2$ (in $\mathbb{R}^2$ or $\mathbb{R}^3$) are orthogonal if $\mathbf{v}_1^T\mathbf{v}_2 = 0$. This can be verified geometrically by just drawing out $\mathbf{v}_1, \mathbf{v}_2$. For example, $[-1, 1]$ and $[1, 1]$ are clearly orthogonal.

More abstractly, we can define for any two vectors $\mathbf{v}_1, \mathbf{v}_2$ in $\mathbb{R}^n$ to be orthogonal if $\mathbf{v}_1^T\mathbf{v}_2 = 0$. Note that we lose the geometrical intuition for $n > 3$. On the otherhand, some fundamental properties of orthogonal vectors are still applicable in higher dimension, such as orthogonal vectors are linearly independent. We can also discuss orthogonality of two functions, as we will see in the chapter on Fourier series. It is of interest then to see discuss properties related to orthogonality of vectors and matrices, onf of which is the diagonalization of matrices.

1.5.2 Orthogonal eigenvectors

Theorem 1.5.1. *Let A be an $n \times n$* **symmetric** *matrix. The eigenvectors corresponding to* **distinct** *eigenvalues are orthogonal.*

Remark:

a.First it is only true for symmetric matrix. For example, the matrix

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$$

has distinct eigenvalues $1, 3$ and two eigenvectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, corresponding to $\lambda = 1$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ corresponding to $\lambda = 3$ but they are not orthogonal.

Second it is only true for eigenvectors corresponding to distinct eigenvalues. For example the matrix

$$A = \begin{bmatrix} 7 & 4 & -4 \\ 4 & -8 & -1 \\ -4 & -1 & -8 \end{bmatrix}$$

has eigenvalues $\lambda_1 = -9$ and $\lambda_2 = 9$. There are two eigenvectors corresponding to $\lambda_1 = -9$

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}; \mathbf{v}_2 = \begin{bmatrix} 1 \\ -4 \\ 0 \end{bmatrix}.$$

The eigenvector corresponding to $\lambda_2 = 0$ is

$$\mathbf{v}_3 = \begin{bmatrix} 4 \\ 1 \\ -1 \end{bmatrix}.$$

Observe that $\mathbf{v}_1, \mathbf{v}_2$ are orthogonal to $\mathbf{v}_3$ but $\mathbf{v}_1$ is not orthogonal to $\mathbf{v}_2$.
Proof. Let A be symmetric and $\lambda_1 \neq \lambda_2$ be two distinct eigenvalues of A. Since they are distinct both of them cannot be zero. WLOG suppose $\lambda_1 \neq 0$. Then

$$\begin{aligned} \mathbf{v}_1^T \mathbf{v}_2 &= (\frac{1}{\lambda_1}(A\mathbf{v}_1)^T)\mathbf{v}_2 \\ &= \frac{1}{\lambda_1} v_1^T A \mathbf{v}_2 = \frac{\lambda_2}{\lambda_1}\mathbf{v}_1^T \mathbf{v}_2. \end{aligned}$$

That is

$$(\frac{\lambda_2}{\lambda_1} - 1)\mathbf{v}_1^T \mathbf{v}_2 = 0.$$

The conclusion follows since $\frac{\lambda_2}{\lambda_1} \neq 1$ because they are distinct.

Next we cover a result for eigenvalues of symmetric matrix.

Theorem 1.5.2. *Let A be an $n \times n$* **symmetric** *matrix. Then all of its eigenvalues are real.*

Proof. Let λ be an eigenvalue of A with associated eigenvector $\mathbf{v}$. WLOG $\lambda \neq 0$ since then it is real. We know that $\bar{\lambda}$ is also an eigenvalue (if λ is real then $\bar{\lambda} = \lambda$ and if it is complex then this is the result from the previous section) with associated eigenvector $\bar{\mathbf{v}}$. Then

$$\bar{\mathbf{v}}^T \mathbf{v} = \frac{1}{\bar{\lambda}}(A\bar{\mathbf{v}})^T \mathbf{v} = \frac{1}{\bar{\lambda}}\bar{\mathbf{v}}^T A \mathbf{v} = \frac{\lambda}{\bar{\lambda}}\bar{\mathbf{v}}^T \mathbf{v}.$$

Now $\mathbf{v}^T \mathbf{v}$ is real and non-zero since $\mathbf{v}$ is not a zero vector. Therefore $\frac{\lambda}{\bar{\lambda}}$ is real. Thus $\lambda = \bar{\lambda}$ is a real number.

1.5.3 Orthogonal matrices

Definition 1.5.3. *A set of vectors* $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ *is orthonormal if* $\mathbf{v}_i^T\mathbf{v}_i = 1$ *and* $\mathbf{v}_i^T\mathbf{v}_j = 0, i \neq j$ *for all* $1 \leq i, j \leq n$*.*

Definition 1.5.4. *A matrix* A *is orthogonal if* $A^TA = I$*. Equivalently,* A *is orthogonal if* $A^T = A^{-1}$*. The last equivlanet definition is if the columns of* A *form an orthonormal set.*

Remark: It is worthwhile to note that an orthogonal matrix is automatically invertible (its inverse is its transpose).

1.5.4 Orthogonal matrix formed by eigenvectors

Let A be a symmetric matrix. We will accept without proof the fact that A has n independent eigenvectors. So we can form a $n \times n$ matrix P whose columns are eigenvectors of A. Since the eigenvectors corresponding to distinct eigenvalues are orthogonal, P is “almost" orthogonal. If we can make the eigenvectors corresponding to the same eigenvalue orthogonal and normalize all columns to length 1 then P will indeed be orthogonal. In fact this can always be done and the procedure used is referred to as Gram-Schimidt process, see e.g. here.

In practice, we do not have to employ Gram-Schmidt but more or less “guess" at what we need. The following example will illustrate.

Example 1.5.5. *Reconsider the matrix*

$$A = \begin{bmatrix} 7 & 4 & -4 \\ 4 & -8 & -1 \\ -4 & -1 & -8 \end{bmatrix}$$

which has eigenvalues $\lambda_1 = -9$ *and* $\lambda_2 = 9$*. There are two eigenvectors corresponding to* $\lambda_1 = -9$

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}; \mathbf{v}_2 = \begin{bmatrix} 1 \\ -4 \\ 0 \end{bmatrix}.$$

The eigenvector corresponding to $\lambda_2 = 0$ *is*

$$\mathbf{v}_3 = \begin{bmatrix} 4 \\ 1 \\ -1 \end{bmatrix}.$$

We mentioned that $\mathbf{v}_1, \mathbf{v}_2$ *are orthogonal to* $\mathbf{v}_3$ *but* $\mathbf{v}_1$ *is not orthogonal to* $\mathbf{v}_2$*. It is easy to choose* $\tilde{\mathbf{v}}_2$ *orthogonal to* v_1 *so that* $\tilde{\mathbf{v}}_2$ *is still an eigenvector indepndent of* $\mathbf{v}_1$ *using Gram-Schmidt. Nevertheless, what we'll do is look at how we find* $\mathbf{v}_1, \mathbf{v}_2$*, namely from the system*

$A + 9I = 0$. *The reduced form is*

$$\begin{bmatrix} 1 & 1/4 & -1/4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus if $\mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ *is an eigenvector coressponding to* $\lambda = 9$ *then* $4x + y - z$ *= 0. We want to choose two solutions for* x, y, z *that are orthogonal to each other. Utilizing* $0, 1$ *helps. So the first choice is* $\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ *and the second choice is* $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$.

Thus

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \end{bmatrix}.$$

1.5.5 Diagonalization of matrices

Diagonal matrices are very simple to work with. Specifically, the product C of two diagonal matrices A, B is diagonal, with diagonal entries equal to the product of the corresponding diagonal entries of A, B. That is $C_{ii} = A_{ii}B_{ii}, \forall i$. Consequently, if A is diagonal, A^n is also diagonal and $(A^n)_{ii} = (A_{ii})^n$.

Thus given a square matrix A, we would like to relate it to a diagonal matrix via the notion of simlarity. Two matrices A, B are similar if there exists an invertible matrix P so that $A = P^{-1}AP$. One motivation for calling such relation simlarity is because two similar matrices have the same eigenvalues (can you see why?). Thus given a matrix A, we would like to see if it is similar to a diagonal matrix D. By the above remark, if this can be done, the diagonal entries of D are exactly the eigenvalues of A. We say in this case that A can be diagonalized and the process of finding such matrix P is called the diagonalization process.

It turns out that not all matrices can be diagonalized. This has to do with the fact that not all matrices has a full set of eigenvectors. If a matrix A does have a full set of eigenvectors, an orthogonal matrix P can be formed using these eigenvectors similar to the above example. Moreover, in this case

$$A = PDP^{-1} = PDP^T,$$

where the column of P is the corresponding eigenvector to the eigenvalue on the diagonal of D.

To see that this is the case, we will show that

$$P^T AP = D.$$

Indeed, $P = [\mathbf{v}_1 \mathbf{v}_2 \cdots \mathbf{v}_n]$ where $\mathbf{v}_i$ is the ith eigenvector of A. We see that

$$AP = [A\mathbf{v}_1 A\mathbf{v}_2 \cdots A\mathbf{v}_n] = [\lambda_1 \mathbf{v}_1 \lambda_2 \mathbf{v}_2 \cdots \lambda_n \mathbf{v}_n].$$

Thus

$$P^T AP = \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} [\lambda_1 \mathbf{v}_1 \lambda_2 \mathbf{v}_2 \cdots \lambda_n \mathbf{v}_n] = D,$$

where $D_{ij} = \mathbf{v}_i \lambda_i \mathbf{v}_j = 0$ if $i \neq j$ and equals λ_i if $i = j$ by the orthogonality of $\mathbf{v}_i, \mathbf{v}_j$.

Example 1.5.6. *Reconsider the matrix*

$$A = \begin{bmatrix} 7 & 4 & -4 \\ 4 & -8 & -1 \\ -4 & -1 & -8 \end{bmatrix}.$$

We have showed that an orthognal matrix formed by eigenvectors of A is

$$P = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \end{bmatrix}.$$

A has eigenvalues $\lambda_1 = -9$ (with multiplicity 2) and $\lambda_2 = 9$. One can verify that

$$\begin{bmatrix} 7 & 4 & -4 \\ 4 & -8 & -1 \\ -4 & -1 & -8 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} -9 & 0 & 0 \\ 0 & -9 & 0 \\ 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \end{bmatrix}$$

Chapter 2

The Laplace transform

2.1 Definition of the Laplace transform

2.1.1 Introduction

In mathematics, there is an abstract notion of an operator: a mapping from a function space to a function space. Let $\mathcal{L}$ denote such an operator. We say $\mathcal{L}$ is a *linear* operator if it satisfies

$$\mathcal{L}(af+bg) = a\mathcal{L}(f) + b\mathcal{L}(g),$$

where a, b are real numbers and f, g are functions. From elementary calculus, the differentiation $\frac{d}{dt}$ and integration $\int dt$ are examples of linear operators.

The Laplace transform that we will study in this chapter is a linear operator that has many important applications, e.g. (but not limited to) in solving ordinary differential equations (ODEs) and partial differential equations (PDEs).

2.1.2 Definition and basic remarks

Definition 2.1.1. *Let f be a function defined for $t \geq 0$. Then the integral*

$$\mathcal{L}(f(t)) = \int_0^\infty e^{-st} f(t)dt$$

is said to be the Laplace transform of f, if the integral **exists.**

Remarks:

a. While f is a function of t, the Laplace transform of f is a function of s.

b. An arbitrary function f does not necessarily have a Laplace transform (if the integral does not converge). E.g. $f(t) = e^{t^2}$ does not have a Laplace transform.

c. Also because of the issue of convergence of the integral, the Laplace transform of an arbitrary function f may not be defined for all values of s. E.g. if $f(t) = e^{at}$, a a constant then $\mathcal{L}(f)$ is only defined for $s > a$. When the domain of s is not explicitly stated, we understand that we only work with the values of s such that $\mathcal{L}(f)$ is well-defined.

d. The Laplace transform is indeed a linear operator:

$$\begin{aligned}\mathcal{L}(af+bg) &= \int_0^\infty e^{-st}[af(t)+bg(t)]dt \\ &= a\int_0^\infty e^{-st}f(t) + b\int_0^\infty e^{-st}g(t) = a\mathcal{L}(f) + b\mathcal{L}(g).\end{aligned}$$

You can see the linearity of the Laplace transfrom comes from the linearity property of the integral operator.

2.1.3 Connection with power series

The Laplace transform can be viewed as a continuous anolog of a power series. Specifically, if $a(n)$ is a discrete function of a positive integer n, the power series associated with $a(n)$ is

$$\sum_{n=0}^{\infty} a(n)x^n,$$

where x is a real variable. For example, the power series representation of e^x is $\sum_{n=0}^{\infty} \frac{x^n}{n!}$. That is $a(n) = \frac{1}{n!}$ in this particular series.

Suppose we formally replace the discrete variable n with a continous variable t. Then the summation over n becomes integration over t. The function $a(n)$ becomes a function $f(t)$ (which is a continuous extension of $a(n)$ to the positve half of the real line). Thus a continuous version of the power series becomes

$$\int_0^\infty f(t)x^t dt.$$

Replacing x with e^{-s} gives

$$\int_0^\infty f(t)x^t dt = \int_0^\infty f(t)(e^{-s})^t dt = \int_0^\infty f(t)e^{-st}dt.$$

Thus, the Laplace transform can be viewed as a continuous version of a power series where the discrete variable n is replaced by the continuous variable t, the discrete coefficients $a(n)$ replaced by $f(t)$ and x replaced by e^{-s}.

2.1.4 The Laplace transform of basic functions

The following are some formulas that you should memorize for the purpose of this course. It is also recommended that you go over their derivation at least once.

$$\begin{aligned}
\mathcal{L}(1) &= \frac{1}{s} \\
\mathcal{L}(t^n) &= \frac{n!}{s^{n+1}}, n = 1, 2, 3 \cdots \\
\mathcal{L}(e^{at}) &= \frac{1}{s-a} \\
\mathcal{L}(\sin(kt)) &= \frac{k}{s^2+k^2} \\
\mathcal{L}(\cos(kt)) &= \frac{s}{s^2+k^2} \\
\mathcal{L}(\sinh(kt)) &= \frac{k}{s^2-k^2} \\
\mathcal{L}(\cosh(kt)) &= \frac{s}{s^2-k^2}.
\end{aligned}$$

Remarks:

a. $\mathcal{L}(1)$ can be seen as a special case of $\mathcal{L}(t^n)$ with $n = 0$ from the fact that $0! = 1$.

b. The Euler formula states $e^{ix} = \cos x + i \sin x$. Hence $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$ and $\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$. Also from definition, $\cosh(x) = \frac{e^x+e^{-x}}{2}, \sinh(x) = \frac{e^x-e^{-x}}{2}$. These identities, coupled with the formula for $\mathcal{L}(e^{at})$ can be used to derive the Laplace transforms of $\sin, \cos, \sinh, \cosh$.

Example 2.1.2. *Laplace transform of a piece-wise defined function Let*

$$\begin{aligned}
f(t) &= 1, \text{ if } 0 \le t \le 3 \\
&\quad\; 2, \text{ if } t \ge 3.
\end{aligned}$$

Then

$$\begin{aligned}
\mathcal{L}(f) &= \int_0^3 e^{-st} dt + 2 \int_3^\infty e^{-st} dt \\
&= \frac{1}{s}(1 - e^{-3s}) - \frac{2}{s} e^{-3s} \\
&= \frac{1}{s} - \frac{3e^{-3s}}{s}.
\end{aligned}$$

2.2 Inverse Laplace Transform and Transform of Derivatives

2.2.1 The inverse problem

Given a function $g(s)$, it is natural to ask if there is a function f such that $\mathcal{L}(f) = g$. That is g is the Laplace transform of f; or f is the *inverse* Laplace transform of g. We denote

$$f(t) = \mathcal{L}^{-1}(g(s)).$$

For example, if $g(s) = \frac{1}{s-a}$ then $f(t) = e^{at}$ is such that $\mathcal{L}(f) = g$ and $f = \mathcal{L}^{-1}(g)$. It is worthwhile to compare $\mathcal{L}$ and $\mathcal{L}^{-1}$ with differentiation and integration in elementary calculus.

Remarks:

a. There is a general formula for the inverse Laplace transform, called the Mellin's inverse formula, see e.g. here. It requires integration in complex coordinates and hence the formula is beyond the scope of this course. We will rely on the formulas for Laplace transform of basic functions and basic manipulations to find inverse Laplace transforms.

b. The inverse Laplace transform is not unique point wise. For example

$$f_1(t) = 0, t \geq 0$$

and

$$\begin{aligned} f_2(t) &= 0, t \neq 1 \\ &= 1, t = 1 \end{aligned}$$

have the same Laplace transforms, that is $\mathcal{L}(f_1) = \mathcal{L}(f_2) = 0$. However, if we know apriori that the inverse Laplace transform is continuous then it is also unique pointwise.

c. The inverse Laplace transform is a *linear* operator. That is

$$\mathcal{L}^{-1}(ag_1(s) + bg_2(s)) = a\mathcal{L}^{-1}(g_1(s)) + b\mathcal{L}^{-1}(g_2(s)).$$

(It follows from the linearity of the Laplace transform). We will utilize this property to find the inverse Laplace transforms.

2.2.2 Examples

Example 2.2.1.

$$\mathcal{L}^{-1}\Big(\frac{1}{s^2} - \frac{1}{s} + \frac{1}{s-2}\Big) = t - 1 + e^{2t}.$$

Example 2.2.2.

$$\begin{aligned} \mathcal{L}^{-1}\Big(\frac{s^2+6s+9}{(s-1)(s-2)(s+4)}\Big) &= \mathcal{L}^{-1}\Big(\frac{-\frac{16}{5}}{s-1} + \frac{\frac{25}{6}}{s-2} + \frac{\frac{1}{30}}{s+4}\Big) \\ &= -\frac{16}{5}e^{t} + \frac{25}{6}e^{2t} + \frac{1}{30}e^{-4t}. \end{aligned}$$

In the second example, we have utilized the technique of partial fraction decomposition. That is we find A, B, C such that

$$\frac{s^2+6s+9}{(s-1)(s-2)(s+4)} = \frac{A}{s-1} + \frac{B}{s-2} + \frac{C}{s+4}.$$

For a review of partial fraction decomposition, see e.g. here.

2.2.3 Laplace transform of the derivatives

Theorem 2.2.3. *Let f be a function defined on $[0,\infty)$ with sufficiently nice derivatives. Then*

$$\mathcal{L}(f^{(n)}) = s^n\mathcal{L}(f)(s) - s^{n-1}f(0) - s^{n-2}f'(0) - \cdots - f^{(n-1)}(0).$$

Example 2.2.4.

$$\begin{aligned}\mathcal{L}(f'(t)) &= s\mathcal{L}(f)(s) - f(0)\\ \mathcal{L}(f''(t)) &= s^2\mathcal{L}(f)(s) - sf(0) - f'(0).\end{aligned}$$

Remark:

a. The Laplace transform of derivatives of f (*of any order*) involves the Laplace transform of f (multiplied with a power of s) subtracting a polynomial in s. Note that the terms $f(0), f'(0), \cdots, f^{(n-1)}(0)$ are constants.

b. This makes the Laplace transform a suitable tool to solve *linear* ODEs with initial conditions. After the transform, the ODEs becomes an equation where the the Laplace transform $\mathcal{L}(f)(s)$ is *the term to solve for*. The linearity of the Laplace transform preserves the linearity of the equations. The initial conditions provide information for $f(0), f'(0), \cdots, f^{(n-1)}(0)$.

2.2.4 Solving linear ODEs

Example 2.2.5. *Solve*

$$\begin{aligned}y'(t) + 3y &= 13\sin(2t),\\ y(0) &= 6.\end{aligned}$$

Sol:

Denoting $Y(s) = \mathcal{L}(y)(s)$. Taking the Laplace transform on both sides yield

$$sY(s) - y(0) + 3Y(s) = 13\frac{2}{s^2+4}.$$

That is

$$Y(s) = \frac{6s^2+50}{(s+3)(s^2+4)}.$$

Using partial fraction yields

$$\frac{6s^2+50}{(s+3)(s^2+4)} = \frac{8}{s+3} + \frac{-2s+6}{s^2+4}.$$

Taking the inverse Laplace transform yields

$$y(t) = 8e^{-3t} - 2\cos(2t) + 3\sin(2t).$$

Example 2.2.6. *Solve*

$$\begin{aligned} y'' - 3y' + 2y &= e^{-4t}, \\ y(0) = 1, y'(0) &= 5. \end{aligned}$$

Sol:
Again denoting $Y(s) = \mathcal{L}(y)(s)$. Taking the Laplace transform on both sides yield

$$[s^2Y(s) - sy(0) - y'(0)] - 3[sY(s) - y(0)] + 2Y(s) = \frac{1}{s+4}.$$

That is

$$(s^2 - 3s + 2)Y(s) = \frac{1}{s+4} + s + 2.$$

Or

$$Y(s) = \frac{s^2 + 6s + 9}{(s^2 - 3s + 2)(s+4)} = \frac{s^2 + 6s + 9}{(s-1)(s-2)(s+4)}.$$

From Example (2.2.2) we have

$$y(t) = \mathcal{L}^{-1}\Big(\frac{s^2 + 6s + 9}{(s-1)(s-2)(s+4)}\Big) \quad = \quad -\frac{16}{5}e^t + \frac{25}{6}e^{2t} + \frac{1}{30}e^{-4t}.$$

2.2.5 Zero-input response and zero-state response

Consider an abstract linear ODE:

$$\begin{aligned} \sum_{i=0}^{n} a_i y^{(i)}(t) &= g(t) \\ y^i(0) &= c_i, i = 0, \cdots, n-1. \end{aligned}$$

Denoting $\mathcal{L}(y)(s) = Y(s)$ and $\mathcal{L}(g)(s) = G(s)$, from the examples in the previous section the Laplace transform of the ODE has the form

$$P(s)Y(s) + Q(s) = G(s),$$

where $P(s) = \sum_{i=0}^{n} a_i s^i$ and $Q(s)$ is a polynomial of degree at most $n-1$ whose coefficients come from the initial conditions $y^{(i)}(0), i = 0, \cdots, n-1$. Thus

$$y(t) = \mathcal{L}^{-1}\Big(\frac{-Q(s)}{P(s)}\Big) + \mathcal{L}^{-1}\Big(\frac{G(s)}{P(s)}\Big) := y_0(t) + y_1(t).$$

This is the abstract formulation of the solution of a linear ODE using Laplace transform. There are two special cases:

a. The input $g(t) = 0$. Then clearly $G(s) = 0$ and $y(t) = y_0(t)$. Thus we refer to $y_0(t)$ as the zero-input response.

b. The initial conditions $c_i = 0, i = 0, \cdots, n-1$. Then clearly $Q(s) = 0$ and $y(t) = y_1(t)$. Thus we refer to $y_1(t)$ as the zero-state response.

In general then, the solution to the initial value ODE is the sum of the zero-input response and the zero-state response.

2.3 Translation Theorems; Unit Step Function (Heaviside Function)

2.3.1 Introduction

The previous section on the Laplace transform of derivatives can be looked at as a study of a particular interaction of linear operators. Indeed, if we denote $\mathcal{D}^n$ as the n-th derivative operator then the previous result says that $\mathcal{L} \circ \mathcal{D}^n$ is also a linear operator, and

$$\mathcal{L} \circ \mathcal{D}^n(f(t)) = s^n \mathcal{L}(f)(s) - s^{n-1} f(0) - s^{n-2} f'(0) - \cdots - f^{(n-1)}(0).$$

(It is a quick but still interesting exercise to verify that $\mathcal{L} \circ \mathcal{D}^n$ is indeed linear).

Another interesting example of a linear operator is the translation operator. Let a be a fixed constant. We denote $\mathcal{T}_a$ to be the linear operator such that

$$\mathcal{T}_a(f(t)) := f(t-a).$$

Again, you can quickly verify that $\mathcal{T}_a$ is indeed linear. In this section we study the interaction between the inverse Laplace transform and the translation operator. That is we study the operator $\mathcal{L}^{-1} \circ \mathcal{T}_a$.

A last interesting example of a linear operator is as followed. Define the **unit step function** or the **Heaviside function** as followed:

$$\begin{aligned} \mathcal{U}(t) &= 1, t \geq 0 \\ &= 0, t < 0. \end{aligned}$$

With a slight abuse of notion, we denote $\mathcal{U}_a$ to be the operator:

$$\mathcal{U}_a(f(t)) := \mathcal{T}_a\Big(\mathcal{U}(t) f(t)\Big) = \mathcal{U}(t-a) f(t-a).$$

Note that the right hand side of the above is just a multiplication of two functions. Roughly speaking, $\mathcal{U}_a(f)$ has the effect of shifting f by a units to the right and "zero" out the part of f to the left of a. The last result we study in this section is the interaction between the Laplace transform and the operator $\mathcal{U}_a$. That is we study the operator $\mathcal{L} \circ \mathcal{U}_a$.

2.3.2 The operator $\mathcal{L}^{-1} \circ \mathcal{T}_a$: translation on the s-axis

Theorem 2.3.1. *Let a be a real number and suppose that $\mathcal{L}(f) = F(s)$ then*

$$\mathcal{L}(e^{at} f(t)) = F(s-a).$$

Equivalently,

$$\mathcal{L}^{-1}(F(s-a)) = e^{at} f(t).$$

Example 2.3.2.

$$\mathcal{L}(e^{5t}t^3) = \frac{3!}{(s-5)^4}$$

Example 2.3.3.

$$\begin{aligned}\mathcal{L}^{-1}\left(\frac{2s+5}{(s-3)^2}\right) &= \mathcal{L}^{-1}\left(\frac{2(s-3)+11}{(s-3)^2}\right)\\ &= \mathcal{L}^{-1}\left(\frac{2}{s-3}\right)+\mathcal{L}^{-1}\left(\frac{11}{(s-3)^2}\right)\\ &= 2e^{3t}+11e^{3t}t.\end{aligned}$$

Remark: Note how we would approach this example from the previous section using partial fraction. Using the result about translation it is much quicker to find the inverse Laplace transform of this example.

Example 2.3.4.

$$\begin{aligned}\mathcal{L}^{-1}\left(\frac{s}{s^2+4s+6}\right) &= \mathcal{L}^{-1}\left(\frac{s}{(s+2)^2+2}\right)\\ &= \mathcal{L}^{-1}\left(\frac{s+2}{(s+2)^2+2}\right)-\mathcal{L}^{-1}\left(\frac{2}{(s+2)^2+2}\right)\\ &= e^{-2t}\cos(\sqrt{2}t)-\sqrt{2}e^{-2t}\sin(\sqrt{2}t).\end{aligned}$$

Remark: Again note how we would not be able to handle this example using partial fraction and the result of the previous section, because s^2+4s+6 is irreducible and it is not of the form s^2+k^2.

2.3.3 The operator $\mathcal{L}\circ\mathcal{U}_a$: translation on the t-axis

Theorem 2.3.5. *Let a be a real number and suppose that $\mathcal{L}(f)=F(s)$ then*

$$\mathcal{L}(f(t-a)\mathcal{U}(t-a)) = e^{-as}F(s).$$

Equivalently,

$$\mathcal{L}^{-1}(e^{-as}F(s)) = f(t-a)\mathcal{U}(t-a).$$

Example 2.3.6.

$$\mathcal{L}^{-1}\left(\frac{1}{s-4}e^{-2s}\right) = e^{4(t-2)}\mathcal{U}(t-2).$$

Remark: The thinking process for the above example is as followed: first we need to identify the function f such that $\mathcal{L}(f) = \frac{1}{s-4}$. This already is an application of the first translation theorem; which gives $f(t) = e^{4t}$. Then we apply the second translation theorem to get the conclusion.

Example 2.3.7.

$$\begin{aligned}\mathcal{L}(t^2\mathcal{U}(t-2)) &= \mathcal{L}([(t-2)^2 + 2(t-2) + 4]\mathcal{U}(t-2)) \\ &= e^{-2s}\Big[\frac{2!}{s^3} + \frac{2}{s^2} + \frac{4}{s}\Big].\end{aligned}$$

Example 2.3.8.

$$\mathcal{L}(\cos(t)\mathcal{U}(t-\pi)) = \mathcal{L}(-\cos(t-\pi)\mathcal{U}(t-\pi)) = -e^{-\pi s}\frac{s}{s^2+1}.$$

Example 2.3.9. *Solve*

$$y' + y = f(t), y(0) = 5$$

where

$$\begin{aligned} f(t) &= 0, 0 \le t < \pi \\ &= 3\cos(t), t \ge \pi. \end{aligned}$$

Remarks:

a. f is a piecewise defined function. It represents an external force that is “kept off" from time 0 to π and then “turned on" afterwards.

b. The procedure to solve this problem is routine; i.e. we apply the Laplace transform on both sides. We skip the details (see also Section 4.3 Example 9 in the textbook). The only thing new is to express f in a convenient form that utilizes the result we developed in section. Namely we write

$$f(t) = -3\cos(t-\pi)\mathcal{U}(t-\pi)$$

and use the result in the previous example.

c. $f(t)$ may also take the form

$$\begin{aligned} f(t) &= 3\cos(t), 0 \le t < \pi \\ &= 0, t \ge \pi. \end{aligned}$$

That is the external force that is “turned on" from time 0 to π and then “turned off" afterwards. We can re-write $f(t)$ using the Heaviside function as

$$\begin{aligned} f(t) &= 3\cos(t)(1 - \mathcal{U}(t-\pi)) \\ &= 3\cos(t) + 3\cos(t-\pi)\mathcal{U}(t-\pi). \end{aligned}$$

2.4 Derivatives of Transforms and Transforms of Integrals

2.4.1 Introduction

In this section, we continue to pursue the theme of the interaction between linear operators we've already known (translation, differentiation $\cdots$) and the Laplace transform (or the inverse Laplace transform). In particular, we will look at the interaction of differentiation and the inverse Laplace transform. Secondly, we look at the interaction between integration (more specifically the *convolution* operation) and the Laplace transform. As an application, we will use the results covered in this section to solve the Volterra Integral Equation).

2.4.2 Derivatives of the Laplace transforms

Theorem 2.4.1. *Let $F(s) = \mathcal{L}(f(t))$. Denote $F^{(n)}$ as the n-th derivative of F. Then*

$$\mathcal{L}(t^n f(t)) = (-1)^n F^{(n)}(s).$$

Equivalently,

$$t^n f(t) = \mathcal{L}^{(-1)}\Big((-1)^n F^{(n)}(s)\Big).$$

Remark: Compare this result with the Laplace transform of the derivative of f:

$$\mathcal{L}(f^{(n)}(t)) = s^n \mathcal{L}(f)(s) - s^{n-1}f(0) - s^{n-2}f'(0) - \cdots - f^{(n-1)}(0).$$

Example 2.4.2.

$$\begin{aligned} \mathcal{L}(t\sin(kt)) &= -\frac{d}{ds}\Big(\frac{k}{s^2+k^2}\Big) \\ &= \frac{2ks}{(s^2+k^2)^2} \end{aligned}$$

Example 2.4.3. *Solve*

$$x'' + 16x = \cos(4t), x(0) = 0, x'(0) = 1.$$

Ans:
Denote $F(s) = \Lambda(x(t))$. Taking the Laplace transform on both sides give

$$s^2F(s) - sx(0) - x'(0) + 16F(s) = \frac{s}{s^2+16}.$$

That is

$$F(s) = \frac{s}{(s^2+16)^2} + \frac{1}{s^2+16}.$$

Observe that

$$\frac{d}{ds}\Big(\frac{1}{s^2+16}\Big) = \frac{-2s}{(s^2+16)^2}.$$

Thus

$$x(t) = \frac{t}{8}\sin(4t) + \frac{1}{4}\sin(4t).$$

2.4.3 Convolution

Definition 2.4.4. *Let $f(t)$ and $g(t)$ be two functions defined on $[0,\infty)$ with nice enough properties. The convolution of f and g, denoted as $f * g$ is defined as*

$$f * g(t) = \int_0^t f(u)g(t-u)du.$$

Remarks:

a. The convolution of two functions is a function.

b. Applying a change of varible by letting $v = t - u$ we have

$$g * f(t) = \int_0^t f(u)g(t-u)du = \int_0^t f(t-v)g(v)dv.$$

That is $f * g = g * f$: the convolution operator is commutative.

Example 2.4.5.

$$e^t * \sin(t) = \int_0^t e^u \sin(t-u)du = \frac{1}{2}(e^t - \sin t - \cos t).$$

2.4.4 Transforms of integrals

Theorem 2.4.6. *Let $f(t)$ and $g(t)$ be two functions defined on $[0,\infty)$ with nice enough properties and $F(s), G(s)$ their corresponding Laplace transforms. Then*

$$\mathcal{L}(f * g(t)) = F(s)G(s).$$

Equivalently,

$$\mathcal{L}^{-1}(F(s)G(s)) = f * g(t).$$

Remarks:

a. We do not have the converse result. That is we do NOT have

$$\mathcal{L}(f(t)g(t)) = F(s) * G(s).$$

b. A useful particular case : the inverse Laplace transform of $F^2(s)$ is $f * f(t)$. Repeating this procedure we can find the inverse Laplace transform of $F^n(s)$ for any n, even though it may be tedious.

Example 2.4.7.

$$\mathcal{L}^{-1}\left(\frac{1}{(s^2+k^2)^2}\right) = \frac{1}{k^2}\sin * \sin(kt) = \frac{\sin(kt) - kt\cos(kt)}{2k^3}.$$

c. Another useful particular case: taking $g(t) = 1$ we have

$$\mathcal{L}\Big(\int_0^t f(u)du\Big) = \frac{F(s)}{s}.$$

Example 2.4.8.

$$\mathcal{L}^{-1}\Big(\frac{1}{s(s^2+1)}\Big) = \int_0^t \sin(u)du = 1 - \cos t.$$

c. Convolution versus partial fractions: sometimes it is convenient to give the answer to an inverse Laplace transform as a convolution. This has the drawback of being not explicit (since the convolution might be hard to calculate). On the other hand a partial fraction may be performed on the Laplace transform which gives an explicit inverse Lapalce transform. The following example will make clear.

Example 2.4.9. *Find*

$$\mathcal{L}^{-1}\left(\frac{s^2}{(s^2+2)(s^2+12)}\right).$$

Sol: The answer can be given simply as

$$f(t) = \frac{1}{\sqrt{24}}\cos(\sqrt{2}t) * \cos(\sqrt{12}t).$$

But notice this is not as explicit as we would like (the convolution would be inconvenient to carry out). On the other hand, using partial fraction we get

$$\frac{s^2}{(s^2+2)(s^2+12)} = \frac{-1/5}{s^2+2} + \frac{6/5}{s^2+12}.$$

Thus

$$f(t) = -\frac{1}{5\sqrt{2}}\sin(\sqrt{2}t) + \frac{6}{5\sqrt{12}}\sin(\sqrt{12}t).$$

Example 2.4.10. *Volterra integral equation Solve for $f(t)$ that satisfies*

$$f(t) = 3t^2 - e^{-t} - \int_0^t f(u)e^{t-u}du.$$

Sol:

Taking the Laplace transform on both sides gives

$$F(s) = \frac{3!}{s^3} - \frac{1}{s+1} - \frac{F(s)}{s-1}.$$

That is

$$\begin{aligned} F(s) &= \frac{(-s^3+6s+6)(s-1)}{s^4(s+1)} \\ &= \frac{-s^4+s^3+6s^2-6}{s^4(s+1)}. \end{aligned}$$

We simplify $F(s)$ as followed:

$$\begin{aligned} \frac{-s^4+s^3}{s^4(s+1)} &= \frac{-1}{s+1} + \frac{1}{s(s+1)} \\ &= \frac{-1}{s+1} + \frac{1}{s} - \frac{1}{s+1} \\ &= \frac{-2}{s+1} + \frac{1}{s} \\ \frac{6s^2-6}{s^4(s+1)} &= \frac{6(s-1)(s+1)}{s^4(s+1)} \\ &= \frac{6s-6)}{s^4} = \frac{6}{s^3} - \frac{6}{s^4}. \end{aligned}$$

We can now find $f(t) = 3t^2 - t^3 + 1 - 2e^{-t}$ by applying the routine inverse transform.

2.5 Transforms of a Periodic Function ; the Dirac Delta Function

2.5.1 Transform of periodic functions

Theorem 2.5.1. *Let f be periodic with period T. Then*

$$\mathcal{L}(f(t)) = \frac{\int_0^T e^{-st} f(t)dt}{1-e^{-sT}}.$$

Proof. By definition

$$\begin{aligned} \mathcal{L}(f(t)) &= \int_0^\infty e^{-st} f(t)dt = \int_0^T e^{-st} f(t)dt + \int_T^\infty e^{-st} f(t)dt \\ &= \int_0^T e^{-st} f(t)dt + \int_0^\infty e^{-s(t+T)} f(t+T)dt \\ &= \int_0^T e^{-st} f(t)dt + e^{-sT}\mathcal{L}(f(t)). \end{aligned}$$

The conclusion follows. ■

Example 2.5.2. *Let $f(t)$ be a square-wave function. Namely:*

$$\begin{aligned} f(t) &= 1, k \leq t \leq k+1; k \textit{ even} \\ &= 0, \textit{ otherwise.} \end{aligned}$$

Then $f(t)$ is periodic with period 2. Thus

$$\begin{aligned} \mathcal{L}(f(t)) &= \frac{\int_0^1 e^{-st}dt}{1-e^{-2s}} = \frac{1-e^{-s}}{s(1-e^{-2s})} \\ &= \frac{1}{s(1+e^{-s})}. \end{aligned}$$

2.5.2 The Dirac Delta function

Unit impulse function

Definition 2.5.3. *Let a, t_0 be given. We define*

$$\begin{aligned} \delta_a(t-t_0) &= \frac{1}{2a}, t_0 - a \leq t < t_0 + a \\ &= 0 \textit{ otherwise}. \end{aligned}$$

We say $\delta_a(t-t_0)$ is a unit impulse function (concentrated around t_0 with width a).

Remarks:
a. $\int_0^\infty \delta_a(t-t_0)dt = 1$. Hence the name *unit* impulse function.
b. We usually thinks of a as very small. $\delta_a(t-t_0)$ represents a force that acts for a very short period of time around t_0 with large magnitude.

The Dirac Delta function

Imagine if we let $a \to 0$ in the unit impulse function $\delta_a(t-t_0)$ and obtain a limiting object, denoted as $\delta(t-t_0)$. Since for any a, the integral $\int_0^\infty \delta_a(t-t_0)dt = 1$ we also expect the same holds for $\delta(t-t_0)$. Also from the definition of $\delta_a(t-t_0)$, the pointwise value of $\delta(t-t_0)$ should be ∞ at $t = t_0$ and 0 otherwise. This expectation is very informal and in fact we cannot justify the existence of $\delta(t-t_0)$ *as a function*. Nevertheless, we can still view $\delta(t-t_0)$ as an object that has certain characteristics when it interacts with other functions. (In more advanced mathematics course you will learn that $\delta(t-t_0)$ is referred to as a *distribution.*)

Definition 2.5.4. *The Dirac Delta "function" concentrated at $t_0 > 0$ denoted by $\delta(t-t_0)$, is characterized by two properties:*

$$\begin{aligned} &a. \quad \delta(t-t_0) = \infty, t = t_0 \\ &\qquad\qquad\qquad = 0, \ t \neq 0. \\ &b. \quad \int_0^\infty \delta(t-t_0)dt = 1. \end{aligned}$$

Remark: Again, we emphasize that the Dirac Delta function is not a function. Its existence is “seen" by its action on other functions as in the following result.

Theorem 2.5.5. *Let $f(t)$ be a function with nice enough properties. Then for $t_0 > 0$*

$$\int_0^\infty f(t)\delta(t-t_0)dt := \lim_{a\to 0}\int_0^\infty f(t)\delta_a(t-t_0)dt = f(t_0).$$

Remarks:

a. Note that the integral on the LHS is defined via the limit as $a \to 0$. By itself it does not necessarily have any meaning because $\delta(t-t_0)$ is not a function.

b. The action of $\delta(t-t_0)$ on the function f is precisely the evaluation of f at t_0.

c. The condition $t_0 > 0$ is because the left limit of the integral is at 0. That is if $t_0 = 0$ then $t_0 - a < 0$ will be outside the left limit of the integral and what we say below in the proof will no longer be correct.

Proof. Denote $F(t)$ as the anti-derivative of $f(t)$. Then by L'Hospital rule:

$$\begin{aligned}\lim_{a\to 0}\int_0^\infty f(t)\delta_a(t-t_0)dt &= \lim_{a\to 0}\frac{F(t_0+a)-F(t_0-a)}{2a}\\ &= \lim_{a\to 0}\frac{f(t_0+a)+f(t_0-a)}{2} = f(t_0).\end{aligned}$$

■

The following corollary is immmediate:

Corollary 2.5.6. *For $t_0 > 0$*

$$\mathcal{L}(\delta(t-t_0)) = e^{-st_0}.$$

Remark: Again note that this corollary only holds for $t_0 > 0$. That is we CANNOT conclude $\mathcal{L}(\delta(t)) = 1$ by plugging in $t_0 = 0$ on both sides of the equation. Observe also that for a nice enough function f,

$$\begin{aligned}\lim_{s\to\infty}\mathcal{L}(f) &= \lim_{s\to\infty}\int_0^\infty e^{-st}f(t)dt\\ &= \int_0^\infty \lim_{s\to\infty} e^{-st}f(t)dt = 0.\end{aligned}$$

Even though $\delta(t-t_0)$ is *not* a function, $\mathcal{L}(\delta(t-t_0))$ is consistent in the sense that it also sastisfies $\lim_{s\to\infty}\mathcal{L}(\delta(t-t_0)) = 0$ for $t_0 > 0$. This property does not hold for $\mathcal{L}(\delta(t))$.

Example 2.5.7. *Solve*

$$\begin{aligned}y'' + y &= 4\delta(t-2\pi)\\ y(0)=1, y'(0) &= 0.\end{aligned}$$

Sol Let $Y(s)$ denote the Laplace transform of y. Then

$$s^2Y - sy(0) - y'(0) + Y = 4e^{2\pi s}.$$

That is

$$Y(s) = \frac{4e^{2\pi s}}{s^2+1} + \frac{s}{s^2+1}.$$

Taking the inverse Laplace transform gives

$$\begin{aligned} y(t) &= 4\sin(t-2\pi)\mathcal{U}(t-2\pi) + \cos(t) \\ &= 4\sin(t)\mathcal{U}(t-2\pi) + \cos(t). \end{aligned}$$

2.5.3 A summary of Laplace transform results

Property	t **domain**	s **domain**
Linearity	$af(t) + bg(t)$	$aF(s) + bG(s)$
s domain derivative	$t^n f(t)$	$(-1)^n F^{(n)}(s)$
t domain derivative	$f^{(n)}(t)$	$s^n F(s) - \sum_{i=1}^{n-1} s^i f^{(n-1-i)}(0)$
t domain convolution	$f * g(t) = \int_0^t f(u)g(t-u)du$	$F(s)G(s)$
t domain integration	$\int_0^t f(u)du$	$\frac{F(s)}{s}$

Property	t **domain**	s **domain**
t domain translation	$f(t-a)\mathcal{U}(t-a)$	$e^{-as}F(s)$
s domain translation	$e^{at}f(t)$	$F(s-a)$
t domain scaling	f(at)	$\frac{F(\frac{s}{a})}{a}$
Periodic function	$f(t+T)=f(t)$	$\frac{1}{1-e^{-Ts}}\int_0^T e^{-st}f(t)dt$
Dirac delta function	$\delta(t-t_0)$	e^{-st_0}

2.6 Systems of Linear Differential Equations

2.6.1 Introduction

In this section we look at system of differential equations. The technique of Laplace transforms will be used to solve such system. Specifically after the transform we get a system of algebraic equations, where the unknowns are the Laplace transforms of the original solutions. The original solutions are obtained by solving for those Laplace transforms and taking the inverse Laplace transform.

2.6.2 Example

Example 2.6.1. *Solve*

$$\begin{aligned} x_1'' + 10x_1 - 4x_2 &= 0 \\ -4x_1 + x_2'' + 4x_2 &= 0 \\ x_1(0) = 0, x_1'(0) = 1, x_2(0) = 0, x_2'(0) &= -1. \end{aligned}$$

Sol: Let $X_1(s), X_2(s)$ *denote the Laplace transforms of* $x_1(t), x_2(t)$ *respectively. Then*

$$\begin{aligned} s^2X_1 - sx_1(0) - x_1'(0) + 10X_1 - 4X_2 &= 0 \\ -4X_1 + s^2X_2 - sx_2(0) - x_2'(0) + 4X_2 &= 0 \end{aligned}$$

Adding the two equations gives

$$(s^2+6)X_1 + s^2X_2 = 0$$

or

$$X_2 = -\frac{s^2+6}{s^2}X_1.$$

Substituing into the first equation gives

$$X_1(s^2 + 10 + 4\frac{s^2+6}{s^2}) = 1.$$

That is

$$X_1(s) = \frac{s^2}{s^4+14s^2+24} = \frac{s^2}{(s^2+2)(s^2+12)}.$$

Thus

$$X_2(s) = -\frac{s^2+6}{(s^2+2)(s^2+12)}.$$

Taking the Laplace transform gives

$$x_1(t) = \frac{1}{\sqrt{24}}\cos(\sqrt{2}t) * \cos(\sqrt{12}t).$$

Alternatively, using partial fraction we get

$$X_1(s) = \frac{-1/5}{s^2+2} + \frac{6/5}{s^2+12}.$$

Thus

$$x_1(t) = -\frac{1}{5\sqrt{2}}\sin(\sqrt{2}t) + \frac{6}{5\sqrt{12}}\sin(\sqrt{12}t).$$

Similarly,

$$X_2(s) = \frac{-2/5}{s^2+2} - \frac{3/5}{s^2+12}.$$

Thus

$$x_2(t) = -\frac{2}{5\sqrt{2}}\sin(\sqrt{2}t) + -\frac{3}{5\sqrt{12}}\sin(\sqrt{12}t).$$

Chapter 3

Fourier Series

3.1 Orthogonal Sets of Functions and Fourier Series

3.1.1 Orthogonal structure and geometry of a vector space

Let V be a vector space with a basis B. For simplicity suppose that B is a finite collection of vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$. This means that any vector $\mathbf{x}$ in V can be expressed as a linear combination of $\mathbf{v}_i, i = 1, \cdots, n$:

$$\mathbf{x} = \sum_{i=1}^{n} c_i \mathbf{v}_i.$$

The coefficients c_i are called the *coordinates* of $\mathbf{x}$ with respect to the basis B. Thus given a vector $\mathbf{x}$, we want to determine the corresponding coefficents c_i. A good choice of the basis B can make this process very simple. For example, in $V = \mathbb{R}^2$ the most natural basis is $B = \{[1,0]^T, [0,1]^T\}$. We can immediately read off the coefficients c_i for any vector $\mathbf{x}$:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Indeed, the coordinates we use to express $\mathbf{x}$ is implicitly understood as using B as the choice of basis. Compare this with the choice of $B' = \{[1,0]^T, [1,1]^T\}$. To find the coordinates of a vector $\mathbf{x}$ with respect to B' requires us to solve a linear system of equations.

What makes the choice of B good and B' not as good? It is the orthonomality of B (and not of B'). Recall that inner-product between two vectors $\mathbf{x} = [x_1, x_2, \cdots, x_n]^T$ and $\mathbf{y} = [y_1, y_2, \cdots, y_n]^T$ is defined as

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i.$$

Inner-product is a bilinear operator. That is

$$\begin{aligned} (a\mathbf{x} + b\mathbf{y}) \cdot \zeta &= a(\mathbf{x} \cdot \zeta) + b(\mathbf{y} \cdot \zeta) \\ \mathbf{x} \cdot (a\mathbf{y} + b\zeta) &= a(\mathbf{x} \cdot \mathbf{y}) + b(\mathbf{x} \cdot \zeta). \end{aligned}$$

Thus if we have an orthonomal basis B, that is

$$\begin{aligned} \mathbf{v}_i \cdot \mathbf{v}_j &= 0, i \neq j \\ &= 1, i = j. \end{aligned}$$

Then

$$\mathbf{x} = \sum_{i=1}^{n} c_i \mathbf{v}_i$$

implies

$$c_i = \mathbf{x} \cdot \mathbf{v}_i, \forall i.$$

Finding the coefficients c_i then requires taking inner-products, a much simpler procedure than solving system of equations.

We can generalize this concept to functions (which is the continuous version of vectors). We want to define the notion of orthogonality (which requires the notion of inner-product). From there we can find candidates for orthonormal bases of functions and express an arbitrary function in terms of the vectors in a basis. A particular choice of basis (the cosine and sine functions) will be the focus of the remainder of the chapter. Fourier series involves studying the representation of functions in this basis and their properties.

3.1.2 Orthogonal functions

Definition 3.1.1. *Let f_1, f_2 be defined on $[a, b]$ with nice enough properties. The inner product of f_1, f_2, denoted as (f_1, f_2) is defined as*

$$(f_1, f_2) = \int_a^b f_1(x) f_2(x) dx.$$

f_1, f_2 are orthogonal if $(f_1, f_2) = 0$. The L_2 norm of f_1, denoted as $\|f_1\|$ is defined as

$$\|f_1\| = \sqrt{(f_1, f_1)} = \sqrt{\int_a^b f_1^2(x) dx}.$$

A set of (possibly infinite) functions $\{\phi_0, \phi_1, \cdots\}$ is **orthogonal** *if $(\phi_i, \phi_j) = 0, \forall i \neq j$. A set of functions $\{\phi_0, \phi_1, \cdots\}$ is* **orthonormal** *if it is orthogonal and $\|\phi_i\| = 1$ for all i.*

Remark: The norm of a function tells us how “big" it is in a certain sense. There are many ways to define a norm, for example the sup norm which returns the maximum of $|f|$; the L1 norm which returns the area under the curve of $|f|$. In this sense, the L2 norm is one of the many norms but it has the advantage of having a natural connection with the inner product.

Example 3.1.2. *The set* $\{1, \cos 2x, \cos 3x, \cdots\}$ *is orthogonal but not orthonormal on* $[-\pi, \pi]$. *Indeed, for any* $m \neq n$

$$\begin{aligned}(1, \cos mx) &= \int_{-\pi}^{\pi} \cos mx dx = \frac{1}{m}(\sin(m\pi) - \sin(-m\pi)) = 0 \\ (\cos mx, \cos nx) &= \int_{-\pi}^{\pi} \cos mx \cos nx dx = \frac{1}{2}\int_{-\pi}^{\pi}[\cos(m+n)x + \cos(m-n)x]dx = 0,\end{aligned}$$

since $\int_{-\pi}^{\pi} \cos mx = 0$ *for any integer* m *as we have seen from the first computation*

On the other hand, for any m

$$\|\cos(mx)\|^2 = \int_{-\pi}^{\pi} (\cos mx)^2 dx = \frac{1}{2}\int_{-\pi}^{\pi}[1 + \cos 2mx]dx = \pi.$$

Therefore the set $\{1, \cos 2x, \cos 3x, \cdots\}$ *is not orthonormal but the set* $\{\frac{1}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \frac{\cos 3x}{\sqrt{\pi}}, \cdots\}$ *is.*

3.1.3 Orthogonality with respect to a weight function

We may have a set of functions that are orthogonal on any finite interval say $[-a, a]$ and want to extend the orthogonality property to the whole real line $(-\infty, \infty)$. A problem that may arise is integrability : the integral on $(-\infty, \infty)$ may not converge. For example, it is clear that $\phi_0 = 1$ and $\phi_1(x) = x$ are orthogonal on any interval $[-a, a]$. We cannot say they are orthogonal on $(-\infty, \infty)$ because $\int_{-\infty}^{\infty} x dx$ is undefined. To proper way to extend the notion of orthogonality is orthogonality with respect to a weight function. The weight function helps with integrability on the whole real line but also preserves the symmetry of the original set.

Definition 3.1.3. *Let* f_1, f_2 *be defined on* $[a, b]$ *with nice enough properties. The inner product of* f_1, f_2 *with respect to a weight function* $w(t)$ *denoted as* $(f_1, f_2)^w$ *is defined as*

$$(f_1, f_2)^w = \int_a^b w(x) f_1(x) f_2(x) dx.$$

The L_2 *norm of* f_1, *denoted as* $\|f_1\|^w$ *is defined as*

$$\|f_1\|^w = \sqrt{(f_1, f_1)^w} = \sqrt{\int_a^b w(x) f_1^2(x) dx}.$$

A set of functions $\{\phi_0, \phi_1, \cdots\}$ *is orthogonal with respect to a weight function* $w(x)$ *on if* $(\phi_i, \phi_j)^w = 0, \forall i \neq j$. *A set of functions* $\{\phi_0, \phi_1, \cdots\}$ *is orthonormal with respect to a weight function* $w(x)$ *if it is orthogonal and* $\|\phi_i\|^w = 1$ *for all* i.

Example 3.1.4. *The set* $\{H_0 = 1, H_1(x) = 2x, H_2(x) = 4x^2 - 2\}$ *are orthogonal on* $(-\infty, \infty)$. *with respect to weight function* e^{-x^2}.

3.1.4 Fourier series

3.1.5 Orthogonal series expansion

Given an orthogonal set of functions $\{\phi_0, \phi_1, \cdots\}$ and a function f, it is natural to ask if we can expand f in terms of the $\phi_i's$. That is can we find coefficients c_i so that

$$f(x) = \sum_{i=1}^{\infty} c_i \phi_i(x)?$$

Note that unlike the example in $\mathbb{R}^n$, we usually expect the expansion to be an infinite series since a general space of functions is usually infinite in dimension. In an ad hoc way, suppose that the equation $f(x) = \sum_{i=1}^{\infty} c_i \phi_i(x)$ *is* true, then it is actually simple to find c_i by the orthogonality of ϕ_i. Indeed, for a specific index k

$$(f, \phi_k) = (\sum_{i=1}^{\infty} c_i \phi_i(x), \phi_k) = \sum_{i=1}^{\infty} c_i (\phi_i(x), \phi_k) = c_k \|\phi_k\|^2.$$

That is $c_k = \frac{(f,\phi_k)}{\|\phi_k\|^2}$ if the representation is true. As we will see in the next section, the representation is true if and only if the series representation converges.

3.1.6 Trigonometric Series

Lemma 3.1.5. *For any $p > 0$, the set of trigonometric functions*

$$\left\{1, \cos\frac{\pi}{p}x, \cos\frac{2\pi}{p}x, \cos\frac{3\pi}{p}x, \cdots, \sin\frac{\pi}{p}x, \sin\frac{2\pi}{p}x, \sin\frac{3\pi}{p}x, \cdots\right\}$$

is orthogonal on $[-p, p]$.

Proof. We first show the orthogonality between 1 and $\cos\frac{k\pi}{p}x$ as well as 1 and $\sin\frac{k\pi}{p}x$

$$\begin{aligned}
\int_{-p}^{p} \cos\frac{k\pi}{p}x dx &= \frac{p}{\pi}\int_{-\pi}^{\pi} \cos kx dx = \frac{p}{k\pi}\sin kx\Big|_{-\pi}^{\pi} = 0 \\
\int_{-p}^{p} \sin\frac{k\pi}{p}x dx &= \frac{p}{\pi}\int_{-\pi}^{\pi} \sin kx dx = -\frac{p}{k\pi}\cos kx\Big|_{-\pi}^{\pi} = 0.
\end{aligned}$$

Now we show the orthogonality between $\cos\frac{m\pi}{p}x$ and $\cos\frac{n\pi}{p}x$

$$\begin{aligned}
\int_{-p}^{p} \cos\frac{m\pi}{p}x \cos\frac{n\pi}{p}x dx &= \frac{p}{\pi}\int_{-\pi}^{\pi} \cos mx \cos nx dx \\
&= \frac{p}{2\pi}\int_{-\pi}^{\pi} [\cos(m+n)x + \cos(m-n)x] dx = 0.
\end{aligned}$$

Similarly, $\sin\frac{m\pi}{p}x$ and $\sin\frac{n\pi}{p}x$ are orthogonal:

$$\begin{aligned}\int_{-p}^{p}\sin\frac{m\pi}{p}x\sin\frac{n\pi}{p}xdx &= \frac{p}{\pi}\int_{-\pi}^{\pi}\sin mx\sin nxdx\\ &= \frac{p}{2\pi}\int_{-\pi}^{\pi}[\cos(m-n)x-\cos(m+n)x]dx=0.\end{aligned}$$

Lastly, $\cos\frac{m\pi}{p}x$ and $\sin\frac{n\pi}{p}x$ are orthogonal:

$$\begin{aligned}\int_{-p}^{p}\cos\frac{m\pi}{p}x\sin\frac{n\pi}{p}xdx &= \frac{p}{\pi}\int_{-\pi}^{\pi}\cos mx\sin nxdx\\ &= \frac{p}{2\pi}\int_{-\pi}^{\pi}[\sin(m+n)x+\sin(n-m)x]dx=0.\end{aligned}$$

We would like to use the trigonometric series as the orthogonal basis for a function expansion as discussed in the previous section. The expansion of a function f, defined on $[-p,p]$, using this basis would read

$$f(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}\Big(a_n\cos\frac{n\pi}{p}x+b_n\sin\frac{n\pi}{p}x\Big).$$

The choice of $\frac{a_0}{2}$ for the constant coefficient is conventional and probably helps to keep the formula for a_0, a_n, b_n in a uniform way (see below). The series is called *the Fouries series* of f and a_n, b_n are called the *the Fourier coefficients* of f.

Now note that

$$\begin{aligned}\int_{-p}^{p}1dx &= 2p\\ \int_{-p}^{p}(\cos\frac{n\pi}{p}x)^2dx &= p\\ \int_{-p}^{p}(\sin\frac{n\pi}{p}x)^2dx &= p.\end{aligned}$$

If the Fourier serires representation is true, the Fourier coefficients are determined as we discussed before:

$$\begin{aligned}a_0 &= \frac{1}{p}\int_{-p}^{p}f(x)dx\\ a_n &= \frac{1}{p}\int_{-p}^{p}f(x)\cos\frac{n\pi}{p}xdx\\ b_n &= \frac{1}{p}\int_{-p}^{p}f(x)\sin\frac{n\pi}{p}xdx.\end{aligned}$$

Whether this representation holds depends on the convergence of the Fourier series, which we discuss in the next section.

3.2 Fourier Series

3.2.1 Orthogonal series expansion

Given an orthogonal set of functions $\{\phi_0, \phi_1, \cdots\}$ and a function f, it is natural to ask if we can expand f in terms of the $\phi_i' s$. That is can we find coefficients c_i so that

$$f(x) = \sum_{i=1}^{\infty} c_i \phi_i(x)?$$

Note that unlike the example in $\mathbb{R}^n$, we usually expect the expansion to be an infinite series since a general space of functions is usually infinite in dimension. In an ad hoc way, suppose that the equation $f(x) = \sum_{i=1}^{\infty} c_i \phi_i(x)$ *is* true, then it is actually simple to find c_i by the orthogonality of ϕ_i. Indeed, for a specific index k

$$(f, \phi_k) = (\sum_{i=1}^{\infty} c_i \phi_i(x), \phi_k) = \sum_{i=1}^{\infty} c_i (\phi_i(x), \phi_k) = c_k \|\phi_k\|^2.$$

That is $c_k = \frac{(f,\phi_k)}{\|\phi_k\|^2}$ if the representation is true. As we will see in the next section, the representation is true if and only if the series representation converges.

3.2.2 Trigonometric Series

Lemma 3.2.1. *For any $p > 0$, the set of trigonometric functions*

$$\left\{1, \cos\frac{\pi}{p}x, \cos\frac{2\pi}{p}x, \cos\frac{3\pi}{p}x, \cdots, \sin\frac{\pi}{p}x, \sin\frac{2\pi}{p}x, \sin\frac{3\pi}{p}x, \cdots\right\}$$

is orthogonal on $[-p, p]$.

Proof. We first show the orthogonality between 1 and $\cos\frac{k\pi}{p}x$ as well as 1 and $\sin\frac{k\pi}{p}x$

$$\begin{aligned}
\int_{-p}^{p} \cos\frac{k\pi}{p}x dx &= \frac{p}{\pi}\int_{-\pi}^{\pi} \cos kx dx = \frac{p}{k\pi}\sin kx\Big|_{-\pi}^{\pi} = 0 \\
\int_{-p}^{p} \sin\frac{k\pi}{p}x dx &= \frac{p}{\pi}\int_{-\pi}^{\pi} \sin kx dx = -\frac{p}{k\pi}\cos kx\Big|_{-\pi}^{\pi} = 0.
\end{aligned}$$

Now we show the orthogonality between $\cos\frac{m\pi}{p}x$ and $\cos\frac{n\pi}{p}x$

$$\begin{aligned}
\int_{-p}^{p} \cos\frac{m\pi}{p}x \cos\frac{n\pi}{p}x dx &= \frac{p}{\pi}\int_{-\pi}^{\pi} \cos mx \cos nx dx \\
&= \frac{p}{2\pi}\int_{-\pi}^{\pi} [\cos(m+n)x + \cos(m-n)x] dx = 0.
\end{aligned}$$

Similarly, $\sin\frac{m\pi}{p}x$ and $\sin\frac{n\pi}{p}x$ are orthogonal:

$$\begin{aligned}\int_{-p}^{p}\sin\frac{m\pi}{p}x\sin\frac{n\pi}{p}xdx &= \frac{p}{\pi}\int_{-\pi}^{\pi}\sin mx\sin nxdx\\ &= \frac{p}{2\pi}\int_{-\pi}^{\pi}[\cos(m-n)x-\cos(m+n)x]dx=0.\end{aligned}$$

Lastly, $\cos\frac{m\pi}{p}x$ and $\sin\frac{n\pi}{p}x$ are orthogonal:

$$\begin{aligned}\int_{-p}^{p}\cos\frac{m\pi}{p}x\sin\frac{n\pi}{p}xdx &= \frac{p}{\pi}\int_{-\pi}^{\pi}\cos mx\sin nxdx\\ &= \frac{p}{2\pi}\int_{-\pi}^{\pi}[\sin(m+n)x+\sin(n-m)x]dx=0.\end{aligned}$$

We would like to use the trigonometric series as the orthogonal basis for a function expansion as discussed in the previous section. Now note that

$$\begin{aligned}\int_{-p}^{p}1dx &= 2p\\ \int_{-p}^{p}(\cos\frac{n\pi}{p}x)^2dx &= p\\ \int_{-p}^{p}(\sin\frac{n\pi}{p}x)^2dx &= p.\end{aligned}$$

Thus, to keep the norm on all basis functions the same, the series we would actually use is

$$\left\{\frac{1}{2},\cos\frac{\pi}{p}x,\cos\frac{2\pi}{p}x,\cos\frac{3\pi}{p}x,\cdots,\sin\frac{\pi}{p}x,\sin\frac{2\pi}{p}x,\sin\frac{3\pi}{p}x,\cdots\right\}.$$

The expansion of a function f, defined on $[-p,p]$, using this basis would read

$$f(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}\Big(a_n\cos\frac{n\pi}{p}x+b_n\sin\frac{n\pi}{p}x\Big).$$

This is called *the Fouries series* of f and a_n, b_n are called the *the Fourier coefficients* of f.

If the Fourier serires representation is true, the Fourier coefficients are determined as we discussed before:

$$\begin{aligned}a_0 &= \frac{1}{p}\int_{-p}^{p}f(x)dx\\ a_n &= \frac{1}{p}\int_{-p}^{p}f(x)\cos\frac{n\pi}{p}xdx\\ b_n &= \frac{1}{p}\int_{-p}^{p}f(x)\sin\frac{n\pi}{p}xdx.\end{aligned}$$

Whether this representation holds depends on the convergence of the Fourier series, which we discuss in the next section.

3.2.3 Convergence of Fourier series

Theorem 3.2.2. *Let f be a piece-wise differentiable function on $[-p, p]$ with piece-wise continuous f'. Then at any point of continuity the Fourier series of f converges to $f(x)$. At a point of discontinuity, the Fourier series for f converges to $\frac{f(x+)+f(x-)}{2}$.*

We will give an example of this convergence at the end of the section.

3.2.4 Periodic extension

Observe that all the functions in

$$\left\{1, \cos\frac{\pi}{p}x, \cos\frac{2\pi}{p}x, \cos\frac{3\pi}{p}x, \cdots, \sin\frac{\pi}{p}x, \sin\frac{2\pi}{p}x, \sin\frac{3\pi}{p}x, \cdots\right\}$$

have a period $2p$ (even qthough this is not neccessarily their fundamental period):

$$\begin{aligned}\cos\frac{k\pi}{p}(x+2p) &= \cos(\frac{k\pi}{p}x + 2k\pi) = \cos\frac{k\pi}{p}x \\ \sin\frac{k\pi}{p}(x+2p) &= \sin(\frac{k\pi}{p}x + 2k\pi) = \sin\frac{k\pi}{p}x.\end{aligned}$$

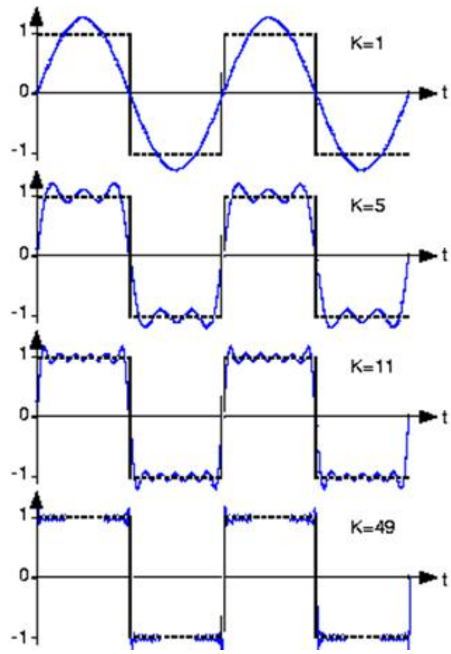
Thus if we have a function f defined on $[-p, p]$ and a Fourier series for f, we can extend f to $(-\infty, \infty)$ by defining that $f(x + 2p) = f(x)$ for all x. That is we do a periodic extension on f with period $2p$. Since all the functions in the trigonometric basis as period $2p$, the same Fourier series will hold for the extension of f to $(-\infty, \infty)$.

3.2.5 An example

Let $f(x)$ be a square wave. That is

$$\begin{aligned}f(x) &= -1, -1 \le x \le 0 \\ &= 1, 0 \le x \le 1 \\ f(x+2) &= f(x).\end{aligned}$$

That is $f(x)$ is defined on $[-p, p]$ with $p = 1$ and we extend f with period 2 to $(-\infty, \infty)$. We will give a picture example of how the partial series converge to the function $f(x)$. For notational purpose, the partial series $\frac{a_0}{2} + \sum_{n=1}^{k}\left(a_n \cos\frac{n\pi}{p}x + b_n \sin\frac{n\pi}{p}x\right)$ will be denoted as S_k. The picture compares S_1, S_5, S_{11}, S_{49} with $f(x)$, respectively. Note that $f(x)$ has discontinuities at $x = 2k + 1$, in particular at $x = 1$. By the convergence theorem, the series converges to $\frac{f(1-)+f(1+)}{2} = 0$. Note how the S_k always attains value 0 at $x = 2k + 1$.



3.3 Fourier Sine and Cosine series, Complex Fourier Series

3.3.1 Even and odd functions

In the previous section, we expand a function f defined on $[-p, p]$ with a series involving cosine and sine. Since cosine and sine are symmetric around 0, we suspect that representing a function with a similar symmetric structure around 0 may involve less trigonemtric functions. We will show that this is indeed the case. First we recall that a function f is odd if $f(x) = -f(-x)$ and even if $f(x) = f(-x)$. The following are some properties of even / odd functions:

a. The product of two even or odd functions is even.
b. The product of an even and an odd function function is odd.
c. The sume (difference) of two even (odd) functions is even (odd).
d. $\int_{-a}^{a} f(x)dx = 2\int_{0}^{a} f(x)dx$ if f is even.
e. $\int_{-a}^{a} f(x)dx = 0$ if f is odd.

3.3.2 Cosine and sine series

Lemma 3.3.1. *Let f be an even function on $[-p, p]$. Then the Fourier series of f only involve cosine functions:*

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos \frac{n\pi}{p} x.$$

On the other hand, if f is an even function on $[-p,p]$ then its Fourier series only involve sine functions:

$$f(x)=\sum_{n=1}^{\infty} b_n \sin\frac{n\pi}{p}x.$$

Proof. Recall that if f has the Fourier series

$$f(x)=\frac{a_0}{2}+\sum_{n=1}^{\infty}\left(a_n\cos\frac{n\pi}{p}x+b_n\sin\frac{n\pi}{p}x\right)$$

then

$$\begin{aligned} a_0 &= \frac{1}{p}\int_{-p}^{p} f(x)dx \\ a_n &= \frac{1}{p}\int_{-p}^{p} f(x)\cos\frac{n\pi}{p}xdx \\ b_n &= \frac{1}{p}\int_{-p}^{p} f(x)\sin\frac{n\pi}{p}xdx. \end{aligned}$$

If f is even then $f(x)f(x)\sin\frac{n\pi}{p}x$ is odd and $b_n = 0, \forall n$ by property e in the previous section. Similarly, if f is odd then $a_0 = 0$ and $f(x)\cos\frac{n\pi}{p}x$ is odd so $a_n = 0, \forall n$ also by property e.

Example 3.3.2. *Consider the square wave function defined on $[-\pi,\pi]$ as*

$$\begin{aligned} f(x) &= -1, -\pi \le x \le 0; \\ &= 1, -0 \le x \le \pi. \end{aligned}$$

This is an odd function. Thus it only involves sine in its expansion. The coefficients are

$$b_n=\frac{2}{\pi}\int_0^{\pi}(1)\sin nxdx=\frac{2}{\pi}\frac{1-(-1)^n}{n}.$$

That is

$$f(x)=\frac{2}{\pi}\sum_{n=1}^{\infty}\frac{1-(-1)^n}{n}\sin nx.$$

3.3.3 Half range expansions

Let f be defined on $[0,p]$. We would like to find a Fourier series representation for f. That is we want to expand f using the basis

$$\left\{1,\cos\frac{\pi}{p}x,\cos\frac{2\pi}{p}x,\cos\frac{3\pi}{p}x,\cdots,\sin\frac{\pi}{p}x,\sin\frac{2\pi}{p}x,\sin\frac{3\pi}{p}x,\cdots\right\}.$$

We cannot proceed immediately because this trigonometric basis is NOT orthogonal on $[0,p]$. Indeed if m is even and n is odd (so that both $m+n$ and $m-n$ are odd)

$$\begin{aligned}\int_0^p \cos\frac{m\pi}{p}x\cos\frac{n\pi}{p}xdx &= \frac{p}{\pi}\int_0^\pi \cos mx\cos nxdx\\ &= \frac{p}{2\pi}\int_0^\pi [\cos(m+n)x+\cos(m-n)x]dx = \frac{-4p}{2\pi}\neq 0.\end{aligned}$$

One can modify the choice of basis to something more suitable for the interval $[0,p]$. On the other hand, there is a very simple way to re-use the results we have developed so far by extending f to the interval $[-p,p]$ using an odd or even extension. Specifically we can define, for an even extension:

$$f(x)=f(-x), -p\leq x\leq 0$$

and for an odd extension

$$f(x)=-f(-x), -p\leq x\leq 0.$$

One last way to extend f is to use the identity expansion. That is we just “copy and paste" the entire graph of f onto the interval $[-p,0]$. In this way, the extension of f is periodic with period p.

Note that for odd and even extension, f is periodic with period $2p$. Therefore, for an odd extension we can use the sine expansion with basis

$$\left\{\sin\frac{\pi}{p}x, \sin\frac{2\pi}{p}x, \sin\frac{3\pi}{p}x, \cdots\right\}.$$

In this case the Fourier coefficients are

$$\begin{aligned}a_0 &= \frac{2}{p}\int_0^p f(x)dx\\ a_n &= \frac{2}{p}\int_0^p f(x)\cos\frac{n\pi}{p}xdx.\end{aligned}$$

Note that the integration is from 0 to p with a factor of 2 because $f(x)\cos\frac{n\pi}{p}x$ is an even function. For an even extension we can use the cosine expansion with basis

$$\left\{1, \cos\frac{\pi}{p}x, \cos\frac{2\pi}{p}x, \cos\frac{3\pi}{p}x, \cdots\right\}.$$

In this case the Fourier coefficients are

$$b_n = \frac{2}{p}\int_0^p f(x)\sin\frac{n\pi}{p}xdx.$$

Again note that the integration is from 0 to p with a factor of 2 because $f(x)\sin\frac{n\pi}{p}x$ is again an even function. For the identity expansion, f is periodic with period p. Thus we would need to use the entire Fourier series expansion with a modified basis

$$\left\{1, \cos\frac{2\pi}{p}x, \cos\frac{4\pi}{p}x, \cos\frac{6\pi}{p}x, \cdots, \sin\frac{2\pi}{p}x, \sin\frac{4\pi}{p}x, \sin\frac{6\pi}{p}x, \cdots\right\}.$$

That is p is replaced by $p/2$ in the basis so that every basis function has a period p (instead of $2p$ as before) which coincides with the period of f.

$$\begin{aligned} a_0 &= \frac{2}{p}\int_0^p f(x)dx \\ a_n &= \frac{2}{p}\int_0^p f(x)\cos\frac{2n\pi}{p}xdx \\ b_n &= \frac{2}{p}\int_0^p f(x)\sin\frac{2n\pi}{p}xdx. \end{aligned}$$

Here again the integration is from 0 to p with a factor of 2 but the reason is not because of odd or even property. It has to do with the fact that all functions in the integrand are periodic with period p. The original Fourier coefficients are computed as integration from $-p$ to p. But since the period is p we can integrate on half the period and double the integral to get the same result.

Latly, since the Fourier series holds for the extension of f on the interval $[-p, p]$ it also holds for f on its original domain $[0, p]$.

Example 3.3.3. *Consider $f(x) = x^2$ defined on [0,p].*

a. Expanding f in a cosine series gives

$$\begin{aligned} a_0 &= \frac{2}{p}\int_0^p x^2dx = \frac{2}{3}p^2 \\ a_n &= \frac{2}{p}\int_0^p x^2\cos(\frac{n\pi}{p}x)dx = \frac{4p^2(-1)^n}{n^2\pi^2}. \end{aligned}$$

Thus

$$f(x) = \frac{p^2}{3} + \frac{4p^2}{\pi^2}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos\frac{n\pi}{p}x.$$

b. Expanding f in a sine series gives

$$b_n = \frac{2}{p}\int_0^p x^2\sin(\frac{n\pi}{p}x)dx = \frac{4p^2(-1)^{n+1}}{n\pi} + \frac{4p^2}{n^3\pi^3}[(-1)^n - 1].$$

Thus

$$f(x) = \sum_{n=1}^{\infty}\left\{\frac{4p^2(-1)^{n+1}}{n\pi} + \frac{4p^2}{n^3\pi^3}[(-1)^n - 1]\right\}\sin\frac{n\pi}{p}x.$$

c. Expanding f using the identity expansion gives

$$\begin{aligned} a_0 &= \frac{2}{p}\int_0^p x^2 dx = \frac{2}{3}p^2 \\ a_n &= \frac{2}{p}\int_0^p x^2 \cos(\frac{2n\pi}{p}x)dx = \frac{p^2}{n^2\pi^2} \\ b_n &= \frac{2}{p}\int_0^p x^2 \sin(\frac{2n\pi}{p}x)dx = \frac{-p^2}{n\pi}. \end{aligned}$$

Thus

$$f(x) = \frac{p^2}{3} + \sum_{n=1}^{\infty} \frac{p^2}{n^2\pi^2} \cos\frac{2n\pi}{p}x - \frac{p^2}{n\pi}\sin\frac{2n\pi}{p}x.$$

3.3.4 Application: Particular solution of a differential equation

Consider a spring mass system that is subject to a driving force f:

$$mx''(t) + kx = f(t),$$

where f is periodic with period 2:

$$\begin{aligned} f(t) &= \pi t, 0 < t < 1 \\ &= \pi(t-2), 1 < t < 2. \end{aligned}$$

To solve for $x(t)$, we first represent f in a half-range sine expansion. With $p = 1$, the Fourier coefficents for f are

$$b_n = 2\int_0^1 \pi t \sin(n\pi t)dt = \frac{2(-1)^{n+1}}{n}.$$

That is

$$f(t) = \sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n}\sin(n\pi t).$$

Since the second derivative of sine is negative sine, we guess the particular solution x_p is of the form

$$x_p(t) = \sum_{n=1}^{\infty} B_n \sin(n\pi t).$$

Plug this form into the differential equation gives

$$\sum_{n=1}^{\infty}(k - mn^2\pi^2)B_n \sin(n\pi t) = \sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n}\sin(n\pi t).$$

That is

$$B_n = \frac{2(-1)^{n+1}}{n(k - mn^2\pi^2)}.$$

The particular solution is

$$x_p(t) = \sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n(k - mn^2\pi^2)} \sin(n\pi t).$$

Remark: It is conceivable to approach this problem using the Laplace transform for periodic function as we discussed before. However, we can anticipate that the Laplace transform $X(s)$ of $x(t)$ is easy to solve for; but the explicit inverse Laplace transform may not be easy to find given the fact that we have found the particular solution to be an infinite sine series. This may be considered as an advantage of the Fourier series method over the Laplace transform.

Resonance

Consider the problem

$$x''(t) + x(t) = f(t), t > 0$$

where $f(t)$ is periodic with period 2π and defined on $[0, 2\pi]$ as followed:

$$\begin{aligned} f(t) &= t, 0 \leq t \leq \pi \\ &= 2\pi - t, \pi \leq t \leq 2\pi. \end{aligned}$$

We want to find a particular solution to this problem. Using even-extension (cosine), we have

$$f(t) = \pi + \sum_{n=1, n \text{ odd}}^{\infty} \frac{-4}{n^2\pi} \cos(nt).$$

Assuming $x_p(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(nt)$ and plugging into the equation we have

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n(1 - n^2)\cos(nt) = \pi + \sum_{n=1, n \text{ odd}}^{\infty} \frac{-4}{n^2\pi} \cos(nt).$$

But here we see a problem: when $n = 1$ the equation reads

$$a_1 \times 0 = \frac{-4}{\pi},$$

which, at least right now, yields no solution for a_1 (and hence for $x_p(t)$). Intuitively, this problem occurs because the period of the forcing function, which is 2π, coincides with the period of the homogenous solution to the problem:

$$x''(t) + x(t) = 0,$$

which is $x_h(t) = C_1 \cos(t) + C_2 \sin(t)$. When this happens, resonance occurs.

To understand better why the coincidence of the two periods causes a problem, consider the ODE

$$x''(t) + x(t) = \cos t,$$

Here we put $\cos t$ for the forcing function, which also has period 2π but it is easier to find a particular solution in this case. Indeed, we can see that $x_p(t) = \frac{1}{2}t \sin t$ is a particular solution. Observe that $\sin t$ is multiplied by t, which causes the amplitude of the solution to go to infinity as $t \to \infty$, hence resonance. This explains why any other forcing function $f(t)$ of fundamental period 2π will also induce this resonance property since the Fourier expansion of $f(t)$ will either have $cos(t)$ or $sin(t)$ term. (Fundamental period because $\sin(nt)$ for $n > 1$ also have period 2π bu won't induce resonance on the system. Clearly 2π is not the fundamental period of $\sin(nt)$ for $n > 1$.)

Now that we know how to solve the problem $x''(t)+x(t) = \cos t$, we revisit the problem

$$x''(t) + x(t) = f(t), t > 0$$

where $f(t)$ is periodic with period 2π and

$$\begin{aligned} f(t) &= t, 0 \le t \le \pi \\ &= 2\pi - t, \pi \le t \le 2\pi. \end{aligned}$$

We now view it as a sum of problems of the form

$$x''(t) + x(t) = A_n \cos(nt), t > 0$$

where A_n is given as above. When $n \neq 1$, this problem has solution

$$x^n(t) = \frac{A_n}{1 - n^2} \cos(nt).$$

When $n = 1$ it has solution

$$x^1(t) = \frac{A_1}{2} t \sin t.$$

Thus a particular solution to the original problem is

$$x(t) = \frac{A_1}{2} t \sin t + \sum_{n=2}^{\infty} \frac{A_n}{1 - n^2} \cos(nt).$$

Finally, we can easily see that the problem

$$x''(t) + \alpha x(t) = f(t), t > 0$$

$\alpha \neq 1$ and $f(t)$ given as above has a particular solution that is representable by Fourier series. Similarly,

$$x''(t) + x(t) = f(t), t > 0$$

also has a particular solution that is representable by Fourier series if we modify $f(t)$ to

$$\begin{aligned} f(t) &= t, 0 \leq t \leq L \\ &= 2L - t, L \leq t \leq 2L, \end{aligned}$$

where $L \neq \pi$. Indeed in this case we need to assume $x_p(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(\frac{n\pi}{L} t)$ and plugging into the equation we have

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} a_n (1 - (\frac{n\pi}{L})^2) \cos(nt) = f(t).$$

The term $1 - (\frac{n\pi}{L})^2 \neq 0$ for all integers n as long as $L \neq \pi$.

3.3.5 Complex Fourier series

Recall the Euler formula

$$e^{ix} = \cos(x) + i \sin(x)$$

and conversely,

$$\begin{aligned} \cos(x) &= \frac{e^{ix} + e^{-ix}}{2} \\ \sin(x) &= \frac{e^{ix} - e^{-ix}}{2i}. \end{aligned}$$

Given this connection, we can expect that the Fourier series can be represented in terms of the complex exponential e^{ix} instead of sine and cosine. Indeed, the collection of functions

$$\left\{ \cdots, e^{\frac{-3i\pi x}{p}}, e^{\frac{-2i\pi x}{p}}, e^{\frac{-i\pi x}{p}}, 1, e^{\frac{i\pi x}{p}}, e^{\frac{2i\pi x}{p}}, e^{\frac{3i\pi x}{p}}, \cdots \right\}$$

is orthogonal on $[-p, p]$. Note that the sequence is double sided : this has to do with the fact that we need e^{-ix} to represent sine and cosine. It also has to do with the fact that if $f(x)$ is a complex valued function (such as e^{ix}) then the L_2 norm of $f(x)$ is defined as

$$\sqrt{\int_{-p}^{p} f(x)\overline{f(x)} dx}.$$

To verify orthogonality, for any $-\infty < m \neq n < \infty$, by Euler's formula

$$\begin{aligned}\int_{-p}^{p} e^{\frac{in\pi x}{p}} e^{\frac{im\pi x}{p}} dx &= \frac{p}{i\pi(n+m)} e^{\frac{i\pi(n+m)x}{p}} \Big|_{-p}^{p} \\ &= \frac{p}{i\pi(n+m)} (e^{i(n+m)\pi} - e^{-i(n+m)\pi}) \\ &= \frac{p}{2\pi(n+m)} \sin(n+m)\pi = 0.\end{aligned}$$

Similarly,

$$\int_{-p}^{p} (1) e^{\frac{in\pi x}{p}} dx = \frac{p}{\pi n} \sin n\pi = 0.$$

Now the L_2 norm squared of $e^{\frac{in\pi x}{p}}$ is

$$\int_{-p}^{p} e^{\frac{in\pi x}{p}} e^{\frac{-in\pi x}{p}} dx = 2p.$$

Therefore, if $f(x)$ has complex Fourier series representation

$$f(x) = c_0 + \sum_{n=-\infty}^{\infty} e^{\frac{in\pi x}{p}},$$

then

$$c_n = \frac{1}{2p} \int_{-p}^{p} f(x) e^{\frac{-in\pi x}{p}} dx, \forall n.$$

Example 3.3.4. *Let $f(x) = e^{-x}$ on $[-\pi, \pi]$. Then*

$$\begin{aligned}c_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-x} e^{-inx} dx = -\frac{1}{2\pi(in+1)} \left(e^{-i(n+1)x} - e^{(in+1)x}\right) \\ &= -\frac{1}{2\pi(in+1)} \left((-1)^n e^{-\pi} - (-1)^n e^{\pi}\right) \\ &= (-1)^n \frac{\sinh \pi}{\pi} \frac{1 - in}{n^2 + 1}.\end{aligned}$$

Thus

$$f(x) = \sum_{n=-\infty}^{\infty} (-1)^n \frac{\sinh \pi}{\pi} \frac{1 - in}{n^2 + 1} e^{inx}.$$

3.3.6 Fundamental period and fundamental frequency

Let $f(x)$ defined on $[-p,p]$ have Fourier series

$$\begin{aligned} f(x) &= \frac{a_0}{2}+\sum_{n=1}^{\infty} a_n \cos\frac{n\pi}{p}x + b_n \sin\frac{n\pi}{p}x \\ &= \sum_{n=-\infty}^{\infty} c_n e^{\frac{n\pi}{p}x}. \end{aligned}$$

In both of the representation, the fundamental period (measured in time, usually seconds) is $T=2p$. The fundamental frequency (measured in Hertz, which is number of revolutions per second) is $f=\frac{1}{T}$. If we map the periodic motion to a circular motion, then each revolution corresponds to a 2π radian. Hence we can define the angular frequency (which serves a similar role to angular speed) as $\omega = 2\pi f$, whose unit is radian per second. Thus the fundamental angular frequency is $\omega = \frac{2\pi}{T} = \frac{p}{\pi}$ and the Fourier series can be written as

$$\begin{aligned} f(x) &= \frac{a_0}{2}+\sum_{n=1}^{\infty} a_n \cos n\omega x + b_n \sin n\omega x \\ &= \sum_{n=-\infty}^{\infty} c_n e^{n\omega x}. \end{aligned}$$

The graph of $(n\omega, |c_n|)$ is referred to as the frequency spectrum of f.

Chapter 4

Basic Partial Differential Equations

4.1 Boundary Value Problems (BVP) and Regular Sturm-Liouville Problems

4.1.1 Introduction - Linear operator in infinite dimensional space

Linear operators on finite dimensional space (such as $\mathbb{R}^n$) can always be represented by matrices. Matrices are relatively “simpler" to study in terms of their properties (this does not mean that theorems about matrices are not deep or sometimes very hard to prove) due to their finite dimensional nature. In infinite dimensional space (such as $C^2[a,b]$, the space of twice differentiable functions on $[a,b]$) we also have linear operators, such as differential operators. In fact, the study of solutions to linear differential equations can be viewed as the search for the eigenfunction corresponding to the eigenvalue 0 of this linear operator. As we have seen, among the matrices symmetric matrices have very nice properties : all their eigenvalues are real and their eigenvectors form an orthogonal basis. The analogy of symmetric matrices in the infinite dimensional space is self-adjoint operator. It is beyond the scope of this course to study such operator in abstract. We present instead a concrete example of such an operator, in the form of a differential equation: the Sturm-Liouville problem.

4.1.2 Regular Sturm-Liouville problem

Let p, q, r be real-valued functions on $[a,b]$ with sufficiently nice property, $r(x) > 0, p(x) > 0 \forall x$. A regular Sturm-Liouville problem with coefficients p, q, r is

$$\frac{d}{dx}(r(x)y') + (q(x) + \lambda p(x))y = 0,$$

subject to

$$\begin{aligned} A_1 y(a) + B_1 y'(a) &= 0 \\ A_2 y(b) + B_2 y'(b) &= 0. \end{aligned}$$

Remarks:

a. This is a boundary value problem, in the sense that the solution is defined on a (usually spatial) interval $[a, b]$ and its (and its derivative's) values are specified on the boundary points a, b. Compare this with the initial value problem where the solution is defined on a (usually time) interval $[0, \infty)$ and values are specified at the initial time $t = 0$.

b. The boundary conditions are said to be homogeneous (it is non-homogenous if the RHS are non-zero). The differential equation is linear. Homogenous boundary conditions and the linear DE imply that a linear combination of solutions is still a solution (which is not the case for the non-homogenous boundary conditions).

4.1.3 Eigenvalues and eigenfunctions

Definition 4.1.1. *The values of λ so that the regular Sturm-Liouville problem has a non-trivial solution are called the eigenvalues of the problem. The solution(s) correspond to a specific eigenvalue is called the eigenfunction(s) associated to this eigenvalue.*

Remark: Let $\mathcal{L}$ be a linear operator (that maps function to function). An example would be the derivative operator. A more interesting example would be

$$\mathcal{L} = \frac{d^2}{dt^2} + 2\frac{d}{dt} - 4I$$

where I is the identity operator. In this sense the LHS of a linear DE can be seen as the image of a linear operator map. One usually define an eigenvalue associated with $\mathcal{L}$ as the value λ so that there exists a non-zero function f such that

$$\mathcal{L}f = \lambda f.$$

f is said to be the eigenfunction associated with eigenvalue λ. In this way it is an extension of the notion of eigenvalue / eigenvector of a matrix.

The way we (and the textbook) define eigenvalue / eigenfunction of a Sturm-Liouville problem above is similar, except for a negative sign and a weight function $p(x)$. That is

$$\frac{d}{dx}(r(x)y') + q(x)y = -\lambda p(x)y,$$

so the linear operator is

$$\mathcal{L} = \frac{d}{dx}(r(x)\frac{d}{dx}) + q(x)I.$$

The function $p(x)$ is interpreted as the weight function that we will use to define orthogonality of eigenfunctions in the next section.

Example 4.1.2. *Consider the regular Sturm-Liouville problem*

$$y'' + \lambda y = 0, y(0) = 0, y(L) = 0.$$

This is a model of the deflection of a beam (see section 3.9 of the textbook). It is a special case of the Sturm-Liouville problem described in the previous section with $r(x) = p(x) = 1, q(x) = 0, A_1 = A_2 = 1, B_1 = B_2 = 0$.

Case 1 : $\lambda = 0$. *The solution to the DE is* $y = c_1 x + c_2$. *Applying the boundary conditions imply* $c_1 = c_2 = 0$. *Thus* $y = 0$ *is the only solution.*

Case 2 : $\lambda = -\alpha^2 < 0$. *The solution to the DE is* $y = c_1 \cosh \alpha x + c_2 \sinh \alpha x$. *Applying the boundary conditions imply* $c_1 = c_2 = 0$. *Thus again* $y = 0$ *is the only solution.*

Case 3 : $\lambda = \alpha^2 > 0$. *The solution to the DE is* $y = c_1 \cos \alpha x + c_2 \sin \alpha x$. *Applying the boundary conditions imply* $c_1 = 0$ *and* $\alpha = \frac{n\pi}{L}$. *Thus a nonzero solution is*

$$y = c_2 \sin \frac{n\pi}{L} x$$

.

Eigenvalues and eigenfunctions: From the three cases we see that the eigenvalues are $\lambda = \left(\frac{n\pi}{L}\right)^2$ *and the associated eigenfunctions are* $\sin \frac{n\pi}{L}$.

4.1.4 Properties of regular Sturm-Liouville problem

a. There exist an infinite number of real eigenvalues that can be arranged in increasing order $\lambda_1 < \lambda_2 < \cdots < \lambda_n < \cdots$ such that $\lambda_N \to \infty$ as $n \to \infty$.

b. The dimension of the eigenspace corresponding to each eigenvalue is 1. That is for each eigenvalue, there is only one eigenfunction (excluding scalar multiples)

c. Eigenfunctions corresponding to distinct eigenvalues are independent.

d. The set of eigenfunctions form an orthogonal set with respect to weight function $p(x)$ on $[a, b]$.

Remarks: Compare how similar these properties are to those of symmetric matrices (except perhaps for property b).

Lemma 4.1.3. *Let* y_1, y_2 *be two eigenfunctions corresponding to two distinct eigenvalues* λ_1, λ_2. *Then*

$$\begin{aligned} y_1(a)y_2'(a) &= y_2(a)y_1'(a) \\ y_1(b)y_2'(b) &= y_2(b)y_1'(b). \end{aligned}$$

Remark: We already have $y_1 \neq y_2$ from property c. This lemma will be useful to show the self-adjoint property of the Sturm-Liouville operator below. *Proof.*

By assumption, y_1, y_2 satisfy

$$\begin{aligned} A_1 y_1(a) + B_1 y_1'(a) &= 0 \\ A_1 y_2(a) + B_1 y_2'(a) &= 0. \end{aligned}$$

Since the boundary condition is not trivial, A_1, B_1 cannot be both 0. That is the determinant of the above system must be 0: $y_1(a)y_2'(a) - y_1'(a)y_2(a) = 0$. The equation at b is similar.

4.1.5 Self-adjoint form

4.1.6 Definition

As mentioned in the introduction, self-adjoint operator is a generalization of symmetric matrices in infinite dimensional space and the Sturm-Liouville operator is an example of it. First, we clarify how to generalize the notion of symmetric matrices into an operator. Recall that we define a matrix to be symmetric if $a_{ij} = a_{ji}, \forall i, j$. This characterization clearly cannot be generalized. Instead, we look at how symmetric interacts with inner-product, as we also have inner product in infinite dimensional space. Indeed, a matrix A is symmetric if and only if for any vectors $\mathbf{u}, \mathbf{v}$

$$(A\mathbf{u})^T\mathbf{v} = u^T(A\mathbf{v}).$$

The proof is quite simple and the direction when A is symmetric is obvious. On the other hand, let $\mathbf{u} = \varepsilon_i, \mathbf{v} = \varepsilon_j, i, j = 1, \cdots, n$ where ε_i denotes the vector with 1 a the ith entry and 0 otherwise. Then we can see the above equation implies that $A_{ij} = A_{ji}$.

Now similarly we can define an operator $\mathcal{L}$ to be self-adjoint if for all functions u, v in the domain of $\mathcal{L}$

$$(\mathcal{L}u, v) = (u, \mathcal{L}v),$$

where $(\cdot, \cdot)$ denotes the inner product between two functions.

4.1.7 Sturm-Liouville operator

We now show that the Sturm-Liouville operator

$$\mathcal{L} = \frac{d}{dx}(r(x)\frac{d}{dx}) + q(x)I$$

is self-adjoint on the domain that is the eigenfunctions of the problem.

First note that since the sum of self-adjoint operators is self-adjoint, and it is clear that $q(x)I$ is self-adjoint, we only need to check the operator

$$\mathcal{L} = \frac{d}{dx}(r(x)\frac{d}{dx}).$$

By integration by parts

$$\int_a^b \frac{d}{dx}(r(x)u'(x))v(x)dx = r(x)u'(x)v(x)\Big|_a^b - \int_a^b r(x)u'(x)v'(x)dx.$$

On the other hand,

$$\int_a^b \frac{d}{dx}(r(x)v'(x))u(x)dx = r(x)v'(x)u(x)\Big|_a^b - \int_a^b r(x)u'(x)v'(x)dx.$$

From (4.1.3)

$$r(x)u'(x)v(x)\Big|_a^b = r(x)v'(x)u(x)\Big|_a^b$$

and the proof is complete.

4.1.8 Converting to self-adjoint form

We have seen that self-adjoint operators enjoy nice properties: real eigenvalues, orthogonal eigenvectors etc. The Sturm-Liouville operator is self-adjoint, thus the eigenvalues and eigenfunctions of the Sturm-Liouville differential equations have the same properties. It is clear that not all differential equations are Sturm-Liouville type and hence not self-adjoint. However, we can transform an linear ODE into a self-adjoint form using the following technique.

Consider the ODE:

$$a(x)y'' + b(x)y' + (c(x) + \lambda d(x))y = 0 \tag{4.1}$$

with appropriate boundary conditions. Recall the Sturm-Liouville problem:

$$\frac{d}{dx}(r(x)y') + (q(x) + \lambda p(x))y = 0$$

which is equivalent to

$$r(x)y'' + r'(x)y' + (q(x) + \lambda p(x))y = 0.$$

Thus the ODE (4.1) would be a Sturm-Liouville problem if

$$\frac{d}{dx}b(x) = a(x).$$

This clearly cannot always be the case. On the other hand, the ODE (4.1) can be turned into a Sturm-Liouville problem if there is a function $m(x)$ so that

$$m(x)[a(x)y'' + b(x)y'] = \frac{d}{dx}\Big(r(x)y'\Big).$$

That is

$$\begin{aligned} m(x)a(x) &= r(x) \\ m(x)b(x) &= r'(x). \end{aligned}$$

Assuming $m(x) \neq 0$, we have

$$\frac{r'(x)}{r(x)} = \frac{b(x)}{a(x)}.$$

That is

$$\frac{d}{dx}(\log(r(x)) = \frac{b(x)}{a(x)}$$

or

$$r(x) = e^{\int \frac{b(x)}{a(x)} dx}.$$

Reversing the above argument, we can see that by multiplying both sides of (4.1) by $m(x) = \frac{e^{\int \frac{b(x)}{a(x)} dx}}{a(x)}$ the ODE (4.1) is transformed a Sturm-Liouville problem with $r(x) = e^{\int \frac{b(x)}{a(x)} dx}$, which is a self-adjoint form.

4.2 Separable Partial Differential Equations

4.2.1 Introduction

A partial differetial equation (PDE) is an equation that involes the partial derivatives of a multi-variate function. The variables can be purely spatial (x, y, z) such as the wave equation or both temporal and spatial (t, x) such as the heat equation. There is no universal technique to find explicit solution of a PDE (even a linear one). In fact, explicit solution is rather the exception than the norm when one investigates a PDE. Nevertheless, we will study several classes of PDE that yield more or less explicit solutions for the remainder of the course.

A useful thing to keep in mind is the techniques we study will yield a *particular solution* to the PDE, but not necessarily the *general solutions* of the PDE. For example, in the equation

$$u_{xx} + u_{yy} = 0$$

a particular solution would be $u(x, y) = cxy$ where c is a constant. But one easily sees that another possible solution is $u(x, y) = ax + by$, a, b are constants.

4.2.2 Separable PDEs

A particularly useful technique when we look for an explicit solution of a PDE is to make an *ansatz*, that is a guess for the functional form of the solution. The guess should certainly be based on the structure of the equation, for example the wave equation mentioned above

$$u_{xx} + u_{yy} = 0$$

can be re-written as

$$u_{xx} = -u_{yy}.$$

So it is natural to guess that a form of $u(x, y)$ is

$$u(x, y) = a(x)b(y), \tag{4.2}$$

for some function a, b (since then the part of x is unaffacted by differentiation with respect to y and vice versa). Any solution of $u(x, y)$ in the form (4.2) is referred to as a *product solution* of the PDE, and the technique of finding a product solution is called *separation of variables*. Lastly the PDE is said to be *separable* if we can use separation of variables to find a solution for it.

An example

Find the product solution of

$$u_{xx} = 4u_y.$$

Sol:
Let $u(x,y) = a(x)b(y)$. Then the equation becomes

$$a_{xx}(x)b(y) = 4a(x)b_y(y).$$

That is

$$\frac{a_{xx}(x)}{4a(x)} = \frac{b_y(y)}{b(y)}.$$

Since the LHS only depends on x and the RHS only on y, it means that they both equals to a constant $-c$. Solving

$$\frac{b_y}{b} = -c$$

gives $b(y) = Ke^{-cy}$ for some arbitrary constant K. The second order ODE

$$a_{xx} + 4ca = 0$$

has solution

$$a(x) = c_1 e^{\sqrt{2|c|}x} + c_2 e^{-\sqrt{2|c|}x}.$$

if $c < 0$. If $c > 0$ then it has solution

$$a(x) = c_1 \cos(\sqrt{2c}x) + c_2 \sin(\sqrt{2c}x).$$

Note: For the particular choice $c_1 = c_2 = \frac{1}{2}$ and $c_1 = -c_2 = \frac{1}{2}$ we have $\sinh(\sqrt{2c}x)$ and $\cosh(\sqrt{2c}x)$ as general solution. Thus it is also possible, and indeed common, to express the general solution when $c > 0$ in terms of $\sinh$ and $\cosh$. Lastly if $c = 0$ then it has solution

$$a(x) = c_1 + c_2 x.$$

4.2.3 Classification of second order linear PDEs

A second order linear PDE has the form

$$Au_{xx} + Bu_{xy} + Cu_{yy} + Du_x + Eu_y + Fu = G,$$

where A, B, C, D, E, F, G are constants. It is *second order* because the highest partial derivative has second order. It is *linear* because a linear combination of solutions to such an equation is also a solution. We have the following classifications:

a. The equation is **elliptic** if $B^2 - 4AC < 0$
b. The equation is **parabolic** if $B^2 - 4AC = 0$
c. The equation is **hyperbolic** if $B^2 - 4AC > 0$.

Example 4.2.1. *The* **heat equation**

$$u_t - ku_{xx} = 0, k > 0$$

is parabolic. The **wave equation**

$$u_{tt} - \alpha^2 u_{xx} = 0$$

is hyperbolic. The **Laplace equation**

$$u_{xx} + u_{yy} = 0$$

is elliptic.

These classifications are important because the techniques to solve different types of PDEs are very different. Also, different types of PDEs model different physical phenomena. For example, the parabolic PDEs usually model the temperature of an object (thus it is called the **heat equation**) while the hyperbolic PDEs usually model the displacement of an object from its equilibrium (thus it is called the **wave equation**).

Remarks:

The names elliptic, parabolic, hyperbolic comes from the analogy of the form of the PDE with the representation of the conic sections in quadratic form:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0.$$

This equation represents an ellipse if $B^2 - 4AC < 0$, a parabola if $B^2 - 4AC = 0$ and a hyperbola if $B^2 - 4AC > 0$. See also here for more details.

4.3 Heat Equation and Boundary Value Problems (BVPs)

4.3.1 Derivation of the heat equation

Consider a rod with length L and cross sectional area A. We assume that this rod is insulated except possibly at the two ends (that is no heat goes in and out anywhere of the rod except possibly at the two ends). We denote the temperature at time $t > 0$ at a point $x, 0 < x < L$ of the rod as $u(t, x)$.

We need to introduce some concepts before deriving the heat equation. First we define the notion of heat. The quantity of heat in an element of mass m and temperature u is defined as

$$Q = \gamma m u. \tag{4.3}$$

That is the heat quantity is proportional to the mass and temperature of the object. The propotional constant γ is called the *specific heat*. This makes sense as the more mass, the higher temperature and / or higher specific heat should all lead to higher heat quantity.

Remark 4.3.1. *It is worth emphasizing that formula (4.3) is an idealized formula for a point mass object. For an extended object that may have different temperature at different points, an integration (adding over the heat quantity of the point mass components) is needed to get the heat quantity of that object.*

Another phenomenon we need to introduce is *heat conduction*, which is the transfer of heat within a body. Specifically, imagine we have a cool rod at the beginning of the experiment:

$$u(0, x) = 0, 0 \leq x \leq L.$$

Suppose we heat the rod at the left end $x = 0 : u(t, 0) = C > 0, t > 0$. This will make the rest of the rod increase in temperature, that is $u(t, a) > 0$ for $a > 0, t > 0$. Obviously some heat quantity has been transfered from $x = 0$ to $x = a > 0$.

The law that governs the heat conduction is called the Fourier's law, and for the one dimensional rod in consideration it is as followed

$$Q_t = -KAu_x, \tag{4.4}$$

where K is the constant called the thermal conductivity (of the rod's material). Approximating u_x with $\frac{u(t,x+dx)-u(t,x)}{dx}$, Fourier's law states that

$$Q_t \approx -KA\frac{u(t, x + dx) - u(t, x)}{dx}.$$

That is, roughly speaking, the rate of change of heat is proportional to the cross sectional area A and the temperature difference $u(t, x + dx) - u(t, x)$. This essentially is Newton's law of cooling. The difference here is Fourier's law states rather that the rate of change of heat is proportional to the cross sectional area A and the tempreature gradient $\frac{u(t,x+dx)-u(t,x)}{dx}$ (in discrete version) or u_x (in continuous version). Either way this makes emperical sense. Lastly the negative sign captures the convention that heat flows from higher temperature to lower temperature. That is if $u(t, x) > u(t, x + dx)$ then there is a positive heat flow from (t, x) to $(t, x + dx)$ thus $Q_t(t, x)$ should be positive. The negative sign guarantees this happens. Conversely, if we compute $Q_t(t, x)$ to be positive then we can conclude that heat flows (instantaneously) from x to the right at time t. For more details on Fourier's law and heat conduction see here.

Now consider a slice of the rod $[x_0, x_0 + \Delta]$ with length Δ. Let the mass density of the rod's material be ρ. We can imagine the slice is composed of (infinitely many) point masses with coordinate x and "length" dx. By formula (4.3), the heat quantity in each of this point mass x at time t is

$$Q(t, x) = \gamma\rho(Adx)u(t, x).$$

The heat quantity of the slice is the (continous) sum of the heat quantity in each of the point mass:

$$Q(t, x_0, \Delta) = \int_{x_0}^{x_0+\Delta} \gamma\rho Au(t, x)dx.$$

We want to analyze the *rate* of heat change in this slice $Q(t, x_0, \Delta)$. That is how fast heat goes in (or out) of this slice at any given time. By definition, it is just the derivative of $Q(t, x_0, \Delta)$ with respect to t:

$$\frac{d}{dt}Q(t, x_0, \Delta) = \int_{x_0}^{x_0+\Delta} \gamma\rho A u_t(t, x)dx. \tag{4.5}$$

Note that we have interchanged the order of differentiation and integration in the RHS of the above equation, which can be justified under certain technical assumptions.

On the other hand, since the rod is insulated, from *conservation of energy* the heat change from the slice must results from the change at the two ends of the slice $[x_0, x_0+\Delta]$. That is the rate of change of the heat in the slice can also be expressed as

$$\frac{d}{dt}Q(t, x_0) - \frac{d}{dt}Q(t, x_0 + \Delta).$$

This is simply saying the rate of change of heat is equal to the rate of heat going in at the left end minus the rate of heat going out at the right end. For example, if $\frac{d}{dt}Q(t, x_0 + \Delta) = \frac{d}{dt}Q(t, x_0)$ then the heat change at one end balances the heat change change at the other end and the rod's heat remains constant (over time). If heat flows in at the left end x_0 then $\frac{d}{dt}Q(t, x_0) > 0$ and no heat flows out at the right end $x_0 + \Delta$ then $\frac{d}{dt}Q(t, x_0) = 0$ and the net heat change in the slice is positive.

By Fourier's law (4.4),

$$\frac{d}{dt}Q(t, x_0) - \frac{d}{dt}Q(t, x_0 + \Delta) = KA\Big\{u_x(t, x_0 + \Delta) - u_x(t, x_0)\Big\}. \tag{4.6}$$

Equating (4.5) and (4.6) gives

$$\int_{x_0}^{x_0+\Delta} \gamma\rho A u_t(t, x)dx = KA\Big\{u_x(t, x_0 + \Delta) - u_x(t, x_0)\Big\}.$$

Dividing both sides by Δ and let $\Delta \to 0$ gives

$$u_t = \frac{K}{\gamma\rho}u_{xx} = ku_{xx}.$$

This is the heat equation. The constant $k = \frac{K}{\gamma\rho}$ is called the *thermal diffusivity* (of the material fo the rod).

4.3.2 Heat equation as a boundary value problem

The boundary value problem

Consider an insulated rod with length L whose temperature at the two ends are always kept at 0:

$$u(t, 0) = u(t, L) = 0, t > 0.$$

Suppose that its initial temperature profile is given by a function $f(x)$:

$$u(0, x) = f(x), 0 < x < L.$$

From our derivation of the heat equation above, the temperature $u(t, x)$ of the rod at any time t and position x is described by the solution to the BVP

$$u_t = ku_{xx} \tag{4.7}$$
$$u(t, 0) = u(t, L) = 0, t > 0 \tag{4.8}$$
$$u(0, x) = f(x), 0 < x < L. \tag{4.9}$$

The general solution

We will now solve this BVP. First we look for the solution to the equation

$$u_t = ku_{xx}$$

without worrying about the boundary and initial conditions. This equation is separable, that is we make the ansatz

$$u(t, x) = A(t)B(x).$$

Note that we have considered this equation in the example of the last section. Plugging in to the equation we have

$$\frac{A_t}{kA} = \frac{B_{xx}}{B} = -\lambda,$$

for some constant λ. Thus A, B satisfy respectively the DEs

$$A_t + k\lambda A = 0$$
$$B_{xx} + \lambda B = 0.$$

The solution for $A(t)$ is

$$A(t) = Ce^{-k\lambda t},$$

for some constant C to be determined. The solution for $B(x)$ depends on the sign of λ and it is

$$B(x) = C_1 x + C_2, \lambda = 0$$
$$B(x) = C_1 \cos(\alpha x) + C_2 \sin(\alpha x), \lambda = \alpha^2 > 0$$
$$B(x) = C_1 \cosh(\alpha x) + C_2 \sinh(\alpha x), \lambda = -\alpha^2 < 0.$$

Remark 4.3.2. *We observe that we look for non-zero solutions in both $A(t), B(x)$. The reason is if either one is zero then $u(t, x) = 0$ and it contradicts the initial condition $u(0, x) = f(x)$ (unless $f(x) = 0$). This will be a general theme in subsequent sections when we investigate the wave and Laplace equations which are also separable.*

4.3.3 The boundary condition

We now consider the boundary condition. $u(t,0) = u(t,L) = 0$ implies

$$A(t)B(0) = A(t)B(L) = 0.$$

Since $A(t)$ is not the zero function, we conclude that $B(0) = B(L) = 0$ and note that $B(x)$ solves the regular Sturm-Liouville problem

$$\begin{aligned} B_{xx} + \alpha^2 B &= 0 \\ B(0) = B(L) &= 0. \end{aligned}$$

The condition $B(0) = B(L) = 0$ implies that $\lambda = \alpha^2 > 0$ and

$$B(x) = C_1 \cos(\alpha x) + C_2 \sin(\alpha x)$$

since the other two possibilities of λ also forces $B(x) = 0$, which again implies $u(t,x) = 0$.

The condition $B(0) = 0$ implies $C_1 = 0$. The condition $B(L) = 0$ implies

$$\sin(\alpha L) = 0.$$

Thus (recalling that we have the freedom to choose what $\lambda = \alpha^2$ is)

$$\alpha = \frac{n\pi}{L},$$

for some natural number n. For each choice of $\alpha = \frac{n\pi}{L}$, it is conceivable that we have a different corresponding constant C_2^n. That is we have a family of solution

$$B_n(x) = C_n \sin(\frac{n\pi}{L}x).$$

Since the boundary condition is homogeneous, the sum of solutions is a solution and thus the general solution to the DE:

$$\begin{aligned} B_{xx} + \alpha^2 B &= 0 \\ B(0) = B(L) &= 0 \end{aligned}$$

is

$$B(x) = \sum_{n=1}^{\infty} C_n \sin(\frac{n\pi}{L}x).$$

Thus the general solution to the heat equation, considering only the boundary condition is

$$u(t,x) = \sum_{n=1}^{\infty} C_n e^{-k\frac{n^2\pi^2}{L^2}t} \sin(\frac{n\pi}{L}x).$$

(The constant C in the general solution of $A(t)$ is absorbed into the C_n in the above representation).

Remark: As $t \to \infty$, each term of the series converges to 0. Thus we may believe that $u(t,x) \to 0$ as $t \to \infty$ (we need to justify exchanging the limit and the summation to make it rigorous). This agrees with our intuitition that the temperature of the rod converges to 0 everywhere due to the fact that the two ends are kept at constant 0 degree.

The initial condition

The condition $u(0,x) = f(x)$ implies

$$\sum_{n=1}^{\infty} C_n \sin(\frac{n\pi}{L}x) = f(x).$$

That is $C_n's$ are the Fourier coefficients in the half-range expansion of $f(x)$ on the interval $[0, L]$ using the sine series. Recalling our results in the Fourier chapter gives

$$C_n = \frac{2}{L}\int_0^L f(x)\sin(\frac{n\pi}{L}x)dx.$$

Thus the solution to the heat equation (4.7) is

$$u(t,x) = \sum_{n=1}^{\infty}\Big(\frac{2}{L}\int_0^L f(x)\sin(\frac{n\pi}{L}x)dx\Big)e^{-k\frac{n^2\pi^2}{L^2}t}\sin(\frac{n\pi}{L}x).$$

An example

Consider the heat equation

$$\begin{aligned} u_t &= u_{xx} \\ u(t,0) &= u(t,\pi) = 0, t > 0 \\ u(0,x) &= 1, 0 < x < \pi. \end{aligned}$$

The Fourier coefficient C_n is

$$C_n = \frac{2}{\pi}\int_0^{\pi}\sin(nx)dx = \frac{2}{\pi}\frac{-1-(-1)^n}{n}.$$

That is

$$u(t,x) = \frac{2}{\pi}\sum_{n=1}^{\infty}\frac{-1-(-1)^n}{n}e^{-n^2t}\sin(nx).$$

4.4 Wave Equation and Boundary Value Problems (BVPs)

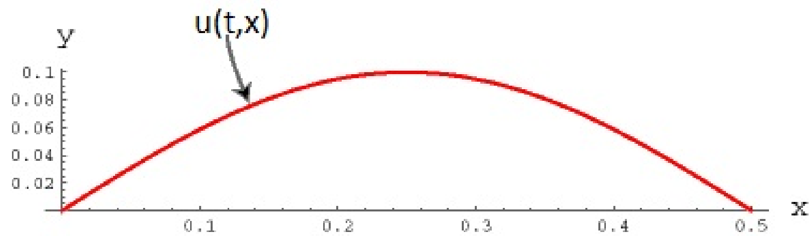
4.4.1 Derivation of the wave equation

Consider a string of length L that originally (*before* time $t = 0$) lies along the interval $[0, L]$ on the x-axis. Suppose at time $t = 0$ we pluck the string and let go. Let $u(t,x)$ be the displacement of the string away from the x-axis at time $t > 0$ and coordinate $x, 0 < x < L$. We suppose that the string is clamped at the two ends:

$$u(t,0) = u(t,L) = 0, t > 0.$$

The “shape" of the string given by the pluck is described by a function $f(x)$:

$$u(0,x) = f(x).$$



We seek to find an equation that $u(t,x)$ follows, under the following assumptions:

a. The displacement $u(t,x)$ is always small compared with the length L of the string. This has the practical consequence that the length of any segment of string whose x coordinates lie between $[x, x+\Delta x]$ for some small Δx can be approximated by Δx.

b. The string is light. That is the effect of gravity on the string is negligible. Consequently, the motion of the string is the result of only the tension of the string itself.

c. The magnitude of the tension of the string is constant in time and uniform through out the string.

Remarks: It goes without saying that the only “motion" of the string is vertical motion. That is if we fix any coordinate x of the string then the only motion at that coordinate is along the y axis.

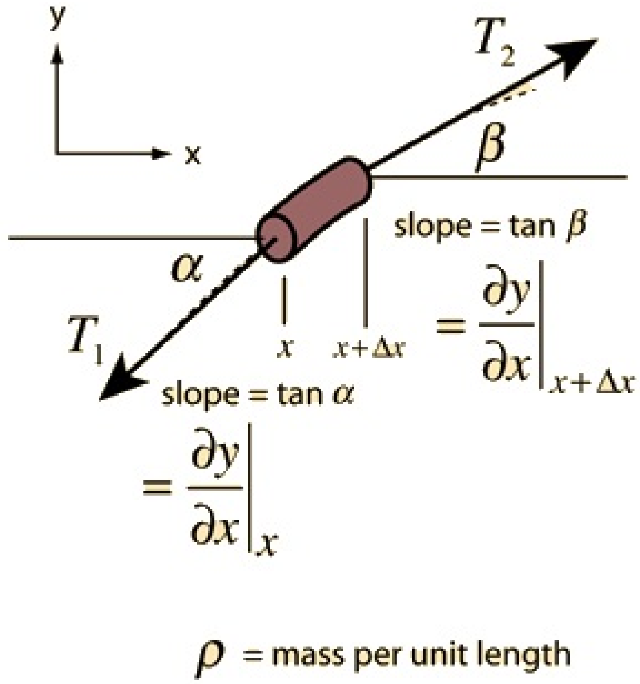
Now we look at the forces that act on a segment of the string whose x coordinates lie between $[x, x+\Delta x]$ for some small Δx. The goal is to apply Newton's second law to this segment, whose (vertical) acceleration can be captured approximately by $u_{tt}(t,x)$ since Δx is small.

From assumptiob b, the only forces acting on this segment are the tensions of the string (the part excluding this segment). Thus there are two tensile forces $\mathbf{T_1}, \mathbf{T_2}$ acting at the two end of the segments as in the picture above. The angles $\mathbf{T_1}, \mathbf{T_2}$ make with the horizontal line are α and β respectively.

Remarks: More technically precise, β is measured counter-clockwise from the positive x-axis to $\mathbf{T_2}$; α is measured counter-clockwise from the negative x-axis to $\mathbf{T_1}$. This would ensure that the expression (4.10) of the net force acting on the segment is correct no matter how $\mathbf{T_1}, \mathbf{T_2}$ point (“up" or "down"). For example, if (contrary to the picture) $\mathbf{T_2}$ points “down" then $\sin\beta$ is negative. Similarly if $\mathbf{T_1}$ points “up" then $\sin\alpha$ is negative.

From assumption a, α, β are close to 0. Thus $\cos(\alpha) \approx \cos(\beta) \approx 1$, which implies $\sin(\alpha) \approx \tan(\alpha)$ ($\sin(\beta) \approx \tan(\beta)$). Finally from assumption c, the maginute of $\mathbf{T_1}$ and $\mathbf{T_2}$ are equal, which we denote as T. Thus the net force that acts on this segment vertically is

$$T\sin\beta - T\sin\alpha \approx T(\tan\beta - \tan\alpha). \tag{4.10}$$



Remark: The negative sign on $T \sin \alpha$ can be thought intuitively as because $\mathbf{T_1}$ always points "left." See also the above remark about how we precisely define α, β.

Geometrically, the tensile forces $\mathbf{T_1}, \mathbf{T_2}$ are tangents to the segment at the two ends. Thus the slope they make with the horizontal line are equal to the derivatives of u with respect to x at these two ends, respectively. Finally, recall that the slope of a line is equal to the tangent of the angle it makes with respect to the horizontal line. That is

$$\begin{aligned} \tan \alpha &= u_x(t, x) \\ \tan \beta &= u_x(t, x + \Delta x) \end{aligned}$$

Thus the net force acting on the segment vertically is

$$T \sin \beta - T \sin \alpha \approx T(\tan \beta - \tan \alpha) = T(u_x(t, x + \Delta x) - u_x(t, x)).$$

Let ρ be the mass per unit length of this segment. By assumption a, the length of this segment is approximately Δx. By Newton's second law, the net vertical force acting on the segment is equal to the mass of the segment, which is approximately $\rho \Delta x$ times the acceleration, which is approximately $u_{tt}(t, x)$. That is

$$T(u_x(t, x + \Delta x) - u_x(t, x)) = \rho(\Delta x) u_{tt}(t, x).$$

All these approximations become exact as $\Delta x \to 0$. Thus dividing both sides by Δx an let $\Delta x \to 0$ we have

$$u_{tt} = a^2 u_{xx},$$

where $a^2 = \frac{T}{\rho}$ is a constant that depends on the material of the string.

4.4.2 Wave equation as a boundary value problem

The boundary value problem

Consider a string of length L that is clamped at the two ends:

$$u(t,0) = u(t,L) = 0, t > 0.$$

Suppose the initial shape of the string is given by a function $f(x)$:

$$u(0,x) = f(x), 0 < x < L.$$

Also suppose that the initial velocity of the string is given by a function $g(x)$:

$$u_t(0,x) = g(x), 0 < x < L.$$

Remark: The initial velocity of the string $u_t(0,x) = g(x)$ tells us the initial velocity of the string at the coordinate x (whether it's going up or down and how fast). In particular, if the string is *plucked*, that is we give it an initial shape $f(x)$ and let go, then $g(x) = 0$.

From our derivation of the wave equation above, the y-coordinate of the string at the point x at time t follows the BVP:

$$\begin{aligned} u_{tt} &= a^2 u_{xx} \\ u(t,0) = u(t,L) &= 0, t > 0 \\ u(0,x) &= f(x), 0 < x < L \\ u_t(0,x) &= g(x), 0 < x < L. \end{aligned}$$

The general solution

Like the heat equation, the wave equation is also separable. Thus again we make the ansatz

$$u(t,x) = A(t)B(x).$$

Plugging in to the equation we have

$$\frac{A_{tt}}{a^2 A} = \frac{B_{xx}}{B} = -\lambda,$$

for some constant λ. Thus A, B satisfy the DEs

$$\begin{aligned} A_{tt} + a^2\lambda A &= 0 \\ B_{xx} + \lambda B &= 0. \end{aligned}$$

The solutions for $A(t), B(x)$ depend on the sign of λ and they are

$$\begin{aligned}
A(t) &= C_1 t + C_2, \lambda = 0 \\
A(t) &= C_1 \cos(a\alpha t) + C_2 \sin(a\alpha t), \lambda = \alpha^2 > 0 \\
A(t) &= C_1 \cosh(a\alpha t) + C_2 \sinh(a\alpha t), \lambda = -\alpha^2 < 0 \\
B(x) &= C_1' x + C_2', \lambda = 0 \\
B(x) &= C_1' \cos(\alpha x) + C_2' \sin(\alpha x), \lambda = \alpha^2 > 0 \\
B(x) &= C_1' \cosh(\alpha x) + C_2' \sinh(\alpha x), \lambda = -\alpha^2 < 0.
\end{aligned}$$

The boundary condition

The boundary condition we are considering is exactly the same as the boundary condition for the heat equation. That is we conclude $B(x)$ satisfies the regular Sturm-Liouville problem

$$\begin{aligned}
B_{xx} + \lambda B &= 0 \\
B(0) = B(L) &= 0.
\end{aligned}$$

Repeating the same argument as in the heat equation, we have a family of solutions

$$B_n(x) = C_n \sin(\frac{n\pi}{L} x),$$

where $n = 1, 2, \cdots$.

From the analysis of the boundary condition, we also have $\lambda_n = \frac{n^2\pi^2}{L^2} > 0$. Thus the form of the general solution for $A(t)$ is:

$$A(t) = C_1 \cos(a\frac{n\pi}{L} t) + C_2 \sin(a\frac{n\pi}{L} t).$$

Thus the general solution for $u(t, x)$ is

$$u(t,x) = \sum_{i=1}^{\infty} \Big\{ C_n^1 \cos(a\frac{n\pi}{L} t) + C_n^2 \sin(a\frac{n\pi}{L} t) \Big\} \sin(\frac{n\pi}{L} x).$$

The initial conditions

Plugging in initial conditions

$$\begin{aligned}
u(0,x) &= f(x), 0 < x < L \\
u_t(0,x) &= g(x), 0 < x < L,
\end{aligned}$$

we have

$$\begin{aligned}
\sum_{i=1}^{\infty} C_n^1 \sin(\frac{n\pi}{L} x) &= f(x) \\
\sum_{i=1}^{\infty} \Big(a\frac{n\pi}{L} C_n^2 \Big) \sin(\frac{n\pi}{L} x) &= g(x).
\end{aligned}$$

That is C_n^1's (resp. $a\frac{n\pi}{L}C_n^2$'s) are the Fourier coefficients in the half-range expansion of $f(x)$ (resp. $g(x)$) on the interval $[0, L]$ using the sine series. Recalling our results in the Fourier chapter gives

$$\begin{aligned} C_n^1 &= \frac{2}{L}\int_0^L f(x)\sin(\frac{n\pi}{L}x)dx \\ C_n^2 &= \frac{2}{an\pi}\int_0^L g(x)\sin(\frac{n\pi}{L}x)dx. \end{aligned}$$

In particular, when the string is *plucked*, $g(x) = 0$ and $C_n^2 = 0, \forall n$.

4.4.3 Standing waves

We first present a technical result:

Lemma 4.4.1. *For all n, there exists C_n^3 and ϕ_n so that*

$$C_n^1\cos(a\frac{n\pi}{L}t) + C_n^2\sin(a\frac{n\pi}{L}t) = C_n^3\sin(a\frac{n\pi}{L}t + \phi_n).$$

Proof: Using trig identity, we have

$$\sin(\frac{n\pi}{L}t + \phi_n) = \sin(\phi_n)\cos(a\frac{n\pi}{L}t) + \sin(a\frac{n\pi}{L}t)\cos(\phi_n).$$

Thus we require

$$\begin{aligned} C_n^3\sin(\phi_n) &= C_n^1 \\ C_n^3\cos(\phi_n) &= C_n^2. \end{aligned}$$

This system has solution

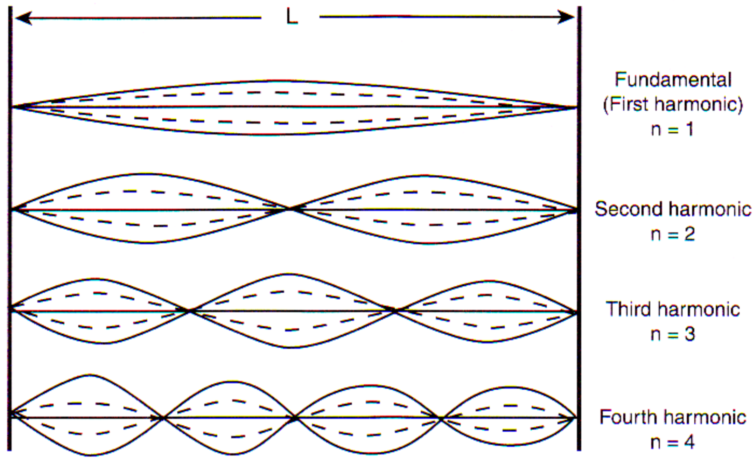
$$\begin{aligned} C_n^3 &= \sqrt{(C_1^n)^2 + (C_2^n)^2} \\ \phi_n &= \tan^{-1}\left(\frac{C_1^n}{C_2^n}\right). \end{aligned}$$

Thus, $u(t, x)$ can be written as

$$\begin{aligned} u(t,x) &= \sum_{i=1}^{\infty} C_n^3\sin(a\frac{n\pi}{L}t + \phi_n)\sin(\frac{n\pi}{L}x) \\ &:= \sum_{i=1}^{\infty} u^n(t,x), \end{aligned}$$

where

$$u^n(t,x) = C_n^3\sin(a\frac{n\pi}{L}t + \phi_n)\sin(\frac{n\pi}{L}x).$$



The $u^n(t, x)$ is referred to as the n-th *standing wave* or the n-th *normal mode*. It is called a standing wave because each u^n has the shape of essentially $\sin(\frac{n\pi}{L}x)$ with time-varying amplitude $C_n^3 \cos(a\frac{n\pi}{L}t + \phi_n)$.

The points on $[0, L]$ for which $\sin(\frac{n\pi}{L}x) = 0$ corresponds to the points of the standing waves where there is no motion. These points are called *nodes*. It is easy to see that the n-th standing wave has $n - 1$ nodes.

The period of the standing wave is the time it takes to complete one cycle (of up and down motion). For each fixed x, the up and down motion is described by the term

$$C_n^3 \sin(a\frac{n\pi}{L}t + \phi_n).$$

Thus the period of the n-th standing wave is

$$T_n = \frac{2L}{an}.$$

The frequency of the n-th standing wave is

$$f_n = \frac{1}{T_n} = \frac{an}{2L} = \frac{n}{2L}\sqrt{\frac{T}{\rho}}.$$

f_1 is called the *fundamental frequency* or the *first harmonic*. For $n > 1$ f_n's are called the *overtones*. In particualr, f_n is referred to as the $n - 1$ overtone.

4.4.4 Wave equation on unbounded domain

Some physical phenomena such as electromagnetic waves also follow the wave equation (see e.g. Maxwell's equation). The difference is they happen on an unbounded domain. In this section we consider the solution of a wave equation on the real line. The difference with the previous section is that there is no boundary condition imposed. (Properly speaking we need to consider electromagnetic waves in 3-d spatial variables. For demonstration purposes, we'll just consider the equation in 1-d spatial variable).

Consider the problem

$$\begin{aligned} u_{tt} &= a^2 u_{xx}, -\infty < x < \infty, t > 0 \\ u(0,x) &= f(x), u_t(0,x) = g(x). \end{aligned}$$

We solve this problem in 3 steps.

Step 1: Change of variables. We let

$$\begin{aligned} \xi &= x + at \\ \eta &= x - at. \end{aligned}$$

By the chain rule:

$$u_x = u_\xi \frac{\partial \xi}{\partial x} + u_\eta \frac{\partial \eta}{\partial x} = u_\xi + u_\eta$$
$$u_t = u_\xi \frac{\partial \xi}{\partial t} + u_\eta \frac{\partial \eta}{\partial t} = a u_\xi - a u_\eta.$$

Similarly

$$u_{xx} = \frac{\partial u_x}{\partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial u_x}{\partial \eta}\frac{\partial \eta}{\partial x} = u_{\xi\xi} + u_{\eta\xi} + u_{\xi\eta} + u_{\eta\eta}$$
$$u_{tt} = \frac{\partial u_t}{\partial \xi}\frac{\partial \xi}{\partial t} + \frac{\partial u_t}{\partial \eta}\frac{\partial \eta}{\partial t} = a^2 u_{\xi\xi} - a^2 u_{\eta\xi} - a^2 u_{\xi\eta} + a^2 u_{\eta\eta}.$$

Multiply u_{xx} with a^2 and subtracting u_{tt}, using the fact that $a^2 u_{xx} - u_{tt} = 0$ we obtain

$$u_{\eta\xi} = 0.$$

Step 2: Integrate $u_{\eta\xi} = 0$

Integrate this equation first with respect to η gives $u_\xi = f(\xi)$, for some function $f(\xi)$ to be determined (so that when we differentiate with respect to η we obtain 0). Integrate again with respect to η gives $u(\xi, \eta) = F(\xi) + G(\eta)$, where $F'(\xi) = f(\xi)$ and $G(\eta)$ is some function of η also to be determined (again so that when we differentiate with respect to ξ we only obtain $f(\xi)$). Converting back to (t, x) variables, we have

$$u(t,x) = F(x + at) + G(x - at).$$

Step 3: Plugging in the initial conditions:

We have

$$\begin{aligned} u(0,x) &= F(x) + G(x) = f(x) \\ u_t(0,x) &= aF'(x) - aG'(x) = g(x). \end{aligned}$$

The second equation can be integrated to give

$$a[F(x) - F(x_0)] - a[G(x) - G(x_0)] = \int_{x_0}^{x} g(s)ds,$$

where x_0 is an arbitrary constant. Letting $c = F(x_0) - G(x_0)$ we obtain the system

$$\begin{aligned} F(x) + G(x) &= f(x) \\ F(x) - G(x) &= \frac{1}{a}\int_{x_0}^{x} g(s)ds + c. \end{aligned}$$

That is

$$F(x) = \frac{1}{2}f(x) + \frac{1}{2a}\int_{x_0}^{x} g(s)ds + c$$
$$G(x) = \frac{1}{2}f(x) - \frac{1}{2a}\int_{x_0}^{x} g(s)ds - c.$$

Finally, plugging in $u(t, x) = F(x + at) + G(x - at)$ gives

$$u(t, x) = \frac{1}{2}[f(x + at) + f(x - at)] + \frac{1}{2a}\int_{x-at}^{x+at} g(s)ds.$$

In particular, when $g(x) = 0$ we see that the solution is the superposition of two traveling waves: $\frac{1}{2}f(x + at)$ traveling to the right and $\frac{1}{2}f(x - at)$ traveling to the left, both with speed a.

4.5 Laplace Equation and BVP in a rectangle

4.5.1 Context for Laplace equations

In this section we study the Laplace equation:

$$\Delta u(x, y) = u_{xx} + u_{yy} = 0, (x, y) \in \Omega$$

where Ω is a rectangle inside $\mathbb{R}^2$ subject to some boundary conditions on the boundary of Ω. Note that unlike the heat or wave equations, conventionally both independent variables of $u(x, y)$ are spatial variables. In this sense there is no time dimension in the Laplace equations. The choice of Ω as a rectangular domain is to facilitate the derivation of explicit solution. We note that one can study the Laplace equation on a general domain (for example, a circle) and in dimension larger than two.

Also unlike the heat or wave equation, we will not present the derivation for the Laplace equation. The reason is there are many contexts where the Laplace equation arises, e.g. describing the steady state solution of the heat or wave equation in higher dimensions (see below). For more information about the situation where the Laplace equation arises, see e.g. here.

4.5.2 The Laplace equation as steady state solution to heat and wave equation in higher dimensions

The main principle behind the heat equation is Fourier's law, which says in one dimension

$$Q_t = -KAu_x.$$

The quantity $q = \frac{Q_t}{A}$ is called the local local heat flux density and in higher dimension we have

$$q = -K\nabla u.$$

(Note: q is a vector in the above equation). Thus we can believe that arguing in a similar way to the one dimensional case, by analyzing the rate of change of the heat in a rectangle, we can derive the heat equation in two dimension as

$$u_t = k(u_{xx} + u_{yy}). \tag{4.11}$$

The wave equation in two dimension requires the analysis of the tension in a membrane (e.g. that of a drum), which is beyond the scope of this note. We will accept that the wave equation in two dimension is

$$u_{tt} = \alpha^2(u_{xx} + u_{yy}). \tag{4.12}$$

A steady state is a state where the system's property (the temperature of the rod, the displacement of the string) becomes steady, i.e. does not change with time. For example, the solution of the heat equation for a rod subject to the zero boundary condition tends to $u(t, x) = 0$ as $t \to \infty$. Thus $u(t, x) = 0$ is the steady state solution for this particular heat equation.

We maynot know whether a system will converge to a steady state solution. On the other hand, if a steady state is reached then by definition we must have

$$u_t = 0$$

(which also implies $u_{tt} = 0$). Plugging this into (4.11) and (4.12) we obtain in both cases

$$u_{xx} + u_{yy} = 0,$$

which is the Laplace equation. The time dimension disappears since the steady state does not depend on time by definition.

4.5.3 Some other contexts for the Laplace equation

We mentioned some other contexts where the Laplace equation may arise. First, in complex analysis if we have a complex function

$$f(z) = u(x, y) + iv(x, y)$$

where $z = x + iy$ and $f(z)$ is analytic (roughly an infinitely differentialbe function) then both u and v satisfy the Laplace equation. Second, the potential of an eletric field follows a non-homogenous Laplace equation, or more precisely referred to as a Poisson equation:

$$u_{xx} + u_{yy} = -\rho.$$

Lastly, the stream function of an incompressible fluid that is also irrotational also follows a Laplace equation.

4.5.4 Laplace equation as BVP in a rectangle

Consider the steady state temperature $u(x, y)$ on a rectangular plate $[0, a] \times [0, b]$ that is kept at 0 at the lower horizontal end $y = 0$ and at temperature profile $f(x)$ at the upper horizontal end $y = b$. The plate is insulated at the two vertical ends $x = 0, x = a$. The insulation condition implies that $Q_t(0, y) = Q_t(a, y) = 0$, where Q is the heat quantity at the point (x, y). Recalling the Fourier's law

$$Q_t = -ku_x,$$

the insulation condition is translated to

$$u_x(0, y) = u_x(a, y) = 0, 0 < y < b.$$

Thus the steady-state temperature is described by the BVP:

$$\begin{aligned} u_{xx} + u_{yy} &= 0 \\ u(x, 0) = 0; u(x, b) &= f(x), 0 < x < a \\ u_x(0, y) = u_x(a, y) &= 0, 0 < y < b. \end{aligned} \tag{4.13}$$

The general solution

Similar to the wave equation we make the ansatz

$$u(x, y) = A(x)B(y).$$

Plugging in to the equation we have

$$\frac{A_{xx}}{A} = -\frac{B_{yy}}{B} = -\lambda,$$

for some constant λ. Thus A, B satisfy the DEs

$$\begin{aligned} A_{xx} + \lambda A &= 0 \\ B_{yy} - \lambda B &= 0. \end{aligned}$$

The solutions for $A(x), B(y)$ depend on the sign of λ and they are

$$\begin{aligned}
A(x) &= C_1 x + C_2, \lambda = 0 \\
A(x) &= C_1 \cos(\alpha x) + C_2 \sin(\alpha x), \lambda = \alpha^2 > 0 \\
A(x) &= C_1 \cosh(\alpha x) + C_2 \sinh(\alpha x), \lambda = -\alpha^2 < 0 \\
B(y) &= C_1' y + C_2', \lambda = 0 \\
B(y) &= C_1' \cosh(\alpha y) + C_2' \sinh(\alpha y), \lambda = \alpha^2 > 0 \\
B(y) &= C_1' \cos(\alpha y) + C_2' \sin(\alpha y), \lambda = -\alpha^2 < 0.
\end{aligned}$$

Note the reverse case for $B(y)$ compared to $A(x)$ for $\lambda > 0$ and $\lambda < 0$. This is because the DE for B has the term $-\lambda B$ instead of $+\lambda A$ in the DE for A.

The boundary condition for x

The boundary condition for x is

$$u_x(0, y) = u_x(a, y) = 0, 0 < y < b.$$

This implies

$$A_x(0)B(y) = A_x(a)B(y) = 0.$$

Since $B(y)$ is not identically 0, we conclude that $A_x(0) = A_x(a) = 0$. That is $A(x)$ solves the regular Sturm-Liouville problem

$$\begin{aligned}
A_{xx} + \lambda A &= 0 \\
A_x(0) = A_x(a) &= 0.
\end{aligned}$$

There are two possible cases for λ that would lead to non-trivial solution here. First $\lambda = 0$ implies $A(x) = C_1 x + C_2$. The boundary condition implies $C_1 = 0$. Thus $A(x) = C_0$ for some constant C_0 is a solution correspond to the $\lambda = 0$ case.

Second, $\lambda > 0$ implies

$$A(x) = C_1 \cos(\alpha x) + C_2 \sin(\alpha x),$$

In this case, $u_x(0, y) = 0$ implies $C_2 = 0$ and $u_x(a, y) = 0$ implies

$$\alpha a = 2n\pi$$

or $\alpha = \frac{2n\pi}{a}$. Thus there is a family of solution

$$A_n(x) = C_n \cos(\frac{2n\pi x}{a}),$$

where $\alpha_n = \frac{2n\pi}{a}, n = 1, 2, \cdots$.

The boundary condition for y

The boundary condition for y is

$$u(x,0)=0; u(x,b)=f(x), 0<x<a.$$

Using a similar argument as the above section, we conclude that $B(y)$ solves the problem

$$\begin{aligned} B_{yy}-\lambda B &= 0 \\ B(0)=0, B(b) &= f(x), \end{aligned}$$

where we have concluded from the above section that $\lambda_n=\alpha_n^2=\left(\frac{2n\pi}{a}\right)^2, n=0,1,\cdots$ are the possible choices for λ.

If $\alpha_n=0$ then $B(y)=C_1'y+C_2'$. $B(0)=0$ implies $C_2'=0$.

If $\alpha_n=\frac{2n\pi}{a}>0$ the general solution for $B(y)$ takes the form

$$B_n(y)=C_1'\cosh(\frac{2n\pi}{a}y)+C_2'\sinh(\frac{2n\pi}{a}y).$$

The condition $B(0)=0$ implies $C_1'=0$.

We now have the general form of $u(x,y)$ as

$$u(x,y)=\sum_{n=0}^{\infty}A_n(x)B_n(y)=C_0y+\sum_{n=1}^{\infty}C_n\sinh(\frac{2n\pi}{a}y)\cos(\frac{2n\pi x}{a}).$$

The condition $u(x,b)=f(x)$ implies

$$C_0b+\sum_{n=1}^{\infty}C_n\sinh(\frac{2n\pi}{a}b)\cos(\frac{2n\pi x}{a}) = f(x).$$

We can rewrite this as

$$C_0b+\sum_{n=1}^{\infty}C_n\sinh(\frac{2n\pi}{a}b)\cos(\frac{2n\pi x}{a})=\sum_{n=0}^{\infty}\widehat{C}_n\cos(\frac{2n\pi x}{a}) = f(x)$$

to make the LHS into a Fourier cosine series, where

$$\begin{aligned} \widehat{C}_0 &= C_0b \\ \widehat{C}_n &= C_n\sinh(\frac{2n\pi}{a}b). \end{aligned}$$

We see that $\widehat{C}_n$ are the Fourier coeficients in the cosine series expansion of $f(x)$. Thus

$$\begin{aligned} \widehat{C}_0 &= \frac{1}{a}\int_0^a f(x)dx \\ \widehat{C}_n &= \frac{2}{a}\int_0^a f(x)\cos(\frac{2n\pi x}{a})dx. \end{aligned}$$

From this it follows that

$$\begin{aligned} C_0 &= \frac{1}{ab}\int_0^a f(x)dx \\ C_n &= \frac{2}{a\sinh(\frac{2n\pi}{a}b)}\int_0^a f(x)\cos(\frac{2n\pi x}{a})dx. \end{aligned}$$

Thus the solution to the BVP (4.13) of the steady-state tempreture on the plate is

$$u(x,y) = \left\{\frac{1}{ab}\int_0^a f(x)dx\right\}y + \sum_{n=1}^{\infty}\left\{\frac{2}{a\sinh(\frac{2n\pi}{a}b)}\int_0^a f(x)\cos(\frac{2n\pi x}{a})dx\right\}\sinh(\frac{2n\pi}{a}y)\cos(\frac{2n\pi x}{a}).$$

Dirichlet problem

The BVP (4.13) we considered has a mixture of boundary condition types. That is it has both Dirichlet type of condition:

$$u(x,0) = 0; u(x,b) = f(x), 0 < x < a$$

and Neumann condition

$$u_x(0,y) = u_x(a,y) = 0, 0 < y < b.$$

This comes from the physical property of the problem we want to solve, that is the plate is kept at certain temperatures at two ends and kept insulated at the other two ends. On the other hand, one can consider problem where all boundary conditions are of Dirichlet type. This problem is, not surprisingly, referred to as the *Dirichlet problem* (for Laplace equation). It will have the following form

$$\begin{aligned} u_{xx} + u_{yy} &= 0 \\ u(x,0) = f(x); u(x,b) &= g(x), 0 < x < a \\ u(0,y) = F(y); u(a,y) &= G(y), 0 < y < b. \end{aligned} \tag{4.14}$$

This problem is not straightforward to solve in this form. Observe that the general solution that we found in (4.5.4) is the same. The difficulty lies in the fitting of the boundary conditions as in (4.5.4) and (4.5.4). There, the zero boundary conditions played a major role in simplifying the process of fitting the boundary conditions.

Nevertheless, we can utilize the superposition principle to help us solve the Dirichlet problem by dividing it to two Dirichlet problems with boundary conditions equal to zero on two sides. More specifically, we can consider the problems

$$\begin{aligned} u_{xx} + u_{yy} &= 0 \\ u(x,0) = f(x); u(x,b) &= g(x), 0 < x < a \\ u(0,y) = 0; u(a,y) &= 0, 0 < y < b. \end{aligned} \tag{4.15}$$

and

$$
\begin{aligned}
u_{xx} + u_{yy} &= 0 \\
u(x,0) = 0; u(x,b) &= 0, 0 < x < a \\
u(0,y) = F(y); u(a,y) &= G(y), 0 < y < b.
\end{aligned} \tag{4.16}
$$

It is straightforward to show that if $u_1(x,y)$ is the solution to (4.15) and $u_2(x,y)$ is the solution to (4.16) then

$$u(x,y) := u_1(x,y) + u_2(x,y)$$

is the solution to (4.14). The problems (4.15) and (4.16) can be solved using a similar techniques to the one we presented above. More spefically, the reader can verify that

$$u_1(x,y) = \sum_{n=1}^{\infty} \left\{ A_n \cosh \frac{n\pi}{a} y + B_n \sinh \frac{n\pi}{a} y \right\} \sin \frac{n\pi}{a} x,$$

where

$$
\begin{aligned}
A_n &= \frac{2}{a} \int_0^a f(x) \sin \frac{n\pi}{a} x dx \\
B_n &= \frac{2}{a \sinh \frac{n\pi}{a} b} \int_0^a g(x) \sin \frac{n\pi}{a} x dx - A_n \cosh \frac{n\pi}{a} b.
\end{aligned}
$$

And

$$u_2(x,y) = \sum_{n=1}^{\infty} \left\{ A_n \cosh \frac{n\pi}{b} x + B_n \sinh \frac{n\pi}{b} x \right\} \sin \frac{n\pi}{b} y,$$

where

$$
\begin{aligned}
A_n &= \frac{2}{b} \int_0^b F(y) \sin \frac{n\pi}{b} y dy \\
B_n &= \frac{2}{b \sinh \frac{n\pi}{b} a} \int_0^b G(y) \sin \frac{n\pi}{b} y dy - A_n \cosh \frac{n\pi}{b} a.
\end{aligned}
$$

4.6 Nonhomogeneous Boundary Value Problem

4.6.1 Non-seperable equations

The heat, wave and Laplace BVPs we considered so far are all separable. That is we can make the ansatz $u(t,x) = A(t)B(x)$ and proceed to solve for A, B. In this section we will consider extension of BVPs such that separability no longer holds and how to solve such problems.

Non-homogenous equations

The term homogenous in mathematics when referred to ODEs and PDEs usually means "the RHS equals to zero and the LHS only involves terms that contains the solution u". For example the heat, wave and Laplace equations we have considered so far are homogenous:

$$\begin{aligned} u_t - ku_{xx} &= 0 \text{ (heat equation)} \\ u_{tt} - \alpha^2 u_{xx} &= 0 \text{ (wave equation)} \\ u_{xx} + u_{yy} &= 0 \text{ (Laplace equation).} \end{aligned}$$

On the other hand, the following heat equation that models heat being generated internally with rate $F(x,t)$ is non-homogenoues:

$$ku_{xx} + F(x,t) = u_t,$$

because if we collect terms that involve u on the LHS it becomes

$$u_t - ku_{xx} = F(x,t).$$

The non-homgeneity makes the equation *no-longer separable* (The ansatz $u(t,x) = A(t)B(x)$ is not appropriate unless it is also the case that $F(x,t) = C(t)D(x)$). Thus we need a different approach for the non-homogeneous case.

Time-dependent boundary conditions

Another extension would be solving the usual homogenous equations we considered so far, subject to *time dependent* boundary conditions instead of constant boundary conditions. For example, consider the problem

$$\begin{aligned} u_t &= u_{xx}, 0 < x < 1, t > 0 \\ u(0,t) &= \cos t, u(1,t) = 0, t > 0 \\ u(x,0) &= 0, 0 < x < 1. \end{aligned}$$

This is similar to the heat BVP we consider except for the boundary condition

$$u(0,t) \quad = \quad \cos t,$$

which depends on t. However, this also makes the problem *non-separable*. To see why, suppose we make the ansatz $u(t) = A(t)B(x)$. Then $u(0,t) = \cos t$ implies that $A(t)B(0) = \cos t$ or $A(t) = \frac{\cos t}{B(0)}$. The equation $u_t = u_{xx}$ becomes

$$\sin tB(x) = \cos tB_{xx}.$$

This implies

$$\frac{B_{xx}}{B(x)} = \tan t$$

which is impossible as B is only dependent on x.

4.6.2 Solving non-homogeneous BVP with homogenous boundary conditions

The abstract procedure

Consider the problem

$$\begin{aligned} u_{xx} + F(x,t) &= u_t, 0 < x < L, t > 0 \\ u(t,0) &= u(t,L) = 0, t > 0 \\ u(0,x) &= f(x), 0 < x < L. \end{aligned}$$

The technique to solve this problem is to assume that both $v(x,t)$ and $F(x,t)$ can be expanded using Fourier sine series with time dependent Fourier coefficients:

$$\begin{aligned} u(t,x) &= \sum_{n=1}^{\infty} u_n(t) \sin(\frac{n\pi}{L}x) \\ F(t,x) &= \sum_{n=1}^{\infty} F_n(t) \sin(\frac{n\pi}{L}x). \end{aligned} \quad (4.17)$$

Remarks:

a. This assumption essentially says that our problem is still separable. Indeed the crucial assumption is on the form of $F(x,t)$. It is conceivable that we can come up with an example of $F(x,t)$ that does not admit the representation as in (4.17). However, for a specific $F(x,t)$ we can easily check to see if the assumption is satisfied and proceed. An easy example where the assumption holds is when F is only dependent on x.

b. The representation $u(t,x) = \sum_{n=1}^{\infty} u_n(t) \sin(\frac{n\pi}{L}$ apriori satisfies the boundary conditions $u(t,0) = u(t,L) = 0$. In order to solve for t we only need to find the coefficients $u_n(t)$. Also observe that $F_n(t)$ is known once we know the specific form of $F(t,x)$.

Plugging in the form as in (4.17) gives

$$\sum_{n=1}^{\infty} \left\{ F_n(t) - \frac{n^2\pi^2}{L^2} u_n(t) \right\} \sin(\frac{n\pi}{L}x) = \sum_{n=1}^{\infty} u_n'(t) \sin(\frac{n\pi}{L}x).$$

Equating the coefficients of $\sin(\frac{n\pi}{L}x)$ gives an ODE in t:

$$u_n'(t) + \frac{n^2\pi^2}{L^2} u_n(t) = F_n(t).$$

This is a linear first order ODE whose general solution can be found explicitly. Each general solution for u_n depends on a constant C_n. Finally these $C_n's$ can be found by the initial condition $u(t,0) = f(x)$.

An example

Consider the BVP

$$\begin{aligned} u_{xx} + (1-x)\sin t &= u_t, 0 < x < 1, t > 0 \\ u(0,t) = 0, u(1,t) &= 0, t > 0 \\ u(x,0) &= x-1, 0 < x < 1. \end{aligned}$$

We first consider the representation

$$(1-x)\sin t = \sum_{n=1}^{\infty} F_n(t)\sin(n\pi x).$$

That is

$$\sum_{n=1}^{\infty} \frac{F_n(t)}{\sin t}\sin(n\pi x) = 1-x.$$

Clearly $\frac{F_n(t)}{\sin t}$ is the Fourier coefficients of the sine series expansion of $1-x$. Thus we have

$$F_n(t) = 2\sin t\int_0^1 (1-x)\sin(n\pi x)dx = \frac{2}{n\pi}\sin t.$$

The ODE in t becomes

$$u_n'(t) + n^2\pi^2 u_n(t) = \frac{2}{n\pi}\sin t.$$

It has a general solution

$$u_n(t) = \frac{2}{n\pi}e^{-n^2\pi^2 t}\int_0^t e^{n^2\pi^2 s}\sin s ds + u_n(0)e^{-n^2\pi^2 t}.$$

To determine $u_n(0)$ we use the initial condition

$$u(x,0) = x-1, 0 < x < 1.$$

That is

$$\sum_{n=0}^{\infty} u_n(0)\sin(n\pi x) = x-1.$$

From the Fourier sine series of $x-1$ we have

$$u_n(0) = 2\int_0^1 (x-1)\sin(n\pi x)dx.$$

4.6.3 Solving homogeneous BVP with time dependent boundary conditions

The abstract procedure

Consider the problem

$$\begin{aligned} u_{xx} &= u_t, 0 < x < L, t > 0 \\ u(t,0) &= u_0(t)\ ,\ u(t,L) = u_1(t), t > 0 \\ u(0,x) &= f(x), 0 < x < L. \end{aligned}$$

We seek to transform this problem into a problem with time independent boundary conditions by making the substitution

$$u(t,x) = v(t,x) + \phi(t,x),$$

where

$$\phi(t,x) = u_0(t) + \frac{x}{L}(u_1(t) - u_0(t)).$$

The function $\phi(t,x)$ enjoys the following nice properties:

$$\begin{aligned} \phi(t,0) &= u_0(t)\ ,\ \phi(t,L) = u_1(t) \\ \phi_{xx} &= 0. \end{aligned}$$

Pluggging in the form $u(t,x) = v(t,x) + \phi(t,x)$ into the original BVP, we see that $v(t,x)$ satisfies the BVP

$$\begin{aligned} v_{xx} + F(t,x) &= v_t, 0 < x < L, t > 0 \\ v(t,0) &= v(t,L) = 0, t > 0 \\ v(0,x) &= g(x), 0 < x < L, \end{aligned}$$

where $F(t,x) = -\phi_t(t,x)$ and $g(x) = f(x) - \phi(0,x)$. This is a non-homogenous BVP with homogenous boundary conditions. We can follow the procedure outlined in the previous section to solve it.

An example

Consider the problem

$$\begin{aligned} u_t &= u_{xx}, 0 < x < 1, t > 0 \\ u(0,t) &= \cos t, u(1,t) = 0, t > 0 \\ u(x,0) &= 0, 0 < x < 1. \end{aligned}$$

We make the substitution

$$u(t,x) = v(t,x) + (1-x)\cos(t).$$

Then $v(t,x)$ solves the BVP

$$\begin{aligned} v_{xx} + (1-x)\sin t &= v_t, 0 < x < 1, t > 0 \\ v(0,t) = 0, v(1,t) &= 0, t > 0 \\ v(x,0) &= x - 1, 0 < x < 1. \end{aligned}$$

This is exactly the problem we solved in the section (4.6.2).

4.6.4 Solving non-homogeneous BVP with time dependent boundary conditions

Lastly, consider the problem

$$\begin{aligned} u_{xx} + F(x,t) &= u_t, 0 < x < L, t > 0 \\ u(t,0) &= u_0(t)\ ,\ u(t,L) = u_1(t), t > 0 \\ u(0,x) &= f(x), 0 < x < L. \end{aligned}$$

We can use the same substitution in the previous section to transform this problem into a problem with time independent boundary conditions:

$$u(t,x) = v(t,x) + \phi(t,x),$$

where

$$\phi(t,x) = u_0(t) + \frac{x}{L}(u_1(t) - u_0(t)).$$

Pluggging in the form $u(t,x) = v(t,x) + \phi(t,x)$ into the original BVP, we see that $v(t,x)$ satisfies the BVP

$$\begin{aligned} v_{xx} + G(t,x) &= v_t, 0 < x < L, t > 0 \\ v(t,0) &= v(t,L) = 0, t > 0 \\ v(0,x) &= g(x), 0 < x < L, \end{aligned}$$

where $F(t,x) = G(t,x) - \phi_t(t,x)$ and $g(x) = f(x) - \phi(0,x)$. This is a non-homogenous BVP with homogenous boundary conditions and we can follow the procedure outlined in the section (4.6.2) to solve it.

4.7 Orthogonal Series Expansions for BVP

4.7.1 Introduction

In this section, we consider several examples of BVPs with nonstandard boundary conditions. The technique to sovle these problems is orthogonal series expansion, the details of which we present below.

4.7.2 A heat equation example

Consider the BVP

$$u_t = ku_{xx}$$

subject to

$$\begin{aligned} u(t,0) &= 0, u_x(t,1) = -hu(x,1), h > 0, t > 0 \\ u(0,x) &= 1, 0 < x < 1. \end{aligned}$$

Remark:

a. This problem models a rod of unit length, whose initial tempreature is 1 everywhere. At the right end there is heat transfer into a surrounding medium, whose temperature is kept at a constant zero. (Hence the rod's left end is at a constant temperature 0). To see this, recall Fourier's law for heat transfer:

$$Q_t = -Ku_x$$

and Newton's law of cooling, which says the rate of change of heat is proportional to the temperature difference. Thus the condition $u_x(t,1) = -hu(x,1)$ can be seen as $Q_t = h(u(x,1)-0)$, which captures the fact that the heat transfer is proportional to the difference of temperature of the right end of the rod and the environment.

b. In terms of boundary condition, the condition $u_x(t,1) = -hu(x,1)$ is a new one and thus the techniques we developed before to solve BVP will need to be modified to solve it.

Solution:

We will still use separation of variables, that is we make the ansatz

$$u(t,x) = A(t)B(x).$$

This is because this ansatz is still consistent with the boundary condition $u_x(t,1) = -hu(x,1)$. Proceed as before, we have

$$\frac{A_t}{kA} = \frac{B_{xx}}{B} = -\lambda,$$

for some constant λ. Thus A, B satisfy respectively the DEs

$$\begin{aligned} A_t + k\lambda A &= 0 \\ B_{xx} + \lambda B &= 0. \end{aligned}$$

The solution for $A(t)$ is

$$A(t) = Ce^{-k\lambda t},$$

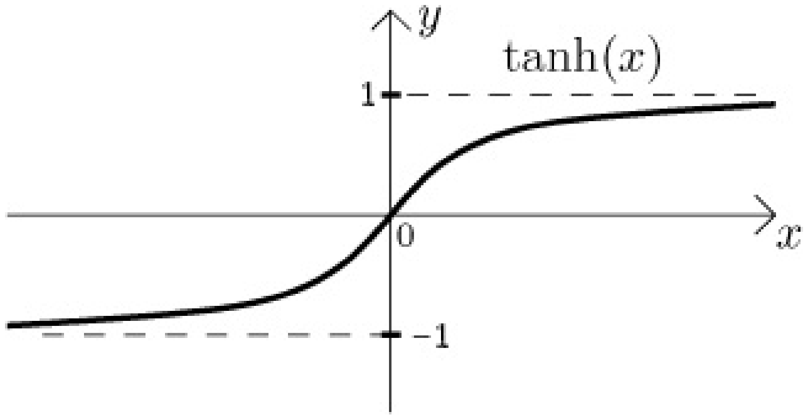
for some constant C to be determined. The solution for $B(x)$ depends on the sign of λ and it is

$$\begin{aligned} B(x) &= C_1 x + C_2, \lambda = 0 \\ B(x) &= C_1 \cos(\alpha x) + C_2 \sin(\alpha x), \lambda = \alpha^2 > 0 \\ B(x) &= C_1 \cosh(\alpha x) + C_2 \sinh(\alpha x), \lambda = -\alpha^2 < 0. \end{aligned}$$

We now consider the boundary conditions

$$u(t, 0) = 0, u_x(t, 1) = -hu(x, 1).$$

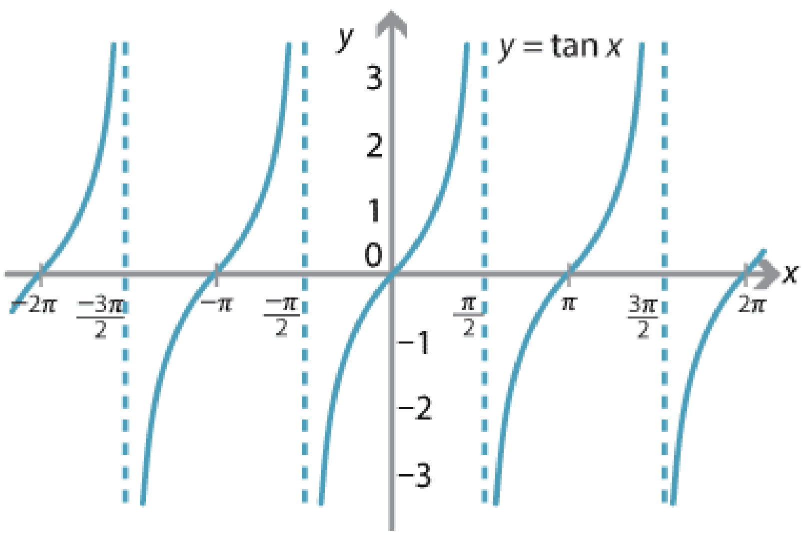
A quick check shows that the choice $B(x) = C_1 x + C_2$ is not possible unless we have $C_1 = C_2 = 0$. The choice $B(x) = C_1 \cosh(\alpha x) + C_2 \sinh(\alpha x)$ also is not possible. First the condition $u(t, 0) = 0$ implies $C_1 = 0$. Second the condition $u_x(t, 1) = -hu(x, 1)$ implies $\tanh(\alpha) = -\frac{\alpha}{h}$. The graph of $\tanh(x)$ intersects the graph of $-\frac{x}{h}$ only at $x = 0$ (note that $h > 0$ so $-\frac{1}{h} < 0$). The choise $\alpha = 0$ is not possible here since this form of $B(x)$ implies $\lambda = \alpha^2 > 0$.



This leaves us with the choice of $B(x) = C_1 \cos(\alpha x) + C_2 \sin(\alpha x)$. The left boundary condition implies that $C_1 = 0$. The right boundary condition implies

$$\tan(\alpha) = -\frac{\alpha}{h}.$$

The graph of $\tan(x)$ intersects the graph of $-\frac{x}{h}$ at infinitely many points $\alpha_n, n = 1, 2, \cdots$. We do not have an explicit representation α_n. Nevetheless, we can still write the general



solution to the BVP as

$$u(t, x) = \sum_{n=1}^{\infty} C_n e^{-k\alpha_n^2 t} \sin(\alpha_n x).$$

Pluggin in the initial condition we have

$$u(0, x) = \sum_{n=1}^{\infty} C_n \sin(\alpha_n x) = 1.$$

The α_n are not of the form $\frac{n\pi}{L}$ so this is not a Fourier expansion of 1. Rather, we recall the following result about eigenvalues and eigenfunctions of a regular Sturm-Liouville problem. Consider the problem

$$\frac{d}{dx}(r(x)y') + (q(x) + \lambda p(x))y = 0, a < x < b.$$

Subject to

$$\begin{aligned} A_1 y(a) + B_1 y'(a) &= 0 \\ A_2 y(b) + B_2 y'(b) &= 0. \end{aligned}$$

Then the eigenfunctions corresponding to distinct eigenvalues are orthogonal with respect to the weight function $p(x)$ on $[a, b]$. In our problem, $\sin(\alpha_n x)$ are eigenfunctions corresponding to eigenvalues α_n for the Sturm-Liouville problem

$$\begin{aligned} B_{xx} + \lambda B &= 0, 0 < x < 1 \\ B(0) &= 0, B_x(1) + hB(1) = 0. \end{aligned}$$

Thus $\sin(\alpha_n x)$ are orthogonal with weight functions $p(X) = 1$. Applying this result gives

$$C_n = \frac{\int_0^1 \sin(\alpha_n x)dx}{\int_0^1 \sin^2(\alpha_n x)dx}.$$

4.7.3 A wave equation example

Consider the BVP

$$u_{tt} = a^2 u_{xx}, 0 < x < 1, t > 0$$

subject to

$$\begin{aligned} u(t, 0) &= 0, u_x(t, 1) = 0, , t > 0 \\ u(0, x) &= x, u_t(0, x) = 0, 0 < x < 1. \end{aligned}$$

Remarks:

a. This problem models the twist angle $u(t, x)$ of a torsionally vibrating shaft of unit length.

b. In terms of boundary condition, we get a Dirichlet type on the left end and a Neumann type on the right end. We will still use separation of variable to solve the problem. Similar to what we have above, the mixed boundary condition has the effect of making the eigenvalues of the regualr Sturm-Liouville problem

$$\begin{aligned} B_{xx} + \lambda B &= 0, 0 < x < 1 \\ B(0) &= 0, B_x(1) = 0 \end{aligned}$$

to have eigenvalues not of “Fourier type", that is they are not of the form $\frac{n\pi}{L}$. Nevetheless, the eigenfunctions corresponding to the eigenvalues are still orthogonal so when it comes to solving for the undetermined coefficients using the initial conditions, we can still use orthogonal series expansion techniques.

c. There are two initial conditions on $u(0,x)$ and $u_t(0,x)$ because the PDE is second order in t.

Solution:

Using separation of variables

$$u(t,x) = A(t)B(x),$$

we have

$$\frac{A_t}{a^2 A} = \frac{B_{xx}}{B} = -\lambda,$$

for some constant λ. Thus A, B satisfy respectively the DEs

$$\begin{aligned} A_{tt} + \alpha^2 \lambda A &= 0 \\ B_{xx} + \lambda B &= 0. \end{aligned}$$

We first discuss the solution for $B(x)$. As mentioned above, $B(x)$ satisfies the regular Sturm-Liouville problem

$$\begin{aligned} B_{xx} + \lambda B &= 0, 0 < x < 1 \\ B(0) &= 0, B_x(1) = 0. \end{aligned}$$

We would not verify the details that if $\lambda \leq 0$, the only solution to this problem is $B(x) = 0$. Thus we conclude $\lambda > 0$ and

$$B_n(x) = C_n \sin(\alpha_n x),$$

where α_n satisfies

$$\cos(\alpha_n) = 0.$$

Here we do have an explit formula for α_n :

$$\alpha_n = \frac{(2n-1)\pi}{2}, n = 1, 2, \cdots$$

Going back to $A(t)$, $A(t)$ solves the DE

$$A_{tt} + a^2 \lambda_n A = 0, n = 1, 2, \cdots$$

Because $a^2 \lambda_n > 0$, we know $A(t)$ has the form

$$A_n(t) = C_n^1 \sin(a\alpha_n t) + C_n^2 \cos(a\alpha_n t).$$

The initial condition $u_t(0,x) = 0$ translates to $A_n'(t) = 0$, which implies $C_n^1 = 0$. Thus

$$\begin{aligned} u(t,x) &= \sum_{n=1}^{\infty} A_n(t) B_n(x) \\ &= \sum_{n=1}^{\infty} C_n \cos(a\alpha_n t)\sin(\alpha_n x). \end{aligned}$$

The initial condition $u(0,x) = x$ implies

$$u(0,x) = \sum_{n=1}^{\infty} C_n \sin(\alpha_n x) = x.$$

As we mentioned, the collection $\{\sin(\alpha_n x), n = 1, 2, \cdots\}$ is orthogonal. Thus

$$C_n = \frac{\int_0^1 x\sin(\alpha_n x)dx}{\int_0^1 \sin^2(\alpha_n x)dx}.$$

4.8 Fourier Series in Two Variables

4.8.1 Introduction

In this section, we explore the techniques of solving two dimensional heat and wave equations. They are

$$\begin{aligned} u_t &= k(u_{xx} + u_{yy}) \text{ heat equation} \\ u_{tt} &= \alpha^2(u_{xx} + u_{yy}) \text{ heat equation.} \end{aligned}$$

At first glance, separation of variables is still applicable. That is it is still reasonable to assume

$$u(t,x,y) = A(t)B(x)C(y),$$

(with appropriate boundary conditions, of course). The difference is when we obtain the general form for $B(x), C(y)$ in terms of a series and use the initial condition to solve for the coefficients, we will have to deal with a series in two variables:

$$\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} A_{mn}\sin(\frac{m\pi}{a}x)\sin(\frac{n\pi}{b}y) = f(x,y), 0 \le x \le a, 0 \le y \le b.$$

We will see how to solve for the coefficients A_{mn} in the following section.

4.8.2 Two dimensional Fourier sine and cosine series

We present the following results without proof. The intuition behind the representation is similar to the one dimensional sine and cosine series. The technical difficulty lies in the convergence of the series to the true function, which we skip.

Let $f(x, y)$ be a sufficiently nice function defined on $0 \le x \le a, 0 \le y \le b$. Then the two dimensional Fourier sine series representation of f is

$$f(x,y) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} A_{mn} \sin(\frac{m\pi}{a}x)\sin(\frac{n\pi}{b}y),$$

where

$$A_{mn} = \frac{4}{ab}\int_0^a\int_0^b f(x,y)\sin(\frac{m\pi}{a}x)\sin(\frac{n\pi}{b}y)dxdy.$$

On the other hand, the two dimensional Fourier cosine series representation of f is

$$\begin{aligned} f(x,y) &= A_{00} + \sum_{m=1}^{\infty} A_{m0}\cos(\frac{m\pi}{a}x) + \sum_{n=1}^{\infty} A_{0n}\cos(\frac{n\pi}{b}y) \\ &+ \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} A_{mn}\cos(\frac{m\pi}{a}x)\cos(\frac{n\pi}{b}y), \end{aligned}$$

where

$$\begin{aligned} A_{00} &= \frac{1}{ab}\int_0^a\int_0^b f(x,y)dxdy \\ A_{m0} &= \frac{2}{ab}\int_0^a\int_0^b f(x,y)\cos(\frac{m\pi}{a}x)dxdy \\ A_{0n} &= \frac{2}{ab}\int_0^a\int_0^b f(x,y)\cos(\frac{n\pi}{b}y)dxdy \\ A_{mn} &= \frac{4}{ab}\int_0^a\int_0^b f(x,y)\cos(\frac{m\pi}{a}x)\cos(\frac{n\pi}{b}y)dxdy. \end{aligned}$$

4.8.3 Heat equation for temperatures in a plate

Consider the BVP

$$\begin{aligned} u_t &= k(u_{xx} + u_{yy}), t > 0, 0 < x < a, 0 < y < b \\ u(t,0,y) &= u(t,a,y) = u(t,x,0) = u(t,x,b) = 0 \\ u(0,x,y) &= f(x,y). \end{aligned}$$

$u(t, x, y)$ describes the temperature on a rectangular plate whose initial temperature profile is $f(x, y)$ and whose temperature at all four sides are kept at 0.

Using separation of variables, we make the ansatz

$$u(t,x,y) = A(t)B(x)C(y).$$

Plugging this form in the PDE we have

$$A_t BC = kA(CB_{xx} + BC_{yy}).$$

That is

$$\frac{A_t}{kA} = \frac{B_{xx}}{B} + \frac{C_{yy}}{C}.$$

The only way for an equation of the form

$$f(t) = g(x) + h(y)$$

to hold is when $f(t) = g(x) + h(y) = -\lambda$. Furthermore, it also must be that $g(x) = -\mu$ and $h(y) = -\lambda + \mu$ for some constants λ, μ. However, from our experience with one dimensional heat equation, we note that $B(x)$ and $C(y)$ will solve a regular Sturm-Liouville problem. Thus, to make it convenient for us, we will actually write

$$\begin{aligned}
\frac{B_{xx}}{B} &= -\lambda, \\
\frac{C_{yy}}{C} &= -\mu, \\
\frac{A_t}{kA} &= -(\lambda + \mu).
\end{aligned}$$

Thus B, C solve the regular Sturm-Liouville problems

$$\begin{aligned}
B_{xx} + \lambda B &= 0, B(0) = B(a) = 0 \\
C_{yy} + \mu C &= 0, C(0) = C(b) = 0
\end{aligned}$$

and A solves the DE

$$A_t + k(\lambda + \mu)A = 0.$$

Using the same argumetn as the one dimensional heat equation case, the general solutions for A, B, C are

$$\begin{aligned}
A(t) &= C^0 e^{-k(\frac{n^2\pi^2}{a^2} + \frac{n^2\pi^2}{b^2})t}, \\
B_n(x) &= C_n^1 \sin(\frac{n\pi x}{a}), \\
C_n(y) &= C_n^2 \sin(\frac{n\pi y}{b}).
\end{aligned}$$

The general solution for the BVP is

$$\begin{aligned} u(t,x,y) &= \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} A_{mn}A(t)B_n(x)C_n(y) \\ &= \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} A_{mn}e^{-k(\frac{n^2\pi^2}{a^2}+\frac{n^2\pi^2}{b^2})t}\sin(\frac{n\pi x}{a})\sin(\frac{n\pi y}{b}). \end{aligned}$$

We use the initial condition to solve for C_n :

$$u(0,x,y) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} A_{mn}\sin(\frac{n\pi x}{a})\sin(\frac{n\pi y}{b}) = f(x,y).$$

Using the two dimensional sine series we have

$$A_{mn} = \frac{4}{ab}\int_0^a\int_0^b f(x,y)\sin(\frac{m\pi}{a}x)\sin(\frac{n\pi}{b}y)dxdy.$$

References

[1] Stanford Engineering Everywhere, *EE261 - The Fourier Transform and its Applications* https://see.stanford.edu/Course/EE261.

[2] Zill, Dennis, Warren S. Wright, and Michael R. Cullen. *Advanced engineering mathematics.* Jones & Bartlett Learning, 2011.