1

Preliminaries

As the present project started as an attempt at rewriting Folland's textbook [31], we begin in the same direction, attempting though to give more details.

1.1 Dimension

The numerical value of many physical quantities depends on the unit one chooses to measure them. My height is 1.8 m, or 180 cm, or 1.90×10^{-16} light-years. The use of light-years here as a unit is weird, but not so much more than the use of centimeters to measure distances at the scale of a nucleon as many textbooks do. (A nucleon has a size of about 10^{-15} m $= 10^{-13}$ cm.) Tradition unfortunately has more weight than rationality in these matters.

The concept of "physical dimension" (which definitely differs from dimension in the mathematical sense) expresses how the numerical value of a physical quantity depends on the units you choose to measure it. A distance has dimension $[l]$ where l stands of course for length. If you increase the unit of length by a factor 100, the corresponding measure *decreases* by a factor 100: 100 cm = 1 m. Then a surface has dimension $[l^2]$: $(100)^2\,\text{cm}^2 = 1\,\text{m}^2$. A volume has dimension $[l^3]$: $1\,\text{km}^3 = (10^3)^3\,\text{m}^3 = 10^9\,\text{m}^3$. The unit of time can be chosen independently from the unit of length. Time has dimension $[t]$, so speed, which is a distance divided by a time, has dimension $[lt^{-1}]$. Thus 1 m/s = 3,600 m/h = 3.6 km/h. Acceleration, which is a change of speed divided by a time, has dimension $[lt^{-2}]$. It is of course a convention to choose time and length as fundamental quantities. One could make other choices, such as choosing time and speed as fundamental quantities. This is indeed basically what is actually done. Since 1983, in the international system the speed of light is *defined* to be *exactly*

$$c = 299{,}792{,}458\ \text{m/s} \tag{1.1}$$

and this serves as a definition of the meter given the unit of time.[1]

A formula in physics *must give a correct result independently of the system of units used.* This is a strong constraint. This is why it often makes sense to multiply or divide quantities of different dimensions, but it *never* makes sense to add them. As we learn in kindergarten,

[1] The reason for this definition is that the speed of light is a fundamental constant of Nature, and that it makes little sense to have its value change as the accuracy of measurements improves.

https://doi.org/10.1017/9781108225144.003 Published online by Cambridge University Press

you do not add pears with bananas. Furthermore, when a quantity occurs in a formula as the argument of, say, an exponential, it must be *dimensionless,* i.e. its value must be independent of the unit system. To understand a formula in physics it always helps to check that it makes sense with respect to dimension, a task we will perform many times.

The unit of mass can be chosen independently of the units of length and time. Momentum,[2] the product of a mass and a speed, then has dimension $[lt^{-1}m]$, and angular momentum, the product of a momentum and a distance, has dimension $[l^2t^{-1}m]$. Energy has the dimension of a mass times the square of a speed, that is $[l^2t^{-2}m]$. Less known is the *action* which occurs in Lagrangian Mechanics as the integral over a time interval of a quantity with the dimension of energy, and thus has the same dimension $[l^2t^{-1}m]$ as the angular momentum.

A fundamental constant of Nature is Planck's constant h, which represents the basic quantum of action (and in particular has the dimension of an action). In physical equations it often occurs in combinations with a factor of $1/2\pi$, so one defines the reduced Planck constant

$$\hbar = \frac{h}{2\pi}, \tag{1.2}$$

whose value is about[3]

$$\hbar = 1.0546 \times 10^{-34}\ \mathrm{J}\cdot\mathrm{s}. \tag{1.3}$$

This is small, as becomes more apparent if this value is expressed in units more related to the microscopic world[4] $\hbar \simeq 6.6 \times 10^{-16}\ \mathrm{eV}\cdot\mathrm{s}$. It is important to note that energy times time, momentum times length and angular momentum all have the same dimension as $\hbar$ so that their quotients by $\hbar$ are dimensionless. These quotients will occur in countless formulas.

Exercise 1.1.1 The Planck-Einstein relation gives the energy E of a photon of frequency ν as $E = h\nu$. Check that this formula makes sense with respect to dimension.

Exercise 1.1.2 The de Broglie momentum–wavelength relation states that to a particle of momentum p is associated a wavelength $\lambda = h/p$. Check that this formula makes sense with respect to dimension.

1.2 Notation

Since to enjoy this topic one has to read the work of physicists, it is best to adopt their notation from the beginning. Complex numbers play a central role, and the conjugate of a complex number a is denoted by a^*. Even some of the best authors let the reader decide

[2] Please do not worry if you do not have a real feeling for the concepts of momentum, energy, etc. (despite the fact that you should experience them every day). This is not going to be an obstacle.

[3] An action has the dimension of an energy times a time. In the International System of Units, the unit of energy is the joule J and the unit of time is the second s, so action is measured in $\mathrm{J}\cdot\mathrm{s}$. From May 2019, the value of h is *defined* to be exactly $6.62607015 \times 10^{-34}\mathrm{J}\cdot\mathrm{s}$.

[4] A joule is a huge energy at the microscopic scale. A more appropriate unit of energy at this scale is the electron-volt eV, the energy acquired by an electron going through a difference of potential of one volt.

https://doi.org/10.1017/9781108225144.003 Published online by Cambridge University Press

whether i denotes a complex number with $i^2 = -1$ or an integer index. Since this requires no extra work, the complex number will be denoted by i, so that $\mathrm{i}^* = -\mathrm{i}$.

When working with complex Hilbert spaces we adopt the convention that the inner product $(\cdot,\cdot)$ is anti-linear in the *first* variable (while often mathematicians use the convention that it is anti-linear in the second variable). One says that the inner product is *sesqui-linear*. That is, as another example of our notation for complex conjugation, we write

$$(ax, y) = a^*(x, y)$$

for any vectors x, y and any complex number a. Moreover

$$(y, x) = (x, y)^*. \tag{1.4}$$

The norm $\|x\|$ of a vector x is given by $\|x\|^2 = (x, x)$, and we recall the Cauchy-Schwarz inequality

$$|(x, y)|^2 \leq \|x\|^2 \|y\|^2,$$

where $|a|$ denotes the modulus of the complex number a. A basic example of a complex Hilbert space[5] is the space $\mathbb{C}^n$, where the inner product is defined by $(x, y) = \sum_{i \leq n} x_i^* y_i$, with the obvious notation $x = (x_i)_{i \leq n}$. Another very important example is the space $L^2(\mathbb{R})$ of complex-valued functions f on the real line for which $\int_{\mathbb{R}} |f|^2 \mathrm{d}x = \int_{\mathbb{R}} |f(x)|^2 \mathrm{d}x < \infty$, where $|f(x)|$ denotes the modulus of $f(x)$. The inner product is then given by $(f, g) = \int_{\mathbb{R}} f^* g \mathrm{d}x$. A physicist would actually write

$$(f, g) = \int_{-\infty}^{\infty} \mathrm{d}^1 x f(x)^* g(x), \tag{1.5}$$

where the superscript 1 refers to the fact that one integrates for a one-dimensional measure. The reason for which the $\mathrm{d}^1 x$ is put before the function to integrate is that this makes the formula easier to parse when there are multiple integrals. We will use this convention systematically. We will not however mention the dimension in which we integrate when this dimension is equal to one.

An *operator* A on a finite-dimensional Hilbert space $\mathcal{H}$ is simply a linear map $\mathcal{H} \to \mathcal{H}$. Its *adjoint* $A^\dagger$ is defined by

$$(A^\dagger(x), y) = (x, A(y)), \tag{1.6}$$

for all vectors x, y. (Mathematicians would use the notation A^* rather than $A^\dagger$.)

Exercise 1.2.1 (a) If A is an operator and α a number, prove the formula

$$(\alpha A)^\dagger = \alpha^* A^\dagger.$$

(b) If A and B are operators, prove that $(AB)^\dagger = B^\dagger A^\dagger$.

An operator (still on a finite-dimensional space) is called *Hermitian* if $A = A^\dagger$. A Hermitian operator A has the crucial property that if a subspace F is such that $A(F) \subset F$ then the orthogonal complement $F^\perp$ of F is also such that $A(F^\perp) \subset F^\perp$.

[5] In this work we consider only finite-dimensional or, more generally, *separable* Hilbert spaces.

https://doi.org/10.1017/9781108225144.003 Published online by Cambridge University Press

Exercise 1.2.2 Deduce this from the fact that $(A(x), y) = (x, A(y))$ for all x, y.

As a consequence, a Hermitian operator has a simple structure: there exists an orthonormal basis of eigenvectors.[6]

Exercise 1.2.3 Give a complete proof of this fact by induction over the dimension of the space.

Moreover the eigenvalues are real since $(x, A(x))$ is real for all x. Indeed,

$$(x, A(x)) = (A(x), x)^* = (x, A(x))^*,$$

where the second equality uses that A is Hermitian. If $A(x) = \lambda x$ then $(x, A(x)) = \lambda(x, x)$, so that $\lambda(x, x) = \lambda^*(x, x)$ and $\lambda = \lambda^*$. It is this property of having a basis of eigenvectors, with real eigenvalues, which makes the class of Hermitian operators so important.

A few times we will need the notion of *anti-linear* operator. Such a map T does satisfy $T(x + y) = T(x) + T(y)$ but for a scalar a we have $T(ax) = a^* T(x)$.

We should also mention that in physics, vectors and inner products are denoted differently, using Dirac's ubiquitous notation, which we will explain later. In most situations however we prefer to use standard mathematical notation, and it is unlikely that anybody reading this will mind.

1.3 Distributions

Laurent Schwartz invented the theory of distributions to give a rigorous meaning to many formal calculations of physicists. The theory of distributions is a fully rigorous part of mathematical analysis. In the main text however we will use only the very basics of this theory at a purely informal level. In Appendix L the reader may find an introduction to rigorous methods.

We will consider distributions on $\mathbb{R}^n$ but here we assume $n = 1$. The central object is the space $\mathcal{S} = \mathcal{S}(\mathbb{R})$ of *rapidly decreasing functions*, called also *test functions* or *Schwartz functions*. A complex-valued[7] function ζ on $\mathbb{R}$ is a test function if it has derivatives of all orders and if for any integers $k, n \geq 0$ one has[8]

$$\sup_x |x^n \zeta^{(k)}(x)| < \infty. \tag{1.7}$$

A distribution is simply a linear functional (which also satisfies certain regularity conditions which will not concern us before we reach Appendix L, as they will be satisfied in all the examples we will consider). That is, a distribution Φ is a complex linear map from $\mathcal{S}$ to $\mathbb{C}$, and for each test function ζ the number $\Phi(\zeta)$ makes sense. Such a distribution should actually be called a *tempered* distribution, but we will simply say "distribution" since we will

[6] Later we will meet a far-reaching extension of this fact, the spectral theorem for self-adjoint operators.

[7] A test function is typically complex-valued. At times, to avoid complications created by anti-linear operators, we also consider real-valued test functions. The space of such functions is denoted by $\mathcal{S}_{\mathbb{R}}$. Please try to remember that the subscript $\mathbb{R}$ on a space of test functions means that we consider only real-valued functions.

[8] The property is often described in words as follows: As $|x| \to \infty$, the function and each of its derivatives decrease faster than $|x|^{-n}$ for each n. This is what motivates the name "rapidly decreasing function".

https://doi.org/10.1017/9781108225144.003 Published online by Cambridge University Press

hardly consider any other type of distribution. Tempered distributions are also known under the name of *generalized functions.* This name has the advantage of explaining the point of the theory of distributions: it generalizes the theory of functions. Indeed, a sufficiently well-behaved function[9] f defines a distribution (= generalized function) Φ_f by the formula

$$\Phi_f(\zeta) = \int \mathrm{d}x\, \zeta(x) f(x). \tag{1.8}$$

Throughout the book *we maintain the convention that when the domain of integration is not mentioned, this domain is the whole space*. Thus (1.8) means $\Phi_f(\zeta) = \int_{\mathbb{R}} \mathrm{d}x f(x)\zeta(x) = \int_{-\infty}^{\infty} \mathrm{d}x f(x)\zeta(x)$.

As we are going to see, distributions can be strange animals. On the other hand, distributions given by a formula such as (1.8) are much better behaved. The short way to describe the situation where the distribution Φ is of the type Φ_f as in (1.8) is to simply say that Φ *is a function*.

In general a distribution is certainly not given by a formula of the type (1.8). However, when dealing with distributions we will maintain the *notational fiction* that they are functions, i.e. for a test function ζ we will write

$$\int \mathrm{d}x\, \zeta(x)\Phi(x) := \Phi(\zeta). \tag{1.9}$$

The use of the symbol := here is to stress that the right-hand side is a definition of the left-hand side, so that you may be reassured that your memory did not fail and that there is no point in looking back for the definition of the left-hand side. Equation (1.9) indeed defines the left-hand side, since the symbol $\Phi(x)$ is a *notation*, and a priori really *makes no sense whatsoever by itself*. It is only the integral $\int \mathrm{d}x\, \zeta(x)\Phi(x)$ which makes sense for any test function, and the value of this integral is (by definition) the quantity $\Phi(\zeta)$, as expressed in (1.9).[10] The central objects of this work, quantum fields, have precisely the previous property. The value of a quantum field cannot be specified at any given point. This value makes sense only "when it is integrated against a test function", or *smeared* in physics-type language.

Even for distributions, it is however sometimes possible to give a meaning to the quantity $\Phi(x)$ for certain values of x. Given an open interval I of the real line, we say that a distribution *is a function on* I if there exists a well-behaved function f on I such that $\Phi(\zeta) = \int \mathrm{d}x\, \zeta(x) f(x)$ whenever the test function ζ has compact support contained in I.[11] It is then reasonable to define $\Phi(x) = f(x)$ for $x \in I$. However, unless Φ is a function, it is not possible to assign a meaning to the symbol $\Phi(x)$ for each value of x.

[9] The meaning of the expression "well-behaved" varies depending on the context. In the present case, f needs to be locally integrable and "not grow too fast at infinity".

[10] The notational fiction (1.9) is however quite useful when one likes to be informal, as we simply do not distinguish distributions from functions until this leads us into trouble.

[11] When $I = \mathbb{R}$ this looks different from the definition (1.8) because then it is not required that ζ has compact support. However, using the regularity properties that are part of the definition of a distribution, and on which we do not dwell here one may show that these definitions coincide, see Appendix L.

https://doi.org/10.1017/9781108225144.003 Published online by Cambridge University Press

Distributions can be added, or multiplied by a scalar, but in general *cannot be multiplied*.[12] The great appeal of distributions is that they can always be differentiated. The derivative of a distribution Φ is the distribution *defined* by the formula

$$\Phi'(\zeta) := -\Phi(\zeta') \tag{1.10}$$

for every test function ζ. The reason behind this definition is best understood by integrating by parts when Φ is given by the formula (1.8) for a well-behaved function f. Then $\Phi'(\zeta) = \int \mathrm{d}x f'(x)\zeta(x)$.

Exercise 1.3.1 Convince yourself from the preceding definition that "if a distribution is actually a nice function, its derivatives as a function and as a distribution coincide". Hint: Recalling (1.8), prove that $(\Phi_f)' = \Phi_{f'}$.

Pretending that Φ' is also a function, we write $\Phi'(\zeta)$ as $\int \mathrm{d}x\zeta(x)\Phi'(x)$, and equally shamelessly we write (1.10) as

$$\int \mathrm{d}x\, \zeta(x)\Phi'(x) := -\int \mathrm{d}x\, \zeta'(x)\Phi(x). \tag{1.11}$$

In several dimensions, the class of test functions is defined as the class of infinitely differentiable functions such that the product of any partial derivative (of any order) and any polynomial in the variables is bounded. The reader may refer to Section L.1 for more about test functions.

1.4 The Delta Function

Besides reviewing the delta function, this section introduces the idea of a smooth cutoff, how to get rid of the troublesome part of an integral.

Mathematically, the delta "function" δ is simply the distribution given by

$$\delta(\zeta) = \zeta(0) \tag{1.12}$$

for any test function $\zeta \in \mathcal{S}$. Pretending that the delta "function" is actually a true function we will shamelessly write (1.12) as

$$\zeta(0) = \int \mathrm{d}x\zeta(x)\delta(x). \tag{1.13}$$

The name "delta function" is historical. Physicists have been using this object long before distributions were invented.

Exercise 1.4.1 (a) Convince yourself that it makes perfect sense to say that $\delta(x) = 0$ if $x \neq 0$.
(b) Make sure that you understand that despite the terminology, the delta function δ is not a function in the mathematical sense and that the quantity $\delta(0)$ *makes no sense*.

[12] One of the reasons for the dreaded infinities which will occur later is that we will have no other choice to proceed than pretending we can multiply certain distributions.

https://doi.org/10.1017/9781108225144.003 Published online by Cambridge University Press

(c) Convince yourself from (1.13) that, in the words of physicists, "the delta function δ is the function of x which is equal to zero for $x \neq 0$ and to infinity for $x = 0$, but in such a way that its integral is 1".
(d) Convince yourself that the *derivative* of the delta function δ, i.e. the distribution δ' given by $\delta'(\zeta) = -\zeta'(0)$ "does not look at all like a function".

For $a \neq 0$ let us *define* $\delta(ax)$ by

$$\int \mathrm{d}x\zeta(x)\delta(ax) := \frac{1}{|a|}\int \mathrm{d}x\zeta(x/a)\delta(x) = \frac{1}{|a|}\zeta(0), \tag{1.14}$$

so that

$$\delta(ax) = \frac{1}{|a|}\delta(x), \tag{1.15}$$

and in particular $\delta(-x) = \delta(x)$.[13]

Proper mathematical terminology requires one to say "the function f" but sooner or later one always says "the function $f(x)$" to carry at the same time the information that the variable is called x. In the same manner we will use expressions such as "the distribution $\Phi(x)$" which *should not* be interpreted as meaning that the quantity $\Phi(x)$ makes sense for a given x.

We will often write the quantity $\delta(x - y)$. It can be seen as a "function" of x depending on the parameter y. This "function" makes sense only when integrated in x against a test function ζ, and one has

$$\zeta(y) = \int \mathrm{d}x\zeta(x)\delta(x - y). \tag{1.16}$$

It can also be seen as a "function" of y depending on the parameter x, and one has

$$\zeta(x) = \int \mathrm{d}y\zeta(y)\delta(x - y). \tag{1.17}$$

Exercise 1.4.2 The quantity $\delta(x - y)$ can also be seen as a distribution Φ in the variables x, y. For a test function $\xi(x, y)$ one has

$$\iint \mathrm{d}x\mathrm{d}y\xi(x, y)\delta(x - y) := \Phi(\xi) := \int \mathrm{d}x\xi(x, x).$$

Convince yourself that this is consistent with (1.16) and (1.17).

Note also that $\delta(x - y) = \delta(y - x)$. We will shift freely between the previous meanings of the quantity $\delta(x - y)$. More generally we will stay very informal. Everything we say at this stage could be made rigorous, but this is not our objective.[14]

[13] Please be prepared: soon we will start manipulating delta functions as if they were functions, for example taking for granted the first equality in (1.14).
[14] It makes no sense to carefully climb molehills when the Himalayas are waiting for us.

https://doi.org/10.1017/9781108225144.003 Published online by Cambridge University Press

Exercise 1.4.3 For a test function ξ convince yourself of the formula (various versions of which we will use many times)

$$\int \mathrm{d}z\delta(x-z)\delta(z-y)\xi(z) = \delta(x-y)\xi(x). \tag{1.18}$$

We will make massive use of the formula[15]

$$\delta(x) = \frac{1}{2\pi}\int \mathrm{d}y \exp(\mathrm{i}xy), \tag{1.19}$$

and we first need to make sense of it. In studying physics, one must keep in mind at all times that the goal is to make predictions about the behavior of the physical world. This is difficult enough. To study the physical world, one makes models of it. One should as far as possible concentrate on problems that arise from the physical world, and stay away from the problems that arise not from the physical world, but from the models we made of it. We may never know for sure whether the physical world is finite or not, but certainly events located very far from our experiments are unlikely to affect very much their outcome, so their inclusion in our model is an idealization, and in (1.19) points very far from the origin should be discounted, for example by a factor $\exp(-ay^2)$ for a very small $a > 0$. This is an example of what is called a "smooth cutoff". Thus, rather than (1.19) we mean

$$\delta(x) = \lim_{a\to 0}\frac{1}{2\pi}\int \mathrm{d}y \exp(\mathrm{i}xy - ay^2). \tag{1.20}$$

Here the limit is "in the sense of distributions". *By definition of convergence in the sense of distributions*, this means that for every test function ζ,

$$\zeta(0) = \int \mathrm{d}x\zeta(x)\delta(x) = \lim_{a\to 0}\int \mathrm{d}x\zeta(x)\psi_a(x), \tag{1.21}$$

where

$$\psi_a(x) := \frac{1}{2\pi}\int \mathrm{d}y \exp(\mathrm{i}xy - ay^2) = \frac{1}{2\sqrt{a\pi}}\exp(-x^2/4a),$$

the second equality resulting from the computation of the Gaussian integral, see Lemma C.3.3. Making the change of variables $x = \sqrt{a}y$, (1.21) becomes

$$\zeta(0) = \lim_{a\to 0}\int \mathrm{d}y\zeta(y\sqrt{a})\psi_1(y),$$

which holds by dominated convergence since ψ_1 has integral 1 and ζ is uniformly bounded.[16]

A regularization procedure as in (1.20) can make sense only if it is robust enough. You might have chosen an origin different from mine, but this does not matter since for all b it follows from (1.20) that we actually have

$$\delta(x) = \lim_{a\to 0}\frac{1}{2\pi}\int \mathrm{d}y \exp(\mathrm{i}xy - a(y-b)^2).$$

[15] $\exp(x)$ is just another notation for e^x.

[16] Certainly the reader has observed the fundamental idea there: a family of functions of integral 1, which peaks more and more narrowly around zero converges to the delta function in the sense of distributions, consistently with Exercise 1.4.1 (a).

https://doi.org/10.1017/9781108225144.003 Published online by Cambridge University Press

Exercise 1.4.4 It would seem at first that the substitution $z = y - b$ in the right-hand side brings a factor $\exp(-\mathrm{i}xb)$. Why is there no such factor? Hint: Where is the delta function different from zero?

We will also need the obvious three-dimensional generalization of (1.19). In this case, not only might you and I have chosen different origins, we might also move at relativistic speed with respect to each other (and consequently we may not agree on the way we measure distances). There is however little point in investigating which specific regularization schemes would take care of this: as we will see later in this section, far more general regularization schemes work.

1.5 The Fourier Transform

Besides reviewing some basic facts about Fourier transforms, this section provides the first example of certain calculations common in physics.

The Fourier transform will play a fundamental role.[17] Let us temporarily denote by $\mathcal{F}_m$ the Fourier transform, that is

$$\mathcal{F}_m(f)(x) = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}y \exp(-\mathrm{i}xy) f(y), \tag{1.22}$$

where the subscript m reminds you that this is the way mathematicians like to define it (whereas our choice of normalization will be different). The right-hand side is defined for f integrable, and in particular for a Schwartz function $f \in \mathcal{S}$. Using integration by parts in the first equality, and differentiation under the integral sign in the second one, we obtain the fundamental facts that for any test function f,

$$\mathcal{F}_m(f')(x) = \mathrm{i}x\mathcal{F}_m(f)(x)\ ;\ \mathcal{F}_m(xf) = \mathrm{i}\mathcal{F}_m(f)', \tag{1.23}$$

where we abuse notation by denoting by xf the function $x \mapsto xf(x)$. An essential fact is that the Fourier transform of a test function is a test function. The details of the proof are a bit tedious, and are given in Section L.1.[18]

The Plancherel formula is the equality

$$(\mathcal{F}_m(f), \mathcal{F}_m(g)) = (f, g), \tag{1.24}$$

for $f, g \in \mathcal{S}$, where $(f, g) = \int \mathrm{d}x f(x)^* g(x)$. It is very instructive to "prove" this formula the way a physicist would, since this is a very simplified occurrence of the type of computations that are ubiquitous in Quantum Field Theory:

$$\begin{aligned}(\mathcal{F}_m(f), \mathcal{F}_m(g)) &= \frac{1}{2\pi} \int \mathrm{d}x \Big(\int \mathrm{d}y_1 \exp(-\mathrm{i}xy_1) f(y_1) \Big)^* \int \mathrm{d}y_2 \exp(-\mathrm{i}xy_2) g(y_2) \\ &= \frac{1}{2\pi} \iint \mathrm{d}y_1 \mathrm{d}y_2 f(y_1)^* g(y_2) \int \mathrm{d}x \exp(\mathrm{i}x(y_1 - y_2))\end{aligned}$$

[17] As will be explained later, it provides a natural correspondence between the "position representation" and the "momentum representation".

[18] The essential point is that iteration of the previous relations and Plancherel's formula show that if f is a test function then for each $n, k \in \mathbb{N}$ the function $x^n \mathcal{F}_m(f)^{(k)}(x)$ belongs to L^2.

https://doi.org/10.1017/9781108225144.003 Published online by Cambridge University Press

$$
\begin{aligned}
&= \iint \mathrm{d}y_1 \mathrm{d}y_2 f(y_1)^* g(y_2) \delta(y_1 - y_2) \\
&= \int \mathrm{d}y_2 f(y_2)^* g(y_2) \\
&= (f, g),
\end{aligned} \tag{1.25}
$$

where we have used (1.19) in the third line and have integrated first in y_1 in the fourth line. Although this type of manipulation might look scary at first to a mathematician, it suffices in fact to insert a factor $\exp(-ax^2)$ in the first line and let $a \to 0$ to make the argument rigorous using (1.20).

As a consequence of Plancherel's formula, for $f \in \mathcal{S}$, the Fourier transform of f has the same L^2 norm as f, i.e. $\mathcal{F}_m$ is an isometry when $\mathcal{S}$ is provided with the L^2 norm. Since $\mathcal{S}$ is dense in L^2 for this norm, an elementary result asserts that we may extend by continuity the Fourier transform as a linear map from L^2 to itself and this extension still satisfies (1.24). Please observe that it is by no means obvious that the right-hand side of the formula (1.22) is well-defined when $f \in L^2$.

One of the miracles of the Fourier transform is that it can be inverted by a formula very similar to (1.22):

$$
\mathcal{F}_m^{-1}(g)(y) = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}x \exp(\mathrm{i}xy) g(x). \tag{1.26}
$$

To justify the notation $\mathcal{F}_m^{-1}$ we observe that, using (1.26) for $g = \mathcal{F}_m(f)$, and using again (1.19),

$$
\begin{aligned}
\mathcal{F}_m^{-1}\left(\mathcal{F}_m(f)\right)(y) &= \frac{1}{2\pi} \int \mathrm{d}x \exp(\mathrm{i}xy) \int \mathrm{d}z \exp(-\mathrm{i}xz) f(z) \\
&= \frac{1}{2\pi} \int \mathrm{d}z f(z) \int \mathrm{d}x \exp(\mathrm{i}x(y - z)) \\
&= \int \mathrm{d}z f(z) \delta(y - z) \\
&= f(y).
\end{aligned} \tag{1.27}
$$

This again can be made rigorous just as (1.25). Let us now look back at (1.21), which we write, using Fubini's theorem,

$$
\zeta(0) = \lim_{a \to 0} \frac{1}{2\pi} \int \mathrm{d}y \exp(-ay^2) \left(\int \mathrm{d}x \exp(\mathrm{i}xy) \zeta(x) \right),
$$

and by dominated convergence we obtain

$$
\zeta(0) = \frac{1}{\sqrt{2\pi}} \int \mathrm{d}y \, \mathcal{F}_m^{-1}(\zeta)(y). \tag{1.28}
$$

Incidentally, it is now quite obvious that the regularization scheme in (1.20) is very robust. To see this, let us investigate for which regularizing families of functions ψ_a, we have in the sense of distributions

$$
\delta(x) = \lim_{a \to 0} \frac{1}{2\pi} \int \mathrm{d}y \exp(\mathrm{i}xy) \psi_a(y).
$$

https://doi.org/10.1017/9781108225144.003 Published online by Cambridge University Press

This means that for any test function $\zeta \in \mathcal{S}$ it holds that

$$\zeta(0) = \lim_{a\to 0} \frac{1}{2\pi} \iint \mathrm{d}x\mathrm{d}y\, \zeta(x) \exp(\mathrm{i}xy)\psi_a(y). \tag{1.29}$$

The right-hand side is

$$\lim_{a\to 0} \frac{1}{\sqrt{2\pi}} \int \mathrm{d}y\, \mathcal{F}_m^{-1}(\zeta)(y)\, \psi_a(y) = \lim_{a\to 0} \int \mathrm{d}y\, \theta(y)\psi_a(y),$$

where $\theta := (2\pi)^{-1/2}\mathcal{F}_m^{-1}(\zeta)$. Since θ is a test function, and since its integral is $\zeta(0)$ by (1.28), (1.29) holds true whenever ψ_a converges to the constant function 1 in the sense of distributions, that is

$$\lim_{a\to 0} \int \mathrm{d}y\, \eta(y)\psi_a(y) = \int \mathrm{d}y\, \eta(y)$$

for each test function η. This is the case for example (using dominated convergence) when $\psi_a(y) = \psi(ay)$ where $\psi \in \mathcal{S}$ satisfies $\psi(0) = 1$, and in particular when $\psi(x) = \exp(-x^2)$ as in (1.20).

Let us also mention that it is possible to define a notion of "Fourier transform of a distribution", and once this is done (1.28) is equivalent to the statement "the delta function is the Fourier transform of the constant function $1/\sqrt{2\pi}$", which is a more elaborate way to describe the way we made sense of (1.19).

Mathematicians love the symmetry between (1.22) and (1.26), but in physics it is better, thinking that x has the dimension of a length and p of a momentum to define the Fourier transform of a function f as

$$\hat{f}(p) = \int \mathrm{d}x \exp(-\mathrm{i}xp/\hbar) f(x) = \sqrt{2\pi}\mathcal{F}_m(f)(p/\hbar) \tag{1.30}$$

and the inverse Fourier transform as

$$\check{\xi}(x) = \int \frac{\mathrm{d}p}{2\pi\hbar} \exp(\mathrm{i}xp/\hbar)\xi(p). \tag{1.31}$$

This makes sense because the quantity $xp/\hbar$ is dimensionless. The Plancherel formula then becomes

$$\int \mathrm{d}x\, |f(x)|^2 = \int \frac{\mathrm{d}p}{2\pi\hbar} |\hat{f}(p)|^2. \tag{1.32}$$

Exercise 1.5.1 Make sure you understand (1.31) by writing out all details. The factor $2\pi\hbar$ will occur constantly.[19]

There are obvious multidimensional versions of these formulas. There is one factor $2\pi\hbar$ per dimension in the analog of (1.31). When integrating in p we will always include these factors.

Exercise 1.5.2 Write the multidimensional versions of these formulas.

[19] According to (1.2), we have $2\pi\hbar = h$ but we write all formulas in term of $\hbar$.

https://doi.org/10.1017/9781108225144.003 Published online by Cambridge University Press

The formula (1.23) now becomes

$$\widehat{-\mathrm{i}\hbar\frac{\mathrm{d}f}{\mathrm{d}x}}(p) = -\mathrm{i}\hbar\widehat{\frac{\mathrm{d}f}{\mathrm{d}x}}(p) = p\,\hat{f}(p). \tag{1.33}$$

The operator $-\mathrm{i}\hbar\mathrm{d}/\mathrm{d}x$ is fundamental in Quantum Mechanics, and (1.33) means that it is much simpler to express on $\hat{f}$ than on f: applying this operator to the function f simply amounts to multiplying the Fourier transform of f by p.

Key facts to remember:

- One should always check that an equation in physics makes sense from the point of view of physical dimension.
- The complex conjugate of a complex number a is denoted by a^* and the adjoint of an operator A by $A^\dagger$.
- Distributions generalize functions but their value is not defined at every point and they make sense only when integrated against a test function. The "delta function" is not a function!
- The ubiquitous Fourier transform is not exactly defined as it would be in mathematics.

https://doi.org/10.1017/9781108225144.003 Published online by Cambridge University Press