Mathematical methods in physics
Part I B

Lectures given by Frank Krauss,
michaelmas terms 2006 and 2007
at University of Durham

Disclaimer:

This is the second half of part I of the course "Mathematical methods in physics". The first half has been given by Prof. V. Khoze. This document does not intend to replace the lecture, it just contains handouts summarizing the lectures. In the preparation of this lecture, a number of excellent books has been used, among them:

- G. G. Stephenson: "Mathematical methods for science students"
- M. L. Boas: "Mathematical methods in phsiycal sciences"
- M. R. Spiegel: "Fourier analysis"

Contents

Chapter 1

Matrices, Determinants, Eigenvalues and Eigenvectors and all that

1.1 Matrices and their properties

1.1.1 Matrices

Systems of Linear Equations

At many occasions in physics and mathematics, systems of linear equations emerge. These are, in general, m equations for n unknowns, i.e.

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m\,. \end{array} \tag{1.1}$$

Introducing the **matrix** $\hat{A}$ with m rows, one for each equation, and n columns, one for each unknown,

$$\hat{A} = \left(\begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn}\,, \end{array} \right) \tag{1.2}$$

and the n-component vectors $\vec{x}$ and $\vec{b}$, allows to represent equation systems of this type like

$$\hat{A}\vec{x} = \vec{b}\,. \tag{1.3}$$

Alternatively, the same set of equations can also be written as

$$\sum_{k=1}^{n} a_{lk}x_k = \sum_{k=1}^{n} x_k a_{lk} = b_l\,, \tag{1.4}$$

where $l \in [1,\, m]$. Using **Einsteins convention of summing over repeated indices** this can be written as

$$a_{lk}x_k = x_k a_{lk} = b_l\,. \tag{1.5}$$

In the case considered here, a matrix with n columns and m rows, the matrix is said to be of type $(m,\, n)$ or just a $(m,\, n)$-matrix. The entries a_{ij} are called the **elements** of the matrix. In the notation here, the first index of the elements (i) denotes the row, whereas the second one (j) denotes the column. Clearly, if there are more unknowns than equations, $n > m$ the system is **under-determined**. If, in contrast, $m > n$, the system is **over-determined** and it is possible that no solution can be constructed. In general, in this case the presence of viable solutions indicates that pairs of rows are multiples of each other or that one row can be re-expressed through a sum of other rows. In other words, in such a case, the rows are not **linearly independent**. If $m = n$, the matrix is called **quadratic** of order n, and in the general case of linearly independent rows and columns, a solution can uniquely be found. If all entries a_{ij} with $i \neq j$ are zero, and entries with $i = j$ are non-zero, the matrix is called **diagonal**, and a solution is simple. If such a diagonal matrix has all diagonal entries equal to one, i.e. $a_{ij} = \delta_{ij}$, it is called the **unit matrix**, often denoted by $\mathbf{1}$.
A standard method to solve such a system is called **Gauss elimination**. In this method, basically the first row/equation is employed to eliminate the first unknown, the (potentially altered) second row/equation is used for the second unknown and so on. Thereby, the allowed operations here are:

- Interchanging two rows;
- Multiplying/dividing a row with/by a constant;

- Adding and/or subtracting multiples of two rows.

Example:

- To see how this works, consider the three equations

$$\begin{array}{ccccccc} 2x & & & - & z & = & 2 \\ 6x & + & 5y & + & 3z & = & 7 \\ 2x & - & y & & & = & 4\,. \end{array} \tag{1.6}$$

To alleviate things, in a first step the first two rows are interchanged, leaving

$$\begin{array}{ccccccc} 2x & & & - & z & = & 2 \\ 2x & - & y & & & = & 4 \\ 6x & + & 5y & + & 3z & = & 7\,. \end{array} \tag{1.7}$$

Eliminating the first unknown with the first equation implies subtracting the first equation from the second, and subtracting three times the first equation from the third, yielding

$$\begin{array}{ccccc} -y & + & z & = & 2 \\ 5y & + & 6z & = & 1\,. \end{array} \tag{1.8}$$

The second unknown is eliminated by adding five times the first row of the emerging system to the second one, such that

$$11z = 11 \quad \text{and} \quad z = 1\,. \tag{1.9}$$

Reinserting then gives $y = -1$ and $x = 3/2$.

Properties of matrices

Matrices, such as $\hat{A}$ in the previous section, allow a number of operations, which will be briefly listed here:

- **Transposition**:
  A matrix can be transposed by interchanging rows and columns, i.e.
  $$\hat{A}^T = \{a_{ij}\}^T = \{a_{ji}\} \ . \tag{1.10}$$
  If $\hat{A}^T = \hat{A}$ the matrix is called **symmetric**, if $\hat{A}^T = -\hat{A}$ the matrix is **antisymmetric**.

- **Adjoint Matrix**:
  Given a matrix $\hat{A}$ with complex elements, its adjoint matrix $\hat{A}^\dagger$ is given by transposing it and complex conjugation of each element:
  $$\hat{A}^\dagger = \{a_{ij}\}^\dagger = \left\{a_{ji}^*\right\} \ . \tag{1.11}$$
  Matrices which satisfy
  $$\hat{A}^\dagger = \hat{A} \tag{1.12}$$
  are called **Hermitian** or **self-adjoint**. Such matrices play a central role in Quantum Mechanics, because their eigenvalues (see below) are guaranteed to be real. Also, their diagonal elements are real numbers.

- **Addition**:
  Two matrices are added to yield another matrix,
  $$\hat{A} + \hat{B} = \hat{C} \,, \tag{1.13}$$
  where the elements of $\hat{C}$ are given by
  $$c_{ij} = a_{ij} + b_{ij} \,. \tag{1.14}$$
  This obviously assumes that all three matrices are of the same type, i.e. have the same number of rows and columns.

- **Multiplication with a number**:
  A matrix can be multiplied by a number by multiplying each element with this number:
  $$c\hat{A} = \hat{A}c = \hat{C} = \{c_{ij}\} = \{ca_{ij}\} \ . \tag{1.15}$$

- **Multiplication**:
  Two matrices are multiplied to yield another matrix,

$$\hat{A}\hat{B} = \hat{C}\,, \tag{1.16}$$

  where the elements of $\hat{C}$ are given by

$$c_{ij} = a_{ik}b_{kj} = b_{kj}a_{ik}\,, \tag{1.17}$$

  i.e. to obtain the elements in the ith row and the jth column of $\hat{C}$, the elements in the ith row of $\hat{A}$ are multiplied one-by-one with the elements in the jth column of $\hat{B}$, and the products are summed over. This automatically implies that if $\hat{A}$ and $\hat{B}$ are of type $(n,\, m)$ and $(k,\, l)$, respectively, then $m = k$ and $\hat{C}$ is a $(n,\, l)$-matrix.

  It is quite clear from the expression above that in general

$$\hat{A}\hat{B} = a_{ik}b_{kj} = b_{kj}a_{ik} \neq b_{ik}a_{kj} = a_{kj}b_{ik} = \hat{B}\hat{A}\,. \tag{1.18}$$

- **Rank**:
  The rank of a matrix is given by a non-negative integer number, which is given in the following way: Consider a $(n,\, m)$-matrix and take the columns as vectors of dimension m. The rank of the matrix is then given by the maximal number of linearly independent vectors in this set. The rank of a matrix does not change under:

  - permutation of the columns,
  - multiplication of a column with any number $c \neq 0$,
  - addition of an arbitrary multiple of one row to any other row,
  - transposition of the matrix,
  - and, hence, permutation of the rows.

In addition, there are some properties which are defined for quadratic matrices only:

- **Trace**:
  The trace of a matrix is defined as the sum of its diagonal elements:

$$\mathrm{Tr}(\hat{A}) = a_{ii}\,. \tag{1.19}$$

- **Inverse**:
  The inverse of a matrix $\hat{A}$, $\hat{A}^{-1}$, is defined such that
  $$\hat{A}\hat{A}^{-1} = \hat{A}^{-1}\hat{A} = \mathbf{1} \tag{1.20}$$
  or, in components,
  $$a_{ij}a^{-1}_{jk} = \delta_{ik}\,. \tag{1.21}$$

- **Determinant**:
  See next section.

More complicated properties

Having defined the inverse and transposed (Hermitian conjugate) of a matrix it is worth to state a few more important properties:

$$(\hat{A}\hat{B})^{-1} = \hat{B}^{-1}\hat{A}^{-1} \tag{1.22}$$

and

$$\begin{aligned}(\hat{A}\hat{B})^T &= \hat{B}^T\hat{A}^T \\ (\hat{A}\hat{B})^\dagger &= \hat{B}^\dagger\hat{A}^\dagger\,.\end{aligned} \tag{1.23}$$

In order to see this, let us start with the transposed:

$$(\hat{A}\hat{B})^T = (a_{ij}b_{jk})_{ki} = a_{ji}b_{kj} = b^T_{jk}a^T_{ji} = \hat{B}^T\hat{A}^T\,. \tag{1.24}$$

For the inverse consider the product

$$1 = (\hat{A}\hat{B})^{-1}(\hat{A}\hat{B}) = \hat{B}^{-1}\hat{A}^{-1}\hat{A}\hat{B} = \hat{B}^{-1}\hat{B} = 1\,. \tag{1.25}$$

1.1.2 Determinants

Definition

One of the most important properties of a quadratic matrix is its **determinant**, a well-defined, unique real or complex number. It is denoted by

$$\det(\hat{A}) = |\hat{A}| = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \sum_\Pi (-1)^{j(\Pi)}\, a_{1i_1}a_{2i_2}\dots a_{ni_n}\,. \tag{1.26}$$

Here, the sum stretches over all $n!$ possible permutations Π of the numbers 1 to n, and $j(\Pi)$ is the sign of the respective permutation. In practise this implies that each product consists of n elements in such a way that it contains exactly one elements per row and column.
To underline this, just think of any matrix as being related to a "checker-board" of the same size filled with $+$ or $-$ signs,

$$\begin{pmatrix} + & - & + & \dots \\ - & + & - & \dots \\ + & - & + & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} . \tag{1.27}$$

For calculating the determinant of a $n \times n$-matrix, **all** possible combinations must be summed of products with n matrix element, such that each product contains exactly one matrix element per column and row. The overall sign of such a product is then determined by the sign of the corresponding checkerboard entries, which are all identical.

Example:

- $$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc\,. \tag{1.28}$$

Properties

Before discussing how a determinant can be calculated using the method of **Laplace**, the properties of determinants will briefly be listed:

- Permutation of rows or columns does not change the **absolute value** of the determinant but only its sign;
- multiplying a row by a number results in multiplying the determinant with the same number;
- adding a multiple of one row to another row does not change the determinant;

- the determinant is zero if the rows are not linearly independent;
- the determinant is invariant under transposition.

There is yet another important property namely the behavior of the determinant under multiplication:

$$\det(\hat{A}\hat{B}) = \det(\hat{A})\det(\hat{B})\,. \tag{1.29}$$

This implies, in particular,

$$\det(\hat{A}\hat{A}^{-1}) = \det(\mathbf{1}) = \det(\hat{A})\det(\hat{A}^{-1}) \tag{1.30}$$

and thus

$$\det(\hat{A}^{-1}) = \frac{1}{\det(\hat{A})}\,. \tag{1.31}$$

Calculating determinants

In principle, determinants of 2×2- or 3×3-matrices can be calculated straightforwardly:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \tag{1.32}$$

and

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}\begin{vmatrix} a_{23} & a_{21} \\ a_{33} & a_{31} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}[a_{22}a_{33} - a_{32}a_{23}] + a_{12}[a_{23}a_{31} - a_{21}a_{33}] + a_{13}[a_{21}a_{32} - a_{31}a_{22}]\,. \end{aligned} \tag{1.33}$$

In the calculation of the latter determinant, **Laplace's** method already has been indicated. The idea here is to use either a row or column (in the case shown here, it was the first row) of an $n \times n$-matrix, and to multiply each of these elements with the determinant of the $(n-1)\times(n-1)$-submatrix which forms when ignoring the row and column of this element in the original matrix. The determinant of this submatrix is sometimes denoted as the **adjoint of the element** a_{ij}, denoted by A_{ij}. This method is called **development**

of the determinant according to the m**th row or column**, formally speaking

$$\det\hat{A} = \sum_j a_{ij} A_{ij} = \sum_i a_{ij} A_{ij} \,, \tag{1.34}$$

where the index i (j) is fixed to the row (column) in question.

Example:

- $$\begin{aligned} \begin{vmatrix} 1 & -5 & 2 \\ 7 & 3 & 4 \\ 2 & 1 & 5 \end{vmatrix} &= 2\begin{vmatrix} 7 & 3 \\ 2 & 1 \end{vmatrix} - 4\begin{vmatrix} 1 & -5 \\ 2 & 1 \end{vmatrix} + 5\begin{vmatrix} 1 & -5 \\ 7 & 3 \end{vmatrix} \\ &= 2\cdot 1 - 4\cdot 11 + 5\cdot 38 = 148\,. \end{aligned} \tag{1.35}$$

- $$\begin{aligned} &\begin{vmatrix} 2 & 9 & 9 & 4 \\ 2 & -3 & 12 & 8 \\ 4 & 8 & 3 & -5 \\ 1 & 2 & 6 & 4 \end{vmatrix} \overset{c_2\to c_2-c_3+c_4}{=} \begin{vmatrix} 2 & 4 & 9 & 4 \\ 2 & -7 & 12 & 8 \\ 4 & 0 & 3 & -5 \\ 1 & 0 & 6 & 4 \end{vmatrix} \\ \overset{c_3\to c_3/3}{=} &\quad 3\begin{vmatrix} 2 & 4 & 3 & 4 \\ 2 & -7 & 4 & 8 \\ 4 & 0 & 1 & -5 \\ 1 & 0 & 2 & 4 \end{vmatrix} \overset{c_1\leftrightarrow c_2}{=} -3\begin{vmatrix} 4 & 2 & 3 & 4 \\ -7 & 2 & 4 & 8 \\ 0 & 4 & 1 & -5 \\ 0 & 1 & 2 & 4 \end{vmatrix} \\ = &\quad -3\left\{4\begin{vmatrix} 2 & 4 & 8 \\ 4 & 1 & -5 \\ 1 & 2 & 4 \end{vmatrix} - (-7)\begin{vmatrix} 2 & 3 & 4 \\ 4 & 1 & -5 \\ 1 & 2 & 4 \end{vmatrix} - 0 + 0\right\} \\ = &\quad -3\left\{4\cdot 0 + 7\begin{vmatrix} 2 & 3 & 4 \\ 4 & 1 & -5 \\ 1 & 2 & 4 \end{vmatrix}\right\} \overset{r_1\to r_1-r_3}{=} -21\begin{vmatrix} 1 & 1 & 0 \\ 4 & 1 & -5 \\ 1 & 2 & 4 \end{vmatrix} \\ = &\quad -21\begin{vmatrix} 1 & -5 \\ 2 & 4 \end{vmatrix} + 21\begin{vmatrix} 4 & -5 \\ 1 & 4 \end{vmatrix} \\ = &\quad -21(4+10) - 21(-5-16) = 147\,. \end{aligned} \tag{1.36}$$

Calculating the inverse of a matrix

The inverse of a matrix is defined through

$$\hat{A}\hat{A}^{-1} = \mathbf{1} \quad \text{or} \quad a_{ij}a^{-1}_{jk} = \delta_{ik} \,. \tag{1.37}$$

For the determination of the inverse of a $n \times n$-matrix, there are actually three methods, which are related to each other:

1. **Direct calculation:**
   The school book method is to calculate it directly through

$$\hat{A}^{-1} = \frac{1}{\det\hat{A}} \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{pmatrix}^T \,, \tag{1.38}$$

   where the A_{ij} are the adjoints, i.e. determinants of $(n-1) \times (n-1)$-matrices.

Examples:

- 2×2-matrix:

$$\begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} 2 & -4 \\ -1 & 3 \end{pmatrix}^T = \frac{1}{2}\begin{pmatrix} 2 & -1 \\ -4 & 3 \end{pmatrix} \,. \tag{1.39}$$

  Obviously

$$\begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix} \cdot \frac{1}{2}\begin{pmatrix} 2 & -1 \\ -4 & 3 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \mathbf{1} \,. \tag{1.40}$$

- 3×3-matrix:

$$\begin{pmatrix} 3 & 2 & 1 \\ 1 & 0 & 2 \\ 4 & 1 & 3 \end{pmatrix}^{-1} = \frac{1}{5}\begin{pmatrix} -2 & -5 & 4 \\ 5 & 5 & -5 \\ 1 & 5 & -2 \end{pmatrix} \,. \tag{1.41}$$

---

2. **Solving a system of linear equations:**
   Going back to the definition,

$$\hat{A}\hat{X} = \hat{A}\hat{A}^{-1} = \mathbf{1} \tag{1.42}$$

   it is clear that for the n^2 entries of the unknown matrix $\hat{X}$ (which is supposed to become the inverse $\hat{A}^{-1}$) there are n^2 systems of linear equations. Solving them yields the desired result.

3. **"Back of the envelope" for small matrices:**
   This method is also known as **Gauss-Jordan** method. The idea here is to have the original matrix and the unit matrix of the same size next to each other, say, the original matrix to the left and the unit matrix to the right. Then, on both matrices an identical series of steps is performed such that the original matrix is manipulated such that the unit matrix results. This is exemplified in Tab. 1.1.

At this point it should be stressed that in general the existence of the inverse of a matrix is not guaranteed. This can be seen directly from the occurrence of the determinant in the denominator in the first calculation method. It also shows under which circumstances an inverse does not exist, namely if the rank of the matrix is smaller than its actual dimension.

1.1.3 Solving systems of linear equations

This allows for the solution of systems of linear equations, i.e. of

$$\hat{A}\vec{x} = \vec{b} \quad \text{or} \quad a_{ij}x_j = b_i\,. \tag{1.43}$$

If all elements of $\vec{b}$ equal zero, $b_i = 0$ then this is called a homogeneous system of linear equations.

Behavior of the solutions

A number of points are worth noting:

1. If the system $\hat{A}\vec{x} = \vec{b}$ is homogeneous, $b_i = 0$, then there always is a trivial solution, namely the null vector, i.e. $x_i = 0$. The relevant question therefore is whether there are other, non-trivial solutions.

Example:

$\begin{pmatrix} 3 & 2 & 1 \\ 1 & 0 & 2 \\ 4 & 1 & 3 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	exchange 1st and 2nd row
$\begin{pmatrix} 1 & 0 & 2 \\ 3 & 2 & 1 \\ 4 & 1 & 3 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	subtract multiples of 1st row
$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & -5 \\ 0 & 1 & -5 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & -3 & 0 \\ 0 & -4 & 1 \end{pmatrix}$	subtract 3rd row from 2nd row
$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 1 & -5 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & -1 \\ 0 & -4 & 1 \end{pmatrix}$	subtract 2nd row from 3rd row
$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & -5 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & -1 \\ -1 & -5 & 2 \end{pmatrix}$	divide 3rd row by -5
$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 1 & -1 \\ 1/5 & 1 & -2/5 \end{pmatrix}$	subtract twice 3rd row
$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} -2/5 & -1 & 4/5 \\ 1 & 1 & -1 \\ 1/5 & 1 & -2/5 \end{pmatrix}$	

Table 1.1: "Back of the envelope"-inversion of a matrix.

2. If $\vec{x}_1$ and $\vec{x}_2$ are linearly independent solutions of the homogeneous problem, then any linear combination of them is a solution of the homogeneous problem. This follows directly from the linearity.

3. If $\vec{x}_i$ is a solution of the inhomogeneous problem and $\vec{x}_h$ is a linearly independent solution of the homogeneous problem, then any linear combination of $\vec{x}_i$ and $\vec{x}_h$ is a solution of the inhomogeneous problem. This follows directly from the linearity.

Let us consider the case where $\hat{A}$ is quadratic. Then, obviously, multiplying with $\hat{A}^{-1}$ from the right yields

$$\hat{A}^{-1}\hat{A}\vec{x} = \vec{x} = \hat{A}^{-1}\vec{b} \tag{1.44}$$

or

$$a_{ij}^{-1}a_{jk}x_k = \delta_{ik}x_k = x_i = a_{ij}^{-1}b_j \,. \tag{1.45}$$

In case the inverse of $\hat{A}$ exists, the solution is unambiguous. This is obviously not the case, if $\det\hat{A} = 0$ - in this case not all rows and columns of the matrix are linearly independent, and therefore the system is under-determined.

Cramer's rule

Another way to solve a system of n equations for n unknowns is based on determinants and known as Cramer's rule. Although, in principle, row reduction is superior in terms of calculation steps for matrices explicitly consisting of numbers, Cramer's method may be advantageous if the matrix elements are not numerical but, e.g., functions. To see how this method works, however, numerical examples are best-suited. Therefore, consider the simplest case, two equations for two unknowns

$$a_{11}x_1 + a_{12}x_2 = b_1 \tag{1.46}$$
$$a_{21}x_1 + a_{22}x_2 = b_2 \,. \tag{1.47}$$

Multiplying the first row by a_{22} and the second one by a_{12} and subtracting both yields

$$x_1 = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{21}a_{12}} \,, \tag{1.48}$$

and, similarly,

$$x_2 = \frac{b_2 a_{11} - b_1 a_{21}}{a_{11} a_{22} - a_{21} a_{12}} \,. \tag{1.49}$$

Obviously, this is a viable solution only, if the denominator in both cases is different from zero. This can be cast into

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \,, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}} \,. \tag{1.50}$$

The recipe which determinants to take for this and the general case reads as follows: Take the matrix of the coefficients, $\hat{A}$. Its determinant in all cases yields the denominator. For the solution of x_m, replace the mth column of $\hat{A}$, the a_{im} with the inhomogeneous coefficients b_i and form the determinant of this new matrix. It yields the numerator.

1.2 Eigenvalues and Eigenvectors

1.2.1 Similarity Transformations

Linear Transformations of vectors

Consider the effect of a matrix $\hat{A}$ on a vector $\vec{x}$: When correctly multiplied, a new vector $\vec{x'}$ emerges, i.e.

$$a_{ij} x_j = x'_i \tag{1.51}$$

or

$$\vec{x}' = \phi(\vec{x}) = \hat{A}\vec{x} \,. \tag{1.52}$$

Clearly, the transformation law ϕ here has the following properties:

$$\phi(\vec{x} + \vec{y}) = \hat{A}\vec{x} + \hat{A}\vec{y} = \phi(\vec{x}) + \phi(\vec{y}) \tag{1.53}$$

and

$$\phi(\lambda\vec{x}) = \hat{A}(\lambda\vec{x}) = \lambda\hat{A}\vec{x} \,, \tag{1.54}$$

where λ is a real number. These are exactly the properties of a linear transformation, see below.
In general, linear transformations of a n-dimensional vector, where n may be infinite, can be represented by $n \times n$ matrices. The result of such a transformation is then obtained by multiplying the vector and the matrix (in Quantum Mechanics the lingo will be that an operator acts on a vector).
To be more specific and, in fact, more intuitive, consider vectors of dimension 3 only, which can be used to specify, e.g., the position or momentum of a particle or similar. The **norm** of such a three-vector is given by

$$||\vec{x}|| = \sqrt{\vec{x}^2} = \sqrt{\vec{x}\cdot\vec{x}} = \sqrt{x_i x_i} = \sqrt{x_i^2} = \sqrt{x_1^2 + x_2^2 + x_3^2}\,. \tag{1.55}$$

Clearly, the norm of such a vector is then identical to its length.

Orthogonal and unitary transformations

Linear transformations, acting on a real vector space, that leave the norm of a vector unchanged are called **Orthogonal Transformations**. Such matrices with real elements which represent transformations of real vectors that leave the norm invariant, i.e.

$$\vec{x}' \cdot \vec{x}' = x_i'^T x_i' = (A_{ij}x_j)^T (A_{ik}x_k) = x_j A_{ji} A_{ik} x_k = x_i x_i \tag{1.56}$$

have the following important property:

$$x_j A_{ji} A_{ik} x_k = x_i x_i \quad \Longrightarrow \quad A_{ji} A_{ik} = \delta_{jk} \quad \Longrightarrow \quad A^T = A^{-1}\,. \tag{1.57}$$

This, in fact, holds true for **all** orthogonal transformations.
Another important case occurs, if the transformations act on complex vector spaces, such as $\mathbf{C}^n$. In this case, the norm of a vector is defined through

$$|\vec{v}|^2 = \sum_{i=1}^{n} |a_i|^2 = a_i a_i^*\,, \tag{1.58}$$

where the asterisk denotes complex conjugation. In this case, the transformations that leave the norm invariant are called **Unitary Transformations**, and the corresponding matrices have the property

$$\hat{M}^\dagger = \hat{M}^{-1}\,. \tag{1.59}$$

In both case, however, linear transformations which leave the norm of vectors invariant have a determinant equal to ± 1.

Example:
As an example, consider for a moment rotations of this vector, clearly leaving its length unchanged and merely altering the direction. Such an operation can be interpreted in two ways: Either the vector actually rotated and the axes remained fixed, or the vector remained fixed, whereas the axes rotated (in the other direction). So, it should not come as a surprise that the norm of vectors remains unchanged.
As an example consider counterclockwise rotations by an angle ϕ around the z-axis. Such transformations are mediated by matrices of the form

$$\hat{T}_\phi = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{1.60}$$

It is simple to check that indeed $T_\phi^T = T_\phi^{-1}$ and that $\det\hat{T}_\phi = 1$.
As a side-product, consider two subsequent rotations around the angles ϕ and ψ around the z-axis. It is well-known that for sequences of **rotations around the same axis**, the angles just add up, such that

$$\begin{aligned} \hat{T}_{\phi+\psi} &= \begin{pmatrix} \cos(\phi+\psi) & -\sin(\phi+\psi) & 0 \\ \sin(\phi+\psi) & \cos(\phi+\psi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \hat{T}_\phi T_\psi = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos\phi\cos\psi - \sin\phi\sin\psi & -\cos\phi\sin\psi - \sin\phi\cos\psi & 0 \\ \sin\phi\cos\psi + \cos\phi\sin\psi & -\sin\phi\sin\psi + \cos\phi\cos\psi & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned} \tag{1.61}$$

From this, some trigonometric relations are readily identified, namely

$$\begin{aligned} \cos(\phi+\psi) &= \cos\phi\cos\psi - \sin\phi\sin\psi \\ \sin(\phi+\psi) &= \sin\phi\cos\psi + \cos\phi\sin\psi\,. \end{aligned} \tag{1.62}$$

Properties of orthogonal and unitary matrices

One of the more interesting properties of orthogonal matrices is that multiplying two of them again yields an orthogonal matrix. Consider two orthogonal matrices $\hat{O}_1$ and $\hat{O}_2$ and their product $(\hat{O}_1\hat{O}_2)$. Clearly,

$$\mathbf{1} = \left(\hat{O}_1\hat{O}_2\right)\left(\hat{O}_1\hat{O}_2\right)^{-1} = \hat{O}_1\hat{O}_2\hat{O}_2^{-1}\hat{O}_1^{-1}\,. \tag{1.63}$$

Inserting that for orthogonal matrices $\hat{O}^T = \hat{O}^{-1}$ then yields

$$\mathbf{1} = \hat{O}_1\hat{O}_2\hat{O}_2^T\hat{O}_1^T = \left(\hat{O}_1\hat{O}_2\right)\left(\hat{O}_1\hat{O}_2\right)^T\,, \tag{1.64}$$

where Eq. (1.23) has been used. Therefore,

$$\left(\hat{O}_1\hat{O}_2\right)^{-1} = \left(\hat{O}_1\hat{O}_2\right)^T\,, \tag{1.65}$$

as advertised.
Similarly, it can be shown that the product of two unitary matrices again is unitary. To this end, assume to unitary matrices $\hat{U}_1$ and $\hat{U}_2$. Obviously,

$$\mathbf{1} \;=\; \left(\hat{U}_1\hat{U}_2\right)\left(\hat{U}_1\hat{U}_2\right)^{-1} = \hat{U}_1\hat{U}_2\hat{U}_2^{-1}\hat{U}_1^{-1} = \hat{U}_1\hat{U}_2\hat{U}_2^\dagger\hat{U}_1^\dagger\,. \tag{1.66}$$

Here the fact that for unitary matrices $\hat{U}^\dagger = \hat{U}^{-1}$ has been used. Since adjoint is - up to complex conjugation - nearly identical to transposition,

$$\hat{U}_2^\dagger\hat{U}_1^\dagger = \left(\hat{U}_1\hat{U}_2\right)^\dagger\,, \tag{1.67}$$

cf. Eqs. (1.23,1.24) and therefore

$$\left(\hat{U}_1\hat{U}_2\right)^{-1} = \left(\hat{U}_1\hat{U}_2\right)^\dagger\,, \tag{1.68}$$

fulfilling the definition of unitary matrices.

Group theory

Using the findings above, it is quite simple to show that the sets of all orthogonal or unitary transformations, respectively, which act on the same vector space, form a non-abelian group under multiplication. The above reasoning shows **closure**, i.e. the fact the product of two such orthogonal (unitary) transformations again yields an orthogonal (unitary) transformation. Since they are represented by matrices, their **associativity**,

$$\hat{A}(\hat{B}\hat{C}) = (\hat{A}\hat{B})\hat{C} \tag{1.69}$$

is simple to show. The identiy element in both cases is the unit matrix of corresponding dimension, and the inverse are uniquely defined and again orthogonal or unitary matrices.

Linear Transformations, general definition

Unique maps ϕ relating two vector spaces $\mathbf{V}$ and $\mathbf{V}'$ are called **linear maps** if they enjoy the following properties:

$$\phi(\vec{x}+\vec{y}) = \phi(\vec{x}) + \phi(\vec{y}) \quad \forall\, \vec{x},\, \vec{y} \in \mathbf{V}\,; \tag{1.70}$$

and

$$\phi(\lambda\vec{x}) = \lambda\phi(\vec{x}) \quad \forall\, \vec{x} \in \mathbf{V} \quad \text{and } \forall \text{ real numbers } \lambda\,. \tag{1.71}$$

If, in addition, for a linear map ϕ the two vector spaces $\mathbf{V}$ and $\mathbf{V}'$ are identical, $\mathbf{V} = \mathbf{V}'$ then the map is called **Linear Transformation** or **Linear Operator**.

Examples:

- The map $\phi(x_1,\, x_2,\, x_3) = (x_1,\, x_2)$, connecting the $\mathbf{R}^3$ with the $\mathbf{R}^2$ is linear. But the vector spaces are not identical, therefore, this is not a linear transformation.

- The map $\phi(x_1,\, x_2,\, x_3) = (x_1,\, 1)$, connecting the $\mathbf{R}^3$ with the $\mathbf{R}^2$ is not linear, because

$$\phi(\vec{x}+\vec{y}) = (x_1+y_1,\, 1) \neq (x_1,\, 1) + (y+1,\, 1) = (x_1+y_1,\, 2)\,. \tag{1.72}$$

From the properties above, it is clear that linear maps ϕ can be represented by matrices:
If $\dim(\mathbf{V}) = m$ and $\dim(\mathbf{V}') = n$ the matrix has dimensions $n \times m$. In order to see why a matrix is a good representation, it suffices to convince yourself that each vector $\vec{x} \in \mathbf{V}$ can be written as a linear combination of the base vectors of $\mathbf{V}$:

$$\vec{v} = \sum_{i=1}^{m} a_i \vec{e}_i^{(\mathbf{V})} \tag{1.73}$$

and thus the essence of the linear transformation is in the information of how the base vectors of $\mathbf{V}$ transform under ϕ. This representation is not quite unique: Each map $\mathbf{V} \to \mathbf{V}'$ is represented by a set of matrices, one per combination of base vectors. The specific matrix then emerges in the following way: The kth column of the matrix $\hat{T}$ is given by the (n-dimensional) vector in $\mathbf{V}'$ emerging as a result of ϕ acting on the kth base vector of $\mathbf{V}$:

$$t_{ik} = [\phi(\vec{e}_k^{(\mathbf{V})}]_i = [\hat{T}\vec{e}_k^{(\mathbf{V})}]_i \,. \tag{1.74}$$

Thus, switching from one set of base vectors to another base set changes the matrix $\hat{T}$ representing ϕ.
However, a pair of matrices representing the same transformation but for different choices of base vectors is called equivalent. Formally speaking, a change of base can be realized by another linear transformation, represented by a matrix $\hat{S}$. Thus,

$$\hat{T}_{\bar{B}} = \hat{S}^{-1}_{B\to\bar{B}} \hat{T}_B \hat{S}_{B\to\bar{B}} \,, \tag{1.75}$$

where the old and new set of bases B and $\bar{B}$ have been made explicit. Such a transformation is called a **Similarity Transformation**.
It can be shown that such **a similarity transformation leaves both the determinant and the trace of a matrix invariant**.

1.2.2 Eigenvalues and eigenvectors

Definition

Eigenvalues $\vec{\lambda}^{(i)}$ and eigenvectors $\lambda^{(i)}$ of a quadratic matrix $\hat{A}$ are defined by

$$\hat{A}\vec{\lambda}^{(i)} = \lambda^{(i)}\vec{\lambda}^{(i)} \,. \tag{1.76}$$

In other words: Eigenvectors are those vectors, which, when being subjected to a matrix $\hat{A}$ yield a multiple of themselves, where the multiple is the corresponding eigenvalue.
It is important to stress here that, if $\vec{\lambda}^{(i)}$ is an eigenvector, then also any multiple $c\vec{\lambda}^{(i)}$ with $c \in \mathbf{R}$ is an eigenvector.
The equation above can also be written as

$$\begin{aligned} (\hat{A} - \lambda^{(i)}\mathbf{1})\vec{\lambda}^{(i)} &= 0 \\ (a_{kl} - \lambda^{(i)}\delta_{kl})\lambda_l^{(i)} &= 0\,. \end{aligned} \tag{1.77}$$

This is a linear equation for each i. Solving it through determinants by Cramer's rule would yield only the trivial solution, $\lambda_m^{(i)} = 0$, unless the determinant of the coefficients was equal to zero,

$$\det(\hat{A} - \lambda^{(i)}\mathbf{1}) = 0\,. \tag{1.78}$$

In this case, the individual solutions would expressions of the type 0/0. However, in this case, the rows of the new matrix $\hat{A} - \lambda^{(i)}\mathbf{1}$ are not linearly independent, resulting in infinitely many solutions for each value $\lambda^{(i)}$ - which is the desired result, since, if any vector is an eigenvector than also real multiples of this vector is an eigenvector.
However, Eq. (1.78) often is called the **characteristic equation** for $\hat{A}$.

Examples:

- Consider

$$\hat{A}\vec{\lambda}^{(i)} = \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda^{(i)}\begin{pmatrix} x \\ y \end{pmatrix} = \lambda^{(i)}\vec{\lambda}^{(i)}\,. \tag{1.79}$$

  Solutions for the eigenvalues are obtained by solving

$$\det(\hat{A} - \lambda^{(i)}\mathbf{1}) = \begin{vmatrix} 5-\lambda & -2 \\ -2 & 2-\lambda \end{vmatrix} = \lambda^2 - 7\lambda + 6 = 0\,. \tag{1.80}$$

  Hence,

$$\lambda^{(1,2)} = \frac{7 \pm \sqrt{49-24}}{2} = \frac{7 \pm 5}{2} \tag{1.81}$$

  or

$$\lambda^{(1)} = 6 \quad \text{and} \quad \lambda^{(2)} = 1\,. \tag{1.82}$$

This fixes the eigenvalues. In order to fix the eigenvectors, the following two systems of linear equations must be solved

$$\begin{array}{rclcrcl} 5x^{(1)} - 2y^{(1)} & = & 6x^{(1)} & \qquad & 5x^{(2)} - 2y^{(2)} & = & x^{(2)} \\ -2x^{(1)} + 2y^{(1)} & = & 6y^{(1)} & & -2x^{(2)} + 2y^{(2)} & = & y^{(2)} \end{array} \tag{1.83}$$

for $(x^{(1,2)}, y^{(1,2)})$. Sets of solutions are given by one equation in each case, namely

$$-x^{(1)} = 2y^{(1)} \quad \text{and} \quad 2x^{(2)} = y^{(2)} \,. \tag{1.84}$$

This shows that the eigenvectors are multiples of

$$\vec{\lambda}^{(1)} = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 \\ 1 \end{pmatrix} \quad \text{and} \quad \vec{\lambda}^{(2)} = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ 2 \end{pmatrix} \,, \tag{1.85}$$

where a the vectors have already been normalized to unit length.

In order to diagonalise the matrix, build a transformation matrix $\hat{T}$ from the eigenvectors and its inverse,

$$\hat{T} = \left(\vec{\lambda}^{(1)} \;\; \vec{\lambda}^{(2)} \right) = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{and} \quad \hat{T}^{-1} = \frac{1}{\sqrt{5}} \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix} \tag{1.86}$$

Then

$$\hat{T}^{-1}\hat{M}\hat{T} = \cdots = \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \lambda^{(1)} & 0 \\ 0 & \lambda^{(2)} \end{pmatrix} \,, \tag{1.87}$$

as anticipated.

- As a 3-dimensional example consider the rotation matrix of the example above,

$$\hat{T}_\phi = \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \,. \tag{1.88}$$

There the characteristic equation reads

$$\begin{array}{rcl} 0 & = & (\cos\phi - \lambda)^2(1-\lambda) + \sin^2\phi(1-\lambda) \\ & = & (1-\lambda)(1 - 2\cos\phi\lambda + \lambda^2) \end{array} \tag{1.89}$$

leading to

$$\lambda^{(1)} = 1 \tag{1.90}$$

with

$$\vec{\lambda}^{(1)} = \vec{e}_z \tag{1.91}$$

and

$$\lambda^{(2,3)} = \cos\phi \pm \sqrt{\cos^2\phi - 1} = \cos\phi \pm i\sin\phi\,. \tag{1.92}$$

The occurrence of complex numbers is slightly disturbing here, but, of course, this may happen. It should be noted that in this case usually also the eigenvectors become complex.

Diagonalizing a matrix

In the most general case, a matrix $\hat{M}$ may be diagonalized by invoking a linear transformation represented by a matrix $\hat{T}$ such that

$$\hat{T}^{-1}\hat{M}\hat{T} = \hat{D}\,, \tag{1.93}$$

where $\hat{D}$ has diagonal form. The question is how the transformation matrix and the diagonal matrix can be determined. To this end, write the equations determining the n eigenvalues and eigenvectors in matrix form,

$$\begin{aligned} &\hat{M}\begin{pmatrix} \lambda_1^{(1)} & \lambda_1^{(2)} & \dots & \lambda_1^{(n)} \\ \lambda_2^{(1)} & \lambda_2^{(2)} & \dots & \lambda_2^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_n^{(1)} & \lambda_n^{(2)} & \dots & \lambda_n^{(n)} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1^{(1)} & \lambda_1^{(2)} & \dots & \lambda_1^{(n)} \\ \lambda_2^{(1)} & \lambda_2^{(2)} & \dots & \lambda_2^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_n^{(1)} & \lambda_n^{(2)} & \dots & \lambda_n^{(n)} \end{pmatrix} \begin{pmatrix} \lambda^{(1)} & 0 & \dots & 0 \\ 0 & \lambda^{(2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda^{(n)} \end{pmatrix}\,. \end{aligned} \tag{1.94}$$

Simple inspection shows that $\hat{T}$ can be constructed as a matrix, where the columns are formed by the eigenvectors of $\hat{M}$, and $\hat{D}$ is a diagonal matrix with the eigenvalues as diagonal elements,

$$\hat{T} = \left(\vec{\lambda}^{(1)} \;\; \vec{\lambda}^{(2)} \;\; \ldots \;\; \vec{\lambda}^{(n)}\right) \begin{pmatrix} \lambda_1^{(1)} & \lambda_1^{(2)} & \ldots & \lambda_1^{(n)} \\ \lambda_2^{(1)} & \lambda_2^{(2)} & \ldots & \lambda_2^{(n)} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_n^{(1)} & \lambda_n^{(2)} & \ldots & \lambda_n^{(n)} \end{pmatrix} \quad \text{and} \quad \hat{D} = \begin{pmatrix} \lambda^{(1)} & 0 & \ldots & 0 \\ 0 & \lambda^{(2)} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda^{(n)} \end{pmatrix} \tag{1.95}$$

Example:

- In order to see how this works, take one of the previous examples, cf. Eq. (1.83):

$$\begin{array}{rcl} 5x^{(1)} - 2y^{(1)} & = & 6x^{(1)} \\ -2x^{(1)} + 2y^{(1)} & = & 6y^{(1)} \end{array} \qquad \begin{array}{rcl} 5x^{(2)} - 2y^{(2)} & = & x^{(2)} \\ -2x^{(2)} + 2y^{(2)} & = & y^{(2)}\,, \end{array} \tag{1.96}$$

which can be written as one matrix equation, namely

$$\begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}. \tag{1.97}$$

Clearly, multiplying from the left with a suitable inverse,

$$\begin{aligned} & \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}^{-1} \begin{pmatrix} 5 & -2 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \\ = & \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix}^{-1} \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix} \end{aligned} \tag{1.98}$$

yields the similarity transformation that transforms the original matrix into a diagonal from with the eigenvalues on the diagonal.

In general, arbitrary $n \times n$ matrices $\hat{A}$, may be brought into diagonal form by a similarity transformation

$$\hat{S}^{-1}\hat{A}\hat{S} = \hat{A}_{\text{diag.}}\,, \tag{1.99}$$

where the matrix $\hat{S}$ corresponding to the similarity transformation is just given by the matrix whose rows are given by the eigenvectors of the original matrix $\hat{A}$ and the diagonal matrix is determined by the eigenvalues of $\hat{A}$. It must be stressed, however, that it is not guaranteed that any matrix $\hat{A}$ may be diagonalized.
In any case, once a matrix can be diagonalized, it is straightforward to see that its determinant equals the product of its eigenvalues. In addition, due to the properties of the similarity transformation diagonalizing the matrix, the trace of a matrix equals the sum of its eigenvalues.

Symmetric matrices and diagonalization

In this section, it will be proved that **symmetric matrices can always be diagonalized** and that **eigenvectors of different eigenvalues are always orthogonal**.
Proof: Consider the eigenvalue equation $\hat{A}\vec{\lambda} = \lambda\vec{\lambda}$ in component notation,

$$A_{ij}\lambda_j = \lambda\lambda_i\,. \tag{1.100}$$

Due to the symmetry property of $\hat{A}$, the eigenvalue equation can also be applied in the form

$$\vec{\kappa}^T\hat{A} = \kappa\vec{\kappa}^T \quad \text{or} \quad \kappa_i A_{ij} = \kappa\kappa_j \tag{1.101}$$

Multiplying the first equation from the left with $\vec{\kappa}^T$ and the second with $\vec{\lambda}$ form the right then yields, in components

$$\kappa_i A_{ij}\lambda_j = \lambda\kappa_i\lambda_i \quad \text{and} \quad \kappa_i A_{ij}\lambda_j = \kappa\kappa_j\lambda_j\,. \tag{1.102}$$

Replacing the summation over j with a summation over i on the r.h.s. of the second equation and subtracting both equations results in

$$0 = \kappa_i\lambda_i(\lambda-\kappa) = \vec{\kappa}\cdot\vec{\lambda}(\lambda-\kappa)\,. \tag{1.103}$$

This can be zero only, if either he two eigenvalues are identical or if the two vectors are orthogonal, proving the theorem.
It is worth noting that a similar theorem also holds true for Hermitian matrices: **The eigenvalues of Hermitian matrices are real**, and **eigenvectors corresponding to different eigenvalues are orthogonal**.

Physical applications of diagonalization

- Consider a body M which is made of a material with a constant mass density ρ. Its inertial tensor is a matrix given by

  $$m_{ij} = \int\limits_M \mathrm{d}^3 r \rho(\vec{r})(r^2\delta_{ij} - x_i x_j)\,. \tag{1.104}$$

  This is a diagonal matrix and hence can be diagonalized. It plays an important role in rotations.

  Since the angular momentum and velocity are related by

  $$\vec{L} = \hat{M}\vec{\omega} \tag{1.105}$$

  this shows that they are parallel only if they point along one of the eigenvectors of the inertial tensor. If, in contrast, they are not parallel to one of these vectors, in general, ω will not be constant but rather move in some precession or similar.

- Consider next an one-dimensional system where two masses are connected by one spring each to a fixed wall and by one spring with each other. Assume the masses and the spring constants to be identical. Then, calling the excursion from rest of the two masses by x_1 and x_2 the systems Lagrangian is given by

  $$\begin{aligned} L &= T - V \\ &= \frac{m}{2}(\dot{x}_1^2 + \dot{x}_2^2) - \frac{k}{2}[x^2 + y^2 + (x-y)^2]\,. \end{aligned} \tag{1.106}$$

  This can be written ion matrix form as

  $$\begin{aligned} L \;=\; &+(\dot{x}_1,\ \dot{x}_2)\begin{pmatrix} m/2 & 0 \\ 0 & m/2 \end{pmatrix}\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} \\ &-(x_1,\ x_2)\begin{pmatrix} 2k/2 & -k/2 \\ -k/2 & 2k/2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\,. \end{aligned} \tag{1.107}$$

  The potential can be diagonalized with the eigenvalues

  $$\lambda^{(1)} = 1 \quad \text{and} \quad \lambda^{(2)} = 3\,. \tag{1.108}$$

  Now, whatever the similarity transformation is that diagonalizes the matrix of the potential, it is clear that it also leaves the matrix of the kinetic energy in diagonal form, since it is proportional to the unit matrix. Denoting the corresponding eigenvectors by $\vec{x}$ and $\vec{y}$, the equations of motion therefore read

$$\ddot{\vec{x}} + \omega^2 \vec{x} = 0 \quad \text{and} \quad \ddot{\vec{y}} + 3\omega^2 \vec{y} = 0\,, \tag{1.109}$$

where $\omega^2 = k/m$. This immediately leads to two oscillation modes with different frequency. The eigenvectors are given by

$$\vec{x} = \frac{1}{\sqrt{2}}(x_1 - x_2) \quad \text{and} \quad \vec{y} = \frac{1}{\sqrt{2}}(x_1 + x_2)\,, \tag{1.110}$$

i.e. a slow oscillations, where both masses are oscillating back and forth in parallel, and a fast oscillation, where the two masses oscillate against each other.

Chapter 2

Fourier Series

2.1 Motivation

2.1.1 Simple harmonic motion

Many problems in physics are related to oscillations/vibrations. Examples include the pendulum or harmonic oscillators, waves of sound or light, AC electric currents, etc.. Other phenomena, such as heat expansion, the physics of electric and magnetic fields, are, not so obvious, described through similar equations. In general, these equations contain a second derivative, something like $\mathrm{d}^2/\mathrm{d}t^2$ for the harmonic oscillator or, in three dimensions, the Laplace operator, ∇^2. Of course, in order to refer to physical situations, such equations are supplemented with boundary conditions. More than often, however, the best way to solve such equations is through an expansion in harmonic functions - a Fourier expansion.

Many ideas and a good part of the terminology used in this part of the lecture will be borrowed from the discussion of simple harmonic motion and wave motion. Hence, to begin with, these two topics will briefly be reviewed.

Suppose a particle P moves with constant speed around a circle with radius R in the xy-plane. At the same time, let another particle, Q, move on the y-axis between $-R$ and R, such that the y-coordinates of P and R coincide at all times. Such a motion is called **simple harmonic motion**, as exhibited, for instance, by a harmonic oscillator without friction.

If the angular velocity of P is ω, and if P started at $\vartheta(t = 0) = 0$, then the

y-coordinate of P and Q read

$$y(t) = R\sin(\omega t)\,, \tag{2.1}$$

and the x-coordinate of P is given by

$$x(t) = R\cos(\omega t)\,. \tag{2.2}$$

In complex coordinates, one could also write the position of P as

$$\begin{aligned} z(t) &= x(t) + iy(t) \\ R\exp(i\omega t) &= R\cos(\omega t) + iR\sin(\omega t)\,. \end{aligned} \tag{2.3}$$

It is often worthwhile to use the complex notation also for Q, understanding that only the real or imaginary component are relevant for the discussion of its motion.
In any case, the velocity of Q is given by the derivative w.r.t. the time t:

$$v(t) = \dot{y}(t) = R\omega\cos\omega t \tag{2.4}$$

or the imaginary component of

$$\dot{z}(t) = Ai\omega e^{i\omega t}\,. \tag{2.5}$$

What do these quantities translate to? To see this, it is worthwhile to draw the function $y(t)$ in dependence of t. Quite obviously this function repeats itself, it is periodic (see below) with a period of $2\pi/\omega$. Physically, it shows the displacement of Q from $y = 0$, its equilibrium position. R, the maximal displacement, is called the **amplitude** of the vibration or the displacement. It is also called the amplitude of the function. Similarly, $R\omega$ is the maximal velocity, or the velocity amplitude, of the motion. Note that the velocity has the same period as the displacement, it is just shifted by some angle - the difference of sin and cos. Now, if the mass of Q is m, the kinetic energy is given by

$$E_{\rm kin} = \frac{m}{2}\dot{y}^2 = \frac{mR^2\omega^2}{2}\cos^2(\omega t)\,, \tag{2.6}$$

and the maximal kinetic energy is given by

$$E_{\rm kin}^{\rm max} = \frac{mR^2\omega^2}{2}\,. \tag{2.7}$$

For an idealized harmonic oscillator which does not use energy due to friction the total energy is always a constant, equal to the largest value of the kinetic energy. It is interesting to note that the total energy is proportional to the square of the velocity amplitude, waves behave in a similar manner.

2.1.2 Waves

Waves are yet another example of an oscillation. To illustrate this, consider highly idealized ("unrealistic" would put it better) water waves, following a sine curve. Taking a photo of the water surface at some time $t = 0$ then would reveal a function

$$z(x,\, t = 0) = A \sin \frac{2\pi x}{\lambda}\,, \tag{2.8}$$

where z represents the height of the water surface w.r.t. the equilibrium, x is the horizontal distance in one dimension, and λ is the distance between wave crests. Usually, λ is called the wave length of the waves.
Suppose now that another photo is taken at some time $t \neq 0$. Assuming the waves to travel with velocity v in positive x direction, the equation representing the water surface now reads

$$z(x,\, t) = A \sin \frac{2\pi(x - vt)}{\lambda}\,. \tag{2.9}$$

There's another interpretation to this: Rather than waiting for a time t and observe the change of the water surface at some fixed point x, one may also fix the time t and consider changes under displacement along x. In other words, the equation for $z(t)$ above yields the function of the water surface in dependence of both x and t and it is a mere question of circumstance which aspect (t fixed, x variable or vice versa) is considered.

2.1.3 More frequencies

To carry this discussion further, consider a string instrument, such as a violin. Again in extreme idealization, sound is produced in such instruments by displacing a string and letting it go again. If there was no energy loss, again, the string would continue to vibrate forever. But what are the allowed modes of this vibration? The answer is quite simple. Assume that the string has a length L. In equilibrium, it stretches along the x-coordinate, and it is fixed at the positions $x = 0$ and $x = L$. Displacements in the, say, y direction, can then be described in an oversimplified way as standing waves $y(x)$ of the string; but the boundary conditions are that $y(0) = y(L) = 0$. Functions which satisfy this are $\sin n\pi x/L$ with $n = 1, 2, 3, dots$. It is not hard to guess that in this oversimplified world all relevant oscillations of the string can be described as a superposition of these elementary modes. It is exactly this transformation of arbitrary oscillations into sums of **eigenmodes** which makes Fourier analysis such an important tool.

2.2 Trigonometric and Exponential functions

2.2.1 General properties of functions

Periodic functions

A function $f(x)$ is called **periodic** with a period $P > 0$, if

$$f(x + P) = f(x) \quad \forall\, x\ , \tag{2.10}$$

where P is a constant. The smallest value of P is called the least period or just the period of f.

Examples:

- the function $\sin x$ has periods 2π, 4π, $\dots$, since, $\sin(x + 2\pi) = \sin(x + 4\pi) = \cdots = \sin x$. Obviously, 2π is the least period of $\sin x$;
- the functions $\sin(nx)$ and $\cos(nx)$ with a constant n have periods $2\pi/n$;
- the function $\exp(ikx)$ with a real constant k and with real values of x has period $2\pi/k$;
- the period of $\tan x$ is π;
- constants have any positive number as period.

Piecewise continuous functions

A function $f(x)$ is called **piecewise continuous** in an interval $x \in I$, if

1. the interval can be divided into a finite number of subintervals, in which $f(x)$ is continuous; and

2. the limits of $f(x)$ are finite as x approaches the subintervals' endpoints.

Another, more pictorial way, of stating these conditions is to demand that $f(x)$ can be patched together with a finite number of patches with only finite jumps in between them.
The limit of $f(x)$ from the right is often denoted as

$$\lim f(x+0) = \lim_{\epsilon\to 0^+} f(x+\epsilon)\,, \tag{2.11}$$

where $\epsilon > 0$. Similarly, the left limit is denoted as

$$\lim f(x-0) = \lim_{\epsilon\to 0^+} f(x-\epsilon) = \lim_{\epsilon\to 0^-} f(x+\epsilon)\,. \tag{2.12}$$

Odd and even functions

A function is called **odd**, if $f(-x) = -f(x)$. If, on the other hand $f(-x) = f(x)$ then the function is called **even**.

Examples include:

- polynomials with even exponents only, such as $a_0 + a_2x^2 + a_4x^4 + \ldots$, where the a_i are constants, are even. In contrast, polynomials with odd exponents only, such as $a_1x + a_3x^3 + a_5x^5 + \ldots$ are odd functions;
- $\sin x$ is an odd function, whereas $\cos x$ is even.

It is quite simple to show that the product of two odd or two even functions is even, whereas the product of an odd and an even function is odd. Also,

$$\int\limits_{-A}^{A} \mathrm{d}x\, f(x) = \begin{cases} 0 & \text{if } f(x) \text{ is odd}\,, \\ 2\int\limits_{0}^{A} \mathrm{d}x\, f(x) & \text{if } f(x) \text{ is even}\,. \end{cases} \tag{2.13}$$

2.2.2 Some trigonometry

Euler's identity

Euler's identity states that

$$e^{i\vartheta} = \cos\vartheta + i\sin\vartheta\,. \tag{2.14}$$

This can be seen geometrically or using the Taylor expansion

$$\begin{aligned} f(x) &= f(x_0) + (x - x_0)\,\frac{\mathrm{d}}{\mathrm{d}x} f(x)\bigg|_{x=x_0} + \frac{(x-x_0)^2}{2}\,\frac{\mathrm{d}^2}{\mathrm{d}x^{\,2}} f(x)\bigg|_{x=x_0} + \dots \\ &= \sum_{n=0}^{\infty} \frac{(x-x_0)^n}{n!}\,\frac{\mathrm{d}^n}{\mathrm{d}x^{\,n}} f(x)\bigg|_{x=x_0}\,, \end{aligned} \tag{2.15}$$

where the obvious definition $\mathrm{d}^0/\mathrm{d}x^{\,0} f(x) = f(x)$ and $0! = 1$ have been used. Applying this on the functions above with $\vartheta_0 = 0$ yields

$$\begin{aligned} e^{i\vartheta} &= 1 + i\vartheta - \frac{1}{2}\vartheta^2 - \frac{1}{6}i\vartheta^3 + \frac{1}{24}\vartheta^4 + \frac{i}{120}\vartheta^5 + \dots \\ \cos(\vartheta) &= 1 \qquad - \frac{1}{2}\vartheta^2 \qquad\qquad + \frac{1}{24}\vartheta^4 + \dots \\ i\sin(\vartheta) &= \qquad i\vartheta \qquad\quad - \frac{i}{6}\vartheta^3 \qquad\qquad + \frac{i}{120}\vartheta^5 + \dots\,. \end{aligned} \tag{2.16}$$

Euler's identity above together with the properties $\sin(-\vartheta) = -\sin\vartheta$ and $\cos(-\vartheta) = \cos\vartheta$ of the trigonometric functions implies that

$$e^{-i\vartheta} = \cos\vartheta - i\sin\vartheta \tag{2.17}$$

and thus

$$\begin{aligned} \cos\vartheta &= \frac{e^{i\vartheta} + e^{-i\vartheta}}{2} \\ \sin\vartheta &= \frac{e^{i\vartheta} - e^{-i\vartheta}}{2i}\,. \end{aligned} \tag{2.18}$$

Trigonometric functions: addition & subtraction

From the identities above, Eq. (2.18), it is easy to see that

$$\begin{aligned} \sin(2\vartheta) &= \frac{e^{2i\vartheta} - e^{-2i\vartheta}}{2i} \\ &= \frac{1}{2i}\left[\left(e^{i\vartheta}\right)^2 - \left(e^{-i\vartheta}\right)^2\right] = \frac{1}{2i}\left[(\cos\vartheta + i\sin\vartheta)^2 - (\cos\vartheta - i\sin\vartheta)^2\right] \\ &= \frac{1}{2i}\left[4i(\cos\vartheta\sin\vartheta\right] = 2\cos\vartheta\sin\vartheta\,. \end{aligned} \tag{2.19}$$

In a similar way,

$$
\begin{aligned}
\sin\vartheta + \sin\phi = \frac{e^{i\vartheta} - e^{-i\vartheta} + e^{i\phi} - e^{-i\phi}}{2i} \\
&= \frac{1}{2i}\left[\exp\left(i\frac{\vartheta+\phi}{2} + i\frac{\vartheta-\phi}{2}\right) - \exp\left(-i\frac{\vartheta+\phi}{2} - i\frac{\vartheta-\phi}{2}\right)\right. \\
&\quad \left. + \exp\left(i\frac{\phi+\vartheta}{2} + i\frac{\phi-\vartheta}{2}\right) - \exp\left(-i\frac{\phi+\vartheta}{2} - i\frac{\phi-\vartheta}{2}\right)\right] \\
&= \frac{1}{2i}\left[\exp\left(i\frac{\vartheta+\phi}{2}\right) - \exp\left(-i\frac{\vartheta+\phi}{2}\right)\right] \\
&\quad \times \left[\exp\left(i\frac{\vartheta-\phi}{2}\right) + \exp\left(-i\frac{\vartheta-\phi}{2}\right)\right] \\
&= 2\sin(\vartheta+\phi)\cos(\vartheta-\phi)\,. \qquad (2.20)
\end{aligned}
$$

Also,

$$
\begin{aligned}
\sin\vartheta - \sin\phi &= 2\cos(\vartheta+\phi)\sin(\vartheta-\phi) \\
\cos\vartheta + \cos\phi &= 2\cos(\vartheta+\phi)\cos(\vartheta-\phi) \\
\cos\vartheta - \cos\phi &= -2\sin(\vartheta+\phi)\sin(\vartheta-\phi)\,. \qquad (2.21)
\end{aligned}
$$

Trigonometric functions: products

Euler's identity can also be employed to express products of trigonometric functions by sums and differences. For instance,

$$
\begin{aligned}
\sin\vartheta\sin\phi = \frac{e^{i\vartheta} - e^{-i\vartheta}}{2i}\frac{e^{i\phi} - e^{-i\phi}}{2i} \\
&= -\frac{1}{4}\left[e^{i(\vartheta+\phi)} - e^{i(\vartheta-\phi)} - e^{-i(\vartheta-\phi)} + e^{-i(\vartheta+\phi)}\right] \\
&= \frac{1}{2}\left[\cos(\vartheta-\phi) - \cos(\vartheta+\phi)\right]\,. \qquad (2.22)
\end{aligned}
$$

Also,

$$
\begin{aligned}
\sin\vartheta\cos\phi &= \frac{1}{2}\left[\sin(\vartheta-\phi) + \sin(\vartheta+\phi)\right] = \frac{1}{2}\left[\sin(\phi+\vartheta) - \sin(\phi-\vartheta)\right] \\
\cos\vartheta\cos\phi &= \frac{1}{2}\left[\cos(\vartheta-\phi) + \cos(\vartheta+\phi)\right]\,. \qquad (2.23)
\end{aligned}
$$

In particular, this implies for the squares of sine or cosine functions that

$$\begin{aligned}\sin^2\vartheta &= \frac{1}{2}(1-\cos 2\vartheta)\\ \cos^2\vartheta &= \frac{1}{2}(1+\cos 2\vartheta)\,.\end{aligned}\tag{2.24}$$

Trigonometric functions: derivatives and integrals

Remember that

$$\frac{\mathrm{d}\sin(kx)}{\mathrm{d}x} = +k\cos(x) \quad \text{and} \quad \frac{\mathrm{d}\cos(kx)}{\mathrm{d}x} = -k\sin(x)\,,\tag{2.25}$$

where k is a constant.
Therefore, their definite integrals read

$$\begin{aligned}\int_a^b \mathrm{d}x\ \sin(kx) &= -\frac{\cos kb - \cos ka}{k}\\ \int_a^b \mathrm{d}x\ \cos(kx) &= +\frac{\sin kb - \sin ka}{k}\,.\end{aligned}\tag{2.26}$$

Choosing $k = 1$ and $b - a = 2\pi$ immediately shows that the integrals of the trigonometric functions over 2π vanish[1].
For the integrals of the squared trigonometric functions, this however does not hold true. There are different ways of calculating, say, the integral of $\sin^2(x)$. In particular, one of them is well worth knowing. It starts by realizing that

$$\sin^2 x + \cos^2 x = 1\,.\tag{2.27}$$

Also, when integrated over the full period, both $\sin^2 x$ and $\cos^2 x$ yield the

[1]It is important to distinguish between definite and indefinite integrals: The former yields numbers as result, whereas the latter yield functions.

same result,

$$\begin{aligned}
\int_0^{2\pi} \mathrm{d}x\ \sin^2 x &= \int_0^{2\pi} \mathrm{d}x\ \cos^2 x = \int_0^{2\pi} \mathrm{d}x\,(1-\sin^2 x) \\
\Longrightarrow \int_0^{2\pi} \mathrm{d}x\ = 2\pi &= 2\int_0^{2\pi} \mathrm{d}x\ \sin^2 x \\
\Longrightarrow \pi &= \int_0^{2\pi} \mathrm{d}x\ \sin^2 x = \int_0^{2\pi} \mathrm{d}x\ \cos^2 x\,. \qquad (2.28)
\end{aligned}$$

This immediately implies that the average value of $\sin^2 x$, denoted by $\langle \sin^2 x\rangle$, is $1/2$:

$$\langle \sin^2 x\rangle = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}x\ \sin^2 x = \frac{\pi}{2\pi}\,. \qquad (2.29)$$

Trigonometric functions: integrals of products

Finally, let us calculate the integral (average value) of products of trigonometric functions over a period. To start with, consider

$$\int_0^{2\pi} \mathrm{d}x\ \sin(nx)\cos(mx) = \frac{1}{2}\int_0^{2\pi} \mathrm{d}x\ \{\sin[(n+m)x] + \sin[(n-m)x]\} = 0\,, \qquad (2.30)$$

where the identities of Eq. (2.23) have been used. With the same procedure,

$$\begin{aligned}
\int_0^{2\pi} \mathrm{d}x\ \sin(nx)\sin(mx) &= \begin{cases} 0 & \text{for } n \neq m \\ \pi & \text{for } n = m \neq 0 \\ 0 & \text{for } n = m = 0 \end{cases} \\
\int_0^{2\pi} \mathrm{d}x\ \cos(nx)\cos(mx) &= \begin{cases} 0 & \text{for } n \neq m \\ \pi & \text{for } n = m \neq 0 \\ 2\pi & \text{for } n = m = 0 \end{cases} \qquad (2.31)
\end{aligned}$$

can be obtained.

2.3 Definition & convergence

2.3.1 Definition

Assume a function $f(x)$ with the following properties:

- $f(x)$ is defined in the interval $[L, -L]$;
- $f(x)$ is periodic with period $2L$, i.e., $f(x+2L) = f(x)$.

Then, the **Fourier series** (or Fourier series expansion) of $f(x)$, here denoted as $\tilde{f}(x)$, is defined through

$$\tilde{f}(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L}\right) \tag{2.32}$$

where the coefficients (also known as **Fourier coefficients**) are given by

$$\begin{aligned} a_n &= \frac{1}{L}\int\limits_{-L}^{L} \mathrm{d}x\, f(x)\cos\frac{n\pi x}{L} \\ b_n &= \frac{1}{L}\int\limits_{-L}^{L} \mathrm{d}x\, f(x)\sin\frac{n\pi x}{L}\,. \end{aligned} \tag{2.33}$$

A motivation of the special form of these coefficients will be given in the next section, cf. Sec. .
This immediately gives an interpretation for the first term in the series expansion, $a_0/2$. Using the form of the Fourier coefficients above, it is clear that this term is nothing but the **mean** of $f(x)$ in the interval $[-L, L]$,

$$\frac{a_0}{2} = \frac{1}{2L}\int\limits_{-L}^{L} \mathrm{d}x\, f(x)\cos\frac{0\pi x}{L} = \frac{1}{2L}\int\limits_{-L}^{L} \mathrm{d}x\, f(x)\,. \tag{2.34}$$

In writing down these expressions for the Fourier expressions, it has implicitly been assumed that the function in question is continuous. If this is not the

case, the coefficients are obtained through integration over the continuous subintervals and adding the results.
It must be emphasized, however, that the Fourier series is only a series that corresponds to $f(x)$. It is not clear a priori (and in fact, counter examples can be constructed), whether it converges or, even if it does so, whether it converges to $f(x)$.

Examples:

- Consider first a constant, like $f(x) = c$ in the interval $[-\pi, \pi]$. Repeating itself with period 2π of course yields a constant function for all values of x. The Fourier coefficients are then given just by the respective integrals,

$$\begin{aligned}
\frac{a_0}{2} &= \frac{c}{2\pi}\int\limits_{-\pi}^{\pi} \mathrm{d}x = c\,;\\
a_n &= \frac{c}{\pi}\int\limits_{-\pi}^{\pi} \mathrm{d}x\,\cos(nx) = \frac{c}{\pi}\,\frac{\sin(nx)}{n}\bigg|_{-\pi}^{\pi} = 0\,;\\
b_n &= \frac{c}{\pi}\int\limits_{-\pi}^{\pi} \mathrm{d}x\,\sin(nx) = -\frac{c}{\pi}\,\frac{\cos(nx)}{n}\bigg|_{-\pi}^{\pi} = 0\,,
\end{aligned} \tag{2.35}$$

  and the Fourier series related to this function is the function itself.

- Consider now the function $f(x) = x\Theta(x)$ in the interval $[-\pi, \pi]$, with period 2π. Then

$$\begin{aligned}
\frac{a_0}{2} &= \frac{1}{\pi}\int\limits_{0}^{\pi} \mathrm{d}x\, x = \frac{\pi}{4}\,;\\
a_n &= \frac{1}{\pi}\int\limits_{0}^{\pi} \mathrm{d}x\, x\cos(nx) = \frac{1}{\pi}\,\frac{x\sin(nx)}{n}\bigg|_{0}^{\pi} - \frac{1}{\pi}\int\limits_{0}^{\pi}\mathrm{d}x\,\frac{\sin(nx)}{n}\\
&= \frac{1}{\pi}\,\frac{\cos(nx)}{n^2}\bigg|_{0}^{\pi} = \frac{1}{\pi n^2}[(-1)^n - 1] = \begin{cases} \dfrac{-2}{n^2\pi} & \text{for } n \text{ odd}\\ 0 & \text{for } n \text{ even}\,; \end{cases}
\end{aligned}$$

$$
\begin{aligned}
b_n &= \frac{1}{\pi}\int_0^\pi \mathrm{d}x\, x\sin(nx) = -\frac{1}{\pi}\,\frac{x\cos(nx)}{n}\bigg|_0^\pi + \frac{1}{\pi}\int_0^\pi \mathrm{d}x\,\frac{\cos(nx)}{n} \\
&= -\frac{1}{\pi}\left(\frac{x\cos(nx)}{n} + \frac{\sin(nx)}{n^2}\right)\bigg|_0^\pi = -\frac{(-1)^n}{n}\,. \qquad (2.36)
\end{aligned}
$$

In all cases, integration by parts has been used. Taken together, the Fourier series then reads

$$
\tilde{f}(x) = \frac{\pi}{4} + \sum_{n=1}^{\infty}\left[-\frac{1-(-1)^n}{\pi n^2}\cos(nx) - \frac{(-1)^n}{n}\sin(nx)\right] . \qquad (2.37)
$$

- As a next example, take the function $f(x) = x$ in the interval $[0,\,\pi]$, repeating itself with period π. Obviously, the interval is not symmetric around 0, but this merely changes the region of integration and the coefficients of the sine and cosine functions. Then, the Fourier coefficients are given by

$$
\begin{aligned}
\frac{a_0}{2} &= \frac{1}{\pi}\int_0^\pi \mathrm{d}x\, x = \frac{\pi}{2}\,; \\
a_n &= \frac{2}{\pi}\int_0^\pi \mathrm{d}x\, x\cos(2nx) = \frac{2}{\pi}\,\frac{x\sin(2nx)}{2n}\bigg|_0^\pi - \frac{2}{\pi}\int_0^\pi \mathrm{d}x\,\frac{\sin(2nx)}{2n} \\
&= \frac{2}{\pi}\,\frac{\cos(nx)}{n^2}\bigg|_0^\pi = 0
\end{aligned}
$$

$$
\begin{aligned}
b_n &= \frac{2}{\pi}\int_0^\pi \mathrm{d}x\, x\sin(nx) = -\frac{2}{\pi}\,\frac{x\cos(2nx)}{2n}\bigg|_0^\pi + \frac{2}{\pi}\int_0^\pi \mathrm{d}x\,\frac{\cos(2nx)}{2n} \\
&= -\frac{2}{\pi}\left(\frac{x\cos(nx)}{2n} + \frac{\sin(2nx)}{(2n)^2}\right)\bigg|_0^\pi = -\frac{1}{n}\,. \qquad (2.38)
\end{aligned}
$$

The related Fourier series thus reads

$$
\tilde{f}(x) = \frac{\pi}{2} - \sum_{n=1}^{\infty}\frac{\sin(2nx)}{n}\,. \qquad (2.39)
$$

- Now, take the function $f(x) = x^2$ in the interval $[-\pi, \pi]$, extended periodically with period 2π. The Fourier coefficients read

$$\begin{aligned} a_0 &= \frac{2\pi^2}{3} \\ a_n &= \frac{4(-1)^n}{n^2} \\ b_n &= 0\,. \end{aligned} \tag{2.40}$$

Thus, the Fourier series is given by

$$\tilde{f}(x) = \frac{\pi^2}{3} + 4\sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos(nx)\,. \tag{2.41}$$

It will be left as a problem to check this.

- Consider the Heavyside (or step) function

$$\Theta(x) := \begin{cases} 0 & \text{for } x \leq 0 \\ 1 & \text{for } x > 0 \end{cases} \tag{2.42}$$

in the interval $[-\pi, \pi]$ and repeat it periodically with period 2π. Its Fourier coefficients read

$$\begin{aligned} \frac{a_0}{2} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{d}x\,\Theta(x) = \frac{1}{2}\,; \\ a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{d}x\,\Theta(x)\cos(nx) = \frac{1}{\pi}\int_{0}^{\pi} \mathrm{d}x\,\cos(nx) = \left.\frac{\sin(nx)}{\pi n}\right|_0^{\pi} = 0\,; \\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{d}x\,\Theta(x)\sin(nx) = \frac{1}{\pi}\int_{0}^{\pi} \mathrm{d}x\,\sin(nx) = -\left.\frac{\cos(nx)}{n\pi}\right|_0^{\pi} \\ &= \begin{cases} \dfrac{2}{n\pi} & \text{for } n \text{ odd} \\ 0 & \text{for } n \text{ even}\,. \end{cases} \end{aligned} \tag{2.43}$$

Hence, the Fourier series related to the step function reads

$$\begin{aligned} \tilde{f}(x) &= \frac{1}{2} + \sum_{n=1}^{\infty} \frac{1}{(2n-1)\pi} \sin((2n-1)x) \\ &= \frac{1}{2} + \frac{2}{\pi}\left[\sin(x) + \frac{\sin(3x)}{3} + \frac{\sin(5x)}{5} + \ldots\right]\,. \end{aligned} \tag{2.44}$$

Note that for $x = 0$ this yields $\tilde{f}(0) = 1/2$, quite in contrast to the value of $f(x = 0) = \Theta(0) = 0$.

2.3.2 The form of the Fourier coefficients

To convince oneself that the form of the Fourier coefficients, cf. Eq. (2.33, is indeed correct, consider the Fourier series of a function $f(x)$ (Eq. (2.32)),

$$\tilde{f}(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi x}{L} + b_n \sin \frac{n\pi x}{L} \right) , \tag{2.45}$$

multiply it with $\cos(m\pi x/L)$ and integrate over x in the interval $[-L, L]$. This yields

$$\begin{aligned}
&\int_{-L}^{L} \mathrm{d}x\, \tilde{f}(x) \cos \frac{m\pi x}{L} \\
&= \int_{-L}^{L} \mathrm{d}x\, \cos \frac{m\pi x}{L} \left[\frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi x}{L} + b_n \sin \frac{n\pi x}{L} \right) \right] \\
&= \frac{La_0}{2m\pi} \sin \frac{m\pi x}{L} \bigg|_{-L}^{L} + \sum_{n=1}^{\infty} \left[a_n \left(\int_{-L}^{L} \mathrm{d}x\, \cos \frac{n\pi x}{L} \cos \frac{m\pi x}{L} \right) \right. \\
&\qquad \left. + b_n \left(\int_{-L}^{L} \mathrm{d}x\, \sin \frac{n\pi x}{L} \cos \frac{m\pi x}{L} \right) \right] \\
&= \sum_{n=1}^{\infty} a_n L \delta_{nm} = a_m L \, ,
\end{aligned} \tag{2.46}$$

where in the last step the properties of Sec. 2.2.2 have been employed. This immediately yields

$$a_m = \frac{1}{L} \int_{-L}^{L} \mathrm{d}x\, \tilde{f}(x) \cos \frac{m\pi x}{L} \, . \tag{2.47}$$

Similar statements of course hold true also for the coefficients b_n. In other words, assuming the Fourier series to coincide with the original function, i.e. $\tilde{f}(x) = f(x)$, fixes the expansion coefficients. So, again, the question remains, when the Fourier expansion coincides with the original function. This is answered in the next section.

2.3.3 Convergence: Dirichlet conditions

The **Dirichlet conditions** below answer the question, whether the Fourier series converges at all.
Suppose $f(x)$ fulfils the following conditions:

1. $f(x)$ is defined and single valued in the interval $[-L, L]$ except at a finite number of values;
2. $f(x)$ has a period of $2L$;
3. both $f(x)$ and its first derivative $f'(x)$ are piecewise continuous in $[-L, L]$.

Then the Fourier series converges such that

$$\tilde{f}(x) \longrightarrow \begin{cases} f(x) & \text{if } x \text{ is a point of continuity;} \\ \dfrac{f(x+0)+f(x-0)}{2} & \text{if } x \text{ is a point of discontinuity.} \end{cases} \tag{2.48}$$

An example for such a point of discontinuity is $x = 0$ for the Heavyside function, $\Theta(x)$. As could be seen from Eq. (2.44), its Fourier series there yields $1/2$, in accordance with the above statements.

The three conditions above are **sufficient but not necessary**: it is possible to have convergence without some of them. However, if they are met by the function $f(x)$, one may substitute $f(x)$ with the series expansion at the function's points of continuity. In any case, it is interesting to note that continuity of f alone does not guarantee the convergence of the series and that rather also its first derivative must be piecewise continuous.

2.3.4 Uniform convergence of series

Suppose an infinite series

$$\tilde{u}(x) = \sum_{n=1}^{\infty} u_n(x)\,. \tag{2.49}$$

The Rth partial sum $S_R(x)$ of the series is then defined as the sum of the R first terms,

$$S_R(x) = \sum_{n=1}^{R} u_n(x)\,. \tag{2.50}$$

Now, by definition, the series $\tilde{u}(x)$ converges to a function $u(x)$ in some interval, if for any value of x in the interval and for any positive number ϵ there exists a positive number N, such that

$$|S_R(x) - u(x)| < \epsilon \quad \forall\, R > N\,. \tag{2.51}$$

In general, this number N depends on both ϵ and the possible values of x. Anyways, $u(x)$ will be called the sum of the series in the following.
An important case, namely **uniform convergence**, occurs, when N does not depend on the range of possible x but on ϵ only. Such uniform convergent series have two important properties, namely

1. If each term of an infinite series (i.e. each of the $u_n(x)$) is continuous in an interval $]a,\, b[$ and if the series $\tilde{u}(x)$ is uniformly convergent to the sum $u(x)$, then

   - $u(x)$ is also continuous in the interval;
   - the series can be integrated term by term in this interval,

   $$\int\limits_a^b \mathrm{d}x \left(\sum_{n=1}^{\infty} u_n(x)\right) = \sum_{n=1}^{\infty}\left(\int\limits_a^b \mathrm{d}x\, u_n(x)\right)\,. \tag{2.52}$$

   In other words, in this case, summation and integration commute - they are interchangeable operations.

2. If, in addition, each term of this series has a derivative and if the series of derivatives is uniformly convergent, then the series can be differentiated term by term,

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{n=1}^{\infty} u_n(x)\right) = \sum_{n=1}^{\infty}\left(\frac{\mathrm{d}u_n(x)}{\mathrm{d}x}\right) . \tag{2.53}$$

2.3.5 The Weierstrass M-test

There are various ways to prove uniform convergence, the most obvious being to actually find $S_R(x)$ in closed form and then apply the definition directly. Another, more powerful way is to use the **Weierstrass M-test**:
If there exists a set of constants M_n with $n = 1, 2, 3,$ *dots* such that

$$|u_n(x)| \leq M_n \quad \forall x \in [a,\, b] \tag{2.54}$$

and if furthermore the sum $\sum_{n=1}^{\infty} M_n$ converges, then

$$\tilde{u}(x) = \sum_{n=1}^{\infty} u_n(x) \tag{2.55}$$

converges uniformly in the interval $[a,\, b]$. In fact, in this case, $\tilde{u}(x)$ is absolutely convergent, i.e. also the sum

$$\sum_{n=1}^{\infty} |u_n(x)| \tag{2.56}$$

of the absolute values of the terms is convergent.

Example: Consider the series $\sum_{n=1}^{\infty} \frac{\sin(nx)}{n^2}$. It converges uniformly in $[-\pi,\, \pi]$ (in fact, everywhere), because a set of constants $M_n = 1/n^2$ can be found, such that

$$\left|\frac{\sin(nx)}{n^2}\right| \leq \frac{1}{n^2} \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} . \tag{2.57}$$

2.4 Simple properties

2.4.1 Parseval's identity

Parseval's identity states that, if a function $f(x)$ satisfies the Dirichlet conditions,

$$\frac{1}{L}\int\limits_{-L}^{L} \mathrm{d}x\,[f(x)]^2 = \frac{a_0^2}{2} + \sum_{n=1}^{\infty}(a_n^2 + b_n^2)\,, \tag{2.58}$$

where the a_n and b_n are the Fourier coefficients defined in Eq. (2.33), which correspond to $f(x)$.
The proof of this identity will be left as a problem.

2.4.2 Sine and Cosine series

Expanding even and off functions

Remember the properties of even and odd functions, when multiplied with each other, cf. Sec. 2.2.1. If the function $f(x)$ is an even function in the interval $[L,\, -L]$ (i.e. $f(-x) = f(x)$) and periodic with period $2L$, its Fourier transform will contain only the cos-terms. This is quite simple to see: An even function (like $f(x)$) when multiplied with an odd function (like $\sin(kx)$) yields an odd function. And odd function vanish, when integrated over a symmetric interval like $[-L,\, L]$. Hence, in this case, all coefficients b_n equal 0, and thus all sin terms vanish. In this case, the Fourier series reduces to a Cosine series.
On the other hand, if $f(x)$ is an odd function in the interval, all its Fourier coefficients a_n equal 0, and thus the Fourier series reduces to a Sine series.

Examples:

- Consider first the function

  $$f(x) = x \tag{2.59}$$

  in the interval $[-\pi,\, \pi]$ and re-express is as either a Sine or Cosine series. Clearly, this is an odd function, therefore it can be expected that all coefficients a_n vanish.

From the definition of the Fourier coefficients, Eq. (2.33), one finds

$$\begin{aligned}
\frac{a_0}{2} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{d}x\, f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{d}x\, x = \left.\frac{x^2}{4\pi}\right|_{-\pi}^{\pi} = 0\,; \\
a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{d}x\, f(x)\cos(nx) = \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{d}x\, x\cos(nx) \\
&= \frac{1}{\pi}\left.\frac{x\sin(nx)}{n}\right|_{-\pi}^{\pi} - \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{d}x\, \frac{\sin(nx)}{n} = \frac{1}{\pi}\left.\frac{\cos(nx)}{n^2}\right|_{-\pi}^{\pi} = 0\,;
\end{aligned}$$

$$\begin{aligned}
b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{d}x\, f(x)\sin(nx) = \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{d}x\, x\sin(nx) \\
&= -\frac{1}{\pi}\left.\frac{x\cos(nx)}{n}\right|_{-\pi}^{\pi} + \frac{1}{\pi}\int_{-\pi}^{\pi} \mathrm{d}x\, \frac{\cos(nx)}{n} = -\frac{2(-1)^n}{n}\,. \qquad (2.60)
\end{aligned}$$

Hence the Fourier expansion of $f(x) = x$ looks like

$$\begin{aligned}
\tilde{f}(x) &= -2\sum_{n=1}^{\infty} \frac{(-1)^n}{n}\sin(nx) \\
&= 2\left[\sin(x) - \frac{\sin(2x)}{2} + \frac{\sin(3x)}{3} + \dots\right]. \qquad (2.61)
\end{aligned}$$

- On the other hand, consider the function

$$f(x) = |x| \qquad (2.62)$$

in the interval $[-\pi,\, \pi]$ and re-express is as either a Sine or Cosine series. Since this function is even, it can be expected that now all coefficients b_n vanish.

From the definition of the Fourier coefficients, Eq. (2.33), one finds

$$\begin{aligned}
\frac{a_0}{2} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{d}x\, f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{d}x\, |x| = -\frac{1}{2\pi}\int_{-\pi}^{0} \mathrm{d}x\, x + \frac{1}{2\pi}\int_{0}^{\pi} \mathrm{d}x\, x \\
&= -\left.\frac{x^2}{4\pi}\right|_{-\pi}^{0} - \left.\frac{x^2}{4\pi}\right|_{0}^{\pi} = \frac{\pi}{2}\,;
\end{aligned}$$

$$
\begin{aligned}
a_n &= \frac{1}{\pi}\int\limits_{-\pi}^{\pi} \mathrm{d}x\, f(x)\cos(nx) = -\frac{1}{\pi}\int\limits_{-\pi}^{0} \mathrm{d}x\, x\cos(nx) + \frac{1}{\pi}\int\limits_{0}^{\pi} \mathrm{d}x\, x\cos(nx) \\
&= -\left.\frac{x\sin(nx)}{n\pi}\right|_{-\pi}^{0} - \int\limits_{-\pi}^{0} \mathrm{d}x\,\frac{\sin(nx)}{n\pi} - \left.\frac{x\sin(nx)}{n\pi}\right|_{0}^{\pi} + \int\limits_{0}^{\pi} \mathrm{d}x\,\frac{\sin(nx)}{n\pi} \\
&= -\left.\frac{\cos(nx)}{n^2\pi}\right|_{-\pi}^{0} + \left.\frac{\cos(nx)}{n^2\pi}\right|_{0}^{\pi} = \frac{2(\cos n\pi - 1)}{n^2\pi} \\
&= \begin{cases} -\dfrac{4}{n^2\pi} & \text{if } n \text{ is odd} \\ 0 & \text{if } n \text{ is even}\,; \end{cases}
\end{aligned}
$$

$$
\begin{aligned}
b_n &= \frac{1}{\pi}\int\limits_{-\pi}^{\pi} \mathrm{d}x\, f(x)\sin(nx) = -\frac{1}{\pi}\int\limits_{-\pi}^{0} \mathrm{d}x\, x\sin(nx) + \frac{1}{\pi}\int\limits_{0}^{\pi} \mathrm{d}x\, x\sin(nx) \\
&= \left.\frac{x\cos(nx)}{n\pi}\right|_{-\pi}^{0} - \int\limits_{-\pi}^{0} \mathrm{d}x\,\frac{\cos(nx)}{n\pi} - \left.\frac{x\cos(nx)}{n\pi}\right|_{0}^{\pi} + \int\limits_{0}^{\pi} \mathrm{d}x\,\frac{\cos(nx)}{n\pi} \\
&= \frac{(-1)^n\pi}{n\pi} - \frac{(-1)^n\pi}{n\pi} - \left.\frac{\sin(nx)}{n^2\pi}\right|_{-\pi}^{0} + \left.\frac{\sin(nx)}{n^2\pi}\right|_{0}^{\pi} = 0\,. \qquad (2.63)
\end{aligned}
$$

Hence, in this case, the Fourier transform of $f(x)$ looks like

$$
\begin{aligned}
\tilde{f}(x) &= \frac{\pi}{2} + \frac{4}{\pi}\sum_{n=0}^{\infty}\frac{\cos[(2n+1)x]}{(2n+1)^2} \\
&= \frac{\pi}{2} + \frac{4}{\pi}\left[\cos(x) + \frac{\cos(3x)}{9} + \frac{\cos(5x)}{25} + \ldots\right]\,. \qquad (2.64)
\end{aligned}
$$

Half-range expansions

Assume now that a function is given as periodic over a range $[-L,\, L]$, but specified only in the interval $[0,\, L]$. In this case, one may use the freedom of **defining** the function in the other half of the period in a way such that

its Fourier expansion can be written entirely as a sine- or cosine series. The trick here is to define the function in $[-L, 0]$ such that it, taken over the full period, $f(x)$ is either even or odd.
Due to the symmetry properties of even and odd functions, $f^{\pm}_{\rm ext}(x)$, being multiplied with cosine or sine functions one finds

$$\left\{ \begin{array}{c} a_n \\ b_n \end{array} \right\} = \frac{1}{L}\int\limits_{-L}^{L} {\rm d}x f^{\pm}_{\rm ext}(x) \left\{ \begin{array}{c} \cos nx \\ \sin nx \end{array} \right\} = \frac{2}{L}\int\limits_{0}^{L} {\rm d}x f(x) \left\{ \begin{array}{c} \cos nx \\ \sin nx \end{array} \right\} , \quad (2.65)$$

potentially alleviating the task of calculating the Fourier coefficients.

Examples:

- Consider first the function $f(x) = x$ in $[0, 2]$ with period 4. An odd expansion is to **define** $f^{-}_{\rm ext}(x) = x$, whereas an even expansion is $f^{+}_{\rm ext}(x) = |x|$. In both cases, the Fourier expansion reads

$$\begin{aligned} \tilde{f}^{-}_{\rm ext} &= -\frac{4}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n}{n}\sin\frac{n\pi x}{2} \\ &= \frac{4}{\pi}\left(\sin\pi x 2 - \frac{1}{2}\sin 2\pi x 2 + \frac{1}{3}\sin 3\pi x 2 - \frac{1}{4}\sin 4\pi x 2 + \dots\right) \end{aligned} \quad (2.66)$$

  and

$$\begin{aligned} \tilde{f}^{+}_{\rm ext} &= 1 + \frac{4}{\pi^2}\sum_{n=1}^{\infty}\frac{(-1)^n - 1}{n^2}\cos\frac{n\pi x}{2} \\ &= 1 - \frac{8}{\pi^2}\left(\cos\pi x 2 + \frac{1}{3^2}\cos 3\pi x 2 + \frac{1}{5^2}\cos 5\pi x 2 + \dots\right) , \end{aligned} \quad (2.67)$$

  respectively.

- Maybe even more instructive is the expansion of the function $f(x) = \sin(x)$, $x \in [0, \pi]$ in terms of a cosine series. To this end the full 2π period is covered by defining $f(x)^{+}_{\rm ext} = |\sin x|$, i.e. by extending the half-range sine to an even full-range function. Then

$$
\begin{aligned}
\frac{a_0}{2} &= \frac{1}{\pi}\int\limits_0^\pi \mathrm{d}x \sin x = \frac{2}{\pi}\,, \\
a_n &= \frac{2}{\pi}\int\limits_0^\pi \mathrm{d}x \sin x \cos(nx) = \frac{1}{\pi}\int\limits_0^\pi \mathrm{d}x\,[\sin(x+nx)+\sin(x-nx)] \\
&= \frac{1}{\pi}\left[-\frac{\cos[(n+1)x])}{n+1}+\frac{\cos[(n-1)x]}{n-1}\right]_0^\pi = -\frac{2[1+(-1)^n]}{\pi(n^2-1)} \quad \text{if } n\neq 1\,, \\
a_1 &= \frac{2}{\pi}\int\limits_0^\pi \mathrm{d}x \sin x\cos x = \frac{2}{\pi}\left.\frac{\sin^2 x}{2}\right|_0^\pi = 0\,.
\end{aligned}
\tag{2.68}
$$

Therefore,

$$
\begin{aligned}
\tilde{f}^{+}_{\mathrm{ext}}(x) = \frac{2}{\pi}&\left[1-\sum_{n=2}^{\infty}\frac{1+(-1)^n}{n^2-1}\cos nx\right] \\
= \quad \frac{2}{\pi}-\frac{4}{\pi}&\left[\frac{\cos 2x}{2^2-1}+\frac{\cos 4x}{4^2-1}+\frac{\cos 6x}{6^2-1}+\ldots\right]\,.
\end{aligned}
\tag{2.69}
$$

2.4.3 Integration and differentiation

Integration and differentiation of Fourier series can be justified through the theorems of Sec. 2.3.4. In other words: If the Fourier series is uniformly convergent in its interval $[-L,\, L]$ it can be integrated term by term. Hence, in this case

$$
\begin{aligned}
\int_{-L}^{L} \mathrm{d}x\, \tilde{f}(x) &= \int_{-L}^{L} \mathrm{d}x\, \left[\frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L}\right)\right] \\
&= \frac{a_0}{2} \int_{-L}^{L} \mathrm{d}x\, + \sum_{n=1}^{\infty} \left[\int_{-L}^{L} \mathrm{d}x\, \left(a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L}\right)\right] \\
&= \frac{a_0 x}{2}\Big|_{-L}^{L} + \sum_{n=1}^{\infty} \left[a_n \int_{-L}^{L} \mathrm{d}x\, \cos\frac{n\pi x}{L} + b_n \int_{-L}^{L} \mathrm{d}x\, \sin\frac{n\pi x}{L}\right] \\
&= a_0 L + \sum_{n=1}^{\infty} \left[\frac{a_n L}{n\pi} \sin\frac{n\pi x}{L} - \frac{b_n L}{n\pi} \cos\frac{n\pi x}{L}\right]_{-L}^{L} = a_0 L\,. \qquad (2.70)
\end{aligned}
$$

This shows that, if the Fourier series is converging to the original function, the integral of both, function and Fourier transform, over the full period coincide and equal the first term in the Fourier series (the mean of the function) times the size of the interval. This should not come as a surprise at this stage.

Example: Consider the Fourier expansion of the function $f(x) = x$, cf. Eq. (2.61),

$$
\tilde{f}(x) = 2\left[\sin(x) - \frac{\sin(2x)}{2} + \frac{\sin(3x)}{3} + \dots\right]. \qquad (2.71)
$$

Integrating this yields

$$
\begin{aligned}
\int \mathrm{d}x\, \tilde{f}(x) &= 2\int \mathrm{d}x\, \left[\sin(x) - \frac{\sin(2x)}{2} + \frac{\sin(3x)}{3} + \dots\right] \\
&= 2\left[-\cos(x) + \frac{\cos(2x)}{4} - \frac{\cos(3x)}{9} + \dots\right] = 2\sum_{n=1}^{n} \frac{(-1)^n \cos(nx)}{n^2}\,.
\end{aligned}
\qquad (2.72)
$$

Since $f(x)$ satisfies the Dirichlet conditions, $\tilde{f}(x)$ can be replaced by $f(x)$

and thus

$$\frac{x^2}{2} = 2\sum_{n=1}^{n} \left.\frac{(-1)^n \cos(nx)}{n^2}\right|_0^x$$
$$= \ 2\sum_{n=1}^{n} \frac{(-1)^n \cos(nx)}{n^2} - 2\sum_{n=1}^{n} \frac{(-1)^n}{n^2}\,. \tag{2.73}$$

Adding in that

$$\sum_{n=1}^{n} \frac{(-1)^n}{n^2} = 1 - \frac{1}{4} + \frac{1}{9} - \frac{1}{16} + \dots = \frac{\pi^2}{12} \tag{2.74}$$

yields

$$x^2 = \frac{\pi^2}{3} + 4\sum_{n=1}^{n} \frac{(-1)^n \cos(nx)}{n^2}\,, \tag{2.75}$$

where the r.h.s. is exactly the Fourier expansion for x^2, cf. Eq. (2.41.
It must be stressed here however that in general the term-by-term integration of a Fourier series does not reproduce the Fourier expansion of the corresponding integrated function; this is due to the presence of the a_0-term which results in something like a_0L, where L is the length of the integration interval.

If in addition, also the derivatives are uniformly convergent, the series can also be differentiated term by term. But it is by no means guaranteed that the differentiated series has anything to do with the derivative of the original function.

Example: As an example for this consider again

$$x = 2\left[\sin(x) - \frac{\sin(2x)}{2} + \frac{\sin(3x)}{3} + \dots\right]\,, \tag{2.76}$$

cf. Eq. (2.61). Differentiating on both sides w.r.t. x clearly yields 1 for the l.h.s.. The right hand side, however, reads

$$2\left[\cos(x) - \cos(2x) + \cos(3x) + \dots\right]\,, \tag{2.77}$$

clearly different from 1. Even worse, this series does not converge at all.

Nevertheless, if both the function itself, $f(x)$ and its derivative, $f'(x)$, satisfy the Dirichlet conditions in the interval in question then term-by-term differentiation of the Fourier series of $f(x)$ is permissible in the sense that it leads to the Fourier series of $f'(x)$.
The catch here is that the Dirichlet conditions demand continuity of the function, even when expanded with its given periodicity. Simple inspection shows that for $f(x) = x^2$ in the interval $[-\pi, \pi]$ this holds true, basically because $f(\pi) = f(-\pi)$. In contrast for the function $f(x) = x$ in the same interval, there are discontinuities at the boundaries of the period.

2.4.4 Complex notation

Another form of the Fourier series can be obtained by using Euler's relation

$$e^{i\vartheta} = \cos\vartheta + i\sin\vartheta \tag{2.78}$$

and its inverse

$$\cos\vartheta = \frac{e^{i\vartheta} + e^{-i\vartheta}}{2} \quad \text{and} \quad \sin\vartheta = \frac{e^{i\vartheta} - e^{-i\vartheta}}{2i}\,. \tag{2.79}$$

These can easily be inserted into the Fourier expansion and, also, in the expressions determining the Fourier coefficients. But, of course, the coefficients can also be found directly and, using Euler's equation, the original sine and cosine terms can be recovered.
It is instructive to see how the coefficients of the complex Fourier series expansion emerge. To that end, consider a series of the form

$$f(x) = c_0 + c_1 e^{ix} + c_{-1}e^{-ix} + c_2 e^{2ix} + c_{-2}e^{-2ix} + \cdots = \sum_{n=-\infty}^{\infty} c_n e^{inx} \tag{2.80}$$

of a function $f(x)$ satisfying Dirichlet's conditions in the interval $[-\pi, \pi]$. Direct calculation (and previous experience with the trigonometric functions) shows that indeed

$$\int\limits_{-\pi}^{\pi} \mathrm{d}x\, e^{ikx} = 0 \tag{2.81}$$

for all integer values of k different from 0. Hence, integrating over the period yields

$$\int_{-\pi}^{\pi} \mathrm{d}x\, f(x) = 2\pi c_0 \tag{2.82}$$

and thus

$$c_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{d}x\, f(x)\,, \tag{2.83}$$

the average value of the function in the interval.
To find any of the c_n the recipe clearly is to compensate for the $\exp(inx)$ term on the r.h.s. of the equation above. This is most easily done by multiplying with $\exp(-inx)$ before integrating. Thus

$$\int_{-\pi}^{\pi} \mathrm{d}x\, f(x) e^{-inx} = \int_{-\pi}^{\pi} \mathrm{d}x\, c_n e^{inx} e^{-inx} = 2\pi c_n \tag{2.84}$$

and therefore

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{d}x\, f(x) e^{-inx} \tag{2.85}$$

for all values of n.
Note that this equation automatically contains the c_0 term as well, there are no factors of $1/2$ to be concerned about. Also,

$$c_{-n} = c_n^* \quad \text{if } f(x) \text{ is real}\,. \tag{2.86}$$

Example: Consider, once more, the Heavyside function in the interval $[-\pi,\, \pi]$.

Its complex Fourier expansion reads

$$
\begin{aligned}
c_0 &= \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} \mathrm{d}x\,\Theta(x) = \frac{1}{2} \\
c_n &= \frac{1}{2\pi}\int\limits_{-\pi}^{\pi} \mathrm{d}x\,\Theta(x) e^{-inx} \\
&= \frac{1}{2\pi}\int\limits_{0}^{\pi} \mathrm{d}x\, e^{-inx} = -\frac{e^{-inx}}{2in\pi}\bigg|_0^\pi \\
&= \frac{1-e^{in\imath}}{2in\pi} = \begin{cases} \dfrac{1}{in\pi} & \text{for } n \text{ odd} \\ 0 & \text{for } n \text{ even and } n \neq 0\,. \end{cases}
\end{aligned} \tag{2.87}
$$

Hence,

$$
\begin{aligned}
f(x) &= \frac{1}{2} + \frac{2}{\pi}\left[\frac{e^{ix}-e^{-ix}}{2i} + \frac{1}{3}\frac{e^{3ix}-e^{-3ix}}{2i} + \frac{1}{5}\frac{e^{5ix}-e^{-5ix}}{2i} + \dots\right] \\
&= \frac{1}{2} + \frac{2}{\pi}\left[\sin x + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \dots\right]\,,
\end{aligned} \tag{2.88}
$$

as before, cf. Eq. (2.44).

2.4.5 Double and multiple Fourier series

It is straightforward to extent the idea of Fourier expansion of a function of one variable to functions with more variables. As an example consider a function $f(x,\,y)$ defined in $\{x,\,y\} \in [-\pi,\,\pi]$ and periodic with period 2π in each dimension. Such a function can be expanded in a double Fourier series, like e.g.

$$
f(x,\,y) = \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} c_{nm} e^{inx} e^{imy} = \sum_{n,m=-\infty}^{\infty} c_{nm} e^{inx+imx}\,, \tag{2.89}
$$

where the coefficients are given by

$$c_{nm} = \frac{1}{(2\pi)^2} \int\limits_{-\pi}^{\pi} \mathrm{d}x \int\limits_{-\pi}^{\pi} \mathrm{d}y f(x,\, y) e^{-inx-imy} . \tag{2.90}$$

2.5 Some applications

2.5.1 Full wave rectifier

Wave rectifiers translate alternating into direct current (AC→DC). An interesting question related to this is how well the resulting current approaches direct current. As a toy example consider a wave rectifier, which lets positive parts of a sine wave pass and which inverts the negative parts,

$$\sin(\omega t) \stackrel{\text{Rect.}}{\longrightarrow} f(t) = \begin{cases} +\sin(\omega t) & \text{for } 0 \leq \omega t < \pi \\ -\sin(\omega t) & \text{for } -\pi \leq \omega t < 0 . \end{cases} \tag{2.91}$$

Since the result is an even function, there won't be any sin-parts in the Fourier series of the result. Consequently, only the a_n need to be calculated,

$$\begin{aligned} \frac{a_0}{2} &= -\frac{1}{2\pi} \int\limits_{-\pi}^{0} \mathrm{d}(\omega t)\, \sin(\omega t) + \frac{1}{2\pi} \int\limits_{0}^{\pi} \mathrm{d}(\omega t)\, \sin(\omega t) = \frac{2}{\pi}\,; \\ a_n &= -\frac{1}{\pi} \int\limits_{-\pi}^{0} \mathrm{d}(\omega t)\, \sin(\omega t) \cos(n\omega t) + \frac{1}{\pi} \int\limits_{0}^{\pi} \mathrm{d}(\omega t)\, \sin(\omega t) \cos(n\omega t) \\ &= \begin{cases} -\dfrac{4}{(n^2-1)\pi} & \text{for } n \text{ even} \\ 0 & \text{for } n \text{ odd} . \end{cases} \end{aligned} \tag{2.92}$$

Hence, the resulting series is

$$f(t) = \frac{2}{\pi} - \frac{4}{\pi} \sum_{n=2,4,6\ldots}^{\infty} \frac{\cos(n\omega t)}{n^2-1} = \frac{2}{\pi} \left[1 - \frac{2\cos(2\omega t)}{3} - \frac{2\cos(3\omega t)}{8} - \ldots \right] , \tag{2.93}$$

which is clearly dominated by the first few terms. Also, the original frequency ω has been eliminated and replaced by its higher harmonics 2ω, 3ω etc..

2.5.2 Quantum mechanical particle in a box

Sketching the physical problem

A good example for the merit of applying Fourier expansions, consider the quantum mechanical problem of a particle in a box. Such problems are often described by the time-independent Schrödinger equation. In one dimension, it reads

$$\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}u_E(x) = [V(x) - E]u_E(x)\,, \tag{2.94}$$

where the $u_E(x)$ are the energy eigenfunctions. They are related to the full wave function by

$$\psi(x,\,t) = \sum_E e^{-iEt}u_E(x) \tag{2.95}$$

with the sums stretching over all allowed energy eigenvalues of the problem. The box can now be represented by a potential of the form

$$V(x) = \begin{cases} 0 & \text{for } x \in [-a,\,a] \\ V & \text{else}\,, \end{cases} \tag{2.96}$$

with $V = \infty$. This infinitely high potential confines the particle to the region between $-a$ and a, it cannot penetrate the potential walls. According to the probabilistic interpretation of the wave function, its value

$$\psi(x,\,t) = e^{-iE/\hbar t}u_E(x) = 0 \quad \text{for } x \notin [-a,\,a] \tag{2.97}$$

translates into the particle being in the region. Therefore, suitable boundary conditions for the wave function inside the region $[-a,\,a]$ are

$$u_E(-a) = u_E(a) = 0\,. \tag{2.98}$$

General solution of the differential equation

A suitable *ansatz* for its solution is

$$u_E(x) = B^+\exp(ikx) + B^-\exp(-ikx)\,, \tag{2.99}$$

where

$$k = \sqrt{\frac{2mE}{\hbar^2}}\,. \tag{2.100}$$

Plugging this into the Schrödinger equation, translates into

$$
\begin{aligned}
& -\frac{\hbar^2}{2m}\,(ik)^2\,[B^+\exp(ikx)+B^-\exp(-ikx)] \\
= \quad & E[B^+\exp(ikx)+B^-\exp(-ikx)]
\end{aligned}
\tag{2.101}
$$

in the region $x \in [-a,\, a]$, proving the validity of the ansatz.

Adding in boundary conditions

Adding in boundary conditions

$$
u_E(\pm a) = B^+\exp(\pm ika) + B^-(\mp ika) = 0 \tag{2.102}
$$

implies there are two sets of solutions, namely

- Sine-type solutions:

$$
\begin{aligned}
B^+ = -B^- \equiv A^- \quad \text{and} \quad k = \frac{2n\pi}{2a} \\
\Longrightarrow \quad u_E(x) = A^-\left[\exp\left(\frac{2ni\pi x}{2a}\right) - \exp\left(-\frac{2ni\pi x}{2a}\right)\right]
\end{aligned}
\tag{2.103}
$$

- Cosine-type solutions:

$$
\begin{aligned}
B^+ = B^- \equiv A^+ \quad \text{and} \quad k = \frac{(2n-1)\pi}{2a} \\
\Longrightarrow \quad u_E(x) = A^+\left[\exp\left(\frac{(2n-1)i\pi x}{2a}\right) + \exp\left(-\frac{(2n-1)i\pi x}{2a}\right)\right]
\end{aligned}
\tag{2.104}
$$

In both cases, $n = 1,\, 2,\, 3,\, \ldots$.

Brief discussion of the physics

These two sets of solutions, +- and −-types, correspond to sinoidal and cosinoidal solutions and, ultimately, to different parities. Here, parity is the property of an object under spatial reflection, i.e. under the transformation $\vec{x} \to -\vec{x}$, or, in one dimension, $x \to -x$. More precisely, parity is defined through

$$
\mathcal{O} \stackrel{\vec{x}\to-\vec{x}}{\longrightarrow} P\mathcal{O}\,. \tag{2.105}
$$

Then clearly, the $\pm$-parts have parity $P = \pm$. Thus the $\pm$-label also labels the parity of the respective eigenmode.
The corresponding energy eigenmodes have kinetic energies only (the potential vanishes), namely

$$\begin{aligned} E_n^- &= \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m}\frac{(2n)^2\pi^2}{(2a)^2} \\ E_n^+ &= \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{2m}\frac{(2n-1)^2\pi^2}{(2a)^2} . \end{aligned} \tag{2.106}$$

Hence, each individual mode can also be written as

$$\begin{aligned} u_{E_n}^-(x) &= 2A_n^- \sin\frac{2n\pi x}{2a} \\ u_{E_n}^+(x) &= 2A_n^+ \cos\frac{(2n-1)\pi x}{2a} . \end{aligned} \tag{2.107}$$

To fix the amplitudes, it must be taken into account that the energy eigenmodes are also normalized, implying

$$1 \equiv \int_{-a}^{a} \mathrm{d}x |u_{E_n}^{\pm}(x)|^2 = 4a|A \pm_n|^2 \tag{2.108}$$

leading to

$$A_n = \frac{1}{2\sqrt{a}} \tag{2.109}$$

Taken everything together yields

$$u_n^{\pm}(x) = \frac{1}{\sqrt{a}} \begin{cases} \cos\dfrac{(2n-1)\pi x}{2a} \\ \sin\dfrac{2n\pi x}{2a} . \end{cases} \tag{2.110}$$

for each allowed value $n \in \mathbf{N}$. This can easily be generalized to the case of more than one dimension, i.e. a three-dimensional box with dimensions a, b and c.

2.5.3 Heat expansion

Sketching the physical problem

Heat conduction can be described by

$$\frac{\partial u(\vec{x},\, t)}{\partial t} = \kappa \nabla^2 u(\vec{x},\, t)\,, \tag{2.111}$$

where the function $u(\vec{x},\, t)$ denotes the temperature at position $\vec{x}$ at a given time t, and the constant κ is called diffusity and is related to

$$\kappa = \frac{K}{\sigma\mu}\,, \tag{2.112}$$

where K is the thermal conductivity, σ is the specific heat, and μ denotes the density of the material in question. The expression ∇^2 is called the **Laplacian**, given by

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\,. \tag{2.113}$$

To see how this works, consider the one-dimensional case of a slab with length L. Both its ends and the entire surface are insulated, translating into no heat transferred beyond the edges of the slab. In mathematical form this can be expressed as

$$u'(0,\, t) = u'(L,\, t) = 0\,. \tag{2.114}$$

Assume also that the slab has some initial temperature, say at time $t = 0$, then

$$u(x,\, 0) = f(x)\,., \tag{2.115}$$

General solution of the differential equation

Let us see now, how Fourier expansion helps us to determine the subsequent temperature, i.e. $u(x,\, t)$ with $t > 0$. The starting point will be an investigation of the general anatomy of the solutions of the differential equation describing heat conduction. For this, the ansatz

$$u(x,\, t) = X(x)T(t) \tag{2.116}$$

will be employed, leading to

$$XT' = \kappa X''T \tag{2.117}$$

when put into the equation of heat conduction, Eq. (2.111). This can be re-expressed as

$$\frac{T'}{\kappa T} = \frac{X''}{X}\,. \tag{2.118}$$

Since T is a function of t only and X merely depends on x, it should be clear that the equality can be realized only, if both sides equal a constant for all values of t and x, respectively. Setting this constant to be $-\lambda^2$, therefore

$$T' + \kappa\lambda^2 T = 0 \quad \text{and} \quad X'' + \lambda^2 X = 0\,. \tag{2.119}$$

Solutions for these two equations are given by

$$T = C\exp(-\kappa\lambda^2 t) \quad \text{and} \quad X = A\cos(\lambda x) + B\sin(\lambda x)\,, \tag{2.120}$$

yielding

$$u(x\,,t) = \exp(-\kappa\lambda^2 t)\left[a\cos(\lambda x) + b\sin(\lambda x)\right]\,, \tag{2.121}$$

with the constants $a = AC$ and $b = BC$.

Adding in boundary conditions

These constants as well as the constant $-\lambda^2$ need to be determined from the boundary conditions:

- The condition $u'(0,\,t) = 0$ enforces all sine terms to vanish, because their derivative would yield a cosine, which does not vanish for its argument being 0. Hence

$$b = 0\,. \tag{2.122}$$

- The condition $u'(L,\,t) = 0$ thus translates into

$$\cos'(\lambda L) \propto \sin(\lambda L) \stackrel{!}{=} 0 \implies \lambda L = m\pi \;\text{ rm with } m = 1,\,2,\,3,\,\ldots \tag{2.123}$$

and therefore

$$\lambda = \frac{m\pi}{L}\,. \tag{2.124}$$

- Finally, the boundary condition $u(x,\,0) = f(x)$ must be met. Hence

$$u(x,\,0) = f(x) = \frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos\left(\frac{m\pi x}{L}\right)\,, \tag{2.125}$$

  the result for $t = 0$. It looks like a Fourier cosine series, with coefficients thus given by

$$\begin{aligned} \frac{a_0}{2} &= \frac{1}{L}\int_0^L \mathrm{d}x\, f(x) \\ a_m &= \frac{1}{L}\int_0^L \mathrm{d}x\, f(x) \sin\left(\frac{m\pi x}{L}\right)\,, \end{aligned} \tag{2.126}$$

  cf. Eq. (2.33).

Hence,

$$\begin{aligned} u(x,\,t) = &\frac{1}{L}\int_0^L \mathrm{d}x\, f(x) \\ &+ \sum_{m=1}^{\infty}\left\{\left[\frac{1}{L}\int_0^L \mathrm{d}x\, f(x)\sin\left(\frac{m\pi x}{L}\right)\right]\cdot\cos\left(\frac{m\pi x}{L}\right)\exp\left(-\frac{\kappa m^2\pi^2 t}{L^2}\right)\right\}. \end{aligned} \tag{2.127}$$

2.5.4 Electrostatic potential

Sketching the physical problem

In the previous example, it has been seen that partial differential equations can sometimes be solved by separating the variables. By using this method, one often encounters sets of orthogonal functions, which are of great importance by themselves. In the following, this method will be applied once more, this time to the case of electrostatics. In the static case, the electric field $\vec{E}$ is determined through Gauss' law. In differential form it is given by

$$\epsilon_0 \vec{\nabla}\vec{E} = \rho\,, \tag{2.128}$$

where $\rho(x)$ is the fixed charge density giving rise to the field. A nice way of solving this problem is to realize that the rotor of the electric field in the static case vanishes, allowing to introduce the electrostatic potential ϕ. It relates to the electric field through

$$\vec{E} = -\vec{\nabla}\phi \,. \tag{2.129}$$

The potential is thus determined through the inhomogeneous Laplace equation

$$\nabla^2\phi = -\epsilon_0\rho \,, \tag{2.130}$$

encountered before. This equation is of great importance in physics.

Homogeneous Laplace equation

Let us now consider a case, where the potential $\phi(\vec{x})$ can be written in Cartesian coordinates as a product of three functions, one for each coordinate,

$$\phi(\vec{x}) = X(x)Y(y)Z(z) \,. \tag{2.131}$$

Setting the charge density to zero for a moment, the (homogeneous) Laplace equation reads

$$\frac{1}{X}\frac{\mathrm{d}^2X}{\mathrm{d}^2x} + \frac{1}{Y}\frac{\mathrm{d}^2Y}{\mathrm{d}^2y} + \frac{1}{Z}\frac{\mathrm{d}^2Z}{\mathrm{d}^2z} = 0 \,. \tag{2.132}$$

Here the partial differentials became total ones, since the functions depend on one variable only. In order to guarantee a solution for all $\vec{x}$, the following system has to be solved

$$\begin{aligned}
\frac{1}{X}\frac{\mathrm{d}^2X}{\mathrm{d}^2x} &= \alpha^2 \\
\frac{1}{Y}\frac{\mathrm{d}^2Y}{\mathrm{d}^2y} &= \beta^2 \\
\frac{1}{Z}\frac{\mathrm{d}^2Z}{\mathrm{d}^2z} &= \gamma^2 \,,
\end{aligned} \tag{2.133}$$

with the constraint that

$$\alpha^2 + \beta^2 + \gamma^2 = 0 \,. \tag{2.134}$$

A general solution for the homogeneous Laplace equation therefore reads

$$\phi(\vec{x}) = \phi_0 e^{\pm\alpha x} e^{\pm\beta y} e^{\pm i\sqrt{\alpha^2+\beta^2}z} \,, \tag{2.135}$$

with ϕ_0 the potential at $\vec{x} = 0$.

A physical problem and its solution

In order to determine α and β, boundary conditions need to be specified. As an example consider a cuboid with dimensions $a \times b \times c$ in x-, y, and z-direction, respectively. Furthermore, assume that all sides are on potential zero, apart from the one at $z = c$, which has a fixed potential $V_{z=c}(x,\, y)$. What is the potential inside the cuboid? The constraint that $\phi = 0$ for $x = 0$, $y = 0$, $z = 0$ implies

$$X = \sin(\alpha x)\ , \quad Y = \sin(\beta y)\ , \quad Z = \sinh(\sqrt{\alpha^2 + \beta^2} z)\,. \tag{2.136}$$

The additional constraint that $\phi = 0$ for $x = a$, $y = b$ yields

$$\alpha a = n\pi \quad \text{and} \quad \beta b = m\pi\,, \tag{2.137}$$

where n and m are integers. From these identification of the frequencies, the expansion coefficients of the Z function can be inferred. They read

$$\gamma_{nm} = \sqrt{\alpha^2 + \beta^2} = \pi\sqrt{\frac{n^2}{a^2} + \frac{m^2}{b^2}}\,. \tag{2.138}$$

Thus, the potential can be written as a sum of components, where each of which is given by

$$\phi_{nm} = A_{nm} \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi y}{b}\right) \sinh\left(\gamma_{nm} z\right)\,. \tag{2.139}$$

By construction this potential vanishes for $x = 0,\, a$, $y = 0,\, b$, and for $z = 0$. The coefficients A_{nm} can now be fixed such that

$$\sum_{nm} \phi_{nm}(x,\, y,\, z = c) = V_{z=c}(x,\, y)\,. \tag{2.140}$$

This seems horrible at first. By taking a closer look, however, it becomes clear that it is just a double Fourier series for the function $V_{z=c}$. Thus, the coefficients A_{nm} are given by

$$A_{nm} = \frac{1}{\sinh\left(\gamma_{nm} c\right)} \cdot \frac{4}{ab} \int\limits_0^a \mathrm{d}x \int\limits_0^b \mathrm{d}y V_{z=c}(x,\, y) \sin\left(\frac{n\pi x}{a}\right) \sin\left(\frac{m\pi y}{b}\right)\,. \tag{2.141}$$

If the potential is different from zero on all sides of the cuboid, then the overall potential can be constructed from the superposition principle: Just repeat this procedure for the other five sides, and add the partial results.

Chapter 3

Fourier transforms

3.1 Definitions and properties

3.1.1 The Fourier integral

The need for Fourier integrals

So far, Fourier series have been defined for periodic functions with period L. A natural question to ask is what happens, if $L \to \infty$. It will turn out that in such cases, the **Fourier series becomes a Fourier integral, where the sum over discrete frequencies is replaced by an integral over a continuum of frequencies**. This Fourier integral may be used to represent also non-periodic functions, such as a single voltage pulse or a flash of light, and it also represents contiuous spectra of frequencies, like whole ranges of tones or colours. In that respect it is often used for the description of a wide range of physical phenomena.
In addition, especially in Quantum Mechanics, it is used to connect the description of physical problems in position space with the corresponding description in momentum space.

Definition

To keep things short at this stage, let us begin by defining the Fourier integral. **Fourier's integral theorem** states that functions $f(x)$ enjoing the following properties

- $f(x)$ and $f'(x)$ are piecewise continuous in any finite interval, and

- the integral of $f(x)$ is absolutely convergent, i.e.

$$\int\limits_{-\infty}^{\infty} \mathrm{d}x |f(x)| = M\,, \quad \text{where } \; |M| < \infty \tag{3.1}$$

can be represented as

$$f(x) = \int\limits_{0}^{\infty} \mathrm{d}k \; \left[A(k)\cos(kx) + B(k)\sin(kx)\right] \,. \tag{3.2}$$

Here, the coefficent functions read

$$\begin{aligned} A(k) &= \frac{1}{\pi}\int\limits_{-\infty}^{\infty} \mathrm{d}x\, f(x)\cos(kx) \\ B(k) &= \frac{1}{\pi}\int\limits_{-\infty}^{\infty} \mathrm{d}x\, f(x)\sin(kx)\,. \end{aligned} \tag{3.3}$$

It must be stressed, however, that the representation above holds true only for points of continuity of $f(x)$. If, in contrast, $f(x)$ is discontinuous at a point x, then the representation above will yield

$$\frac{\lim\limits_{y\to x^+} f(y) + \lim\limits_{y\to x^-} f(y)}{2} = \int\limits_{0}^{\infty} \mathrm{d}k \; \left[A(k)\cos(kx) + B(k)\sin(kx)\right] \,. \tag{3.4}$$

The similarity of the Fourier integral defined here with the Fourier series considerred in the previous chapter is apparent - this is why more than often the term Fourier expansion is used for both the discrete series and for the integral presented here.

Equivalent forms & a simple motivation

There is a number of equivalent forms for the above Fourier integral, one of which will be discussed below. As a side-product, some further motivation for the Fourier integral and its specific form will emerge. It should be kept in mind, however, that the following is just a sketch of a more formal derivation.

The starting point of the reasoning is the Fourier series of a function $f(x)$,

$$f(x) = \frac{1}{2L}\int\limits_{-L}^{L} \mathrm{d}z\, f(z) \quad + \quad \frac{1}{L}\sum_{m=1}^{\infty} \cos\frac{m\pi x}{L}\left[\int\limits_{-L}^{L} \mathrm{d}z f(z)\cos\frac{m\pi z}{L}\right]$$
$$+ \quad \frac{1}{L}\sum_{m=1}^{\infty} \sin\frac{m\pi x}{L}\left[\int\limits_{-L}^{L} \mathrm{d}z f(z)\sin\frac{m\pi z}{L}\right]. \tag{3.5}$$

Combining two of the equations in Eq. (2.21), namely

$$\begin{aligned} \cos\vartheta + \cos\phi &= 2\cos(\vartheta+\phi)\cos(\vartheta-\phi) \\ \cos\vartheta - \cos\phi &= -2\sin(\vartheta+\phi)\sin(\vartheta-\phi) \end{aligned} \tag{3.6}$$

yields

$$\cos\alpha\cos\beta + \sin\alpha\sin\beta = \cos(\alpha-\beta) \tag{3.7}$$

and thus

$$f(x) = \frac{1}{2L}\int\limits_{-L}^{L} \mathrm{d}z\, f(z) \quad + \quad \frac{1}{L}\sum_{m=1}^{\infty}\left[\int\limits_{-L}^{L} \mathrm{d}z f(z)\cos\frac{m\pi(x-z)}{L}\right]. \tag{3.8}$$

Consider now the case, where $L \to \infty$. Setting

$$k = \frac{m\pi}{L} \quad \text{and} \quad \Delta k = \frac{\pi}{L} \tag{3.9}$$

and using

$$\cos\alpha = \frac{e^{i\alpha} + e^{-i\alpha}}{2} \tag{3.10}$$

yields

$$\begin{aligned} f(x) \quad = \quad & \frac{1}{2L}\int\limits_{-L}^{L} \mathrm{d}z\, f(z)\cos\frac{m\pi(x-z)}{L}\Big|_{m=0} \\ & + \frac{1}{2L}\sum_{m=1}^{\infty}\left[\int\limits_{-L}^{L} \mathrm{d}z f(z)\left(\cos\frac{im\pi(x-z)}{L} + \cos\frac{-im\pi(x-z)}{L}\right)\right], \end{aligned}$$

allowing to write

$$f(x) = \frac{1}{2L} \sum_{m=0}^{\infty} \left[\int_{-L}^{L} \mathrm{d}z f(z) \left(\exp \frac{im\pi(x-z)}{L} + \exp \frac{-im\pi(x-z)}{L} \right) \right] .$$

In other words, the sum stretches now from 0 rather than from 1 to infinity. Then, transforming the exponentials back to cosines,

$$\begin{aligned} f(x) \quad &\longrightarrow \quad \sum_{m=0}^{\infty} \frac{\Delta k}{\pi} \int_{-\infty}^{\infty} \mathrm{d}z f(z) \cos[k(x-z)] \\ &\xrightarrow{\Delta k \to 0} \quad \frac{1}{\pi} \int_{0}^{\infty} \mathrm{d}k \int_{-\infty}^{\infty} \mathrm{d}z f(z) \cos[k(x-z)] \, . \end{aligned} \tag{3.11}$$

Employing symmetry arguments, i.e. $\cos(-x) = \cos x$, this can be cast into

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}k \int_{-\infty}^{\infty} \mathrm{d}z f(z) \cos[k(x-z)] \, . \tag{3.12}$$

Using again the identity

$$\cos(\alpha - \beta) = \cos\alpha \cos\beta + \sin\alpha \sin\beta \tag{3.13}$$

yields the desired result.
It is straightforward to convince oneself that application of the same relations of sine and cosine function will yield the original form of the Fourier integral above.

Complex form

At this point it is worthwhile to notice that the integrand above, $f(z)\cos[k(x-z)]$ is an even function of k and, hence, yields a non-vanishing contribution when integrated over k. In contrast, $f(z)\sin[k(x-z)]$ is an odd function in k, yielding a zero result after integration over k. Therefore, there's no harm in adding, say, $f(z) i \sin[k(x-z)]$ to the integrand above. Thus,

$$\begin{aligned}
f(x) &= \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k \int\limits_{-\infty}^{\infty} \mathrm{d}z f(z) \cos[k(x-z)] \\
&= \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k \int\limits_{-\infty}^{\infty} \mathrm{d}z f(z) \left\{\cos[k(x-z)] + i \sin[k(x-z)]\right\} \\
&= \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k \int\limits_{-\infty}^{\infty} \mathrm{d}z f(z) \exp[ik(x-z)] = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k e^{ikx} \int\limits_{-\infty}^{\infty} \mathrm{d}z f(z) e^{-ikz} .
\end{aligned} \tag{3.14}$$

Thus, in its complex form, the Fourier transform can be written as

$$f(x) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k e^{ikx} \int\limits_{-\infty}^{\infty} \mathrm{d}z f(z) e^{-ikz} \equiv \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k e^{ikx} F(k) . \tag{3.15}$$

Due to its definition, the Fourier transform

$$F(k) = \int\limits_{-\infty}^{\infty} \mathrm{d}z f(z) e^{-ikz} \tag{3.16}$$

enjoys the following property

$$F(-k) = F^*(k) , \tag{3.17}$$

where the $*$ denotes complex conjugation, i.e. the operation $i \to -i$.

3.1.2 Detour: The δ-function

Let's go back to the derivation above, specifically, consider

$$\begin{aligned}
f(x) &= \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k \int\limits_{-\infty}^{\infty} \mathrm{d}z f(z) \exp[ik(x-z)] \\
&= \int\limits_{-\infty}^{\infty} \mathrm{d}z\, f(z) \left\{ \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k\, e^{ik(x-z)} \right\}
\end{aligned} \tag{3.18}$$

Apparently, the expression in the curly brackets depends on the difference $z-x$; morevore it obviously projects out the value of $f(z=x)$ in the remaining z integral. This is exactly the behaviour of Dirac's δ-function (which in fact for mathematicians is a distribution rather than a function). It has the property that

$$\int_{-\infty}^{\infty} \mathrm{d}z\, f(z)\delta(x-z) = f(x)\,, \tag{3.19}$$

if f is continuous at the point x. Anyway, written in this fashion, the way the δ-function behaves becomes more obvious: it acts under integration in the same way a Kronecker-δ does under summation.
To properly define the δ function:

$$\int_a^b \mathrm{d}x\, f(x)\delta(x-y) = \begin{cases} f(y) & \text{if } y \in [a,b] \\ 0 & \text{else}\,. \end{cases} \tag{3.20}$$

Representations

The form given above in the curly brackets serves as one way (among others) of representing this δ-function:

$$\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}k\, e^{ikx}\,. \tag{3.21}$$

It must be stressed, however, that anything said here about this function is valid only under an integration over x.
There's yet another way to look at it. To this end, consider a sequence of $\delta_n(z-x)$ defined through

$$\delta_n(z-x) = \frac{\sin[n(z-x)]}{\pi(z-x)} = \frac{1}{2\pi}\int_{-n}^{n} \mathrm{d}k\, e^{ik(z-x)}\,. \tag{3.22}$$

With some effort, it can be shown that

$$f(x) = \lim_{n\to\infty}\int_{-\infty}^{\infty} \mathrm{d}z\, f(z)\delta_n(z-x)\,. \tag{3.23}$$

Taking this as a starting point, inserting the definition of the δ_n from above, and reshuffling the order of integration yields

$$\begin{aligned} f(x) &= \frac{1}{2\pi} \lim_{n\to\infty} \int\limits_{-\infty}^{\infty} \mathrm{d}z\, f(z) \int\limits_{-n}^{n} \mathrm{d}k\, e^{ik(z-x)} \\ &= \frac{1}{2\pi} \lim_{n\to\infty} \int\limits_{-n}^{n} \mathrm{d}k \int\limits_{-\infty}^{\infty} \mathrm{d}z\, f(z) e^{ik(z-x)} = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k e^{-ikx} \int\limits_{-\infty}^{\infty} \mathrm{d}z\, f(z) e^{ikz}\,, \end{aligned} \tag{3.24}$$

exactly the form of the Fourier integral encountered in Eq. (3.14).

Important properties

The first property, concerns the integral over a product of a function and the derivative of the δ-function. Understanding the δ function with the help of the representation above, it is clear that partial integration can be applied, leading to

$$\int \mathrm{d}x f(x) \delta'(x-a) = -f'(a)\,. \tag{3.25}$$

If the argument of the δ-function itself is again a function, say $g(x)$, then

$$\int\limits_{-\infty}^{\infty} \mathrm{d}x\, f(x) \delta(g(x)) = \sum_{x_i} \left| \left[\frac{\mathrm{d}g(x)}{\mathrm{d}x} \right]_{x=x_i} \right|^{-1} f(x_i)\,, \tag{3.26}$$

where the x_i are the (simple) zeroes of $g(x)$: $g(x_i) = 0$. In particular,

$$\int\limits_{-\infty}^{\infty} \mathrm{d}x\, f(x) \delta(x^2 - a^2) = \frac{f(x-a) + f(x+a)}{2|a|}\,. \tag{3.27}$$

In more dimensions, the δ-function is just the product of δ-functions in the corresponding dimensions:

$$\delta(\vec{x}) = \delta(x)\delta(y)\delta(z)\,. \tag{3.28}$$

3.1.3 Fourier transforms

Fourier transforms and their inverse

In the previous section the Fourier integral has been discussed; in its complex form it has been given by

$$f(x) \;=\; \int\limits_{-\infty}^{\infty} \mathrm{d}k\, F(k) e^{ikx}\,, \tag{3.29}$$

where the coefficient function $F(k)$ has been defined through

$$F(k) \;=\; \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}x\, f(x) e^{-ikx}\,. \tag{3.30}$$

In order to pronounce the symmetry of the two forms, the normalisation factor $1/2\pi$ neccessary to go from arguments x to the arguments k and back, can be re-distributed. The weay how to that is purely conventional, any choice will do, as long as one sticks to it. A natural choice is to distribute it evenly between the two:

$$\begin{aligned} f(x) \;&=\; \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} \mathrm{d}k\, F(k) e^{ikx} \\ F(k) \;&=\; \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} \mathrm{d}x\, f(x) e^{-ikx}\,. \end{aligned} \tag{3.31}$$

This is actually the choice that will be used in the following.
In any case, the function $F(k)$ is known as the Fourier transform of $f(x)$, and, vice versa, $f(x)$ is the Fourier transform of $F(k)$.

A cool example

Consider the function

$$f(x) = \begin{cases} 1 & \text{for } |x| < a \\ 0 & \text{for } |x| > a\,, \end{cases} \tag{3.32}$$

where $a > 0$. Its Fourier transform is

$$
\begin{aligned}
F(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}x f(x) e^{-ikx} \\
&= \frac{1}{\sqrt{2\pi}} \int_{-a}^{a} \mathrm{d}x e^{-ikx} = \frac{1}{\sqrt{2\pi}} \left. \frac{e^{-ikx}}{-ik} \right|_{-a}^{a} \\
&= \frac{1}{\sqrt{2\pi}} \frac{2\sin(ka)}{k} \quad \text{if } k \neq 0\,. \qquad (3.33)
\end{aligned}
$$

For $k = 0$, $F(k) = 2a/\sqrt{2\pi}$ can be obtained by taking the limit and applying the rule of l'Hopital.

So far, so good. But how can this result be used for something more interesting? After all, the title of this paragraph is "A cool example". To see this, remember the inverse Fourier transform. It connects the integral

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}k F(k) e^{ikx} \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}k \frac{2\sin(ka)}{k} e^{ikx} = \begin{cases} 1 & \text{for } |x| < a \\ 1/2 & \text{for } |x| = a \\ 0 & \text{for } |x| > a\,, \end{cases} \qquad (3.34)
\end{aligned}
$$

with a finite value. The catch here is that there seems to be a pole at $k = 0$. To proceed, let us go back to the inverse Fourier transform and replace the exponential first with sines and cosines:

$$
\begin{aligned}
f(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}k \frac{2\sin(ka)}{k} e^{ikx} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}k \frac{2\sin(ka)}{k} \left[\cos(kx) + i\sin(kx)\right] \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}k \frac{2\sin(ka)\cos(kx)}{k} \qquad (3.35)
\end{aligned}
$$

because the integrand in the second term is an odd function of k. This implies that

$$
\int_{-\infty}^{\infty} \mathrm{d}k \frac{\sin(ka)\cos(kx)}{k} = \begin{cases} \pi & \text{for } |x| < a \\ \pi/2 & \text{for } |x| = a \\ 0 & \text{for } |x| > a\,. \end{cases} \qquad (3.36)
$$

Choosing now $x = a$, using

$$\sin(ka)\cos(ka) = \frac{1}{2}\left[\sin(ka - ka) + \sin(ka + ka)\right] = \frac{\sin(2ka)}{2} \tag{3.37}$$

and employing the symmetry of the integral implies that

$$\pi = \int_0^\infty \mathrm{d}k \frac{\sin(ka)\cos(ka)}{k} = \int_0^\infty \mathrm{d}k \frac{\sin(2ka)}{2k} \tag{3.38}$$

and hence

$$\frac{\pi}{2} = \int_0^\infty \mathrm{d}k \frac{\sin(ka)}{k}\,. \tag{3.39}$$

So far, $a > 0$ has been assumed. Similar reasoning shows that for $a < 0$, the value of the integral merely changes sign and therefore

$$\int_0^\infty \mathrm{d}k \frac{\sin(ka)}{k} = \left\{ \begin{array}{ll} +\frac{\pi}{2} & \text{for } a > 0 \\ -\frac{\pi}{2} & \text{for } a < 0 \end{array} \right\} = \frac{\pi}{2}[\Theta(a) - \Theta(-a)]\,. \tag{3.40}$$

Fourier Sine and Cosine transforms

Similar to the case of Fourier series, it is tempting to use the symmetry properties of the function $f(x)$ in order to alleviate the Fourier transformation; the symmetry in question is, of course, the parity of the function.
For even or odd functions $f^{\pm}(x)$ (i.e. functions with positive or negative parity), it is sufficient to consider **Fourier Cosine** or **Fourier Sine Transforms**, respectively. They are given by

$$\begin{aligned} f^+(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}k\, F^+(k)\cos(kx) = \frac{2}{\sqrt{2\pi}} \int_0^\infty \mathrm{d}k\, F^+(k)\cos(kx) \\ F^+(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}k\, f^+(x)\cos(kx) = \frac{2}{\sqrt{2\pi}} \int_0^\infty \mathrm{d}k\, f^+(x)\cos(kx) \end{aligned} \tag{3.41}$$

and

$$\begin{aligned}
f^-(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}k\, F^-(k)\sin(kx) = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} \mathrm{d}k\, F^-(k)\sin(kx) \\
F^-(k) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{d}k\, f^-(x)\sin(kx) = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} \mathrm{d}k\, f^-(x)\sin(kx)
\end{aligned} \tag{3.42}$$

However, there are cases, discussed below, where Fourier Sine and Cosine Transforms may be sensible independent of the parity of the function.

Example:

- The Fourier Cosine Transform of

$$f(x) = e^{-mx}\,, \quad \text{with } m > 0 \tag{3.43}$$

is given by

$$\begin{aligned}
F_c(k) = F^+(k) &= \sqrt{\frac{2}{\pi}} \int_0^{\infty} \mathrm{d}x e^{-mx}\cos(kx) \\
&= \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \mathrm{d}x e^{-mx}[e^{ikx} + e^{-ikx}] \\
&= -\frac{1}{\sqrt{2\pi}} \left[\frac{e^{-(m-ik)x}}{m-ik} + \frac{e^{-(m+ik)x}}{m+ik}\right]_0^{\infty} = \sqrt{\frac{2}{\pi}}\frac{m}{m^2+k^2}\,.
\end{aligned} \tag{3.44}$$

This can be used to show that

$$\int_0^{\infty} \mathrm{d}k \frac{\cos(kx)}{m^2+k^2} = \sqrt{\frac{\pi}{2}}\frac{1}{m}\sqrt{\frac{2}{\pi}} \int_0^{\infty} \mathrm{d}k \frac{m\cos(kx)}{m^2+k^2} = \frac{\pi}{2m} e^{-mx}\,, \tag{3.45}$$

where in the last transformation the inverse Fourier Transformation has been used.

With similar tricks a large number of integrals involving trigonometric functions can be solved.

Parseval's identity, once again

In Sec. 2.4.1 Parseval's identity for Fourier series has been discussed. In this section, an analogous expression for Fourier transforms will be presented: If $F(k)$ and $G(k)$ are Fourier transforms of $f(x)$ and $g(x)$, respectively, then

$$\begin{aligned}
&\int\limits_{-\infty}^{\infty} \mathrm{d}x f(x) g(x) \\
&= \int\limits_{-\infty}^{\infty} \mathrm{d}x \left\{ \left[\frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} \mathrm{d}k F(k) e^{ikx} \right] \left[\frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} \mathrm{d}k' G(k') e^{ik'x} \right] \right\} \\
&= \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k \mathrm{d}k' F(k) G(k') \int\limits_{-\infty}^{\infty} \mathrm{d}x e^{i(k+k')x} .
\end{aligned}$$

At this point, it is worthwhile to remember the representations of the δ-function to yield

$$\begin{aligned}
\int\limits_{-\infty}^{\infty} \mathrm{d}x f(x) g(x) &= \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k \mathrm{d}k' F(k) G(k') \delta(k+k') \\
&= \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k F(k) G(-k) = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k F(k) G^*(k) ,
\end{aligned} \tag{3.46}$$

where the $*$ denotes the complex conjugate, obtained by replacing i with $-i$. In particular, if $f(x) = g(x)$ is a real function, then

$$\int\limits_{-\infty}^{\infty} \mathrm{d}x |f(x)|^2 = \frac{1}{2\pi} \int\limits_{-\infty}^{\infty} \mathrm{d}k |F(k)|^2 . \tag{3.47}$$

Convolution theorem

In a similar fashion, the convolution theorem can be proved. A convolution of two functions, $(f * g)(x)$ is defined through

$$(f * g)(x) \equiv \int_{-\infty}^{\infty} \mathrm{d}y f(y) g(x-y)\,. \tag{3.48}$$

The Fourier Transform of such a convolution reads

$$(F * G)(k) = (2\pi) F(k) G(k)\,. \tag{3.49}$$

To see this consider

$$\begin{aligned}(2\pi)F(k)G(k) &= \frac{2\pi}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}y e^{iky} f(y) \int_{-\infty}^{\infty} \mathrm{d}z e^{ikz} g(z) \\ &= \int_{-\infty}^{\infty} \mathrm{d}y \mathrm{d}z e^{ik(y+z)} f(y) g(z)\end{aligned} \tag{3.50}$$

and transform from y, z to $x = y + z$, y. Then

$$\begin{aligned}(2\pi)F(k)G(k) &= \int_{-\infty}^{\infty} \mathrm{d}x \mathrm{d}y e^{ikx} f(y) g(x-y) \\ &= \int_{-\infty}^{\infty} \mathrm{d}x e^{ikx} \left[\int_{-\infty}^{\infty} \mathrm{d}y f(y) g(x-y) \right] \\ &= \int_{-\infty}^{\infty} \mathrm{d}x e^{ikx} (f * g)(x) = (F * G)(k)\,.\end{aligned} \tag{3.51}$$

In other words, up to a factor of 2π the Fourier Transform of a convolution of two functions equals the product of their Fourier Transforms taken individually,

$$(F * G)(k) = 2\pi F(k) G(k)\,. \tag{3.52}$$

3.2 Applications

3.2.1 Finite wave train

An important application of Fourier transforms is the resolution of a finite pulse into a continous spectrum of sinusoidal waves. To highlight this, imagine an infinite wave train with frequency ω_0 is cut such that

$$f(t) = \begin{cases} \sin(\omega_0 t) & \text{for } |t| < \dfrac{n\pi}{\omega_0} \\ 0 & \text{else}\,. \end{cases} \tag{3.53}$$

This corresponds to cutting out n cyxcles of the original wave train.
Since, obviously, $f(t)$ is odd, it is sufficient to analyse its Fourier sine transform, given by

$$\begin{aligned} F^-(\omega) &= \sqrt{\frac{2}{\pi}} \int\limits_0^\infty \mathrm{d}t \sin(\omega t) f(t) \\ &= \sqrt{\frac{2}{\pi}} \int\limits_0^{n\pi/\omega_0} \mathrm{d}t \sin(\omega t)\sin(\omega_0 t) \\ &= \frac{1}{\sqrt{2\pi}} \int\limits_0^{n\pi/\omega_0} \mathrm{d}t \left[\cos(\omega-\omega_0)t - \cos(\omega+\omega_0)t\right] \\ &= \frac{1}{\sqrt{2\pi}} \left[\frac{\sin \dfrac{n\pi(\omega-\omega_0)}{\omega_0}}{\omega-\omega_0} - \frac{\sin \dfrac{n\pi(\omega+\omega_0)}{\omega_0}}{\omega+\omega_0} \right]. \end{aligned} \tag{3.54}$$

It is quite interesting, how this depends on frequency: For $\omega = \omega_0 + \delta$ and δ small compared to ω and ω_0, the first term dominates. Then

$$F^-(\omega) \approx \frac{1}{\sqrt{2\pi}\delta} \sin \frac{n\pi\delta}{\omega}\,, \tag{3.55}$$

yielding zeroes for

$$\delta = \omega - \omega_0 = \pm\frac{\omega}{n},\ \pm\frac{2\omega}{n},\ \ldots\,. \tag{3.56}$$

In fact, this corresponds very well to the amplitude curve for the single slit diffraction pattern, exhibiting the same zeroes. It is worth noting that the maximum is given for $\delta = 0$, i.e. for a resonance. In this case,

$$F^-_{\rm max}(\omega) = \frac{n\pi}{\sqrt{2\pi}\omega}\,. \tag{3.57}$$

In any case, it is tempting to take the dfistance to the first zero as a measure for the spread of the wave pulse. It is given by

$$\Delta\omega = \frac{\omega}{n} \tag{3.58}$$

and is small for large trains and large for small trains.

3.2.2 Infinite string

Sketching the physical problem

Consider an infinitely long string, stretched along the x-axis. Displacements, say in y-direction at one point of this string will then propagate along it, such that they satisfy a wave equation with

$$\frac{1}{\alpha^2}\frac{\partial^2 y}{\partial t^2} = \frac{\partial^2 y}{\partial x^2}\,. \tag{3.59}$$

Here, α plays the role of an inverse wave velocity.
A simple way to initiate a wave on this string is to just displace it at some time $t = 0$ according to a function $y(x,\,0) = f(x)$ with $\partial y(x,\,0)/\partial t = 0$. It makes perfect sense to assume that this initial displacement is of only limited amplitude, i.e. $|y(x,\,t) < M$.

General solution

To find a general solution, let's again use a separation ansatz,

$$y(x,\,t) = X(x)T(t) = \left[A^+(k)e^{ikx} + A^-(k)e^{-ikx}\right]\left[B^+(\omega)e^{i\omega t} + B^-(\omega)e^{-i\omega t}\right] \tag{3.60}$$

where, in principle the coefficients $A^\pm$ and $B^\pm$ are functions of k and ω, respectively. However, plugging this into the differential equation above yields

$$\frac{\omega^2}{\alpha^2} = k^2\,. \tag{3.61}$$

In other words,

$$\frac{1}{\alpha} = \frac{k}{\omega}\,. \tag{3.62}$$

Since k and ω are connected by a mere constant, the dependence on the former can easily be replaced by a dependence on the latter (or vice versa).

Including boundary conditions

The boundary conditions read

$$\begin{aligned} v(x,\,0) = f(x) &= \left[A^+(k)e^{ikx} + A^-(k)e^{-ikx}\right]\left[B^+(k) + B^-(k)\right] \\ v_t(x,\,0) = 0 &= \left[A^+(k)e^{ikx} + A^-(k)e^{-ikx}\right] i\omega \left[B^+(k) - B^-(k)\right] \end{aligned} \tag{3.63}$$

yielding

$$B^+(k) = B^-(k) = B(k) \tag{3.64}$$

and, renaming $A^\pm(k) \to A^\pm(k) = B(k)A^\pm(k)$ yields

$$f(x) = A^+(k)e^{ikx} + A^-(k)e^{-ikx} = X(x) \tag{3.65}$$

and something like

$$y(x,\,t) = f(x)\cos(k\alpha t)\,. \tag{3.66}$$

This already looks very nice, but there's even more that can be gained. In order to do so, let us Fourier expand $f(x)$:

$$f(x) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} \mathrm{d}k \left[C(k)\cos(kx) + S(k)\sin(kx)\right] \tag{3.67}$$

with

$$\begin{aligned} C(k) &= \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} \mathrm{d}x f(x)\cos(kx) \\ S(k) &= \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} \mathrm{d}x f(x)\sin(kx)\,. \end{aligned} \tag{3.68}$$

Putting everything together yields

$$y(x,\,t) \;=\; \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\mathrm{d}k\int\limits_{-\infty}^{\infty}\mathrm{d}x'\,f(x')\left[\cos(kx)\cos(kx')+\sin(kx)\sin(kx')\right]\cos(k\alpha t)\,. \tag{3.69}$$

Employing, once more,

$$\cos\alpha\cos\beta+\sin\alpha\sin\beta=\cos(\alpha-\beta)\,, \tag{3.70}$$

this yields

$$\begin{aligned} y(x,\,t) &= \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\mathrm{d}k\int\limits_{-\infty}^{\infty}\mathrm{d}x'\,f(x')\cos[k(x-x')]\cos(k\alpha t) \\ &= \frac{1}{4\pi}\int\limits_{-\infty}^{\infty}\mathrm{d}k\int\limits_{-\infty}^{\infty}\mathrm{d}x'\,f(x')\left\{\cos[k(x-x'+\alpha t)]+\cos[k(x-x'-\alpha t)]\right\}\,, \end{aligned} \tag{3.71}$$

where in the last step

$$\cos\alpha\cos\beta=\frac{1}{2}[\cos(\alpha-\beta)+\cos(\alpha+\beta)] \tag{3.72}$$

has been used. Changing the order of integration results in

$$\begin{aligned} y(x,\,t) \;=\; & \frac{1}{4\pi}\int\limits_{-\infty}^{\infty}\mathrm{d}x'\int\limits_{-\infty}^{\infty}\mathrm{d}k\,f(x')\cos[k(x-x'+\alpha t)] \\ & +\frac{1}{4\pi}\int\limits_{-\infty}^{\infty}\mathrm{d}x'\int\limits_{-\infty}^{\infty}\mathrm{d}k\,f(x')\cos[k(x-x'-\alpha t)] \\ =\; & \frac{1}{2}f(x+\alpha t)+\frac{1}{2}f(x-\alpha t)\,. \end{aligned} \tag{3.73}$$

This last line shows that, if the string is distorted for instance by a δ-like "kick", this kick will propagate with half the amplitude to the left and the right.

3.3 Green functions

Green functions are extremely useful in the solution of partial differential equations, but of course they can also be used in order to solve ordinary differential equations.

3.3.1 Green function for the driven harmonic oscillator

What is the Green function?

This will be discussed employing the example of a driven harmonic oscillator. Ignoring damping, the driven harmonic oscialltor is described by

$$y'' + \omega^2 y = f(t)\,. \tag{3.74}$$

Let us discuss the case where the oscillator initially is at rest, i.e.

$$y(0) = y'(0) = 0\,. \tag{3.75}$$

Since the driving force $f(t)$ can be rewritten as a sequence of "kicks" through

$$f(t) = \int_0^\infty \mathrm{d}t' f(t')\delta(t-t')\,, \tag{3.76}$$

it is sufficient to construct the most general solution for

$$y'' + \omega^2 y = \delta(t-t') \tag{3.77}$$

instead. Denoting this solution by $G(t,\,t')$, i.e.

$$\frac{\mathrm{d}^2 G(t,\,t')}{\mathrm{d}t^2} + \omega^2 G(t,\,t') = \left(\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega^2\right) G(t,\,t') = \delta(t-t') \tag{3.78}$$

the true solution $y(t)$ can be obtained by "adding up the kicks",

$$y(t) = \int_0^\infty \mathrm{d}t' G(t,\,t') f(t')\,. \tag{3.79}$$

In order to see this, insert this form of $y(t)$ into the original differential equation:

$$y''(t) + \omega^2 y(t) = \left(\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega^2\right) \int\limits_0^\infty \mathrm{d}t' G(t,\, t') f(t')$$
$$= \int\limits_0^\infty \mathrm{d}t' \left(\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega^2\right) G(t,\, t') f(t') = \int\limits_0^\infty \mathrm{d}t' \delta(t - t') f(t') = f(t)\,. \tag{3.80}$$

The function $G(t,\, t')$ is called the **Green function** of the problem. It is the **response of the system to a unit impulse at** $t = t'$.
It is quite simple to show that

$$\begin{aligned} G(t,\, t') = G(t - t') &= \begin{cases} 0 & \text{for } 0 < t < t' \\ \dfrac{\sin[\omega(t - t')]}{\omega} & \text{for } t' < t \end{cases} \\ &= \frac{\sin[\omega(t - t')]}{\omega} \Theta(t - t') \end{aligned} \tag{3.81}$$

is a solution of

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega^2\right) G(t,\, t') = \delta(t - t')\,. \tag{3.82}$$

For $0 < t < t'$

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega^2\right) G(t,\, t') = 0\,, \tag{3.83}$$

because in iths case, $G(t - t') = 0$, whereas for $t' < t$

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}t^2} + \omega^2\right) G(t,\, t') = \left(-\omega^2 + \omega^2\right) G(t,\, t') = 0\,, \tag{3.84}$$

as demanded.

Explicit construction of the Green function

But how can such a solution be constructed? In order to do so, let us Fourier transform the full equation. As has been seen in Sec. 3.1.2, a representation of the δ-function is given by the exponential

$$\delta(t-t') = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}E e^{iE(t-t')} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}E e^{iE(t-t')}\Delta(E) \tag{3.85}$$

such that its Fourier transform $\Delta(E)$ is readily identified with $1/\sqrt{2\pi}$. For the l.h.s. of Eq. (3.78), it is useful to realise first that it must be a function of $t-t'$, since also the r.h.s. of this equation is a function of $t-t'$. Then, Fourier transforming $G(t,\,t') = G(t-t')$ yields

$$G(t-t') \;=\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}E e^{iE(t-t')}G(E)\,. \tag{3.86}$$

Plugging this into the equation above results in

$$\left(-E^2+\omega^2\right)G(E) = 1/\sqrt{2\pi} \tag{3.87}$$

and thus

$$G(E) = -\frac{1}{\sqrt{2\pi}}\frac{1}{E^2-\omega^2}\,. \tag{3.88}$$

This now needs to be inserted into the equation for $G(t-t')$:

$$\begin{aligned}
G(t-t') &= -\frac{1}{(\sqrt{2\pi})^2}\int_{-\infty}^{\infty} \mathrm{d}E\,\frac{e^{iE(t-t')}}{E^2-\omega^2}\\
&= -\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}E\,\frac{1}{2\omega}\left[\frac{e^{iE(t-t')}}{E-\omega}-\frac{e^{iE(t-t')}}{E+\omega}\right]\\
&= -\frac{1}{4\pi\omega}\int_{-\infty}^{\infty} \mathrm{d}y\,\left[\frac{e^{iy(t-t')}e^{i\omega(t-t')}}{y}-\frac{e^{iy(t-t')}e^{-i\omega(t-t')}}{y}\right]\\
&= -\frac{2i\sin[\omega(t-t')]}{4\pi\omega}\int_{-\infty}^{\infty} \mathrm{d}y\,\frac{e^{iy(t-t')}}{y}\,.
\end{aligned} \tag{3.89}$$

The remaining integral is, formally speaking, ill-defined. For our purpose, however, it suffices to realise that the integral of the δ-function is the Heavyside- or Θ-function[1],

$$\Theta(x) = \int \mathrm{d}x \delta(x) \,. \tag{3.94}$$

Adding in one of the representations of the δ-function, namely

$$\delta(t - t') = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}y \, e^{iy(t-t')} \tag{3.95}$$

immediately shows that

$$\begin{aligned} i \int_{-\infty}^{\infty} \mathrm{d}y \, \frac{e^{iy(t-t')}}{iy} &= i \int_{-\infty}^{\infty} \mathrm{d}y \int \mathrm{d}(t-t') e^{iy(t-t')} \\ &= 2\pi i \int \mathrm{d}(t-t')\delta(t-t') = 2\pi i \Theta(t-t') \,, \end{aligned} \tag{3.96}$$

---

[1] Alternatively, it can be done by help of residues,

$$\int_{-\infty}^{\infty} \mathrm{d}y \, \frac{e^{iy(t-t')}}{y} = 2\pi i \sum_{\nu=1}^{k} \mathrm{Res} \left[\frac{e^{iy(t-t')}}{y}; \, y_\nu \right] , \tag{3.90}$$

where the sum stretches over all zeroes of the function (i.e. over the zero at $y = 0$), and the residuum is given by

$$\mathrm{Res} \left[f(x); \, x_\nu \right] = \frac{1}{(m-1)!} \lim_{x \to x_\nu} \frac{\mathrm{d}^{m-1} [(x - x_\nu)^m f(x)]}{\mathrm{d}x^{m-1}} \,. \tag{3.91}$$

In the case considered here,

$$\mathrm{Res} \left[\frac{e^{iy(t-t')}}{y}; \, 0 \right] = \frac{1}{0!} \lim_{y \to 0} \frac{\mathrm{d}^0 [(y-0)^1 \frac{e^{iy(t-t')}}{y}]}{\mathrm{d}x^0} = \lim_{y \to 0} e^{iy(t-t')} = 1 \,. \tag{3.92}$$

Then,

$$\int_{-\infty}^{\infty} \mathrm{d}y \, \frac{e^{iy(t-t')}}{y} = 2\pi i \tag{3.93}$$

and thus

$$G(t-t') = -\frac{2i\sin[\omega(t-t')]}{4\pi\omega}2\pi i\Theta(t-t') = \frac{\sin[\omega(t-t')]}{\omega}\Theta(t-t')\,, \tag{3.97}$$

as advertised.

Using the Green function

Let us go back to Eq. (3.79),

$$y(t) = \int\limits_0^\infty \mathrm{d}t' G(t,\,t')f(t') = \int\limits_0^t \mathrm{d}t' \frac{\sin[\omega(t-t')]}{\omega}f(t') \tag{3.98}$$

after the form of the Green function of the driven harmonic oscillator has been inserted, cf. Eq. (3.81).
This enables us to immedaitely write the solution for the driving force

$$f(t) = \sin(\omega_0 t) \quad \text{for } t > 0\,, \tag{3.99}$$

where again the boundary conditions $y(0) = y'(0) = 0$ have been employed. It reads

$$\begin{aligned}
y(t) &= \int\limits_0^t \mathrm{d}t' \frac{\sin[\omega(t-t')]}{\omega}\sin(\omega_0 t') \\
&= \frac{1}{2\omega}\int\limits_0^t \mathrm{d}t' \left\{\cos[\omega(t-t')-\omega_0 t'] - \cos[\omega(t-t')+\omega_0 t']\right\} \\
&= \frac{1}{2\omega}\left[-\frac{\sin[\omega t-(\omega+\omega_0)t']}{\omega+\omega_0} + \frac{\sin[\omega t-(\omega-\omega_0)t']}{\omega-\omega_0}\right]_0^t \\
&= \frac{1}{2\omega}\left[\frac{\sin(\omega_0 t)+\sin(\omega t)}{\omega+\omega_0} + \frac{\sin(\omega_0 t)-\sin(\omega_t)]}{\omega-\omega_0}\right] \\
&= \frac{1}{\omega}\left[\frac{\sin\frac{(\omega+\omega_0)t}{2}\cos\frac{(\omega-\omega_0)t}{2}}{\omega+\omega_0} - \frac{\sin\frac{(\omega-\omega_0)t}{2}\cos\frac{(\omega+\omega_0)t}{2}}{\omega-\omega_0}\right]
\end{aligned} \tag{3.100}$$

It is tedious but simple to prove that this is indeed a solution to the problem.