

# Week ①

PHY 303 Quantum Mechanics

*Instructor: Sebastian Wüster, IISER Bhopal, 2021*

*These notes are provided for the students of the class above only. There is no guarantee for correctness, please contact me if you spot a mistake.*

## 1.5 Mathematical Foundations of Quantum mechanics

In the previous week, we got reminded that quantum physics involves probability theory and wavefunctions in a central way. Let us thus briefly review (or provide) the essential math to deal with both of these. For this we have preponed some of Griffith chapter 3, particularly 3.2. You can find many other QM books that do all the math at the beginning, we will still leave most of it for our chapter 3 later, as Griffith does.

### 1.5.1 Complex numbers

*I shall assume you know complex numbers from math courses. The following are the absolute bare essentials provided for sake of completeness and to define notation. If we later use anything related to complex numbers that you had not seen before, please let me know. I might then include those points in this section. See also PHy106, week 8.*

Quantum mechanics involves complex numbers in an essential way. A complex number can be written as  $z = x + iy$ , where the imaginary unit  $i$  is defined via  $i = \sqrt{-1}$ . We then call  $x = \Re[z]$  the real part and  $y = \Im[z]$  the imaginary part of  $z$ . Since it has two components, we can draw each complex number as a 2D vector  $(x, y)$  in a 2D space called the complex plane. That also makes apparent that we can write it in polar notation  $z = re^{i\varphi}$ , where  $0 \leq r$  is a real amplitude and  $0 \leq \varphi \leq 2\pi$  a real phase.

Each complex number has what is called its complex conjugate denoted by  $z^*$  which is  $z^* = x - iy$  (i.e. we just flip the sign of the imaginary part). With this we can find the modulus of a complex number as  $|z| = \sqrt{z^*z} = \sqrt{x^2 + y^2}$ , just as would be the case for a 2D vector.

A very useful equation is Euler's equation:

$$e^{iz} = e^{-y+ix} = e^{-y}[\cos(x) + i \sin(x)]. \quad (1.4)$$

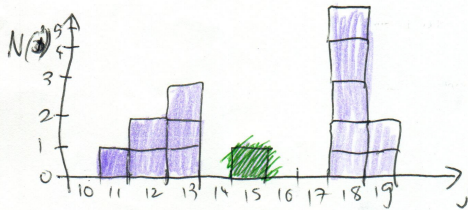
### 1.5.2 Probability theory

Here we only establish the most important concepts of probability theory, to have everyone on the same page.

**Probabilities** Consider a random variable  $j$  with discrete/integer outcomes. Suppose we draw from a set of  $N$  elements, where outcome  $j$  is contained  $N(j)$  times. Clearly  $N = \sum_j N(j)$ . The probability of outcome  $j$  is then  $P(j) = N(j)/N$ . All the values  $P(j)$  together are called a probability distribution. The average (mean) outcome is

$$\langle j \rangle = \sum_j P(j)j. \quad (1.5)$$

**Example 2, Age of computer players:** (discrete) Consider the following set of players of some online computer game, with ages: 11, 12 ( $\times 2$ ), 13 ( $\times 3$ ), 15, 18 ( $\times 5$ ), 19 ( $\times 2$ ). For this set, we can draw the histogram below.



**left:** Histogram of the number of occurrences  $N(j)$  of event  $j$  (here  $j = \text{age}$ ), within a sample of 13 players. [Ignore the green box at 15]

Using  $P(j) = N(j)/N$  we find e.g. the probability that a randomly chosen player has age 13 to be  $P(13) = 3/13 = 30\%$ . The mean age turns out to be 14.4, it gives an indication of the centre of the probability distribution. The mean is not to be confused with the most likely outcome, which is here 18. In fact, there is no outcome possible near the mean value (for 14,15) at all! Another concept would be the median which is the value  $j$  for which the probability of a higher result equals that of a lower result<sup>a</sup>

<sup>a</sup>This is not used as much in quantum mechanics.

Above we dealt with random variables that take integer values. For continuous outcomes  $x$  we define

**Probability densities** Consider a random variable  $x$  with continuous outcomes. We describe those with a probability density  $\rho(x)$  such that the probability for the outcome to lie between  $x$  and  $x + dx$  is given by  $\rho(x)dx$ . The probability for the result to lie between  $a$  and  $b$  then is  $P_{ab} = \int_a^b \rho(x)dx$ .

The

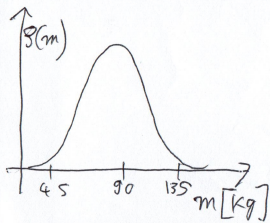
**Mean** of the probability density is

$$\langle x \rangle = \int_{-\infty}^{\infty} x\rho(x)dx \quad (1.6)$$

- In quantum mechanics, we usually refer to the mean as expectation value.

- Clearly the probability to have an arbitrary (unspecified) outcome has to be 1. This implies  $\sum_j P(j) = 1$  (discrete) or  $\int_{-\infty}^{\infty} \rho(x) dx = 1$  (continuous).
- From the definition  $P(j) = N(j)/N$  it is obvious that a probability has to be a positive real number (since the same is true for the “number of some items”). Saying the probability for an event is -5% makes as little sense as saying “I own -5 cars”. We state probabilities are positive and real here so clearly, since it provides a very important sanity check for many quantum mechanical calculations, that DO involve negative and complex numbers in between (but not in the final result if the latter is a probability).

**Example 3, Weight of computer players:** (continuous) Weight can take any value, not just integers.



**left:** An exemplary continuous probability distribution for the weight of an average human is shown on the left (possibly unrealistic).

Using e.g. the formula  $P_{ab} = \int_a^b \rho(x) dx$ , we can infer information such as “the probability for the weight to be  $> 88$  kg is 50%”.

- Comparing example 2 and example 3: Physically, of course the age was also a continuous quantity. We had referred to the custom of giving an integer age for a person, rounding down.
- Be aware, that the figures are showing two quantities of slightly different character, example 2 shows a histogram for the number of occurrences within a finite sample size of  $N = 13$ , while example 3 discusses a probability distribution that does not refer to a specific sample size.

An important property of a histogram, probability distribution or probability density is how “scattered it is”. For example the set of outcomes  $[3,50,123,555]$  varies very widely (large scatter), while outcomes  $[122,123,123,124]$  are much less scattered. To quantify this we use the following

### Measures for the width of a probability distribution

The variance  $\sigma^2$  of a probability distribution is

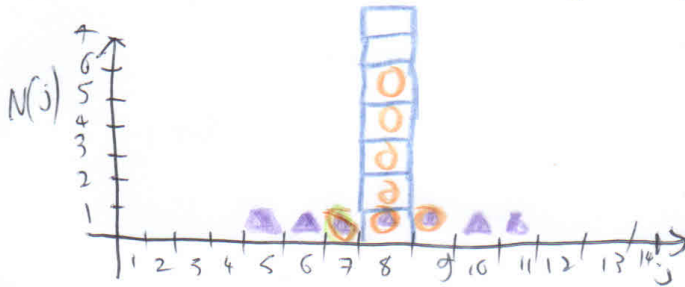
$$\sigma^2 = \text{Var}[P] = \langle (j - \langle j \rangle)^2 \rangle = \sum_j P(j)(j - \langle j \rangle)^2, \quad \sigma^2 = \int dx \rho(x)(x - \langle x \rangle)^2 \quad (1.7)$$

in the discrete and continuous cases respectively. From these, we define the standard deviation

$$\sigma = \sqrt{\text{Var}} = \sqrt{\langle (j - \langle j \rangle)^2 \rangle}, \quad \sigma = \sqrt{\int dx \rho(x)(x - \langle x \rangle)^2} \quad (1.8)$$

- The variance adds up the squares of the deviation  $j - \langle j \rangle$  from the mean.
- Exercise: One can show that  $\sigma^2 = \langle j^2 \rangle - \langle j \rangle^2$  (for both discrete and continuous). This formula is usually easier to use than the direct definition.
- The standard deviation gives a measure of the width of a probability density. Not to be confused with the standard error  $\sigma/\sqrt{N}$ , which gives the error of the mean  $\langle j \rangle$ .

**Example 4, Varying width of probability distribution:** Below are three age distributions such as in example 2,

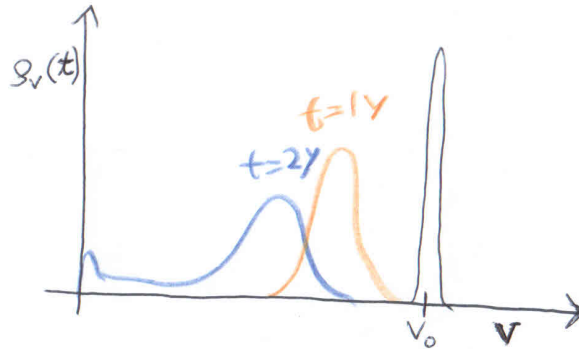


**left:** drawn with symbols (i) blue squares (ii) red circles (iii) violet triangles.

If you would calculate the mean according to Eq. (1.5), you'd find the same answer ( $\langle j \rangle = 8$ ) for each. However the standard deviations using (1.8) are (i)  $\sigma = 0$  (ii)  $\sigma = 1/\sqrt{3}$  (iii)  $\sigma = \sqrt{14/3} \approx 2$ . You can see how the standard deviation gets larger for wider distributions. It roughly correspond to the (half) width of the distribution.

*Time dependence of probability distributions:* Nature offers many examples of probability distributions that (may) depend on time, we will see many of those in quantum mechanics. An example in a totally different context is given below:

**Example 5, Value degradation of a car:** Cars quickly lose value on the second hand market after purchasing them. The figure below shows the value distribution of a certain model at three different times.



**left:** Value  $v$  of a car model at the moment of purchase (black), after one year (brown) and after two years (blue).

initially there is only a small spread due to different vendor margins. After one years, cars loose value since they age, but also the width of the distribution increases as cars experiences different mishaps such as scratches.

### 1.5.3 Vector spaces and matrices

For handling wavefunctions in quantum mechanics, it is very useful to realize that these wavefunctions can be viewed as vectors, which we shall explain in section 1.5.4. You will be familiar with the concept of a vector space from linear algebra, most often dealing with 3-component (3D) vectors  $\mathbf{v} = [v_x, v_y, v_z]^T$  that can label a point in 3D space, let us first briefly review those concepts in the present section.

**Vector space:** A vector space  $V$  over a field<sup>a</sup>  $F$  is a set of objects with an addition operation “+” (of two vectors  $\mathbf{v} \in V$ ) and a multiplication operation “ $\times$ ” (of an element in the field  $f \in F$  and a vector). The operations have to fulfill the following axioms (for  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$  and  $a, b \in F$ )

- (i)  $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$  (associativity of +)
- (ii)  $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$  (commutativity of +)
- (iii)  $\exists \mathbf{0}$  such that  $\mathbf{v} + \mathbf{0} = \mathbf{v} \forall \mathbf{v}$  (identity element of +)
- (iv)  $\exists -\mathbf{v}$  such that  $\mathbf{v} + (-\mathbf{v}) = \mathbf{0} \forall \mathbf{v}$  (inverse element for +)
- (v)  $a(\mathbf{u} + \mathbf{v}) = a\mathbf{u} + a\mathbf{v}$ , and  $(a + b)\mathbf{u} = a\mathbf{u} + b\mathbf{u}$  (distributivity of scalar multiplication)
- (vi)  $1\mathbf{v} = \mathbf{v}$  (where 1 denotes the multiplicative identity in the field)
- (vii)  $a(b\mathbf{v}) = (ab)\mathbf{v}$  (compatibility of scalar multiplication with field multiplication)

We have used the usual notation not to write the  $\times$  symbol explicitly.

<sup>a</sup>e.g. type of numbers, such as real or complex numbers

- The example you probably know best is  $\mathbb{R}^3$ , the set of three-component real column vectors  $\mathbf{v} = [v_x, v_y, v_z]^T$ , that can define a point in 3D space<sup>3</sup>. Here we use the operation “+” as  $\mathbf{v} + \mathbf{u} \equiv [v_x + u_x, v_y + u_y, v_z + u_z]^T$  and  $a \times \mathbf{v} \equiv [au_x, au_y, au_z]^T$ . You can verify that these objects and operations fulfill all the axioms in the definition above.
- Note that the complex numbers also form a field, so the first generalisation we will do in the following is to allow  $a, b$  and the column entries of the vectors to be complex.
- An important concept in a vector space is that of orthogonality, which we define via the scalar product

$$\mathbf{v}^* \cdot \mathbf{u} = v_x^* u_x + v_y^* u_y + v_z^* u_z, \quad (1.9)$$

calling vectors orthogonal when  $\mathbf{v}^* \cdot \mathbf{u} = 0$ .

- Another important concept is the norm of a vector, which we define as  $\|\mathbf{v}\| = |\mathbf{v}| = \sqrt{|\mathbf{v}^* \cdot \mathbf{v}|}$ .
- The last two dotpoints can also be defined via a set of axioms (which we defer to section 3.2). The advantage of all these axioms, is that they then allow the generalisation of these concepts to any objects fulfilling them. We want to use this very shortly, to apply them all to wavefunctions.

Every vector space  $V$  has a

<sup>3</sup>Here  $T$  indicates the “transpose”. I.e.  $[v_x, v_y, v_z]^T = \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix}$

**Basis:** as a set  $\mathcal{B}$  of  $d$  linearly independent vectors  $\mathcal{B} = \{\mathbf{b}_n\}$  that spans the space, which means we can write every  $\mathbf{v} \in V$  as

$$\mathbf{v} = \sum_{n=1}^{d-1} v_n \mathbf{b}_n, \quad (1.10)$$

for some suitable coefficients  $v_n \in F$ .

- Each vector space can have many different bases.
- We call the number of elements  $d$  in the basis  $\{\mathbf{b}_n\}$  the dimension of the vector space.
- Special bases are orthonormal bases, that fulfill  $\|\mathbf{b}_n\| = 1$  and  $\mathbf{b}_n^* \cdot \mathbf{b}_m = \delta_{nm}$ . We can then find the coefficients for Eq. (1.10) using  $v_m = \mathbf{b}_m^* \cdot \mathbf{v}$  (exercise).
- Again going to the example  $\mathbb{R}^3$ , you would usually think of the orthonormal basis  $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$  for  $\mathbf{i} = [1, 0, 0]^T$ ,  $\mathbf{j} = [0, 1, 0]^T$ ,  $\mathbf{k} = [0, 0, 1]^T$ , but could equally well choose  $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$  with  $\mathbf{b}_1 = \frac{1}{\sqrt{3}}[1, 1, 1]^T$ ,  $\mathbf{b}_2 = \frac{1}{\sqrt{6}}[1, -2, 1]^T$ ,  $\mathbf{b}_3 = \frac{1}{\sqrt{2}}[1, 0, -1]^T$ , which is also orthonormal (exercise).

When working with vectors, one often also ends up dealing with matrices. You may know an  $n \times m$  matrix firstly as a rectangular array (tableaux, grid) of numbers with  $n$  rows and  $m$  columns. However mathematicians would rather define a matrix via a:

**Linear transformation:** A linear transformation  $T$  is a map from an  $N$  dimensional vector space  $V$  to an  $M$  dimensional one  $W$ :

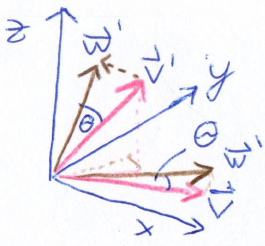
$$T : \mathbf{v} \rightarrow T(\mathbf{v}) = \mathbf{w}, \quad (1.11)$$

that maps vectors  $\mathbf{v} \in V$  onto those  $\mathbf{w} \in W$ , and is linear, which means:

$$T(a\mathbf{v} + b\mathbf{w}) = aT(\mathbf{v}) + bT(\mathbf{w}). \quad (1.12)$$

For any deeper discussion of linear transformations we refer to linear algebra books or courses.

**Example 6, Rotations:** One example of a linear map are rotations in  $\mathbb{R}^3$ :



**left:** A rotation by an angle  $\theta$  around the  $z$ -axis uniquely allocates a rotated vector  $\mathbf{w}$  (brown) to each input vector  $\mathbf{v}$  (pink).

We mostly require the case in the following, where both vector spaces are identical  $V = W$ , so we shall assume that in the following. It turns out that we know everything about a linear transformation, if we know what it does with each orthonormal basis vector. To see this, let us apply it to one of them:

$$T(\mathbf{b}_j) = \sum_{i=1}^{d=N} T_{ij} \mathbf{b}_i. \quad (1.13)$$

Here we have used the fact that, regardless of what the map does, we can write the result again in the basis  $\mathcal{B}$ . Using the orthonormality of the basis, we can find the coefficients  $T_{ij} = \mathbf{b}_i^* \cdot T(\mathbf{b}_j)$ , i.e. by “projecting” Eq. (1.14) onto the basis vector  $\mathbf{b}_i$ . Let us now write both, an arbitrary input vector  $\mathbf{v}$  and output vector  $\mathbf{w}$  of the transformation in the basis  $\mathcal{B}$ , i.e.  $\mathbf{v} = \sum_k v_k \mathbf{b}_k$  and  $\mathbf{w} = \sum_i w_i \mathbf{b}_i$ . We then see that:

$$T(\mathbf{v}) = T\left(\sum_k v_k \mathbf{b}_k\right) \stackrel{\text{Eq. (1.12)}}{=} \sum_k v_k T(\mathbf{b}_k) \stackrel{\text{Eq. (1.14)}}{=} \sum_k v_k \sum_i T_{ik} \mathbf{b}_i \equiv \sum_i \underbrace{\left(\sum_k T_{ik} v_k\right)}_{w_i} \mathbf{b}_i. \quad (1.14)$$

We thus know all the coefficients  $w_i = \sum_k T_{ik} v_k$  in the basis expansion for  $\mathbf{w}$  in terms of those of  $\mathbf{v}$  and the numbers  $T_{ik}$ .

We can write these as a

**Matrix:** A matrix  $\underline{M}$  is an  $N \times M$  table of numbers  $M_{ik}$ , where  $i$  numbers the row and  $k$  the column. The numbers  $ik$  are called matrix elements.

- As per discussion before, we can view every matrix as a representation of a linear transformation using specific bases.
- You recognise  $w_i = \sum_k T_{ik} v_k$  as the component notation for the vector matrix product  $\mathbf{w} = \underline{T} \cdot \mathbf{v}$ .
- Griffith reminds you of all these things only in section 3.1.3. I would encourage you to jump ahead there if you wish to. Do not be disturbed by the notation used, where they write  $|v\rangle$  instead of vector  $\mathbf{v}$  (just swap the notation in your head). We shall discuss the reason for that notation in section 3.1 later.

**Example 7, Rotation matrices:** We can again look at example 6 and thus now infer, that for the linear transformation “rotation around  $z$ -axis by angle  $\theta$ ”, there exists a specific matrix (called rotation matrix)  $\underline{O}_z(\theta)$ , which provides  $\mathbf{w} = \underline{O}_z(\theta)\mathbf{v}$  and  $\mathbf{w}' = \underline{O}_z(\theta)\mathbf{v}'$ . Note that this is a single matrix, that rotates all vectors correctly. *If you have a deeper interest, see my notes for PHY305, section 3.4 .*

Throughout physics, we make extensive use of the



**Matrix eigenvalue equation:**

$$\underline{\underline{M}} \cdot \mathbf{v}_k = \lambda_k \mathbf{v}_k, \quad (1.15)$$

for a square matrix  $\underline{\underline{M}}$ , with eigenvectors  $\mathbf{v}_k$  for eigenvalues  $\lambda_k$ .

- The subscript  $k$  indicates that an  $n \times n$  matrix can have multiple eigenvectors and eigenvalues, up to  $n$ .
- Sometimes, multiple eigenvectors share the same eigenvalue. That eigenvalue is then called degenerate.
- We skip a review here on how to practically find the eigenvalues and eigenvectors of a given matrix. However it will be important for this course. If you are not 100% comfortable, please consult your favorite linear algebra book or course, or Griffith section 3.1.4. For assignments it will typically be permitted and recommended to find eigenvalues and eigenvectors with a computer, for this get your copy of `mathematica` from the CC webpage and learn about the command `EigenSystem`.

We will deal here mostly with complex square matrices. Even for that subset, there are lots of different types of matrices that have their own name, which you are reminded of in the following<sup>4</sup>:

**Special matrices:**

$$\begin{array}{ll} \text{Symmetric matrices: } \underline{\underline{S}} = \underline{\underline{S}}^T, & \text{Hermitian matrices: } \underline{\underline{Q}} = \underline{\underline{Q}}^\dagger, \\ \text{Orthogonal matrices: } \underline{\underline{O}} \cdot \underline{\underline{O}}^T = \mathbf{1}, & \text{Unitary matrices: } \underline{\underline{U}} \cdot \underline{\underline{U}}^\dagger = \mathbf{1}. \end{array}$$

Finally we are in a position to make the statement for which we have included section 1.5.3. If you don't like any of the above, just remember the following

**Eigenvectors of a Hermitian matrix:** A  $N \times N$  Hermitian matrix  $\underline{\underline{Q}}$  with  $O_{ij} \in \mathbb{C}$  has exactly  $N$  eigenvectors with real eigenvalues (not necessarily different). These eigenvectors  $\{\mathbf{v}_k\}$  form a basis of the vector space  $\mathbb{C}^N$ .

Finally, for many<sup>5</sup> matrices one can do a

---

<sup>4</sup>Here  $T$  indicates the "transpose". I.e.  $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$  and  $\dagger$  the "complex adjoint" (transpose plus complex conjugation)  $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^\dagger = \begin{bmatrix} a^* & c^* \\ b^* & d^* \end{bmatrix}$

<sup>5</sup>but not all, see linear algebra

**Matrix diagonalisation:** We write the matrix  $\underline{M}$  as a product

$$\underline{M} = \underline{Q} \cdot \underline{D} \cdot \underline{Q}^{-1} \quad (1.16)$$

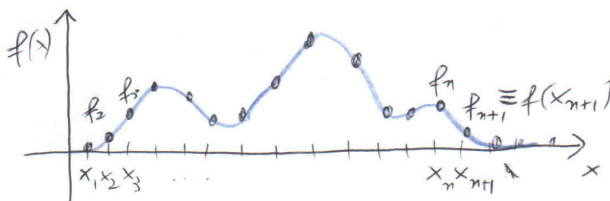
where  $\underline{D}$  is a diagonal matrix which contains the eigenvalues  $\lambda_k$  on the diagonal (see Eq. (1.15)), and  $\underline{Q}$  is an orthogonal matrix (see box above), which contains the eigenvectors as columns, i.e.  $\underline{Q} = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$ .

- Importantly for us, later, Hermitian and Orthogonal matrices are always diagonalizable.

### 1.5.4 Hilbertspaces and Operators

As we promised, it turns out “the space of all functions  $f(x)$  with certain properties”, is also a vector space. For these properties we shall take that  $\int dx |f(x)|^2 < \infty$ , which is called  $\mathbb{L}_2$  the set of all square integrable functions.

The mathematically minded can see that functions are also “vectors” by ticking off the definitions of a vector space in section 1.5.3 one by one (exercise, see also Griffith section 3.2). A more intuitive way to understand why functions can be treated as vectors is to discretize the space they are defined on as in the figure below. You would have to do this anyway, whenever you want to handle functions with a computer.



**left:** A continuous real function  $f(x)$  (blue) of one real number can be discretised by allowing only certain input values  $x_n$  with integer index  $n$ , and thus requiring function values at those points only  $f_n = f(x_n)$ . When we make the spacing  $\Delta x = |x_{n+1} - x_n|$  between adjacent discrete points infinitely small, we recover the continuous function.

Once discretized, we can represent the function  $f(x)$  as a vector  $\mathbf{f} = [f_1, f_2, f_3, \dots, f_n, f_{n+1} \dots]^T$  just as in section 1.5.3. However, since we formally would need to define infinitely many points  $x_k$  to cover  $-\infty < x < \infty$ , this is an infinite dimensional vector. The advantage of realizing that e.g.  $\mathbb{L}_2$  forms a vector space, is that we can generalize all the other useful concepts of section 1.5.3 to functions as well, as we shall do in the following.

We start with the scalar product  $\mathbf{v} \cdot \mathbf{u}$ , where we can instead use a

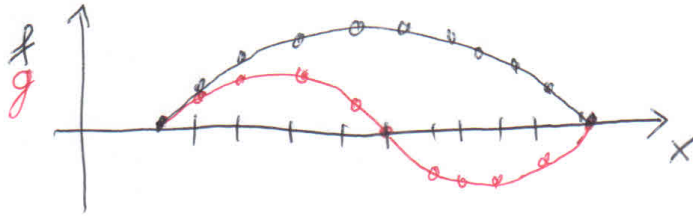
**Scalar product** of functions. Consider two functions in  $\mathbb{L}_2$ :  $f(x), g(x)$ . We take

$$(f, g) = \int_{-\infty}^{\infty} dx f^*(x)g(x) \quad (1.17)$$

as their scalar product.

- Mathematically one can again use a formal definition of the properties of a scalar product, which  $\mathbf{v} \cdot \mathbf{u}$  and  $(f, g)$  both fulfill, see linear-algebra books.
- Using the function scalar product, we can most importantly directly define the norm of a function  $\|f\| = (f, f) = \int_{-\infty}^{\infty} dx |f(x)|^2$ , in direct analogy to the norm of a vector  $\mathbf{v} \cdot \mathbf{v}$ . We then also call two functions  $f$  and  $g$  orthogonal, if  $(f, g) = 0$ .

**Example 8, Discretized orthogonal functions:** Consider the two (real) functions below, for which we also indicate a possible discretisation.



**left:** Two real functions  $f$  (black) and  $g$  (red), with their discretisation •

Since the integration  $(f, g) = \int_{-\infty}^{\infty} dx f^*(x)g(x) \equiv I$  is originally anyway defined via the discrete sum  $I = \sum_n \Delta x f_n^* g_n$ , we can see that this is directly proportional to a vector scalar product. We can also see how these two functions are orthogonal, since each possible product  $f_n^* g_n$  is compensated by one other term of opposite sign.

We can combine both prior concepts and define the

**Hilbert space:** as a complete<sup>a</sup> vector space with a certain scalar product.

<sup>a</sup>See math books, not so crucial here

- The primary example here is  $\mathbb{L}_2$  with the scalar product (1.17).
- We later see that all quantum wavefunctions form what we call the Hilbert-space in this lecture. This will be enough for chapter 2 of this lecture. Later, in chapter 3, we introduce even more abstract Hilbert spaces.

Since it is a vector space, also the Hilbert space  $\mathbb{L}_2$  has a basis:

**Function space basis:** We can write every function  $f(x)$  in  $\mathbb{L}_2$  in terms of an infinite dimensional basis  $\{b_n(x)\}$ , such that

$$f(x) = \sum_{n=0}^{\infty} f_n b_n(x). \quad (1.18)$$

- As for column vector-space bases, there are many different bases.

**Example 9, Function space bases:** A basis that you have already seen in other courses, is the basis of polynomials, which is used in the Taylor expansion:

$$f(x) = \sum_{n=0}^{\infty} \underbrace{\left( \frac{f^{(n)}(0)}{n!} \right)}_{=f_n} \underbrace{x^n}_{=b_n(x)}. \quad (1.19)$$

Another one is used in the Fourier series for even functions with period  $L$  (See PHY106, section 2.3.2):

$$f(x) = \sum_{n=0}^{\infty} \tilde{f}_n \cos\left(\frac{2\pi n}{L}x\right). \quad (1.20)$$

We will see many more examples in this course.

Next, we want to generalize the concept of a matrix (or linear transformation) to function vector spaces. We use the

**Linear operator:** as a linear map  $O : \mathbb{L}_2 \rightarrow \mathbb{L}_2$  of one function onto another, i.e.

$$\hat{O}f(x) = g(x). \quad (1.21)$$

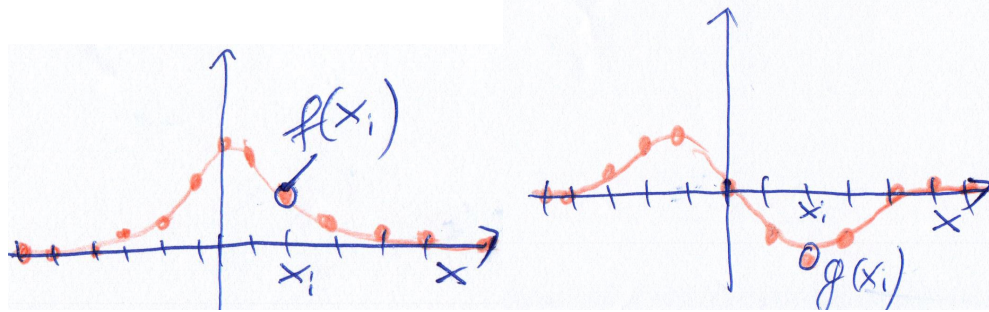
We say in (1.21) the operator  $\hat{O}$  is applied onto the function  $f(x)$  to yield the function  $g(x)$ .

- It is often helpful when doing quantum mechanics, to denote all operators with hats to distinguish them from the variables they describe (eigenvalues, see below). We shall consistently use the hat in this lecture.
- Note that (1.21) is structurally equivalent to a matrix-vector multiplication  $\underline{\underline{M}}\mathbf{v} = \mathbf{w}$  or the linear map (1.11) (in fact it IS a linear map, in the vector space of functions).
- Griffith calls Operators also “linear transformations”, but many other books call them operators.

To once again understand the relation between matrices and operators through a discretisation of space, see this example

**Example 10, Differential operator as a matrix:** A simple example is the derivative operator  $\hat{O} \rightarrow \frac{\partial}{\partial x}$ . Then, for example  $\hat{O} \underbrace{e^{-x^2}}_{=f(x)} = \underbrace{-2xe^{-x^2}}_{=g(x)}$ .

Both functions are drawn below in a discretized form as used in example 1.5.4.



First we can write each function as a vector  $\mathbf{f} = [f(x_1), f(x_2), \dots, f(x_n), \dots]^T$ ,  $\mathbf{g} = [g(x_1), g(x_2), \dots, g(x_n), \dots]^T$ . We recall (one possible) definition of a derivative  $f'(x) = \lim_{\Delta x \rightarrow 0} [f(x + \Delta x) - f(x - \Delta x)] / (2\Delta x)$ , but use it without the limit  $\Delta x \rightarrow 0$ , for a small but finite  $\Delta x = x_2 - x_1$ . You can convince yourself that it is then possible to write a matrix

$$\underline{Q} = \frac{1}{2\Delta x} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ \vdots & & & & & \end{bmatrix}, \quad (1.22)$$

such that  $\mathbf{g} = \underline{Q} \cdot \mathbf{f}$ . We call  $\underline{Q}$  a matrix representation of the operator  $\hat{O}$ .

We will see later (in section 3) that in quantum mechanics indeed any operator can be represented by a matrix.

At this point we can write the

**Operator eigenvalue equation:**

$$\hat{O}f_n(x) = o_n f_n(x), \quad (1.23)$$

for a linear operator  $\hat{O}$ , and eigenvalues  $o_n$  for eigenfunction  $f_n(x)$ .

- An example would be  $\underbrace{\frac{\partial}{\partial x}}_{=\hat{O}} \underbrace{\exp[kx]}_{=f_k(x)} = \underbrace{k}_{=\lambda_k} \exp[kx]$ .

Due to the analogy between matrices and operators discussed in example 10, all names for matrices that we introduced earlier generalize also for operators. In particular we require the notion of a

**Hermitian operator** as an operator  $\hat{O}$  for which

$$\left( \int_{-\infty}^{\infty} dx f^*(x) \hat{O} g(x) \right)^* = \int_{-\infty}^{\infty} dx g^*(x) \hat{O} f(x). \quad (1.24)$$

- Note, the definition makes use of the scalar product (1.17) between  $f(x)$  and  $\hat{O}g(x)$ .
- An example that you met in PHY106 (week 8), is  $\hat{O} = -i\hbar \frac{\partial}{\partial x}$ . Note, without the  $i$  it would not be Hermitian.

We now have at last reached the point which we wanted to deliver before week 2, that is the

**Eigenfunctions of a Hermitian operator** form an orthonormal basis of the Hilbertspace. Let  $\hat{O}f_n(x) = o_n f_n(x)$ . That the  $f_n(x)$  form a basis means that we can write any arbitrary function  $g(x)$  as

$$g(x) = \sum_{n=0}^{\infty} g_n f_n(x), \quad (1.25)$$

for some suitable coefficients  $g_n \in \mathbb{C}$ . That the  $f_n(x)$  are orthonormal implies

$$(f_n, f_k) = \int dx f_n^*(x) f_k(x) = \delta_{nk} \quad (1.26)$$

- We will prove this statement in week 6 and discuss this much more.
- Being able to make the statement above was the reason that we preponed some element from Griffith chapter 3. It is possible to understand most of the material from weeks 2-5 also without it, and they will provide many illustrative examples. Hence if the last parts of this week have confused you, don't worry for now, but please revisit this after week 6.

The above is a lot of material for one “week”, however I assume that the elements on complex numbers, vectors, matrices and probability theory are mostly familiar to you. Wherever it was not, it is crucial that you revise this from other sources. In practice, quantum mechanics very heavily requires linear algebra and probability theory. For the part of the week above starting from section 1.5.4 (which the mathematicians call “functional analysis”), I do NOT assume that you knew it before. Hence, if that was too fast, please ask plenty of questions.