

Introduction to Constructive Quantum Field Theory

Christian Brennecke*

Abstract

In these notes, we discuss basic mathematical results in constructive quantum field theory. We start with a recap of classical mechanics, quantum mechanics, special relativity and quantum field theory, concluding with a precise notion of a quantum field theory based on the Wightman axioms. In the first main part we then focus on the explicit construction of several important free field theories (scalar, vector and Dirac spinor fields) which presupposes a detailed discussion of distributions and basic aspects of the representation theory of the Lorentz and Poincaré groups. The second part focuses on general properties of quantum field theories. This includes the reconstruction from Wightman functions, locality results, the spin-statistics theorem and Haag's theorem, which is related to the interaction picture. Many of these results rely on the analyticity properties of the Wightman functions which follow from a general discussion of the Laplace transform of tempered distributions. The third and final part focuses on basic aspects of the Euclidean approach to quantum field theory. Here, we discuss Gaussian measures on locally convex spaces, the Osterwalder-Schrader reconstruction theorem, aspects of the construction of the ϕ_2^4 model and some recent results related to Euclidean lattice Yang-Mills gauge theories.

*Institute for Applied Mathematics, University of Bonn, Endenicher Allee 60, 53115 Bonn, Germany

Contents

1	Introduction	1
1.1	Classical Mechanics	1
1.1.1	Canonical Transformations, Flows and Symmetries	5
1.2	Quantum Mechanics	11
1.2.1	Basic Examples via Canonical Quantization	15
1.2.2	Symmetries and Their Representations	25
1.3	Special Relativity	41
1.3.1	Examples of Classical Fields	48
1.4	Quantum Fields and the Wightman Axioms	53
2	Mathematical Interlude	58
2.1	Locally Convex Topological Vector Spaces	58
2.2	Distributions and Tempered Distributions	73
2.2.1	Tempered Distributions on $\mathbb{R}^n$	74
2.2.2	Distributions on Open Subsets $\Omega \subset \mathbb{R}^n$	77
2.2.3	Regularity and Nuclear Theorems	87
2.3	The Lorentz and Poincaré Groups	94
2.4	Operators on Fock Space	101
3	Construction of Free Quantum Fields	107
3.1	The Free Massive Scalar Field	107
3.1.1	Quantization of a Massive, Relativistic Particle	107
3.1.2	Quantization of the Massive Klein-Gordon Field	111
3.2	Representations of the Poincaré Group	114
3.2.1	Finite Dimensional Unitary Representations of $SU(2)$	114
3.2.2	Finite Dimensional Representations of $SL(2, \mathbb{C})$	114
3.3	The Free Massive Vector Field	114
3.4	The Free Massive Dirac Field	114
4	Mathematical Interlude	114
4.1	Laplace Transforms and Holomorphic Functions	114
4.2	Analytic Continuation and Edge of the Wedge Theorem	123
5	General Properties of Quantum Field Theories	123
5.1	Wightman Functions and Reconstruction Theorem	123
5.2	Locality and Irreducibility	123
5.3	Spin and Statistics	123
5.4	The Interaction Picture and Haag's Theorem	123
6	Mathematical Interlude	123
6.1	Gaussian Measures on Locally Convex Vector Spaces	124
6.2	The Euclidean Massive Klein-Gordon Field	124

7	The Euclidean Approach to Quantum Field Theory	124
7.1	Path Integral Formulation of Quantum Field Theories	124
7.2	Osterwalder-Schrader Reconstruction Theorem	124
7.3	Construction of the Euclidean ϕ_2^4 Field Theory	124
8	Basic Results on Lattice Yang-Mills Theories	124
8.1	Gauge Theories in Physics	124
8.2	Leading Order Partition Function of U(N) Lattice Yang-Mills	124

1 Introduction

In this section, we introduce basic notions and discuss standard examples from classical and quantum physics. The main purpose of this section is to motivate a certain notion of a quantum field theory which is based on the so called Wightman axioms. This is the central notion of a quantum field discussed in these notes. The main references on which these notes are based are [5, 31, 30], which contain extensive reference lists to background, advanced and the original research literature. For detailed physics background on classical mechanics, special relativity, quantum mechanics and quantum field theory, we refer the interested reader to [11, 12, 13, 33].

1.1 Classical Mechanics

A basic starting point for the discussion of classical mechanics is Newton's law. Consider a massive point particle of mass $m > 0$ which moves in some Euclidean space $\mathbb{R}^d$. In classical mechanics, our goal is to describe the trajectory of the particle at every given instance of time. This can, in principle, be determined based on Newton's law. If the position of the particle is described by the curve $t \mapsto x(t) \in \mathbb{R}^d$ and its momentum is $t \mapsto p(t) = m(dx/dt)(t) \in \mathbb{R}^d$, then Newton's law states that the change of momentum in time is equal to the force that acts on the particle, that is

$$\frac{dp}{dt} = m \frac{d^2x}{dt^2} = F(x). \quad (1.1)$$

Here, $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$ denotes the force that acts on the particle. Note that (1.1) describes a second order ODE and admits, under suitable assumptions on F , a unique regular (e.g. twice continuously differentiable) solution $t \mapsto x(t)$ with initial data $x_0, p_0 \in \mathbb{R}^d$ such that $x(0) = x_0, m(dx/dt)(0) = p_0$. In typical situations of interest, the force F is conservative which means that $F = -\nabla V$ for some potential $V : \mathbb{R}^d \rightarrow \mathbb{R}$. For instance, the gravitational force between a particle of mass $m_1 > 0$ at $x \in \mathbb{R}^3$ and a particle of mass $m_2 > 0$ fixed at the origin $0 \in \mathbb{R}^3$ is described by $F_{\text{gravity}}(x) = -Gm_1m_2x/|x|^3$ so that $F_{\text{gravity}} = -\nabla V_{\text{gravity}}$ for $V_{\text{gravity}}(x) = -Gm_1m_2/|x|$ (G denotes the gravitational constant). In the following we restrict our attention to conservative forces.

The dynamical law (1.1) implies that when no force ($F = 0$) acts on the particle, the particle moves at constant speed (and the trace of its position is a straight line). This is called, more specifically, Newton's first law or the law of inertia. Classical mechanics assumes this law to be correct under the assumption that the coordinate system that we use to describe our particle is a so called inertial frame. Such frames are rather vaguely described (see e.g. [11, Chapter 1]) as coordinate frames in which space is homogeneous and isotropic (no point and no direction play a mechanically distinct role) and time is homogeneous (no instance of time plays a mechanically distinct role). The existence of such frames is assumed (an example of a frame which is not an inertial system is one which accelerates with regards to a particle at rest that does not interact with anything else) and the Galilean relativity principle states that the mechanical laws can not distinguish one inertial system from another. Based on these considerations,

it follows that different inertial systems move with constant speed with regards to each other (consider a massive particle at rest upon which no force acts, centered at the origin of $\mathbb{R}^d$ with regards to one coordinate system, and apply the Galilean principle so that $x(t) = vt + x_0$ for all $t \in \mathbb{R}$ in the transformed system). From a more practical point of view, what this discussion implies is that if K and K' denote two inertial systems moving with relative speed $v \in \mathbb{R}^d$ with regards to one another and the position of our particle in K is described by the curve $t \mapsto x(t)$, then its position with regards to the coordinate system in K' is described by

$$x'(t) = x(t) - vt,$$

assuming that there is a universal time, parametrized by $t \in \mathbb{R}$, that is used in both systems K and K' (that is $t' = t$). In other words, within the framework of classical mechanics, we assume that time is measured independently of the (inertial) reference frame that we choose to describe mechanical phenomena.

Given a potential $V : \mathbb{R}^d \rightarrow \mathbb{R}$ and a solution $t \mapsto x(t)$ of (1.1), one observes that

$$H(x, p) = \frac{|p|^2}{2m} + V(x)$$

is preserved in time. Indeed, we readily find that

$$\frac{d}{dt}H(x, p) = p(t) \cdot \frac{d^2x}{dt^2} + \nabla V(x) \cdot \frac{dx}{dt} = 0.$$

The function $H : \mathbb{R}^{2d} \rightarrow \mathbb{R}$ is called Hamiltonian and it is identified with the energy of the system (the sum of kinetic and potential energies). One can reformulate the dynamics described by (1.1) as a system of first order ODE that involve H . This is the starting point for the Hamiltonian formulation of classical mechanics in which the possible mechanical states are described by points $(x, p) \in U \times \mathbb{R}^n$ for an open set $U \subset \mathbb{R}^n$ and some $n \in \mathbb{N}$ (more generally, the states are described by points in a $2n$ -dimensional manifold). The dynamics of the states is determined by Hamilton's equations

$$\frac{dx}{dt} = \nabla_p H, \quad \frac{dp}{dt} = -\nabla_x H, \tag{1.2}$$

where $\nabla_x = (\partial_{x_1}, \dots, \partial_{x_n})$ and $\nabla_p = (\partial_{p_1}, \dots, \partial_{p_n})$ denote conventionally the gradients in $U \times \mathbb{R}^n$ with regards to the first and, respectively, last n coordinates.

Example 1.1 (Single Particle in External Field). The Hamiltonian of a single particle of mass $m > 0$ that moves in $U \subset \mathbb{R}^d$ under the influence of an external potential $V : U \rightarrow \mathbb{R}$, whose position is $x \in U$ and whose momentum is $p \in \mathbb{R}^d$ is given by

$$H(x, p) = \frac{|p|^2}{2m} + V(x).$$

The Hamiltonian equations read

$$\frac{dx}{dt} = \nabla_p H = \frac{p}{m}, \quad \frac{dp}{dt} = -\nabla_x H = -\nabla V = F,$$

which is equivalent to our previous definition $p = m(dx/dt)$ and Newton's law (1.1).

Example 1.2 (System of Interacting Particles). Consider a collection of $N \in \mathbb{N}$ particles of mass $m > 0$ that move in $\mathbb{R}^d$ and interact with each other through some interaction potential $V : \mathbb{R}^d \rightarrow \mathbb{R}$. Their positions and momenta can be described collectively by $(x_1, \dots, x_N, p_1, \dots, p_N) \in \mathbb{R}^{dN} \times \mathbb{R}^{dN}$. The many-body energy is described by

$$H(x_1, \dots, x_N, p_1, \dots, p_N) = \sum_{i=1}^N \frac{|p_i|^2}{2m} + \sum_{1 \leq i < j \leq N} V(x_i - x_j).$$

The first contribution is called the kinetic energy while the second term on the right hand side describes the interaction energy among the particles. The dynamics reads

$$\frac{dx_i}{dt} = \frac{p_i}{m}, \quad \frac{dp_i}{dt} = \sum_{1 \leq j \leq N: j \neq i} F(x_i - x_j) \quad (\forall i = 1, \dots, N).$$

Example 1.3 (Particle in Static Magnetic and Electric Fields). In classical physics, the electromagnetic phenomena are described by Maxwell's equations. If we denote the (time-dependent) electric and magnetic fields by $E : \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ and, respectively, $B : \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$, then in the presence of charge and current densities $\rho : \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}$ and, respectively, $j : \mathbb{R} \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$, Maxwell's equations read

$$\begin{aligned} \operatorname{div} E &= \rho, & \nabla \times E &= -\frac{1}{c} \partial_t B, \\ \operatorname{div} B &= 0, & \nabla \times B &= \frac{1}{c} (\partial_t E + j). \end{aligned} \tag{1.3}$$

Here, c denotes the speed with which electromagnetic waves travel in the vacuum. The charge density ρ (which describes the electric charge per unit volume) describes the distribution of all electric charges so that the total charge Q in $U \subset \mathbb{R}^d$ at time t equals

$$Q(t) = \int_U dx \rho(t, x).$$

By (1.3), the densities ρ and j are necessarily related by the continuity equation

$$\partial_t \rho + \operatorname{div} j = 0. \tag{1.4}$$

In particular, if the current density has compact support, say in $B_R(0) \subset \mathbb{R}^d$ for some $R > 0$, this implies that the total charge in this region

$$\frac{dQ}{dt}(\cdot) = \int_{B_R(0)} dx \partial_t \rho(\cdot, x) = - \int_{\partial B_R(0)} \sigma(dx) j(\cdot, x) = 0$$

is constant. Here, the second step follows from integration by parts. In other words, Maxwell's equations explain the conservation of the total electric charge. In case of a single particle of charge e moving in $\mathbb{R}^3$ whose position is described by $t \mapsto x(t)$, the charge and current densities are equal to the distributions (cf. Section 2.2.2 below)

$$t \mapsto \rho(t, x) = e \delta_{x(t)} \in \mathcal{D}'(\mathbb{R}^3), \quad t \mapsto j(x, t) = e \delta_{x(t)} \frac{dx}{dt}(t) \in \mathcal{D}'(\mathbb{R}^3, \mathbb{R}^3), \tag{1.5}$$

where here and in the following, $\delta_y \in \mathcal{D}'(U)$ describes the Dirac distribution centered at $y \in U \subset \mathbb{R}^d$. On the other hand, the force exerted by electric and magnetic fields E and B on a particle of charge e is equal to the Lorentz force so that by Newton's law

$$m \frac{d^2 x}{dt^2} = eE + \frac{e}{c} \frac{dx}{dt} \times B. \quad (1.6)$$

The non-trivial system of equations (1.3), (1.5), (1.6) and generalizations thereof to a system of interacting charged particles (including possibly not only electromagnetic, but also other interactions such as gravity) models an enormous range of phenomena.

To relate the dynamics (1.6) to a Hamiltonian system, we need to relate the electric and magnetic fields to certain potentials. In typical situations of interest, this can be done as follows. Consider a particle moving in $\mathbb{R}^3$ whose influence on the charge and current densities ρ and j is negligible. Moreover, assume that the particle travels through a smooth magnetic field which is time-independent, that is $\partial_t B = 0$. Then (1.3) implies

$$(\partial_i E_j - \partial_j E_i) = 0$$

for all $i, j \in \{1, 2, 3\}, i \neq j$. Assuming E to be smooth and identifying the rotation $\nabla \times E$ with the exterior differential dE of the one form $E = \sum_{i=1}^3 E_i dx_i \in \Omega^1(\mathbb{R}^3)$, note that

$$dE = \sum_{1 \leq i < j \leq 3} (\partial_i E_j - \partial_j E_i) dx_i \wedge dx_j = 0.$$

In other words, $\partial_t B = 0$ and (1.3) imply that E is closed. By the Poincaré Lemma (see e.g. [14, Prop. 6.30 & Theorem 15.14]), we conclude that $E = d\Phi$ for some potential $\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}$ or, equivalently by slight abuse of notation, $E = \nabla \Phi$. Identifying similarly the magnetic field B with the differential two form

$$B = B_1 dx_2 \wedge dx_3 - B_2 dx_1 \wedge dx_3 + B_3 dx_1 \wedge dx_2 \in \Omega^2(\mathbb{R}^3),$$

the second identity in (1.3) is equivalent to $dB = 0$, so that $B = dA$ for $A = \sum_{i=1}^3 A_i dx_i$ or, equivalently, $B = \nabla \times A$ for a magnetic vector potential $A = (A_1, A_2, A_3) : \mathbb{R}^3 \rightarrow \mathbb{R}^3$.

In terms of the electric and magnetic potentials (Φ, A) , the identity (1.6) reads

$$m \frac{d^2 x}{dt^2} = e \nabla \Phi + \frac{e}{c} \frac{dx}{dt} \times \nabla \times A$$

and setting $p = m dx/dt + eA/c$, this dynamics can be reformulated as

$$\frac{dx}{dt} = \frac{1}{m} \left(p - \frac{e}{c} A \right), \quad \frac{dp_i}{dt} = \frac{e}{mc} \left(p - \frac{e}{c} A \right) \cdot \partial_i A - e \partial_i \Phi \quad (i = 1, 2, 3).$$

Notice that this corresponds to the Hamiltonian dynamics for the energy function

$$(x, p) \mapsto H(x, p) = \frac{1}{2m} \left| p - \frac{e}{c} A(x) \right|^2 + e \Phi(x).$$

1.1.1 Canonical Transformations, Flows and Symmetries

Consider now a general Hamiltonian system with smooth energy $H \in C^\infty(\mathcal{P})$ and (flat) phase space $\mathcal{P} = U \times \mathbb{R}^n$. As already remarked in the previous section, the Hamiltonian H is preserved in time under the Hamiltonian dynamics (1.2). Indeed, this follows from

$$\frac{dH}{dt} = \nabla_x H \cdot \frac{dx}{dt} + \nabla_p H \cdot \frac{dp}{dt} = \sum_{i=1}^n (\partial_{x_i} H \partial_{p_i} H - \partial_{p_i} H \partial_{x_i} H) = 0.$$

In typical examples the Hamiltonian H is not the only quantity that is preserved in time, but also other quantities such as the momentum or angular momentum may be preserved. This is related to specific symmetries of the Hamiltonian. In order to make this more precise, and in foresight of the following sections, let us introduce some additional machinery. Let $F \in C^\infty(\mathcal{P})$ be a smooth function and denote by $t \mapsto (x(t), p(t)) \in \mathcal{P}$ a (local) solution of Hamilton's equations (1.2). Then a straightforward calculation shows

$$\frac{d}{dt} F \circ (x, p) = \left(\nabla_x F \cdot \frac{dx}{dt} + \nabla_p F \cdot \frac{dp}{dt} \right) \circ (x, p) = \{F, H\} \circ (x, p), \quad (1.7)$$

where for $F, G \in C^\infty(\mathcal{P})$, we introduced the Poisson bracket

$$\{F, G\} = \sum_{i=1}^n (\partial_{x_i} F \partial_{p_i} G - \partial_{p_i} F \partial_{x_i} G).$$

It is bilinear as a map $\{\cdot, \cdot\} : C^\infty(\mathcal{P}) \times C^\infty(\mathcal{P}) \rightarrow C^\infty(\mathcal{P})$, it satisfies

$$\{F, G\} + \{G, F\} = 0, \quad \{\{F, G\}, H\} + \{\{G, H\}, F\} + \{\{H, F\}, G\} = 0$$

and an explicit calculation shows that for all $i, j \in \{1, \dots, n\}$, we have

$$\{x_i, x_j\} = 0, \quad \{p_i, p_j\} = 0, \quad \{x_i, p_j\} = \delta_{ij}. \quad (1.8)$$

By (1.7), the vanishing of $\{F, H\} = 0$ implies that F is a conserved quantity under the Hamiltonian dynamics (1.2). It is thus natural to look for criteria that ensure the vanishing $\{F, G\} = 0$ of the Poisson bracket. Below, we relate this to certain observable-associated flows, for which we need to introduce a few additional tools.

Denote by $\mathbf{J} \in \mathbb{R}^{2n \times 2n}$ the matrix

$$\mathbf{J} = \begin{pmatrix} 0 & \mathbf{1}_{\mathbb{R}^n} \\ -\mathbf{1}_{\mathbb{R}^n} & 0 \end{pmatrix}$$

and recall that the symplectic group $\mathrm{Sp}(2n) = \mathrm{Sp}(2n, \mathbb{R})$ is defined by

$$\mathrm{Sp}(2n) = \{M \in \mathbb{R}^{2n \times 2n} : M^T \mathbf{J} M = \mathbf{J}\}.$$

Matrices $M \in \mathrm{Sp}(2n)$ are called symplectic and they leave the symplectic bilinear form $\sigma : \mathbb{R}^{2n} \times \mathbb{R}^{2n} \rightarrow \mathbb{R}$, defined by $\sigma(\zeta_1, \zeta_2) = \zeta_1 \cdot \mathbf{J} \zeta_2$, invariant. That is

$$\sigma(M\zeta_1, M\zeta_2) = \sigma(\zeta_1, \zeta_2)$$

for all $\zeta_1, \zeta_2 \in \mathbb{R}^{2n}$. Symplectic matrices are invertible, because $1 = \det(M^T \mathbf{J} M) = |\det M|^2$ and we have that $M^{-1} = -\mathbf{J} M^T \mathbf{J}$, because $\mathbf{J}^2 = -\mathbf{1}_{\mathbb{R}^{2n}}$.

Symplectic matrices occur naturally in the question for which coordinate transformations $\phi \in C^\infty(\mathcal{P}, \mathbb{R}^{2n})$ the form of the Hamiltonian equations (1.2) is preserved. For a precise statement in the following lemma, let us say that $\phi \in C^\infty(\mathcal{P}, \mathbb{R}^{2n})$ is a canonical transformation if its differential satisfies $(D\phi)(\zeta) \in \text{Sp}(2n)$, for every $\zeta \in \mathcal{P}$. Note that this implies in particular that ϕ is a local diffeomorphism, by the inverse function theorem. In other words, we can think of ϕ as a (local) coordinate transformation.

Lemma 1.1. *Let $t \mapsto \zeta(t)$ be a solution to Hamiltonian's equations so that*

$$\frac{d\zeta}{dt} = \mathbf{J} \nabla H(\zeta).$$

If $\zeta' = \phi \circ \zeta$ for a canonical transformation $\phi \in C^\infty(\mathcal{P})$, then

$$\frac{d\zeta'}{dt} = \mathbf{J} \nabla (H \circ \phi^{-1})(\zeta').$$

Moreover, setting $\phi^ F = F \circ \phi$, ϕ is canonical if and only if for all $F, G \in C^\infty(\mathbb{R}^{2n})$*

$$\phi^* \{F, G\} = \{\phi^* F, \phi^* G\}.$$

Remark 1.1. *Note that if H describes the energy of a particle system described within a given reference frame whose coordinates are $\zeta = (x, p) \in \mathcal{P}$, then if $\phi \in C^\infty(\mathcal{P}, \mathbb{R}^{2n})$ is a (possibly local) coordinate transformation (a local diffeomorphism), then $H \circ \phi^{-1}$ describes the energy in terms of the new coordinates $\zeta' = \phi \circ \zeta$.*

Remark 1.2. *Geometrically, $\phi^* F$ is the pullback of the form $F \in \Omega^0(\mathbb{R}^{2n}) = C^\infty(\mathbb{R}^{2n})$.*

Proof. Using the chain rule, we verify that

$$\begin{aligned} \frac{d\zeta'}{dt} &= D\phi(\zeta) \frac{d\zeta}{dt} = D\phi(\zeta) \mathbf{J} \nabla H(\zeta) = D\phi(\zeta) \mathbf{J} \nabla (H \circ \phi^{-1})(\zeta') \\ &= D\phi(\zeta) \mathbf{J} D\phi^T(\zeta) \nabla (H \circ \phi^{-1})(\zeta') \\ &= \mathbf{J} \nabla (H \circ \phi^{-1})(\zeta'), \end{aligned}$$

where we used that $\nabla H = (DH)^T$ and that $M^T \in \text{Sp}(2n)$ if $M \in \text{Sp}(2n)$ (*exercise*).

For the second statement, notice that $\{F, G\} = \nabla F \cdot \mathbf{J} \nabla G$ as well as $\nabla(\phi^* F) = (D\phi)^T(\nabla F) \circ \phi$, so that the pullback identity is equivalent to the statement that

$$(\nabla F \cdot \mathbf{J} \nabla G) \circ \phi = ((\nabla F) \circ \phi) \cdot D\phi \mathbf{J} (D\phi)^T (\nabla G) \circ \phi.$$

Now, on the one hand, if ϕ is canonical, then $D\phi \mathbf{J} (D\phi)^T = \mathbf{J}$ and we conclude the invariance of the Poisson bracket under taking the pullback by ϕ . On the other hand, assuming that the invariance of the Poisson bracket, we can choose the canonical coordinate functions x_i, p_j (for F, G), for $i, j = 1, \dots, n$, in $\mathbb{R}^{2n}$ to deduce that $D\phi \mathbf{J} (D\phi)^T = \mathbf{J}$. $\square$

Next, let us recall the notion of the flow and the Lie derivative with regards to a vector field. Let $X \in C^\infty(\mathcal{P}, \mathbb{R}^{2n})$ be a smooth vector field (identifying $\mathcal{P} = U \times \mathbb{R}^n$ with a flat, smooth manifold and $\mathbb{R}^{2n} \simeq T\mathcal{P} = \bigsqcup_{(x,p) \in \mathcal{P}} \mathbb{R}^{2n}$ with its tangent space). By basic existence and uniqueness theory for ordinary differential equations, recall that for every $\zeta \in \mathcal{P}$ there exists a (non-empty) time interval $(-t_\zeta, t_\zeta) \subset \mathbb{R}$, an open set $U_\zeta \subset \mathcal{P}$ that contains ζ and a smooth map $\Phi_X : (-t_\zeta, t_\zeta) \times U_\zeta \mapsto \mathcal{P}$ such that for every $\xi \in U_\zeta$, the map $t \mapsto \Phi_X(t, \xi)$ is equal to the unique, smooth solution of the initial value problem

$$\begin{cases} \frac{df}{dt} = X(f), \\ f(0) = \xi. \end{cases}$$

In particular, $\Phi_X(t, \cdot)$ is a local diffeomorphism with inverse $\Phi_X(-t, \cdot)$, because

$$\Phi_X(s, \cdot) \circ \Phi_X(t, \cdot) = \Phi_X(t, \cdot) \circ \Phi_X(s, \cdot) = \Phi_X(s+t, \cdot).$$

Given $F \in C^\infty(\mathcal{P})$, a vector field X and its flow Φ_X , the Lie derivate $\mathcal{L}_X(F) \in C^\infty(\mathcal{P})$ of F in the direction of X is defined by

$$\mathcal{L}_X(F)(\zeta) = \frac{d}{dt} F(\Phi_X(t, \zeta))|_{t=0} = \lim_{t \rightarrow 0} \frac{1}{t} (F(\Phi_X(t, \zeta)) - F(\zeta)).$$

Recall that this is a natural geometric definition of the directional derivative of F (or, suitably extended, of a tensor field) in direction X . In standard coordinates $X = (X_1, \dots, X_{2n})$, we find from the chain rule that

$$\mathcal{L}_X(F)(\zeta) = \sum_{i=1}^{2n} (X_i \partial_{\zeta_i} F)(\zeta) = (\nabla F \cdot X)(\zeta) = (XF)(\zeta)$$

and, based on the group property of the flow, we find that

$$\begin{aligned} \mathcal{L}_X(\Phi_X(t, \cdot)^* F) &= \lim_{s \rightarrow 0} \frac{1}{s} (F(\Phi_X(t, \cdot) \circ \Phi_X(s, \cdot)) - F(\Phi_X(t, \cdot))) \\ &= \lim_{s \rightarrow 0} \frac{1}{s} (F(\Phi_X(s, \Phi_X(t, \cdot))) - F(\Phi_X(t, \cdot))) = \Phi_X(t, \cdot)^* (\mathcal{L}_X(F)), \end{aligned}$$

as well as

$$\frac{d}{dt} (\Phi_X(t, \cdot)^* F) = (\nabla F \cdot X) \circ \Phi_X(t, \cdot) = \Phi_X(t, \cdot)^* (\mathcal{L}_X(F)) = \mathcal{L}_X(\Phi_X(t, \cdot)^* F).$$

Given these notions, observe that Hamilton's equations (1.2) are equivalent to the flow dynamics with regards to the Hamiltonian vector field $X_H = \mathbf{J} \nabla H \in C^\infty(\mathcal{P}, \mathbb{R}^{2n})$. Consequently, the Hamiltonian dynamics (1.7) of observables can be rewritten as

$$\frac{d}{dt} \Phi_{X_H}(t, \cdot)^* F = \Phi_{X_H}(t, \cdot)^* \{F, H\} = \{\Phi_{X_H}(t, \cdot)^* F, H\}.$$

Here, the first equality is equivalent to (1.7). For the second equality, on the other hand, we used that $\mathcal{L}_{X_H}(F) = \{F, H\}$, which follows from evaluating the first equality at $t = 0$, and that $\frac{d}{dt} \Phi_{X_H}(t, \cdot)^* F = \mathcal{L}_{X_H}(\Phi_{X_H}(t, \cdot)^* F)$, proved above. The connection of the previous observations to canonical coordinates is the content of the next lemma.

Lemma 1.2. Denote by $(t, \zeta) \mapsto \Phi_t(\zeta) = \Phi_{X_H}(t, \zeta)$ the flow of the Hamiltonian vector field $X_H = \mathbf{J}\nabla H$. Then, for every $t \in \mathbb{R}$ for which Φ_t exists, $\Phi_t(\cdot)$ is canonical.

Proof. By Lemma 1.1, it suffices to show that for all $F, G \in C^\infty(\mathcal{P})$, we have that

$$\Phi_t^*\{F, G\} = \{\Phi_t^*F, \Phi_t^*G\}. \quad (1.9)$$

To this end, let us set $F_t = \Phi_t^*F$ and $G_t = \Phi_t^*G$ so that $dF_t/dt = \{F_t, H\}$ and thus

$$\frac{d}{dt}\{F_t, G_t\} = \{\{F_t, H\}, G_t\} + \{F_t, \{G_t, H\}\} = \{\{F_t, G_t\}, H\}$$

with $\{F_t, G_t\}|_{t=0} = \{F, G\}$ (all identities holding true pointwise in $\mathcal{P}$). Since

$$\frac{d}{dt}\Phi_t^*\{F, G\} = \{\Phi_t^*\{F, G\}, H\}$$

with $(\Phi_t^*\{F, G\})|_{t=0} = \{F, G\}$, the map $t \mapsto \Phi_t^*\{F, G\}$ solves the same initial value problem. By standard results for ordinary differential equations, we conclude (1.9). $\square$

Let us now get back to the question of characterizing quantities that are conserved under (1.2). We noted earlier that an observable $F \in C^\infty(\mathcal{P})$ is conserved if $\{F, H\} = 0$. From the previous discussion, we conclude that this is equivalent to the statement that $\Phi_{X_H}(t, \cdot)^*F = F$, that is, F is invariant under the flow associated to $X_H = \mathbf{J}\nabla H$.

Now, based on the characterization $\{F, H\} = 0$, the key observation is that we can also turn this picture around: F is conserved under the dynamics (1.2) if the Hamiltonian

$$H = \Phi_{X_F}(t, \cdot)^*H \iff H(x, p) = H(\Phi_{X_F}(t, x, p)), \forall (x, p) \in \mathcal{P}$$

is invariant under the flow $\Phi_{X_F}(t, \cdot)$ generated by the Hamiltonian vector field $X_F = \mathbf{J}\nabla F$. To find the conserved quantities, this naturally suggests to look at the symmetries of H , as first established by E. Noether. Indeed, in typical examples, the symmetries of H are described by smooth, linear Lie group actions. Recall that a Lie group G is a smooth manifold that is also a group and such that the group multiplication and inverse maps are smooth as well. Typical examples are matrix groups such as the group of linear, orthogonal transformations

$$\mathrm{O}(n) = \{R \in \mathbb{R}^{n \times n} : R^T R = \mathbf{1}_{\mathbb{R}^n}\}.$$

For simplicity of notation, let's identify in this brief discussion Lie group elements with their linear representations that act on $\mathcal{P}$. If H is invariant under such an action, it is invariant under the action of all its one-parameter subgroups (that is, continuous group homomorphisms) $\gamma : \mathbb{R} \rightarrow G$. In particular, if the Lie algebra $\mathfrak{g}$ of G is generated by $g_1, \dots, g_l$, then H is invariant under the integral curves $(-\epsilon, \epsilon) \ni t \mapsto \gamma_i(t) = \exp(tg_i)$, where $\exp : \mathfrak{g} \rightarrow G$ denotes the exponential map associated to G . To associate to each generator a conserved quantity $Q_i \in C^\infty(\mathcal{P})$, the task is then to solve for Q_i in

$$\frac{d}{dt}\gamma_i(\cdot)(x, p) = \mathbf{J}\nabla Q_i(\gamma_i(\cdot)(x, p)) \iff \mathbf{J}g_i(x, p) = \nabla Q_i(x, p), \forall (x, p) \in \mathcal{P}.$$

If $\mathcal{P}$ has vanishing first de Rham cohomology group, we can solve for Q_i if the differential form $\omega = \sum_{j=1}^k (\omega_j dx_j + \omega'_j dp_j)$ that is associated with the vector field $\mathbf{J}g_i$ is closed, that is $d\omega = 0$. Below, we illustrate this strategy through some explicit examples. For more on the geometric background and a general discussion of Noether's theorem in the context of symplectic manifolds, see e.g. [14, Chapters 9, 15, 16, 18 & 20].

Example 1.4 (Momentum Conservation). Consider a single particle moving in $U \subset \mathbb{R}^d$ in an external field $V : \mathbb{R}^d \rightarrow \mathbb{R}$. The Hamiltonian is of the form

$$H(x, p) = \frac{|p|^2}{2m} + V(x),$$

as in Example 1.1. Consider the translation group $\mathcal{G} = (\mathbb{R}^d, +)$ acting on $\mathcal{P} = U \times \mathbb{R}^d$ as $a(x, p) = (x + a, p)$, for $a \in \mathcal{G}$, and consider its one-parameter subgroups $\mathcal{G}_n = (\{nt : t \in \mathbb{R}\}, +)$, for directions $n \in \mathbb{R}^d$ with $|n| = 1$. Then, the flow Φ_{X_n} associated to the (constant, Hamiltonian) vector field $X_n = (n, 0) = \mathbf{J}\nabla P_n \in C^\infty(\mathcal{P}, \mathbb{R}^{2d})$, for $P_n \in C^\infty(\mathcal{P})$ denoting the momentum $P_n(x, p) = p \cdot n$ in direction n , is equal to

$$\Phi_{X_n}(t, x, p) = (x + tn, p).$$

Indeed, $\Phi_{X_n}(0, x, p) = (x, p)$ and $d\Phi_{X_n}(\cdot, x, p)/dt = (n, 0) = X_n \circ \Phi_{X_n}(\cdot, x, p)$.

Thus, P_n is conserved under (1.2) if

$$\Phi_{X_n}(t, \cdot)^* H = H \iff V(x) = V(x + tn), \forall (x, p) \in \mathcal{P}, t \in \mathbb{R}.$$

Generalized to the N -body setting as in Example 1.2, consider the Hamiltonian vector field $\mathbf{X}_n = (n, \dots, n, 0, \dots, 0) = \mathbf{J}\nabla \mathbf{P}_n \in C^\infty(\mathcal{P}^N, \mathbb{R}^{2dN})$, for $\mathbf{P}_n \in C^\infty(\mathcal{P}^N)$ denoting the momentum $\mathbf{P}_n(x, p) = \sum_{i=1}^N p_i \cdot n$. Arguing as above, the total momentum $\mathbf{P} = \sum_{i=1}^N p_i$ is a conserved quantity, because of the invariance

$$\Phi_{\mathbf{X}_n}(t, \cdot)^* H_N = H_N \iff \Phi_{\mathbf{X}_n}(t, \cdot)^* V(x_i - x_j) = V(x_i - x_j), \forall i, j, n \in \mathbb{R}^d, |n| = 1,$$

where we recall that $H_N(x_1, \dots, x_N, p_1, \dots, p_N) = \sum_{i=1}^N \frac{|p_i|^2}{2m} + \sum_{1 \leq i < j \leq N} V(x_i - x_j)$.

Example 1.5 (Angular Momentum Conservation). Consider a single particle moving in $\mathbb{R}^3$ with Hamiltonian $H : \mathcal{P} \rightarrow \mathbb{R}$ defined by

$$H(x, p) = \frac{|p|^2}{2m} + V(x),$$

as in Example 1.4, for $\mathcal{P} = \mathbb{R}^3 \times \mathbb{R}^3$. Recall that the special orthogonal group

$$SO(3) = \{R \in \mathbb{R}^{3 \times 3} : R^T R = \mathbf{1}_{\mathbb{R}^3}, \det R = 1\}$$

describes rotations in $\mathbb{R}^3$ (for more on $\mathfrak{so}(3)$, see Section 3.2). It is a well-known fact from linear algebra that a general rotation matrix $R \in SO(3)$ can be written as

$$R = e^{\omega n \cdot X}, \text{ where } n \cdot X = \sum_{i=1}^3 n_i X_i,$$

for an angle $\omega \in [0, 2\pi)$, a rotation axis $n \in \mathbb{R}^3$, $|n| = 1$ and where the matrices

$$X_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \quad X_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (1.10)$$

form a basis of the (real) linear space of skew-symmetric matrices

$$\mathfrak{so}(3) = \{X \in \mathbb{R}^{3 \times 3} : X = -X^T\}.$$

One can show, moreover, that every $R \in \text{SO}(3)$ can be decomposed into a product

$$R = e^{\varphi X_3} e^{\theta X_1} e^{\psi X_3}$$

of three rotations around the so called Euler angles φ, θ, ψ . We assume these basic facts in the sequel without proof. For detailed explanations, see e.g. [16, Sections 2.1 and 7.1] and Problem ?? below.

In geometric terms, $\mathfrak{so}(3) \simeq T_{1_{\mathbb{R}^3}}\text{SO}(3)$, equipped with the usual matrix commutator as the Lie bracket $[\cdot, \cdot] : \mathfrak{so}(3) \times \mathfrak{so}(3) \rightarrow \mathfrak{so}(3)$, represents the Lie algebra of the Lie group $\text{SO}(3)$ and the corresponding exponential map $\exp : \mathfrak{so}(3) \rightarrow \text{SO}(3)$, defined by

$$\exp(X) = e^X = \sum_{k=0}^{\infty} \frac{X^k}{k!},$$

is surjective. Matrices $R \in \text{SO}(3)$ act naturally on points in $\mathcal{P}$ as $R(x, p) = (Rx, Rp)$. Consider then e.g. the one-parameter subgroup $\{R_{e_3, \omega} : \omega \in [0, 2\pi)\} \subset \text{SO}(3)$ of rotations around the axis $e_3 = (0, 0, 1)$, where an explicit calculation yields (*exercise*)

$$R_{e_3, \omega} = \exp(\omega X_3) = \begin{pmatrix} \cos(\omega) & -\sin(\omega) & 0 \\ \sin(\omega) & \cos(\omega) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then, $\omega \mapsto R_{e_3, \omega}(x, p) = (e^{X_3 \omega} x, e^{X_3 \omega} p) = \Phi_{X_{e_3}}(\omega, x, p)$ is equal to the flow of

$$(x, p) \mapsto X_{e_3}(x, p) = (X_3 x, X_3 p) = \mathbf{J} \nabla L_3(x, p) \in C^\infty(\mathcal{P}, \mathcal{P}) \quad \text{for } L_3(x, p) = x_1 p_2 - x_2 p_1.$$

In particular, if $V(R_{e_3, \omega} \cdot) = V(\cdot)$ is invariant under rotations around e_3 , the angular momentum L_3 around e_3 is a conserved quantity. If, more generally, $V(R \cdot) = V(\cdot)$ for all $R \in \text{SO}(3)$, then the angular momentum $(x, p) \mapsto L(x, p) = x \times p$ is conserved. We leave the detailed verification of the last two statements as a basic *exercise*.

Problem 1.1. *Prove that $\text{O}(3)$ and $\text{SO}(3)$ are Lie groups and that $\mathfrak{so}(3) \simeq T_{1_{\mathbb{R}^3}}\text{SO}(3)$.*

1.2 Quantum Mechanics

Despite describing a vast range of phenomena, there are various physical observations that can not be explained based on the principles of classical mechanics. This includes e.g. the discreteness of atomic spectra or internal particle properties such as spin. Quantum theory generalizes classical mechanics in order to describe such phenomena. The mathematical setting is quite different, so let us summarize a few of the basic axioms.

First of all, possible states of a quantum mechanical system are normalized vectors, the so called wave functions, $\psi \in \mathcal{H}$, $\|\psi\| = 1$, in a complex, separable Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle)$. In accordance with standard physics notation, we assume throughout these notes that inner products are conjugate linear in the first slot and linear in the second slot. Wave functions $\psi_1, \psi_2 \in \mathcal{H}$ which only differ by a complex phase so that $\psi_1 = e^{i\omega} \psi_2$ for some $\omega \in [0, 2\pi)$, are identified as describing the same physics. In other words, one may identify the state space more precisely with the space of unit rays $\mathcal{H}/\sim$, where $\psi_1 \sim \psi_2$ if and only if $\psi_1 = e^{i\omega} \psi_2$ for some $\omega \in [0, 2\pi)$. This is sometimes referred to as a (global) gauge invariance (the phase describing the choice of a specific gauge).

Physical observables, such as the position or momentum of a massive particle, correspond to suitable self-adjoint operators $A : D_A \rightarrow \mathcal{H}$, D_A denoting the linear, dense domain of A . Recall that the adjoint A^* of such an operator $A : D_A \rightarrow \mathcal{H}$ has domain

$$D(A^*) = \{\varphi \in \mathcal{H} : D(A) \ni \psi \mapsto \langle \varphi, A\psi \rangle \in \mathbb{C} \text{ extends to a linear functional in } \mathcal{H}^*\}.$$

Riesz' lemma shows that if $\varphi \in D(A^*)$, then there exists a unique $\xi_\varphi \in \mathcal{H}$ such that $\langle \varphi, A\psi \rangle = \langle \xi_\varphi, \psi \rangle$ for every $\psi \in D(A)$ and one thus defines $A^*\varphi = \xi_\varphi$. An operator is self-adjoint if $D(A) = D(A^*)$ and $A = A^*$. Such operators have a canonical form which is the content of the spectral theorem. For its proof, see e.g. [21, Chapters VII & VIII].

Theorem 1.1 (Spectral Theorem). *Let $A : D_A \rightarrow \mathcal{H}$ be self-adjoint. Then, there is a measure space $\mathcal{M} = (\mathcal{X}, \mathcal{B}(\mathcal{X}), \mu)$ with a finite Borel measure μ , a unitary map $U : \mathcal{H} \rightarrow L^2(\mathcal{M})$ and a real-valued, μ -measurable function $f : \mathcal{X} \rightarrow \mathbb{R}$ such that*

$$UD_A = \{\varphi \in L^2(\mathcal{M}) : f\varphi \in L^2(\mathcal{M})\} \quad \text{and} \quad UAU^*\varphi = f\varphi \in L^2(\mathcal{M}), \quad \forall \varphi \in UD_A.$$

Define $g(A) = U^*(g \circ f)U$ for every bounded, measurable function $g : \mathbb{R} \rightarrow \mathbb{R}$, where $g \circ f$ is interpreted as multiplication operator in $L^2(\mathcal{M})$, and denote by $\chi_\Omega : \mathbb{R} \rightarrow \mathbb{R}$ the characteristic function of $\Omega \in \mathcal{B}(\mathbb{R})$. Then, the family of bounded, self-adjoint operators $(\chi_\Omega(A))_{\Omega \in \mathcal{B}(\mathbb{R})}$ forms a projection valued measure, which means that

$$i) \quad \chi_\emptyset(A) = 0 \text{ and } \chi_{\mathbb{R}}(A) = \mathbf{1}_{\mathcal{H}},$$

$$ii) \quad \chi_{\Omega_1}(A)\chi_{\Omega_2}(A) = \chi_{\Omega_1 \cap \Omega_2}(A) \text{ for every } \Omega_1, \Omega_2 \in \mathcal{B}(\mathbb{R}) \text{ and}$$

$$iii) \quad \chi_\Omega(A) = \sum_{i=1}^{\infty} \chi_{\Omega_i}(A) \text{ strongly in } \mathcal{H}, \text{ if } \Omega = \cup_{i=1}^{\infty} \Omega_i \text{ with } \Omega_i \cap \Omega_j = \emptyset, \forall i \neq j.$$

In terms of this projection-valued measure, $A : D_A \rightarrow \mathcal{H}$ has the spectral decomposition

$$A = \int_{\sigma(A)} \lambda \chi_{d\lambda}(A), \tag{1.11}$$

where $\sigma(A) \subset \mathbb{R}$ denotes the spectrum of A .

Remark 1.3. Recall that $\sigma(A) = \mathbb{C} \setminus \rho(A)$ where the resolvent set $\rho(A) \subset \mathbb{C}$ is given by

$$\rho = \{z \in \mathbb{C} : (A - z) \text{ admits a bounded inverse } (A - z)^{-1} : \mathcal{H} \rightarrow D_A\}.$$

We split $\sigma(A) = \sigma_d(A) \cup \sigma_{\text{ess}}(A)$ into a discrete part $\sigma_d(A)$, the set of isolated eigenvalues of A of finite multiplicity, and its complement $\sigma_{\text{ess}}(A)$, the essential spectrum.

Remark 1.4. The theorem generalizes the well-known fact from linear algebra that every Hermitian matrix $H = H^* \in \mathbb{C}^{n \times n}$ can be diagonalized and admits an orthonormal eigenbasis $(\varphi_i)_{i=1}^n$ so that $H\varphi_i = \lambda_i\varphi_i$, for suitable (real) eigenvalues $\lambda_i \in \mathbb{R}$. In this case, the spectral projection valued measure representation of H is simply given by

$$H = \sum_{i=1}^n \lambda_i |\varphi_i\rangle\langle\varphi_i|.$$

The map U can be defined by linearly extending $\mathbb{C}^n \ni \varphi_i \mapsto \chi_{\{\lambda_i\}} \in L^2(\Omega, \mathcal{B}(\Omega), \mu)$, where $\Omega = \{\lambda_i : i = 1, \dots, n\}$ and where μ denotes the counting measure on Ω .

Based on the spectral theorem, let us point out how, in the context of quantum mechanics, observables like the position or the momentum of a particle are connected with self-adjoint operators. Suppose that $A : D_A \rightarrow \mathcal{H}$ represents some observable $\mathcal{O}$ and that the system is in state $\psi \in \mathcal{H}$. Then, based on the normalization $\|\psi\| = 1$ and on the spectral decomposition (1.11) of A , one identifies $\mathcal{O}$ with a real-valued random variable (ranging almost surely in the spectrum $\sigma(A) \subset \mathbb{R}$ of A) and the probability $\mathbb{P}$ that $\mathcal{O}$ takes a specific value in some measurable set $\Omega \in \mathcal{B}(\mathbb{R})$ is defined as

$$\mathbb{P}(\mathcal{O} \in \Omega) = \int_{\Omega} \langle \psi, \chi_{d\lambda}(A)\psi \rangle = \langle \psi, \chi_{\Omega}(A)\psi \rangle. \quad (1.12)$$

Notice that the law $\Omega \mapsto \mathcal{O}_*(\mathbb{P})(\Omega) = \mathbb{P}(\mathcal{O} \in \Omega)$ defines indeed a Borel probability measure on $\mathbb{R}$. In other words, quantum mechanics only provides probabilistic predictions for the outcomes of physical measurements. Within this probabilistic interpretation, the expectation value $\mathbb{E}\mathcal{O}$ of the observable $\mathcal{O}$ represented by A is then equal to

$$\mathbb{E}\mathcal{O} = \int d\mathbb{P} \mathcal{O} = \int_{\mathbb{R}} \mathcal{O}_*(\mathbb{P})(d\lambda) \lambda = \int_{\sigma(A)} \lambda \langle \psi, \chi_{d\lambda}(A)\psi \rangle = \langle \psi, A\psi \rangle \quad (1.13)$$

and for this reason, we refer in the sequel to inner products like that on the r.h.s. of the previous equation as expectation values. Other basic statistical quantities from probability theory can be similarly related to suitable inner products, e.g. the variance

$$\mathbb{E}(\mathcal{O} - \mathbb{E}\mathcal{O})^2 = \mathbb{E}\mathcal{O}^2 - (\mathbb{E}\mathcal{O})^2 = \langle \psi, A^2\psi \rangle - \langle \psi, A\psi \rangle^2. \quad (1.14)$$

Observe that while (1.12) is well-defined for every $\psi \in \mathcal{H}$ (as it should be if every state $\psi \in \mathcal{H}$ is a possible state of the system), this need not be the case for (1.13) and (1.14).

Before switching to some concrete examples, let us explain how one describes the dynamics of quantum systems. Here, one postulates the existence of a strongly continuous one-parameter unitary group $(U_t)_{t \in \mathbb{R}}$ acting on $\mathcal{H}$. If the system is in state $\psi \in \mathcal{H}$

at time $t = 0$, then the system is in state $U_t\psi$ at every other time $t \in \mathbb{R}$. Note that $U_t\psi$ is indeed a valid state, because $\|U_t\psi\| = \|\psi\| = 1$, for all $t \in \mathbb{R}$. By (1.2), the classical dynamics is determined by the energy (the Hamiltonian). By analogy, one defines the Hamilton operator $H : D_H \rightarrow \mathbb{R}$ in quantum mechanics as the generator of the quantum dynamics $(U_t)_{t \in \mathbb{R}}$. This relies on the following fundamental result.

Theorem 1.2 (Stone's Theorem). *Let $(U_t)_{t \in \mathbb{R}}$ be a strongly continuous one-parameter unitary group on $\mathcal{H}$. Then, there exists a self-adjoint operator $H : D_H \rightarrow \mathcal{H}$ such that*

$$U_t = e^{-itH}, \quad \forall t \in \mathbb{R}.$$

Proof. We follow [21]. Before defining our candidate for H , we first need to find a suitable dense domain on which we can differentiate $t \mapsto U_t(\cdot)$. Using that, heuristically, $\phi \approx e^{-itH}\phi$ for small t (assuming we knew the existence of H already), it is useful to consider for $f \in C_c^\infty(\mathbb{R})$ and $\phi \in \mathcal{H}$ the vector space generated by vectors of the form

$$\phi_f = \int_{\mathbb{R}} dt f(t) U_t \phi \in \mathcal{H}.$$

Here, the integral on the r.h.s. can be defined as a vector-valued Riemann integral (and coincides with the usual Bochner integral). Set

$$D = \text{span}\{\phi_f : f \in C_c^\infty(\mathbb{R}), \phi \in \mathcal{H}\}.$$

Then $D \subset \mathcal{H}$ is dense, because for a standard approximation of the identity $(f_n)_{n \in \mathbb{N}}$ in $C_c^\infty(\mathbb{R})$ (choose e.g. $f = nf(n\cdot)$ for some $f \in C_c^\infty(\mathbb{R})$ such that $0 \leq f \leq 1$, $\int_{\mathbb{R}} f = 1$ so that $\text{supp}(f_n) \subset (-1/n, 1/n)$, $\forall n \in \mathbb{N}$), we have that

$$\limsup_{n \rightarrow \infty} \|\phi_{f_n} - \phi\| = \limsup_{n \rightarrow \infty} \left\| \int_{\mathbb{R}} dt f_n(t) (U_t \phi - \phi) \right\| \leq \limsup_{n \rightarrow \infty} \sup_{|t| \leq 1/n} \|U_t \phi - \phi\| = 0.$$

Next, we define H (initially on D) by differentiating $t \mapsto U_t$. For $\phi_f \in D$, we compute

$$\lim_{t \rightarrow 0} \frac{1}{t} (U_t \phi_f - \phi_f) = \lim_{t \rightarrow 0} \frac{1}{t} \int_{\mathbb{R}} ds f(s) (U_{t+s} - U_s) \phi = - \int_{\mathbb{R}} f'(s) U(s) \phi = -\phi_{f'},$$

where in the last step we applied the dominated convergence theorem. This suggests to define the operator $H : D \rightarrow D$ by

$$H\phi_f = i \lim_{t \rightarrow 0} \frac{1}{t} (U_t \phi_f - \phi_f) = -i\phi_{f'}.$$

By definition of $\phi_f \in D$, note that $U_t : D \rightarrow D$ for each $t \in \mathbb{R}$ ($U_t \phi_f = \phi_{f(-t)}$),

$H : D \rightarrow D$, $[U_t, H] = 0$ in D and that H is a symmetric operator, because

$$\begin{aligned}
\langle H\phi_f, \psi_g \rangle &= \langle -i\phi_{f'}, \psi_g \rangle \\
&= i \int_{\mathbb{R}^2} ds dt f'(t) g(s) \langle \phi, U_{-t+s} \psi \rangle \\
&= i \int_{\mathbb{R}^2} ds dt f'(t) g(s+t) \langle \phi, U_s \psi \rangle \\
&= -i \int_{\mathbb{R}^2} ds dt f(t) g'(s+t) \langle \phi, U_s \psi \rangle \\
&= \langle \phi_f, H\psi_g \rangle.
\end{aligned}$$

To conclude the theorem, it suffices to show that H is essentially self-adjoint and that the exponential of its (self-adjoint) closure is equal to U_t . For the first part, suppose that $\psi \in D(H^*)$ with $H^*\psi = i\psi$. Then, for each $\phi \in D$, we compute

$$\partial_t \langle U_t \phi, \psi \rangle = \langle -iH U_t \phi, \psi \rangle = -\langle U_t \phi, \psi \rangle.$$

This implies that $\langle U_t \phi, \psi \rangle = \langle \phi, \psi \rangle e^{-t}$ so that $\langle \phi, \psi \rangle = 0$, because $e^{-t} \rightarrow \infty$ as $t \rightarrow -\infty$ while $|\langle -U_t \phi, \psi \rangle| \leq \|\phi\| \|\psi\|$. Since $\phi \in D$ was arbitrary and $\overline{D} = \mathcal{H}$, this implies that $\psi = 0$. Repeating an analogous argument for the case $H^*\psi = -i\psi$, we deduce that $H : D \rightarrow D$ is essentially self-adjoint. Now, denote by $\overline{H} : D(\overline{H}) \rightarrow \mathcal{H}$ the self-adjoint closure of H and set $V_t = e^{-it\overline{H}}$. Given $\phi \in D$, we compute that

$$\partial_t (U_t \phi - V_t \phi) = -iH U_t \phi - i\overline{H} V_t \phi = -i\overline{H} (U_t - V_t) \phi,$$

which implies

$$\partial_t \|U_t \phi - V_t \phi\|^2 = 2 \operatorname{Im} \langle \overline{H} (U_t \phi - V_t \phi), U_t \phi - V_t \phi \rangle = 0.$$

Thus, $U_t \phi = V_t \phi$ for all $t \in \mathbb{R}$ and $\phi \in D$, so that $U_t = V_t$, using $\overline{D} = \mathcal{H}$. $\square$

In quantum mechanics, one identifies the dynamics conventionally with $(U_t)_{t \in \mathbb{R}} = (e^{-iHt/\hbar})_{t \in \mathbb{R}}$, where $\hbar$ is a fundamental small constant, called Planck's constant. The unitary dynamics is then related to the Schrödinger equation: if H has domain $D_H \subset \mathcal{H}$, then an application of the spectral theorem shows that for every $\varphi \in D_H$, the map

$$t \mapsto \psi(t) = e^{-itH/\hbar} \varphi \in C(\mathbb{R}, D_H) \cap C^1(\mathbb{R}, \mathcal{H}),$$

is the unique $C(\mathbb{R}, D_H) \cap C^1(\mathbb{R}, \mathcal{H})$ solution of the initial value problem

$$\begin{cases} i\hbar \frac{d\psi}{dt} = H\psi, \\ \psi(0) = \varphi. \end{cases} \quad (1.15)$$

For general initial data $\varphi \in \mathcal{H}$, not necessarily in D_H , the dynamics $(e^{-itH/\hbar} \varphi)_{t \in \mathbb{R}}$ solves (1.15) weakly (e.g. in form sense by testing the equation with elements from D_H).

Similarly as in (1.7) in the context of classical mechanics, one can study the dynamics of observables. Suppose that a physical observable $\mathcal{O}$ is represented by a self-adjoint operator $A : D_A \rightarrow \mathcal{H}$ and that, at time $t = 0$, the system is in state $\psi \in \mathcal{H}$, $\|\psi\| = 1$. As mentioned earlier in (1.12), the probability that $\mathcal{O} \in \Omega$ is defined as

$$\mathbb{P}(\mathcal{O} \in \Omega) = \langle \psi, \chi_\Omega(A)\psi \rangle.$$

If the system evolves in time until $t \in \mathbb{R}$ via $(U_t)_{t \in \mathbb{R}}$, this probability changes to

$$\mathbb{P}(\mathcal{O}_t \in \Omega) = \langle \psi, U_{-t} \chi_\Omega(A) U_t \psi \rangle = \langle \psi, \chi_\Omega(U_{-t} A U_t) \psi \rangle.$$

This clearly motivates to view $A_t = U_{-t} A U_t : U_{-t} D_A \rightarrow \mathcal{H}$ as representing the dynamical observable $(\mathcal{O}_t)_{t \in \mathbb{R}}$ at time $t \in \mathbb{R}$. Note that $A_t = A_t^*$, because unitary conjugation preserves self-adjointness. The formal quantum analogue of (1.7) thus becomes

$$\frac{d}{dt} A = \frac{1}{i\hbar} [A, H]. \quad (1.16)$$

Due to potential domain constraints, the rigorous interpretation of (1.16) requires some care. However, assuming e.g. that $A \in \mathcal{L}(\mathcal{H})$ extends in fact to a self-adjoint, bounded operator on $\mathcal{H}$, the operator dynamics (1.16) always admits a rigorous, weak formulation in quadratic form sense (by taking the expectation of (1.16) w.r.t. vectors in D_H).

1.2.1 Basic Examples via Canonical Quantization

In Section 1.1, we discussed several basic examples of classical mechanical systems. In this section, we describe the quantum mechanical versions of these models.

In order to generalize Examples 1.1, 1.2 and 1.3, we need to understand which self-adjoint operators should represent the position and the momentum of a particle. In our examples, the corresponding operators can be found based on a recipe called canonical quantization. This method is an important tool to guess and to formulate quantum mechanical (or quantum field theoretical) models with desirable properties that generalize their classical counterparts. It should be kept in mind, however, that this method has certain limitations (some of which are discussed below) and that, after all, quantum mechanics (or quantum field theory) is a more accurate theory to describe interacting particle systems than classical mechanics. For a quick introductory discussion of quantization schemes and their limitations, see e.g. [7, Chapters 13, 22 & 23].

Keeping the previous remarks in mind, let us explain the canonical quantization scheme. Comparing (1.7) with (1.16) motivates to replace the Poisson bracket $\{\cdot, \cdot\}$ of two classical observables by the commutator $\frac{1}{i\hbar}[\cdot, \cdot]$ of the corresponding quantum operators when passing from a classical to a quantum mechanical description. The constant of proportionality $i\hbar$ ensures consistency of the dynamics (1.7) and (1.16).

Let us apply the quantization scheme by reconsidering Examples 1.1, 1.2 and 1.3. Consider first a single, massive particle moving in $\mathbb{R}^d$, similarly as in Example 1.1. Its classical dynamics is determined by its position and momentum $(x, p) \in \mathcal{P} = \mathbb{R}^d \times \mathbb{R}^d$. Recalling (1.8), these satisfy as observables (that is, as coordinate functions in $C^\infty(\mathcal{P})$)

$$\{x_i, x_j\} = 0, \quad \{p_i, p_j\} = 0, \quad \{x_i, p_j\} = \delta_{ij}.$$

By the canonical quantization scheme, we now want to find a Hilbert space $\mathcal{H}$ and self-adjoint operators $X = (X_1, \dots, X_d), P = (P_1, \dots, P_d)$ that represent the position and, respectively, momentum of the particle so that (on a suitable domain $D \subset \mathcal{H}$)

$$[X_i, X_j] = 0, \quad [P_i, P_j] = 0, \quad [X_i, P_j] = i\hbar\delta_{ij}. \quad (1.17)$$

Let us start to look for a suitable position operator X . In classical probability, the position $\mathcal{X}$ of an object can be modeled as a random variable with distribution $\mathbb{P} \in \mathcal{P}(\mathbb{R}^d)$ (the set of Borel probability measures on $\mathbb{R}^d$) through the map $x \mapsto \mathcal{X}(x) = x$, defined on the probability space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P})$. If $\mathbb{P}$ is absolutely continuous with regards to the Lebesgue measure dx , it admits a density $\rho : \mathbb{R}^d \rightarrow [0, \infty)$ and we have that

$$\mathbb{P}(\mathcal{X} \in \Omega) = \int_{\Omega} dx \rho(x).$$

Comparing this with (1.12), it is natural to define $\mathcal{H} = L^2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), dx) = L^2(\mathbb{R}^d)$ and to define, for $i \in \{1, \dots, d\}$, the position coordinate operators $X_i : D_{X_i} \rightarrow L^2(\mathbb{R}^d)$ by

$$(X_i \psi)(x) = x_i \psi(x) \text{ for a.e. } x \in \mathbb{R}^d, \quad \psi \in D_{X_i} = \{\psi \in L^2(\mathbb{R}^d) : x \mapsto x_i \psi(x) \in L^2(\mathbb{R}^d)\}.$$

Then each X_i is self-adjoint by Theorem 1.1 and there exists a dense domain $D \subset L^2(\mathbb{R}^d)$, e.g. $D = \mathcal{S}(\mathbb{R}^d)$, such that for all $i, j \in \{1, \dots, d\}$, we have $D \subset D_{X_j}, X_i D \subset D_{X_j}$ and

$$[X_i, X_j]|_D = 0.$$

In this so called Schrödinger representation, the modulus square $|\psi|^2$ of an element $\psi \in L^2(\mathbb{R}^d)$ has the interpretation of a probability density. This means that $|\psi(x)|^2 dx$ describes the probability to find the particle near the point $x \in \mathbb{R}^d$.

The momentum operators P_j can certainly be no functions of $X = (X_1, \dots, X_d)$ and one may attempt instead to look for differential operators that satisfy (1.17). A quick computation shows that we can define $P_j = -i\hbar\partial_{x_j} : D_{P_j} \rightarrow L^2(\mathbb{R}^d)$ on the domain

$$D_{P_j} = \{\psi \in L^2(\mathbb{R}^d) : \partial_{x_j} \psi \in L^2(\mathbb{R}^d)\} = \{\psi \in L^2(\mathbb{R}^d) : p \mapsto p_j \widehat{\psi}(p) \in L^2(\mathbb{R}^d)\},$$

where ∂_{x_j} denotes the j -th distributional derivate in $\mathbb{R}^d$ and where $\widehat{\psi}$ denotes the $L^2(\mathbb{R}^d)$ -Fourier transform of ψ . Recall that for $\psi, \widehat{\psi} \in L^1 \cap L^2(\mathbb{R}^d)$, we have $\mathbb{R}^d$ -a.s. that

$$\widehat{\psi}(p) = \int_{\mathbb{R}^d} dx e^{-2\pi i p x} \psi(x), \quad \psi(x) = \int_{\mathbb{R}^d} dp e^{2\pi i p x} \widehat{\psi}(p). \quad (1.18)$$

Moreover, the map $L^2(\mathbb{R}^d) \ni \psi \mapsto \widehat{\psi} \in L^2(\mathbb{R}^d)$ is unitary and it maps $\mathcal{S}(\mathbb{R}^d)$ to itself.

It is straightforward to verify that the above choices imply that, on a suitable dense domain $D \subset L^2(\mathbb{R}^d)$ that is preserved by all the X_i, P_j (e.g. $D = \mathcal{S}(\mathbb{R}^d)$), we have that

$$[X_i, X_j] = 0, \quad [P_i, P_j] = 0, \quad [X_i, P_j] = i\hbar\delta_{ij}.$$

The relations (1.17) are called canonical commutation relations (CCR) and they have a fundamental physical consequence, called the Heisenberg uncertainty principle. For the following bounds, fix e.g. some normalized state $\psi \in \mathcal{S}(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$. Recalling the identity (1.14), notice that (1.17) and Cauchy-Schwarz imply that

$$\begin{aligned} \frac{\hbar}{2} &= \frac{1}{2} |\langle \psi, [X_j - \langle \psi, X_j \psi \rangle, P_j - \langle \psi, P_j \psi \rangle] \psi \rangle| \\ &\leq \langle \psi, (X_j - \langle \psi, X_j \psi \rangle)^2 \psi \rangle^{\frac{1}{2}} \langle \psi, (P_j - \langle \psi, P_j \psi \rangle)^2 \psi \rangle^{\frac{1}{2}} \\ &= (\langle \psi, X_j^2 \psi \rangle - \langle \psi, X_j \psi \rangle^2)^{\frac{1}{2}} (\langle \psi, P_j^2 \psi \rangle - \langle \psi, P_j \psi \rangle^2)^{\frac{1}{2}}. \end{aligned}$$

In other words, the variances of the position and momentum distributions of a particle cannot simultaneously be small, at the scale $\hbar \ll 1$. This means that if the position distribution of a particle is localized (the particle position is known), its momentum distribution is necessarily rather wide (the particle momentum is completely uncertain, because all momenta are equally likely to be measured). Obviously, this is in strong contrast to classical mechanics which does not imply such constraints. The smaller $\hbar$, the weaker the uncertainty constraints and in the semiclassical limit $\hbar \rightarrow 0$ one recovers classical behavior in the sense that the position and momentum operators commute. For an introductory discussion of basic, semiclassical results, see e.g. [7, Chapter 15].

Having set up a model for the position and momentum operators of a particle moving in $\mathbb{R}^d$, the Hamilton operator H that describes the energy of a particle of mass $m > 0$ in an external field $V : \mathbb{R}^d \rightarrow \mathbb{R}$ is, by analogy to (1.1), given by

$$H = \frac{|i\hbar\nabla|^2}{2m} + V(x) = \frac{\hbar^2}{2m}(-\Delta) + V(x).$$

Here and in the following, we write $V(x)$ for the multiplication operator $\psi \mapsto V\psi$ in $L^2(\mathbb{R}^d)$. Note that H is a well-defined, symmetric operator on $\mathcal{S}(\mathbb{R}^d)$. Under suitable conditions on V , it is essentially self-adjoint and thus admits a unique self-adjoint extension. See e.g. [22] for a thorough discussion of self-adjointness of Schrödinger operators.

It should be obvious how the above considerations can be generalized to formulate a quantum model of a particle moving in some subregion $U \subset \mathbb{R}^d$. Similarly, we can generalize Examples 1.2 and 1.3. If we want to describe a system of N particles of masses $m_j > 0$ that move in $\mathbb{R}^d$ and interact through a potential $V : \mathbb{R}^d \rightarrow \mathbb{R}$, a possible state space is $\mathcal{H} = L^2(\mathbb{R}^{dN}, \mathcal{B}(\mathbb{R}^{dN}), dx) = \bigotimes_{j=1}^N L^2(\mathbb{R}^d)$ and the Hamiltonian takes the form

$$H = \sum_{j=1}^N \frac{\hbar^2}{2m_j} (-\Delta_{x_j}) + \sum_{1 \leq i < j \leq N} V(x_i - x_j).$$

Similarly, if a particle moves in $\mathbb{R}^3$ in a static electromagnetic field that is generated by the electric and magnetic potentials $\Phi : \mathbb{R}^3 \rightarrow \mathbb{R}$ and, respectively, $A : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, its energy is described by a suitable self-adjoint extension of the operator

$$H = \frac{1}{2m} |i\hbar\nabla + \frac{e}{c}A(x)|^2 + e\Phi(x).$$

The previous examples were obtained via canonical quantization, described at the beginning of this section. Although one can not hope for an intuitive derivation of the quantization scheme, it is quite natural to wonder whether it allows for other choices of the position and momentum operators. The following fundamental result implies that the choice is essentially unique for systems of finitely many degrees of freedom.

Theorem 1.3 (Stone-von Neumann). *Let $(U_s)_{s \in \mathbb{R}^d}, (V_t)_{t \in \mathbb{R}^d}$ be two strongly continuous families of unitary operators on a separable Hilbert space $\mathcal{H}$ that satisfy the integrated canonical commutation relations (ICCR)*

$$U_s U_t = U_{s+t}, \quad V_s V_t = V_{s+t}, \quad V_t U_s = e^{is \cdot t} U_s V_t, \quad \forall s, t \in \mathbb{R}^d. \quad (1.19)$$

Assume, moreover, that $\mathcal{H}$ has no non-trivial closed subspace D (i.e., $D \neq \{0\}$ and $D \neq \mathcal{H}$) that is invariant under $(U_s)_{s \in \mathbb{R}^d}$ and $(V_t)_{t \in \mathbb{R}^d}$ and call such a pair an irreducible realization of the ICCR. Then, any two non-trivial irreducible realizations of the ICCR are unitarily equivalent.

Problem 1.2. *Suppose that $A = A^*, B = B^* \in \mathcal{L}(\mathcal{H})$. Show that A and B satisfy the commutation relation $[A, B] = i$ if and only if $e^{isB} e^{itA} = e^{its} e^{itA} e^{isB}$ for every $s, t \in \mathbb{R}$.*

Remark 1.5. *The previous problem motivates why we can view the relations (1.19) as an integrated version of (1.17) for the strongly continuous families of unitary maps*

$$U_s = e^{is \cdot X}, \quad V_t = e^{it \cdot P/\hbar}, \quad \forall s, t \in \mathbb{R}^d.$$

The integrated variant (1.19) of the CCR is easier to analyze, since all involved operators are bounded so that no domain issues (due to the unboundedness of X and P) arise.

Proof of Theorem 1.3. We prove the theorem for $d = 1$. The general case follows along the same lines and is left as an *exercise* for the interested reader.

In order to prove the theorem, it suffices to consider the case $\mathcal{H} = L^2(\mathbb{R})$ (any two separable Hilbert spaces are unitarily equivalent) and to show that every irreducible realization $(U'_s)_{s \in \mathbb{R}}, (V'_t)_{t \in \mathbb{R}}$ of (1.19) is equivalent to the canonical one, defined by

$$U_s = e^{isX}, \quad V_t = e^{itP/\hbar}, \quad \forall s, t \in \mathbb{R}.$$

To see that this defines indeed an irreducible realization of the integrated canonical commutation relations (1.19), note first of all that for every $\varphi, \psi \in L^2(\mathbb{R})$, we have that

$$\langle \varphi, V_t \psi \rangle = \int_{\mathbb{R}} dp \, \widehat{\overline{\varphi}}(p) e^{2\pi i t p} \widehat{\psi}(p) = \int_{\mathbb{R}} dx \, \overline{\varphi}(x) \psi(x+t) = \langle \varphi, \psi(\cdot + t) \rangle,$$

so that $(V_t \psi)(\cdot) = \psi(\cdot + t)$, for every $t \in \mathbb{R}$. Based on this observation, one readily verifies the identities (1.19), the non-trivial case being

$$(V_t U_s \psi)(\cdot) = V_t (e^{is \cdot} \psi(\cdot)) = e^{ist} e^{is \cdot} \psi(\cdot + t) = e^{ist} (U_s V_t \psi)(\cdot), \quad \forall \psi \in L^2(\mathbb{R}).$$

Furthermore, the pair $(U_s)_{s \in \mathbb{R}^d}$ and $(V_t)_{t \in \mathbb{R}^d}$ is irreducible. Indeed, suppose there is a non-trivial closed subspace $D \subset L^2(\mathbb{R})$ that is both invariant under all maps U_t and all

maps V_s , for $s, t \in \mathbb{R}$. Then we can pick a normalized element $\varphi \in D$ and consider a vector $\psi \in D^\perp$ in the orthogonal complement of D . By the invariance of D under all maps U_s and V_t , we find that for every $\zeta \in \mathcal{S}(\mathbb{R})$, it holds true that

$$\int_{\mathbb{R}} dx \widehat{\zeta}(x) \overline{\psi}(x) \varphi(x) = \int_{\mathbb{R}} dp \zeta(p) \left(\int_{\mathbb{R}} dx \overline{\psi}(x) e^{-2\pi i x p} \varphi(x) \right) = 0.$$

A standard density argument implies that $|\psi(x)\varphi(x)| = 0$ for *a.e.* $x \in \mathbb{R}$. Since $V_t\varphi \in D$, the same argument implies $|\psi(x)\varphi(x+t)| = 0$ *a.s.* in $\mathbb{R}$, for every $t \in \mathbb{R}$. But then

$$0 = \int_{\mathbb{R}^2} dx dy |\psi(x)\varphi(x-y)| \zeta(y) = \int_{\mathbb{R}} dx |\psi(x)| (|\varphi| * \zeta)(x)$$

for every $\zeta \in \mathcal{S}(\mathbb{R})$. Choosing a strictly positive $\zeta \in \mathcal{S}(\mathbb{R})$, so that $\zeta * |\varphi| > 0$ everywhere since $\|\varphi\| = 1$, we conclude $|\psi(x)| = 0$ for *a.e.* $x \in \mathbb{R}$ and thus $\psi = 0 \in L^2(\mathbb{R})$. This means that $D = \overline{D} = \mathcal{H}$ and it proves that $(U_s)_{s \in \mathbb{R}}, (V_t)_{t \in \mathbb{R}}$ is irreducible. We leave it as an *exercise* to show that a simple generalization of the previous analysis implies that the realization of the ICCR from Remark 1.5 is irreducible, for every $d \in \mathbb{N}$.

In order to proceed with the proof, it turns out convenient to set

$$W_{s,t} = e^{\frac{ist}{2}} U_t V_s = e^{-\frac{its}{2}} V_s U_t, \quad W'_{s,t} = e^{\frac{ist}{2}} U'_t V'_s = e^{-\frac{its}{2}} V'_s U'_t,$$

so that our goal is equivalent to finding a unitary map $\mathcal{U} : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ that satisfies

$$\mathcal{U} W_{s,t} \mathcal{U}^* = W'_{s,t}, \quad \forall s, t \in \mathbb{R}. \quad (1.20)$$

Now, observe that for every normalized $\psi, \psi' \in L^2(\mathbb{R})$, the linear spaces

$$D_\psi = \text{span}\{W_{s,t}\psi : s, t \in \mathbb{R}\}, \quad D'_{\psi'} = \text{span}\{W'_{s,t}\psi' : s, t \in \mathbb{R}\}$$

are dense in $L^2(\mathbb{R})$, because they are invariant under the irreducible realizations of the ICCR. To construct $\mathcal{U}$, we look for two normalized elements $\psi, \psi' \in L^2(\mathbb{R})$ that satisfy

$$\langle \psi, W_{s,t}\psi \rangle = \langle \psi', W'_{s,t}\psi' \rangle, \quad \forall s, t \in \mathbb{R}. \quad (1.21)$$

Once such elements ψ, ψ' are found, we conclude from (1.19) that

$$\begin{aligned} \left\langle \sum_{j=1}^n \lambda_j W_{s_j, t_j} \psi, \sum_{k=1}^m \mu_k W_{s_k, t_k} \psi \right\rangle &= \sum_{j=1}^n \sum_{k=1}^m \bar{\lambda}_k \mu_k \langle \psi, W_{s_j, t_j}^* W_{s_k, t_k} \psi \rangle \\ &= \sum_{j=1}^n \sum_{k=1}^m \bar{\lambda}_k \mu_k \langle \psi, W_{-s_j, -t_j} W_{s_k, t_k} \psi \rangle \\ &= \sum_{j=1}^n \sum_{k=1}^m \bar{\lambda}_k \mu_k e^{-\frac{i}{2}(s_j t_k - s_k t_j)} \langle \psi, W_{s_k - s_j, t_k - t_j} \psi \rangle \\ &= \left\langle \sum_{j=1}^n \lambda_j W'_{s_j, t_j} \psi', \sum_{k=1}^m \mu_k W'_{s_k, t_k} \psi' \right\rangle. \end{aligned}$$

This and the density of $D_\psi, D_{\psi'}$ in $L^2(\mathbb{R})$ imply that $\mathcal{U} : D_\psi \mapsto D_{\psi'}$, defined by

$$\mathcal{U}\left(\sum_{k=1}^n \lambda_k W_{s_k, t_k} \psi\right) = \sum_{k=1}^n \lambda_k W'_{s_k, t_k} \psi', \quad \forall \lambda_k \in \mathbb{C}, s_k, t_k \in \mathbb{R}, k \in \mathbb{N},$$

is an isometry and extends to a unitary map that satisfies, by definition, (1.20).

It thus remains to find two vectors $\psi, \psi' \in L^2(\mathbb{R})$, $\|\psi\| = \|\psi'\| = 1$, that satisfy (1.21). Finding such elements is not obvious and to explain the key idea, let us rewrite

$$\langle \psi, W_{s,t} \psi \rangle = \text{tr} |\psi\rangle \langle \psi| W_{s,t}.$$

To find suitable ψ and ψ' , it would be useful if the trace on the r.h.s. would only involve the operators $(W_{s,t})_{s,t \in \mathbb{R}}$ and this suggests to look for averages of the form

$$|\psi\rangle \langle \psi| = \int_{\mathbb{R}^2} ds dt \varphi(s, t) W_{s,t} \quad (1.22)$$

for some $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$, say $\varphi \in \mathcal{S}(\mathbb{R}^2)$. In particular, this would imply that

$$\psi(x) \int_{\mathbb{R}} ds \bar{\psi}(s) f(s) = \int_{\mathbb{R}^2} ds dt e^{\frac{i}{2}st + itx} \varphi(s, t) f(x+s) = \int_{\mathbb{R}^2} ds dt e^{\frac{i}{2}(s+x)t} \varphi(s-x, t) f(s)$$

for every $f \in \mathcal{S}(\mathbb{R})$ and for *a.e.* $x \in \mathbb{R}$. Recalling the well-known fact that the Fourier transform of a Gaussian density remains a Gaussian, that is

$$\int_{\mathbb{R}} dx e^{-ax^2 + bx} = \sqrt{\frac{\pi}{a}} e^{\frac{b^2}{4a}}, \quad \forall a \in (0, \infty), b \in \mathbb{C},$$

it is left as a *exercise* to verify that (1.22) is indeed satisfied for the explicit choice

$$x \mapsto \psi(x) = \pi^{-\frac{1}{4}} e^{-\frac{1}{2}x^2} \in \mathcal{S}(\mathbb{R}), \quad (s, t) \mapsto \varphi(s, t) = \frac{1}{2\pi} e^{-\frac{1}{2}(s^2 + t^2)} \in \mathcal{S}(\mathbb{R}^2). \quad (1.23)$$

Similarly, it is left as an elementary *exercise* to verify that for the choice (1.23), we have

$$\langle \psi, W_{s,t} \psi \rangle = e^{-\frac{1}{4}(s^2 + t^2)}, \quad \forall s, t \in \mathbb{R}.$$

Now, to find a corresponding state $\psi' \in L^2(\mathbb{R})$, consider by analogy to (1.22) the operator

$$P' = \int_{\mathbb{R}^2} ds dt \varphi(s, t) W'_{s,t} \in \mathcal{L}(L^2(\mathbb{R}))$$

whose boundedness is a direct consequence of the integrability of φ and the unitarity of U'_s, V'_t . Recalling that $(W'_{s,t})^* = W'_{-s, -t}$ and that $\varphi(s, t) = \varphi(-s, -t)$, we find that $P' = (P')^*$ is self-adjoint. Motivated by (1.22), we may expect that $P' \neq 0$ is a non-trivial projection. To see this, assume first by contradiction that $P' = 0$. Then, using

$$W'_{-s_1, -t_1} W'_{s_2, t_2} W'_{s_1, t_1} = e^{i(s_1 t_2 - s_2 t_1)} W'_{s_2, t_2},$$

we conclude that for every $\zeta \in \mathcal{S}(\mathbb{R}^2)$, we also have that

$$\begin{aligned} \int_{\mathbb{R}^2} ds dt \zeta(s, t) \varphi(s, t) W'_{s,t} &= \frac{1}{4\pi^2} \int_{\mathbb{R}^4} ds_1 dt_1 ds_2 dt_2 \widehat{\zeta}(s_1, t_1) e^{is_1 s_2 + it_1 t_2} \varphi(s_2, t_2) W'_{s_2, t_2} \\ &= -\frac{1}{4\pi^2} \int_{\mathbb{R}^2} ds dt \widehat{\zeta}(-t, s) W'_{-s, -t} P' W'_{s,t} = 0. \end{aligned}$$

But $\varphi > 0$ in $\mathbb{R}^2$, so pick e.g. a non-negative bump function $\zeta_\epsilon \in C_c^\infty(B_\epsilon(0))$ which satisfies $\int_{\mathbb{R}^2} ds dt \zeta_\epsilon(s, t) \varphi(s, t) = 1$. Moreover, pick a normalized $\tau' \in L^2(\mathbb{R}^2)$. Then, by the strong continuity of the map $(s, t) \mapsto W_{s,t}$, we find for $\epsilon > 0$ small enough that

$$\left\| \int_{\mathbb{R}^2} ds dt \zeta_\epsilon(s, t) \varphi(s, t) W_{s,t} \tau' - \tau' \right\| \leq \sup_{|s'|, |t'| \leq \epsilon} \|W_{s',t'} \tau' - \tau'\| \int_{\mathbb{R}^2} ds dt \zeta_\epsilon(s, t) \varphi(s, t) < 1$$

and thus $\int_{\mathbb{R}^2} ds dt \zeta_\epsilon(s, t) \varphi(s, t) W_{s,t} \neq 0$, a contradiction. This shows that $P' \neq 0$. To see that $(P')^2 = P'$ is a projection, we compute explicitly that

$$\begin{aligned} (P')^2 &= \int_{\mathbb{R}^4} ds_1 dt_1 ds_2 dt_2 \varphi(s_1, t_1) \varphi(s_2, t_2) W'_{s_1, t_1} W'_{s_2, t_2} \\ &= \int_{\mathbb{R}^4} ds_1 dt_1 ds_2 dt_2 \varphi(s_1, t_1) \varphi(s_2, t_2) e^{\frac{i}{2}(s_1 t_2 - s_2 t_1)} W'_{s_1 + s_2, t_1 + t_2} \\ &= \int_{\mathbb{R}^4} ds dt ds' dt' \varphi(s - s', t - t') \varphi(s', t') e^{\frac{i}{2}((s-s')t' - s'(t-t'))} W'_{s,t} \\ &= \int_{\mathbb{R}^2} ds dt \varphi(s, t) W'_{s,t} \left(\int_{\mathbb{R}^2} \frac{ds' dt'}{2\pi} e^{-(s')^2 - (t')^2 + \frac{1}{2}(s-it)s' + \frac{i}{2}(s-it)t'} \right) = P'. \end{aligned}$$

More generally, motivated once again by (1.22), a similar computation (*exercise*) yields

$$P' W'_{s,t} P' = P' e^{-\frac{1}{4}(s^2 + t^2)} P' = e^{-\frac{1}{4}(s^2 + t^2)} P'.$$

Given these observations, we conclude the proof by finding some normalized vector $\psi' = P(\varphi')$, $\|\psi'\| = 1$, for some $0 \neq \varphi' \in L^2(\mathbb{R})$, so that $P'\psi' = \psi'$ and thus

$$\langle \psi', W'_{s,t} \psi' \rangle = \langle \psi', P' W'_{s,t} P' \psi' \rangle = e^{-\frac{1}{4}(s^2 + t^2)} = \langle \psi, W_{s,t} \psi \rangle, \quad \forall s, t \in \mathbb{R}.$$

This proves (1.21) which, as explained earlier, implies the unitary equivalence (1.20). $\square$

The canonical quantization scheme has been quite important historically, because it leads to natural candidates for quantum variants of well-known classical models in a relatively straightforward way. This applies in particular to the development of quantum field theories, which describe quantum versions of classical fields. A classical field Φ , such as the electromagnetic field, can be understood as a collection of observables $(\Phi_x)_{x \in U \subset \mathbb{R}^3}$ labelled by the space coordinates (Φ is typically tensor-valued, e.g. scalar- or vector-valued). In this sense, it consists of an infinite number of degrees of freedom. When applying the canonical quantization scheme to it, one associates, loosely speaking, to each observable Φ_x a pair of classical canonical variables that satisfy (1.8) (like position

and momentum in case of a single particle) and replaces it with a corresponding pair of operators that satisfy the canonical commutation relations (1.17). Theorem 1.3 shows that in case of finitely many degrees of freedom (including e.g. all interacting many-body systems of finitely many particles), this procedure yields a quantum model which is essentially unique. In the remaining part of this section we explain that, in contrast to the finite particle case, for an infinite number of degrees of freedom, uniqueness is lost.

For a simple motivation that suggests the loss of uniqueness for infinitely many variables, consider the pair $(U_s)_{s \in \mathbb{R}^d}, (V_t)_{t \in \mathbb{R}^d}$ as in Remark 1.5 and set for $\sigma > 0$

$$U'_s = e^{is \cdot X/\sigma} = U_{s/\sigma}, \quad V'_t = e^{i\sigma t \cdot P/\hbar} = V_{\sigma t}, \quad \forall s, t \in \mathbb{R}^d.$$

With analogous notation as in the proof of Theorem 1.3, a basic change of variables shows that $\mathcal{U}W_{s,t}\mathcal{U}^* = W'_{s,t}$ for $\mathcal{U} : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)$, defined by

$$(\mathcal{U}\varphi)(\cdot) = \sigma^{-\frac{d}{2}}\varphi(\cdot/\sigma).$$

In other words, rescaling by $\sigma > 0$ yields a unitarily equivalent, irreducible realization of the ICCR (1.19), for every $d \in \mathbb{N}$. Heuristically, we may expect the loss of unitary equivalence if $d = \infty$, in which case the above expression for $\mathcal{U}$ does no longer make sense. To set up this observation rigorously, we first need to define a reasonable version of $L^2(\prod_{j \in \mathbb{N}} \mathbb{R})$ that generalizes in a suitable sense the Lebesgue spaces $L^2(\mathbb{R}^d)$, for $d \in \mathbb{N}$. To this end, observe that for every $d \in \mathbb{N}$, the map

$$\mathcal{V}_d : L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mu^{\otimes d}), \quad \mu(dx) = \frac{1}{\sqrt{\pi}} e^{-x^2} dx, \quad (\mathcal{V}_d \varphi)(\cdot) = \pi^{\frac{1}{4}} e^{\frac{1}{2}|\cdot|^2} \varphi(\cdot)$$

is a unitary map and that $\mu^{\otimes d}$ is a Gaussian probability measure on $\mathbb{R}^d$.

Problem 1.3. *Let $d \in \mathbb{N}$ and set $U'_s = \mathcal{V}_d U_s \mathcal{V}_d^*$, $V'_t = \mathcal{V}_d V_t \mathcal{V}_d^*$ and $W'_{s,t} = \mathcal{V}_d e^{\frac{i}{2}s \cdot t} U_t V_s \mathcal{V}_d^*$ as well as $x \mapsto \psi'(x) = 1 \in L^2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mu^{\otimes d})$. Prove that*

$$\begin{aligned} (U'_s \varphi)(x) &= e^{is \cdot x} \varphi(x), \\ (V'_t \varphi)(x) &= e^{-t \cdot x - \frac{1}{2}|t|^2} \varphi(x + t), \\ \langle \psi', W'_{s,t} \psi' \rangle &= e^{-\frac{1}{4}(|s|^2 + |t|^2)}, \end{aligned}$$

for every $s, t \in \mathbb{R}^d$, $\varphi \in L^2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mu^{\otimes d})$ and a.e. $x \in \mathbb{R}^d$.

Problem 1.3 suggests that instead of considering the canonical models in the different spaces $L^2(\mathbb{R}^d)$ for each fixed $d \in \mathbb{N}$, we can set up a model with an infinite number of degrees of freedom into which one can naturally embed all finite particle models. More precisely, applying the Kolmogorov extension theorem, see e.g. [4, Chapter 10], we define μ_∞ as the unique (regular Borel, see [3, Chapter 1]) probability measure on $(\Omega, \mathcal{B}(\Omega))$ for $\Omega = \prod_{j=1}^\infty \mathbb{R}$ (equipped with the product topology, which is metrizable) such that

$$\int_\Omega \mu_\infty(dx) f(x) = \int_\Omega \mu_\infty(dx) f(x_1, x_2, \dots, x_d) = \int_{\mathbb{R}^d} \mu^{\otimes d}(dx) f(x_1, x_2, \dots, x_d)$$

for every integrable $f : \Omega \rightarrow \mathbb{C}$ that only depends on the first $d \in \mathbb{N}$ coordinates in Ω . Then, identifying for $d \in \mathbb{N}$ an element $\varphi \in L^2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ with the function $x \mapsto (\varphi \otimes 1)(x) = \varphi(x_1, \dots, x_d) \in L^2(\Omega, \mathcal{B}(\Omega), \mu_\infty)$, we have the isometric embeddings

$$L^2(\mu^{\otimes d}) \hookrightarrow L^2(\mu^{\otimes d+1}) \hookrightarrow L^2(\mu^{\otimes d+2}) \hookrightarrow \dots \hookrightarrow L^2(\mu_\infty)$$

for $L^2(\mu^{\otimes d}) = L^2(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mu^{\otimes d})$ and $L^2(\mu_\infty) = L^2(\Omega, \mathcal{B}(\Omega), \mu_\infty)$. Moreover, since integrable functions in $(\Omega, \mathcal{B}(\Omega), \mu_\infty)$ can be approximated up to arbitrarily small errors by linear combinations of characteristic functions of open sets in Ω (equipped with the product topology admitting a countable basis for its topology), we verify that

$$L^2_{<\infty} = \bigcup_{d=1}^{\infty} L^2(\mu^{\otimes d}) \subset L^2(\mu_\infty)$$

is dense in $L^2(\mu_\infty)$. As a consequence, the canonical projection $\Pi_d : L^2(\mu_\infty) \rightarrow L^2(\mu^{\otimes d})$ satisfies $\lim_{d \rightarrow \infty} \|\psi - \Pi_d \psi\| = 0$, for every $\psi \in L^2(\mu_\infty)$. We leave the careful proof of the last two statements as an *exercise*. In view of Problem 1.3, we then define

$$(U_s \varphi)(x) = e^{is \cdot x} \varphi(x), \quad (V_t \varphi)(x) = e^{-t \cdot x - \frac{1}{2}|t|^2} \varphi(x+t), \quad (W_{s,t} \varphi)(x) = e^{\frac{is \cdot t}{2}} (U_t V_s \varphi)(x)$$

for every $s, t \in \mathbb{R}^{<\infty} = \bigcup_{d \in \mathbb{N}} \mathbb{R}^d \hookrightarrow \Omega$ and $\varphi \in L^2(\mu_\infty)$. Notice that, heuristically, the pair $(U_s)_{s \in \mathbb{R}^{<\infty}}, (V_t)_{t \in \mathbb{R}^{<\infty}}$ corresponds to a countably infinite sequence of canonical position and momentum operators $(X_j, P_j)_{j \in \mathbb{N}}$ that satisfy the CCR (1.17).

As in the proof of Theorem 1.3, we obtain for every $s, t \in \mathbb{R}^{<\infty}$ that

$$U_s U_t = U_{s+t}, \quad V_s V_t = V_{s+t}, \quad V_t U_s = e^{is \cdot t} U_s V_t \quad (1.24)$$

and that

$$\langle \psi, W_{s,t} \psi \rangle = e^{-\frac{1}{4}(|s|^2 + |t|^2)}.$$

Here, $x \mapsto \psi(x) = 1 \in L^2(\mu_\infty)$ with $\|\psi\| = 1$. Since $(U_s)_{s \in \mathbb{R}^{<\infty}}, (V_t)_{t \in \mathbb{R}^{<\infty}}$ define isometries on $L^2_{<\infty}$, they must be unitary on $L^2(\mu_\infty)$ (alternatively, unitarity can be checked explicitly based on (1.24)). The pair $(U_s)_{s \in \mathbb{R}^{<\infty}}, (V_t)_{t \in \mathbb{R}^{<\infty}}$ is also irreducible. Indeed, assuming $D \subset L^2(\mu_\infty)$ to be a closed, non-empty invariant subspace, we can pick $\varphi \in D$, $\|\varphi\| = 1$ and $\psi \in D^\perp$, as in the proof of Theorem 1.3. As before

$$0 = \int \mu_\infty(dx) \zeta(x) \bar{\psi}(x) V_t \varphi(x)$$

for every $\zeta \in \mathcal{S}(\mathbb{R}^d) \hookrightarrow L^2_{<\infty}$ and every $t \in \mathbb{R}^{<\infty}$ so that $|\psi(x) e^{-t \cdot x - |t|^2/2} \varphi(x+t)| = 0$ and hence $|\psi(x) \varphi(x+t)| = 0$ a.s. in Ω . Now, note that for $d \in \mathbb{N}$ sufficiently large and $0 \neq \xi \in L^2(d\mu_\infty)$ we have explicitly

$$(\Pi_d \xi)(x, X) = \frac{\int_{\Omega_{>d}} \mu^{\otimes >d}(dX) \xi(x, X)}{\left| \int_{\mathbb{R}^d} \mu^{\otimes d}(dx) \int_{\Omega_{>d}} \mu^{\otimes >d}(dX) \xi(x, X) \right|^2}$$

for *a.e.* $x \in \mathbb{R}^d$, $\Omega_{>d} = \prod_{j=d+1}^{\infty} \mathbb{R}$ and $\mu^{\otimes >d}$ defined such that $\mu_{\infty} = \mu^{\otimes d} \otimes \mu^{\otimes >d}$ (based on another application of Kolmogorov's theorem). If we assume by contradiction that $\psi \neq 0$, Cauchy-Schwarz implies for suitable $C > 0$ (by normalization) that

$$\begin{aligned} |\langle \Pi_d \psi, \Pi_d U_s V_t \varphi \rangle| &\leq C \int_{\mathbb{R}^d} \mu^{\otimes d}(dx) \left| \int_{\Omega_{>d}} \mu^{\otimes >d}(dX) \psi(x, X) \int_{\Omega_{>d}} \mu^{\otimes >d}(dY) V_t \varphi(x, Y) \right| \\ &\leq C \int_{\mathbb{R}^d} \mu^{\otimes d}(dx) \langle |\psi|(x, \cdot), 1 \rangle_{L^2(\mu^{\otimes >d})} \langle 1, |U_s V_t \varphi|(x) \rangle_{L^2(\mu^{\otimes >d})} \\ &\leq C \int_{\Omega} \mu_{\infty}(dx) e^{-t \cdot x - \frac{1}{2}|t|^2} |\psi(x) \varphi(x+t)| = 0 \end{aligned}$$

for every $t \in \mathbb{R}^{<\infty}$. But now arguing as in the finite dimensional case (using that U_s and V_t commute with Π_d if $s, t \in \mathbb{R}^d$), we conclude that $\Pi_d \psi = 0 \in L^2(\mu^{\otimes d})$ for every $d \in \mathbb{N}$. This shows $\psi = 0 \in L^2(\mu_{\infty})$, a contradiction, and the irreducibility follows.

Now, note that for every other irreducible pair $(U'_s)_{s \in \mathbb{R}^{<\infty}}, (V'_t)_{t \in \mathbb{R}^{<\infty}}$ that satisfies (1.24) and that is unitarily equivalent to $(U_s)_{s \in \mathbb{R}^{<\infty}}, (V_t)_{t \in \mathbb{R}^{<\infty}}$ (via some map $\mathcal{U}$, say), there must exist a normalized state $\psi' (= \mathcal{U}\psi) \in L^2(\mu_{\infty})$ such that

$$\langle \psi', W'_{s,t} \psi' \rangle = e^{-\frac{1}{4}(|s|^2 + |t|^2)}$$

for every $s, t \in \mathbb{R}^{<\infty}$. Our previous considerations then suggest to fix $\sigma > 0$ and to set

$$U'_s = U_{s/\sigma}, \quad V'_t = V_{\sigma t}.$$

Clearly, $(U'_s)_{s \in \mathbb{R}^{<\infty}}, (V'_t)_{t \in \mathbb{R}^{<\infty}}$ satisfies (1.24) and is irreducible. But this pair can not be unitarily equivalent to $(U_s)_{s \in \mathbb{R}^{<\infty}}, (V_t)_{t \in \mathbb{R}^{<\infty}}$. Otherwise, there exists some $\psi' \in L^2(\mu_{\infty})$, $\|\psi'\| = 1$, such that $\langle \psi', V'_{t_d} \psi' \rangle = e^{-\frac{1}{4}}$ for every $t_d = (0, \dots, 0, 1, 0, \dots)$ (the non-zero entry being at the d -th slot) and every $d \in \mathbb{N}$. But if $\psi'_d = \Pi_d \psi'$ is the projection of ψ' onto $L^2(\mu^{\otimes d}) \hookrightarrow L^2(\mu_{\infty})$, we infer from a small variation (*exercise*) of Problem 1.3 that

$$\langle \psi'_d, V'_{t_{d+1}} \psi'_d \rangle = e^{-\frac{\sigma^2}{4}} \|\psi'_d\| \rightarrow e^{-\frac{\sigma^2}{4}}$$

as $d \rightarrow \infty$. On the other hand, $\langle \psi', V'_{t_{d+1}} \psi' \rangle = e^{-\frac{1}{4}}$ and strong continuity implies

$$|\langle \psi', V'_{t_{d+1}} \psi' \rangle - \langle \psi'_d, V'_{t_{d+1}} \psi'_d \rangle| \rightarrow 0$$

as $d \rightarrow \infty$. Thus, we obtain for every $\sigma \neq 1$ the contradiction that $e^{-\frac{\sigma^2}{4}} = e^{-\frac{1}{4}}$. In these cases, the pair $(U'_s)_{s \in \mathbb{R}^{<\infty}}, (V'_t)_{t \in \mathbb{R}^{<\infty}}$ is not unitarily equivalent to $(U_s)_{s \in \mathbb{R}^{<\infty}}, (V_t)_{t \in \mathbb{R}^{<\infty}}$.

The previous counter examples to uniqueness rely on a simple scaling argument so that, heuristically, also the corresponding position and momentum operators are merely rescaled versions of the canonical ones. For a more detailed discussion including a characterization of unitary equivalence in the setting of $L^2(\mu_{\infty})$, see [31, Section C.4].

1.2.2 Symmetries and Their Representations

Following [33, Section 2.2], a symmetry transformation is a transformation that does not change the physics of a system, but rather our point of view of how to describe it. Basic examples include Euclidean transformations (such as rotations) that describe the change of the coordinate system used to measure e.g. the position of a particle. In the previous sections, we learned that physical observables are described mathematically through self-adjoint operators, based on the spectral theorem. How do we describe symmetries in the context of quantum mechanics? This section discusses this question based on a result due to Wigner which says that symmetries can always be modeled by unitary operators on the Hilbert space that contains the physical states.

Let us first make precise what we mean by a symmetry transformation. Suppose that we consider a quantum system whose state is described by a normalized wave function $\psi \in \mathcal{H}$ in a Hilbert space $\mathcal{H}$. The probabilistic interpretation of quantum mechanics, summarized in (1.12), implies that all probabilities one may potentially be interested in reduce to the computation of squares of inner products

$$|\langle \varphi, \psi \rangle|^2 \quad \text{for } \varphi, \psi \in \mathcal{H}, \|\varphi\| = \|\psi\| = 1. \quad (1.25)$$

Assume that T describes a symmetry transformation. What are reasonable conditions to impose on T ? Since ψ and $e^{i\omega}\psi$, for $\omega \in [0, 2\pi)$, describe the same physical state, it is natural to assume that $T : \mathcal{R}_1 \rightarrow \mathcal{R}_1$ is a map on the space of unit rays. Here, we set

$$\mathcal{R}_1 = \bigcup_{\psi \in \mathcal{H}, \|\psi\|=1} \{\varphi \in \mathcal{H} : \varphi \sim \psi\} = \bigcup_{\psi \in \mathcal{H}, \|\psi\|=1} [\psi],$$

recalling that $\psi_1 \sim \psi_2$ if and only if $\psi_1 = e^{i\omega}\psi_2$ for some $\omega \in [0, 2\pi)$. This means that T maps one physical state to another physical state. Furthermore, T must not influence measurement outcomes which means that

$$T[\varphi] \cdot T[\psi] = [\varphi] \cdot [\psi] \quad (1.26)$$

for every $[\varphi], [\psi] \in \mathcal{R}_1$, where we set $[\varphi] \cdot [\psi] = |\langle \varphi, \psi \rangle|$. It is left as a basic *exercise* to show that (1.26) implies that T is injective. Since we can usually switch from one coordinate system to another and back, it is also reasonable to assume that T is onto.

In the sequel, we call a map $T : \mathcal{R}_1 \rightarrow \mathcal{R}_1$ with the previous properties a symmetry transformation. Notice in particular that every linear, unitary and every antilinear, antiunitary operator on $\mathcal{H}$ induces a symmetry transformation. It is a fundamental result that these operators constitute all possible symmetry transformations.

Theorem 1.4 (Wigner). *Let $T : \mathcal{R}_1 \rightarrow \mathcal{R}_1$ be a symmetry transformation. Then there exists a linear, unitary or an antilinear, antiunitary operator $U : \mathcal{H} \rightarrow \mathcal{H}$ such that $[U\psi] = T[\psi]$ for every $\psi \in \mathcal{H}$, $\|\psi\| = 1$. In particular, T is equal to the symmetry transformation that is induced by U . If $\dim \mathcal{H} \geq 2$, the map $U : \mathcal{H} \rightarrow \mathcal{H}$ is unique up to multiplication by a constant of modulus one.*

Proof. We follow [2] and begin with a few preliminary remarks.

We assume throughout the proof that $\dim \mathcal{H} \geq 3$ (see Problem 1.4 for the remaining, simpler cases). Since the equivalence $\varphi \sim \psi$ for two non-zero vectors implies that $[\varphi/\|\varphi\|] = [\psi/\|\psi\|] \in \mathcal{R}_1$, we can extend $T : \mathcal{R} \rightarrow \mathcal{R}$ via $T[\psi] = \|\psi\|T[\psi/\|\psi\|]$ for every non-zero ray $\psi \in \mathcal{R} = \mathcal{H}/\sim$. Then (1.26) remains valid for every $\varphi, \psi \in \mathcal{R}$.

Next, observe that if $([\psi]_i)_{i=1}^n$ is a sequence of orthonormal rays so that $[\psi]_i \cdot [\psi]_j = \delta_{ij}$, then we have by (1.26) for every pair of representatives $\psi_i \in [\psi]_i, \psi_j \in [\psi]_j$ and $\psi'_i \in T[\psi]_i, \psi'_j \in T[\psi]_j$ that

$$|\langle \psi_i, \psi_j \rangle| = |\langle \psi'_i, \psi'_j \rangle| = \delta_{ij} = \langle \psi_i, \psi_j \rangle = \langle \psi'_i, \psi'_j \rangle.$$

As a consequence, if $\varphi = \sum_{i=1}^n \lambda_i \psi_i (= \sum_{i=1}^n \langle \psi_i, \varphi \rangle \psi_i)$, then we have for every $\varphi' \in T[\varphi]$

$$\begin{aligned} \left\| \varphi' - \sum_{i=1}^n \langle \psi'_i, \varphi' \rangle \psi'_i \right\|^2 &= \|\varphi'\|^2 - \sum_{i=1}^n |\langle \psi'_i, \varphi' \rangle|^2 = (T[\varphi] \cdot T[\varphi]) - \sum_{i=1}^n (T[\psi_i] \cdot T[\varphi])^2 \\ &= \left\| \varphi - \sum_{i=1}^n \langle \psi_i, \varphi \rangle \psi_i \right\|^2 = 0. \end{aligned}$$

In other words, $\varphi' = \sum_{i=1}^n \lambda'_i \psi'_i$ for constants $|\lambda'_i| = |\lambda_i|$ for each $i \in \{1, \dots, n\}$. Notice that this looks already close to linearity. Loosely speaking, our goal is to make sure that the transformed phases λ'_i can be chosen in a consistent (linear or antilinear) way.

Now, let's begin to construct the map U . This is done in three main steps. In the first step, we have some freedom how to set the direction of U on a fixed unit vector. To be more precise, let $[e] \in \mathcal{R}_1$ and choose some $e \in [e], e' \in T[e] \in \mathcal{R}_1$. Then we may set

$$Ue = e'. \quad (1.27)$$

In the second step, we extend U to $\{e\}^\perp$ in a way that is consistent with (1.27) and with the statement of the theorem. To this end, consider $\psi = e + \varphi$ for some $0 \neq \varphi \in \{e\}^\perp$. Then, we can write equivalently $\psi = \lambda_1 e + \lambda_2 e_\varphi$ for the normalized vector $e_\varphi = \varphi/\|\varphi\| \in \{e\}^\perp$, $\lambda_1 = 1$ and $\lambda_2 = \|\varphi\|$. Fixing some $e'_\varphi \in T[e_\varphi]$, we know based on the preliminary remarks that for every $\psi' \in T[\psi]$, we have that

$$\psi' = \lambda'_1 e' + \lambda'_2 e'_\varphi$$

for some $\lambda'_1, \lambda'_2 \in \mathbb{C}$ with $|\lambda'_1| = 1, |\lambda'_2| = \|\varphi\|$. In particular, by orthogonality of e' and e'_φ , there is one and only one element $\psi' = e' + \lambda'_2 e'_\varphi \in T[\psi]$ such that $\lambda'_1 = 1, |\lambda'_2| = \|\varphi\|$. Based on this specific choice (which determines λ'_2), we define

$$U_\perp \varphi = \lambda'_2 e'_\varphi \in T[e_\varphi], \quad U(e + \varphi) = e' + U_\perp \varphi \in T[e + \varphi], \quad U\varphi = U_\perp \varphi \in T[e_\varphi]. \quad (1.28)$$

This implies in particular the additivity $U(e + \varphi) = Ue + U\varphi$ for every $\varphi \in \{e\}^\perp$.

Next, we analyze the map $U_\perp$ in detail. Our goal is to verify that $U_\perp$ defines a linear or an antilinear map on $\{e\}^\perp$. To this end, pick $\varphi_1, \varphi_2 \in \{e\}^\perp$. By (1.26), the preliminary remarks and (1.28), we have that $|\langle U_\perp \varphi_1, U_\perp \varphi_2 \rangle|^2 = |\langle \varphi_1, \varphi_2 \rangle|^2$ as well as

$$|1 + \langle U_\perp \varphi_1, U_\perp \varphi_2 \rangle|^2 = |\langle e' + U_\perp \varphi_1, e' + U_\perp \varphi_2 \rangle|^2 = |\langle e + \varphi_1, e + \varphi_2 \rangle|^2 = |1 + \langle \varphi_1, \varphi_2 \rangle|^2.$$

By expanding the squares, this implies that $\operatorname{Re} \langle U_\perp \varphi_1, U_\perp \varphi_2 \rangle = \operatorname{Re} \langle \varphi_1, \varphi_2 \rangle$. In particular, if $\langle \varphi_1, \varphi_2 \rangle = \operatorname{Re} \langle \varphi_1, \varphi_2 \rangle$ is real, we must have $\operatorname{Im} \langle U_\perp \varphi_1, U_\perp \varphi_2 \rangle = 0$ (recalling that $|\langle U_\perp \varphi_1, U_\perp \varphi_2 \rangle| = |\langle \varphi_1, \varphi_2 \rangle|$) so that in this case

$$\langle U_\perp \varphi_1, U_\perp \varphi_2 \rangle = \langle \varphi_1, \varphi_2 \rangle. \quad (1.29)$$

Next, let us pick some vector $\psi_2 \in \{e\}^\perp$ (at this point, we make use of the assumption that $\dim \mathcal{H} \geq 3$) that is orthonormal to $\psi_1 = e_{\varphi_1} = \varphi_1 / \|\varphi_1\| \in \{e\}^\perp$ (assuming in the sequel without loss of generality that $\varphi_1 \neq 0$) and such that

$$\varphi_1 = \lambda_1 \psi_1, \quad \varphi_2 = \mu_1 \psi_1 + \mu_2 \psi_2. \quad (1.30)$$

From the preliminary remarks, we recall that if $\psi'_i = U_\perp \psi_i \in T[\psi_i]$, then ψ'_1 and ψ'_2 are still orthonormal. By definition of $U_\perp$, we have for every $\psi \in \{e\}^\perp, \nu \in \mathbb{C}$ that

$$U_\perp(\nu\psi) = \chi_\psi(\nu)\psi' \quad \text{with} \quad |\chi_\psi(\nu)| = |\nu|.$$

In the sequel, let's abbreviate $\chi_i(\nu) = \chi_{\psi_i}(\nu)$. Note in particular that $\chi_i(1) = 1$, by definition of ψ'_i . Moreover, by the previous observations, we find for $\nu_1, \nu_2 \in \mathbb{C}$ that

$$\operatorname{Re} \overline{\chi_i(\nu_1)} \chi_i(\nu_2) = \operatorname{Re} \langle U_\perp(\nu_1 \psi_i), U_\perp(\nu_2 \psi_i) \rangle = \operatorname{Re} \langle \nu_1 \psi_i, \nu_2 \psi_i \rangle = \operatorname{Re} \bar{\nu}_1 \nu_2$$

so that

$$\operatorname{Re} \chi_i(\nu) = \operatorname{Re} \overline{\chi_i(1)} \chi_i(\nu) = \operatorname{Re} \nu \quad \text{and} \quad \chi_i(\nu) = \nu \quad \text{if } \nu \in \mathbb{R} \quad (\text{since } |\chi_i(\nu)| = |\nu|).$$

Now, pick an arbitrary linear combination $\psi = \nu_1 \psi_1 + \nu_2 \psi_2$. Then, by the preliminary remarks, we know that $U_\perp \psi = \nu'_1 \psi'_1 + \nu'_2 \psi'_2$ for suitable coefficients $|\nu'_i| = |\nu_i|, i \in \{1, 2\}$. If $\nu_i = 0$, then $\nu'_i = 0 = \chi_i(\nu_i)$. Otherwise, if $\nu_i \neq 0$, then $\langle (\bar{\nu}_i)^{-1} \psi_i, \nu_i \psi_i \rangle = 1$ so that

$$\langle (\bar{\nu}_i)^{-1} \psi_i, \nu_i \psi_i \rangle = \operatorname{Re} \langle (\bar{\nu}_i)^{-1} \psi_i, \nu_i \psi_i \rangle = \langle U_\perp(\bar{\nu}_i)^{-1} \psi_i, U_\perp \nu_i \psi_i \rangle = \overline{\chi_i((\bar{\nu}_i)^{-1})} \chi_i(\nu_i).$$

At the same time, $\langle (\bar{\nu}_i)^{-1} \psi_i, \nu_i \psi_i \rangle = \langle (\bar{\nu}_i)^{-1} \psi_i, \psi \rangle$, so that by the same argument

$$\langle (\bar{\nu}_i)^{-1} \psi_i, \psi \rangle = \langle U_\perp(\bar{\nu}_i)^{-1} \psi_i, U_\perp \psi \rangle = \langle U_\perp(\bar{\nu}_i)^{-1} \psi_i, \nu'_1 \psi'_1 + \nu'_2 \psi'_2 \rangle = \overline{\chi_i((\bar{\nu}_i)^{-1})} \nu'_i.$$

Combining this, we find that $\nu'_i = \chi_i(\nu_i)$. In other words, for every $\nu_1, \nu_2 \in \mathbb{C}$, we have

$$U_\perp(\nu_1 \psi_1 + \nu_2 \psi_2) = \chi_1(\nu_1) U_\perp \psi_1 + \chi_2(\nu_2) U_\perp \psi_2.$$

Since this implies

$$\chi_{\psi_1 + \psi_2}(\nu)(\psi'_1 + \psi'_2) = \chi_{\psi_1 + \psi_2}(\nu) U_\perp(\psi_1 + \psi_2) = U_\perp(\nu \psi_1 + \nu \psi_2) = \chi_1(\nu) \psi'_1 + \chi_2(\nu) \psi'_2,$$

we have by orthogonality in fact that $\chi_1(\nu) = \chi_2(\nu)$ for all $\nu \in \mathbb{C}$, that is

$$U_\perp(\nu_1 \psi_1 + \nu_2 \psi_2) = \chi(\nu_1) U_\perp \psi_1 + \chi(\nu_2) U_\perp \psi_2.$$

where from now on we set $\chi(\nu) = \chi_1(\nu)$ for all $\nu \in \mathbb{C}$. This leads us to asking what possibilities we have for the function χ ? By the previous observations, $|\chi(i)| = 1$ and $\operatorname{Re} \chi(i) = \operatorname{Re} i = 0$ so that $\chi(i) = \sigma i$ for $\sigma \in \{-1, 1\}$. For a general $\nu \in \mathbb{C}$, on the other hand, we know that $\operatorname{Re} \chi(\nu) = \operatorname{Re} \nu$ and, similarly, that

$$\operatorname{Im} \chi(\nu) = \operatorname{Re}(-i\chi(\nu)) = \sigma \operatorname{Re}(\overline{\chi(i)}\chi_1(\nu)) = \sigma \operatorname{Re}(-i\nu) = \sigma \operatorname{Im} \nu$$

so that, in conclusion, $\chi(\nu) = \nu$ for all $\nu \in \mathbb{C}$ or $\chi(\nu) = \bar{\nu}$ for all $\nu \in \mathbb{C}$. In particular, $\chi : \mathbb{C} \rightarrow \mathbb{C}$ is either a linear or an antilinear map.

Let us collect the previous observations and draw some conclusions on $U_\perp$. Recalling that $\varphi_1 = \lambda_1\psi_1$ and $\varphi_2 = \mu_1\psi_1 + \mu_2\psi_2$, the previous arguments imply that

$$\begin{aligned} U_\perp(\varphi_1 + \varphi_2) &= U_\perp((\lambda_1 + \mu_1)\psi_1 + \mu_2\psi_2) = \chi_1(\lambda_1 + \mu_1)U_\perp\psi_1 + \chi_1(\mu_2)U_\perp\psi_2 \\ &= \chi_1(\lambda_1)U_\perp\psi_1 + \chi_1(\mu_1)U_\perp\psi_1 + \chi_1(\mu_2)U_\perp\psi_2 \\ &= U_\perp\varphi_1 + U_\perp\varphi_2, \end{aligned}$$

that $U_\perp(\nu\varphi_i) = \chi(\nu)\varphi_i$, $i \in \{1, 2\}$ for $\chi(\nu) = \nu$ or $\chi(\nu) = \bar{\nu}$ and, finally, that

$$\langle U_\perp\varphi_1, U_\perp\varphi_2 \rangle = \overline{\chi(\lambda_1)}\chi(\mu_1) = \chi(\bar{\lambda}_1)\chi(\mu_1) = \chi(\bar{\lambda}_1\mu) = \chi(\langle \varphi_1, \varphi_2 \rangle).$$

Since $\varphi_1 (\neq 0)$ and φ_2 were arbitrary vectors in $\{e\}^\perp$, this proves that $U_\perp$ extends either to a linear, isometric or to an antilinear, anti-isometric map on $\{e\}^\perp$ that is consistent with the symmetry transformation $T : \mathcal{R}_1 \rightarrow \mathcal{R}_1$, by construction.

We are now ready to finish the construction of the map U . To this end, it only remains to define its action on vectors of the form $\alpha e + \varphi$ for $\varphi \in \{e\}^\perp$ and for some $\alpha \in \mathbb{C} \setminus \{0, 1\}$. Here, we finally set

$$U(\alpha e + \varphi) = \chi(\alpha)U(e + \alpha^{-1}\varphi) = \chi(\alpha)e' + \chi(\alpha)U_\perp(\alpha^{-1}\varphi) = \chi(\alpha)e' + U_\perp\varphi = \chi(\alpha)Ue + U\varphi$$

and it is readily verified that this yields a map $U : \mathcal{H} \rightarrow \mathcal{H}$ that is either linear and unitary or antilinear and antiunitary (depending on the function $\chi : \mathbb{C} \rightarrow \mathbb{C}$) with the consistency property that $U\psi \in T[\psi]$ for every $\psi \in \mathcal{H}$.

We conclude the proof by showing the uniqueness of U , up to multiplication by a constant phase. Suppose that $U_1, U_2 : \mathcal{H} \rightarrow \mathcal{H}$ are compatible with the symmetry transformation and that each is either linear, unitary or antilinear, antiunitary. Then, both U_1 and U_2 are necessarily additive. By compatibility with T , we have $U_1\psi \sim U_2\psi$ for all $\psi \in \mathcal{H}$. This means that for every $\psi \in \mathcal{H}$, there exists $\omega_\psi \in [0, 2\pi)$ so that

$$U_2\psi = e^{i\omega_\psi}U_1\psi.$$

If $\varphi, \psi \in \mathcal{H}$ are two linearly independent vectors (which exist if $\dim(\mathcal{H}) \geq 2$), the previously mentioned additivity of U_1 and U_2 implies

$$e^{i\omega_\varphi}U_1\varphi + e^{i\omega_\psi}U_1\psi = U_2(\varphi + \psi) = e^{i\omega_{\varphi+\psi}}(U_1\varphi + U_1\psi) = e^{i\omega_{\varphi+\psi}}U_1\varphi + e^{i\omega_{\varphi+\psi}}U_1\psi$$

so that $e^{i\omega_\varphi} = e^{i\omega_{\varphi+\psi}} = e^{i\omega_\psi}$. Note here that linear independence of φ and ψ is preserved by U_1 , by (anti-)unitarity and the fact that φ and ψ are linearly independent if and only

if $|\langle \varphi, \psi \rangle| < \|\varphi\| \|\psi\|$. On the other hand, if φ, ψ are linearly dependent, we can find some $\zeta \in \mathcal{H}$ independent of both φ and ψ (using once more $\dim(\mathcal{H}) \geq 2$) so that

$$e^{i\omega\varphi} = e^{i\omega\varphi+\zeta} = e^{i\omega\zeta} = e^{i\omega\psi+\zeta} = e^{i\omega\psi}$$

by the previous argument. In conclusion, this shows $U_2 = e^{i\omega}U_1$ for some $\omega \in [0, 2\pi)$. $\square$

Problem 1.4. *Verify Theorem 1.4 for $\dim \mathcal{H} = 2$. For $\dim \mathcal{H} = 1$, prove the existence of a linear, unitary or antilinear, antiunitary map that is compatible with T as in Theorem 1.4, but show that the uniqueness statement as for $\dim(\mathcal{H}) \geq 2$ fails.*

Theorem 1.4 shows that unitary (or antiunitary) operators are appropriate to model symmetries in quantum mechanics where states are described by suitable vectors in a Hilbert space. We conclude this section by presenting some basic examples.

Example 1.6 (Translations). Consider a particle of mass $m > 0$ moving in $\mathbb{R}^d$ in an external field $V : \mathbb{R}^d \rightarrow \mathbb{R}$. As discussed in Section 1.2.1, we can model this by choosing the Hilbert space $\mathcal{H} = L^2(\mathbb{R}^d)$ and the Hamiltonian

$$H = \frac{|i\hbar\nabla|^2}{2m} + V(x).$$

As usual, we assume that V satisfies suitable constraints so that H can be realized as a self-adjoint operator on a dense domain in $L^2(\mathbb{R}^d)$. A basic symmetry transformation is to translate the coordinate system that we use to measure positions. If the state of the system is described in the original coordinate system by $\psi \in L^2(\mathbb{R}^d)$ and the system is translated by $a \in \mathbb{R}^d$, the state from the viewpoint of the transformed system reads

$$x \mapsto (U_a\psi)(x) = \psi(x - a) \in L^2(\mathbb{R}^d) = U_a L^2(\mathbb{R}^d).$$

Clearly, the maps $(U_a)_{a \in \mathbb{R}^d}$ form a family of unitary transformations.

In analogy to Example 1.4, consider a direction $n \in \mathbb{R}^d, |n| = 1$. In the classical setting, the group of translations $(tn)_{t \in \mathbb{R}}$ leads to the conservation of the momentum in direction n if V is translation invariant in direction n , i.e. $V(\cdot + tn) = V(\cdot)$. An analogue remains valid in the quantum mechanical setting. To see this, note that Theorem 1.2 applied to $(U_{tn})_{t \in \mathbb{R}}$ implies the existence of a self-adjoint operator $p \cdot n$ that satisfies

$$U_{tn} = e^{-itp \cdot n/\hbar}, \quad \forall t \in \mathbb{R}.$$

A direct calculation shows that

$$(p \cdot n \psi)(x) = i\hbar \lim_{t \rightarrow 0} \frac{1}{t} (U_{tn} - U_0 \psi)(x) = i\hbar \lim_{t \rightarrow 0} \frac{\psi(x - tn) - \psi(x)}{t} = (-i\hbar \nabla \cdot n \psi)(x)$$

for all $\psi \in L^2(\mathbb{R}^d)$ that satisfy $\nabla \cdot n \psi \in L^2(\mathbb{R}^d)$. Choosing for $n \in \mathbb{R}^d$ the canonical base vectors e_j in $\mathbb{R}^d$, for $j = 1, \dots, d$, the corresponding translations are generated by

$$p_j : D_{p_j} = \{\psi \in L^2(\mathbb{R}^d) : \partial_{x_j} \psi \in L^2(\mathbb{R}^d)\} \rightarrow L^2(\mathbb{R}^d), \quad p_j \psi = -i\hbar \partial_{x_j} \psi, \quad j \in \{1, \dots, d\}.$$

In other words, the generators of translations in $\mathbb{R}^d$ are the momentum operators that we introduced earlier in Section 1.2.1, based on the canonical quantization method.

Now, let us explain why translation invariance $V(\cdot + tn) = V(\cdot)$ of V implies the conservation of $p \cdot n$ under the Schrödinger dynamics (1.15). In fact, a direct calculation shows the invariance $U_{sn} H U_{sn}^* = H$, for all $s \in \mathbb{R}$. The functional calculus thus implies

$$U_{sn} e^{-itH/\hbar} U_{sn}^* = e^{-itH/\hbar} \iff e^{-is(p \cdot n)_t/\hbar} U_{sn}^* = \mathbf{1}_{L^2(\mathbb{R}^d)} \iff e^{-is(p \cdot n)_t/\hbar} = e^{-isp \cdot n/\hbar}$$

for every $s, t \in \mathbb{R}$ and this implies by (the proof of) Theorem 1.2 that

$$(p \cdot n)_t = e^{itH/\hbar} p \cdot n e^{-itH/\hbar} = p \cdot n, \quad \forall t \in \mathbb{R}.$$

Example 1.7 (Rotations). Consider a particle in $\mathbb{R}^3$, in the same setting as in the previous Example 1.6. Symmetry transformations that describe rotations of the coordinate system can be implemented through the unitary operators

$$U_R : L^2(\mathbb{R}^3) \rightarrow L^2(\mathbb{R}^3), \quad (U_R \psi)(x) = \psi(R^{-1}x), \quad \text{for } R \in SO(3).$$

In analogy to Example 1.5, the quantum angular momentum $L \cdot n$ in direction n , $|n|$, is the generator of the one-parameter unitary group $(U_{e^{\omega n \cdot x}})_{\omega \in \mathbb{R}}$, where the matrices X_1, X_2, X_3 denote the generators of $\mathfrak{so}(3)$, defined in (1.10). For example, the angular momentum in direction e_3 takes the explicit form

$$\begin{aligned} (L_3 \psi)(x) &= i\hbar \lim_{\omega \rightarrow 0} \frac{\psi(e^{-\omega X_3} x) - \psi(x)}{\omega} \\ &= i\hbar \lim_{\omega \rightarrow 0} \frac{1}{\omega} \left(\psi((\cos(\omega)x_1 + \sin(\omega)x_2, -\sin(\omega)x_1 + \cos(\omega)x_2, x_3)) - \psi(x) \right) \\ &= (x_2(i\hbar \partial_{x_1})\psi - x_1(i\hbar \partial_{x_2})\psi)(x) = ((x_1 p_2 - x_2 p_1)\psi)(x). \end{aligned}$$

More generally, one readily verifies that $L = (L_1, L_2, L_3) = x \times p$ on a suitable, dense domain, where x and p denote the quantum mechanical position and momentum operators. Once again, this is consistent with the definition of quantum angular momentum via canonical quantization of the classical angular momentum. Recall that the latter corresponds to the classical observable $(x, p) \mapsto L(x, p) = x \times p \in C^\infty(\mathbb{R}^3 \times \mathbb{R}^3)$.

The two previous examples illustrate that a symmetry that corresponds to a Lie group may give rise to a strongly continuous unitary representation of the group on the state space. At this point, let us introduce some basic definitions. Let G be a group. Then, a group representation $\rho : G \rightarrow \text{GL}(V) = \{ A : V \rightarrow V : A \text{ linear}, A^{-1} \text{ exists} \}$ of G on a vector space V is a group homomorphism to $\text{GL}(V)$, the vector space of invertible linear maps on V . That is, ρ satisfies

$$\rho(g_1 g_2) = \rho(g_1) \rho(g_2), \quad \forall g_1, g_2 \in G.$$

Unless mentioned otherwise, all considered vector spaces V are assumed to be vector spaces over $\mathbb{C}$. We call $\dim(V)$ the dimension of V . If G is a topological group and V is a

normed space, we call ρ strongly continuous if the map $G \ni g \mapsto \rho(g)v \in V$ is continuous, for every $v \in V$. The representation ρ is called faithful if it is injective and it is called irreducible if there does not exist a non-trivial, closed subspace $S \subset V$, so that $S \neq \{0\}$ and $S \neq V$, which is invariant under ρ , that is, $\rho(g)S \subset S$ for all $g \in G$. Otherwise, ρ is called reducible. From a physical point of view, according to the ideas of E. Wigner [34, 35], if a quantum system is equipped with a symmetry group that is described by an irreducible representation, then the system can be considered elementary in the sense that the symmetry can not be used to resolve some finer structure of the model. If it is reducible, on the other hand, one may factorize the state space accordingly into finer, elementary subspaces.

In view of Theorem 1.4, in the context of quantum mechanics typical symmetry groups, like e.g. $\text{SO}(3)$ in Example 1.7, lead to representations that take values in the space of unitary operators $\mathcal{U}(\mathcal{H}) = \{U : \mathcal{H} \rightarrow \mathcal{H} : U \text{ is unitary}\}$, if $\mathcal{H}$ denotes the Hilbert space describing the possible states. Unitary operators (as opposed to antiunitary operators) naturally occur if the symmetry is described by a connected Lie group that contains the identity (which is unitary). A representation $\rho : G \rightarrow \mathcal{U}(\mathcal{H})$ that takes its values in $\mathcal{U}(\mathcal{H})$ is called a unitary representation of G .

Lemma 1.3. *A unitary representation $\rho : G \rightarrow \mathcal{U}(\mathcal{H})$ is irreducible if and only if for every $\psi \in \mathcal{H}, \psi \neq 0$, we have that*

$$\mathcal{H} = \overline{\text{span}\{\rho(g)\psi : g \in G\}} \iff \{0\} = \{\varphi \in \mathcal{H} : \langle U(g)\psi, \varphi \rangle = 0, \forall g \in G\}.$$

Proof. The space $\overline{\text{span}\{U(g)\psi : g \in G\}}$ is clearly invariant under ρ . If ρ is irreducible, this implies it must be equal to $\{0\}$ ($\psi = 0$) or $\mathcal{H}$ ($\psi \neq 0$). On the other hand, if $\mathcal{H} = \overline{\text{span}\{\rho(g)\psi : g \in G\}}$ for every non-zero $\psi \in \mathcal{H}$ and if there exists a non-zero, closed linear subspace $S \subset \mathcal{H}$ invariant under ρ , we pick $\varphi \in S, \varphi \neq 0$ and find $\mathcal{H} = \overline{\text{span}\{\rho(g)\varphi : g \in G\}} \subset S \subset \mathcal{H}$. In other words, $S = \mathcal{H}$ or equivalently $S^\perp = \{0\}$. $\square$

Lemma 1.4 (Schur). *Consider a finite-dimensional, unitary representation $\rho : G \rightarrow \mathcal{H}$. Then ρ is irreducible if and only if every linear map A that commutes with every operator $\rho(g)$, for $g \in G$, is a multiple of the identity $\mathbf{1}_{\mathcal{H}}$.*

Proof. Recall that every linear map A has at least one eigenvalue $\lambda \in \mathbb{C}$ in a complex vector space $\mathcal{H}$. If $\varphi \neq 0$ is a normalized eigenvector and ρ is irreducible, we know that $\overline{\text{span}\{\rho(g)\varphi : g \in G\}} = \mathcal{H}$. In particular, if $\psi = U(g)\varphi$, we see that

$$A\psi = A\rho(g)\varphi = \rho(g)A\varphi = \lambda\rho(g)\varphi = \lambda\psi,$$

which implies by continuity that $A = \lambda\mathbf{1}_{\mathcal{H}}$. On the other hand, if ρ reducible, pick a non-trivial subspace $S \subset \mathcal{H}, S \neq \mathcal{H}$, that is invariant under ρ . This induces a unitary representation $\rho' : G \rightarrow \mathcal{U}(S)$, because $\rho(g)S \subset S$. Without loss of generality, we may assume that ρ' is irreducible, because otherwise we can repeat the previous step until it is (recalling that $\dim(\mathcal{H}) < \infty$). By the previous step, we may thus assume that

$S = \overline{\text{span}\{\rho(g)\varphi : g \in G\}}$ for some $0 \neq \varphi \in S$. Notice that the invariance of S under ρ also implies the invariance of $S^\perp$ under ρ , which readily follows from unitarity and

$$\langle \rho(g)\psi, \zeta \rangle = \langle \psi, \rho(g)^*\zeta \rangle = \langle \psi, \rho(g^{-1})\zeta \rangle.$$

Now, choose $P : \mathcal{H} \rightarrow S$ to be the orthogonal projection onto S and let $\psi \in \mathcal{H}, \psi \neq 0$ so that $\psi = \psi_1 + \psi_2 = \sum_{i=1}^\infty \mu_i \rho(g_i)\varphi + \psi_2$ for $\psi_1 \in S, \psi_2 \in S^\perp$. Then, using that $P\rho(h)\varphi = \rho(h)\varphi \in S$, continuity of P as well as $P\rho(h)\psi_2 = 0$ for every $h \in G$, we obtain

$$\begin{aligned} P\rho(h)\psi &= \sum_{i=1}^\infty \mu_i P\rho(h)\rho(g_i)\varphi + P\rho(h)\psi_2 = \sum_{i=1}^\infty \mu_i \rho(hg_i)\varphi \\ &= \rho(h) \left(\sum_{i=1}^\infty \mu_i P\rho(g_i)\varphi + P\psi_2 \right) = \rho(h)P\psi. \end{aligned}$$

This implies that $[P, \rho(g)] = 0$ for every $g \in G$, but $P \neq \lambda \mathbf{1}_{\mathcal{H}}$, for every $\lambda \in \mathbb{C}$. $\square$

In Example 1.7, we analysed the unitary representation $\rho : \text{SO}(3) \rightarrow \mathcal{U}(\mathcal{H})$, given by

$$\rho(R)\psi = \psi(R^{-1}.)$$

for $\psi \in \mathcal{H} = L^2(\mathbb{R}^3)$, to describe the rotation symmetry of a system that describes a massive particle in $\mathbb{R}^3$. This representation is induced by the group action

$$\text{SO}(3) \times \mathbb{R}^3 \ni (R, x) \mapsto Rx \in \mathbb{R}^3$$

of $\text{SO}(3)$ on $\mathbb{R}^3$ and the $\text{SO}(3)$ -invariance of the Lebesgue measure. In the same way, every other finite-dimensional representation, $T : \text{SO}(3) \rightarrow \text{SL}(n, \mathbb{C})$ (mapping to the space $\text{SL}(n, \mathbb{C})$ of $n \times n$ matrices with determinant one) induces a unitary representation $\rho : \text{SO}(3) \rightarrow \mathcal{U}(L^2(\mathbb{R}^3, \mathbb{C}^n))$ on $L^2(\mathbb{R}^3, \mathbb{C}^n)$ through

$$\rho(R)\psi = T(R)\psi(R^{-1}.), \quad \forall \psi \in L^2(\mathbb{R}^3, \mathbb{C}^n). \quad (1.31)$$

Studying the representations of $\text{SO}(3)$ therefore provides important information on the possible quantum systems with rotational symmetry.

Up to this point, we discussed exclusively unitary representations as a tool to represent symmetries in quantum mechanics. Due to the global gauge invariance in quantum mechanics, we may allow more generally for so called projective unitary representations. In view of Wigner's Theorem 1.4, note indeed that if $\rho(g_1) \in \mathcal{U}(\mathcal{H})$ and $\rho(g_2) \in \mathcal{U}(\mathcal{H})$ represent two symmetry transformations, their composition (as a symmetry transformation) is represented by $\rho(g_1g_2) \in \mathcal{U}(\mathcal{H})$ so that

$$[\rho(g_1g_2)\psi] = [\rho(g_1)\rho(g_2)\psi], \quad \forall \psi \in \mathcal{H}, \|\psi\| = 1.$$

A priori, we can therefore only say that the operators $\rho(g_1g_2)$ and $\rho(g_1)\rho(g_2)$ are equal up to a constant phase. This important observation leads naturally to the notion of projective representations, a generalization of ordinary representations. While referring

for a general introduction to this topic to e.g. [17, Chapter 12], here we follow a hands on approach and define a projective unitary representation $\rho : G \rightarrow \mathcal{U}(\mathcal{H})$ to be a map such that for every $g_1, g_2 \in G$, there is some phase $\lambda(g_1, g_2) \in \mathbb{C}$, $|\lambda(g_1, g_2)| = 1$ so that

$$\rho(g_1 g_2) = \lambda(g_1, g_2) \rho(g_1) \rho(g_2). \quad (1.32)$$

In view of the classification of quantum systems with a given symmetry into elementary systems, it is therefore an important task to study the projective representations of the symmetry group. Although the class of projective representations is in general larger than that of the ordinary representations, in several physically relevant cases the projective representations can be understood in terms of the representations of its covering group. This applies in particular to $\mathrm{SO}(3)$ as explained below. To this end, we use the following result that relates $\mathrm{SO}(3)$ to the group $\mathrm{SU}(2)$, which is defined by

$$\begin{aligned} \mathrm{SU}(2) &= \{A \in \mathbb{C}^{2 \times 2} : A^* A = \mathbf{1}_{\mathbb{C}^2}, \det A = 1\} \\ &= \{A = (a_{ij})_{i,j=1}^2 \in \mathbb{C}^{2 \times 2} : a_{11} = \bar{a}_{22}, a_{21} = -\bar{a}_{12}, |a_{11}|^2 + |a_{12}|^2 = 1\}. \end{aligned} \quad (1.33)$$

Observe that the second equality implies that $\mathrm{SU}(2)$ is a (locally path) connected, compact and simply connected Lie group, for it is diffeomorphic to the unit sphere $S^3 \subset \mathbb{R}^4$.

Proposition 1.1. *There exists a two-to-one group homomorphism $R : \mathrm{SU}(2) \rightarrow \mathrm{SO}(3)$, which is a local homeomorphism.*

Proof. Denote by $\sigma_1, \sigma_2, \sigma_3 \in \mathbb{C}^{2 \times 2}$ the Pauli spin matrices, defined by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.34)$$

Note that $i\sigma_1, i\sigma_2, i\sigma_3 \in \mathbb{C}^{2 \times 2}$ form a basis of the space of traceless, antisymmetric matrices, which, as a side remark, equals the Lie algebra $\mathfrak{su}(2)$ of $\mathrm{SU}(2)$ (*exercise*).

Now, consider the map

$$\mathbb{R}^3 \ni x \mapsto \sigma \cdot x = \sum_{j=1}^3 \sigma_j x_j = \begin{pmatrix} x_3 & x_1 + ix_2 \\ x_1 - ix_2 & -x_3 \end{pmatrix} = (\sigma \cdot x)^* \in \mathbb{C}^{2 \times 2}$$

so that $\mathrm{tr} \sigma \cdot x = 0$ for every $x \in \mathbb{R}^3$ and observe that

$$\det(\sigma \cdot x) = -x_3^2 - |x_1 - ix_2|^2 = -|x|^2.$$

Given that $A(\sigma \cdot x)A^*$ is self-adjoint and traceless, and since the $\sigma_1, \sigma_2, \sigma_3 \in \mathbb{C}^{2 \times 2}$ form a basis of the traceless, self-adjoint matrices in $\mathbb{C}^{2 \times 2}$, we find for every $A \in \mathrm{SU}(2)$ a unique $y_A(x) \in \mathbb{R}^3$ such that

$$A(\sigma \cdot x)A^* = \sigma \cdot y_A(x), \quad \text{with} \quad (y_A(x))_j = \frac{1}{2} \mathrm{tr} \sigma_j A(\sigma \cdot x)A^*, j \in \{1, 2, 3\}.$$

The map $x \rightarrow y_A(x)$ is clearly linear, continuous and can be written as $y_A(x) = R(A)x$ for some $R(A) \in \mathbb{R}^{3 \times 3}$. We claim that the continuous map

$$\mathrm{SU}(2) \ni A \mapsto R(A) = (R(A)_{ij})_{i,j=1}^3 = \frac{1}{2}(\mathrm{tr} \sigma_i A \sigma_j A^*) \in \mathbb{R}^{3 \times 3}$$

is a two-to-one homomorphism from $\mathrm{SU}(2)$ to $\mathrm{SO}(3)$. To see this, note first that

$$|R(A)x|^2 = -\det \sigma \cdot (R(A)x) = -\det A(\sigma \cdot x)A^* = -\det \sigma \cdot x = |x|^2.$$

Combined with the fact that $R(\mathbf{1}_{\mathbb{C}^2}) = \mathbf{1}_{\mathbb{R}^3}$ with $\det \mathbf{1}_{\mathbb{R}^3} = 1$, that $\mathrm{SU}(2) \ni A \mapsto \det R(A) \in \{-1, 1\}$ is continuous and the fact that $\mathrm{SU}(2)$ is connected, we conclude that $R(A) \in \mathrm{SO}(3)$. The group homomorphism property follows from

$$\sigma \cdot (R(AB)x) = AB(\sigma \cdot x)B^*A^* = A(\sigma \cdot (R(B)x))A^* = \sigma \cdot (R(A)R(B)x)$$

so that $R(AB) = R(A)R(B)$. That $\mathrm{SU}(2) \ni A \mapsto R(A) \in \mathrm{SO}(3)$ is surjective is discussed in Problem 1.5. Finally, suppose that $R(A) = \mathbf{1}_{\mathbb{R}^3}$. This is the case if and only if

$$A \sigma \cdot x = \sigma \cdot x A$$

for every $x \in \mathbb{R}^3$. In particular, this implies that $[A, \sigma_i] = 0$ for all $i \in \{1, 2, 3\}$. Choosing $i = 3$, an explicit computation (*exercise*) then verifies that $b = 0$ if

$$A = \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}.$$

Combining this with the condition $[A, \sigma_1] = 0$ implies that $\mathrm{Im}(a) = 0$ (*exercise*) and since $A^*A = \mathbf{1}_{\mathbb{C}^2}$, we conclude $A = \mathbf{1}_{\mathbb{C}^2}$ or $A = -\mathbf{1}_{\mathbb{C}^2}$. Combined with the homomorphism property, we see that $R(A_1) = R(A_2)$ if and only if $A_1 = A_2$ or $A_1 = -A_2$ so that the map $\mathrm{SU}(2) \ni A \mapsto R(A) \in \mathrm{SO}(3)$ is a continuous two-to-one homomorphism. That R is a local homeomorphism is discussed in Problem 1.5. $\square$

Problem 1.5. Consider the map $\mathrm{SU}(2) \ni A \mapsto R(A) \in \mathrm{SO}(3)$ from the previous proposition and recall the definition of the generators $X_1, X_2, X_3 \in \mathbb{R}^{3 \times 3}$ of $\mathfrak{so}(3)$ from (1.10).

a) Show that for every $\omega \in [0, 2\pi)$, we have that

$$R_{e_3, \omega} = \begin{pmatrix} \cos(\omega) & -\sin(\omega) & 0 \\ \sin(\omega) & \cos(\omega) & 0 \\ 0 & 0 & 1 \end{pmatrix} = R(A_{e_3, \omega}), \text{ for } A_{e_3, \omega} = \begin{pmatrix} e^{\frac{i}{2}\omega} & 0 \\ 0 & e^{-\frac{i}{2}\omega} \end{pmatrix} \in \mathrm{SU}(2).$$

b) Let $R \in \mathrm{SO}(3)$. Prove that there exists a direction $v_3 \in \mathbb{R}^3, |v_3| = 1$, such that $Rv_3 = v_3$. If $\mathcal{B} = (v_1, v_2, v_3)$ is an orthonormal basis of $\mathbb{R}^3$, show that the matrix representation $R_{\mathcal{B}}$ of R with respect to $\mathcal{B}$ equals $R_{\mathcal{B}} = R_{e_3, \omega}$ for some $\omega = \omega_{\mathcal{B}} \in [0, 2\pi)$.

c) Show that the Hilbert space $\mathbb{R}^3$ is isometric to the space of matrices $\{\sigma \cdot x : x \in \mathbb{R}^3\} \subset \mathbb{C}^{2 \times 2}$ with inner product $\langle A, B \rangle = \mathrm{tr} AB, \forall A, B \in \{\sigma \cdot x : x \in \mathbb{R}^3\}$.

- d) Let $R \in \text{SO}(3)$ and $v_3 \in \mathbb{R}^3, |v_3| = 1$, be as in part b). Analyze the spectra of $\sigma \cdot v_3$ and σ_3 to conclude that there exists a unitary map $U \in \text{U}(2) \subset \mathbb{C}^{2 \times 2}$ such that $U\sigma \cdot v_3 U^* = \sigma_3$. Use part c) to find an orthonormal basis $\mathcal{B} = (v_1, v_2, v_3)$ such that $U\sigma \cdot v_j U^* = \sigma_j$ for each $j \in \{1, 2, 3\}$.
- e) Prove that $\text{SU}(2) \ni A \mapsto R(A) \in \text{SO}(3)$ is surjective and locally invertible with continuous inverse.
- f) Prove explicitly that for every $x \in \mathbb{R}^3$ and $\omega \in [0, 2\pi)$, it holds true that

$$e^{\frac{i}{2}\omega\sigma_3}(\sigma \cdot x)e^{-\frac{i}{2}\omega\sigma_3} = \sigma \cdot (e^{\omega X_3}x). \quad (1.35)$$

Our next goal is to show that every finite-dimensional projective unitary representation of $\text{SO}(3)$ corresponds to an ordinary unitary representation of $\text{SU}(2)$. In other words, by enlarging the symmetry group, the classification problem of its projective representations can be reduced to the classification of the ordinary representations of its covering group. To make this more precise, we say that a projective representation ρ can be lifted to an ordinary representation ρ' , sometimes called de-projectification of ρ , if there exists $\lambda : G \rightarrow \{z \in \mathbb{C} : |z| = 1\}$ such that $\rho(g) = \lambda(g)\rho'(g)$, for every $g \in G$.

Proposition 1.2. *Let G be a compact, connected and simply connected Lie group. Then, every finite dimensional strongly continuous projective unitary representation of G can be lifted to a strongly continuous unitary representation of G .*

Remark 1.6. *The reason we assume compactness is that under certain additional assumptions, non-compact Lie groups generally do not admit non-trivial finite-dimensional unitary representations. For more details on this, see e.g. [17, Chapter 12], [9].*

Remark 1.7. *Prop. 1.2 is a simplified version of a more general result from [1].*

Before proving Prop. 1.2, we record its consequences for $\text{SO}(3)$ -representations.

Corollary 1.1. *Let $\rho : \text{SO}(3) \rightarrow \mathcal{U}(\mathcal{H})$ be a finite-dimensional strongly continuous projective unitary representation of $\text{SO}(3)$. Then, there exists an ordinary representation τ of $\text{SU}(2)$ and a continuous map $\lambda : \text{SU}(2) \rightarrow \{z \in \mathbb{C} : |z| = 1\}$ such that*

$$\rho(R(A)) = \lambda(A)\tau(A), \quad \forall A \in \text{SU}(2).$$

Here, $R : \text{SO}(3) \rightarrow \text{SU}(2)$ denotes the homomorphism constructed in Proposition 1.1.

Proof. As noted earlier, $\text{SU}(2)$ is diffeomorphic to the unit sphere $S^3 \subset \mathbb{R}^4$ and is thus compact, connected and simply connected. Applying Prop. 1.2 to

$$\text{SU}(2) \ni A \mapsto \rho(R(A)) \in \mathcal{U}(\mathcal{H}),$$

which defines a projective representation of $\text{SU}(2)$, implies the claim. $\square$

Proof of Prop. 1.2. Assume w.l.o.g. $\mathcal{H} = \mathbb{C}^n$ and let $\rho : G \rightarrow \mathcal{U}(\mathcal{H}) = \mathcal{U}(\mathbb{C}^n) \subset \mathrm{GL}(\mathbb{C}^n)$ be a projective unitary representation so that

$$\rho(g_1 g_2) = \mu(g_1, g_2) \rho(g_1) \rho(g_2) \quad \text{with} \quad |\mu(g_1, g_2)| = 1, \quad \forall g_1, g_2 \in G.$$

Since ρ is assumed to be strongly continuous and finite-dimensional, it is continuous in the usual sense if we equip $\mathcal{U}(\mathbb{C}^n) \subset \mathrm{GL}(\mathbb{C}^n)$ with the standard Euclidean topology. Notice that this, combined with the continuity of the maps $G^2 \ni (g_1, g_2) \mapsto g_1 g_2 \in G$ as well as $G^2 \ni (g_1, g_2) \mapsto \rho(g_1) \rho(g_2) \in \mathcal{U}(\mathbb{C}^n)$, implies also the continuity of

$$G^2 \ni (g_1, g_2) \mapsto \frac{1}{n} \mathrm{tr} \left(\rho(g_1 g_2) (\rho(g_1) \rho(g_2))^{-1} \right) = \mu(g_1, g_2) \in \{z \in \mathbb{C} : |z| = 1\}.$$

In the sequel, we assume furthermore that $\rho(\mathbf{1}_G) = \mathbf{1}_{\mathbb{C}^n}$ (if this is not the case, multiply ρ by a suitable constant of modulus one, which yields again a projective representation).

The proof consists of two main steps. In the first step, we prove the claim locally at the identity $\mathbf{1}_G$ of G . In the second step we bootstrap the construction based on the connectedness properties of G and the local homomorphism from the first step, to obtain a global representation of G with the desired properties.

Let's start with the local step. Among the key observations is the simple fact that

$$G \ni g \mapsto \det \rho(g) \in \{z \in \mathbb{C} : |z| = 1\}$$

is continuous. Since $\det \rho(\mathbf{1}_G) = 1$, continuity implies the existence of a small open neighborhood $V \subset G$, $\mathbf{1}_G \in V$, in which $|\det \rho(g) - 1|$ is small enough, so that $V \ni g \mapsto \lambda(g) = (\det \rho(g))^{-1/n}$ is well-defined and continuous. We then simply set

$$\tau(g) = \lambda(g) \rho(g) \in \mathcal{U}(\mathcal{H}) \tag{1.36}$$

for every $g \in V$ so that by the properties of ρ , we have that

$$\tau(g_1 g_2) = \frac{\lambda(g_1 g_2) \mu(g_1 g_2)}{\lambda(g_1) \lambda(g_2)} \tau(g_1) \tau(g_2).$$

Using that $\det \tau(g) = 1$, by definition of λ , this implies that

$$\left(\frac{\lambda(g_1 g_2) \mu(g_1 g_2)}{\lambda(g_1) \lambda(g_2)} \right)^n = 1.$$

Assuming $V \subset G$ to so small such that the last step implies $\frac{\lambda(g_1 g_2) \mu(g_1 g_2)}{\lambda(g_1) \lambda(g_2)} = 1$, we find

$$\tau(g_1 g_2) = \tau(g_1) \tau(g_2), \quad \forall g_1, g_2 \in V. \tag{1.37}$$

Loosely speaking, this means ρ can be lifted locally to a homomorphism $\tau : V \rightarrow \mathcal{U}(\mathcal{H})$.

Our next goal is to show that $\tau : V \rightarrow \mathcal{U}(\mathcal{H})$ can be extended to a global homomorphism $\tau : G \rightarrow \mathcal{U}(\mathcal{H})$ in such a way that ρ is induced by τ . Here, we use the connectedness properties of G . In particular, without loss of generality, we assume in the following that $V \subset G$ is path connected.

The key step is how to extend τ to G . Afterwards, straightforward arguments imply that the extension corresponds to a de-projectification of ρ . So, let's first extend τ from V to all of G . Since $\tau(\mathbf{1}_G) = \mathbf{1}_{\mathbb{C}^n}$, a natural strategy is to choose for general $g \in G$ a continuous path $\gamma : [0, 1] \rightarrow G$ with $\gamma(0) = \mathbf{1}_G$, $\gamma(1) = g$ and to define $\tau(g)$ as

$$\tau(g, \gamma) = \tau(\gamma_{t_k} \gamma_{t_{k-1}}^{-1}) \tau(\gamma_{t_{k-1}} \gamma_{t_{k-2}}^{-1}) \dots \tau(\gamma_{t_1} \gamma_{t_0}^{-1}), \quad (1.38)$$

where $0 = t_0 < t_1 < t_2 < \dots < t_k = 1$ is a partition that is fine enough to ensure that $\gamma_{t_i} \gamma_{t_{i-1}}^{-1} \in V$ for every $i \in \{1, \dots, k\}$. We leave it as a simple *exercise* to show that (1.37) implies that a refinement of the partition does not change the r.h.s. in (1.38). Thus, $\tau(g, \gamma)$ only depends on γ and not on the specific partition that we choose. In the next step, we show that the simple connectedness of G implies that

$$\tau(g, \gamma_1) = \tau(g, \gamma_2) \quad (1.39)$$

for every pair of continuous paths $\gamma_i : [0, 1] \rightarrow G$ s.t. $\gamma_i(0) = \mathbf{1}_G$, $\gamma_i(1) = g$, $i \in \{1, 2\}$. Assuming for the moment the validity of (1.39), we can then define

$$\tau(g) = \tau(g, \gamma_1)$$

for some (and hence all) $\gamma : [0, 1] \rightarrow G$ s.t. $\gamma(0) = \mathbf{1}_G$, $\gamma(1) = g$. Since V is path connected, note that this coincides with our earlier definition $\tau : V \rightarrow \mathcal{U}(\mathcal{H})$. Furthermore, note that for two paths $\gamma_1 : [0, 1] \rightarrow G$ with $\gamma_1(0) = \mathbf{1}_G$, $\gamma_1(1) = g_1 \in G$ and $\gamma_2 : [0, 1] \rightarrow G$ with $\gamma_2(0) = \mathbf{1}_G$, $\gamma_2(1) = g_2 \in G$, we can form a path $\gamma : [0, 1] \rightarrow G$ with $\gamma(0) = \mathbf{1}_G$, $\gamma(1) = g_1 g_2 \in G$ as the concatenation $\gamma = \gamma_1 \circ (-\gamma_2) \circ \gamma_2$. Then (1.38) and (1.39) imply (*exercise*) that

$$\tau(g_1 g_2) = \tau(g_1 g_2, \gamma) = \tau(g_1, \gamma_1) \tau(g_2, \gamma_2) = \tau(g_1) \tau(g_2),$$

so that $\tau : G \rightarrow \mathcal{U}(\mathcal{H})$ is a representation of G . Similarly, (1.36) and (1.38) imply that for every $g \in G$, there exists $\lambda'(g) \in \mathbb{C}$ with $|\lambda'(g)| = 1$ such that

$$\begin{aligned} \tau(g) &= \tau(\gamma_{t_k} \gamma_{t_{k-1}}^{-1}) \tau(\gamma_{t_{k-1}} \gamma_{t_{k-2}}^{-1}) \dots \tau(\gamma_{t_1} \gamma_{t_0}^{-1}) \\ &= \lambda^{-1}(\gamma_{t_k} \gamma_{t_{k-1}}^{-1}) \rho(\gamma_{t_k} \gamma_{t_{k-1}}^{-1}) \dots \lambda^{-1}(\gamma_{t_1} \gamma_{t_0}^{-1}) \rho(\gamma_{t_1} \gamma_{t_0}^{-1}) \\ &= \lambda'(g) \rho(g). \end{aligned}$$

In other words, $\tau : G \rightarrow \mathcal{U}(\mathcal{H})$ is a de-projectification of $\rho : G \rightarrow \mathcal{U}(\mathcal{H})$. Observe, moreover, that the map $G \ni g \mapsto \lambda'(g)$ is continuous.

To finish the proof, it remains to verify that $\tau(g, \gamma)$ is independent of $\gamma : [0, 1] \rightarrow G$ with $\gamma(0) = \mathbf{1}_G$, $\gamma(1) = g$. By standard concatenation, this is equivalent to showing that $\tau(\mathbf{1}_G, \gamma) = \mathbf{1}_{\mathbb{C}^n}$ for every loop $\gamma : [0, 1] \rightarrow G$ starting and ending at $\mathbf{1}_G$. So consider such a loop and choose a homotopy $\psi : [0, 1] \times [0, 1] \rightarrow G$ between γ and the constant loop $[0, 1] \ni t \mapsto \gamma_0(t) = \mathbf{1}_G$, by the simple connectedness of G . Note that trivially $\tau(\mathbf{1}_G, \gamma_0) = \mathbf{1}_{\mathbb{C}^n}$. We now show that for $|u_1 - u_2|$ sufficiently small, we also have that

$$\tau(g, \psi_{u_1}) = \tau(g, \psi_{u_2}), \quad (1.40)$$

where $\psi_u(t) = \psi(u, t)$ such that $\psi_u(0) = \psi_u(1) = \mathbf{1}_G$, for every $u \in [0, 1]$, and such that $\psi_0 = \gamma_0$, $\psi_1 = \gamma$. Once (1.40) is proved, the path independence of $\tau(g, \gamma)$ follows.

To prove (1.40), we choose a partition $0 = t_0 < t_1 < t_2 < \dots < t_k = 1$ so that

$$\tau(g, \psi_{u_1}) = \tau(\psi_{u_1}(t_k)\psi_{u_1}(t_{k-1})^{-1}) \dots \tau(\psi_{u_1}(t_1)\psi_{u_1}(t_0)^{-1}), \quad \psi_{u_1}(t_i)\psi_{u_1}(t_{i-1})^{-1} \in V$$

and

$$\tau(g, \psi_{u_2}) = \tau(\psi_{u_2}(t_k)\psi_{u_2}(t_{k-1})^{-1}) \dots \tau(\psi_{u_2}(t_1)\psi_{u_2}(t_0)^{-1}), \quad \psi_{u_2}(t_i)\psi_{u_2}(t_{i-1})^{-1} \in V$$

for every $i \in \{1, \dots, k\}$. Continuity and (1.37) then also imply that

$$\psi_{u_1}(s)\psi_{u_1}(t)^{-1}, \quad \psi_{u_2}(s)\psi_{u_2}(t)^{-1} \in V, \quad \forall t_i \leq s, t \leq t_{i+1}, i \in \{0, 1, \dots, k-1\}.$$

Now, for every fixed $i \in \{1, \dots, k\}$, write

$$\psi_{u_1}(t_i)\psi_{u_1}(t_{i-1})^{-1} = (\psi_{u_1}(t_i)\psi_{u_2}(t_i)^{-1})(\psi_{u_2}(t_i)\psi_{u_2}(t_{i-1})^{-1})(\psi_{u_2}(t_{i-1})\psi_{u_1}(t_{i-1})^{-1})$$

and notice that for $|u_1 - u_2|$ sufficiently small, the continuity of the homotopy ψ implies that all three elements in the brackets on the r.h.s. belong to V . Applying (1.37) yields

$$\begin{aligned} & \tau(\psi_{u_1}(t_i)\psi_{u_1}(t_{i-1})^{-1}) \\ &= \tau(\psi_{u_1}(t_i)\psi_{u_2}(t_i)^{-1})\tau(\psi_{u_2}(t_i)\psi_{u_2}(t_{i-1})^{-1})\tau(\psi_{u_2}(t_{i-1})\psi_{u_1}(t_{i-1})^{-1}). \end{aligned}$$

Similarly, continuity of ψ implies that

$$(\psi_{u_2}(t_{i-1})\psi_{u_1}(t_{i-1})^{-1})^{-1} = \psi_{u_1}(t_{i-1})\psi_{u_2}(t_{i-1})^{-1} \in V$$

whenever $|u_1 - u_2|$ is small so that another application of (1.37) shows that

$$\begin{aligned} & \tau(\psi_{u_1}(t_k)\psi_{u_1}(t_{k-1})^{-1}) \dots \tau(\psi_{u_1}(t_1)\psi_{u_1}(t_0)^{-1}) \\ &= \tau(\mathbf{1}_G)\tau(\psi_{u_2}(t_k)\psi_{u_2}(t_{k-1})^{-1})\tau(\psi_{u_2}(t_{k-1})\psi_{u_1}(t_{k-1})^{-1}) \\ & \quad \times \tau(\psi_{u_1}(t_{k-1})\psi_{u_2}(t_{k-1})^{-1})\tau(\psi_{u_2}(t_{k-1})\psi_{u_2}(t_{k-2})^{-1})\tau(\psi_{u_2}(t_{k-2})\psi_{u_1}(t_{k-2})^{-1}) \\ & \quad \times \dots \times \tau(\psi_{u_1}(t_1)\psi_{u_2}(t_1)^{-1})\tau(\psi_{u_2}(t_1)\psi_{u_2}(t_0)^{-1})\tau(\mathbf{1}_G) \\ &= \tau(\psi_{u_2}(t_k)\psi_{u_2}(t_{k-1})^{-1}) \dots \tau(\psi_{u_2}(t_1)\psi_{u_2}(t_0)^{-1}). \end{aligned}$$

This shows that $\tau(g, \psi_{u_1}) = \tau(g, \psi_{u_2})$ whenever $|u_1 - u_2|$ is sufficiently small. $\square$

Consider now once more the transformation behavior of a rotation invariant quantum system whose symmetry is described by a projective unitary representation

$$\rho(R)\psi = T(R)\psi(R^{-1}.), \quad \forall \psi \in L^2(\mathbb{R}^3, \mathbb{C}^n) \quad (1.41)$$

as in (1.31), where $T : \text{SO}(3) \rightarrow \text{SL}(n, \mathbb{C})$ is some finite-dimensional projective representation of $\text{SO}(3)$. Applying the unitarian trick, i.e. replacing $T : \text{SO}(3) \rightarrow \text{SL}(n, \mathbb{C})$ by the equivalent unitary representation

$$\text{SO}(3) \ni R \mapsto \left(\int_{\text{SO}(3)} \mu(dG) T(G)T(G)^* \right)^{-\frac{1}{2}} T(R) \left(\int_{\text{SO}(3)} \mu(dG) T(G)T(G)^* \right)^{\frac{1}{2}},$$

where μ denotes the normalized Haar measure on $\text{SO}(3)$ (see e.g. [14, Chapter 14]), we may assume without loss of generality that $T : \text{SO}(3) \rightarrow \mathcal{U}(\mathbb{C}^n)$ is unitary. Corollary 1.1 then tells us that we can identify T with an ordinary representation $U : \text{SU}(2) \rightarrow \mathcal{U}(\mathcal{H})$. Instead of T , we may thus study the ordinary unitary representation

$$\text{SU}(2) \ni A \mapsto U(A)\psi(R(A)^{-1}\cdot), \quad \forall \psi \in L^2(\mathbb{R}^3, \mathbb{C}^n). \quad (1.42)$$

Conversely, notice that every (projective) representation $U : \text{SU}(2) \rightarrow \mathcal{U}(\mathcal{H})$ such that $U(A) = U(B)$ whenever $R(A) = R(B)$ gives rise to a projective representation of $\text{SO}(3)$ (cf. [17, Prop. 12.69]). In other words, in view of the classification of quantum systems with rotational symmetry $\text{SO}(3)$, we can directly analyze the ordinary representations of $\text{SU}(2)$, without losing any physically relevant information. Although the original symmetry group is replaced by a more abstract group, it is mathematically easier to work with ordinary representations. In the latter case, it is for instance straightforward to associate to each Lie algebra generator a quantum observable, based on Stone's theorem. There are also fundamental results like Nelson's theorem (see e.g. [20] and [17, Prop. 12.85]) that support the viewpoint to take the covering group as a starting point. Finally, generalized transformation laws such as (1.42) naturally lead to the idea of abstract groups as models for internal symmetries that have no relation to classical coordinate transformations. This idea is central for the formulation of modern gauge theories which are quantizations of classical field theories equipped with an internal symmetry group G , where fields are modeled as connection forms on a principal G -bundle.

The previous discussion motivates to study the unitary representations of $\text{SU}(2)$. We defer this to Section 3.2 and conclude this section with some further examples instead.

Example 1.8 (Spin- $\frac{1}{2}$ Particles). Consider the state space $\mathcal{H} = L^2(\mathbb{R}^3, \mathbb{C}^2)$, the defining representation $\text{SU}(2) \ni A \mapsto A \in \mathbb{C}^{2 \times 2}$ of $\text{SU}(2)$ and the induced representation ρ on $\mathcal{H}$, defined through (1.42) for $n = 2$. This model is used to describe elementary particles such as electrons, positrons or neutrons (non-relativistically). Based on ρ , we obtain a generalized angular momentum, say in the direction of $e_3 = (0, 0, 1) \in \mathbb{R}^3$, by

$$(J_3\psi)(x) = i \lim_{\omega \rightarrow 0} \frac{1}{\omega} \left(e^{-\frac{i}{2}\sigma_3\omega} \psi(e^{-\omega X_3}x) - \psi(x) \right) = \left(L_3 + \frac{\sigma_3}{2} \right) \psi(x), \quad \text{for a.e. } x \in \mathbb{R}^3.$$

Here, we used Problem 1.5 which shows that $R(e^{-\frac{i}{2}\sigma_3\omega}) = e^{-\omega X_3}$. The operator $S_3 = \frac{1}{2}\sigma_3$ is called the spin operator in direction e_3 . Similarly, one defines spin operators S_1 and S_2 and collects them into the spin vector $S = (S_1, S_2, S_3)$ which measures a quantum mechanical property that has no classical counterpart. In analogy to spatial angular momentum, one interprets spin as some internal angular momentum. It is readily verified that the spectrum $\text{spec}(S_j) = \{-s, s\}$ for $s = \frac{1}{2}$ so that the spin in each coordinate direction is quantized. Particles as above are said to be of spin $s = \frac{1}{2}$, a terminology related to the representations of $\text{SU}(2)$ and explained in more detail in Section 3.2.

Spin is an important property used to explain e.g. the Stern-Gerlach experiment or the fine structure of atomic spectra. To describe energetic effects in the non-relativistic setting, a suitable Hamiltonian that takes spin into account is the Pauli-Hamiltonian

$$H = \frac{1}{2m} |i\hbar\nabla + \frac{e}{c}A(x)|^2 + e\Phi(x) - \frac{e}{mc} S \cdot B(x),$$

where $B(x) = \nabla \times A(x)$ denotes the magnetic field generated by A . It is worth to note that for $A(x) = (B \times x)/2$, so that $B(x) = B = \text{const.}$ for all $x \in \mathbb{R}^3$, one has that

$$\frac{1}{2m} |i\hbar \nabla + \frac{e}{c} A(x)|^2 = \frac{\hbar^2}{2m} (-\Delta) - \frac{\hbar e}{2mc} L \cdot B + O(B^2).$$

What this shows is that the spin S couples in the Pauli-Hamiltonian to the magnetic field B like a small constant magnetic field couples to the spatial angular momentum in the Hamiltonian of a massive, charged particle that moves in an electromagnetic field.

Example 1.9 (Indistinguishable Particles). Despite the spin, which is related to the internal symmetry group $\text{SU}(2)$, there exist also other symmetries that have no classical counterpart. Consider for instance a system of $N \in \mathbb{N}$ particles moving in $\mathbb{R}^d$. We saw earlier that a natural state space is $\mathcal{H} = \bigotimes_{i=1}^N L^2(\mathbb{R}^d) = L^2(\mathbb{R}^{dN})$. If $\psi \in \mathcal{H}$ describes the state of the system, it enables us via (1.12) to predict the probabilities for certain measurement outcomes. But what if the N particles are indistinguishable, based on their basic properties such as their mass, spin or charge?

In this case, it seems reasonable to postulate that permuting the particles is a symmetry of the system. Since the permutation group $\mathcal{S}_N$ of N elements acts on $\mathbb{R}^{dN}$ via

$$\pi(x_1, \dots, x_N) = (x_{\pi(1)}, \dots, x_{\pi(N)}), \quad \forall \pi \in \mathcal{S}_N, (x_1, \dots, x_N) \in \mathbb{R}^{dN},$$

we can define the associated unitary permutation operators $U_\pi : \mathcal{H} \rightarrow \mathcal{H}$ by

$$(U_\pi \psi)(x_1, \dots, x_N) = \psi(x_{\pi(1)}, \dots, x_{\pi(N)}).$$

Now, in general, we clearly have $\psi \not\sim U_\pi \psi$. So, in the setting of indistinguishable particles, it is natural to restrict the state space to those wave functions ψ that satisfy

$$U_\pi \psi \sim \psi \iff U_\pi \psi = \lambda_\pi \psi$$

for every $\pi \in \mathcal{S}_N$, where $\lambda_\pi \in \{z \in \mathbb{C} : |z| = 1\}$. Since $U_{\pi_1 \circ \pi_2} = U_{\pi_1} U_{\pi_2}$, this implies

$$\lambda_{\pi_1 \circ \pi_2} = \lambda_{\pi_1} \lambda_{\pi_2}, \quad \forall \pi_1, \pi_2 \in \mathcal{S}_N,$$

so that $\lambda : \mathcal{S}_N \rightarrow \{z \in \mathbb{C} : |z| = 1\}$ defines a one dimensional representation of $\mathcal{S}_N$. Now, it is well-known that every permutation $\pi = \tau_1 \circ \dots \circ \tau_k$ is equal to a product of transpositions $(\tau_j)_{j=1}^k$, $k \leq N$, swapping exactly two elements. For a transposition τ , on the other hand, the homomorphism property implies $\lambda_{\tau \circ \tau} = \lambda_\tau^2 = \lambda(\mathbf{1}_{\mathcal{S}_N}) = 1$, so that $\lambda_\tau \in \{-1, 1\}$ and thus $\lambda_\pi \in \{-1, 1\}$ for every $\pi \in \mathcal{S}_N$. Furthermore, based on the simple observation $(j, k) \circ (j, l) = (j, l) \circ (l, k)$ for every $j, k, l \in \{1, \dots, N\}$, it follows that

$$\lambda_{(j,k)} = \lambda_{(l,k)} = \lambda_{(k,l)} = \lambda_{(n,l)} = \lambda_{(l,n)}$$

for all $j, k, l, n \in \{1, \dots, N\}$, and thus $\lambda_\tau = 1$ for all $\tau \in \mathcal{S}_N$ or $\lambda_\tau = -1$ for all transpositions $\tau \in \mathcal{S}_N$. This determines the one dimensional representations of $\mathcal{S}_N$.

According to each representation, we can now build appropriate physical state spaces (recall that the latter ought to be linear). One calls the particles described by states in

$$L_s^2(\mathbb{R}^{dN}) = \{\psi \in L^2(\mathbb{R}^{dN}) : U_\pi \psi = \psi, \forall \pi \in \mathcal{S}_N\} = \bigotimes_{\text{sym}}^N L^2(\mathbb{R}^d)$$

bosons while fermions are particles that are described by states in

$$L_a^2(\mathbb{R}^{dN}) = \{\psi \in L^2(\mathbb{R}^{dN}) : U_\pi \psi = (-1)^{\deg(\pi)} \psi, \forall \pi \in \mathcal{S}_N\} = \bigwedge^N L^2(\mathbb{R}^d).$$

For more on indistinguishability, see also the discussion in [33, Sections 4.1 & 9.7].

1.3 Special Relativity

In Section 1.1, we described some of the basic ideas of classical mechanics. In particular, Examples 1.1, 1.2 and 1.3 describe the standard Hamiltonians for charged, massive particles interacting with electromagnetic fields. Although a vast range of mechanical and electromagnetic phenomena is well-described by these models, towards the end of the 19th century, fundamental difficulties emerged in the unification of Newton's and Maxwell's theories. These problems were related to the understanding of the propagation of light (e.g. what is the medium in which electromagnetic waves propagate, interpreting electromagnetic waves as a mechanical phenomenon?) and, consequently, of the classical notions of space and time. This led in particular to a reconsideration of the concept of inertial systems (for which Newton's law is supposed to be valid), briefly described in Section 1.1, and the relativity principle which states that the form of physical laws does not depend on the choice of the inertial frame used to describe the physical phenomena.

Anticipating the formalism of special relativity, it is useful to make the assumptions of Newtonian mechanics on space and time more precise using a geometric formulation; we follow [19, Chapter 6]. According to Newtonian mechanics, absolute space is a three dimensional Riemannian manifold (E, g) with a symmetric, positive definite metric $g \in C^\infty(E, (T^*E)^{\otimes 2})$. Newtonian mechanics then postulates that (E, g) is isometric to $(\mathbb{R}^3, \cdot)$, where $\cdot$ denotes the Euclidean inner product. This means that there exists a diffeomorphism $\Phi : E \rightarrow \mathbb{R}^3$ so that

$$g_p(v, w) = d\Phi_p(v) \cdot d\Phi_p(w), \forall v, w \in T_p E.$$

Particles are described by trajectories $[t_0, t_1] \ni t \mapsto \gamma(t) \in E$ and Newton's law states that the force acting on a particle is equal to its acceleration times its mass. In particular, time is absolute in the sense that it is independent of any specific coordinate frame.

Observe that the postulate that space is isometric to $(\mathbb{R}^3, \cdot)$ singles out isometric coordinate frames as preferred coordinates. Such frames correspond to the previously mentioned inertial frames. Indeed, if a trajectory $[t_0, t_1] \ni t \mapsto \gamma(t) \in E$ of a particle of mass $m > 0$ is described in terms of isometric coordinates $y = (y_1, y_2, y_3) : E \rightarrow \mathbb{R}^3$ and

if no force acts on the particle, then Newton's law states that

$$\begin{aligned} m \frac{d^2 \gamma}{dt^2} = 0 &\iff m \left(\frac{d^2(y \circ \gamma)}{dt^2} + \sum_{1 \leq i, j \leq 3} \Gamma_{ij}(\gamma) \frac{d(y_i \circ \gamma)}{dt} \frac{d(y_j \circ \gamma)}{dt} \right) = 0, \\ &\iff \frac{d^2(y \circ \gamma)}{dt^2} = 0. \end{aligned} \quad (1.43)$$

Here, we used that the metric g is locally constant and diagonal in isometric coordinates so that the Christoffel symbols $(\Gamma_{ij}^k)_{i,j,k=1}^3$ vanish identically (for their definition and other basic notions from (semi-) Riemannian geometry, see e.g. [19, Chapter 3]). (1.43) implies that in isometric coordinates particles travel on straight lines $(y \circ \gamma)(t) = at + b$ for suitable $a, b \in \mathbb{R}^3$. By the last identity in (1.43), this describes equivalently the geodesics of (E, g) in isometric coordinates. As a consequence, the Riemannian distance $d_g(p, q)$ between two points $p = y^{-1}(v) \in E$ and $q = y^{-1}(w) \in E$, defined by

$$d_g(p, q) = \inf \{ L(\gamma) : [0, 1] \ni t \mapsto \gamma(t) \in E \text{ piecewise smooth, } \gamma(0) = p, \gamma(1) = q \}$$

for $L(\gamma) = \int_0^1 dt \sqrt{g_{\gamma(t)}(d\gamma/dt, d\gamma/dt)}$, is equal to the standard Euclidean distance

$$d_g(p, q) = d_g(y^{-1}(v), y^{-1}(w)) = |v - w|.$$

Now, suppose we consider a change of coordinates from one inertial frame to another. This corresponds to a Riemannian isometry $\Psi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ which, by the last observation and the invariance of the Riemannian distance under isometries, satisfies

$$|\Psi(v) - \Psi(w)| = |v - w|, \quad \forall v, w \in \mathbb{R}^3.$$

In other words, $\Psi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ is a Euclidean isometry and thus equal to a Euclidean transformation $\Psi(x) = Rx + a$ for some translation $a \in \mathbb{R}^3$ and some orthogonal matrix $R \in O(3)$ (*exercise*). If we can reach the new inertial frame continuously from the original frame, we may assume additionally that $R \in SO(3)$. Put in geometric terms, the isometry group, defined as the group of isometries on (E, g) (with the composition of maps as group multiplication) is isomorphic to the Euclidean group

$$E(3) = \{(a, R) : a \in \mathbb{R}^3, R \in O(3)\} \text{ with } (a, R)(a', R') = (Ra' + a, RR'). \quad (1.44)$$

Finally, in anticipation of the axiomatic transformation behavior of quantum fields, let us also recall the transformation behavior of classical fields with regards to inertial frames in Newtonian mechanics. By definition, classical fields include all tensor fields on (E, g) . Suppose that w.r.t. a fixed inertial frame $u : E \rightarrow \mathbb{R}^3$ a field is represented by

$$\Phi = \Phi^{i_1 \dots i_p; j_1 \dots j_q} du_{i_1} \otimes \dots \otimes du_{i_p} \otimes \frac{\partial}{\partial u_{j_1}} \otimes \dots \otimes \frac{\partial}{\partial u_{j_q}} \in C^\infty(E, (T^*E)^{\otimes p} \otimes (TE)^{\otimes q}), \quad (1.45)$$

where here and in the following we use the Einstein summation convention (indices appearing once as upper and once as lower index are summed over). Recall that the vector fields $\partial/\partial u_j \in C^\infty(E, TE)$, for $j \in \{1, 2, 3\}$, in (1.45) are defined by

$$\frac{\partial}{\partial u_j}|_p = (du^{-1})|_{u(p)} \frac{\partial}{\partial x_j}|_{u(p)}, \quad \forall p \in E.$$

This means that $\partial/\partial u_j \in C^\infty(E, TE)$ is equal to the pushforward of the standard Euclidean vector field $\partial/\partial x_j \in C^\infty(\mathbb{R}^3, T\mathbb{R}^3)$ (which as a vector corresponds to the standard basis vector $e_j \in \mathbb{R}^3$ and as a derivation to the directional derivative in direction $e_j \in \mathbb{R}^3$) by the diffeomorphism $u^{-1} : \mathbb{R}^3 \rightarrow E$. Put differently, if $U \in C^\infty(E, TE)$ is a general vector field of the form

$$U = U^j \frac{\partial}{\partial u_j} \in C^\infty(E, TE),$$

its pushforward $(duU) \circ u^{-1} \in C^\infty(\mathbb{R}^3, T\mathbb{R}^3)$ by $u : E \rightarrow \mathbb{R}^3$, which is interpreted as the coordinate representation of U with regards to the chart $u : E \rightarrow \mathbb{R}^3$, takes values

$$(duU) \circ u^{-1}(v) = du|_{u^{-1}(v)} \left((U^j \circ u^{-1})(v) \frac{\partial}{\partial u_j}|_{u^{-1}(v)} \right) = (U^j \circ u^{-1})(v) \frac{\partial}{\partial x_j}|_v, \quad \forall v \in \mathbb{R}^3.$$

The one forms $du_j \in C^\infty(E, (T^*E))$ in (1.45), on the other hand, correspond to the dual elements which are defined so that $du_j(\partial/\partial u_k) = \delta_{jk}$. This means explicitly that

$$(du_j)|_p = (dx_j)|_{u(p)}(du)|_p, \quad \forall p \in E,$$

so that the pullback $(u^{-1})^*\omega \in C^\infty(\mathbb{R}^3, (T^*\mathbb{R}^3))$ of a general one form

$$\omega = \omega^j du_j \in C^\infty(E, (T^*E)),$$

interpreted as the coordinate representation of ω w.r.t. the chart $u : E \rightarrow \mathbb{R}^3$, satisfies

$$(u^{-1})^*\omega|_v = (\omega^j \circ u^{-1})(v)(du_j)_{u^{-1}(v)} du|_v^{-1} = (\omega^j \circ u^{-1})(v)(dx_j)|_v$$

Problem 1.6. *In the same notation as above, suppose $u : E \rightarrow \mathbb{R}^3$ is a global chart and let $U = U^j \frac{\partial}{\partial u_j} \in C^\infty(E, TE)$, $f \in C^\infty(E)$. Verify that*

$$(Uf) \circ u^{-1} = (U^j \circ u^{-1}) \partial_{x_j}(f \circ u^{-1}).$$

Generalizing the previous remarks to Φ , we see that the coordinate representation Φ_u of the field Φ with regards to the coordinate frame $u : E \rightarrow \mathbb{R}^3$ takes the form

$$\Phi_u = (\Phi^{i_1 \dots i_p; j_1 \dots j_q} \circ u^{-1}) dx_{i_1} \otimes \dots \otimes dx_{i_p} \otimes \frac{\partial}{\partial x_{j_1}} \otimes \dots \otimes \frac{\partial}{\partial x_{j_q}} \in C^\infty(\mathbb{R}^3, (T^*\mathbb{R}^3)^{\otimes p} \otimes (T\mathbb{R}^3)^{\otimes q}).$$

Now, consider a different inertial frame $u' : E \rightarrow \mathbb{R}^3$ w.r.t. which Φ has the form

$$\Phi = (\Phi')^{i_1 \dots i_p; j_1 \dots j_q} du'_{i_1} \otimes \dots \otimes du'_{i_p} \otimes \frac{\partial}{\partial u'_{j_1}} \otimes \dots \otimes \frac{\partial}{\partial u'_{j_q}} \in C^\infty(\mathbb{R}^3, (T^*\mathbb{R}^3)^{\otimes p} \otimes (T\mathbb{R}^3)^{\otimes q}).$$

Then, the chain rule, applied to $(u' \circ u^{-1})(\cdot) = R(\cdot) + a$, for suitable $(a, R) \in E(3)$, implies that the component functions $(\Phi^{i_1 \dots i_p; j_1 \dots j_q} \circ u^{-1})$ and $((\Phi')^{i_1 \dots i_p; j_1 \dots j_q} \circ (u')^{-1})$ of the coordinate representations Φ_u and, respectively, $\Phi_{u'}$ of the field Φ are related by

$$\begin{aligned} & (\Phi^{i_1 \dots i_p; j_1 \dots j_q} \circ u^{-1})(\cdot) \\ &= R_{k_1}^{i_1} \dots R_{k_p}^{i_p} (R^{-1})^{j_1}_{l_1} \dots (R^{-1})^{j_p}_{l_p} ((\Phi')^{k_1 \dots k_p; l_1 \dots l_q} \circ (u')^{-1})(R(\cdot) + a). \end{aligned} \quad (1.46)$$

As mentioned earlier, the transformation behavior (1.46) has a quantum field analogue that is collected as part of the Wightman axioms (see Section 1.4 below).

Despite its simplicity and intuitive character, the Newtonian perspective on space and time outlined above led to fundamental, conceptual difficulties (related for instance to the attempts to explain the Michelson-Morley experiment; for a brief historical overview see e.g. [27, Section 1.6]). The resolution of these problems led ultimately to the theory of special and general relativity by Einstein. In special relativity, one still holds up the relativity principle, but one modifies the notion of inertial frame. A central postulate of special relativity is that the speed of propagation of light, or more generally of any interaction between interacting particles (cf. the discussion in [12, Chapter 1]), is a universal constant that is independent of the inertial frame. Note that in the context of Newtonian mechanics, interactions act instantaneously at an arbitrarily large distance which seems counterintuitive. The postulate of the constancy of the speed of light remedies this and it leads to the following, modified geometric formulation of spacetime.

On a heuristic level, suppose a ray of light is described in two different inertial frames with coordinates $x = (x_0, x_1, x_2, x_3)$ and, respectively, $x' = (x'_0, x'_1, x'_2, x'_3)$ where x_0 and x'_0 measure the time in each frame via the identification $t = x_0/c$, $t' = x'_0/c$, $c > 0$ denoting the speed of light. Note in particular that time is now considered relative to the inertial frame, like the spatial coordinates. Special relativity then postulates that

$$c = \left| \frac{dx}{dt} \right| = \left| \frac{dx'}{dt'} \right| \iff 0 = c^2 |dt|^2 - |dx|^2 = c^2 |dt'|^2 - |dx'|^2.$$

The second equality on the r.h.s. is an infinitesimal version of the identity

$$\eta(x, x) = \eta^{ij} x_i x_j = \eta(x', x'), \quad \text{for } \eta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \in \mathbb{R}^{4 \times 4}. \quad (1.47)$$

In special relativity, this observation is generalized to the following point of view on space and time. Spacetime is a four dimensional time-oriented Lorentz manifold (E, g) which is isometric to the (flat) semi-Riemannian manifold $(\mathbb{R}^4, \eta)$. Here, η is identified with

the constant (symmetric, non-degenerate) metric tensor with components as in (1.47) with regards to the Euclidean standard basis and it is called the Minkowski metric. Accordingly, the space $(\mathbb{R}^4, \eta)$ is called Minkowski space. Identifying the tangent spaces of $(\mathbb{R}^4, \eta)$ canonically with $T_x \mathbb{R}^4 \simeq \mathbb{R}^4$, for all $x \in \mathbb{R}^4$, tangent vectors $v \in \mathbb{R}^4$ are future-oriented if $x_0 > 0$ and past-oriented if $x_0 < 0$. This defines the time-orientation of $(\mathbb{R}^4, \eta)$ (cf. [19, Chapter 5]).

As an aside, let us also mention that the theory of general relativity generalizes the previous postulate and interprets spacetime as a general, time-oriented Lorentz manifold whose curvature determines gravity. Freely falling bodies follow the geodesics of spacetime while its curvature itself is dynamically linked to the distribution of matter within it through Einstein's equation (for more details, see e.g. [19, Chapter 12]).

Following the geometric discussion on Newtonian mechanics above, let us proceed similarly for special relativity and record some of its basic aspects. First of all, notice that g defines indeed a symmetric, non-degenerate bilinear form. Its non-degeneracy follows e.g. by isometric equivalence to η and by noting that $\eta(x, y) = 0$ for all $y \in \mathbb{R}^4$ implies that $|x|^2 = \eta(x, Px) = 0$, for $P(x_0, x_1, x_2, x_3) = (x_0, -x_1, -x_2, -x_3)$, so that $x = 0$. According to its semi-definiteness properties, one calls tangent vectors $v \in T_g E$ timelike if $g_p(v, v) > 0$, lightlike if $g_p(v, v) = 0$ and spacelike if $g_p(v, v) < 0$. Physically, a timelike vector corresponds to a velocity smaller than $c > 0$ and a lightlike vector is a velocity of the speed of light. Spacelike vectors correspond to velocities greater than the speed of light. By definition, spacelike tangent vectors are excluded in the description of massive ($m > 0$) and massless ($m = 0$) particles. Massive particles are described by future directed timelike curves $[\tau_0, \tau_1] \ni \tau \mapsto \gamma(\tau) \in E$, so that $g_{\gamma(\tau)}(d\gamma/d\tau, d\gamma/d\tau) > 0$ for all $\tau \in [\tau_0, \tau_1]$. Massless particles are described by future directed lightlike curves $[\tau_0, \tau_1] \ni \tau \mapsto \gamma(\tau) \in E$, so that $g_{\gamma(\tau)}(d\gamma/d\tau, d\gamma/d\tau) = 0$ for all $\tau \in [\tau_0, \tau_1]$. Finally, as a generalization of Newton's law of inertia, special relativity postulates that freely falling particles (those on which no external forces act) are described by the geodesics of (E, g) .

As in Newtonian mechanics, the isometric coordinates correspond to the preferred inertial frames. Given such a coordinate frame $x = (x_0, x_1, x_2, x_3) : E \rightarrow \mathbb{R}^4$, geodesics $[\tau_0, \tau_1] \ni \tau \mapsto (x \circ \gamma)(\tau)$ take the simple form

$$(x \circ \gamma)(\tau) = a\tau + b$$

for suitable $a, b \in \mathbb{R}^4$. The defining property $d^2\gamma/d\tau^2 = 0$ implies furthermore (*exercise*) that, as in the Riemannian setting, geodesics minimize the so called proper time

$$T(\gamma) = \int_{\tau_0}^{\tau_1} d\tau \sqrt{g_{\gamma(\tau)}(d\gamma/d\tau, d\gamma/d\tau)}$$

among all piecewise smooth, timelike curves $\varphi : [\tau_0, \tau_1] \rightarrow E$ and that the semi-Riemannian distance between two timelike separated points $p = x^{-1}(v), q = x^{-1}(w) \in E$ equals

$$d_g(p, q) = \sqrt{\eta(v - w, v - w)}.$$

Here, timelike, lightlike and spacelike separated points refer to points that can be connected by a timelike, lightlike and, respectively, spacelike curve. In Minkowski space

$(\mathbb{R}^4, \cdot)$, this corresponds to points $v, w \in \mathbb{R}^4$ such that $\eta(v - w, v - w) > 0$ (timelike), $\eta(v - w, v - w) = 0$ (lightlike) and $\eta(v - w, v - w) < 0$ (spacelike). Spacelike separated events cannot causally influence each other (if $\eta(x, x) = |ct|^2 - |\mathbf{x}|^2 < 0$ for $x = (ct, \mathbf{x})$, a light ray starting at $0 \in \mathbb{R}^3$ travels the distance ct until time t , so it does not reach $\mathbf{x}$).

As in the Riemannian setting, if $\Psi = (x' \circ x^{-1}) : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ describes a change of inertial frames (that is, a semi-Riemannian isometry), the above observations imply

$$\eta(\Psi(v) - \Psi(w), \Psi(v) - \Psi(w)) = \eta(v - w, v - w), \quad \forall v, w \in \mathbb{R}^4.$$

Assuming w.l.o.g. that $\Psi(0) = 0$ (by shifting Ψ to $\Psi' = \Psi - \Psi(0)$ if necessary), bilinearity of the metric implies $\eta(\Psi(v), \Psi(w)) = \eta(v, w)$, for every $v, w \in \mathbb{R}^4$. From here, it is straightforward to deduce (*exercise*) that $\Psi : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ is affine-linear and that the isometry group $\text{ISO}(1, 3)$ of (E, g) is isomorphic to the Poincaré group $\mathcal{P}$, defined by

$$\mathcal{P} = \{(a, L) : a \in \mathbb{R}^4, L \in \text{O}(1, 3)\} \quad \text{with} \quad (a, L)(a', L') = (La' + a, LL'). \quad (1.48)$$

Here, $\text{O}(1, 3)$ denotes the group of linear maps $L : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ that satisfy

$$\eta(L(v), L(w)) = \eta(v, w), \quad \forall v, w \in \mathbb{R}^4. \quad (1.49)$$

The elements of $\text{O}(1, 3)$ are called Lorentz transformations. In special relativity, fields transform as in (1.46), replacing $(a, R) \in \text{E}(3)$ by $(a, L) \in \mathcal{P}$.

The analogue of $\text{SO}(3)$ in Newtonian mechanics (rotations that are path-connected to the identity $\mathbf{1}_{\mathbb{R}^3} \in \text{SO}(3)$), consists of the proper Lorentz group $\mathcal{L}_+^\uparrow$, defined by

$$\mathcal{L}_+^\uparrow = \{L = (L_{\mu\nu})_{\mu,\nu=0}^3 \in \text{O}(1, 3) : L_{00} > 0, \det L = 1\}. \quad (1.50)$$

The condition $\det L = 1$ is motivated as in Newtonian mechanics, noting the embedding

$$\text{O}(3) \ni R \mapsto \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \in \text{O}(1, 3).$$

The condition $L_{00} \geq 0$ ensures that L maps future oriented, time- or lightlike vectors to themselves. Observe here that $L \in \text{O}(1, 3)$ implies that L^{-1} exists and that both $L^T, L^{-1} \in \text{O}(1, 3)$ (*exercise*) as well as

$$\eta^{\mu\nu} L_{\mu\lambda} L_{\nu\kappa} = \eta_{\lambda\kappa} = \eta^{\mu\nu} L_{\lambda\mu} L_{\kappa\nu} \quad \xrightarrow{\lambda, \kappa=0} \quad L_{00}^2 = 1 + \sum_{j=1}^3 L_{j0}^2 = 1 + \sum_{j=1}^3 L_{0j}^2.$$

Hence, if $x = (x_0, x_1, x_2, x_3) = (x_0, \mathbf{x})$ is future oriented (i.e. $x_0 > 0$) with $\eta(x, x) = \eta(Lx, Lx) \geq 0$, we also have that

$$(Lx)_0 = L_{00}x_0 + \sum_{j=1}^3 L_{0j}x_j > |\mathbf{L}_0 \cdot \mathbf{x}| - \sum_{j=1}^3 L_{0,j}x_j \geq 0.$$

Moreover, if $L, M \in \mathcal{L}_+^\uparrow$, the fact that $L_{00}, M_{00} \geq 1$ and Cauchy-Schwarz imply that

$$\left| \sum_{j=1}^3 L_{0j} M_{j0} \right|^2 \leq \left(\sum_{j=1}^3 L_{0j}^2 \right) \left(\sum_{j=1}^3 M_{j0}^2 \right) = (L_{00}^2 - 1)(M_{00}^2 - 1) < L_{00}^2 M_{00}^2,$$

which implies

$$(LM)_{00} = L_{00} M_{00} + \sum_{j=1}^3 L_{0j} M_{j0} > L_{00} M_{00} - \left| \sum_{j=1}^3 L_{0j} M_{j0} \right| > 0.$$

Since $\det ML = \det M \det L = 1$, this verifies that $\mathcal{L}_+^\uparrow$ is indeed a subgroup of $O(1, 3)$. Further properties of $\mathcal{L}_+^\uparrow$ as well as the associated group $\mathcal{P}_+^\uparrow = \{(a, L) \in \mathcal{P} : L \in \mathcal{L}_+^\uparrow\}$ are discussed in detail in Sections 2.3 and 3.2.

Let us conclude this section by connecting special relativity to Newtonian mechanics. As already mentioned, in special relativity freely falling bodies are postulated to follow the geodesics of spacetime. Choosing a fixed inertial frame, suppose that $[\tau_0, \tau_1] \mapsto x(\tau) \in \mathbb{R}^4$ describes such a geodesic. If the particle is massless, we have by definition $\eta(dx/d\tau, dx/d\tau) = 0$. Since linear reparametrizations $\tau \mapsto a\tau + b$ map geodesics to geodesics, we have some freedom to choose a particular parametrization for massive particles. In order for the theory to be consistent with Newton mechanics, one parametrizes $[\tau_0, \tau_1] \mapsto x(\tau) \in \mathbb{R}^4$ conventionally in such a way that

$$\eta(dx/d\tau, dx/d\tau) = m^2 c^2, \quad \forall \tau \in [\tau_0, \tau_1].$$

This is always possible (*exercise*) and in such a parametrization one refers to the derivative $p = (p_0, \mathbf{p}) = dx/d\tau$ as the four-momentum. The geodesic equation can then be formulated as the Hamiltonian dynamics

$$\frac{dx}{d\tau} = p, \quad \frac{dp}{d\tau} = 0. \quad (1.51)$$

Recalling that massive particles are described by forward oriented curves, this implies

$$\eta(p(\tau), p(\tau)) = p_0^2 - |\mathbf{p}|^2 = m^2 c^2, \quad p_0(\tau) > 0, \quad \forall \tau \in [\tau_0, \tau_1]. \quad (1.52)$$

The connection to Newtonian mechanics follows from combining the dynamics (1.51) and the relativistic energy relation (1.52) to eliminate p_0 and rescale proper time τ in terms of the local time $x_0/c = t$. This yields

$$p_0 = \omega(\mathbf{p}) = \sqrt{|\mathbf{p}|^2 + m^2 c^2}, \quad \frac{dx_0}{d\tau} = \omega(\mathbf{p}) (> 0).$$

By the inverse function theorem, the second identity implies that we can express the relativistic time τ as a function of $x_0 = ct$ such that $\frac{d\tau}{dt} = \frac{c}{\omega(\mathbf{p})}$, and the chain rule then implies that the remaining, spatial variables $(\mathbf{x}, \mathbf{p})$ satisfy

$$\frac{d\mathbf{x}}{dt} = \frac{d\mathbf{x}}{d\tau} \frac{d\tau}{dt} = \frac{c\mathbf{p}}{\omega(\mathbf{p})}, \quad \frac{d\mathbf{p}}{dt} = 0.$$

This corresponds to a Hamiltonian system on $\mathbb{R}^3 \times \mathbb{R}^3$ with Hamiltonian $(\mathbf{x}, \mathbf{p}) \mapsto H(\mathbf{x}, \mathbf{p}) = c\omega(\mathbf{p})$. Conservation of the energy $E = H(\mathbf{x}, \mathbf{p})$ under the Hamiltonian flow $t \mapsto (\mathbf{x}(t), \mathbf{p}(t)) \in \mathbb{R}^3 \times \mathbb{R}^3$ implies the identity $p = (p_0, \mathbf{p}) = (E/c, \mathbf{p})$. For this reason, p is also called the energy-momentum vector. Finally, observe that

$$E = c\omega(\mathbf{p}) = mc^2 \sqrt{1 + \frac{|\mathbf{p}|^2}{m^2 c^2}} = mc^2 + \frac{|\mathbf{p}|^2}{2m} + O((|\mathbf{v}|/c)^4)$$

for $\mathbf{v} = d\mathbf{x}/dt$. In other words, we recover the Newtonian expression for the kinetic energy of a free particle (up to the so called rest mass) to leading order in $|\mathbf{v}|/c$. The classical Newtonian dynamics for massive particles thus follows in the non-relativistic limit $|\mathbf{v}|/c \rightarrow 0$. For massless particles, on the other hand, the non-relativistic approximation can not be applied. While the Hamiltonian formulation remains valid, we obtain from $\eta(p, p) = 0$ that $E = |\mathbf{p}|c$ and hence $\mathbf{v} = d\mathbf{x}/dt = c\mathbf{p}/|\mathbf{p}|$ so that $|\mathbf{v}|/c = 1$. In other words, massless particles travel with the speed of light.

1.3.1 Examples of Classical Fields

Compared to Newtonian mechanics whose initial focus lies on describing the trajectories of interacting massive particles, modern particle physics focuses rather on fields as fundamental entities while particles are interpreted as quantum excitations of the fields (once the latter are quantized). In Section 3, this viewpoint is illustrated for three important types of fields. In this section, we introduce the corresponding classical fields.

In Example (1.3), we introduced the electromagnetic field so let's discuss it first in the context of special relativity. Throughout this section, we work without loss of generality directly in Minkowski space $(\mathbb{R}^4, \eta)$ and recall that changes of inertial frames are determined by the Poincaré group $\mathcal{P}$. As mentioned in the previous section, classical fields correspond to tensor fields on $(\mathbb{R}^4, \eta)$. In practice, on the other hand, fields are typically described as solutions of suitable partial differential equations such as (1.3). If such solutions are supposed to describe a physical field, the equations must be invariant under Poincaré transformations. By definition, this means that if a field is represented by components $(\Phi^{i_1 \dots i_p; j_1 \dots j_q})_{i_1, \dots, i_p, j_1, \dots, j_q} \in C^\infty(\mathbb{R}^4)$, its transformed components, determined by (1.46) (with $(a, R) \in \mathbf{E}(3)$ replaced by $(a, L) \in \mathcal{P}$), must solve the same equations in $\mathbb{R}^4$. This is in particular the case if the field equations admit a geometric (i.e. coordinate free) formulation. Let's make this precise for the electromagnetic field.

In case of the electromagnetic field, the field equations are Maxwell's equations (1.3). Based on a relativistic generalization of the Lorentz force (recall (1.6) and see e.g. [5, Section 7.2.3]), it is natural to associate the field with a rank two tensor (as a relativistic force, both its input and output are smooth vector fields so that, by duality, it can be identified with a type-two tensor which turns out to be antisymmetric). It turns out that the identification of the field with a two form $F \in \Omega^2(\mathbb{R}^4)$ leads to a geometric reformulation of Maxwell's equations (1.3).

So, consider smooth (time-dependent) electric $\mathbf{E} = (E_1, E_2, E_3) \in C^\infty(\mathbb{R}^4, \mathbb{R}^3)$ and magnetic $\mathbf{B} = (B_1, B_2, B_3) \in C^\infty(\mathbb{R}^4, \mathbb{R}^3)$ fields as well as smooth charge $\rho \in C^\infty(\mathbb{R}^4)$

and current densities $\mathbf{j} = (j_1, j_2, j_3) \in C^\infty(\mathbb{R}^4, \mathbb{R}^3)$. Then, one defines the electromagnetic field tensor $F \in \Omega^2(\mathbb{R}^4)$ by

$$\begin{aligned} F = & E_1 dx_0 \wedge dx_1 + E_2 dx_0 \wedge dx_2 + E_3 dx_0 \wedge dx_3 \\ & + B_1 dx_2 \wedge dx_3 - B_2 dx_1 \wedge dx_3 + B_3 dx_1 \wedge dx_2, \end{aligned} \quad (1.53)$$

its Hodge- $\star$ (see e.g. [18, Section 7.9.2]) dual form $\star F \in \Omega^2(\mathbb{R}^4)$ by

$$\begin{aligned} \star F = & -B_1 dx_0 \wedge dx_1 - B_2 dx_0 \wedge dx_2 - B_3 dx_0 \wedge dx_3 \\ & - E_1 dx_2 \wedge dx_3 + E_2 dx_1 \wedge dx_3 - E_3 dx_1 \wedge dx_2 \end{aligned} \quad (1.54)$$

and, finally, the current differential form $j \in \Omega^3(\mathbb{R}^4)$ by

$$\begin{aligned} j = & \rho dx_1 \wedge dx_2 \wedge dx_3 - j_1 dx_0 \wedge dx_2 \wedge dx_3 \\ & + j_2 dx_0 \wedge dx_1 \wedge dx_3 - j_3 dx_0 \wedge dx_1 \wedge dx_2. \end{aligned} \quad (1.55)$$

Recalling $x_0 = ct$ so that $\partial_{x_0} = \frac{1}{c}\partial_t$, Maxwell's equations take the following elegant form.

Problem 1.7. *Show explicitly that the equations (1.3), (1.4) are equivalent to*

$$dF = 0, \quad d\star F = j, \quad dj = 0. \quad (1.56)$$

Use this to show that the equations (1.3), (1.4) are invariant under Poincaré transformations (alternatively, verify the invariance explicitly based on (1.46), for $(a, L) \in \mathcal{P}$).

In geometric terms, Problem 1.7 implies that the electromagnetic field can be interpreted as a classical field $F_{\text{EM}} \in \Omega^2(E)$ on spacetime (E, g) satisfying $dF_{\text{EM}} = 0$ and $d\star F_{\text{EM}} = j_{\text{EM}}$ (note that both equations are invariant under taking pullbacks with regards to isometric coordinates) and that the usual Maxwell's equations (1.3) in $\mathbb{R}^4$ describe the corresponding coordinate expressions with regards to a fixed inertial frame. In particular, this yields a coordinate free, physical law for the electromagnetic field which is of fundamental importance in view of the relativity principle.

Generalizing Example 1.3, the equations (1.56) can also be reformulated in terms of the electromagnetic vector potential. Indeed, applying the Poincaré lemma to the closed form $F \in \Omega^2(\mathbb{R}^4)$, we find that there exists a one form $A \in \Omega^1(\mathbb{R}^4)$ such that $F = dA$. We can identify this with a vector $A = (\Phi, \mathbf{A}) = (A_0, A_1, A_2, A_3) \in C^\infty(\mathbb{R}^4)$ via

$$A = \Phi dx_0 + A_1 dx_1 + A_2 dx_2 + A_3 dx_3$$

so that

$$F = \sum_{0 \leq i < j \leq 3} (\partial_{x_i} A_j - \partial_{x_j} A_i) dx_i \wedge dx_j. \quad (1.57)$$

This connects Maxwell's equations with the wave equation in a straightforward way. To see this, notice first that if $F = dA$ solves (1.56), so does $F' = dA'$ for $A' = A + d\chi$ for every $\chi \in C^\infty(\mathbb{R}^4)$, because $dA' = dA + d^2\chi = dA$. This is referred to as the gauge invariance of the electromagnetic field. Since electromagnetic forces are encoded by F , a

gauge transform does not influence the physics of the system. This observation is further developed in modern gauge theories (for more on this, see Section 8.1).

Abbreviating $\partial_\nu = \partial_{x_\nu}$, $\partial^\nu = \eta^{\nu\lambda} \partial_\lambda$ and choosing $\chi \in C^\infty(\mathbb{R}^4)$ such that

$$\partial^\nu \partial_\nu \chi = -\partial^\nu A_\nu \in C^\infty(\mathbb{R}^4),$$

note that $A' = A + d\chi$ satisfies

$$\partial^\nu A'_\nu = \partial^\nu A_\nu + \partial^\nu \partial_\nu \chi = 0.$$

This choice of χ is called Lorenz gauge and can always be arranged. Note indeed that

$$\partial^\nu \partial_\nu = \eta^{\nu\lambda} \partial_\lambda \partial_\nu = \partial_{x_0}^2 - \sum_{j=1}^3 \partial_{x_j}^2 = \frac{1}{c^2} \partial_t^2 - \sum_{j=1}^3 \partial_{x_j}^2 = \square$$

corresponds to the d'Alembertian, the generator of solutions of the wave equation. The upshot is that potentials $A \in \Omega^1(\mathbb{R}^4)$ in Lorenz gauge that solve (1.56) satisfy

$$\square A = j, \quad \partial^\nu A_\nu = 0, \quad (1.58)$$

where $\square A = (\square A_\mu)_{\mu=0}^3$ and $j = (j_\mu)_{\mu=0}^3$. This corresponds componentwise to the wave equation with propagation speed c , the speed of light, and with a vector constraint.

Problem 1.8. *Prove the following statements:*

a) Let $f \in C^\infty(\mathbb{R}^4)$. Define the function $u \in C^\infty(\mathbb{R}^4)$ pointwise by

$$u(x) = u(x_0, \mathbf{x}) = \frac{1}{4\pi x_0} \int_0^{x_0} dy_0 \int_{|\mathbf{x}-\mathbf{y}|=x_0} \sigma(d\mathbf{y}) f(y_0, \mathbf{y}),$$

where $\sigma(\cdot)$ denotes the uniform measure on $\{\mathbf{y} \in \mathbb{R}^3 : |\mathbf{x}-\mathbf{y}| = x_0\} \subset \mathbb{R}^3$. Verify that $\square u = f$. For more on wave equations and a derivation of this formula, see [28].

b) Use Maxwell's equations in Lorenz gauge to deduce the wave equation in (1.58).

c) Let $A \in C^\infty(\mathbb{R}^4, \mathbb{R}^4)$ satisfy (1.58) for $j = 0$ and define for $(a, L) \in \mathcal{P}$

$$A'_\mu(x) = L^\nu{}_\mu A_\nu(Lx + a), \quad \forall x \in \mathbb{R}^4.$$

Show that $A' \in C^\infty(\mathbb{R}^4, \mathbb{R}^4)$ also satisfies (1.58).

Without sources ($j = 0$), the system of equations (1.56) describes the so called free (non-interacting) electromagnetic field. Motivated by the Lorenz gauge formulation, it also turns out to be of independent interest to study scalar valued solutions $\varphi \in C^\infty(\mathbb{R}^4, \mathbb{R})$ (real scalar field) or $\varphi \in C^\infty(\mathbb{R}^4, \mathbb{C})$ (complex scalar field) of the so called Klein-Gordon equation with mass $m > 0$. Such fields solve the generalized wave equation

$$(\square + m^2)\varphi = (\partial^\nu \partial_\nu + m^2)\varphi = 0. \quad (1.59)$$

In quantum field theory, the quantum excitations of the electromagnetic field describe photons while the quantum excitations of fields solving (1.59) correspond to bosons of mass m and spin zero (the connection to spin as well as between m and the mass of the field quanta is explained in Section 3.1). Notice that (1.59) is invariant under inertial coordinate changes. This follows by noting that $\partial^\nu \partial_\nu = \Delta_\eta$ is equal to the Laplace-Beltrami operator related to the Minkowski metric. Alternatively, we readily compute

$$\begin{aligned} (\partial^\nu \partial_\nu + m^2) \varphi(L(\cdot) + a) &= \eta^{\nu\mu} L^\lambda{}_\nu L^\kappa{}_\mu (\partial_\lambda \partial_\kappa) \varphi(L(\cdot) + a) + m^2 \varphi(L(\cdot) + a) \\ &= ((\square + m^2) \varphi)(L(\cdot) + a) = 0. \end{aligned}$$

The quantization of the scalar field is explained in detail in Section 3.1.

Problem 1.9. On $H^1(\mathbb{R}^3, \mathbb{R}) \oplus L^2(\mathbb{R}^3, \mathbb{R})$, define the field Hamiltonian

$$H(\phi, \pi) = \frac{1}{2} \int_{\mathbb{R}^3} dx (|\pi(x)|^2 + |\nabla \phi(x)|^2 + m^2 |\phi(x)|^2).$$

Show that its Fréchet derivative $DH|_{(\phi, \pi)} \in (H^1(\mathbb{R}^3, \mathbb{R}) \oplus L^2(\mathbb{R}^3, \mathbb{R}))^*$ at the point $(\phi, \pi) \in H^1(\mathbb{R}^3, \mathbb{R}) \oplus L^2(\mathbb{R}^3, \mathbb{R})$ is given by

$$DH|_{(\phi, \pi)}(\rho, \sigma) = \langle \nabla \phi, \nabla \rho \rangle_{L^2(\mathbb{R}^3)} + m^2 \langle \phi, \rho \rangle_{L^2(\mathbb{R}^3)} + \langle \pi, \sigma \rangle_{L^2(\mathbb{R}^3)}.$$

Heuristically, if we interpret $(\pi(x))_{x \in \mathbb{R}^3}$ and $(\phi(x))_{x \in \mathbb{R}^3}$ as continuous families of canonical momentum and, respectively, position variables, this means that formally

$$\partial_{\pi(x)} H = DH|_{(\phi, \pi)}(0, \delta_x) = \pi(x), \quad \partial_{\phi(x)} H = DH|_{(\phi, \pi)}(\delta_x, 0) = (-\Delta + m^2)\phi(x).$$

Setting $c = 1$, verify that the (1.59) equals the corresponding Hamiltonian dynamics that is generated by H on the formal phase space $\mathcal{P} = \prod_{x \in \mathbb{R}^3} \mathbb{R} \times \prod_{x \in \mathbb{R}^3} \mathbb{R}$.

We close this section with a brief discussion of classical Dirac fields. Dirac fields are spinor valued fields on Minkowski space and as such beyond the classical notion of a tensor field. The detailed mathematical introduction of spinors is beyond the scope of these notes and we refer the interested reader to e.g. [32, Chapter 13] and [16, Section 8.4]. As the name suggests, spinor fields are related to the quantum property of spin. In particular, there is no field in classical physics that is described as a Dirac field. As a mathematical object, it is nevertheless useful in view of the correct relativistic description of massive quantum particles of half-integer spin (such as electrons, protons or neutrons). These particles can be described through the quantization of classical Dirac fields. In Section 3.4, we discuss an alternative motivation for the introduction of spinors based on the inclusion of parity as a fundamental symmetry.

Consider the Dirac γ -matrices $(\gamma^\mu)_{\mu=0}^3$, defined by

$$\gamma^0 = \begin{pmatrix} 0 & \mathbf{1}_{\mathbb{C}^2} \\ \mathbf{1}_{\mathbb{C}^2} & 0 \end{pmatrix} \in \mathbb{C}^4, \quad \gamma^\mu = \begin{pmatrix} 0 & \sigma_\mu \\ -\sigma_\mu & 0 \end{pmatrix} \in \mathbb{C}^4, \quad \forall \mu \in \{1, 2, 3\}, \quad (1.60)$$

where we recall the definition of the Pauli spin matrices $\sigma_\mu \in \mathbb{C}^{2 \times 2}$ from (2.14). The Dirac matrices form what's called a Clifford algebra.

Problem 1.10. Denote by $[A, B]_+ = AB + BA$ the anticommutator. Prove that

$$[\gamma^\mu, \gamma^\nu]_+ = 2\eta_{\mu\nu}, \forall \mu, \nu \in \{0, 1, 2, 3\}.$$

Conclude that for every $x, y \in \mathbb{R}^4$, we have that

$$[\gamma^\mu x_\mu, \gamma^\nu y_\nu]_+ = 2\eta(x, y).$$

Problem 1.10 implies in particular the isometric property $(\gamma^\mu x_\mu)^2 = \eta(x, x)\mathbf{1}_{\mathbb{C}^4}$ for every $x \in \mathbb{R}^4$ which leads to the observation that

$$\square \mathbf{1}_{\mathbb{C}^4} = (\gamma^\mu \partial_\mu)^2.$$

In other words, the linear operator $\gamma^\mu \partial_\mu$ can be considered a square-root of $\square$. By definition, classical Dirac fields $\psi : \mathbb{R}^4 \rightarrow \mathbb{C}^4$ correspond to solutions of the Dirac equation

$$(i\gamma^\mu \partial_\mu - m)\psi = 0. \quad (1.61)$$

In particular, every solution $\psi \in C^\infty(\mathbb{R}^4, \mathbb{C}^4)$ also solves componentwise

$$0 = (i\gamma^\mu \partial_\mu - m)^2 \psi = ((-\square + m^2)\mathbf{1}_{\mathbb{C}^4} - 2mi\gamma^\mu \partial_\mu)\psi = -(\square + m^2)\psi$$

the Klein-Gordon equation.

As mentioned at the beginning of this section, the previous examples describe classical free fields. As one can check explicitly (*exercise*) it turns out that the defining equations for the above examples can be obtained as critical point equations of suitable non-linear functionals (*exercise*). For instance, the functionals $\mathcal{S}_{\text{SC}}$, $\mathcal{S}_{\text{EM}}$ and $\mathcal{S}_{\text{D}}$ for the real scalar, electromagnetic and, respectively, Dirac fields take the form

$$\begin{aligned} \mathcal{S}_{\text{SC}}(\varphi) &= \int_{\mathbb{R}^4} (\partial^\nu \varphi \partial_\nu \varphi + m^2 \varphi^2) dx, \\ \mathcal{S}_{\text{EM}}(A, j) &= \int_{\mathbb{R}^4} (dA \wedge \star dA - 2A \wedge j), \\ \mathcal{S}_{\text{D}}(\psi) &= \int_{\mathbb{R}^4} \bar{\psi} \gamma^0 (i\gamma^\mu \partial_\mu - m)\psi dx. \end{aligned}$$

Classical interacting theories can be obtained by combining fields of different type. The fundamental principle on which possible interactions can be obtained is called the local gauge principle, see e.g. [5, Section 7.4] for a brief introduction. This is discussed in detail in Section 8.1. Of independent mathematical interest are also presumably simpler interacting theories such as the ϕ_{d}^4 -theories whose action functional takes the form

$$\mathcal{S}_{\phi_{\text{d}}^4}(\varphi) = \int_{\mathbb{R}^d} (\partial^\nu \varphi \partial_\nu \varphi + m^2 \varphi^2 + \lambda \varphi^4) dx.$$

Critical points of $\mathcal{S}_{\phi_{\text{d}}^4}$ satisfy a cubic, non-linear wave equation (*exercise*). In these models, the non-linearity is usually referred to as the self-interaction term.

1.4 Quantum Fields and the Wightman Axioms

The previous sections summarized basic concepts from classical and non-relativistic quantum physics. Quantum field theory combines the principles of quantum mechanics with special relativity. Constructive quantum field theory in particular aims at the rigorous construction of non-trivial interacting quantum field theories and the detailed analysis of their properties. To date, it remains a major open challenge in mathematical physics to construct a non-trivial interacting quantum field theory in spacetime dimension $d = 4$. In lack of such examples, we continue our discussion in the following sections within an axiomatic framework for general quantum field theories. The explicit construction of non-interacting theories in Section 3 shows that the axioms are certainly consistent. This is further confirmed by the rigorous construction of non-trivial interacting theories in lower spacetime dimensions $d < 4$, outlined in Section 7.

Let us now list the Wightman axioms that describe the central notion of a quantum field theory discussed in these notes. From now on, we typically ignore universal physical constants and set in particular $c = \hbar = 1$.

0. Relativistic Quantum Theory. The possible states of the theory are described by the unit rays in a complex, separable Hilbert space $\mathcal{H}$. The state transformation law with regards to coordinate changes from one inertial frame to another is given by a strongly continuous projective unitary representation

$$\mathcal{P}_+^\uparrow \ni (a, L) \mapsto U(a, L) \in \mathcal{U}(\mathcal{H}) = \{U \in \mathcal{L}(\mathcal{H}) : U^*U = UU^* = \mathbf{1}_{\mathcal{H}}\}$$

of the proper Poincaré group $\mathcal{P}_+^\uparrow$. It turns out that every such representation can be lifted to a strongly continuous unitary representation of the inhomogeneous $\mathbb{R}^4 \rtimes \mathrm{SL}(2, \mathbb{C})$ (consisting of pairs $(a, A) \in \mathbb{R}^4 \times \mathrm{SL}(2, \mathbb{C})$, see Section 2.3 for the details), where

$$\mathrm{SL}(2, \mathbb{C}) = \{L \in \mathbb{C}^{2 \times 2} : \det L = 1\} \quad (1.62)$$

denotes the special linear group. This is a consequence of [1] of which Corollary 1.1 was a simple special case. In the sequel, unless mentioned otherwise, we therefore consider without loss of generality representations of $\mathbb{R}^4 \rtimes \mathrm{SL}(2, \mathbb{C})$ and we denote the covering map by $\mathrm{SL}(2, \mathbb{C}) \ni A \mapsto L(A) \in \mathcal{L}_+^\uparrow$ (the covering map is discussed in Section 2.3).

The relativistic four momentum operator $(P_\mu)_{\mu=0}^3$ is identified through an application of Stone's Theorem 1.2 with the generators of spacetime translations $a \in \mathbb{R}^4$, that is

$$U(a, \mathbf{1}_{\mathbb{R}^4}) = e^{iP^\mu a_\mu}.$$

To be more precise, one applies the following generalization of Theorem 1.2, whose proof is left as a problem (alternatively, one may consult [21, Chapter VIII]), to the strongly continuous family of unitary operators $(U(a, \mathbf{1}_{\mathbb{R}^4}))_{a \in \mathbb{R}^4}$.

Theorem 1.5. *Let $\mathbb{R}^n \ni x \mapsto U(x)$ be a strongly continuous map of $\mathbb{R}^n$ into the set of unitary operators on some separable Hilbert space $\mathcal{H}$ and such that*

$$U(x + y) = U(x)U(y), \quad \forall x, y \in \mathbb{R}^n.$$

Set $D = \text{span}(\int_{\mathbb{R}^n} dx f(x) U(x) \phi : f \in C_c^\infty(\mathbb{R}^n), \phi \in \mathcal{H})$. Then D is a domain of essential self-adjointness for each of the generators A_j corresponding to the strongly continuous unitary groups

$$\mathbb{R} \ni x_j \mapsto U(0, \dots, 0, x_j, 0, \dots, 0) = e^{iA_j x_j},$$

their projection valued measures $(\chi_\Omega(A_j))_{\Omega \in \mathcal{B}(\mathbb{R}^n)}$ commute, each $A_j : D \rightarrow D$ and as a consequence $[A_j, A_k] = 0$ in D , for all $j, k \in \{1, \dots, n\}$. Furthermore, there exists a projection-valued measure $(\chi_\Omega)_{\Omega \in \mathcal{B}(\mathbb{R}^n)}$ such that

$$U(x) = \int_{\mathbb{R}^n} e^{ixy} \chi_{dy} = \prod_{j=1}^n \int_{\mathbb{R}} e^{ix_j y_j} \chi_{dy}(A_j) = \prod_{j=1}^n e^{ix_j A_j}.$$

As a consequence, $e^{iAx} = \prod_{j=1}^n e^{ix_j A_j}$, where Ax denotes the generator of the strongly continuous, one-parameter group $\mathbb{R} \ni t \mapsto U(tx)$, for fixed $x \in \mathbb{R}^n$.

Problem 1.11. Prove Theorem 1.5 by generalizing the arguments from Theorem 1.2.

In other words, $P = (P_0, P_1, P_2, P_3) : \mathcal{D}_P \rightarrow \mathcal{H}^4$ is equal to a densely defined, self-adjoint operator. We typically suppress the P -dependence of its spectral measure $(\chi_\Omega(P))_{\Omega \in \mathcal{B}(\mathbb{R}^4)} = (\chi_\Omega)_{\Omega \in \mathcal{B}(\mathbb{R}^4)}$. The component $P_0 = H$ is interpreted as the energy (the Hamiltonian) of the system and the operator $P^\mu P_\mu$ equals the square of its mass. We remark that in the non-interacting examples discussed below, the latter is by construction always a positive constant multiple $P^\mu P_\mu = m^2 > 0$ of the identity in $\mathcal{H}$. Motivated by the energy relation (1.52) for a single particle of mass $m > 0$ (see also the related discussions in [30, Section 1.4]), one assumes the spectral condition that $(\chi_\Omega)_{\Omega \in \mathcal{B}(\mathbb{R}^4)}$ is supported in the closure

$$\overline{V}^+ = \{p \in \mathbb{R}^4 : p^\mu p_\mu = p_0^2 - |\mathbf{p}|^2 \geq 0, p_0 \geq 0\}$$

of the forward light cone V^+ , defined by

$$V^+ = \{p \in \mathbb{R}^4 : p^\mu p_\mu > 0, p_0 > 0\}. \quad (1.63)$$

In other words, $\chi_\Omega = 0$ for every $\Omega \subset (\overline{V}^+)^C$ so that for every bounded, measurable $g : \mathbb{R}^4 \rightarrow \mathbb{C}$, we have that

$$g(P) = \int_{\mathbb{R}^4} g(p) \chi_{dp} = \int_{\overline{V}^+} g(p) \chi_{dp}.$$

In particular, $P_0 \geq 0$ and $P^\mu P_\mu \geq 0$ (positivity of the energy and mass).

Finally, one assumes the existence of a unique (up to multiplication by a phase) state $\psi_0 \in \mathcal{H}$, $\|\psi_0\| = 1$, which represents the vacuum and which satisfies

$$U(a, A)\psi_0 = \psi_0, \quad \forall (a, A) \in \mathbb{R}^4 \rtimes \text{SL}(2, \mathbb{C}). \quad (1.64)$$

The vacuum state has the lowest possible energy $P_0\psi_0 = 0$ and, more generally, zero four momentum $P\psi_0 = 0$ and is usually interpreted as describing the state with no physical particles. As such, it takes the same form in all inertial systems. If $0 \in \sigma(P)$ turns out to be a discrete eigenvalue (that is, an isolated eigenvalue of finite multiplicity) of P , we say that the theory has a mass gap of size $\Delta = \inf \sigma(P^\mu P_\mu) \cap (0, \infty) > 0$. The terminology is motivated by (1.52). If one interprets the quantum field to consist of elementary excitations that carry its force (like e.g. photons as the excitations of the electromagnetic field in quantum electrodynamics), a mass gap means that these excitations have a mass at least as large as Δ .

It is worth to remark that $0 \in \sigma(P)$ is the only value that can be an eigenvalue of P . For suppose that $0 \neq p \in \mathbb{R}^4$ is another eigenvalue with normalized eigenvector $\psi_p \in \mathcal{H}$, i.e. $P\psi_p = p\psi_p$. Then, for every $A \in \text{SL}(2, \mathbb{C})$, we have that

$$\begin{aligned} P_\mu U(0, A)\psi_p &= \lim_{\epsilon \rightarrow 0} \frac{1}{i\epsilon} (U(\epsilon_\mu, \mathbf{1}_{\mathbb{R}^4}) - \mathbf{1}_{\mathcal{H}}) U(0, A)\psi_p \\ &= U(0, A) \lim_{\epsilon \rightarrow 0} \frac{1}{i\epsilon} (U(L(A)^{-1}\epsilon_\mu, \mathbf{1}_{\mathbb{R}^4}) - \mathbf{1}_{\mathcal{H}}) \psi_p \\ &= \eta(p, L(A)^{-1}e_\mu) U(0, A)\psi_p \\ &= \eta(L(A)p, e_\mu) U(0, A)\psi_p = (L(A)p)_\mu U(0, A)\psi_p, \end{aligned}$$

where $\epsilon_\mu \in \mathbb{R}^4$ is such that $\eta(x, \epsilon_\mu) = \epsilon x_\mu, \forall x \in \mathbb{R}^4, \mu \in \{0, 1, 2, 3\}$. Thus, $U(0, A)\psi_p$ is a normalized eigenvector of P with eigenvalue $0 \neq L(A)p \in \mathbb{R}^4$. In particular, if $L(A_1)p \neq L(A_2)p$, $U(0, A_1)\psi_p$ and $U(0, A_2)\psi_p$ are orthonormal and we get an uncountable family of orthonormal vectors (*exercise*) which can not exist in the separable space $\mathcal{H}$.

As explained e.g. in [22, Section IX.8], it also turns out that (1.64) is equivalent to

$$U(a, \mathbf{1}_{\mathbb{R}^4})\psi_0 = \psi_0, \forall a \in \mathbb{R}^4. \quad (1.65)$$

Using as above the identity

$$U(a, A) = U(a, \mathbf{1}_{\mathbb{R}^4})U(0, A) = U(0, A)U(L(A)^{-1}a, \mathbf{1}_{\mathbb{R}^4}),$$

for all $(a, A) \in \mathbb{R}^4 \times \text{SL}(2, \mathbb{C})$, we obtain indeed that

$$\begin{aligned} f(P)U(0, A) &= \int_{\mathbb{R}^4} \chi_{dx} f(x)U(0, A) = \int_{\mathbb{R}^4} \chi_{dx} \left(\int_{\mathbb{R}^4} dp \hat{f}(p) e^{2\pi i p x} \right) U(0, A) \\ &= \int_{\mathbb{R}^4} dp \hat{f}(p) U(2\pi(p_0, -\mathbf{p}), \mathbf{1}_{\mathbb{R}^4}) U(0, A) \\ &= U(0, A) f(L(A)P) \end{aligned}$$

for every $f \in \mathcal{S}(\mathbb{R}^4)$, $A \in \text{SL}(2, \mathbb{C})$. Using a dominated convergence argument to approximate the characteristic function of the point $0 \in \mathbb{R}^4$, this implies (*exercise*)

$$[U(0, A), \chi_{\{0\}}] = 0.$$

By the uniqueness assumption on ψ_0 , (1.65), we know that the range $\text{ran}(\chi_{\{0\}}) = \text{span}(\psi_0)$ is one dimensional, for $\phi \in \text{ran}(\chi_{\{0\}})$ implies that $P\phi = 0$ and thus (1.65) with ψ_0 replaced by ϕ . Thus, the previous commutator identity implies that

$$U(0, A)|_{\text{ran}(\chi_{\{0\}})} : \text{ran}(\chi_{\{0\}}) \rightarrow \text{ran}(\chi_{\{0\}})$$

so that $\text{SL}(2, \mathbb{C}) \ni A \mapsto U(0, A)|_{\text{ran}(\chi_{\{0\}})}$ is a one-dimensional (and hence commutative, irreducible) unitary representation of $\text{SL}(2, \mathbb{C})$. As a consequence, $U(0, A)|_{\text{ran}(\chi_{\{0\}})}$ is a multiple of the identity operator in $\text{ran}(\chi_{\{0\}})$, by Lemma 1.4. Since $\text{SL}(2, \mathbb{C})$ is connected (see Section 2.3) and $U(0, 0) = \mathbf{1}_{\mathcal{H}}$, we get $U(a, A)\psi_0 = \psi_0$ for all $(a, A) \in \mathbb{R}^4 \rtimes \text{SL}(2, \mathbb{C})$.

I. Regularity of the Field. For $n \in \mathbb{N}$, we consider n -component quantum fields $\Phi = (\Phi_1, \dots, \Phi_n)$ whose transformation law is related to an n -dimensional matrix representation $S : \text{SL}(2, \mathbb{C}) \rightarrow \mathbb{C}^{n \times n}$ of $\text{SL}(2, \mathbb{C})$. These representations are determined and exemplified in Section 3 below. The field components are assumed to form operator-valued distributions which, by definition, means the following:

For every $f \in \mathcal{S}(\mathbb{R}^4)$, there exist operators $\Phi(f) = (\Phi_1(f), \dots, \Phi_n(f))$, defined on a linear, dense domain $D \subset \mathcal{H}$ such that $\psi_0 \in D$. The domains of the adjoint components $\Phi(f)^* = (\Phi_1(f)^*, \dots, \Phi_n(f)^*)$ also contain D and we assume, moreover, that

$$U(a, A)D \subset D, \quad \Phi_k(f)D \subset D, \quad \Phi_k(f)^*D \subset D, \quad \forall (a, A) \in \mathbb{R}^4 \rtimes \text{SL}(2, \mathbb{C}), k \in \{1, \dots, n\}.$$

Finally, we assume that for every $\varphi, \psi \in D$ and $j \in \{1, \dots, n\}$, the linear functional

$$\mathcal{S}(\mathbb{R}^4) \ni f \mapsto \langle \varphi, \Phi_k(f)\psi \rangle \in \mathbb{C}$$

defines a tempered distribution (see Section 2.2.2 for the definition and basic properties).

Let us add two comments on the regularity assumptions. First, in view of basic quantum theory, a quantum field $(\Phi_x)_{x \in \mathbb{R}^4}$ should correspond most naturally to a family of self-adjoint operators labeled by the points in Minkowski space (to every spacetime point we attach an observable). We allow for more general fields which are not necessarily self-adjoint. Such fields are useful and occur quite naturally in quantum field theory (for instance, in the form of the so called creation and annihilation operators, see Section 2.4). Second, as becomes already clear from the construction of the simplest case of a massive non-interacting scalar field, it is usually not possible to define a quantum field as an operator-valued function (that is, a self-adjoint operator Φ_x for each spacetime point $x \in \mathbb{R}^4$), but rather as an operator-valued distribution. In this sense, one interprets the operators $(\Phi(f))_{f \in \mathcal{S}(\mathbb{R}^4)}$ via

$$\Phi(f) = \int_{\mathbb{R}^4} dx f(x) \Phi_x \tag{1.66}$$

as a smeared version of the field $(\Phi_x)_{x \in \mathbb{R}^4}$, which is only to be understood symbolically. Mathematically, there is of course no necessity to restrict to tempered distributions and one might consider more general fields in this regard as well. The temperedness assumption is nevertheless useful and allows to work with the Fourier transform.

II. Transformation Behavior of the Field. With the same notation as in **I**, the transformation law of the field Φ takes the form

$$U(a, A)\Phi_k(f)U(a, A)^* = S(A)_k^l \Phi_l((a, A)f), \quad \forall (a, A) \in \mathbb{R}^4 \rtimes \text{SL}(2, \mathbb{C}), f \in \mathcal{S}(\mathbb{R}^4), \quad (1.67)$$

where we set

$$((a, A)f)(\cdot) = f(L(A)^{-1}(\cdot - a)), \quad \forall (a, A) \in \mathbb{R}^4 \rtimes \text{SL}(2, \mathbb{C}). \quad (1.68)$$

The identity (1.67) is assumed to hold in $D \subset \mathcal{H}$. The left hand side in (1.67) is interpreted as representing the field components with regards to a transformed inertial frame. With the identification (1.66), note that heuristically (1.67) is equivalent to

$$\begin{aligned} U(a, A)(\Phi_k)_x U(a, A)^* &= U(a, A) \left(\int_{\mathbb{R}^4} dy \delta_x(y) (\Phi_k)_y \right) U(a, A)^* \\ &= U(a, A) \Phi_k(\delta_x) U(a, A)^* \\ &= S(A)_k^l \int_{\mathbb{R}^4} dy \delta_x(L(A)^{-1}(y - a)) (\Phi_l)_y \\ &= S(A)_k^l (\Phi_l)_{a+L(A)x} \end{aligned} \quad (1.69)$$

for every $(a, A) \in \mathbb{R}^4 \rtimes \text{SL}(2, \mathbb{C})$, $x \in \mathbb{R}^4$. This is motivated by the classical field transformation law (1.46) (with $(a, R) \in \text{E}(3)$ replaced by $(a, A) \in \mathbb{R}^4 \rtimes \text{SL}(2, \mathbb{C})$).

III. Local Commutativity (Microscopic Causality). For a fixed choice of $-$ or $+$, the following holds true: let $f, g \in \mathcal{S}(\mathbb{R}^4)$ satisfy $f(x)g(y) = 0$ if $\eta(x - y, x - y) \geq 0$. In this case, we call f, g spacelike separated. Then, for every such pair of functions

$$[\Phi_k(f), \Phi_l(g)]_{\pm} = [\Phi_k(f), \Phi_l^*(g)]_{\pm} = 0 \quad \text{in } D, \quad \forall k, l \in \{1, \dots, n\}, \quad (1.70)$$

where $[A, A']_{\pm} = AA' \pm A'A$. The identities (1.70) are motivated by the causality relations of events in $(\mathbb{R}^4, \eta)$. As was noted in Section 1.3, spacelike separated events can not influence each other. It is therefore natural to assume that the corresponding fields can be jointly diagonalized (in case of self-adjoint quantum fields), which is defined mathematically by requiring that their projection valued measures commute. The possibility for anticommutators in (1.70) is related to the quantum field theoretic discovery that half-integer spin fields can only be appropriately quantized if one imposes anticommutation relations. There is no classical or non-relativistic motivation of this assumption.

IV. Quantum Field Theory. A relativistic quantum theory satisfying **0** with a quantum field satisfying axioms **I, II, III**, is called a quantum field theory if the vacuum $\psi_0 \in \mathcal{H}$ is cyclic for the smeared fields. By definition, this means that the linear space

$$D_0 = \text{span}\{\Phi_{i_1}(f_1) \dots \Phi_{i_k}(f_k)\psi_0 \in D : f_j \in \mathcal{S}(\mathbb{R}^4) \forall j \in \mathbb{N}, \mathbf{i} \in \{1, \dots, n\}^k, k \in \mathbb{N}\}$$

is dense, that is, $\overline{D_0} = \mathcal{H}$. Loosely speaking, a non-trivial quantum field should have an impact on a sufficiently large class of states in order to be physically relevant.

Axioms **O** to **IV** summarize precisely what is commonly understood as a minimal set of requirements of a quantum field theory. In view of particle physics, one should also include axioms that relate the theory to scattering experiments. This topic is not discussed in these notes. For more details on this, including references on possible extensions related to scattering, see the remarks in [30, Section 3.1].

2 Mathematical Interlude

In this chapter, we discuss basic aspects of the theory of distributions which are properly introduced in the context of locally convex, topological vector spaces, of the Lorentz and Poincaré groups as well as of basic Fock space operators which are used subsequently to construct explicit examples of free quantum field theories. The relevance of these topics in view of the description of quantum fields should be clear from the Wightman axioms described in Section 1.4. In addition to [30, 31], we follow [26, Chapters 1, 6, 7], [21, Chapter V] and [22, Chapter IX].

2.1 Locally Convex Topological Vector Spaces

From a functional analytic point of view, distributions are most naturally defined in the context of locally convex, topological vector spaces. In this section, we first discuss some general definitions and results in this context. In the next Section 2.2, we get more concrete by studying distributions and tempered distributions on $\mathbb{R}^n$.

A topological vector space is a vector space X equipped with a topology such that the vector space operations $+: X \times X \rightarrow X$ and $\cdot: \mathbb{K} \times X \rightarrow X$ are continuous and such that the point sets $\{x\}$ are closed, for every $x \in X$. Throughout this section, we always consider real ($\mathbb{K} = \mathbb{R}$) or complex ($\mathbb{K} = \mathbb{C}$) vector spaces.

Problem 2.1. *Let X be a topological vector space and define for $x \in X, 0 \neq \alpha \in \mathbb{K}$ the translation $\tau_x: X \rightarrow X$ and, respectively, multiplication operators $\mu_\alpha: X \rightarrow X$ by*

$$\tau_x(y) = x + y, \quad \mu_\alpha(y) = \alpha \cdot y, \quad \forall y \in X.$$

Show that both τ_x and μ_α map X homeomorphically onto itself.

We call a family $\mathcal{B}$ of subsets a local base at $0 \in X$ if every $B \in \mathcal{B}$ is open, contains $0 \in B$ and if every other open neighborhood of $0 \in X$ contains an element $B \in \mathcal{B}$. Given a local base $\mathcal{B}$ at $0 \in X$, Problem 2.1 implies that every open set in X is equal to a union of translates of elements of $\mathcal{B}$. In the following, we therefore refer to a local base at $0 \in X$ simply as a local base. Any such base determines the topology of X .

We say that X is locally convex if there exists a local base $\mathcal{B}$ whose elements are convex. We say that (X, τ) is metrizable if there exists a metric $d: X \times X \rightarrow [0, \infty)$ on X whose induced topology is equal to τ . We call X a Fréchet space if it is a locally convex, metrizable, topological vector space whose metric $d: X \times X \rightarrow [0, \infty)$ is translation invariant (that is, $d(x, y) = d(\tau_z x, \tau_z y) = d(x + z, y + z)$ for every $x, y, z \in X$) and such that (X, d) is complete (every Cauchy sequence converges to a limit in X). We say that X has the Heine-Borel property if every closed and bounded subset of X is compact.

Proposition 2.1. *Suppose X is a topological vector space, $K \subset X$ is compact and $C \subset X$ is closed such that $K \cap C = \emptyset$. Then $0 \in X$ has a neighborhood $U \subset X$ such that*

$$(K + U) \cap (C + U) = \emptyset.$$

As a consequence, if $\mathcal{B}$ is a local base for X , then every element in $\mathcal{B}$ contains the closure of some element of $\mathcal{B}$. Furthermore, X is a Hausdorff space.

Proof. By assumption, single point sets are closed and they are certainly compact. Applying the claim on K and C to two single point sets in place of K and C thus shows that X is a Hausdorff space. To prove the general claim on K and C , on the other hand, we may assume w.l.o.g. that $K \neq \emptyset$ (so that $K + U \neq \emptyset$) and proceed as follows.

First of all, suppose $W \subset X$ is an open neighborhood of $0 \in X$. Then W contains another neighborhood U such that $U = -U$ and such that $U + U \subset W$. Indeed, since $0 + 0 = 0 \in W$ and $+: X \times X \rightarrow X$ is continuous, by assumption, we find open sets V_1, V_2 such that $0 \in V_1, V_2$ and such that $V_1 + V_2 \subset W$. Choosing $U = V_1 \cap V_2 \cap (-V_1) \cap (-V_2)$, we see that $U = -U$ and that $U + U \subset V_1 + V_2 \subset W$, as desired.

Now, let $x \in K$. Then $x \in C^c$ and $C^c - x$ is an open neighborhood of $0 \in X$. By the preceding observation, we find $\tilde{U}_x \subset X$ open, $0 \in \tilde{U}_x = -\tilde{U}_x$ such that $\tilde{U}_x + \tilde{U}_x \subset C^c - x$. Applying the preceding observation once more, now for $\tilde{U}_x$, we find an open neighborhood $U_x \subset X$, $0 \in U_x = -U_x$ such that $U_x + U_x + U_x + U_x \subset \tilde{U}_x + \tilde{U}_x \subset C^c - x$. In particular, $x + U_x + U_x + U_x \subset x + (U_x + U_x + U_x + U_x) \subset C^c$ has empty intersection with the closed set C . Moreover, since $U_x = -U_x$, we also have that

$$(x + U_x + U_x) \cap (C + U_x) = \emptyset \quad (\iff \quad (x + U_x + U_x - U_x) \cap C = \emptyset).$$

Now, K is compact so that we find finitely many points $x_1, \dots, x_n \in K$ such that

$$K \subset \bigcup_{j=1}^n (x_j + U_{x_j}).$$

Then setting $U = \bigcap_{j=1}^n U_{x_j}$, we conclude that

$$K + U \subset \bigcup_{j=1}^n (x_j + U_{x_j} + U) \subset \bigcup_{j=1}^n (x_j + U_{x_j} + U_{x_j})$$

has empty intersection with $C + U$. Finally, for the statement about $\mathcal{B}$, notice that

$$C + U = \bigcup_{x \in C} (x + U)$$

is open, because the topology of X is translation invariant. The previous observation thus shows that $K + U \subset (C + U)^c$ so that also its closure $\overline{K + U} \cap (C + U) = \emptyset$. Thus, if $\mathcal{B}$ is a local base for X and $\tilde{B} \in \mathcal{B}$ is an open neighborhood of $0 \in X$, choose $K = \{0\}$

and $C = \tilde{B}^c$ to find an open neighborhood $U \subset X$ of $0 \in X$ and thus, by definition of a local base, an element $B \subset U$ with $0 \in B$ such that

$$\overline{B} = \overline{K+B} \subset \overline{K+U} \subset (\tilde{B}^c + U)^c = \left(\bigcup_{x \in U} (x + \tilde{B}^c) \right)^c = \bigcap_{x \in U} (x + \tilde{B}) \subset \tilde{B}.$$

□

We call a subset $S \subset X$ balanced if $\alpha S \subset S$ for every $\alpha \in \mathbb{K}$ with $|\alpha| \leq 1$.

Proposition 2.2. *Let X be a topological vector space. Then, every neighborhood of $0 \in X$ contains a balanced neighborhood of $0 \in X$ and every convex neighborhood $0 \in X$ contains a balanced, convex neighborhood of $0 \in X$. In particular, every locally convex space X admits a balanced, convex local base.*

Proof. Let $B \subset X$ be an open neighborhood of $0 \in X$. First we show that B contains a balanced neighborhood W . To see this, recall that multiplication by $\alpha \in \mathbb{K}$ is a homeomorphism and in particular continuous. This means that $\cdot (U_\delta \times V) \subset B$ for some $U_\delta \subset \mathbb{K}$ and $V \subset X$ open with $0 \in V$ or, in other words, that $\alpha V \subset B$ for all $|\alpha| < \delta$ for some $\delta > 0$. Thus

$$W = \bigcup_{0 \leq |\alpha| < \delta: \alpha V \subset B} \alpha V \subset B$$

is equal to a balanced, open neighborhood contained in B .

In the next step, we assume additionally that B is convex and we construct a balanced convex neighborhood $S \subset B$. To this end, consider $A = \bigcap_{|\alpha|=1} \alpha B$ and choose $W \subset B$ as above. Then, for $|\alpha| = 1$, we have that $\alpha^{-1}W \subset W$ so that $W \subset \bigcap_{|\alpha|=1} \alpha B = A \subset B$. In particular, the interior A° is non-trivial is a neighborhood of $0 \in X$, because $0 \in W \subset A^\circ$. Since A is an intersection of convex sets, it is convex and this implies that A° is convex, too (*exercise*). A° (containing $0 \in X$) is also balanced if A is (*exercise*). To show that A is balanced, on the other hand, let $|\beta| \leq 1$ and notice that $|\beta|\alpha B \subset \alpha B$ for every $|\alpha| = 1$, because αB is convex and contains $0 \in X$. Thus, we have for all $0 \neq |\beta| \leq 1$

$$\beta A = |\beta| \frac{\beta}{|\beta|} A = |\beta| \bigcap_{|\alpha|=1} \alpha B \subset \bigcap_{|\alpha|=1} \alpha B = A.$$

This implies that $A^\circ \subset B$ is an open, convex, balanced neighborhood of $0 \in X$. □

In a topological vector space X , we call a set $S \subset X$ bounded if for every neighborhood $U \subset X$ of $0 \in X$ we have that $S \subset tU$ for every $t > s$, for some $s > 0$.

Proposition 2.3. *Let B be an open neighborhood of $0 \in X$ in the topological vector space X . Assume that $(a_n)_{n \in \mathbb{N}}$ is increasing with $\lim_{n \rightarrow \infty} a_n = \infty$. Then*

$$X = \bigcup_{n \in \mathbb{N}} a_n B.$$

If $(b_n)_{n \in \mathbb{N}}$ is decreasing with $\lim_{n \rightarrow \infty} b_n = 0$ and B is in addition bounded, then

$$\mathcal{B} = \{b_n B : n \in \mathbb{N}\}$$

forms a local base for X .

Proof. For the first part, fix $x \in X$ and notice that $\{\alpha \in \mathbb{K} : \alpha x \in B\}$ is open, by continuity of scalar multiplication, and contains $0 \in \mathbb{K}$, by assumption on B . Since $\lim_{n \rightarrow \infty} a_n^{-1} = 0$, we must have $a_n^{-1}x \in B$, that is $x \in a_n B$, for $n \in \mathbb{N}$ large enough.

For the second part, let U be an open neighborhood of $0 \in X$. Since B is assumed to be bounded, we have that $B \subset tU$ for all $t > 0$ large enough. This shows that for all $n \in \mathbb{N}$ large enough, we have $B \subset b_n^{-1}U$, that is $b_n B \subset U$. $\square$

Problem 2.2. Let X be a topological vector space. Show that the closure of every bounded set, every compact set and every convergent sequence are bounded.

Next, let us recall some basic definitions and results on linear maps. We say that a linear map $\Lambda : X \rightarrow Y$ between topological vector spaces is bounded if $\Lambda(E) \subset Y$ is a bounded set whenever $E \subset X$ is a bounded set. When $\Lambda : X \rightarrow \mathbb{K}$ is a linear function that maps to $\mathbb{K}$, we usually call it a functional. The dual space X^* of a topological vector space X consists of the continuous linear functionals on X , which is itself a vector space with the standard addition and scalar multiplication of functions.

Proposition 2.4. Let $\Lambda : X \rightarrow Y$ be a linear map between topological vector spaces X, Y and consider the following four properties:

- i) Λ is continuous.
- ii) Λ is bounded.
- iii) If $\lim_{n \rightarrow \infty} x_n = 0$, then $\{\Lambda x_n : n \in \mathbb{N}\}$ is bounded.
- iv) If $\lim_{n \rightarrow \infty} x_n = 0$, then $\lim_{n \rightarrow \infty} \Lambda x_n = 0$.

Then, i) implies ii) which implies iii). If we assume in addition that X is metrizable with a translation invariant metric, we also have that iii) implies iv) which implies i).

Proof. Assume i), let $E \subset X$ be bounded and let $W \subset Y$ be a bounded neighborhood of $0 \in Y$. Since $\Lambda 0 = 0$ and Λ is continuous, the preimage of W under Λ is an open neighborhood V of $0 \in X$, in particular $\Lambda(V) \subset W$. Since E is bounded, $E \subset tV$ for $t > 0$ sufficiently large and thus $\Lambda(t^{-1}E) = t^{-1}\Lambda(E) \subset \Lambda(V) \subset W$ for $t > 0$ large enough. Thus, $\Lambda(E)$ is bounded. This shows that i) implies ii). The latter implies iii), on the other hand, because convergent sequences are bounded by Problem 2.2.

Now, assume iii), suppose $\lim_{n \rightarrow \infty} x_n = 0$ and assume additionally that X admits a translation invariant metric $d : X \times X \rightarrow [0, \infty)$ that is compatible with its topology. We first claim that then there exists a sequence $(a_n)_{n \in \mathbb{N}}$ in $\mathbb{R} \subset \mathbb{K}$ such that $\lim_{n \rightarrow \infty} a_n = \infty$

and such that still $\lim_{n \rightarrow \infty} a_n x_n = 0$. To see this, we use that $\lim_{n \rightarrow \infty} x_n = 0$ is equivalent to $\lim_{n \rightarrow \infty} d(x_n, 0) = 0$. Thus, for every $k \in \mathbb{N}$ we find some $n_k \in \mathbb{N}$ such that

$$d(x_n, 0) \leq \frac{1}{k^2}, \forall n \geq n_k.$$

We may then set $a_n = 1$ if $n \leq n_k$ and $a_n = k$ if $n_k \leq n < n_{k+1}$. This implies for $n_k \leq n < n_{k+1}$ that

$$d(a_n x_n, 0) = d(k x_n, 0) \leq \sum_{j=1}^k d(j x_n, (j-1) x_n) \leq k d(x_n, 0) \leq \frac{1}{k} \rightarrow 0$$

as $k \rightarrow \infty$ and thus $\lim_{n \rightarrow \infty} a_n x_n = 0$. In particular, by *iii*), this sequence is bounded so that also $(\Lambda(a_n x_n))_{n \in \mathbb{N}}$ is bounded.

Next, we claim that if $(y_n)_{n \in \mathbb{N}}$ is bounded in Y , we have that $\lim_{n \rightarrow \infty} b_n y_n = 0$ whenever $\lim_{n \rightarrow \infty} b_n = 0$ in $\mathbb{K}$. Indeed, let $V \subset Y$ be an open neighborhood of $0 \in Y$. Without loss of generality, we may assume in addition that V is balanced, by Prop. 2.2. Since $(y_n)_{n \in \mathbb{N}}$ is bounded, we have that $y_n \in tV$ for some $t > 0$. Choosing $n \in \mathbb{N}$ large enough such that $|b_n t| \leq 1$, we conclude that $b_n y_n \in b_n t V \subset V$. Thus, $\lim_{n \rightarrow \infty} b_n y_n = 0$.

Now, apply the previous statement to $b_n = a_n^{-1}$ and $y_n = \Lambda(a_n x_n)$. Then

$$\lim_{n \rightarrow \infty} \Lambda x_n = \lim_{n \rightarrow \infty} a_n^{-1} \Lambda(a_n x_n) = 0,$$

which proves *iv*).

Finally, assume *iv*). By linearity, *iv*) implies that Λ is sequentially continuous. Let $U \subset Y$ be open and pick $x \in \Lambda^{-1}(U)$. If there exists no metric ball $B_\delta(x) \subset \Lambda^{-1}(U)$, we find a sequence $(x_n)_{n \in \mathbb{N}}$ such that $\lim_{n \rightarrow \infty} x_n = x$, but such that $\Lambda x_n \notin U$ for all $n \in \mathbb{N}$, that is, Λx is not the limit of $(\Lambda x_n)_{n \in \mathbb{N}}$, in contradiction to our assumption. Thus, for every $U \subset Y$ open and every $x \in \Lambda^{-1}(U)$, there exists some $B_\delta(x) \subset \Lambda^{-1}(U)$, showing that $\Lambda^{-1}(U) \subset X$ is open. This proves the continuity of $\Lambda : X \rightarrow Y$. $\square$

In the context of distributions, we always deal with locally convex, topological vector spaces. As we explain next, such spaces have topologies that are generated by continuous families of seminorms. Recall that a seminorm on a vector space X is a real-valued function $p : X \rightarrow \mathbb{R}$ such that

- a) $p(x + y) \leq p(x) + p(y)$ for every $x, y \in X$ and
- b) $p(\lambda x) = |\lambda| p(x)$ for every $x \in X, \lambda \in \mathbb{K}$.

We say that a family of seminorms $\mathcal{P}$ on X is separating if for every $x \neq 0$ there exists some $p \in \mathcal{P}$ such that $p(x) \neq 0$. We say that a convex subset $A \subset X$ is absorbing if for every $x \in X$ there exists $t > 0$ such that $x \in tA$. For example, by Prop. 2.3, every convex, open neighborhood $B \subset X$ of $0 \in X$ is absorbing. Notice, that $0 \in A$ for every absorbing A (because $0 = tx$ for some $t > 0$ and $x \in X$ implies $x = 0 \in A$). To each convex, absorbing A , we associate its Minkowski functional $\mu_A : X \rightarrow [0, \infty)$, defined by

$$\mu_A(x) = \inf \{t > 0 : x \in tA\}. \quad (2.1)$$

It turns out that every seminorm on X corresponds to a Minkowski functional of a balanced ($\alpha B \subset B$ for $|\alpha| \leq 1$), convex, absorbing set $B \subset X$. This is proved in the next lemma which collects basic properties of seminorms.

Lemma 2.1. *Let $p : X \rightarrow [0, \infty)$ be a seminorm on the vector space X . Then, the following holds true:*

- i) $p(0) = 0$, $|p(x) - p(y)| \leq p(x - y)$ and $p(x) \geq 0$ for all $x, y \in X$.*
- ii) The set $\{x \in X : p(x) = 0\}$ is a linear subspace of X .*
- iii) $B = \{x \in X : p(x) < 1\}$ is convex, balanced, absorbing and such that $p = \mu_B$.*

Proof. We start with *i*). $p(0) = 0p(0) = 0$ by scalar homogeneity of seminorms and $|p(x) - p(y)| \leq p(x - y)$ follows from the triangle inequality and $p(x - y) = p(y - x)$, since $p(-x) = p(x)$. Choosing $y = 0$, this yields $0 \leq |p(x)| \leq p(x)$ and concludes *i*). Note in particular that p defines a norm if and only if $p(x) \neq 0$ whenever $x \neq 0$.

Part *ii*) follows directly from

$$p(\lambda_1 x_1 + \lambda_2 x_2) \leq |\lambda_1|p(x_1) + |\lambda_2|p(x_2)$$

for every $\lambda_1, \lambda_2 \in \mathbb{K}$ and $x_1, x_2 \in X$.

Finally, consider *iii*). Convexity of B is a direct consequence of the previous bound for $\lambda_1 = t, \lambda_2 = 1 - t$, for $t \in [0, 1]$. Similarly, its balancedness follows from $p(\alpha x) = |\alpha|p(x) < 1$ whenever $|\alpha| \leq 1$ and $x \in B$ and B being absorbing follows, because for every $x \in X$, we have $(p(x) + \delta)^{-1}x \in B$ for every $\delta > 0$, that is $x \in (p(x) + \delta)B$.

It remains to show that $p = \mu_B$. To this end, we use that $x \in tB$ if and only if $t^{-1}x \in B$, that is $p(x) < t$, for $t > 0$. In other words, $\mu_B(x) = \inf_{t > p(x)} \{t\} = p(x)$. $\square$

Part *iii*) of the previous lemma motivates to collect some general properties of Minkowski functionals as well.

Lemma 2.2. *Suppose $A \subset X$ is a convex, absorbing set in the vector space X . Then the following holds true.*

- i) $\mu_A(x + y) \leq \mu_A(x) + \mu_A(y)$ for every $x, y \in X$.*
- ii) $\mu_A(tx) = t\mu_A(x)$ for every $t \geq 0, x \in X$*
- iii) If A is balanced, then μ_A defines a seminorm.*
- iv) If $B = \{x \in X : \mu_A(x) < 1\}$ and $C = \{x \in X : \mu_A(x) \leq 1\}$, then $B \subset A \subset C$ and $\mu_A = \mu_B = \mu_C$.*

Proof. For part *i*), choose $s = \mu_A(x) + \delta, t = \mu_A(y) + \delta$ for $\delta > 0$ so that $x \in sA, y \in tA$. Then, we get by the convexity of A that

$$\frac{x + y}{s + t} = \frac{s}{s + t} \frac{x}{s} + \frac{t}{s + t} \frac{y}{t} \in A$$

and thus $\mu_A(x+y) \leq s+t = \mu_A(x) + \mu_A(y) + 2\delta$. Since $\delta > 0$ was arbitrary, we get *i*).

For *ii*), note that $\mu_A(0) = 0$, so the equality certainly holds true for $t = 0$. On the other hand, for $t > 0$, we have that

$$\mu_A(tx) = \inf\{u > 0 : tx \in uA\} = \inf\{tu > 0 : x \in uA\} = t\mu_A(x).$$

For *iii*), assume that A is balanced, that is, $\alpha A \subset A$ whenever $|\alpha| \leq 1$. Subadditivity of μ_A follows from *i*), so it suffices to prove that $\mu_A(\lambda x) = |\lambda| \mu_A(x)$ for all $0 \neq \lambda \in \mathbb{K}$. To this end, note that A balanced implies tA is balanced as well, for every $t > 0$. This implies $\lambda x \in tA$ if and only if $|\lambda|x \in tA$ so that by part *ii*), we conclude

$$\mu_A(\lambda x) = \inf\{t > 0 : \lambda x \in tA\} = \inf\{t > 0 : |\lambda|x \in tA\} = \mu_A(|\lambda|x) = |\lambda|\mu_A(x).$$

Finally, consider *iv*). Notice that $U \subset V$ implies generally that $\mu_V \leq \mu_U$. Now

$$\begin{aligned} B = \{x \in X : \mu_A(x) < 1\} &\subset \{x \in X : \exists 0 < t < 1 : x \in tA\} \\ &\subset A \subset \{x \in X : x \in tA \text{ for } t = 1\} \subset C \end{aligned}$$

so that $\mu_C \leq \mu_A \leq \mu_B$. Notice here that both B and C are convex and absorbing, by parts *i*) and *ii*), so μ_B and μ_C are well-defined (*exercise*). To conclude *iv*), it is enough to show that $\mu_B \leq \mu_C$. To this end, choose $s, t > 0$ such that

$$\mu_C(x) < s < t \implies \frac{x}{s} \in C \implies \mu_A(x/s) \leq 1 \implies \mu_A(x/t) \leq \frac{s}{t} < 1 \implies \frac{x}{t} \in B.$$

The last conclusion implies that $\mu_B(x) \leq t$ and since $t > \mu_C(x)$ was arbitrary, we must have that $\mu_B \leq \mu_C$ and thus $\mu_A = \mu_B = \mu_C$. This concludes *iv*). $\square$

The following two results show that the topology in a locally convex topological vector space is always generated by a separating family of continuous seminorms. In the following sections, this is used to define the space of distributions and tempered distributions. For the first result, recall that every locally convex, topological vector space X admits a convex, balanced local base, by Proposition 2.2.

Theorem 2.1. *Suppose $\mathcal{B}$ is a convex, balanced local base in the topological vector space X . Then, the following holds true.*

i) $V = \{x \in X : \mu_V(x) < 1\}$ for every $V \in \mathcal{B}$.

ii) $\{\mu_V : V \in \mathcal{B}\}$ is a separating family of continuous seminorms on X .

Proof. Recall first that every $V \in \mathcal{B}$ is absorbing, by Prop. 2.3, so μ_V is well-defined. Now, if $\mu_V(x) < 1$ this means that $x \in tV$ for some $t < 1$, in particular $x \in V$. On the other hand, if $x \in V$, then $\frac{x}{t} \in V$ for some $t < 1$, by the continuity of scalar multiplication and because V is open. This proves *i*).

For the second part *ii*), we first note that μ_V defines a seminorm by Lemma 2.2 *iii*). The family is separating, because if $x \neq 0$, we find an open neighborhood $V \subset X$ of

$0 \in X$ such that $x \notin V$ (X is Hausdorff) and by *i*), this shows that $\mu_V(x) \geq 1$. Finally, the continuity follows from Lemma 2.1 *i*), which implies that

$$|\mu_V(x) - \mu_V(y)| \leq \mu_V(x - y) < \epsilon$$

whenever $x - y \in \epsilon V$. In other words, for every $y \in X$, $\epsilon > 0$, we find an open set $y + \epsilon V \subset \mu_V^{-1}((\mu_V(y) - \epsilon, \mu_V(y) + \epsilon))$, which proves the continuity of $\mu_V : X \rightarrow [0, \infty)$. $\square$

Theorem 2.2. *Suppose $\mathcal{P}$ is a separating family of seminorms on a vector space X . For every $p \in \mathcal{P}$ and every $n \in \mathbb{N}$, define*

$$V(p, n) = \{x \in X : p(x) < n^{-1}\}$$

and consider the family $\mathcal{B}$ of finite intersections of translates of these sets $V(p, n)$, for $p \in \mathcal{P}, n \in \mathbb{N}$. Then, $\mathcal{B}$ is a convex, balanced local base for a topology τ on X which turns X into a locally convex, topological vector space. Moreover, the following holds true.

- i) Every $p \in \mathcal{P}$ is continuous.*
- ii) $E \subset X$ is bounded with regards to τ if and only if every $p \in \mathcal{P}$ is bounded on E .*
- iii) If (X, τ_0) is assumed to be a topological vector space with a convex, balanced local base $\mathcal{B}_0$ and if $\mathcal{P}$ denotes the family of continuous seminorms defined as in Theorem 2.1, then the topology τ is equal to $\tau = \tau_0$.*
- iv) If $\mathcal{P}$ is countable, (X, τ) is metrizable with a translation invariant metric.*

Proof. The topology τ is defined as the topology generated by $\mathcal{B}$, that is, a set is open if and only if it is an arbitrary union of translates of the elements in $\mathcal{B}$. Notice that every element in $\mathcal{B}$ is convex and balanced, by the standard seminorm properties. By definition, τ is also a translation invariant topology.

We now show that (X, τ) is a topological vector space. We have to verify that single point sets are closed and that addition and scalar multiplication are continuous. So, consider $x \in X$ and assume $x \neq 0$. Since $\mathcal{P}$ separates points, we find some $p \in \mathcal{P}$ such that $p(x) > 0$. This implies that $x \notin V(p, n)$ if $n \in \mathbb{N}$ is large enough so that $p(x) \geq n^{-1}$. This means that $0 \in X$ does not lie in the open neighborhood $x - V(p, n) \subset X$ of $x \in X$ and thus $x \notin \overline{\{0\}}$. Since $x \neq 0$ was arbitrary, this shows that $\{0\} = \overline{\{0\}}$ is closed. By translation invariance of τ , we conclude that every single point set is closed.

Next, let us show the continuity of addition and scalar multiplication. Let $U \subset X$ be an open neighborhood of $0 \in X$. Then, by definition of τ , we find elements $x_j \in X$, $p_j \in \mathcal{P}$, $n'_j \in \mathbb{N}$ and $n_j \in \mathbb{N}$ for $j \in \{1, \dots, M\}$ such that

$$0 \in \bigcap_{j=1}^M V(p_j, n_j) \subset \bigcap_{j=1}^M (x_j + V(p_j, n'_j)) \subset U.$$

Then, defining

$$V = \bigcap_{j=1}^M V(p_j, 2n_j),$$

we see that $p_j(x + y) \leq p_j(x) + p_j(y) < \frac{1}{n_j}$ for all $j \in \{1, \dots, M\}$ if $x, y \in V$ so that $V + V \subset U$. For a general open set $U' \subset X$, $U' \neq \emptyset$, we thus conclude that $(x + V) + (y + V) \subset U'$ whenever $x + y \in U'$. Thus, $+$: $X \times X \rightarrow X$ is continuous. Similarly, suppose that $\alpha x \in U'$ for some $\alpha \in \mathbb{K}$, $x \in X$ and $U' \subset X$ open. Then $\beta y \in U'$ whenever $(\beta, y) \in \{\gamma \in \mathbb{K} : |\gamma - \alpha| < s^{-1}\} \times (x + tV)$ if $t = s/(1 + |\alpha|s)$ and $s > 0$ is chosen such that $x \in sV$. Indeed, we have that

$$\beta y - \alpha x = \beta(y - x) + (\beta - \alpha)x \in |\beta|tV + |\beta - \alpha|sV \subset \frac{|\beta - \alpha|s + |\alpha|s}{1 + |\alpha|s}V + V \subset V + V \subset U$$

with $V \subset X$ as above, for the open neighborhood $U = U' - \alpha x$ of $0 \in X$. Here, we used that V is balanced. This shows that $\cdot$: $\mathbb{K} \times X \rightarrow X$ is continuous.

We conclude that (X, τ) is a locally convex, topological vector space with a convex, balanced local base for its topology. Now, let's switch to the proof of properties *i*) to *iv*). By Lemma 2.1, we have that $x + \epsilon V(p, 1) \subset p^{-1}((p(x) - \epsilon, p(x) + \epsilon))$ for every $\epsilon > 0$ and $x \in X$. Thus, every $p \in \mathcal{P}$ is continuous with respect to τ , which proves *i*).

Assume next that $E \subset X$ is bounded and let $p \in \mathcal{P}$. By definition of boundedness, $E \subset kV(p, 1)$ for some $k \in \mathbb{N}$ so that $\sup_{x \in E} p(x) < k$. That is, p is bounded on E . On the other hand, let $U \subset X$ be an open neighborhood of $0 \in X$ and assume that every $p \in \mathcal{P}$ is bounded on E . Choosing $p_j \in \mathcal{P}$ and $n_j \in \mathbb{N}$ as above such that $\sup_{x \in E} p_j(x) < k_j$, we see that for $n \geq k_j n_j$

$$E \subset \bigcap_{j=1}^M V(p_j, k_j^{-1}) \subset \bigcap_{j=1}^M V(p_j, n^{-1}n_j) = n \bigcap_{j=1}^M V(p_j, n_j) \subset nU$$

Thus, E is bounded and this concludes part *ii*).

Let's switch to part *iii*). Let $\mathcal{P}$ denote the family of τ_0 -continuous seminorms on X and let τ denote the topology constructed as above. By τ_0 -continuity of $p \in \mathcal{P}$, we see that $V(p, n) = p^{-1}([0, n^{-1}]) \in \tau_0$ for every $n \in \mathbb{N}$. Thus $\tau \subset \tau_0$. On the other hand, if $V \in \mathcal{B}$ and $p = \mu_V$, then by Prop. 2.1, we have that $V = \{x \in X : \mu_V(x) < 1\} = V(p, 1)$. Thus $V \in \tau$ for every $V \in \mathcal{B}$ and thus $\tau_0 \subset \tau$. This proves *iii*).

Finally, assume that $\mathcal{P} = \{p_n : n \in \mathbb{N}\}$ is a countable family of seminorms and set

$$d(x, y) = \max_{n \in \mathbb{N}} \frac{c_n p_n(x - y)}{1 + p_n(x - y)} \quad (2.2)$$

for some fixed, positive sequence $(c_n)_{n \in \mathbb{N}}$ such that $\lim_{n \rightarrow \infty} c_n = 0$. It is straightforward to check that $d : X \times X \rightarrow [0, \infty)$ defines a translation invariant metric (*exercise*). Moreover, we claim that the family of metric balls

$$\left(B_\delta(0) = \{x \in X : d(0, x) < \delta\} \right)_{\delta > 0}$$

forms a convex, balanced local base (at $0 \in X$) for τ . Clearly, the base elements are convex and balanced, so it suffices to show that they form a local base for τ . To this

end, fix $\delta > 0$. Since $\lim_{n \rightarrow \infty} c_n = 0$, we have that $c_n < \delta$ for all, but finitely many $n_1, \dots, n_l \in \mathbb{N}$ so that

$$\begin{aligned} B_\delta &= \bigcap_{n \in \mathbb{N}} \left\{ x \in X : \frac{c_n p_n(x)}{1 + p_n(x)} < \delta \right\} = \bigcap_{j=1}^l \left\{ x \in X : \frac{c_{n_j} p_{n_j}(x)}{1 + p_{n_j}(x)} < \delta \right\} \\ &= \bigcap_{j=1}^l p_{n_j}^{-1} \left(\left[0, \frac{\delta}{c_j - \delta} \right) \right). \end{aligned}$$

The right hand side is equal to a finite intersection of open sets, because all $p \in \mathcal{P}$ are continuous, so that $B_\delta \in \tau$. Hence, the topology generated by d is a subset of τ .

On the other hand, suppose $U \subset X$ is an open neighborhood of $0 \in X$. As in the first part of the proof, we find $p_j \in \mathcal{P}$ and $n_j > 1 \in \mathbb{N}$ for $j \in \{1, \dots, M\}$ such that

$$\bigcap_{j=1}^M V(p_j, n_j) \subset U.$$

Setting $\delta_j = n_j^{-1} < 1$ and choosing $\delta < \frac{1}{2} \min\{c_j \delta_j : j = 1, \dots, M\}$, we see that $\delta < \frac{c_j \delta_j}{1 + \delta_j}$ for every $j \in \{1, \dots, M\}$. This is equivalent to $\frac{\delta}{c_j - \delta} < \delta_j$ for all $j \in \{1, \dots, M\}$ so that

$$\begin{aligned} B_\delta(0) &\subset \bigcap_{j=1}^M \left\{ x \in X : \frac{c_j p_j(x)}{1 + p_j(x)} < \delta \right\} = \bigcap_{j=1}^M \left\{ x \in X : p_j(x) < \frac{\delta}{c_j - \delta} \right\} \\ &\subset \bigcap_{j=1}^M \left\{ x \in X : p_j(x) < \delta_j \right\} \subset U. \end{aligned}$$

Thus, every open set $U \subset X$ with $0 \in U$ contains a metric open ball, proving *iv*). $\square$

In view of the last result, it is useful to revisit Prop. 2.4 which analyses the continuity of linear maps $\Lambda : X \rightarrow Y$ between topological vector spaces. In case that the topologies on X and Y are generated by separating families of seminorms $\mathcal{P}$ and $\mathcal{Q}$, respectively, we can characterize continuity of a linear map $\Lambda : X \rightarrow Y$ more precisely.

Proposition 2.5. *Let X and Y be locally convex, topological vector spaces whose topologies are generated by separating families $\mathcal{P}$ and $\mathcal{Q}$ of seminorms on X and, respectively, on Y . Then, a linear map $\Lambda : X \rightarrow Y$ is continuous if and only if for every $q \in \mathcal{Q}$ there exist $p_1, \dots, p_k \in \mathcal{P}$ for some $k \in \mathbb{N}$ and a constant $C > 0$ such that*

$$q(\Lambda x) \leq C \sum_{j=1}^k p_j(x), \forall x \in X. \quad (2.3)$$

Proof. Suppose $\Lambda : X \rightarrow Y$ is continuous and let $q \in \mathcal{Q}$. Then $q \circ \Lambda : X \rightarrow \mathbb{R}$ is continuous so that $(q \circ \Lambda)^{-1}([0, 1]) \subset X$ is an open neighborhood of $0 \in X$. By definition of the

topology on X , this implies there exist $p_1, \dots, p_k \in \mathcal{P}$ and $n_1, \dots, n_k \in \mathbb{N}$ such that

$$\bigcap_{j=1}^k V(p_j, n_j) \subset (q \circ \Lambda)^{-1}([0, 1)).$$

Now, define $p = \sum_{j=1}^k n_j p_j : X \rightarrow [0, \infty)$. Then every $x \in p^{-1}([0, 1))$ satisfies

$$\sum_{j=1}^k n_j p_j(x) < 1 \implies x \in \bigcap_{j=1}^k V(p_j, n_j) \subset (q \circ \Lambda)^{-1}([0, 1)).$$

Thus, $p^{-1}([0, 1)) \subset (q \circ \Lambda)^{-1}([0, 1))$ so that for every $x \in X$ and $\epsilon > 0$, we have

$$\frac{x}{p(x) + \epsilon} \in p^{-1}([0, 1)) \implies q\left(\Lambda \frac{x}{p(x) + \epsilon}\right) = \frac{q(\Lambda x)}{p(x) + \epsilon} \leq 1.$$

In other words, $q(\Lambda x) \leq p(x) + \epsilon$ for every $x \in X$ and $\epsilon > 0$. Sending $\epsilon \rightarrow 0$, this proves the bound (2.3) for $C = 2 \max_{j=1, \dots, k} n_j$.

Now assume that (2.3) holds true. We want to show that this implies continuity of $\Lambda : X \rightarrow Y$. By definition of the topologies on X and Y , it suffices to show that the preimage of every subbase element $\bigcap_{j=1}^N V(q_j, n_j)$ for $q_j \in \mathcal{Q}$ and $n_j \in \mathbb{N}$ under Λ contains some element $\bigcap_{k=1}^M V(p_k, m_k)$ for $p_k \in \mathcal{P}$ and $n_k \in \mathbb{N}$. Now, for every $j \in \{1, \dots, M\}$, pick $p_{j1}, \dots, p_{jN_j} \in \mathcal{P}$ according to (2.3) such that

$$q_j(\Lambda x) \leq C_j \sum_{k=1}^{N_j} p_{jk}(x)$$

and set $m_j = n_j N_j C_j$. This implies

$$x \in \bigcap_{j=1}^N \bigcap_{k=1}^{N_j} V(p_{jk}, m_j) \implies q_j(\Lambda x) \leq C_j \sum_{k=1}^{N_j} p_{jk}(x) < \frac{1}{n_j}, \forall j \implies \Lambda x \in \bigcap_{j=1}^N V(q_j, n_j),$$

as desired. Hence, $\Lambda : X \rightarrow Y$ is continuous. $\square$

In the last part of this section, we draw a few conclusions from the Baire and Hahn-Banach theorems in the context of locally convex, topological vector spaces. We start with a useful variant of the uniform boundedness principle. Before that, let's recall a variant of Baire's theorem (for its proof, see e.g. [26, Chapter 2]).

Theorem 2.3 (Baire). *Let (X, d) be a complete metric space and assume that*

$$X = \bigcup_{n=1}^{\infty} S_n$$

for a sequence of closed sets $S_n \subset X$, $n \in \mathbb{N}$. Then $\text{int}(S_n) \neq \emptyset$ for some $n \in \mathbb{N}$.

The following result is a direct consequence of Theorem 2.3 for linear maps between Fréchet spaces, that is, metrizable, locally convex, topological vector spaces with a translation invariant metric which are complete. Notice that if X is a Fréchet space, it admits a countable, convex, balanced local base for its topology so that by Propositions 2.1 and 2.2, we can assume without loss of generality that the topology of X is generated by a countable, separating family $(p_n)_{n \in \mathbb{N}}$ of continuous seminorms. Moreover, we may assume without loss of generality that the metric $d : X \times X \rightarrow [0, \infty)$ takes the form as in (2.2), which is a consequence of the following problem.

Problem 2.3. *Let (X, τ) be a topological vector space with local base $\mathcal{B}$ for τ . We say that $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence if and only if for every $V \in \mathcal{B}$, there exists an $N \in \mathbb{N}$ such that $x_n - x_m \in V$ for all $n, m \geq N$. Now, suppose that X is metrizable with a translation invariant metric $d : X \times X \rightarrow [0, \infty)$. Show that $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence if and only if it is a Cauchy sequence in the usual metric sense. Infer that any two invariant metrics $d_1, d_2 : X \times X \rightarrow [0, \infty)$ that are compatible with τ have the same Cauchy sequences and that (X, d_1) is complete if and only if (X, d_2) is complete.*

Theorem 2.4 (Uniform Boundedness Principle). *Let X and Y be Fréchet spaces whose topologies are generated by separating families $(p_n)_{n \in \mathbb{N}}$ and $(q_n)_{n \in \mathbb{N}}$ of continuous seminorms on X and, respectively, Y . Suppose that $\mathcal{F}$ is family of continuous maps from X to Y such that for every $x \in X$, the orbit $\{\Lambda x : \Lambda \in \mathcal{F}\} \subset Y$ is bounded. Then, for every $n \in \mathbb{N}$, there exist $m_1, \dots, m_N \in \mathbb{N}$ and some constant $C > 0$ such that*

$$\sup_{\Lambda \in \mathcal{F}} q_n(\Lambda x) \leq C \sum_{j=1}^N p_{m_j}(x), \quad \forall x \in X.$$

Proof. Consider the complete metric space (X, d_X) with d_X e.g. as in (2.2) and define

$$S_{n,k} = \{x \in X : q_n(\Lambda x) \leq k, \quad \forall \Lambda \in \mathcal{F}\}, \quad \forall n \in \mathbb{N}.$$

For each $k \in \mathbb{N}$, the set $S_{n,k}$ is closed as it is the intersection over $\Lambda \in \mathcal{F}$ of the closed sets $(q_n \circ \Lambda)^{-1}([0, k])$ and by Theorem 2.2 *ii*), the orbit $\{\Lambda x : \Lambda \in \mathcal{F}\} \subset Y$ is bounded if and only if the orbits $\{q_n(\Lambda x) : \Lambda \in \mathcal{F}\}$ are bounded for every $n \in \mathbb{N}$. Thus.

$$X = \bigcup_{k \in \mathbb{N}} S_{n,k}$$

for every $n \in \mathbb{N}$. By Theorem 2.3 at least one $S_{n,k}$ has non-empty interior so we can find some open, metric ball $B_\delta(x_0) \subset \text{int}(S_{n,k})$. This means that

$$\sup_{\Lambda \in \mathcal{F}} q_n(\Lambda x_0 + \delta \Lambda x) \leq k, \quad \forall x \in B_1(0) \implies \sup_{\Lambda \in \mathcal{F}} q_n(\Lambda x) \leq C = C_{\delta,k,x_0}, \quad \forall x \in B_1(0) \subset X.$$

Finally, $B_1(0) \subset X$ is an open neighborhood of $0 \in X$ and contains some base element $\bigcap_{j=1}^N V(p_{m_j}, n_j) \subset B_1(0)$. As in the proof of Prop. 2.5, this implies that

$$\sup_{\Lambda \in \mathcal{F}} q_n(\Lambda x) \leq C, \quad \forall x \in p^{-1}([0, 1]) \implies \sup_{\Lambda \in \mathcal{F}} q_n(\Lambda x) \leq C \sum_{j=1}^N n_j p_{m_j}(x), \quad \forall x \in X,$$

where, as in the proof of Prop. 2.5, we set $p = \sum_{j=1}^N n_j p_{m_j} : X \rightarrow [0, \infty)$. $\square$

A useful consequence used in Section 2.2.3 below is the following.

Corollary 2.1. *Let $X_1, \dots, X_k$ be Fréchet spaces and let $L : X_1 \times X_2 \times \dots \times X_k \rightarrow \mathbb{K}$ be a separately continuous, multilinear functional. Then, L is jointly continuous, i.e. there exist continuous seminorms $p_j : X_j \rightarrow [0, \infty)$ and some $C > 0$ such that*

$$|L(x_1, \dots, x_k)| \leq C p_1(x_1) \dots p_k(x_k), \quad \forall (x_1, \dots, x_k) \in X_1 \times \dots \times X_k.$$

Proof. For simplicity of notation, consider the case $k = 2$ (the general case follows with the same arguments and is left as an *exercise*). L is a multilinear map on the product $X_1 \times X_2$, which is metrizable if both X_1 and X_2 are metrizable. Arguing similarly as in the proof of Prop. 2.4, we readily see (*exercise*) that $L : X_1 \times X_2 \rightarrow \mathbb{K}$ is jointly continuous if and only if it is sequentially continuous at $0 \in X_1 \times X_2$. This uses that $L(\cdot, x_2)$ is continuous (and thus sequentially continuous) on X_1 for each fixed $x_2 \in X_2$ and that $L(x_1, \cdot)$ is continuous (and thus sequentially continuous) on X_2 for each fixed $x_1 \in X_1$. Hence, suppose that $\lim_{n \rightarrow \infty} (x_n, y_n) = 0 \in X_1 \times X_2$. Then, by separate continuity, $\lim_{n \rightarrow \infty} L(x_n, y) = 0$ for every fixed $y \in X_2$, so that the orbits $\{L(x_n, y) : n \in \mathbb{N}\}$ are bounded for every $y \in X_2$. Applying Theorem 2.4 to the sequence $(\Lambda_n = L(x_n, \cdot))_{n \in \mathbb{N}}$ of linear maps on X_2 implies that

$$\sup_{n \in \mathbb{N}} |L(x_n, y)| \leq q(y), \quad \forall y \in X_2,$$

for some continuous seminorm $q : X_2 \rightarrow [0, \infty)$ and therefore that

$$\limsup_{n \rightarrow \infty} |L(x_n, y_n)| \leq \limsup_{n \rightarrow \infty} q(y_n) = 0.$$

This proves that $L : X_1 \times X_2 \rightarrow \mathbb{K}$ is jointly continuous. Finally, arguing as in the proof of Prop. 2.5, we know that $L^{-1}(B_1(0))$ is an open neighborhood of $0 \in X_1 \times X_2$ so that, by definition of the product topology on $X_1 \times X_2$, we find a suitable base element

$$\bigcap_{i=1}^k V(u_i, n_i) \times \bigcap_{j=1}^l V(v_j, m_j) \subset L^{-1}(B_1(0))$$

for seminorms u_i and v_j on X_1 and, respectively, X_2 , and for suitable integers $n_i \in \mathbb{N}$ and $m_j \in \mathbb{N}$, for $i \in \{1, \dots, k\}, j \in \{1, \dots, l\}$. This implies that

$$|L(x, y)| \leq p_1(x) p_2(y), \quad \forall (x, y) \in X_1 \times X_2,$$

where $p_1 = \sum_{i=1}^k n_i u_i : X_1 \rightarrow [0, \infty)$ and $p_2 = \sum_{j=1}^l n_j v_j : X_2 \rightarrow [0, \infty)$. $\square$

We close this section with basic results related to the dual space X^* of a locally convex, topological vector space X . Recall that X^* denotes the linear space of continuous, linear functionals from X to $\mathbb{K}$ if X is a topological vector space. This is a vector space with the standard addition and scalar multiplication of functions. We also would

like to view X^* as a topological vector space. One possible, natural topology is the weak* topology, sometimes denoted by $\sigma(X^*, X)$. This is defined as the topology that is induced by the family $(f_x)_{x \in X}$ of seminorms which are defined by

$$f_x(\Lambda) = |\Lambda x|, \quad \forall \Lambda \in X^*.$$

This family is clearly separating, because $f_x(\Lambda) = 0$ for all $x \in X$ is equivalent to $\Lambda = 0 \in X^*$. By Theorem 2.2, X^* equipped with this topology defines a locally convex, topological vector space. Recall in particular that convergence of functionals

$$\Lambda_n \xrightarrow{\text{weak}^*} \Lambda \text{ as } n \rightarrow \infty \iff \lim_{n \rightarrow \infty} \Lambda_n x = \Lambda x, \quad \forall x \in X$$

corresponds to pointwise convergence, sometimes also referred to as simple convergence. The last result of this section shows that linear functionals on X^* that are continuous w.r.t. the weak* topology are in one to one correspondence with X . Before proving this, recall the following two fundamental results (for the proofs, see e.g. [26, Chapter 3]).

Theorem 2.5 (Hahn-Banach). *Suppose that M is a linear subspace of a vector space X , that $p : X \rightarrow [0, \infty)$ is a seminorm on X and that $f : M \rightarrow \mathbb{K}$ is a linear functional such that $|f(x)| \leq p(x)$, $\forall x \in M$. Then, f extends to a linear functional on $\Lambda : X \rightarrow \mathbb{K}$ such that $|\Lambda x| \leq p(x)$, $\forall x \in X$.*

A well-known consequence is the following result.

Theorem 2.6. *Suppose A and B are disjoint, non-empty, convex sets in a topological vector space X .*

i) *If A is open, there exist $\Lambda \in X^*$ and $\gamma \in \mathbb{R}$ such that*

$$\operatorname{Re} \Lambda x < \gamma \leq \operatorname{Re} \Lambda y, \quad \forall x \in A, y \in B.$$

ii) *If X is a locally convex, topological vector space, A is compact and B closed, there exist $\Lambda \in X^*$ and $\gamma_1, \gamma_2 \in \mathbb{R}$ such that*

$$\operatorname{Re} \Lambda x < \gamma_1 < \gamma_2 < \operatorname{Re} \Lambda y, \quad \forall x \in A, y \in B.$$

In particular, X^ separates points.*

Proposition 2.6. *Let X be a locally convex, topological vector space and consider its dual $(X^*, \sigma(X^*, X))$. Then, for every $\Phi : X^* \rightarrow \mathbb{K}$ that is linear and continuous with respect to $\sigma(X^*, X)$, there exists $x \in X$ such that $\Phi(\Lambda) = \Lambda x$, $\forall \Lambda \in X^*$. As a consequence, the topological dual space of X^* is isomorphic to X .*

Proof. First of all, by definition of $\sigma(X^*, X)$, it is clear that $X^* \ni \Lambda \mapsto \Lambda x \in \mathbb{K}$, defines a linear, continuous functional on X^* , for every $x \in X$. Conversely, suppose that $\Phi : X^* \rightarrow \mathbb{K}$ is a continuous, linear functional on X^* . Then $\Phi^{-1}(B_1(0))$ is open so that

$$\bigcap_{j=1}^M \{ \Lambda \in X^* : |\Lambda x_j| < n_j^{-1} \} \subset \Phi^{-1}(B_1(0))$$

for suitable $x_j \in X$ and $n_j \in \mathbb{N}$, for $j \in \{1, \dots, M\}$. In particular, $\Phi(\Lambda) = 0$ whenever $\Lambda x_j = 0$ for all $j \in \{1, \dots, M\}$ (because then also $|\Phi(n\Lambda)| < 1$ for every $n > 0$), that is

$$\bigcap_{j=1}^M \ker(\Psi_{x_j}) \subset \ker(\Phi),$$

where $\Psi_{x_j}(\Lambda) = \Lambda x_j$. In analogy to the finite dimensional setting, this suggests that

$$\Phi(\Lambda) = \sum_{j=1}^M \alpha_j \Psi_{x_j}(\Lambda) = \Lambda \left(\sum_{j=1}^M \alpha_j x_j \right)$$

for suitable $\alpha = (\alpha_1, \dots, \alpha_M) \in \mathbb{K}^M$ and for all $\Lambda \in X^*$. To prove this, notice indeed that $\Phi(\Lambda)$ is determined by

$$\pi(\Lambda) = (\Lambda x_1, \dots, \Lambda x_M) \in \mathbb{K}^M,$$

in the sense that $\pi(\Lambda) = \pi(\Lambda')$ implies that $\pi(\Lambda - \Lambda') = 0$ implying in turn that $\Phi(\Lambda - \Lambda') = 0$, i.e. $\Phi(\Lambda) = \Phi(\Lambda')$. In other words, $f : \pi(X^*) \rightarrow \mathbb{K}$, defined by

$$f(\pi(\Lambda)) = \Phi(\Lambda)$$

is a well-defined, linear functional from $\pi(X^*) \subset \mathbb{K}^M$ to $\mathbb{K}$. By Theorem 2.5, f extends to a linear functional on $\mathbb{K}^M$ so that, denoting the extension again by f , we get

$$f(y_1, \dots, y_M) = \sum_{j=1}^M \alpha_j y_j$$

for suitable $\alpha = (\alpha_1, \dots, \alpha_M) \in \mathbb{K}^M$ and all $y = (y_1, \dots, y_M) \in \mathbb{K}^M$. As a consequence

$$\Phi(\Lambda) = f(\pi(\Lambda)) = \sum_{j=1}^M \alpha_j \Lambda x_j = \Lambda \left(\sum_{j=1}^M \alpha_j x_j \right), \quad \forall \Lambda \in X^*,$$

as desired. This proves that $\Phi = \Psi_x$ for $x = \sum_{j=1}^M \alpha_j x_j \in X$.

Notice, finally, that $\Psi : X \rightarrow (X^*)^*$, defined by

$$x \mapsto \Psi_x = (X^* \ni \Lambda \mapsto \Lambda x) \in (X^*)^*,$$

is an isomorphism. Indeed, by the previous remarks Ψ is surjective. It is also injective, because $\Psi(x) = \Psi(y)$ implies that $\Lambda x = \Lambda y$ for all $\Lambda \in X^*$ and an application of Theorem 2.6 ii) (to the closed, compact sets $\{x\}, \{y\} \subset X$) shows that this implies $x = y$. $\square$

The last result we collect concerns uniform convergence in the weak* topology.

Proposition 2.7. *Let X be a Fréchet space with dual $(X^*, \sigma(X^*, X))$. Assume that the sequence $(\Lambda_n)_{n \in \mathbb{N}}$ in X^* converges to $\Lambda \in X^*$. Then, the convergence is uniform on compact subsets of X .*

Proof. The proof consists of a standard $\frac{\epsilon}{3}$ argument. Suppose that $K \subset X$ is compact. By assumption, we know that $\lim_{n \rightarrow \infty} \Lambda_n x = \Lambda x$ for every $x \in X$. In particular, $(\Lambda_n x)_{n \in \mathbb{N}}$ is a bounded sequence, for every $x \in X$. By Theorem 2.4, applied to the Fréchet spaces X and $Y = \mathbb{R}$, we find some continuous seminorm $\rho : X \rightarrow [0, \infty)$ so that

$$\sup_{n \in \mathbb{N}} |\Lambda_n x| \leq \rho(x), \quad \forall x \in X.$$

Notice that this also implies that $|\Lambda x| \leq \rho(x)$ for every $x \in X$. Now, let $\epsilon > 0$. By continuity of ρ , we can cover $K \subset X$ by finitely many open sets $(U_j)_{j=1}^M$ such that

$$\rho(x - y) < \frac{\epsilon}{3}, \quad \forall x, y \in U_j$$

for all $j \in \{1, \dots, M\}$. Now, pick points $x_j \in U_j$ and $N \in \mathbb{N}$ such that

$$|\Lambda_n x_j - \Lambda x_j| < \frac{\epsilon}{3}, \quad \forall n \geq N, j \in \{1, \dots, M\}.$$

Since $K \subset \bigcup_{j=1}^M U_j$, this implies that for all $n \geq N$, we have

$$\begin{aligned} \sup_{x \in K} |\Lambda_n x - \Lambda x| &\leq \sup_{x \in K} \min_{j \in \{1, \dots, M\}} (|\Lambda_n x - \Lambda_n x_j| + |\Lambda_n x_j - \Lambda x_j| + |\Lambda x_j - \Lambda x|) \\ &\leq 2 \sup_{x \in K} \min_{j \in \{1, \dots, M\}} \rho(x - x_j) + \frac{\epsilon}{3} < \epsilon, \end{aligned}$$

which concludes the uniform convergence. $\square$

2.2 Distributions and Tempered Distributions

In this section, we study distributions on open subsets $\Omega \subset \mathbb{R}^n$ which generalize the notion of a classical function. As already pointed out in Section 1.4, the construction of quantum fields (and their expectations) necessitates to work with singular objects that can not be described as classical functions. Mathematically, such objects can be obtained as linear functionals on a suitable test function space, identifying a (sufficiently regular, e.g. locally integrable) function $f : \Omega \rightarrow \mathbb{C}$ via duality with the linear functional

$$\varphi \mapsto \int_{\Omega} dx f(x) \varphi(x).$$

Clearly, for sufficiently large test function space, this fully determines the function f , but there are also many functionals that do not correspond to (e.g. locally integrable) functions (e.g. the Dirac δ -functional). It may be useful to keep in mind here that the smaller the test function space, the more irregularity is allowed for the distributions.

In this section, we first study a class of distributions on $\Omega = \mathbb{R}^n$ whose elements are called tempered distributions. Afterwards, we discuss basic properties of general distributions whose definition is a bit more involved compared to tempered distributions (the latter are easily defined using the tools from the previous Section 2.1). Tempered distributions are more regular than general distributions and admit in particular a Fourier transform, which turns out useful in view as remarked in Section 1.4.

2.2.1 Tempered Distributions on $\mathbb{R}^n$

The space of tempered distributions corresponds to the topological dual space of the space $\mathcal{S}(\mathbb{R}^n)$ of Schwartz functions, viewed as a locally convex, topological vector space. Let us make this more precise and collect various properties of $\mathcal{S}(\mathbb{R}^n)$. For a multi-index $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$, we abbreviate in the sequel

$$x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}, \quad \partial^\alpha = \partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n}.$$

We call a smooth map $\varphi \in C^\infty(\mathbb{R}^n, \mathbb{C})$ a Schwartz function, or a function of rapid decrease, if for every pair of multi-indices $\alpha, \beta \in \mathbb{N}_0^n$, we have that

$$|\varphi|_{\alpha, \beta} = \sup_{x \in \mathbb{R}^n} |(x^\alpha \partial^\beta \varphi)(x)| < \infty. \quad (2.4)$$

The set of Schwartz functions is denoted by $\mathcal{S}(\mathbb{R}^n)$. With standard addition and scalar multiplication of functions, $\mathcal{S}(\mathbb{R}^n)$ forms a vector space. It is straightforward to verify (*exercise*) that $|\cdot|_{\alpha, \beta} : \mathcal{S}(\mathbb{R}^n) \rightarrow [0, \infty)$ is a seminorm for every $\alpha, \beta \in \mathbb{N}_0^n$. Moreover, the family $(|\cdot|_{\alpha, \beta})_{\alpha, \beta \in \mathbb{N}_0^n}$ is separating. Indeed, if $\varphi \neq 0$, then $\varphi(x) \neq 0$ for some $x \in \mathbb{R}^n$ and therefore $|\varphi|_{0,0} > 0$. According to Theorem 2.2, we can thus view $\mathcal{S}(\mathbb{R}^n)$ as a locally convex, topological vector space. In the remaining sections, all topological statements about $\mathcal{S}(\mathbb{R}^n)$ refer to this topology.

Proposition 2.8. $\mathcal{S}(\mathbb{R}^n)$ is a Fréchet space.

Proof. By Problem 2.3, it suffices to show that every Cauchy sequence $(\varphi_n)_{n \in \mathbb{N}}$ has a limit $\varphi \in \mathcal{S}(\mathbb{R}^n)$. The Cauchy property and the definition of the topology imply that

$$|\varphi_k - \varphi_l|_{\alpha, \beta} \rightarrow 0$$

as $k, l \rightarrow \infty$, for every $\alpha, \beta \in \mathbb{N}_0^n$. Choosing $\alpha = \beta = 0$ shows that $\|\varphi_k - \varphi_l\|_\infty = \sup_{x \in \mathbb{R}^n} |\varphi_k(x) - \varphi_l(x)| \rightarrow 0$ as $k, l \rightarrow \infty$. By completeness of $(C_b(\mathbb{R}^n), \|\cdot\|_\infty)$, we find some $\varphi \in C(\mathbb{R}^n)$ with $\|\varphi\|_\infty < \infty$ such that $\lim_{k \rightarrow \infty} \|\varphi - \varphi_k\|_\infty = 0$. Analogously, we find for every other pair $\alpha, \beta \in \mathbb{N}_0^n$ a continuous function $\varphi_{\alpha, \beta} \in C(\mathbb{R}^n)$ such that $\lim_{k \rightarrow \infty} \|\varphi_{\alpha, \beta} - x^\alpha \partial^\beta \varphi_k\|_\infty = 0$. Completeness of $\mathcal{S}(\mathbb{R}^n)$ thus follows if we can prove that $\varphi \in C^\infty(\mathbb{R}^n)$ and that $x^\alpha \partial^\beta \varphi = \varphi_{\alpha, \beta}$ for all $\alpha, \beta \in \mathbb{N}_0^n$. For concreteness, consider first the case $\alpha = 0, \beta = (1, 0, \dots, 0)$. Then, since

$$\varphi_k(x) = \varphi_k(0, x_2, \dots, x_n) + \int_0^{x_1} dt (\partial_{x_1} \varphi_k)(t, x_2, \dots, x_n),$$

the uniform convergence $\lim_{k \rightarrow \infty} \|\varphi_{0, (1, 0, \dots, 0)} - \partial_{x_1} \varphi_k\|_\infty = 0$ implies that

$$\varphi(x) = \varphi(0, x_2, \dots, x_n) + \int_0^{x_1} dt (\varphi_{0, (1, 0, \dots, 0)})(t, x_2, \dots, x_n),$$

so that φ admits the continuous partial derivative $\varphi_{0, (1, 0, \dots, 0)} \in C(\mathbb{R}^n)$ in direction $e_1 = (1, 0, \dots, 0) \in \mathbb{R}^n$. Similarly, we argue for the other directions and conclude that $\varphi \in$

$C^1(\mathbb{R}^n)$ with $\partial^\beta \varphi = \varphi_{0,\beta}$ for every $\beta \in \mathbb{N}_0^n$ with $|\beta| = \sum_{j=1}^n \beta_j = 1$. Repeating the argument inductively (replacing φ by $\partial^\beta \varphi$ for $|\beta| = 1$), we readily conclude that $\varphi \in C^\infty(\mathbb{R}^n)$ with $\partial^\beta \varphi = \varphi_{0,\beta}$ for every $\beta \in \mathbb{N}_0^n$ and, furthermore, that $x^\alpha \partial^\beta \varphi = \varphi_{\alpha,\beta}$ for every $\alpha, \beta \in \mathbb{N}_0^n$. We leave the detailed verification of these facts as an *exercise*. $\square$

Problem 2.4. Based on Proposition 2.5, show that the embedding $\mathcal{S}(\mathbb{R}^n) \hookrightarrow L^p(\mathbb{R}^n)$ is continuous, for every $p \geq 1$.

We denote the topological dual space of $\mathcal{S}(\mathbb{R}^n)$ by $\mathcal{S}'(\mathbb{R}^n)$. Its elements are called tempered distributions. Let us list a few basic examples.

Example 2.1 ($L^p(\mathbb{R}^n)$ Functions). For $f \in L^p(\mathbb{R}^n)$, $p \geq 1$, consider the functional

$$\mathcal{S}(\mathbb{R}^n) \ni \varphi \mapsto \Lambda_f(\varphi) = \int_{\mathbb{R}^n} dx f(x) \varphi(x) \in \mathbb{C}.$$

Then, Λ_f is well-defined by Hölder's inequality, using $\mathcal{S}(\mathbb{R}^n) \subset L^q(\mathbb{R}^n)$ for $\frac{1}{p} + \frac{1}{q} = 1$. The fact that $\Lambda_f \in \mathcal{S}'(\mathbb{R}^n)$ follows from Prop. 2.5 and the bound

$$|\Lambda_f(\varphi)| \leq \|f\|_{L^p} \left(\int_{\mathbb{R}^n} dx \frac{(1+|x|)^{n+1} |\varphi(x)|^q}{(1+|x|)^{n+1}} \right)^{\frac{1}{q}} \leq \|f\|_{L^p} \sum_{|\alpha| \leq \lceil \frac{n+1}{q} \rceil} |\varphi|_{\alpha,0}.$$

Example 2.2 (Polynomially Bounded Measures). Suppose $\mu : \mathcal{B}(\mathbb{R}^n) \rightarrow [0, \infty)$ is a polynomially bounded measure which means that $\mu([-L, L]^n) \leq C(L^{pn} + 1)$ for some $C > 0$, $p \in \mathbb{N}$ and for all $L > 0$. Then

$$\mathcal{S}(\mathbb{R}^n) \ni \varphi \mapsto \Lambda_\mu(\varphi) = \int_{\mathbb{R}^n} \mu(dx) \varphi(x) \in \mathbb{C}$$

defines a tempered distribution. Indeed, arguing as in the previous example, we see that

$$|\Lambda_\mu(\varphi)| \leq \int_{\mathbb{R}^n} \frac{\mu(dx)}{(1+|x|)^{np+2}} \sum_{|\alpha| \leq np+2} |\varphi|_{\alpha,0} \leq \left(\sum_{L=1}^{\infty} \frac{2C}{L^2} \right) \sum_{|\alpha| \leq np+2} |\varphi|_{\alpha,0}.$$

Example 2.3 (Dirac δ and its Derivatives). The Dirac functional centered at $x_0 \in \mathbb{R}^n$, denoted by $\delta_{x_0} : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{C}$, is defined $\mathcal{S}(\mathbb{R}^n)$ by

$$\delta_{x_0}(\varphi) = \varphi(x_0)$$

so that in particular $|\delta_{x_0}(\varphi)| \leq |\varphi|_{0,0}$ and thus $\delta_{x_0} \in \mathcal{S}'(\mathbb{R}^n)$, by Prop. 2.5. It is well-known that there does not exist a locally integrable function $f_{x_0} \in L^1_{\text{loc}}(\mathbb{R}^n)$ such that

$$\delta_{x_0}(\varphi) = \int dx f_{x_0}(x) \varphi(x), \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^n).$$

That is, δ_{x_0} does not correspond to a classical function (*exercise*). It can be identified, however, with the finite, positive Borel measure on $\mathbb{R}^n$, defined by

$$\delta_{x_0}(\Omega) = \begin{cases} 1 & : x_0 \in \Omega, \\ 0 & : \text{else.} \end{cases}$$

There are also distributions which do not correspond to Borel measures. For $\alpha \in \mathbb{N}_0^n$, define the α -th derivative $\partial^\alpha \delta_{x_0} : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{C}$ of δ_{x_0} via duality by

$$(\partial^\alpha \delta_{x_0})(\varphi) = (-1)^{|\alpha|} \delta_{x_0}(\partial^\alpha \varphi), \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^n).$$

We leave it as an *exercise* to show that $\partial^\alpha \delta_{x_0} \in \mathcal{S}'(\mathbb{R}^n)$ and that, for $\alpha \neq (0, 0, \dots, 0)$, $\partial^\alpha \delta_{x_0} \in \mathcal{S}'(\mathbb{R}^n)$ does not correspond to a finite, positive Borel measure on $\mathbb{R}^n$.

Like in the last Example 2.3, duality can be used to extend many common operations from classical to generalized functions. Let us collect several examples.

Example 2.4 (Distributional Derivatives). Let $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$, then for $\alpha \in \mathbb{N}_0^n$, we define its α -th $\partial^\alpha \Lambda : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{C}$ weak, or distributional, derivative by

$$(\partial^\alpha \Lambda)(\varphi) = (-1)^{|\alpha|} \Lambda(\partial^\alpha \varphi), \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Symbolically, this generalizes the well-known integration by parts formula in $\mathbb{R}^n$, i.e.

$$\int_{\mathbb{R}^n} dx (\partial^\alpha \Lambda)(x) \varphi(x) = (-1)^{|\alpha|} \int_{\mathbb{R}^n} dx \Lambda(x) (\partial^\alpha \varphi)(x).$$

Example 2.5 (Translations and Linear Coordinate Transformations). Let $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ and $(a, L) \in \mathbb{R}^n \times \text{GL}(n)$, then we define $\tau_a \Lambda : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{C}$ and $L\Lambda : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{C}$ by

$$(\tau_a \Lambda)(\varphi) = \Lambda(\varphi(\cdot - a)), \quad (L\Lambda)(\varphi) = |\det L| \Lambda(\varphi(L\cdot)), \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Notice that this is consistent with the action $L\varphi = \varphi(L^{-1}\cdot)$ on function $\varphi \in \mathcal{S}(\mathbb{R}^n)$.

Example 2.6 (Multiplication by Polynomially Bounded Functions). Let $f \in C^\infty(\mathbb{R}^n)$ be a polynomially bounded function which means, by definition, that for every $\alpha \in \mathbb{N}_0^n$, there exist $m_\alpha \in \mathbb{N}_0, C_\alpha > 0$ such that

$$|\partial^\alpha f(x)| \leq C_\alpha (1 + |x|^2)^{m_\alpha}, \quad \forall x \in \mathbb{R}^n.$$

Then, for $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$, we define $f\Lambda : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{C}$ by

$$f\Lambda(\varphi) = \Lambda(f\varphi), \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Example 2.7 (Convolution with Schwartz Functions). Let $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ and $\psi \in \mathcal{S}(\mathbb{R}^n)$. Then, we define the convolution $\Lambda * \psi : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{C}$ by

$$(\Lambda * \psi)(\varphi) = \Lambda(\psi(\cdot - \cdot) * \varphi), \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Problem 2.5. Prove that the examples 2.4, 2.5, 2.6 and 2.7 are all elements in $\mathcal{S}'(\mathbb{R}^n)$.

As a last example, we want to define the Fourier transform of tempered distributions. Before doing that, let's record some basic properties of the Fourier transform as a map on $\mathcal{S}(\mathbb{R}^n)$ (for the proof of the next theorem, see e.g. [22]). For $\varphi \in \mathcal{S}(\mathbb{R}^n)$, we define

$$\widehat{\varphi}(p) = \int_{\mathbb{R}^n} dx e^{-2\pi i p x} \varphi(x), \quad \check{\varphi}(p) = \int_{\mathbb{R}^n} dx e^{2\pi i p x} \varphi(x), \quad \forall p \in \mathbb{R}^n. \quad (2.5)$$

Theorem 2.7. The maps $\widehat{\cdot}: \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n)$ and $\check{\cdot}: \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n)$ are continuous, linear maps. Furthermore, the following holds true:

- i) $\widehat{\cdot}: \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n)$ is the inverse map of $\check{\cdot}: \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n)$.
- ii) For every $\alpha, \beta \in \mathbb{N}_0^n$ and $\varphi \in \mathcal{S}(\mathbb{R}^n)$, the function $\mathbb{R}^n \ni p \mapsto ip^\alpha \partial^\beta \widehat{\varphi}(p) \in \mathbb{C}$ is equal to the Fourier transform of $\mathbb{R}^n \ni x \mapsto \partial^\alpha ((-ix)^\beta \varphi(x)) \in \mathbb{C}$.
- iii) For every $\varphi, \chi \in \mathcal{S}(\mathbb{R}^n)$, we have that

$$\langle \varphi, \chi \rangle_2 = \int_{\mathbb{R}^n} dx \overline{\varphi}(x) \chi(x) = \langle \widehat{\varphi}, \widehat{\chi} \rangle_2$$

so that, in particular, $\|\varphi\|_2 = \|\widehat{\varphi}\|_2$.

- iv) For every $\varphi, \chi \in \mathcal{S}(\mathbb{R}^n)$, we have that $\widehat{\varphi * \chi} = \widehat{\varphi} \widehat{\chi}$.

Example 2.8 (Fourier Transform). Let $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$, then we define $\widehat{\Lambda}: \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbb{C}$ by

$$\widehat{\Lambda}(\varphi) = \Lambda(\widehat{\varphi}).$$

Symbolically, this generalizes Plancherel's formula

$$\int_{\mathbb{R}^n} dx \widehat{\Lambda}(x) \varphi(x) = \int_{\mathbb{R}^n} dx \widehat{\Lambda}(x) \overline{(\widehat{\varphi})}(x) = \int_{\mathbb{R}^n} dx \widehat{\Lambda}(x) \overline{(\widehat{\widehat{\varphi}})}(x) = \int_{\mathbb{R}^n} dx \Lambda(x) \widehat{\varphi}(x).$$

Note that $\widehat{\Lambda} \in \mathcal{S}'(\mathbb{R}^n)$, because $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ and $\widehat{\cdot}: \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n)$ is continuous.

In Sections 2.2.3 and 4.1 we study $\mathcal{S}'(\mathbb{R}^n)$ in more detail and analyze the previous operations as maps on $\mathcal{S}'(\mathbb{R}^n)$, viewed as a locally convex, topological vector space.

2.2.2 Distributions on Open Subsets $\Omega \subset \mathbb{R}^n$

As another application of the results on locally convex spaces, we discuss in this section general distributions on open subsets $\Omega \subset \mathbb{R}^n$.

When studying for instance the analyticity properties of (the Laplace transforms) of tempered distributions (see Section (4.1) below), it is useful to have a less restrictive notion of a distribution that allows for more irregular behavior. We may obtain such objects by considering a test function space smaller than e.g. the Schwartz space. A

natural choice is $C_c^\infty(\Omega)$, the space of smooth, compactly supported functions in Ω . In the setting of distributions, this space is typically denoted by $\mathcal{D}(\Omega)$ and identified with

$$\mathcal{D}(\Omega) = \bigcup_{\substack{K \subset \Omega, \\ K \text{ compact}}} \mathcal{D}_K(\Omega), \quad \mathcal{D}_K(\Omega) = \{\varphi \in C^\infty(\Omega) : \text{supp } \varphi \subset K\}. \quad (2.6)$$

For each compact $K \subset \Omega$, the space $\mathcal{D}_K(\Omega)$ carries a locally convex topology as defined in Theorem 2.2, using the countable family of norms $(|\cdot|_j)_{j \in \mathbb{N}_0}$, defined by

$$|\varphi|_j = \max \{|\partial^\alpha \varphi(x)| : x \in \Omega, \alpha \in \mathbb{N}_0^n \text{ with } |\alpha| \leq j\}, \quad \forall \varphi \in \mathcal{D}_K(\Omega).$$

In the sequel, topological results about $\mathcal{D}_K(\Omega)$ refer to this topology, denoted by τ_K .

Problem 2.6. *Prove that $\mathcal{D}_K(\Omega)$ is a Fréchet space and that it has the Heine-Borel property, for each compact $K \subset \Omega$.*

As a union of, say, a countable family of spaces $\mathcal{D}_{K_i}(\Omega)$ for increasing, compact subsets $K_i \subset \Omega$, it is natural to equip $\mathcal{D}(\Omega)$ with the locally convex, metrizable topology induced by the seminorms defined by $p_{ij}(\varphi) = \max_{x \in K_i, |\alpha| \leq j} |\partial^\alpha \varphi(x)|$, for $\varphi \in \mathcal{D}(\Omega)$, $i \in \mathbb{N}, j \in \mathbb{N}_0$. This topology has, however, the disadvantage that it is not complete. For example, choose $\Omega = \mathbb{R}$, $K_i = [-i, i]$, choose some function $\varphi \in \mathcal{D}(\mathbb{R})$ with $\text{supp } \varphi \subset (0, 1)$ such that $\varphi|_{(\epsilon, 1-\epsilon)} > 0$ (for some small $\epsilon > 0$) and define for $j \in \mathbb{N}$

$$\psi_n = \sum_{j=1}^n \varphi(\cdot - j) \in \mathcal{D}(\mathbb{R}).$$

Then, we have for every $i \in \mathbb{N}, j \in \mathbb{N}_0$ and $m \leq n$ that

$$|p_{ij}(\psi_n - \psi_m)| \leq \sum_{k=m+1}^n |p_{ij}(\varphi(\cdot - k))| = 0$$

whenever $m > i$, so $(\psi_n)_{n \in \mathbb{N}}$ is Cauchy in $\mathcal{D}(\Omega)$. Moreover, the pointwise limit $\psi(x) = \lim_{n \rightarrow \infty} \psi_n(x)$ exists for every $x \in \mathbb{R}$, but $\psi \notin \mathcal{D}(\Omega)$. Indeed, its support is not compact, because $\psi(x+j) = \varphi(x) > 0$ for every $x \in (\epsilon, 1-\epsilon)$, $j \in \mathbb{N}$. The property that prohibits completeness is that each seminorm p_{ij} only measures what happens in K_i .

We now introduce another locally convex topology which, among other things, forces Cauchy sequences to have support in a fixed compact set. This property implies in particular its completeness, based on the completeness of $\mathcal{D}_K(\Omega)$. To this end, set

$$\beta = \{W \subset \mathcal{D}(\Omega) : W \text{ convex, balanced; } \mathcal{D}_K(\Omega) \cap W \in \tau_K \forall \text{ compact } K \subset \Omega\}$$

and define τ to be the family of unions of sets $\varphi + W$, for $\varphi \in \mathcal{D}(\Omega)$ and $W \in \beta$. We remark that the general concept behind this definition is the so called strict inductive limit topology (see e.g. [21] for more on this and also Problem 2.7 below).

Theorem 2.8. τ defines a topology on $\mathcal{D}(\Omega)$ and $\beta \subset \tau$ is a local base for τ . Moreover, the space $(\mathcal{D}(\Omega), \tau)$ is a locally convex, topological vector space.

Proof. Clearly, $\emptyset, \mathcal{D}(\Omega) \in \beta$, hence $\emptyset, \mathcal{D}(\Omega) \in \tau$ and τ is stable under taking arbitrary unions of translates of elements of β , which follows immediately from its definition. To conclude that τ defines a topology and that β is a local base for it, it suffices (*exercise*) to show that for every $V_1, V_2 \in \tau$, $\varphi \in V_1 \cap V_2$, there exists some $W \in \beta$ such that

$$\varphi + W \subset V_1 \cap V_2.$$

By assumption, we find $\varphi_i \in \mathcal{D}(\Omega)$, $W_i \in \beta$ such that $\varphi \in \varphi_i + W_i \subset V_i$, for $i \in \{1, 2\}$. Then, choose $K \subset \Omega$ compact such that $\varphi, \varphi_1, \varphi_2 \in \mathcal{D}_K(\Omega)$. Now, $\varphi - \varphi_i \in W_i$ and $W_i \cap \mathcal{D}_K(\Omega) \in \tau_K$, so the continuity of scalar multiplication in $\mathcal{D}_K(\Omega)$ implies also that

$$\frac{1}{1 - \delta_i}(\varphi - \varphi_i) \in W_i$$

for suitable, small $\delta_i > 0$, $i \in \{1, 2\}$. By convexity of W_i , this implies that

$$\varphi - \varphi_i + \delta_i W_i \subset (1 - \delta_i)W_i + \delta_i W_i \subset W_i$$

and therefore that $\varphi + \delta_i W_i \subset \varphi_i + W_i \subset V_i$ for $i \in \{1, 2\}$. In other words, $\varphi + W \subset V_1 \cap V_2$ for $W = \delta_1 W_1 \cap \delta_2 W_2$, which concludes the proof that τ defines a topology.

Clearly, τ has a convex, local base so it only remains to prove that single point sets are closed and that addition as well as scalar multiplication are continuous maps. Let $\varphi, \psi \in \mathcal{D}(\Omega)$ be such that $\psi \neq \varphi$. We construct an open set $V \in \tau$ that contains ψ , but not φ . This implies that the complement of $\{\varphi\}$ is open and hence $\{\varphi\}$ is closed. One way to choose V is as follows. Set

$$W = \left\{ \zeta \in \mathcal{D}(\Omega) : |\zeta|_0 = \sup_{x \in \Omega} |\zeta(x)| < |\varphi - \psi|_0 \right\}$$

so that $W \cap \mathcal{D}_K(\Omega) \in \tau_K$ for every compact $K \subset \Omega$. Since W is balanced and convex, we conclude that $W \in \beta$ and thus $\psi + W \in \tau$. Finally, $\varphi - \psi \notin W$, i.e. $\varphi \notin \psi + W \in \tau$.

Next, let's prove the continuity of addition. Let $\varphi_1 + \varphi_2 \in V$ for some $\varphi_1, \varphi_2 \in \mathcal{D}(\Omega)$, $V \in \tau$. Then $0 \in V - \varphi_1 - \varphi_2$ is open and contains $0 \in \mathcal{D}(\Omega)$, so we find a local base element $W \subset V - \varphi_1 - \varphi_2$ with $W \in \beta$, i.e. $\varphi_1 + \varphi_2 + W \subset V$. By convexity, we have

$$(\varphi_1 + \frac{1}{2}W) + (\varphi_2 + \frac{1}{2}W) \subset \varphi_1 + \varphi_2 + W \subset V,$$

so that $+^{-1}(V) \subset \mathcal{D}(\Omega) \times \mathcal{D}(\Omega)$ is open. Thus, $+: \mathcal{D}(\Omega) \times \mathcal{D}(\Omega) \rightarrow \mathcal{D}(\Omega)$ is continuous.

Finally, consider the scalar multiplication in $\mathcal{D}(\Omega)$. Arguing similarly as in the previous step, let $\lambda \in \mathbb{K}$, $\varphi \in \mathcal{D}(\Omega)$, $W \in \beta$ and $V \in \tau$ such that

$$\lambda\varphi + W \subset V.$$

Now, if $|\kappa - \lambda| < \delta$, $\delta\varphi \in \epsilon_1 W$ and $\psi - \varphi \in \epsilon_2 W$, we have that

$$\kappa\psi - \lambda\varphi = (\kappa - \lambda)\varphi + \kappa(\psi - \varphi) \in \epsilon_1 W + (|\lambda| + \delta)\epsilon_2 W,$$

by balancedness of W . Now, $\epsilon_1 W \in \beta$ so for $K \subset \Omega$ compact such that $\varphi \in \mathcal{D}_K(\Omega)$, we have that $\epsilon_1 W \cap \mathcal{D}_K(\Omega) \in \tau_K$. By Prop. 2.3, this implies that $\delta\varphi \in \epsilon_1 W$ if $\delta > 0$ is sufficiently small. Choosing e.g. $\epsilon_1 = \frac{1}{2}$ and $\delta > 0$ accordingly, and then $\epsilon_2 > 0$ such that $(|\lambda| + \delta)\epsilon_2 = \frac{1}{2}$, the convexity of W implies that

$$\cdot(B_\delta(\lambda) \times \varphi + \epsilon_2 W) \subset \lambda\varphi + \epsilon_1 W + (|\lambda| + \delta)\epsilon_2 W \subset \lambda\varphi + W.$$

This implies that $\cdot : \mathbb{K} \times \mathcal{D}(\Omega) \rightarrow \mathcal{D}(\Omega)$ is continuous. $\square$

In the sequel, whenever we speak of $\mathcal{D}(\Omega)$ as a topological space, we refer to the topology τ . Possible motivations for this choice are discussed in the next two problems.

Problem 2.7. *Verify that τ is the strongest locally convex topology on $\mathcal{D}(\Omega)$ such that all the embeddings $\mathcal{D}_K(\Omega) \hookrightarrow \mathcal{D}(\Omega)$ are continuous.*

Problem 2.8. *Denote by $\mathcal{P}$ the set of seminorms on the linear space $\mathcal{D}(\Omega)$. Show that τ is equal to the locally convex topology generated by all seminorms $p \in \mathcal{P}$ such that $p \circ \iota_K : \mathcal{D}_K(\Omega) \rightarrow [0, \infty)$ is continuous for every compact $K \subset \Omega$. Here, $\iota_K : \mathcal{D}_K(\Omega) \hookrightarrow \mathcal{D}(\Omega)$ denotes the canonical embedding.*

Remark 2.1. *For a concrete family of seminorms that generates τ , see e.g. [10].*

The key properties of $\mathcal{D}(\Omega)$ are summarized in the next theorem.

Theorem 2.9. *The following statements hold true:*

- i) *A convex, balanced subset $V \subset \mathcal{D}(\Omega)$ is open if and only if $V \in \beta$.*
- ii) *For compact $K \subset \Omega$, the topology τ_K equals the subspace topology of $\mathcal{D}_K(\Omega)$ inherited from $\mathcal{D}(\Omega)$.*
- iii) *A set $E \subset \mathcal{D}(\Omega)$ is bounded if and only if $E \subset \mathcal{D}_K(\Omega)$ for some compact $K \subset \Omega$ and there are numbers C_j , $j \in \mathbb{N}_0$, such that*

$$\sup_{\varphi \in E} |\varphi|_j = \sup_{\varphi \in E} \sup_{x \in \Omega, |\alpha| \leq j} |\partial^\alpha \varphi(x)| \leq C_j, \quad \forall j \in \mathbb{N}_0.$$

- iv) *$\mathcal{D}(\Omega)$ has the Heine-Borel property.*
- v) *If $(\varphi_i)_{i \in \mathbb{N}}$ is a Cauchy sequence in $\mathcal{D}(\Omega)$, then $\varphi_i \in \mathcal{D}_K(\Omega), \forall i \in \mathbb{N}$, for some compact $K \subset \Omega$ and we have that*

$$\lim_{i, j \rightarrow \infty} |\varphi_i - \varphi_j|_k = 0, \quad \forall k \in \mathbb{N}_0.$$

- vi) *If $\lim_{i \rightarrow \infty} \varphi_i = 0$ in $\mathcal{D}(\Omega)$, then $\varphi_i \in \mathcal{D}_K(\Omega), \forall i \in \mathbb{N}$, for some compact $K \subset \Omega$ and*

$$\lim_{i \rightarrow \infty} |\varphi_i|_k = 0, \quad \forall k \in \mathbb{N}_0.$$

vii) If $(\varphi_i)_{i \in \mathbb{N}}$ is a Cauchy sequence in $\mathcal{D}(\Omega)$, then there exists a limit $\varphi \in \mathcal{D}(\Omega)$ such that $\lim_{i \rightarrow \infty} \varphi_i = \varphi$ in $\mathcal{D}(\Omega)$.

Remark 2.2. The last part vii) shows that $\mathcal{D}(\Omega)$ has the important property of being topologically complete. $\mathcal{D}(\Omega)$ is, however, not metrizable (as explained below).

Proof. We start with i). By definition of τ , we know that $\beta \subset \tau$. So assume, on the other hand, that $V \subset \mathcal{D}(\Omega)$ is convex, balanced and open. We want to show that $\mathcal{D}_K(\Omega) \cap V \in \tau_K$, for every compact $K \subset \Omega$. Assuming without loss of generality $\mathcal{D}_K(\Omega) \cap V \neq \emptyset$, pick $\varphi \in \mathcal{D}_K(\Omega) \cap V$. Then, since β is a local base for τ , we find some $W \in \beta$ such that $\varphi + W \subset V$ and therefore that

$$\varphi + (\mathcal{D}_K(\Omega) \cap W) \subset \mathcal{D}_K(\Omega) \cap V.$$

Since $W \in \beta$, we have that $\mathcal{D}_K(\Omega) \cap W \in \tau_K$. The previous step thus means that for every $\varphi \in \mathcal{D}_K(\Omega) \cap V$, we find an open neighborhood of φ in τ_K that is contained in $\mathcal{D}_K(\Omega) \cap V$, i.e. $\mathcal{D}_K(\Omega) \cap V \in \tau_K$. This proves i).

To prove ii), note first that the previous part shows that the restriction of τ to $\mathcal{D}_K(\Omega)$ is contained in τ_K . On the other hand, assume that $U \in \tau_K$. By definition of τ_K , we find for every $\varphi \in U$ some $j_\varphi \in \mathbb{N}_0$ and $\delta_\varphi > 0$ such that

$$\varphi + \{\psi \in \mathcal{D}_K(\Omega) : |\psi|_{j_\varphi} < \delta_\varphi\} = \varphi + \mathcal{D}_K(\Omega) \cap W_\varphi \subset U,$$

where we set $W_\varphi = \{\psi \in \mathcal{D}(\Omega) : |\psi|_{j_\varphi} < \delta_\varphi\}$. Clearly, W_φ is convex, balanced and such that $W_\varphi \cap \mathcal{D}_{K'}(\Omega) \in \tau_{K'}$ for each compact $K' \subset \Omega$. Now, setting

$$V = \bigcup_{\varphi \in U} (\varphi + W_\varphi) \in \tau,$$

we obtain that $U = \mathcal{D}_K(\Omega) \cap V$. Thus $\tau_K \subset \tau|_{\mathcal{D}_K(\Omega)} = \{V \cap \mathcal{D}_K(\Omega) : V \in \tau\}$.

To prove iii), we argue by contraposition. Suppose $E \subset \mathcal{D}(\Omega)$ is such that $E \not\subset \mathcal{D}_K(\Omega)$ for every compact $K \subset \Omega$. Then, e.g. by using an increasing sequence of compact subsets of Ω , we find a sequence $(\varphi_k)_{k \in \mathbb{N}}$ in E and a sequence $(x_k)_{k \in \mathbb{N}}$ which does not have a limit point in Ω such that $\varphi_k(x_k) \neq 0, \forall k \in \mathbb{N}$ (*exercise*). Then, define

$$S_k = \{\psi \in \mathcal{D}(\Omega) : |\psi(x_k)| < k^{-1}|\varphi_k(x_k)|\}, \forall k \in \mathbb{N}.$$

Each set S_k is clearly convex and balanced. Moreover, the linear functional

$$\mathcal{D}_K(\Omega) \ni \psi \mapsto \Lambda_k(\psi) = \delta_{x_k}(\psi) = \psi(x_k) \in \mathbb{K}$$

lies in $\mathcal{D}_K(\Omega)'$, the topological dual of $\mathcal{D}_K(\Omega)$, because

$$|\Lambda_k(\psi)| \leq |\psi|_0, \forall \psi \in \mathcal{D}_K(\Omega).$$

and since $\mathcal{D}_K(\Omega)$ is a Fréchet space, for every compact $K \subset \Omega$. This implies that $S_k \cap \mathcal{D}_K(\Omega) = (|\cdot| \circ \Lambda_k)^{-1}([0, k^{-1}|\varphi_k(x_k)|]) \in \tau_K$ for every compact $K \subset \Omega$, and since

every such $K \subset \Omega$ contains at most finitely many points of the sequence $(x_k)_{k \in \mathbb{N}}$ (note that $S_k \cap \mathcal{D}_K(\Omega) = \mathcal{D}_K(\Omega) \in \tau_K$ if $x_k \notin K$), it also implies that

$$S = \bigcap_{k \in \mathbb{N}} S_k \in \beta \subset \tau.$$

But then $\varphi_k \notin kS$ for every $k \in \mathbb{N}$, so no constant multiple of S contains E , which means that E is not bounded. Hence, if E is bounded, it is necessarily contained in some $\mathcal{D}_K(\Omega)$ for some compact $K \subset \Omega$, in particular it is bounded as a subset of $\mathcal{D}_K(\Omega)$, by part *ii*). By the definition of τ_K and Theorem 2.2 *ii*), this shows that

$$\sup_{\varphi \in E} |\varphi|_j \leq C_j,$$

for every $j \in \mathbb{N}_0$, concluding the first direction of *iii*). Conversely, if $E \subset \mathcal{D}_K(\Omega)$ is bounded in $\mathcal{D}_K(\Omega)$, then for every open neighborhood $S_K \subset \mathcal{D}_K(\Omega)$ of $0 \in \mathcal{D}_K(\Omega)$, there exists some $s > 0$ such that $E \subset tS_K$, for all $t > s$. In particular, if $S \subset \mathcal{D}(\Omega)$ is an open neighborhood of $0 \in \mathcal{D}(\Omega)$, we conclude from *ii*) that $E \subset tS \cap \mathcal{D}_K(\Omega) \subset tS$ for all $t > s$, so that $E \subset \mathcal{D}_K(\Omega) \subset \mathcal{D}(\Omega)$ is bounded in $\mathcal{D}(\Omega)$, too.

The remaining properties *iv*) to *vii*) strongly rely on the previous parts. For *iv*), suppose $E \subset \mathcal{D}(\Omega)$ is closed and bounded. Then, by *iii*), $E \subset \mathcal{D}_K(\Omega)$ for some compact $K \subset \Omega$ and by *ii*), it is closed and bounded as a subset of $\mathcal{D}_K(\Omega)$. The space $\mathcal{D}_K(\Omega)$ has the Heine-Borel property by Problem 2.6. Invoking once again *ii*), this shows that $E \subset \mathcal{D}(\Omega)$ is compact. Similarly, by Problem 2.2, every Cauchy and every convergent sequence is bounded. Such sequences are therefore contained in some $\mathcal{D}_K(\Omega)$ and Cauchy and, respectively, converge with regards to τ_K by *ii*). This implies *v*) and *vi*) (whose seminorm statements are just reformulations of the Cauchy and convergence properties in $\mathcal{D}_K(\Omega)$). Using additionally the completeness of $\mathcal{D}_K(\Omega)$, we also conclude *vii*). $\square$

Based on the previous definitions, we denote by $\mathcal{D}'(\Omega)$ the dual space of $\mathcal{D}(\Omega)$, i.e.

$$\mathcal{D}'(\Omega) = \{\Lambda : \mathcal{D}(\Omega) \rightarrow \mathbb{K} : \Lambda \text{ linear and continuous}\}.$$

Its elements are called distributions on $\Omega \subset \mathbb{R}^n$. Continuity is characterized as follows.

Theorem 2.10. *Let $\Lambda : \mathcal{D}(\Omega) \rightarrow Y$ be a linear map into a locally convex, topological vector space Y . Then, the following properties are equivalent:*

- i) Λ is continuous.*
- ii) Λ is bounded.*
- iii) If $\lim_{k \rightarrow \infty} \varphi_k = 0$ in $\mathcal{D}(\Omega)$, then $\lim_{k \rightarrow \infty} \Lambda \varphi_k = 0$ in Y .*
- iv) $\Lambda|_{\mathcal{D}_K(\Omega)} : \mathcal{D}_K(\Omega) \rightarrow Y$ is continuous for every compact $K \subset \Omega$.*

In particular, $\Lambda \in \mathcal{D}'(\Omega)$ if and only if for every compact $K \subset \Omega$ there exists some $j \in \mathbb{N}_0$ and some constant $C > 0$ such that $|\Lambda \varphi| \leq C |\varphi|_j$, $\forall \varphi \in \mathcal{D}_K(\Omega)$.

Proof. By Prop. 2.4, we know that *i*) implies *ii*). To see that *ii*) implies *iii*), assume $\lim_{k \rightarrow \infty} \varphi_k = 0$ in $\mathcal{D}(\Omega)$. Then, Theorem 2.9 *vi*) tells us that there is some compact set $K \subset \Omega$ such that $\varphi_k \in \mathcal{D}_K(\Omega)$, for all $k \in \mathbb{N}$, and that $\lim_{k \rightarrow \infty} \varphi_k = 0$ in $\mathcal{D}_K(\Omega)$. Since $\mathcal{D}_K(\Omega)$ is a Fréchet space, Prop. 2.4 applied to the restriction $\Lambda|_{\mathcal{D}_K(\Omega)} : \mathcal{D}_K(\Omega) \rightarrow Y$ shows that $\lim_{k \rightarrow \infty} \Lambda \varphi_k = 0$. Hence, *ii*) implies *iii*).

If we assume *iii*), then it also clearly holds true that if $\lim_{k \rightarrow \infty} \varphi_k = 0$ in $\mathcal{D}_K(\Omega)$, then $\lim_{k \rightarrow \infty} \Lambda \varphi_k = 0$ in Y , for every compact $K \subset \Omega$. Note indeed that $\lim_{k \rightarrow \infty} \varphi_k = 0$ in $\mathcal{D}_K(\Omega)$ implies that $\lim_{k \rightarrow \infty} \iota_K(\varphi_k) = 0$ in $\mathcal{D}(\Omega)$, by continuity of $\iota_K : \mathcal{D}_K(\Omega) \rightarrow \mathcal{D}(\Omega)$, so that $\lim_{k \rightarrow \infty} \varphi_k = 0$ in $\mathcal{D}(\Omega)$. By Prop. 2.4 and the fact that $\mathcal{D}_K(\Omega)$ is Fréchet, this implies that $\Lambda|_{\mathcal{D}_K(\Omega)} : \mathcal{D}_K(\Omega) \rightarrow Y$ is continuous, i.e. part *iv*).

Next, let $U \subset Y$ be a convex, balanced, open neighborhood of $0 \in Y$. Then, by linearity, $\Lambda^{-1}(U)$ is a convex and balanced set (*exercise*). By Theorem 2.9 *ii*), it is open if and only if $\Lambda^{-1}(U) \cap \mathcal{D}_K(\Omega) \in \tau_K$, for every compact $K \subset \Omega$. This is the case if $\Lambda|_{\mathcal{D}_K(\Omega)} : \mathcal{D}_K(\Omega) \rightarrow Y$ is continuous. Since the topology on Y is generated by unions of translates of convex, balanced sets, this proves that *iv*) implies *i*).

Finally, the characterization of $\mathcal{D}'(\Omega)$ is a reformulation of part *iv*) for the special case $Y = \mathbb{K}$. This uses once more that $\mathcal{D}_K(\Omega)$ is Fréchet and Prop. 2.5. $\square$

Example 2.9 (Differential Operators). Every differential operator $\partial^\alpha : \mathcal{D}(\Omega) \rightarrow \mathcal{D}(\Omega)$, $\alpha \in \mathbb{N}_0^n$, is continuous. This follows from Theorem 2.10 *iv*) and the bounds

$$|\partial^\alpha \varphi|_j \leq |\varphi|_{j+|\alpha|}, \quad \forall \varphi \in \mathcal{D}(\Omega), j \in \mathbb{N}_0,$$

which imply that $\partial^\alpha|_{\mathcal{D}_K(\Omega)} : \mathcal{D}_K(\Omega) \rightarrow \mathcal{D}_K(\Omega)$ is continuous for every compact $K \subset \Omega$.

Example 2.10 (Dirac δ). The Dirac δ functional, introduced in Example 2.3, naturally defines an element in $\mathcal{D}'(\Omega)$ by setting $\delta_x(\varphi) = \varphi(x)$, $\forall \varphi \in \mathcal{D}(\Omega)$. Then

$$|\delta_x(\varphi)| \leq |\varphi|_0, \quad \forall \varphi \in \mathcal{D}(\Omega),$$

so continuity follows from Theorem 2.10 *iv*). It is interesting to note that this observation can be used to show that $\mathcal{D}(\Omega)$ is not metrizable. To this end, notice that

$$\mathcal{D}_K(\Omega) = \bigcap_{x \in K^c} \{\varphi \in \mathcal{D}(\Omega) : \delta_x(\varphi) = 0\} \subset \mathcal{D}(\Omega)$$

is closed, for every compact $K \subset \Omega$. Moreover, the interior $(\mathcal{D}_K(\Omega))^\circ = \emptyset$, for every such $K \subset \Omega$. Indeed, assume by contradiction there was some element $\varphi \in (\mathcal{D}_K(\Omega))^\circ$. Choosing a suitable, compact $K' \subset \Omega$ such that $K \subset K' \subset \Omega$, $(K' \setminus K)^\circ \neq \emptyset$, we know that $\emptyset \neq \mathcal{D}_{K'}(\Omega) \cap (\mathcal{D}_K(\Omega))^\circ \in \tau_{K'}$ (as $(\mathcal{D}_K(\Omega))^\circ$ is open in $\mathcal{D}(\Omega)$) so that in particular

$$U_\epsilon = \{\psi \in \mathcal{D}_{K'}(\Omega) : |\psi - \varphi|_0 < \epsilon\} \subset \mathcal{D}_{K'}(\Omega) \cap (\mathcal{D}_K(\Omega))^\circ$$

for some $\epsilon > 0$. Choosing a small bump function $\psi_\epsilon \in C_c^\infty((K' \setminus K)^\circ)$ with $|\psi_\epsilon|_0 < \epsilon$, we find the contradiction that $\varphi + \psi_\epsilon \in U_\epsilon \subset (\mathcal{D}_K(\Omega))^\circ \subset \mathcal{D}_K(\Omega)$.

The upshot is that $\mathcal{D}(\Omega)$ is not metrizable. For it is equal to a countable union

$$\mathcal{D}(\Omega) = \bigcup_{j=1}^{\infty} \mathcal{D}_{K_j}(\Omega)$$

of closed subspaces and since $\mathcal{D}(\Omega)$ is complete, Theorem 2.3 would imply that at least one of the subspaces $\mathcal{D}_{K_j}(\Omega)$ would have non-empty interior if $\mathcal{D}(\Omega)$ was metrizable.

Example 2.11 (Tempered Distributions). For $\Omega = \mathbb{R}^n$, we have $\mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$. That is, every tempered distribution is a distribution. To see this, pick $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$, then

$$|\Lambda(\varphi)| \leq C_{\alpha,\beta} |\varphi|_{\alpha,\beta}, \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^n)$$

for every $\alpha, \beta \in \mathbb{N}_0^n$ and suitable constants $C_{\alpha,\beta} > 0$. This implies that

$$|\Lambda(\varphi)| \leq C_{\alpha,\beta} |\varphi|_{\alpha,\beta} \leq C'_{K,\alpha,\beta} |\varphi|_{|\beta|}, \quad \forall \varphi \in \mathcal{D}_K(\Omega)$$

and thus $\Lambda \in \mathcal{D}'(\Omega)$. Conversely, there are $\Lambda \in \mathcal{D}'(\Omega)$ which are not tempered, e.g.

$$\Lambda = \sum_{k=0}^{\infty} x^k \delta_{-k} \in \mathcal{D}'(\mathbb{R}), \quad \Lambda' = \sum_{k=0}^{\infty} \partial^k \delta_{-k} \in \mathcal{D}'(\mathbb{R}).$$

Like for tempered distributions, we can generalize many operations for classical functions to distributions. We leave it as an *exercise* to show that, for instance, Examples 2.4, 2.5, 2.6 and 2.7 readily generalize to $\mathcal{D}'(\Omega)$ (for general, open $\Omega \subset \mathbb{R}^n$).

We conclude this section with two important results on $\mathcal{D}'(\Omega)$. While the first deals with its completeness, the second provides a precise, structural description of its elements (locally). In the sequel, we turn $\mathcal{D}'(\Omega)$ into a locally convex, topological vector space by equipping it with its weak* topology, based on Theorem 2.2. Recall that this implies

$$\Lambda_j \xrightarrow{\text{weak}^*} \Lambda \in \mathcal{D}'(\Omega) \iff \lim_{j \rightarrow \infty} \Lambda_j \varphi = \Lambda \varphi, \quad \forall \varphi \in \mathcal{D}(\Omega).$$

Similarly, $(\Lambda_j)_{j \in \mathbb{N}}$ is Cauchy in $\mathcal{D}'(\Omega)$ if and only if $(\Lambda_j \varphi)_{j \in \mathbb{N}}$ is (in $\mathbb{K}$), for all $\varphi \in \mathcal{D}(\Omega)$.

Theorem 2.11. *$\mathcal{D}'(\Omega)$ is complete. If $(\Lambda_j)_{j \in \mathbb{N}}$ converges to $\Lambda \in \mathcal{D}'(\Omega)$, the convergence is uniform on bounded subsets of $\mathcal{D}(\Omega)$ and, furthermore, $(\partial^\alpha \Lambda_j)_{j \in \mathbb{N}}$ converges to $\partial^\alpha \Lambda$, for every $\alpha \in \mathbb{N}_0^n$.*

Proof. Assume that $(\Lambda_j \varphi)_{j \in \mathbb{N}}$ is Cauchy in $\mathbb{K}$, for every $\varphi \in \mathcal{D}(\Omega)$. By the completeness of $\mathbb{K}$, we define the map $\Lambda : \mathcal{D}(\Omega) \rightarrow \mathbb{K}$ by

$$\Lambda \varphi = \lim_{j \rightarrow \infty} \Lambda_j \varphi, \quad \forall \varphi \in \mathcal{D}(\Omega).$$

This map is clearly linear, so it only remains to show that Λ is continuous. To this end, we show that $\Lambda|_{\mathcal{D}_K(\Omega)} : \mathcal{D}_K(\Omega) \rightarrow \mathbb{K}$ is continuous for every compact $K \subset \Omega$. Then, the claim follows from Theorem 2.10 *iv*).

So, let $K \subset \Omega$ be compact. Similarly as in the proof of Theorem 2.4, define the sets

$$S_k = \{\varphi \in \mathcal{D}_K(\Omega) : |\Lambda_j(\varphi)| \leq k, \forall j \in \mathbb{N}\}, \forall k \in \mathbb{N}.$$

Then $S_k \subset \mathcal{D}_K(\Omega)$ is a (countable intersection of closed subsets and thus) closed in the Fréchet space $\mathcal{D}_K(\Omega)$. Since $\Lambda\varphi = \lim_{j \rightarrow \infty} \Lambda_j\varphi$ for all $\varphi \in \mathcal{D}_K(\Omega)$, we have that

$$\mathcal{D}_K(\Omega) = \bigcup_{k=1}^{\infty} S_k.$$

By Theorem 2.3, there exists at least one set S_k such that $S_k^\circ \neq \emptyset$. Arguing as in the proof of Theorem 2.4 (*exercise*), this implies that

$$\sup_{j \in \mathbb{N}} |\Lambda_j\varphi| \leq C|\varphi|_m, \forall \varphi \in \mathcal{D}_K(\Omega)$$

for some $m \in \mathbb{N}$ and some constant $C > 0$. Hence, also $|\Lambda\varphi| \leq C|\varphi|_m$ for all $\varphi \in \mathcal{D}_K(\Omega)$, which proves the continuity of $\Lambda|_{\mathcal{D}_K(\Omega)} : \mathcal{D}_K(\Omega) \rightarrow \mathbb{K}$. Thus, $\Lambda \in \mathcal{D}'(\Omega)$.

Now, let $\lim_{j \rightarrow \infty} \Lambda_j = \Lambda$ in $\mathcal{D}'(\Omega)$ and let $E \subset \mathcal{D}(\Omega)$ be bounded. By Problem 2.2 and Theorem 2.9, $\overline{E}$ is bounded, hence compact in $\mathcal{D}(\Omega)$, and we also know that $E \subset \overline{E} \subset \mathcal{D}_K(\Omega)$ for some compact $K \subset \Omega$. In particular, $\overline{E}$ is compact in the Fréchet space $\mathcal{D}_K(\Omega)$. By Prop. 2.7, this implies $(\Lambda_j)_{j \in \mathbb{N}}$ converges uniformly to Λ on $E \subset \overline{E}$.

Finally, the convergence of the derivatives follows from the pointwise convergence

$$\lim_{j \rightarrow \infty} (\partial^\alpha \Lambda_j)(\varphi) = \lim_{j \rightarrow \infty} \Lambda_j((-1)^{|\alpha|} \partial^\alpha \varphi) = \Lambda((-1)^{|\alpha|} \partial^\alpha \varphi) = (\partial^\alpha \Lambda)(\varphi)$$

for every $\varphi \in \mathcal{D}(\Omega)$ and every $\alpha \in \mathbb{N}_0^n$. □

At the beginning of this section, we introduced distributions as a generalization of classical functions in order to allow for singular behavior which can not be captured by regular (e.g. locally integrable) functions. How much irregularity is possible in the context of distributions? This question is answered by the following result. A precise characterization of tempered distributions is discussed in the next Section 2.2.3.

Theorem 2.12. *Let $\Lambda \in \mathcal{D}'(\Omega)$ and let $K \subset \Omega$ be compact. Then, there exists a continuous function $f : \Omega \rightarrow \mathbb{K}$ and some $\alpha \in \mathbb{N}_0^n$ such that $\Lambda|_{\mathcal{D}_K(\Omega)} = \partial^\alpha f$, that is*

$$\Lambda\varphi = (-1)^{|\alpha|} \int_{\Omega} dx f(x) (\partial^\alpha \varphi)(x), \forall \varphi \in \mathcal{D}_K(\Omega).$$

Remark 2.3. *Theorem 2.12 describes the local regularity of distributions. Some further analysis can be used to obtain their global structure, too - see e.g. [26, Theorem 6.27].*

Proof. Let $Q \subset \mathbb{R}^n$ be some open cube, centered at $0 \in \mathbb{R}^n$, that contains $K \subset \Omega$. In the following, constants that depend on K or Q are typically denoted by C and may change from line to line, as usual. For every $\psi \in \mathcal{D}_Q = \mathcal{D}_Q(\mathbb{R}^n)$, we have that

$$|\psi|_0 \leq C \max_{x \in Q} |\partial_i \psi(x)|$$

for each $i \in \{1, \dots, n\}$, by a first order Taylor expansion. Iterating this, we obtain

$$|\psi|_m \leq C \max_{x \in Q} |T^m \psi(x)|$$

for every $m \in \mathbb{N}$, where $T = \partial_1 \partial_2 \dots \partial_n$. On the other hand, we know that

$$\psi(y) = \int_{Q_y} dx (T\psi)(x),$$

where $y \in Q$, $Q_y = \{x \in Q : x_i \leq y_i, \forall i = 1, \dots, n\}$. In particular, T is injective on $\mathcal{D}_Q$.

Now, since $\Lambda \in \mathcal{D}'(\Omega)$, we have for some $m \in \mathbb{N}$ and some $C > 0$ that

$$|\Lambda(\varphi)| \leq C |\varphi|_m, \quad \forall \varphi \in \mathcal{D}_K(\Omega).$$

Combined with the previous observations, this shows that

$$|\Lambda\varphi| \leq C \max_{x \in Q} |T^m \varphi(x)| \leq C \int_Q dx |(T^{m+1}\varphi)(x)|, \quad \forall \varphi \in \mathcal{D}_K(\Omega).$$

By injectivity of T on $\mathcal{D}_Q$, note that $T^{m+1} : \mathcal{D}_K(\Omega) \rightarrow \mathcal{D}_K(\Omega)$ is injective as well. We can therefore define a linear operator $\Lambda' : \text{ran}(T^{m+1}) \rightarrow \mathcal{D}_K(\Omega)$ by setting

$$\Lambda'(T^{m+1}\varphi) = \Lambda\varphi, \quad \forall \varphi \in \mathcal{D}_K(\Omega).$$

Theorem 2.5 shows that Λ' extends to some element $\Lambda' \in (L^1(Q))^*$ so that, by duality, there exists an element $g \in L^\infty(Q)$ such that

$$\Lambda\varphi = \Lambda'(T^{m+1}\varphi) = \int_K dx g(x) (T^{m+1}\varphi)(x), \quad \forall \varphi \in \mathcal{D}_K(\Omega).$$

Finally, defining $h \in C(\mathbb{R}^n)$ by

$$h(y) = \int_{-\infty}^{y_1} dx_1 \dots \int_{-\infty}^{y_n} dx_n g(x) \chi_Q(x), \quad \forall y \in \mathbb{R}^n,$$

we conclude that $T\Lambda_h = \partial_1 \partial_2 \dots \partial_n \Lambda_h = (-1)^n \Lambda_g$ in $\mathcal{D}'_Q$ and thus

$$\begin{aligned} \Lambda\varphi &= (-1)^n \int_K dx h(x) (T^{m+2}\varphi)(x) \\ &= (-1)^{n(m+2)} \int_K dx ((-1)^{n+n(m+2)} h(x)) (\partial_1^{m+2} \dots \partial_n^{m+2} \varphi)(x), \quad \forall \varphi \in \mathcal{D}_K(\Omega). \end{aligned}$$

This proves the claim for $f = (-1)^{n+n(m+2)} h|_\Omega \in C(\Omega)$. □

Problem 2.9. Express the Dirac functional $\delta = \delta_0 \in \mathcal{D}'(\mathbb{R})$ as a derivative of a continuous function such that $\delta(\varphi) = \Lambda_f(\partial^\alpha \varphi)$, $\forall \varphi \in \mathcal{D}(\mathbb{R})$, for some $f \in C(\mathbb{R})$ and $\alpha \in \mathbb{N}$.

2.2.3 Regularity and Nuclear Theorems

In this section, we characterize the regularity of tempered distributions. This is based on the so called **N**-representation of $\mathcal{S}(\mathbb{R}^n)$ and $\mathcal{S}'(\mathbb{R}^n)$ which has many additional, interesting consequences such as the Nuclear Theorem. We follow [21, Appendix V.3].

The **N**-representation describes the spaces $\mathcal{S}(\mathbb{R}^n)$ and $\mathcal{S}'(\mathbb{R}^n)$ as suitable sequence spaces. This is based on the expansion of Schwartz elements in terms of the products of the Hermite functions, which form an orthonormal basis of $L^2(\mathbb{R})$. To make this precise, we first collect some preliminary facts. Recall the seminorms $(|\cdot|_{\alpha,\beta})_{\alpha,\beta \in \mathbb{N}_0^n}$ introduced in Section 2.2.1 and define another family of seminorms $(|\cdot|_{\alpha,\beta,2})_{\alpha,\beta \in \mathbb{N}_0^n}$ by

$$|\varphi|_{\alpha,\beta,2} = \|x^\alpha \partial^\beta \varphi\|_{L^2(\mathbb{R}^n)}, \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Lemma 2.3. *The families $(|\cdot|_{\alpha,\beta})_{\alpha,\beta \in \mathbb{N}_0^n}$ and $(|\cdot|_{\alpha,\beta,2})_{\alpha,\beta \in \mathbb{N}_0^n}$ are equivalent.*

Proof. On the one hand, we have that

$$\begin{aligned} |\varphi|_{\alpha,\beta,2}^2 &= \int_{\mathbb{R}^n} dx |x^\alpha (\partial^\beta \varphi)(x)|^2 \leq \int_{\mathbb{R}^n} dx \frac{(1 + |x|^{n+1}) |x|^{2|\alpha|} |\partial^\beta \varphi(x)|^2}{1 + |x|^{n+1}} \\ &\leq C \sum_{\gamma \in \mathbb{N}_0^n: |\gamma| \leq |\alpha| + \frac{n+1}{2}} |\varphi|_{\gamma,\beta}^2 \end{aligned}$$

for every $\varphi \in \mathcal{S}(\mathbb{R}^n)$. Conversely, by Cauchy-Schwarz, a similar estimate shows that

$$\begin{aligned} |\varphi|_{\alpha,\beta} &= \sup_{x \in \mathbb{R}^n} \left| \int_{-\infty}^{x_1} dy_1 \dots \int_{-\infty}^{x_n} dy_n ((\partial_1 \dots \partial_n)(y^\alpha \partial^\beta \varphi))(y) \right| \\ &\leq \sup_{x \in \mathbb{R}^n} \int_{\mathbb{R}} \frac{dy_1}{1 + |y_1|} \dots \int_{\mathbb{R}} \frac{dy_n}{1 + |y_n|} \prod_{j=1}^n (1 + |y_j|) |(\partial_1 \dots \partial_n)(y^\alpha \partial^\beta \varphi))(y)| \\ &\leq C \sum_{\substack{\gamma \in \mathbb{N}_0^n: |\gamma| \leq |\alpha| + n, \\ \delta \in \mathbb{N}_0^n: |\delta| \leq |\beta| + n}} |\varphi|_{\gamma,\delta,2}. \end{aligned}$$

for every $\varphi \in \mathcal{S}(\mathbb{R}^n)$, concluding the proof. $\square$

The previous lemma shows that the locally convex topology on $\mathcal{S}(\mathbb{R}^n)$ can equivalently be characterized through the seminorms $(|\cdot|_{\alpha,\beta,2})_{\alpha,\beta \in \mathbb{N}_0^n}$. Below, we introduce yet another family of equivalent seminorms that are related to the Hermite functions $(\phi_j)_{j \in \mathbb{N}_0}$. The latter are explicitly given by

$$\phi_j(x) = \frac{(-1)^j}{\pi^{\frac{1}{4}} \sqrt{2^j j!}} e^{\frac{1}{2}x^2} \left(\frac{d}{dx} \right)^j e^{-x^2}, \quad \forall x \in \mathbb{R}, j \in \mathbb{N}_0$$

and correspond to the eigenfunctions of the harmonic oscillator $H_{\text{osc}} = -\frac{d^2}{dx^2} + x^2$ on $L^2(\mathbb{R})$. Standard results on Schrödinger operators imply that H_{osc} is essentially self-adjoint on $C_c^\infty(\mathbb{R})$, that the spectrum $\sigma(H_{\text{osc}}) = \sigma_d(H_{\text{osc}}) \subset [0, \infty)$ is purely discrete and

that H_{osc} admits a unique, positive ground state (for the proofs, see e.g. [22, Chapter X.4] and [24, Chapter XIII.12]). Now, $\phi_0 \in L^2(\mathbb{R})$ is positive, because $\phi_0(x) = \pi^{-\frac{1}{4}} e^{-\frac{1}{2}x^2} > 0$ for all $x \in \mathbb{R}$, and it solves the eigenvalue equation

$$\left(-\frac{d^2}{dx^2} + x^2\right)\phi_0 = \phi_0.$$

Hence, it is equal to the unique, positive ground state of H_{osc} (note that $\phi_0 \in D_{H_{\text{osc}}}$ lies in the domain of $H_{\text{osc}} = (H_{\text{osc}})|_{C_c^\infty(\mathbb{R})}$). To find all eigenfunctions of H_{osc} in the first place, one notices (e.g. by elliptic regularity) that they are smooth and furthermore that

$$(H_{\text{osc}})|_{D_{H_{\text{osc}}} \cap C^\infty(\mathbb{R})} = 2A^*A + 1, \quad \text{where} \quad A = \frac{1}{\sqrt{2}}\left(x + \frac{d}{dx}\right), \quad A^* = \frac{1}{\sqrt{2}}\left(x - \frac{d}{dx}\right).$$

The so called ladder (or creation and annihilation) operators, A^* and A satisfy on $C^\infty(\mathbb{R})$ the canonical commutation relations $[A, A^*] = 1$, $[A, A] = 0$, $[A^*, A^*] = 0$, so that

$$(2A^*A + 1)(A^*)^j = (A^*2A^*A + 2A^*)(A^*)^{j-1} + (A^*)^j = (A^*)^j(2A^*A) + (2j+1)(A^*)^j, \quad \forall j \in \mathbb{N}.$$

This suggests to find the eigenfunctions of H_{osc} by solving $2A^*A\phi_0 = 0$ (which implies $A\phi_0 = 0$), to verify that $(A^*)^j\phi_0 \in L^2(\mathbb{R})$ does not vanish for each $j \in \mathbb{N}$ and to normalize the resulting eigenfunctions. It implies that for suitable C_j , $j \in \mathbb{N}_0$, we have

$$\phi_j = C_j(A^*)^j\phi_0 \quad \text{with} \quad H_{\text{osc}}\phi_j = (2A^*A + 1)\phi_j = (2j+1)\phi_j, \quad \forall j \in \mathbb{N}_0.$$

These observations, whose verification is left as an *exercise*, imply that $(\phi_j)_{j \in \mathbb{N}_0}$ is an orthonormal sequence in $L^2(\mathbb{R})$. The fact that it is an orthonormal basis of $L^2(\mathbb{R})$, i.e. complete, follows by showing that $\sigma_d(H_{\text{osc}}) \setminus \{2j+1 : j \in \mathbb{N}_0\} = \emptyset$. This is also left as an *exercise*, see e.g. [22, Chapter X, Problems 30 & 31] for a detailed strategy.

Now, let $(\phi_\alpha)_{\alpha \in \mathbb{N}_0^n} = (\phi_{\alpha_1} \otimes \dots \otimes \phi_{\alpha_n})_{\alpha \in \mathbb{N}_0^n}$ denote the corresponding product basis of $L^2(\mathbb{R}^n)$, set $\mathbf{N} = A^*A : \mathcal{S}(\mathbb{R}) \rightarrow \mathcal{S}(\mathbb{R})$ and define for $\beta \in \mathbb{N}_0^n$

$$\|\cdot\|_\beta = \|(\mathbf{N} + 1)^\beta \cdot\|_{L^2(\mathbb{R}^n)}, \quad (\mathbf{N} + 1)^\beta = \prod_{j=1}^n (\mathbf{N}_j + 1)^{\beta_j} : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n).$$

Here, $\mathbf{N}_j = \mathbf{1}_{L^2(\mathbb{R})} \otimes \dots \otimes \mathbf{N} \otimes \dots \otimes \mathbf{1}_{L^2(\mathbb{R})}$ acts as $\mathbf{N}$ on the j -th coordinate in $\mathcal{S}(\mathbb{R}^n)$.

Problem 2.10. *Prove that $(\|\cdot\|_\beta)_{\beta \in \mathbb{N}_0^n}$ defines a family of seminorms on $\mathcal{S}(\mathbb{R}^n)$ that is equivalent to the family $(|\cdot|_{\alpha, \beta, 2})_{\alpha, \beta \in \mathbb{N}_0^n}$.*

Now, we are ready to characterize $\mathcal{S}(\mathbb{R}^n)$ as a sequence space. To this end, define

$$\mathbf{S}_n = \left\{ a = (a_\alpha)_{\alpha \in \mathbb{N}_0^n} : \sup_{\alpha \in \mathbb{N}_0^n} |\alpha|^m |a_\alpha| < \infty, \quad \forall m \in \mathbb{N}_0 \right\}$$

and equip it with the locally convex topology that is induced by the seminorms

$$\|(a_\alpha)_{\alpha \in \mathbb{N}_0^n}\|_\beta = \|((\alpha + 1)^\beta a_\alpha)_{\alpha \in \mathbb{N}_0^n}\|_{\ell^2(\mathbb{N}_0^n)} \quad \text{for} \quad (\alpha + 1)^\beta = \prod_{j=1}^n (\alpha_j + 1)^{\beta_j}, \quad \forall \alpha \in \mathbb{N}_0^n.$$

Finally, recall that every $\varphi \in \mathcal{S}(\mathbb{R}^n)$ admits an $L^2(\mathbb{R}^n)$ -basis expansion of the form

$$\varphi = \sum_{\alpha \in \mathbb{N}_0^n} \langle \phi_\alpha, \varphi \rangle \phi_\alpha \quad \text{such that} \quad \|\varphi\|_{L^2(\mathbb{R}^n)} = \left\| (\langle \phi_\alpha, \varphi \rangle)_{\alpha \in \mathbb{N}_0^n} \right\|_{\ell^2(\mathbb{N}_0^n)}.$$

Theorem 2.13 (**N**-representation of $\mathcal{S}(\mathbb{R}^n)$). *The map $\iota : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathbf{S}_n$, given by*

$$\mathcal{S}(\mathbb{R}^n) \ni \varphi \mapsto \iota(\varphi) = (\langle \phi_\alpha, \varphi \rangle)_{\alpha \in \mathbb{N}_0^n} \in \mathbf{S}_n$$

defines a topological isomorphism between $\mathcal{S}(\mathbb{R}^n)$ and $\mathbf{S}_n$.

Proof. Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$ so that $\|\varphi\|_\beta < \infty$ for every $\beta \in \mathbb{N}_0^n$. Then, observe that

$$\|\varphi\|_\beta^2 = \left\langle \varphi, \prod_{j=1}^n (\mathbf{N}_j + 1)^{2\beta_j} \varphi \right\rangle = \sum_{\alpha \in \mathbb{N}_0^n} |a_\alpha|^2 \prod_{j=1}^n (\alpha_j + 1)^{2\beta_j} = \|(a_\alpha)_{\alpha \in \mathbb{N}_0^n}\|_\beta^2 \quad (2.7)$$

for $a_\alpha = \langle \phi_\alpha, \varphi \rangle$, $\alpha \in \mathbb{N}_0^n$. Since $\beta \in \mathbb{N}_0^n$ is arbitrary, this clearly implies

$$\sup_{\alpha \in \mathbb{N}_0^n} |\alpha|^m |a_\alpha| < \infty$$

for every $m \in \mathbb{N}^0$, so $(a_\alpha)_{\alpha \in \mathbb{N}_0^n} \in \mathbf{S}_n$. By the isometric property (2.7), we also conclude that ι is injective. To prove that it is onto, pick an arbitrary $(b_\alpha)_{\alpha \in \mathbb{N}_0^n} \in \mathbf{S}_n$. Then, for every $m \in \mathbb{N}$, we have that

$$\varphi_m = \sum_{\alpha \in \mathbb{N}_0^n : |\alpha| \leq m} b_\alpha \phi_\alpha \in \mathcal{S}(\mathbb{R}^n)$$

and it is straightforward to verify (*exercise*) for $m_1 \leq m_2$ that

$$\|\varphi_{m_1} - \varphi_{m_2}\|_\beta^2 = \sum_{\alpha \in \mathbb{N}_0^n : m_1 < |\alpha| \leq m_2} (\alpha + 1)^{2\beta} |b_\alpha|^2 \leq \sum_{\alpha \in \mathbb{N}_0^n : m_1 < |\alpha|} (\alpha + 1)^{2\beta} |b_\alpha|^2 \rightarrow 0$$

as $m_1 \rightarrow \infty$. Thus, $(\varphi_m)_{m \in \mathbb{N}}$ is a Cauchy sequence that has a limit $\varphi \in \mathcal{S}(\mathbb{R}^n)$. The convergence $\lim_{m \rightarrow \infty} \varphi_m = \varphi$ in $\mathcal{S}(\mathbb{R}^n)$ implies the convergence in $L^2(\mathbb{R}^n)$, too, so that

$$\varphi = \sum_{\alpha \in \mathbb{N}_0^n} b_\alpha \phi_\alpha,$$

that is, $\iota(\varphi) = (b_\alpha)_{\alpha \in \mathbb{N}_0^n} \in \mathbf{S}_n$. This shows that ι is bijective. The property (2.7) implies that ι is a homeomorphism, by definition of the topologies on $\mathcal{S}(\mathbb{R}^n)$ and $\mathbf{S}_n$. $\square$

The space $\mathcal{S}'(\mathbb{R}^n)$ of tempered distributions admits a sequence representation as well. Recall that we view $\mathcal{S}'(\mathbb{R}^n)$ as a locally convex space equipped with the weak* topology. To formulate the sequence representation precisely, define the linear space $\mathbf{S}'_n$ by

$$\mathbf{S}'_n = \left\{ a = (a_\alpha)_{\alpha \in \mathbb{N}_0^n} : \exists \beta \in \mathbb{N}_0^n \text{ so that } \sup_{\alpha \in \mathbb{N}_0^n} \frac{|a_\alpha|}{(\alpha + 1)^\beta} < \infty \right\}.$$

Theorem 2.14 (**N**-representation of $\mathcal{S}'(\mathbb{R}^n)$). *The map $\iota' : \mathcal{S}'(\mathbb{R}^n) \rightarrow \mathbf{S}'_n$, defined by*

$$\mathcal{S}'(\mathbb{R}^n) \ni \Lambda \mapsto \iota'(\Lambda) = (\Lambda\phi_\alpha)_{\alpha \in \mathbb{N}_0^n} \in \mathbf{S}'_n$$

is a linear isomorphism. Furthermore, for every $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$, we have that

$$\Lambda = \lim_{m \rightarrow \infty} \sum_{\alpha \in \mathbb{N}_0^n : |\alpha| \leq m} (\Lambda\phi_\alpha)\phi_\alpha,$$

interpreted symbolically as the Hermite expansion $\Lambda = \sum_{\alpha \in \mathbb{N}_0^n} (\Lambda\phi_\alpha)\phi_\alpha$ in $\mathcal{S}'(\mathbb{R}^n)$.

Proof. Let $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ and define $a_\alpha = \Lambda\phi_\alpha$, for all $\alpha \in \mathbb{N}_0^n$. From Prop. 2.5, we infer that for some $\beta \in \mathbb{N}_0^n$ (s.t. $|\beta_j| \in \mathbb{N}_0$ are sufficiently large for all $j \in \{1, \dots, n\}$) we have

$$|a_\alpha| = |\Lambda\phi_\alpha| \leq C \|\phi_\alpha\|_\beta = C(\alpha + 1)^\beta, \quad \forall \alpha \in \mathbb{N}_0^n.$$

Hence, $\iota'(\Lambda) \in \mathbf{S}'_n$. The map is clearly linear and it is injective, because $\Lambda\phi_\alpha = \Lambda'\phi_\alpha$ for all $\alpha \in \mathbb{N}_0^n$ and $\Lambda, \Lambda' \in \mathcal{S}'(\mathbb{R}^n)$ implies that

$$\Lambda\varphi = \sum_{\alpha \in \mathbb{N}_0^n} \langle \phi_\alpha, \varphi \rangle \Lambda\phi_\alpha = \Lambda'\varphi,$$

where we used the fact that $\lim_{m \rightarrow \infty} \sum_{\alpha \in \mathbb{N}_0^n : |\alpha| \leq m} \langle \phi_\alpha, \varphi \rangle \phi_\alpha = \varphi$, $\forall \varphi \in \mathcal{S}(\mathbb{R}^n)$ (*exercise*). Furthermore, ι' is onto. For let $(b_\alpha)_{\alpha \in \mathbb{N}_0^n} \in \mathbf{S}'_n$ and define the linear map $L : \mathbf{S}_n \rightarrow \mathbb{K}$ by

$$L(a_\alpha)_{\alpha \in \mathbb{N}_0^n} = \sum_{\alpha \in \mathbb{N}_0^n} b_\alpha a_\alpha.$$

Then, an application of Cauchy-Schwarz implies

$$|L(a_\alpha)_{\alpha \in \mathbb{N}_0^n}| \leq \sum_{\alpha \in \mathbb{N}_0^n} |b_\alpha a_\alpha| \leq C \sum_{\alpha \in \mathbb{N}_0^n} (\alpha + 1)^\beta |a_\alpha| \leq C \|(a_\alpha)_{\alpha \in \mathbb{N}_0^n}\|_{\beta'},$$

for suitable $\beta, \beta' \in \mathbb{N}_0^n$ so that $L : \mathbf{S}_n \rightarrow \mathbb{K}$ defines in fact a continuous, linear functional on $\mathbf{S}_n$. By Theorem 2.13 (and using the notation from its proof), this yields a continuous, linear functional $\Lambda = L \circ \iota \in \mathcal{S}'(\mathbb{R}^n)$ which is represented by

$$\Lambda\varphi = L(\langle \phi_\alpha, \varphi \rangle)_{\alpha \in \mathbb{N}_0^n} = \sum_{\alpha \in \mathbb{N}_0^n} b_\alpha \langle \phi_\alpha, \varphi \rangle.$$

In particular, $\Lambda\phi_\alpha = b_\alpha, \forall \alpha \in \mathbb{N}_0^n$ so that $\iota'(\Lambda) = (b_\alpha)_{\alpha \in \mathbb{N}_0^n}$.

Finally, the Hermite expansion of $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ follows from the fact that

$$\left(\sum_{\alpha \in \mathbb{N}_0^n : |\alpha| \leq m} (\Lambda\phi_\alpha)\phi_\alpha \right)(\varphi) = \sum_{\alpha \in \mathbb{N}_0^n : |\alpha| \leq m} \langle \phi_\alpha, \varphi \rangle \Lambda\phi_\alpha \rightarrow \sum_{\alpha \in \mathbb{N}_0^n} \langle \phi_\alpha, \varphi \rangle \Lambda\phi_\alpha = \Lambda\varphi$$

as $m \rightarrow \infty$, for every $\varphi \in \mathcal{S}(\mathbb{R}^n)$. □

Based on Theorems 2.13 and 2.14, we can readily draw a large number of interesting consequences. An immediate consequence deals with the separability of $\mathcal{S}(\mathbb{R}^n)$ and $\mathcal{S}'(\mathbb{R}^n)$ (recall that a topological space $\mathcal{X}$ is separable if it contains a countable, dense set, i.e. every non-empty open neighborhood in $\mathcal{X}$ contains elements of that sequence).

Corollary 2.2. *$\mathcal{S}(\mathbb{R}^n)$ is dense in $\mathcal{S}'(\mathbb{R}^n)$. Both $\mathcal{S}(\mathbb{R}^n), \mathcal{S}'(\mathbb{R}^n)$ are separable.*

Proof. The density of $\mathcal{S}(\mathbb{R}^n)$ in $\mathcal{S}'(\mathbb{R}^n)$ follows immediately from the fact that

$$\mathcal{S}(\mathbb{R}^n) \ni \sum_{\alpha \in \mathbb{N}_0^n : |\alpha| \leq m} (\Lambda \phi_\alpha) \phi_\alpha \rightarrow \Lambda \in \mathcal{S}'(\mathbb{R}^n)$$

as $m \rightarrow \infty$, for every $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$. A slight modification of the argument also shows that

$$\Sigma = \left\{ \sum_{\alpha \in \mathbb{N}_0^n} b_\alpha \phi_\alpha : b_\alpha \in \mathbb{Q} + i\mathbb{Q} \wedge \exists m \in \mathbb{N} : b_\alpha = 0 \text{ if } |\alpha| > m \right\} \subset \mathcal{S}(\mathbb{R}^n)$$

is dense in $\mathcal{S}'(\mathbb{R}^n)$. Since Σ is countable, this proves the separability of $\mathcal{S}'(\mathbb{R}^n)$. The separability of $\mathcal{S}(\mathbb{R}^n)$ follows by observing that Σ is in fact also dense in $\mathcal{S}(\mathbb{R}^n)$. $\square$

The next result describes the regularity of tempered distributions. It is a (global) analogue of (the local) Theorem 2.12.

Theorem 2.15. *Let $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$. Then, there exists a polynomially bounded, continuous function $f \in C(\mathbb{R}^n)$ and some $\alpha \in \mathbb{N}_0^n$ such that $\Lambda = \partial^\alpha f$, that is*

$$\Lambda \varphi = (-1)^{|\alpha|} \int_{\mathbb{R}^n} dx f(x) (\partial^\alpha \varphi)(x), \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Proof. Let $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ with Hermite expansion $\Lambda = \sum_{\alpha \in \mathbb{N}_0^n} (\Lambda \phi_\alpha) \phi_\alpha$ such that

$$|\Lambda \phi_\alpha| \leq C(\alpha + 1)^\beta, \quad \forall \alpha \in \mathbb{N}_0^n,$$

for some $\beta \in \mathbb{N}_0^n$. By Problem 2.10, we find some $\gamma \in \mathbb{N}_0^n$ such that

$$|\phi_\alpha|_{0,0} = \sup_{x \in \mathbb{R}^n} |\phi_\alpha(x)| \leq C(\alpha + 1)^\gamma, \quad \forall \alpha \in \mathbb{N}_0^n.$$

Now, set $a_\alpha = (\alpha + 1)^{-\kappa} (\Lambda \phi_\alpha)$, $\forall \alpha \in \mathbb{N}_0^n$, where $\kappa = (\beta_1 + \gamma_1 + 2, \dots, \beta_n + \gamma_n + 2) \in \mathbb{N}_0^n$. Then, $F = \sum_{\alpha \in \mathbb{N}_0^n} a_\alpha \phi_\alpha \in C_b(\mathbb{R}^n)$ by uniform convergence that follows from the bound

$$\sum_{\alpha \in \mathbb{N}_0^n} |a_\alpha| \sup_{x \in \mathbb{R}^n} |\phi_\alpha(x)| \leq C \sum_{\alpha \in \mathbb{N}_0^n} (\alpha + 1)^{\gamma + \beta - \kappa} \leq C \sum_{\alpha \in \mathbb{N}_0^n} \prod_{j=1}^n \frac{1}{(\alpha_j + 1)^2} < \infty.$$

By definition of $(a_\alpha)_{\alpha \in \mathbb{N}_0^n}$, we infer that

$$\Lambda(\varphi) = \sum_{\alpha \in \mathbb{N}_0^n} \langle \phi_\alpha, \varphi \rangle (\Lambda \phi_\alpha) = \sum_{\alpha \in \mathbb{N}_0^n} (\alpha + 1)^\kappa \langle \phi_\alpha, \varphi \rangle a_\alpha = F((\mathbf{N} + 1)^\kappa \varphi), \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^n).$$

In other words, $\Lambda = (\mathbf{N} + 1)^\kappa F = \prod_{j=1}^n (\mathbf{N}_j + 1)^{\kappa_j} F$ in the sense of $\mathcal{S}'(\mathbb{R}^n)$. Expanding

$$(\mathbf{N} + 1)^\kappa = \prod_{j=1}^n 2^{-\kappa_j} (-\partial_{x_j}^2 + x_j^2 + 1)^{\kappa_j} = \sum_{\delta, \varepsilon \in \mathbb{N}_0^n: \delta_j + \varepsilon_j \leq 2\kappa_j, \forall j} c_{\delta\varepsilon} x^\delta \partial^\varepsilon$$

as a linear operator on $\mathcal{S}(\mathbb{R}^n)$ for a suitable sequence $(c_{\delta\varepsilon})_{\delta, \varepsilon \in \mathbb{N}_0^n}$, using that $x^\delta F \in C(\mathbb{R}^n)$ is polynomially bounded for every $\delta \in \mathbb{N}_0^n$ (since $F \in C_b(\mathbb{R}^n)$) and, finally, using that every polynomially bounded $\psi \in C(\mathbb{R}^n)$ can be written as a k -th partial derivative

$$\psi(x) = (\partial_{x_j}^k \zeta)(x), \text{ where } \zeta(x) = \int_0^{x_j} dy_1 \dots \int_0^{y_{k-1}} dy_k \psi(x_1, \dots, x_{j-1}, y_k, x_{j+1}, \dots, x_n),$$

with $\zeta \in C(\mathbb{R}^n)$ being polynomially bounded as well, we readily conclude (*exercise*) that

$$\Lambda = (\mathbf{N} + 1)^\kappa F = \sum_{\delta, \varepsilon \in \mathbb{N}_0^n: \delta_j + \varepsilon_j \leq 2\kappa_j, \forall j} (-1)^{|\varepsilon|} c_{\delta\varepsilon} \partial^\varepsilon (x^\delta F) = \partial^\alpha f$$

for some polynomially bounded $f \in C(\mathbb{R}^n)$ and $\alpha = 2\kappa \in \mathbb{N}_0^n$. $\square$

Another interesting consequence deals with multilinear functionals on $\mathcal{S}(\mathbb{R}^n)$, which occur below in the form of the Wightman functions of a quantum field theory.

Theorem 2.16 (Nuclear Theorem for $\mathcal{S}'$). *Let $L : \mathcal{S}(\mathbb{R}^{n_1}) \times \dots \times \mathcal{S}(\mathbb{R}^{n_k}) \rightarrow \mathbb{K}$ be a separately continuous, multilinear functional. Then, there exists a unique tempered distribution $\Lambda \in \mathcal{S}'(\mathbb{R}^{n_1 + \dots + n_k})$ such that*

$$\Lambda(\varphi_1 \otimes \dots \otimes \varphi_k) = L(\varphi_1, \dots, \varphi_k), \quad \forall \varphi_j \in \mathcal{S}(\mathbb{R}^{n_j}), j \in \{1, \dots, k\}.$$

Proof. For simplicity of notation, consider the bilinear case. The general case follows along the same lines and is left as an *exercise*.

By Prop. 2.1, $L : \mathcal{S}(\mathbb{R}^n) \times \mathcal{S}(\mathbb{R}^m) \rightarrow \mathbb{K}$ is jointly continuous so that

$$|L(\phi_\alpha, \phi_\beta)| \leq C \|\phi_\alpha\|_\gamma \|\phi_\beta\|_\delta = C \prod_{j=1}^n (\alpha_j + 1)^{\gamma_j} \prod_{k=1}^m (\alpha_k + 1)^{\gamma_k} = C((\alpha, \beta) + 1)^{(\gamma, \delta)}$$

for all $(\alpha, \beta) \in \mathbb{N}_0^{n+m}$ and some fixed $(\gamma, \delta) \in \mathbb{N}_0^{n+m}$. By the characterization of $\mathcal{S}'(\mathbb{R}^{n+m})$, we conclude that $\Lambda = \sum_{(\alpha, \beta) \in \mathbb{N}_0^{n+m}} L(\phi_\alpha, \phi_\beta) \phi_\alpha \otimes \phi_\beta \in \mathcal{S}'(\mathbb{R}^{n+m})$ so that

$$\Lambda(\varphi \otimes \psi) = \sum_{\alpha \in \mathbb{N}_0^n, \beta \in \mathbb{N}_0^m} \langle \phi_\alpha, \varphi \rangle \langle \phi_\beta, \psi \rangle L(\phi_\alpha, \phi_\beta) = L(\varphi, \psi), \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^n), \psi \in \mathcal{S}(\mathbb{R}^m).$$

Uniqueness follows from the fact that the Hermite coefficients determine Λ uniquely. $\square$

Note that Theorem 2.16 is typically wrong in other common function space settings. Consider e.g. the functional $L^2(\mathbb{R}^n) \times L^2(\mathbb{R}^n) \ni (\varphi, \psi) \mapsto \langle \varphi, \psi \rangle \in \mathbb{K}$. Its kernel equals $\Lambda \in \mathcal{S}'(\mathbb{R}^{2n})$ with $\Lambda(x, y) = \delta(x - y)$, in particular $\Lambda \notin (L^2(\mathbb{R}^{2n}))^* \simeq L^2(\mathbb{R}^{2n})$. The theorem is, however, also true in $\mathcal{D}'(\mathbb{R}^n)$ which we state without proof.

Theorem 2.17 (Nuclear Theorem for $\mathcal{D}'$). *Let $L : \mathcal{D}(\mathbb{R}^{n_1}) \times \dots \times \mathcal{D}(\mathbb{R}^{n_k}) \rightarrow \mathbb{K}$ be a separately continuous, multilinear functional. Then, there exists a unique distribution $\Lambda \in \mathcal{D}'(\mathbb{R}^{n_1+\dots+n_k})$ such that*

$$\Lambda(\varphi_1 \otimes \dots \otimes \varphi_k) = L(\varphi_1, \dots, \varphi_k), \quad \forall \varphi_j \in \mathcal{D}(\mathbb{R}^{n_j}), j \in \{1, \dots, k\}.$$

Proof. We leave it as an *exercise* to prove that L is jointly continuous. Then, one can proceed similarly as in the proof of Theorem 2.16 after setting up an analogue of the $\mathbf{N}$ -representation, based on the Fourier series expansion. In dimension $n = 1$, this is outlined in [21, Problems 59 & 60]. Alternatively, see e.g. [6, Chapter 1.1]. $\square$

Finally, let us reconsider the operations on $\mathcal{S}'(\mathbb{R}^n)$ defined in Examples 2.4, 2.5, 2.7 and 2.8. A consequence of Corollary 2.2 is that all these operators, viewed as maps from $\mathcal{S}'(\mathbb{R}^n)$ to itself are continuous. This follows from the following general considerations. Recall that $\mathcal{S}(\mathbb{R}^n)$ embeds continuously into $\mathcal{S}'(\mathbb{R}^n)$. Let's denote the linear, injective embedding by $\iota : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}'(\mathbb{R}^n)$. The continuity follows from

$$|\iota(\varphi)(\psi)| = \left| \int_{\mathbb{R}^n} dx \varphi(x) \psi(x) \right| \leq |\psi|_{0,0} \|\varphi\|_1 \leq C |\psi|_{0,0} \|\varphi\|_\alpha$$

for suitable $\alpha \in \mathbb{N}_0^n$, the definition of the topology on $\mathcal{S}'(\mathbb{R}^n)$ and Prop. 2.5. Suppose that $S : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n)$ is a linear, continuous map and define $S' : \mathcal{S}'(\mathbb{R}^n) \rightarrow \mathcal{S}'(\mathbb{R}^n)$ by

$$S'(\Lambda)(\varphi) = \Lambda(S\varphi), \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^n).$$

Then, a similar argument as before implies that S' is continuous, too (*exercise*).

Now, suppose that $T : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n)$ is linear and continuous so that $\iota \circ T \circ \iota^{-1} : \iota(\mathcal{S}(\mathbb{R}^n)) \rightarrow \mathcal{S}'(\mathbb{R}^n)$ is linear and continuous as well. Since $\iota(\mathcal{S}(\mathbb{R}^n))$ is dense in $\mathcal{S}'(\mathbb{R}^n)$, $\iota \circ T \circ \iota^{-1}$ admits at most one continuous extension $T' : \mathcal{S}'(\mathbb{R}^n) \rightarrow \mathcal{S}'(\mathbb{R}^n)$. In particular, if there exists some continuous, linear $S : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n)$ as above that satisfies

$$S'_{|\iota(\mathcal{S}(\mathbb{R}^n))} = \iota \circ T \circ \iota^{-1} \iff \int_{\mathbb{R}^n} dx \varphi(x) (S\psi)(x) = \int_{\mathbb{R}^n} dx (T\varphi)(x) \psi(x), \quad \forall \varphi, \psi \in \mathcal{S}(\mathbb{R}^n),$$

then $T' = S' : \mathcal{S}'(\mathbb{R}^n) \rightarrow \mathcal{S}'(\mathbb{R}^n)$ defines this unique, continuous extension of $\iota \circ T \circ \iota^{-1}$. We leave it as an instructive *exercise* to verify that the operations from Examples 2.4, 2.5, 2.7 and 2.8 are defined exactly in such a way to meet the previous criterion. As a consequence, all the actions defined in these examples are continuous on $\mathcal{S}'(\mathbb{R}^n)$. In case of the Fourier transform, we record the following corollary.

Corollary 2.3. *The maps $\widehat{\cdot} : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n)$ and $\widetilde{\cdot} : \mathcal{S}(\mathbb{R}^n) \rightarrow \mathcal{S}(\mathbb{R}^n)$ extend to linear, continuous maps $\widehat{\cdot} : \mathcal{S}'(\mathbb{R}^n) \rightarrow \mathcal{S}'(\mathbb{R}^n)$ and, respectively, $\widetilde{\cdot} : \mathcal{S}'(\mathbb{R}^n) \rightarrow \mathcal{S}'(\mathbb{R}^n)$ such that*

$$\widetilde{(\widehat{\Lambda})} = \widehat{(\widetilde{\Lambda})} = \Lambda, \quad \forall \Lambda \in \mathcal{S}'(\mathbb{R}^n).$$

Proof. Continuity of $\widehat{\cdot}$ and $\widetilde{\cdot}$ follows from the preceding remarks and the inversion property is true on $\mathcal{S}(\mathbb{R}^n) \hookrightarrow \mathcal{S}'(\mathbb{R}^n)$ which is dense in $\mathcal{S}'(\mathbb{R}^n)$. $\square$

2.3 The Lorentz and Poincaré Groups

In Section 1.3, the Lorentz and Poincaré groups emerged as the special relativistic analogues of the orthogonal and, respectively, Euclidean groups in Newton mechanics. In this section, we define these groups precisely and we collect important properties of the Lorentz group, following [16]. Most importantly, we show that the special linear group $\mathrm{SL}(2, \mathbb{C})$ covers $\mathcal{L}_+^\uparrow$ by generalizing the proof of Prop. 1.1 that relates $\mathrm{SU}(2)$ with $\mathrm{SO}(3)$. In Section 3.2, we continue our discussion by analyzing the unitary representations of the Poincaré group (or, more precisely, of the inhomogeneous group $\mathbb{R}^4 \rtimes \mathrm{SL}(2, \mathbb{C})$).

We recall that $\mathrm{O}(1, 3)$ is defined as the group of isometries on Minkowski space, i.e.

$$\mathrm{O}(1, 3) = \{L \in \mathbb{R}^{4 \times 4} : \eta(Lx, Ly) = \eta(x, y) \forall x, y \in \mathbb{R}^4\}, \quad (2.8)$$

where η is the Minkowski metric. With respect to the standard basis of $\mathbb{R}^4$, η has the form (1.47). In matrix language, the invariance of η under $L \in \mathrm{O}(1, 3)$ is equivalent to

$$L^T \eta L = \eta. \quad (2.9)$$

It is clear that $\mathbf{1}_{\mathbb{R}^4} \in \mathrm{O}(1, 3)$ and since $\eta^2 = \eta^T \eta = \eta \eta^T = \mathbf{1}_{\mathbb{R}^4} = (-\eta)^2$, we also have that $\pm \eta \in \mathrm{O}(1, 3)$. By non-degeneracy of η , every $L \in \mathrm{O}(1, 3)$ is injective and hence invertible. The (left inverse and hence the) inverse L^{-1} of L is equal to

$$L^{-1} = \eta L^T \eta \in \mathbb{R}^{4 \times 4},$$

which immediately implies that both $L^T \in \mathrm{O}(1, 3)$ and $L^{-1} \in \mathrm{O}(1, 3)$. Equipped with standard matrix multiplication, $\mathrm{O}(1, 3)$ is a group, because

$$(L_1 L_2)^T \eta (L_1 L_2) = L_2^T L_1^T \eta L_1 L_2 = L_2^T \eta L_2 = \eta, \quad \forall L_1, L_2 \in \mathrm{O}(1, 3).$$

Clearly, matrix multiplication $\mathrm{O}(1, 3) \times \mathrm{O}(1, 3) \ni (L_1, L_2) \mapsto L_1 L_2 \in \mathrm{O}(1, 3)$ and taking the inverse $\mathrm{O}(1, 3) \ni L \mapsto L^{-1} = \eta L^T \eta \in \mathrm{O}(1, 3)$ define continuous maps on $\mathrm{O}(1, 3)$ with respect to the subspace topology inherited from $\mathbb{R}^{4 \times 4}$ (*exercise*). Here, we view $\mathbb{R}^{4 \times 4}$ as a real Hilbert space with the matrix inner product $\langle A, B \rangle = \mathrm{tr}(A^T B)$. Notice that the induced norm corresponds to the standard Euclidean norm if we identify $\mathbb{R}^{4 \times 4} \simeq \mathbb{R}^{16}$.

It turns out that both multiplication $\mathrm{O}(1, 3) \times \mathrm{O}(1, 3) \ni (L_1, L_2) \mapsto L_1 L_2 \in \mathrm{O}(1, 3)$ and taking the inverse $\mathrm{O}(1, 3) \ni L \mapsto L^{-1} = \eta L^T \eta \in \mathrm{O}(1, 3)$ are indeed smooth when we view $\mathrm{O}(1, 3)$ as a Lie group, which is made precise in Lemma 2.4 below. To motivate part of the statement of Lemma 2.4, observe that if $e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} \in \mathrm{O}(1, 3)$ for some $A \in \mathbb{R}^{4 \times 4}$, then for small $A \in \mathbb{R}^{4 \times 4}$, we have that

$$\eta = (e^A)^T \eta e^A = \eta + A^T \eta + \eta A + O(A^2) \implies A^T \approx -\eta A \eta.$$

Lemma 2.4. *$\mathrm{O}(1, 3)$ is a six-dimensional real Lie group smoothly embedded into $\mathbb{R}^{4 \times 4}$ and with a real analytic atlas. Its Lie algebra is isomorphic to $\mathfrak{so}(1, 3)$, defined by*

$$\mathfrak{so}(1, 3) = \{A \in \mathbb{R}^{4 \times 4} : A^T = -\eta A \eta\}. \quad (2.10)$$

Proof. We use the matrix exponential and its inverse to construct an analytic atlas. Before we make this precise, notice first of all that $\mathfrak{so}(1, 3)$ defines in fact a Lie algebra with Lie bracket given by the usual matrix commutator. This follows from

$$[A_1, A_2]^T = [A_2^T, A_1^T] = [\eta A_2 \eta, \eta A_1 \eta] = -\eta [A_1, A_2] \eta, \quad \forall A_1, A_2 \in \mathfrak{so}(1, 3).$$

Setting $\mathbf{a}_0 = (a_{01}, a_{02}, a_{03})$, $\mathbf{a}_\cdot = (a_{10}, a_{20}, a_{30}) \in \mathbb{R}^3$, it follows from the constraint

$$A^T = \begin{pmatrix} a_{00} & \mathbf{a}_0^T \\ \mathbf{a}_\cdot & (a_{ij})_{1 \leq i, j \leq 3}^T \end{pmatrix} = \begin{pmatrix} a_{00} & \mathbf{a}_0^T \\ \mathbf{a}_\cdot & -(a_{ij})_{1 \leq i, j \leq 3} \end{pmatrix} = -\eta A \eta,$$

that a general element $A \in \mathfrak{so}(1, 3)$ takes the form

$$A = \begin{pmatrix} 0 & b_1 & b_2 & b_3 \\ b_1 & 0 & -x_3 & x_2 \\ b_2 & x_3 & 0 & -x_1 \\ b_3 & -x_2 & x_1 & 0 \end{pmatrix} = \sum_{k=1}^3 (b_k B_k + x_k X_k),$$

for some $(b_1, b_2, b_3, x_1, x_2, x_3) \in \mathbb{R}^6$. Here, the generators B_k are obtained by setting $b_k = 1$ and all other entries to zero while the X_k are obtained by setting $x_k = 1$ and all other entries to zero, for $k \in \{1, 2, 3\}$. This shows that $\mathfrak{so}(1, 3) \simeq \mathbb{R}^6$ is six-dimensional.

Next, let's recall some basic properties of the matrix exponential $\exp : \mathbb{R}^{4 \times 4} \rightarrow \mathbb{R}^{4 \times 4}$. It is clearly an analytic function with $\exp(0) = \mathbf{1}_{\mathbb{R}^4} \in \mathrm{O}(1, 3) \subset \mathbb{R}^{4 \times 4}$. Observe that

$$(e^A)^T \eta e^A = \sum_{k=0}^{\infty} \frac{(A^T)^k}{k!} \eta e^A = \sum_{k=0}^{\infty} \frac{(-\eta A \eta)^k}{k!} \eta e^A = \eta e^{-A} e^A = \eta, \quad \forall A \in \mathfrak{so}(1, 3).$$

Conversely, it is a standard fact that for $\|X\| < 1$, the map defined by

$$\log(\mathbf{1}_{\mathbb{R}^4} + X) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{X^k}{k}$$

is analytic, too, and satisfies $e^{\log(\mathbf{1}_{\mathbb{R}^4} + X)} = \mathbf{1}_{\mathbb{R}^4} + X$ for all $X \in \mathbb{R}^{4 \times 4}$ with $\|X\| < 1$ as well as $\log e^A = A$ for all A with $\|A\| < \log 2$. Furthermore, notice that for every $L \in \mathrm{O}(1, 3)$ with $\|\mathbf{1}_{\mathbb{R}^4} - L\| < 1$, it holds true that

$$L^T \eta = \eta L^{-1} \iff L^T = e^{\log L^T} = \eta e^{-\log L} \eta \iff (\log L)^T = \log(L^T) = -\eta \log(L) \eta.$$

In other words, $\log(L) \in \mathfrak{so}(1, 3)$ for every $L \in \mathrm{O}(1, 3)$ with $\|\mathbf{1}_{\mathbb{R}^4} - L\| < 1$.

To construct an analytic atlas of $\mathrm{O}(1, 3)$, we can now proceed as follows. For an arbitrary $L \in \mathrm{O}(1, 3)$, we pick an open neighborhood

$$U_{L, \epsilon} = \{X \in \mathbb{R}^{4 \times 4} : \|\mathbf{1}_{\mathbb{R}^4} - L^{-1}X\| < \epsilon\} = \{LY \in \mathbb{R}^{4 \times 4} : \|\mathbf{1}_{\mathbb{R}^4} - Y\| < \epsilon\} \subset \mathbb{R}^{4 \times 4}$$

at $L \in \mathrm{O}(1, 3)$ and we define a local chart $\varphi_{L, \epsilon} : U_{L, \epsilon} \rightarrow \mathbb{R}^{4 \times 4}$ with $\varphi_{L, \epsilon}(L) = 0$ by

$$\varphi_{L, \epsilon}(X) = \log(\mathbf{1}_{\mathbb{R}^4} + (L^{-1}X - \mathbf{1}_{\mathbb{R}^4})) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(L^{-1}X - \mathbf{1}_{\mathbb{R}^4})^k}{k}.$$

As composition of analytic functions, $\varphi_{L,\epsilon}$ is an analytic function and we have that

$$\varphi_{L,\epsilon}(U_{L,\epsilon} \cap \mathrm{O}(1,3)) = \varphi_{L,\epsilon}(U_{L,\epsilon}) \cap \mathfrak{so}(1,3)$$

is an open slice in $\mathfrak{so}(1,3) \subset \mathbb{R}^{4 \times 4}$. Using the exponential map as inverse, it is clear that the transition functions $\varphi_{\epsilon,L_1}^{-1} \circ \varphi_{\epsilon,L_2}$ for $L_1, L_2 \in \mathrm{O}(1,3)$ are analytic (in the usual sense as maps from subsets of $\mathfrak{so}(1,3) \simeq \mathbb{R}^6$ to $\mathfrak{so}(1,3)$) and that the corresponding smooth structure ensures smoothness of the embedding $\mathrm{O}(1,3) \hookrightarrow \mathbb{R}^{4 \times 4}$ (*exercise*). In other words, $\mathrm{O}(1,3)$ is a smoothly embedded six-dimensional submanifold of $\mathbb{R}^{4 \times 4}$. As topological space with subspace topology inherited from $\mathbb{R}^{4 \times 4}$, it is in particular a six-dimensional topological manifold and the smooth structure that turns the embedding $\mathrm{O}(1,3) \hookrightarrow \mathbb{R}^{4 \times 4}$ into a smooth map is unique (recall e.g. [14, Theorem 8.2]).

To conclude the lemma, it remains to determine the Lie algebra of $\mathrm{O}(1,3)$, which is isomorphic to its tangent space $T_{\mathbf{1}_{\mathbb{R}^4}}\mathrm{O}(1,3)$ at $\mathbf{1}_{\mathbb{R}^4}$. With the same notation as above, consider for every $k \in \{1,2,3\}$ the smooth paths

$$(-\epsilon, \epsilon) \ni t \mapsto \gamma_{B_k}(t) = e^{tB_k} \in \mathrm{O}(1,3), \quad (-\epsilon, \epsilon) \ni t \mapsto \gamma_{X_k}(t) = e^{tX_k} \in \mathrm{O}(1,3).$$

Then, clearly $\frac{d}{dt}(\gamma_{B_k})|_{t=0} = B_k$ and $\frac{d}{dt}(\gamma_{X_k})|_{t=0} = X_k$. This shows that $T_{\mathbf{1}_{\mathbb{R}^4}}\mathrm{O}(1,3) \simeq \mathfrak{so}(1,3)$ and concludes the proof of the lemma. $\square$

The proof of Lemma 2.4 can be adapted to many other matrix groups including e.g. $\mathrm{O}(3)$ and $\mathrm{SO}(3)$ with Lie algebra $\mathfrak{so}(3)$ (alternatively, one can also apply general results on Lie groups such as the fact that closed subgroups of Lie groups are automatically Lie subgroups, see e.g. [14, Chapter 20]). In the sequel, we therefore typically omit the detailed verification that a matrix group is a Lie group and simply refer to Lemma 2.4. Our next goal is to find a parametrization of $\mathrm{O}(1,3)$, analogous to the parametrization of $\mathrm{SO}(3)$ below. This is used to prove that $\mathrm{SL}(2, \mathbb{C})$ is a double cover of $\mathcal{L}_+^\uparrow$.

Problem 2.11. Recall the notation from Prop. 1.1 and prove the following.

a) Extend the identity (1.35) from Problem 1.5 by showing that

$$R(e^{\frac{i}{2}\omega\sigma_j}) = e^{\omega X_j}, \quad \forall \omega \in [0, 2\pi), j \in \{1, 2, 3\}. \quad (2.11)$$

b) Verify explicitly that for every $A \in \mathrm{SU}(2)$, it holds true that

$$A = e^{\frac{i}{2}\varphi\sigma_3} e^{\frac{i}{2}\theta\sigma_1} e^{\frac{i}{2}\psi\sigma_3}$$

for suitable $\varphi \in [0, 2\pi), \theta \in [0, \pi], \psi \in [-2\pi, 2\pi]$, the so called Euler angles.

c) Use the previous parts and Prop. 1.1 to show that every $R \in \mathrm{SO}(3)$ takes the form

$$R = \begin{pmatrix} c_\varphi c_\psi - s_\varphi s_\psi c_\theta & -c_\varphi s_\psi - s_\varphi c_\psi c_\theta & s_\varphi s_\theta \\ s_\varphi c_\psi + c_\varphi s_\psi c_\theta & -s_\varphi s_\psi + c_\varphi c_\psi c_\theta & -c_\varphi s_\theta \\ s_\psi s_\theta & c_\psi s_\theta & c_\theta \end{pmatrix} = R(\varphi, \theta, \psi),$$

for suitable $\varphi \in [0, 2\pi), \theta \in [0, \pi], \psi \in [-2\pi, 2\pi]$, where $c_\omega = \cos(\omega), s_\omega = \sin(\omega)$.

The parametrization of $O(1, 3)$ can be reduced to one of $\mathcal{L}_+^\uparrow$, which is defined by

$$\mathcal{L}_+^\uparrow = \{L \in O(1, 3) : \det L = 1, L_{00} > 0\}.$$

Recall that the computations after (1.50) actually show that $\mathcal{L}_+^\uparrow$ is a subgroup, that

$$\mathcal{L}_+^\uparrow = \{L \in O(1, 3) : \det L = 1, L_{00} \geq 1\} \quad (2.12)$$

and that $SO(3) \hookrightarrow \mathcal{L}_+^\uparrow$ embeds continuously into $\mathcal{L}_+^\uparrow$ through the map

$$SO(3) \ni R \mapsto \iota(R) = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix} \in \mathcal{L}_+^\uparrow.$$

The rotations are generated by the elements $X_1, X_2, X_3 \in \mathfrak{so}(1, 3)$, as defined in the proof of Lemma 2.4. The generators $B_1, B_2, B_3 \in \mathfrak{so}(1, 3)$, on the other hand, are called relativistic boosts and describe coordinate changes from inertial systems that move with constant speed with regards to each other (see [16, Section 8.2] for the details). In the following, by slight abuse of notation, whenever we say $R \in SO(3)$ lies in $O(1, 3)$, we are actually referring to its embedding $\iota(R) \in O(1, 3)$. We also define the sets

$$\begin{aligned} \mathcal{L}_-^\uparrow &= \{L \in O(1, 3) : \det L = -1, L_{00} \geq 1\}, \\ \mathcal{L}_+^\downarrow &= \{L \in O(1, 3) : \det L = 1, L_{00} \leq -1\}, \\ \mathcal{L}_-^\downarrow &= \{L \in O(1, 3) : \det L = -1, L_{00} \leq -1\}. \end{aligned}$$

As shown below, the sets $\mathcal{L}_+^\uparrow, \mathcal{L}_-^\uparrow, \mathcal{L}_+^\downarrow$ and $\mathcal{L}_-^\downarrow$ are the connected components of $O(1, 3)$. The subgroup $\mathcal{L}_+^\uparrow$ is the connected component that contains the identity $\mathbf{1}_{\mathbb{R}^4}$. Notice that the other components are non-empty, because $\eta \in \mathcal{L}_-^\uparrow$, $-\eta \in \mathcal{L}_-^\downarrow$ and $-\mathbf{1}_{\mathbb{R}^4} \in \mathcal{L}_+^\downarrow$.

As a preparation for the parametrization of $O(1, 3)$, we need the following lemma.

Lemma 2.5. *Denote by $\mathcal{P} \in \mathbb{R}^{4 \times 4}$ and $\mathcal{T} \in \mathbb{R}^{4 \times 4}$ space and, respectively, time inversion, defined by $\mathcal{P}(x_0, \mathbf{x}) = (x_0, -\mathbf{x})$ and $\mathcal{T}(x_0, \mathbf{x}) = (-x_0, \mathbf{x})$, for every $x = (x_0, \mathbf{x}) \in \mathbb{R}^4$. The operator $\mathcal{P}$ is commonly referred to as the parity operator on $\mathbb{R}^4$. Then:*

$$i) \quad \mathcal{L}_-^\uparrow = \mathcal{P}\mathcal{L}_+^\uparrow = \mathcal{L}_+^\uparrow\mathcal{P}, \quad \mathcal{L}_+^\downarrow = -\mathcal{L}_+^\uparrow \quad \text{and} \quad \mathcal{L}_-^\downarrow = -\mathcal{P}\mathcal{L}_+^\uparrow = -\mathcal{L}_+^\uparrow\mathcal{P}.$$

$$ii) \quad \text{An element } L \in \mathcal{L}_+^\uparrow \text{ lies in } SO(3) \text{ if and only if } L_{00} = 1.$$

$$iii) \quad \text{Let } e_0 = (1, 0, 0, 0) \in \mathbb{R}^4 \text{ and suppose that } L_1, L_2 \in \mathcal{L}_+^\uparrow. \text{ Then } L_1 e_0 = L_2 e_0 \text{ if and only if there exists a unique element } R \in SO(3) \text{ such that } L_1 = L_2 R.$$

Proof. We start with the proof of *i*). Note that, with regards to the standard basis, $\mathcal{P} = \eta$. Therefore, $(\eta L)_{00} = (L\eta)_{00} = L_{00}$ and $\det(\mathcal{P}L) = \det(L\mathcal{P}) = -\det L$ for every $L \in \mathbb{R}^{4 \times 4}$. Combined with $\mathcal{P}^2 = \mathbf{1}_{\mathbb{R}^4}$, we immediately conclude that $\mathcal{L}_-^\uparrow = \mathcal{P}\mathcal{L}_+^\uparrow = \mathcal{L}_+^\uparrow\mathcal{P}$. The remaining two identities are proved similarly and are left as an *exercise*.

For part *ii*), let $L \in \text{SO}(3)$. Thus, by convention, $L = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}$ for some $R \in \text{SO}(3)$, so that $L_{00} = 1$. Conversely, if $L_{00} = 1$, the computations after (1.50) imply that

$$L_{00}^2 = 1 + \sum_{j=1}^3 L_{j0}^2 = 1 + \sum_{j=1}^3 L_{0j}^2 = 1$$

so that $L = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}$ for some $R \in \mathbb{R}^{3 \times 3}$. Note that $\det R = \det L = 1$ and that $L^T \eta L = \eta$ implies $R^T \mathbf{1}_{\mathbb{R}^3} R = \mathbf{1}_{\mathbb{R}^3}$. This shows that $R \in \text{SO}(3)$.

Finally, for part *iii*), let $L_1, L_2 \in \mathcal{L}_+^\uparrow$. If $L_1 e_0 = L_2 e_0$, then $L_2^{-1} L_1 e_0 = e_0$ and $R = L_2^{-1} L_1 \in \mathcal{L}_+^\uparrow$ satisfies $R_{00} = (R e_0)_0 = 1$ so that $R \in \text{SO}(3)$. Conversely, for $R \in \text{SO}(3) \hookrightarrow \mathcal{L}_+^\uparrow$, we have $R e_0 = e_0$. Thus, $L_1 = L_2 R$ implies that $L_1 e_0 = L_2 e_0$. $\square$

Lemma 2.6. *Let $L \in \mathcal{L}_+^\uparrow$. Then, $L = R_1 e^{bB_3} R_3$ for some $b \in \mathbb{R}$ and $R_1, R_2 \in \text{SO}(3)$.*

Proof. Let $L \in \mathcal{L}_+^\uparrow$ and consider $L(1, 0, 0, 0) = (L_{00}, L_{10}, L_{20}, L_{30}) \in \mathbb{R}^4$. If $L_{00} = 1$, then $L \in \text{SO}(3)$ and we are done. On the other hand, if $L_{00} > 1$, then as recalled previously

$$\sum_{j=1}^3 L_{j0}^2 = L_{00}^2 - 1 = r^2 > 0,$$

where we choose $r = (L_{00}^2 - 1)^{\frac{1}{2}} > 0$. By standard properties of the hyperbolic functions, we find $b > 0$ such that $r = \sinh(b) = -i \sin(ib)$ and $L_{00} = \sqrt{1 + r^2} = \cosh(b) = \cos(ib)$. Then, a straightforward computation (*exercise*) shows that

$$e^{bB_3}(1, 0, 0, 0) = (L_{00}, 0, 0, r).$$

On the other hand, the vector $\mathbf{L} = (L_{10}, L_{20}, L_{30}) \in \mathbb{R}^3$ can be represented in standard spherical coordinates by

$$\mathbf{L} = (r \cos(\varphi) \sin(\theta), r \sin(\theta) \sin(\varphi), r \cos(\theta)),$$

for some $\varphi \in [0, 2\pi), \theta \in [0, \pi]$. In the notation of Problem 2.11 c), this implies (*exercise*)

$$\mathbf{L} = R(\varphi + \frac{\pi}{2}, \theta, 0)(0, 0, r)$$

and hence

$$R(\varphi + \frac{\pi}{2}, \theta, 0) e^{bB_3}(1, 0, 0, 0) = (L_{00}, \mathbf{L}) = L(1, 0, 0, 0).$$

By Lemma 2.5, this proves $L = R_1 e^{bB_3} R_2$ for $b > 0$ as above, $R_1 = R(\varphi + \frac{\pi}{2}, \theta, 0) \in \text{SO}(3)$ and for a unique $R_2 = (R_1 e^{bB_3})^{-1} L \in \text{SO}(3)$. $\square$

Notice that the proof of Lemma 2.6 implies that $\mathcal{L}_+^\uparrow$ is path connected. Indeed, if for suitable Euler angles $L \in \mathcal{L}_+^\uparrow$ is parametrized by

$$L = R(\varphi_1, \theta_1, 0)e^{bB_3}R(\varphi_2, \theta_2, \psi),$$

then the path $[0, 1] \ni t \mapsto R(t\varphi_1, t\theta_1, 0)e^{tbB_3}R(t\varphi_2, t\theta_2, t\psi) \in \mathcal{L}_+^\uparrow$ connects $\mathbf{1}_{\mathbb{R}^4}$ analytically to L . In particular, $\mathcal{L}_+^\uparrow$ is connected and, by Lemma 2.5, the group $O(1, 3)$ splits into the four connected components

$$O(1, 3) = \mathcal{L}_+^\uparrow \cup \mathcal{L}_+^\downarrow \cup \mathcal{L}_-^\uparrow \cup \mathcal{L}_-^\downarrow.$$

In view of the projective unitary representations of $\mathcal{L}_+^\uparrow$, the most important task is to find its universal cover. Generalizing Proposition 1.1, we show that this is equal to the special linear group $SL(2, \mathbb{C})$, which is defined by

$$SL(2, \mathbb{C}) = \{A \in \mathbb{C}^{2 \times 2} : \det A = 1\}. \quad (2.13)$$

Problem 2.12. *Show that $SL(2, \mathbb{C})$ is a six-dimensional real Lie group with Lie algebra $\mathfrak{sl}(2, \mathbb{C}) = \{A \in \mathbb{C}^{2 \times 2} : \text{tr } A = 0\}$. Determine an explicit isomorphism between $\mathfrak{sl}(2, \mathbb{C})$ (as a real Lie algebra) and $\mathfrak{so}(1, 3)$. Based on the polar decomposition $A = U|A|$ for general matrices, where U is a partial isometry and $|A| = \sqrt{A^*A}$ is non-negative, show that $SL(2, \mathbb{C})$ is path connected.*

In addition to the properties discussed in Problem 2.12, let us point out that $SL(2, \mathbb{C})$ is simply connected (see e.g. [8, Chapter 13]). With this in mind, the next proposition justifies to refer to it as the universal covering group of $\mathcal{L}_+^\uparrow$.

Proposition 2.9. *There exists a two-to-one group homomorphism $L : SL(2, \mathbb{C}) \rightarrow \mathcal{L}_+^\uparrow$ which is a local homeomorphism.*

Proof. As mentioned before, we generalize the proof of Proposition 1.1. To this end, recall that the self-adjoint Pauli matrices $\sigma_1, \sigma_2, \sigma_3 \in \mathbb{C}^{2 \times 2}$ are defined by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.14)$$

Setting, moreover, $\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, consider the map

$$\mathbb{R}^4 \ni x \mapsto -x^\mu \sigma_\mu = \sum_{j=1}^3 \sigma_j x_j - \sigma_0 x_0 = \begin{pmatrix} x_3 - x_0 & x_1 + ix_2 \\ x_1 - ix_2 & -x_3 - x_0 \end{pmatrix} = (-x^\mu \sigma_\mu)^* \in \mathbb{C}^{2 \times 2}$$

and notice that for every $x \in \mathbb{R}^4$, it holds true that

$$\det(-x^\mu \sigma_\mu) = x^\mu x_\mu.$$

Now, for every $A \in \mathbb{C}^{2 \times 2}$, the conjugated matrix $A(-x^\mu \sigma_\mu)A^*$ is self-adjoint and the matrices $(\sigma_j)_{j=0}^3$ form an orthogonal basis of $\mathbb{C}^{2 \times 2}$. This implies that

$$A(-x^\mu \sigma_\mu)A^* = -(y_A(x))^\mu \sigma_\mu,$$

where

$$(y_A(x))_\nu = \frac{1}{2} \eta_{\nu\nu} \text{tr } \sigma_\nu A(x^\mu \sigma_\mu)A^* \in \mathbb{R}, \quad \forall \nu \in \{0, 1, 2, 3\}.$$

In other words, $y_A(x) = L(A)x$ for some $L(A) \in \mathbb{R}^{4 \times 4}$. If $A \in \text{SL}(2, \mathbb{C})$, we have that

$$\eta(L(A)x, L(A)x) = (Lx)^\mu (Lx)_\mu = \det(-L(A)x^\mu \sigma_\mu) = \det(-x^\mu \sigma_\mu) = x^\mu x_\mu = \eta(x, x),$$

because $\det A = \det A^* = 1$. Thus, $L(A) \in \text{O}(1, 3)$. Since $\text{SL}(2, \mathbb{C})$ is connected, $L(\mathbf{1}_{\mathbb{C}^2}) = \mathbf{1}_{\mathbb{R}^4}$ and since $\text{SL}(2, \mathbb{C}) \ni A \mapsto L(A) \in \text{O}(1, 3)$ is continuous, $L(A) \in \mathcal{L}_+^\uparrow$ for every $A \in \text{SL}(2, \mathbb{C})$. That $L(AB) = L(A)L(B)$ follows exactly as in Proposition 1.1.

Next, we show that $L(A) = \mathbf{1}_{\mathbb{R}^4}$ if and only if $A \in \{-\mathbf{1}_{\mathbb{C}^2}, \mathbf{1}_{\mathbb{C}^2}\}$, implying the double covering property. So, suppose that $L(A) = \mathbf{1}_{\mathbb{R}^4}$. Then, for $e_0 = (1, 0, 0, 0)$, we have

$$\mathbf{1}_{\mathbb{C}^2} = \sigma_0 = e_0^\mu \sigma_\mu = (L(A)e_0)^\mu \sigma_\mu = A\sigma_0 A^* = AA^*.$$

Combined with $\det A = 1$, we see that $A \in \text{SU}(2)$ and could conclude that $A \in \{-\mathbf{1}_{\mathbb{C}^2}, \mathbf{1}_{\mathbb{C}^2}\}$ by arguing as in Prop. 1.1. An alternative argument goes as follows. Writing

$$A = U \begin{pmatrix} e^{i\omega} & 0 \\ 0 & e^{-i\omega} \end{pmatrix} U^* = U D U^*$$

for a some $U \in \text{U}(2)$ and $\omega \in [0, 2\pi)$, we have that $L(A) = L(D) = \mathbf{1}_{\mathbb{R}^4}$. Since

$$\begin{aligned} D(-x^\mu \sigma_\mu)D^* &= -x_0 \sigma_0 + x_1 \begin{pmatrix} 0 & e^{2i\omega} \\ e^{-2i\omega} & 0 \end{pmatrix} + ix_2 \begin{pmatrix} 0 & e^{2i\omega} \\ -e^{-2i\omega} & 0 \end{pmatrix} + x_3 \sigma_3 \\ &= -x_0 \sigma_0 + (\cos(2\omega)x_1 - \sin(2\omega)x_2)\sigma_1 + (\sin(2\omega)x_1 + \cos(2\omega)x_2)\sigma_2 + x_3 \sigma_3, \end{aligned}$$

the assumption $L(D) = \mathbf{1}_{\mathbb{R}^4}$ implies that $\cos(2\omega) = 1$ and $\sin(2\omega) = 0$, i.e. $\omega \in \{0, \pi\}$. This means that $D \in \{-\mathbf{1}_{\mathbb{C}^2}, \mathbf{1}_{\mathbb{C}^2}\}$ and hence $A = U D U^* \in \{-\mathbf{1}_{\mathbb{C}^2}, \mathbf{1}_{\mathbb{C}^2}\}$.

Next, let us show that $L : \text{SL}(2, \mathbb{C}) \rightarrow \mathcal{L}_+^\uparrow$ is surjective. To this end, observe first that for every $A \in \text{SU}(2) \subset \text{SL}(2, \mathbb{C})$, we have that

$$L(A)_{00} = \frac{1}{2} \text{tr } AA^* = 1$$

so that $L(A) \in \text{SO}(3)$, by Lemma 2.5. Noting that $L|_{\text{SU}(2)} : \text{SU}(2) \rightarrow \text{SO}(3)$ is equivalent to the covering map from Prop. 1.1, we conclude that this restriction is onto. By the group homomorphism property and Lemma 2.6, it therefore suffices to prove that for every $b > 0$, there exists $A \in \text{SL}(2, \mathbb{C})$ such that

$$L(A) = e^{bB_3} = \begin{pmatrix} \cosh(b) & 0 & 0 & \sinh(b) \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh(b) & 0 & 0 & \cosh(b) \end{pmatrix}.$$

By the defining property of $L(A)$, we thus need to find $A \in \mathrm{SL}(2, \mathbb{C})$ such that

$$-Ax^\mu \sigma_\mu A^* = -(\cosh(b)x_0 + \sinh(b)x_3)\sigma_0 + x_1\sigma_1 + x_2\sigma_2 + (\sinh(b)x_0 + \cosh(b)x_3)\sigma_3$$

for every $x \in \mathbb{R}^4$. We leave it as an *exercise* to verify that $L(A) = e^{bB_3}$ for

$$A = i \begin{pmatrix} 0 & e^{\frac{b}{2}} \\ e^{-\frac{b}{2}} & 0 \end{pmatrix} = i \cosh(b/2)\sigma_1 + \sinh(b/2)\sigma_2 \in \mathrm{SL}(2, \mathbb{C}).$$

Finally, the smoothness of the local inverse of $L : \mathrm{SL}(2, \mathbb{C}) \rightarrow \mathcal{L}_+^\uparrow$ follows from the fact that $\mathcal{L}_+^\uparrow$ is locally diffeomorphic to $\mathfrak{so}(1, 3)$ (via the exponential map) which is diffeomorphic to $\mathfrak{sl}(2, \mathbb{C})$, by Problem 2.12, which is locally diffeomorphic to $\mathrm{SL}(2, \mathbb{C})$. $\square$

We conclude this section with a few further definitions which become relevant when discussing the transformation and analytic continuation behavior of the Wightman functions. The complex Lorentz group $\mathcal{L}(\mathbb{C})$ is the complexified $\mathrm{O}(1, 3)$, i.e.

$$\mathcal{L}(\mathbb{C}) = \{L \in \mathbb{C}^{4 \times 4} : L^T \eta L = \eta\} \quad (2.15)$$

It is a Lie group that decomposes into two disconnected parts $\mathcal{L}_\pm(\mathbb{C}) = \{L \in \mathcal{L}(\mathbb{C}) : \det L = \pm 1\}$ (which turn out to be the connected components of $\mathcal{L}(\mathbb{C})$) and the next problem determines the double cover of the proper complex Lorentz group $\mathcal{L}_+(\mathbb{C})$.

Problem 2.13. *Prove that $\mathcal{L}_+(\mathbb{C})$ is a Lie group and determine its Lie algebra. Show that the direct product $\mathrm{SL}(2, \mathbb{C}) \times \mathrm{SL}(2, \mathbb{C})$ is a double cover of $\mathcal{L}_+(\mathbb{C})$ by proceeding as in the proof of Prop. 2.9 and defining $L : \mathrm{SL}(2, \mathbb{C}) \times \mathrm{SL}(2, \mathbb{C}) \rightarrow \mathcal{L}_+(\mathbb{C})$ via*

$$A z^\mu \sigma_\mu B^T = (L(A, B)z)^\mu \sigma_\mu, \quad \forall z \in \mathbb{C}^4.$$

Finally, to include the spacetime translations we associate to the groups $\mathcal{L}_+^\uparrow$, $\mathrm{SL}(2, \mathbb{C})$, $\mathcal{L}_+(\mathbb{C})$ and $\mathrm{SL}(2, \mathbb{C}) \times \mathrm{SL}(2, \mathbb{C})$ their inhomogeneous versions

$$\mathcal{P}_+^\uparrow = \mathbb{R}^4 \rtimes \mathcal{L}_+^\uparrow, \quad \mathbb{R}^4 \rtimes \mathrm{SL}(2, \mathbb{C}), \quad \mathbb{C}^4 \rtimes \mathcal{L}_+(\mathbb{C}), \quad \mathbb{C}^4 \rtimes (\mathrm{SL}(2, \mathbb{C}) \times \mathrm{SL}(2, \mathbb{C})), \quad (2.16)$$

which correspond to the outer direct products. In case of the covering groups $\mathrm{SL}(2, \mathbb{C})$ and $\mathrm{SL}(2, \mathbb{C}) \times \mathrm{SL}(2, \mathbb{C})$, the group laws are

$$(a, A)(b, B) = (a + L(A)b, AB) \quad \text{and} \quad (a, A, B)(b, A', B') = (a + L(A, B)b, AA', BB').$$

for all $(a, A), (b, B) \in \mathbb{R}^4 \rtimes \mathrm{SL}(2, \mathbb{C})$ and $(a, A, B), (b, A', B') \in \mathbb{C}^4 \rtimes (\mathrm{SL}(2, \mathbb{C}) \times \mathrm{SL}(2, \mathbb{C}))$.

2.4 Operators on Fock Space

In this section, we introduce basic examples of operator-valued distributions on the so called bosonic and fermionic Fock spaces. These operators, the so called creation and

annihilation operators, can be used to provide explicit infinite dimensional representations of the canonical commutation and, respectively, anticommutation relations. In the next Section 3, this is used to define concrete examples of free quantum fields.

The Fock space built over a given one-particle Hilbert space is a simple model of a state space that allows for an arbitrary number of particles. In these notes, we are particularly interested in describing the bosonic and fermionic Fock spaces. For particles whose wave functions are elements in $L^2(\mathcal{M})$, for some $\mathcal{M} = (\mathcal{X}, \mathcal{B}(\mathcal{X}), \mu)$, the Fock space $\mathcal{F} = \mathcal{F}(L^2(\mathcal{M}))$ built over $L^2(\mathcal{M})$ is defined as

$$\begin{aligned} \mathcal{F} &= \left\{ \psi = (\psi^{(0)}, \psi^{(1)}, \dots) : \psi_n \in L_s^2(\mathcal{X}^n) \forall n \in \mathbb{N}_0, \sum_{n=0}^{\infty} \|\psi^{(n)}\|_{L^2(\mathcal{X}^n)}^2 < \infty \right\} \\ &= \mathbb{C} \oplus \bigoplus_{n=1}^{\infty} \bigotimes_{j=1}^n L^2(\mathcal{M}) = \mathbb{C} \oplus \bigoplus_{n=1}^{\infty} L^2(\mathcal{X}^n). \end{aligned} \quad (2.17)$$

Here, we set $L^2(\mathcal{X}^0) = \mathbb{C}$. The space $\mathcal{F}$ is a Hilbert space (*exercise*) with inner product

$$\langle \varphi, \psi \rangle = \overline{\varphi^{(0)}} \psi^{(0)} + \sum_{j=1}^{\infty} \langle \varphi^{(j)}, \psi^{(j)} \rangle_{L^2(\mathcal{X}^j)}.$$

Based on our description of many-body systems, the sector $L^2(\mathcal{X}^n) \hookrightarrow \mathcal{F}$ describes states with exactly $n \in \mathbb{N}_0$ particles. The space describing states of zero particles is one-dimensional and generated by what's called the vacuum $\Phi = (1, 0, 0, \dots) \in \mathcal{F}$.

The bosonic and fermionic Fock spaces $\mathcal{F}_s$ and, respectively, $\mathcal{F}_a$ correspond to the symmetrizations and, respectively, antisymmetrizations of $\mathcal{F}$ as in Example 1.9, that is

$$\begin{aligned} \mathcal{F}_s &= \mathbb{C} \oplus \bigoplus_{n=1}^{\infty} \bigotimes_{\text{sym}}^n L^2(\mathcal{M}) = \mathbb{C} \oplus \bigoplus_{n=1}^{\infty} L_s^2(\mathcal{X}^n), \\ \mathcal{F}_a &= \mathbb{C} \oplus \bigoplus_{n=1}^{\infty} \bigwedge^n L^2(\mathcal{M}) = \mathbb{C} \oplus \bigoplus_{n=1}^{\infty} L_a^2(\mathcal{X}^n), \end{aligned} \quad (2.18)$$

In a Fock space setting, we can create and annihilate a particle in a state $f \in L^2(\mathcal{M})$ by tensorizing and, respectively, taking the inner product with $f \in L^2(\mathcal{M})$. It is useful to encode this in a set of operators which turn out to form operator-valued distributions. We start with the bosonic case. Given $f \in L^2(\mathcal{M})$, we define the bosonic creation operator $a^*(f)$ and the bosonic annihilation operator $a(g)$ by

$$\begin{aligned} (a^*(f)\psi)^{(n)}(x_1, \dots, x_n) &= \frac{1}{\sqrt{n}} \sum_{j=1}^n f(x_j) \psi^{(n-1)}(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n), \\ (a(g)\psi)^{(n)}(x_1, \dots, x_n) &= \sqrt{n+1} \int_{\mathcal{X}} \mu(dx) \overline{g(x)} \psi^{(n+1)}(x, x_1, \dots, x_n) \end{aligned} \quad (2.19)$$

for every

$$\psi \in \mathcal{F}_s^{<\infty} = \bigcup_{n=0}^{\infty} \mathcal{F}_s^{\leq n}, \quad \mathcal{F}_s^{\leq n} = \mathbb{C} \oplus \bigoplus_{j=1}^n L_s^2(\mathcal{X}^j) \hookrightarrow \mathcal{F}_s$$

Here, in (2.19), we set for $n = 0$ that $(a^*(f)\psi)^{(0)} = 0$. Note that $\mathcal{F}_s^{<\infty} \subset \mathcal{F}_s$ is dense and that $a^*(f) : \mathcal{F}_s^{<\infty} \rightarrow \mathcal{F}_s^{<\infty}$ and $a(f) : \mathcal{F}_s^{<\infty} \rightarrow \mathcal{F}_s^{<\infty}$ preserve the bosonic symmetry.

Lemma 2.7. *Let $f, g \in L^2(\mathcal{M})$, $\psi, \varphi \in \mathcal{F}_s^{<\infty}$. Then, $\langle \psi, a(f)\varphi \rangle = \langle a^*(f)\psi, \varphi \rangle$ and*

$$[a(f), a^*(g)] = \langle f, g \rangle, \quad [a(f), a(g)] = [a^*(f), a^*(g)] = 0. \quad (2.20)$$

Moreover, both $a(f)$ and $a^*(f)$ are closable and their closures are adjoint to each other.

Finally, for $\mathcal{M} = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), dx)$, we have that $\mathcal{S}(\mathbb{R}^d) \ni f \mapsto \langle \varphi, a^*(f)\psi \rangle \in \mathbb{C}$ and $\mathcal{S}(\mathbb{R}^d) \ni f \mapsto \langle \varphi, a(f)\psi \rangle \in \mathbb{C}$ are elements in $\mathcal{S}'(\mathbb{R}^d)$.

Proof. The computations in the first part are straightforward. For definiteness, we verify the first commutator equation in (2.20) and leave the remaining identities as an exercise. To prove the first equation in (2.20), we can pick without loss of generality a state $\psi \in L_s^2(\mathcal{X}^n) \hookrightarrow \mathcal{F}_s$, for some $n \in \mathbb{N}_0$, and we find explicitly that

$$\begin{aligned} & ([a(f), a^*(g)]\psi)(x_1, \dots, x_n) \\ &= \sqrt{n+1} \int_{\mathcal{X}} \mu(dx) \bar{f}(x) \frac{1}{\sqrt{n+1}} g(x) \psi(x_1, \dots, x_n) \\ & \quad + \sqrt{n+1} \int_{\mathcal{X}} \mu(dx) \bar{f}(x) \left(\frac{1}{\sqrt{n+1}} \sum_{j=1}^n g(x_j) \psi(x, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \right) \\ & \quad - \frac{1}{\sqrt{n}} \sum_{j=1}^n g(x_j) \left(\sqrt{n} \int_{\mathcal{X}} \mu(dx) \bar{f}(x) \psi(x, x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \right) \\ &= \langle f, g \rangle \psi(x_1, \dots, x_n), \end{aligned}$$

as claimed. Similarly, one verifies the remaining identities.

Next, recall that $a(f)$ is closable if it admits a smallest closed extension which exists if and only if $\psi_n \rightarrow 0, a(f)\psi_n \rightarrow \phi$ as $n \rightarrow \infty$, for $\phi, \psi_n \in \mathcal{F}_s, \forall n \in \mathbb{N}$, implies that $\phi = 0$. So, consider such a sequence. Then, by definition of the norm in $\mathcal{F}_s$, we have

$$\lim_{n \rightarrow \infty} \|\psi_n^{(m)}\|_{L_s^2(\mathcal{X}^m)} = 0, \quad \forall m \in \mathbb{N}_0,$$

and hence also $\phi = 0$, because

$$\lim_{n \rightarrow \infty} \|(a(f)\psi_n)^{(m)}\|_{L_s^2(\mathcal{X}^m)} = \sqrt{m} \left\| \int_{\mathcal{X}} \mu(dx) \bar{f}(x) \psi_n^{(m+1)}(x, \cdot) \right\|_{L_s^2(\mathcal{X}^m)} = 0, \quad \forall m \in \mathbb{N}_0.$$

Similarly, one shows that $a^*(f)$ is closable (*exercise*). To see that the closures $\overline{a(f)}$ and $\overline{a^*(f)}$ are adjoint to each other, notice first that the identity

$$\langle \varphi, a(f)\psi \rangle = \langle a^*(f)\varphi, \psi \rangle$$

extends from $\varphi, \psi \in \mathcal{F}_s^{<\infty}$ to $\varphi \in D_{\overline{a^*(f)}}, \psi \in D_{\overline{a(f)}}$. Indeed, $\psi \in D_{\overline{a(f)}}$ if and only if there exists a sequence $(\psi_n)_{n \in \mathbb{N}}$ such that $\lim_{n \rightarrow \infty} \psi_n = \psi, \lim_{n \rightarrow \infty} a(f)\psi_n$ exists (and equals $\overline{a(f)}\psi$), and similarly for $\varphi \in D_{\overline{a^*(f)}}$. Choosing such sequences, we get

$$\langle \varphi, \overline{a(f)}\psi \rangle = \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \langle \varphi_m, a(f)\psi_n \rangle = \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \langle a^*(f)\varphi_m, \psi_n \rangle = \langle \overline{a^*(f)}\varphi, \psi \rangle$$

By definition of the adjoint, the last identity implies immediately that $\overline{a^*(f)} \subset \overline{a(f)}^*$. On the other hand, suppose that $\psi = (\psi^{(0)}, \psi^{(1)}, \dots) \in D_{\overline{a(f)}}^*$ and let

$$\psi_n = (\psi^{(0)}, \dots, \psi^{(n)}, 0, \dots) \in \mathcal{F}_s^{<\infty}, \forall n \in \mathbb{N}_0.$$

Then, $\lim_{n \rightarrow \infty} \psi_n = \psi$ in $\mathcal{F}_s$ and we have that

$$\begin{aligned} \limsup_{n \rightarrow \infty} \|a^*(f)\psi_n\|^2 &= \limsup_{n \rightarrow \infty} \sum_{j=1}^n \|a^*(f)\psi^{(j)}\|_{L_s^2(\mathcal{X}^j)}^2 \\ &= \limsup_{n \rightarrow \infty} \sup_{\varphi \in \mathcal{F}_s^{<\infty}: \|\varphi\|=1} |\langle \varphi, a^*(f)\psi_n \rangle|^2 \\ &= \limsup_{n \rightarrow \infty} \sup_{\varphi \in \mathcal{F}_s^{<\infty}: \|\varphi\|=1} |\langle \varphi_{n+1}, a^*(f)\psi \rangle|^2 \\ &= \limsup_{n \rightarrow \infty} \sup_{\varphi \in \mathcal{F}_s^{<\infty}: \|\varphi\|=1} |\langle a(f)\varphi_{n+1}, \psi \rangle|^2 < \infty, \end{aligned}$$

where the last step follows by the assumption that $\psi \in D_{\overline{a(f)}}^*$. Consequently,

$$\|a^*(f)(\psi_n - \psi_m)\| \rightarrow 0$$

as $n, m \rightarrow \infty$ so that $(a^*(f)\psi_n)_{n \in \mathbb{N}}$ converges in $\mathcal{F}_s$. This shows that $\psi \in D_{\overline{a^*(f)}}$ and similar considerations conclude that $\overline{a(f)}^* \subset \overline{a^*(f)}$ (*exercise*) so that $\overline{a(f)}^* = \overline{a^*(f)}$.

Finally, if $\mathcal{M} = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), dx)$ and $\varphi, \psi \in \mathcal{F}_s^{<\infty}$, assume without loss of generality that $\varphi = \varphi_n, \psi = \psi_n$ for some $n \in \mathbb{N}$. Then, Cauchy-Schwarz implies that

$$|\langle \varphi, a(f)\psi \rangle| = |\langle \psi, a^*(f)\varphi \rangle| \leq \sqrt{n} \|\psi\| \|\varphi\| \|f\|_2.$$

Since in addition $\mathcal{S}(\mathbb{R}^d) \hookrightarrow L^2(\mathbb{R}^d)$ with continuous embedding, by Problem 2.4, we conclude that $f \mapsto \langle \varphi, a^*(f)\psi \rangle \in \mathcal{S}'(\mathbb{R}^d)$ and $f \mapsto \langle \varphi, a(\overline{f})\psi \rangle \in \mathcal{S}'(\mathbb{R}^d)$. $\square$

From now on, for simplicity of notation, we identify $a^*(f)$ and $a(g)$ with their closures. Before switching to the fermionic setting, let us point out that the creation operators provide us with an important example of an operator-valued distribution. Indeed, let's view $(a^*(f))_{f \in \mathcal{S}(\mathbb{R}^d)}$ as a quantum field and suppose that one could identify it with an operator-valued field $(a_x^*)_{x \in \mathbb{R}^d}$ so that

$$a^*(f) = \int_{\mathbb{R}^d} dx f(x) a_x^*, \forall f \in \mathcal{S}(\mathbb{R}^d).$$

Without giving a precise meaning to the operator integral, a minimal regularity requirement for such an identification should certainly be that $a_x^* = \lim_{\epsilon \rightarrow 0} a^*(\varphi_{\epsilon, x})$, where $\varphi_{\epsilon, x} = \frac{1}{\epsilon^d} \varphi(\frac{1}{\epsilon}(\cdot - x))$ for a suitable $\varphi \in C_c^\infty(\mathbb{R}^d) \subset \mathcal{S}(\mathbb{R}^d)$, $0 \leq \varphi \leq 1$ with $\int_{\mathbb{R}^d} \varphi = 1$, on a sufficiently large domain in $\mathcal{F}_s$. If this was possible, it would lead to divergences as in

$$\langle a^*(\varphi_{\epsilon, x})\psi, a^*(\varphi_{\epsilon, x})\psi \rangle = \langle \psi, a^*(\varphi_{\epsilon, x})a(\varphi_{\epsilon, x})\psi \rangle + \|\varphi_{\epsilon, x}\|^2 \geq \frac{1}{\epsilon^d} \xrightarrow{\epsilon \rightarrow 0} \infty.$$

Such problems are reminiscent of divergences that occur when trying to define powers of distributions (like e.g. the Dirac δ) and in fact, heuristically, we can identify

$$(a_x^* \psi)^{(n)}(x_1, \dots, x_n) = \frac{1}{\sqrt{n}} \sum_{j=1}^n \delta_x(x_j) \psi^{(n-1)}(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n), \quad \forall \psi \in \mathcal{F}_s,$$

which does not correspond to a densely defined, linear operator. Identifications such as

$$a(f) = \int_{\mathbb{R}^d} dx \bar{f}(x) a_x, \quad a^*(g) = \int_{\mathbb{R}^d} dx g(x) a_x^* \quad (2.21)$$

are in practice nevertheless quite useful and can in many instances given a precise, mathematical meaning in terms of statements about densely defined, quadratic forms. Indeed, in case of the annihilation operators, we can define

$$(a_x \psi)^{(n)}(x_1, \dots, x_n) = \sqrt{n+1} \psi^{(n+1)}(x, x_1, \dots, x_n), \quad \forall n \in \mathbb{N}_0, \psi \in \mathcal{D}_S,$$

where

$$\mathcal{D}_S = \left\{ \psi \in \mathcal{F}_s^{<\infty} : \psi^{(n)} \in \mathcal{S}(\mathbb{R}^{nd}), \quad \forall n \in \mathbb{N} \right\}.$$

Then, $a_x \mathcal{D}_S \subset \mathcal{D}_S$ for every $x \in \mathbb{R}^d$ so that arbitrary powers of the field operators $(a_x)_{x \in \mathbb{R}^d}$ are well-defined, too. By duality, this enables us to define e.g. for every suitable, measurable integral kernel $\Lambda : \mathbb{R}^{dk+dl} \rightarrow \mathbb{C}$ the quadratic form

$$\begin{aligned} & \left(\int_{\mathbb{R}^{dk+dl}} dx_1 \dots dx_k dy_1 \dots dy_l \Lambda(x_1, \dots, x_k, y_1, \dots, y_l) a_{x_1}^* \dots a_{x_k}^* a_{y_1} \dots a_{y_l} \right) (\varphi, \psi) \\ &= \int_{\mathbb{R}^{dk+dl}} dx_1 \dots dx_k dy_1 \dots dy_l \Lambda(x_1, \dots, x_k, y_1, \dots, y_l) \langle a_{x_1} \dots a_{x_k} \varphi, a_{y_1} \dots a_{y_l} \psi \rangle \end{aligned} \quad (2.22)$$

on the dense domain $\mathcal{D}_S \times \mathcal{D}_S \subset \mathcal{F}_s \times \mathcal{F}_s$. In practice, (2.22) even turns out useful if $\Lambda \in \mathcal{S}'(\mathbb{R}^{dk+dl})$ can be identified with a (formal) integral kernel of some distribution. For example, if $k = l = 1$ and $\Lambda(x, y) = \delta(x - y)$ which represents the integral kernel of the identity operator $\mathbf{1}_{\mathcal{S}(\mathbb{R}^d)}$ on $\mathcal{S}(\mathbb{R}^d)$, then (2.22) corresponds to the quadratic form

$$\left(\int_{\mathbb{R}^{2d}} dx dy \Lambda(x, y) a_x^* a_y \right) (\varphi, \psi) = \left(\int_{\mathbb{R}^d} dx a_x^* a_x \right) (\varphi, \psi) = \int_{\mathbb{R}^d} dx \langle a_x \varphi, a_x \psi \rangle = \langle \varphi, \mathcal{N} \psi \rangle,$$

where $\mathcal{N} : D_{\mathcal{N}} \rightarrow \mathcal{F}_s$ is equal to the so called number of particles operator, defined by

$$\mathcal{N} \varphi = (n \varphi^{(n)})_{n \in \mathbb{N}_0} \in D_{\mathcal{N}}, \quad \forall \varphi \in D_{\mathcal{N}} = \{ \varphi \in \mathcal{F}_s : (n \varphi^{(n)})_{n \in \mathbb{N}_0} \in \mathcal{F}_s \}. \quad (2.23)$$

Note that $\mathcal{N}$ is self-adjoint by Theorem 1.1.

Problem 2.14. Let $\Lambda \in L^2(\mathbb{R}^{dk+dl})$. Consider the quadratic form defined in (2.22) and denoted by $Q_{\Lambda} : \mathcal{D}_S \times \mathcal{D}_S \rightarrow \mathbb{C}$. Prove that the quadratic form

$$(\mathcal{N} + 1)^{-\frac{k}{2}} Q_{\Lambda} (\mathcal{N} + 1)^{-\frac{l}{2}}$$

can be identified with the form associated to a bounded operator $A_{\Lambda} : \mathcal{F}_s \rightarrow \mathcal{F}_s$ whose operator norm is bounded by $\|A_{\Lambda}\| \leq \|\Lambda\|_{L^2(\mathbb{R}^{dk+dl})}$. Conclude that Q_{Λ} corresponds to the quadratic form associated with a densely defined, linear operator on $\mathcal{F}_s$.

The previous considerations apply to the bosonic setting. In the fermionic Fock space $\mathcal{F}_a$, defined in (2.18), the fermionic creation and annihilation operators $a^*(f)$ and, respectively, $a(g)$ are defined similarly on $\mathcal{F}_a^{<\infty}$ by

$$\begin{aligned} (a^*(f)\psi)^{(n)}(x_1, \dots, x_n) &= \frac{1}{\sqrt{n}} \sum_{j=1}^n (-1)^j f(x_j) \psi^{(n-1)}(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n), \\ (a(g)\psi)^{(n)}(x_1, \dots, x_n) &= \sqrt{n+1} \int_{\mathcal{X}} \mu(dx) \bar{g}(x) \psi^{(n+1)}(x, x_1, \dots, x_n). \end{aligned} \quad (2.24)$$

In these notes, we work exclusively with purely bosonic or purely fermionic models. For simplicity, we therefore use the same symbols for the creation and annihilation operators in both cases. The fermionic operators preserve the fermionic antisymmetry and they satisfy the so called canonical anticommutation relations

$$[a(f), a^*(g)]_+ = \langle f, g \rangle, \quad [a(f), a(g)]_+ = [a^*(f), a^*(g)]_+ = 0. \quad (2.25)$$

Problem 2.15. For $f, g \in L^2(\mathcal{M})$, prove that the fermionic creation and annihilation operators, defined in (2.24), extend to bounded operators $a^*(f), a(g) : \mathcal{F}_a \rightarrow \mathcal{F}_a$. Verify the anticommutation relations (2.25) and show that $a^*(f) = (a(f))^*$.

We conclude this section with some further notation. Recalling the definition of the bosonic number of particles operator $\mathcal{N}$, defined in (2.23), notice that

$$\mathcal{N} = 0 \oplus \bigoplus_{n=1}^{\infty} \sum_{j=1}^n (\mathbf{1}_{L^2(\mathcal{M})})_{x_j} = d\Gamma(\mathbf{1}_{L^2(\mathcal{M})}).$$

By convention, one therefore calls $\mathcal{N}$ the second quantization $d\Gamma(\mathbf{1}_{L^2(\mathcal{M})})$ of the one-body operator $\mathbf{1}_{L^2(\mathcal{M})}$ (although this is not related to a quantization as in the sense of Section 1.2.1, say). More generally, if $A : D_A \rightarrow L^2(\mathcal{M})$ is self-adjoint, its second quantization $d\Gamma(A)$ is defined (on a suitable, dense domain in $\mathcal{F}_{s/a}$) by

$$d\Gamma(A) = 0 \oplus \bigoplus_{n=1}^{\infty} \sum_{j=1}^n A_{x_j}, \quad A_{x_j} = \mathbf{1}_{L^2(\mathcal{M})} \otimes \dots \otimes \mathbf{1}_{L^2(\mathcal{M})} \otimes A \otimes \mathbf{1}_{L^2(\mathcal{M})} \otimes \dots \otimes \mathbf{1}_{L^2(\mathcal{M})},$$

where the factor A appears in the j -th slot in A_{x_j} , for $j \in \mathbb{N}$. If $U \in \mathcal{U}(L^2(\mathcal{M}))$ is a unitary operator, on the other hand, its second quantization $\Gamma(U)$ is defined by

$$\Gamma(U) = \mathbf{1}_{\mathbb{C}} \oplus \bigoplus_{n=1}^{\infty} U^{\otimes n}.$$

In the concrete examples below, we specify the domain properties of such operators precisely. The notation $\Gamma(\cdot)$ and $d\Gamma(\cdot)$ stems from the observation (*exercise*) that

$$i \frac{d}{dt} \Gamma(U_t) = d\Gamma(A) \Gamma(U_t) \quad \text{for } U_t = e^{-itA}.$$

Problem 2.16. Prove that $\Gamma(U) : \mathcal{F}_{s/a} \rightarrow \mathcal{F}_{s/a}$ maps the bosonic and, respectively, fermionic Fock spaces to themselves.

3 Construction of Free Quantum Fields

In this section, we construct several important, non-interacting quantum field theories and verify that they satisfy the Wightman axioms, discussed in Section 1.4. As a warm-up, we start with the quantization of the massive Klein-Gordon field and discuss its physical interpretation. Next, we analyze the finite-dimensional representations of the Lorentz and Poincaré groups which is used in the last section to construct the massive vector and Dirac quantum fields.

3.1 The Free Massive Scalar Field

The goal of this section is to construct a quantization of the real, massive, scalar (or equivalently Klein-Gordon) field. From Section 1.3.1, we recall that scalar fields solve

$$(\square + m^2)\varphi = 0, \quad (3.1)$$

where

$$\square = \partial^\nu \partial_\nu = \partial_t^2 - \sum_{j=1}^3 \partial_{x_j}^2 = \partial_t^2 - \Delta$$

denotes the d'Alembertian in $\mathbb{R}^4$. In Section 1.2.1, we discussed how basic systems of finitely many degrees of freedom can be quantized based on the canonical quantization strategy. Before generalizing this method to classical, relativistic fields, we first recall two perspectives on how to quantize a massive, relativistic particle. Afterwards, we quantize the Klein-Gordon field and we discuss its interpretation.

3.1.1 Quantization of a Massive, Relativistic Particle

Suppose we consider a single particle of mass $m > 0$ evolving freely in Minkowski space-time $(\mathbb{R}^4, \eta)$. As discussed in detail at the end of Section 1.3, classically this reduces to the Hamiltonian dynamics on $\mathcal{P} = \mathbb{R}^3 \times \mathbb{R}^3$ with Hamiltonian $H \in C^\infty(\mathcal{P})$, defined by

$$H(\mathbf{x}, \mathbf{p}) = \omega(\mathbf{p}) = \sqrt{|\mathbf{p}|^2 + m^2}.$$

According to the canonical quantization scheme, the corresponding quantum mechanical model describes the state of the system by a wave function in $L^2(\mathbb{R}^3)$, position and momentum are described by the multiplication operator $x = (x_1, x_2, x_3)$ and, respectively, by $p = (-i\partial_{x_1}, -i\partial_{x_2}, -i\partial_{x_3})$ on their canonical domains, and the energy is described by the Hamilton operator $H : D_H \rightarrow L^2(\mathbb{R}^3)$, defined by

$$H = \sqrt{-\Delta + m^2}, \quad D_H = \{\psi \in L^2(\mathbb{R}^3) : \mathbf{p} \mapsto \sqrt{|\mathbf{p}|^2 + m^2} \hat{\psi}(\mathbf{p}) \in L^2(\mathbb{R}^3)\}.$$

Note that $H : D_H \rightarrow L^2(\mathbb{R}^3)$ is self-adjoint by Theorem 1.1. In particular, the strongly continuous, unitary dynamics $(e^{-iHt})_{t \in \mathbb{R}}$ is well-defined.

The previous quantization strategy determines the Hamiltonian, that generates the quantum dynamics, based on the classical Hamiltonian formulation of the model. One

may also proceed differently here, based on a conceptually equally important point of view, namely the assumption that every relativistic quantum theory contains a projective unitary representation of the proper Poincaré group $\mathcal{P}_+^\uparrow$. This is consistent with Axiom **O** in Section 1.4. For a single, massive particle, such a representation can be obtained as follows. Define the so called mass shell V_m^+ of mass $m > 0$ by

$$V_m^+ = \{p \in \mathbb{R}^4 : p^2 = p^\mu p_\mu = p_0^2 - |\mathbf{p}|^2 = m^2, p_0 > 0\}. \quad (3.2)$$

Trajectories of classical, relativistic particles of mass $m > 0$ are contained in V_m^+ so it is natural to build a unitary representation of $\mathcal{P}_+^\uparrow$ on a suitable L^2 -version of V_m^+ . To proceed similarly as in Section 1.2.2, where we obtained a unitary representation of $\text{SO}(3)$ on $L^2(\mathbb{R}^3)$, we first need to construct a measure on V_m^+ that is invariant under $\mathcal{L}_+^\uparrow$. To this end, notice that V_m^+ is a smooth manifold that is diffeomorphic to $\mathbb{R}^3$ via

$$\mathbb{R}^3 \ni \mathbf{p} \mapsto \iota^{-1}(\mathbf{p}) = (\omega(\mathbf{p}), \mathbf{p}) \in V_m^+, \quad \omega(\mathbf{p}) = \sqrt{|\mathbf{p}|^2 + m^2}.$$

Since V_m^+ is an embedded submanifold in $\mathbb{R}^4$, its topology coincides with the subspace topology inherited from $\mathbb{R}^4$, i.e. $\Omega \subset V_m^+$ is open if and only if there exists some $U \subset \mathbb{R}^4$ open such that $\Omega = U \cap V_m^+$. As explained shortly, a $\mathcal{L}_+^\uparrow$ -invariant measure $\lambda_m : \mathcal{B}(V_m^+) \rightarrow [0, \infty)$, that in fact turns out to be unique (see [22, Chapter IX]), can be defined by

$$\lambda_m(\Omega) = \int_{\iota(\Omega)} \frac{d\mathbf{p}}{2\omega(\mathbf{p})} = \int_{\iota(\Omega)} \frac{d\mathbf{p}}{2\sqrt{|\mathbf{p}|^2 + m^2}}, \quad \forall \Omega \in \mathcal{B}(V_m^+). \quad (3.3)$$

Lemma 3.1. *Let λ_m be as in (3.3). Then $\lambda_m(L\Omega) = \lambda_m(\Omega)$ for all $\Omega \in \mathcal{B}(V_m^+)$, $L \in \mathcal{L}_+^\uparrow$.*

Proof. Since the Borel sets are generated by open sets, it suffices to show that

$$\lambda_m(\Omega) = \int_{\iota(\Omega)} \frac{d\mathbf{p}}{2\omega(\mathbf{p})} = \int_{\iota(L\Omega)} \frac{d\mathbf{p}}{2\omega(\mathbf{p})} = \lambda_m(L\Omega)$$

for all open sets $\Omega \subset V_m^+$ and for all $L \in \mathcal{L}_+^\uparrow$. Furthermore, it is enough to consider open sets that are bounded. So, pick some open, bounded $\Omega \subset V_m^+$ and let $U \subset \mathbb{R}^4$ be open and bounded such that $\Omega = U \cap V_m^+$. Choosing a standard mollifying sequence $(\psi_\delta)_{\delta>0}$, such that $\psi_\delta = \delta^{-4}\psi(\delta^{-1}\cdot) \in C_c^\infty(B_\delta(0))$ for some $\psi \in C_c^\infty(B_1(0)) \subset C_c^\infty(\mathbb{R}^4)$ with $0 \leq \psi \leq 1$, $\int_{\mathbb{R}^4} dx \psi(x) = 1$, an application of dominated convergence implies that

$$\lambda_m(\Omega) = \int_{\mathbb{R}^3} d\mathbf{p} \frac{\chi_{\iota(\Omega)}(\mathbf{p})}{2\omega(\mathbf{p})} = \int_{\mathbb{R}^3} d\mathbf{p} \frac{\chi_U(\omega(\mathbf{p}), \mathbf{p})}{2\omega(\mathbf{p})} = \lim_{\delta \rightarrow 0} \int_{\mathbb{R}^3} d\mathbf{p} \frac{(\chi_U * \psi_\delta)(\omega(\mathbf{p}), \mathbf{p})}{2\omega(\mathbf{p})}.$$

On the other hand, defining for $\epsilon > 0$ the thickening

$$V_{m,\epsilon}^+ = \{p \in \mathbb{R}^4 : p^2 \leq m^2 < p^2 + \epsilon, p_0 > 0\},$$

we have for every $\varphi \in C_c^\infty(\mathbb{R}^4)$ that

$$\begin{aligned} \int_{V_{m,\epsilon}^+} dp \varphi(p) &= \int_{\mathbb{R}^3} d\mathbf{p} \int_{\sqrt{|\mathbf{p}|^2+m^2}}^{\sqrt{|\mathbf{p}|^2+m^2+\epsilon}} dp_0 \varphi(p_0, \mathbf{p}) \\ &= \int_{\mathbb{R}^3} d\mathbf{p} (\sqrt{|\mathbf{p}|^2+m^2+\epsilon} - \sqrt{|\mathbf{p}|^2+m^2}) (\varphi(\omega(\mathbf{p}), \mathbf{p}) + O(\epsilon)) \\ &= \epsilon \int_{\mathbb{R}^3} d\mathbf{p} \frac{\varphi(\omega(\mathbf{p}), \mathbf{p})}{2\omega(\mathbf{p})} + O(\epsilon^2). \end{aligned}$$

This shows that

$$\lambda_m(\Omega) = \lim_{\delta \rightarrow 0} \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_{V_{m,\epsilon}^+} dp (\chi_U * \psi_\delta)(p).$$

Now, $V_{m,\epsilon}^+$ is clearly invariant under $\mathcal{L}_+^\uparrow$ and the standard Lebesgue measure on $\mathbb{R}^4$ is also invariant under $\mathcal{L}_+^\uparrow$, because $L \in \mathcal{L}_+^\uparrow$ means in particular that $\det L = 1$. Therefore, the restriction of Lebesgue measure to $V_{m,\epsilon}^+$ is invariant as well. Consequently

$$\int_{V_{m,\epsilon}^+} dp \varphi(p) = \int_{V_{m,\epsilon}^+} dp \varphi(Lp), \quad \forall L \in \mathcal{L}_+^\uparrow, \varphi \in C_c^\infty(\mathbb{R}^4).$$

Finally, using that $(\chi_U * \psi_\delta)(L^{-1} \cdot) = \chi_{LU} * \psi_\delta(L^{-1} \cdot)$ (using again the invariance of Lebesgue measure under $\mathcal{L}_+^\uparrow$), that $(\psi_\delta(L^{-1} \cdot))_{\delta > 0}$ is another standard mollifying sequence and that $LU \cap V_m^+ = LU \cap LV_m^+ = L\Omega$, we conclude that

$$\lambda_m(\Omega) = \lim_{\delta \rightarrow 0} \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_{V_{m,\epsilon}^+} dp (\chi_{LU} * \psi_\delta(L^{-1} \cdot))(p) = \lambda_m(L\Omega)$$

for every open, bounded $\Omega \subset V_m^+$, and thus every $\Omega \in \mathcal{B}(V_m^+)$, and for all $L \in \mathcal{L}_+^\uparrow$. $\square$

Based on the previous lemma, we obtain a representation of $\mathcal{P}_+^\uparrow$ on $L^2(V_m^+, \mathcal{B}(V_m^+), \lambda_m)$ by setting for $\psi \in L^2(V_m^+, \mathcal{B}(V_m^+), \lambda_m)$ that

$$(U(a, L)\psi)(p) = e^{ip^\mu a_\mu} \psi(L^{-1}p), \quad \text{for a.e. } p \in V_m^+, \quad \forall (a, L) \in \mathcal{P}_+^\uparrow. \quad (3.4)$$

Differentiation w.r.t. time translations shows that the Hamiltonian H takes the form

$$(H\psi)(p) = -i \frac{d}{dt} (U(te_0, \mathbf{1}_{\mathbb{R}^4})\psi)_{|t=0}(p) = \omega(\mathbf{p})\psi(\omega(\mathbf{p}), \mathbf{p}), \quad \text{for a.e. } p = (\omega(\mathbf{p}), \mathbf{p}) \in V_m^+.$$

Identifying similarly with $\mathbf{p} = (p_1, p_2, p_3)$ the momentum operators, we conclude that

$$H^2 = |\mathbf{p}|^2 + m^2 \quad (3.5)$$

on a suitable, dense domain in $L^2(V_m^+, \mathcal{B}(V_m^+), \lambda_m)$ (*exercise*). Eq. (3.5) generalizes the energy-momentum relation (1.52) of a relativistic particle to the quantum setting. In Section 3.2, we also explain in which sense the model describes a spin-zero particle.

Problem 3.1. For $(a, L) \in \mathcal{P}_+^\uparrow$, let $U(a, L)$ be defined as in (3.4).

- a) Show that $\mathcal{P}_+^\uparrow \ni (a, L) \mapsto U(a, L) \in \mathcal{U}(L^2(V_m^+))$ defines a strongly continuous, unitary representation of $\mathcal{P}_+^\uparrow$ on $L^2(V_m^+) = L^2(V_m^+, \mathcal{B}(V_m^+), \lambda_m)$.
- b) Find a unitary map from $L^2(V_m^+)$ to $L^2(\mathbb{R}^3, \mathcal{B}(\mathbb{R}^3), (2\omega(\mathbf{p}))^{-1}d\mathbf{p})$ and compute the transformed unitary representation of $\mathcal{P}_+^\uparrow$ on $L^2(\mathbb{R}^3, \mathcal{B}(\mathbb{R}^3), (2\omega(\mathbf{p}))^{-1}d\mathbf{p})$ as well as the associated Hamiltonian explicitly.
- c) Find a unitary map from $L^2(\mathbb{R}^3, \mathcal{B}(\mathbb{R}^3), (2\omega(\mathbf{p}))^{-1}d\mathbf{p})$ to $L^2(\mathbb{R}^3, \mathcal{B}(\mathbb{R}^3), d\mathbf{p})$ and compute the transformed unitary representation of $\mathcal{P}_+^\uparrow$ on $L^2(\mathbb{R}^3, \mathcal{B}(\mathbb{R}^3), d\mathbf{p})$ as well as the associated Hamiltonian explicitly. Connect this to the quantization at the beginning of this section.

Based on (3.5), we would like to interpret the representation $\mathcal{P}_+^\uparrow \ni (a, L) \mapsto U(a, L)$ as describing a relativistic massive particle of mass $m > 0$. If that is the case, we should not be able to resolve any finer structure related to the mass $m > 0$ within this representation. This is made precise in the next lemma.

Problem 3.2. For $m > 0$, show that $V_m^+ = \{L(m, 0, 0, 0) : L \in \mathcal{L}_+^\uparrow\}$ by arguing as in the proof of Lemma 2.6.

Lemma 3.2. The representation $\mathcal{P}_+^\uparrow \ni (a, L) \mapsto U(a, L)$ on $L^2(V_m^+)$ is irreducible.

Proof. By Lemma 1.3, given some $\psi \neq 0$, it suffices to show that

$$\langle \varphi, U(a, L)\psi \rangle = 0, \quad \forall (a, L) \in \mathcal{P}_+^\uparrow \implies \varphi = 0 \in L^2(V_m^+) = L^2(V_m^+, \mathcal{B}(V_m^+), \lambda_m)$$

To this end, pick some $\zeta \in \mathcal{S}(\mathbb{R}^4)$ so that by Fourier inversion $\zeta(x) = \int dp e^{-ip^\mu x_\mu} \tilde{\zeta}(p)$ for all $x \in \mathbb{R}^4$ and $\tilde{\zeta}(p) = \widehat{\zeta}(-p_0, \mathbf{p})$, in particular $\tilde{\zeta} \in \mathcal{S}(\mathbb{R}^4)$. Then

$$0 = \int_{\mathbb{R}^4} da \zeta(a) \left(\int_{V_m^+} \lambda_m(dp) \overline{\varphi}(p) e^{ip^\mu a_\mu} \psi(L^{-1}p) \right) = \int_{V_m^+} \lambda_m(dp) \tilde{\zeta}(p) \overline{\varphi}(p) \psi(L^{-1}p).$$

Then, since $\tilde{\zeta} \in \mathcal{S}(\mathbb{R}^4)$ was arbitrary, this implies (*exercise*) that $\overline{\varphi}(p)\psi(L^{-1}p) = 0$ and hence $|\varphi(p)\psi(L^{-1}p)| = 0$ for *a.e.* $p \in V_m^+$ and for every $L \in \mathcal{L}_+^\uparrow$.

To conclude from here that $\varphi = 0 \in L^2(V_m^+)$, we argue as follows. First of all, we may assume that $\psi \neq 0$ is strictly positive in an open neighborhood $U \subset V_m^+$ that contains the vector $me_0 = (m, 0, 0, 0) \in V_m^+$. Thus, in particular $\varphi|_U = 0$. Then, we find for every $p \in V_m^+$ a Lorentz transformation $L = L_p \in \mathcal{L}_+^\uparrow$ such that $p = Lme_0$, by Problem 3.2. Since $|\varphi(q)\psi(L_p q)| = 0$ for *a.e.* $q \in V_m^+$, we conclude that $\varphi = 0$ in an open neighborhood around $p \in V_m^+$ and since $p \in V_m^+$ was arbitrary, it follows that $\varphi = 0 \in L^2(V_m^+)$. $\square$

3.1.2 Quantization of the Massive Klein-Gordon Field

In this section, we quantize the real, massive, scalar field. Like in the previous section, we discuss two approaches to the problem. We start with the first one which is based on a Hamiltonian formulation (recall in particular Problem 1.9). Given a classical solution of (3.1), the most natural interpretation of its quantization corresponds (heuristically) to an operator-valued field $(\phi_x)_{x \in \mathbb{R}^4}$ which is Lorentz invariant and which solves (3.1) in an operator-valued sense. For the rigorous construction, it turns out that both the field and the equation (3.1) have to be interpreted in a distributional sense. Given an operator-valued distribution $\phi = (\phi(f))_{f \in \mathcal{S}(\mathbb{R}^4)}$ (in the sense of Axiom **I** in Section 1.4) whose domains include a dense subspace $D \subset \mathcal{H}$, we say that ϕ solves (3.1) if

$$\phi((\partial^\mu \partial_\mu + m^2)f)|_D = 0, \quad \forall f \in \mathcal{S}(\mathbb{R}^4). \quad (3.6)$$

Formally, by integration by parts, this is equivalent to

$$0 = \int_{\mathbb{R}^4} dx \phi_x (\partial^\mu \partial_\mu f + m^2 f)(x) = \int_{\mathbb{R}^4} dx (\partial^\mu \partial_\mu \phi + m^2 \phi)_x f(x), \quad \forall f \in \mathcal{S}(\mathbb{R}^4)$$

in the domain $D \subset \mathcal{H}$. Analogously, we understand all of the other differential equations that occur in the following analysis involving quantum fields.

Before stating the definition of $\phi = (\phi(f))_{f \in \mathcal{S}(\mathbb{R}^4)}$, let us outline the general strategy that leads to the construction of solutions of (3.6). Motivated by Problem 1.9, we interpret (3.1) as a Hamiltonian dynamics (with canonical coordinates $(\varphi_x)_{x \in \mathbb{R}^3}, (\pi_x)_{x \in \mathbb{R}^3}$)

$$\frac{d\varphi}{dt} = \pi, \quad \frac{d\pi}{dt} = -(-\Delta + m^2)\varphi \iff \frac{d\Phi}{dt} = \mathbf{J}H(\Phi), \quad (3.7)$$

where we define on the real Hilbert space $L^2(\mathbb{R}^3) \oplus L^2(\mathbb{R}^3)$ the operators

$$\mathbf{J} = \begin{pmatrix} 0 & \mathbf{1}_{L^2(\mathbb{R}^3)} \\ -\mathbf{1}_{L^2(\mathbb{R}^3)} & 0 \end{pmatrix}, \quad H(\varphi, \pi) = \begin{pmatrix} -\Delta + m^2 & 0 \\ 0 & \mathbf{1}_{L^2(\mathbb{R}^3)} \end{pmatrix}.$$

Problem 3.3. Consider the dynamics (3.7) and prove the following statements.

a) Setting $\omega(\nabla) = \sqrt{|\nabla|^2 + m^2}$, show that for every $\Phi = (\varphi, \pi) \in \mathcal{S}(\mathbb{R}^3) \times \mathcal{S}(\mathbb{R}^3)$ that

$$\mathbb{R} \ni t \mapsto T_t \Phi = \begin{pmatrix} \cos(\omega(\nabla)t) & \omega(\nabla)^{-1} \sin(\omega(\nabla)t) \\ -\omega(\nabla) \sin(\omega(\nabla)t) & \cos(\omega(\nabla)t) \end{pmatrix} \begin{pmatrix} \varphi \\ \pi \end{pmatrix} \in L^2(\mathbb{R}^3) \oplus L^2(\mathbb{R}^3)$$

is a strong solution of (3.7) with initial data Φ at $t = 0$.

b) Show that $(T_t)_{t \in \mathbb{R}}$ is strongly continuous family of bounded operators on $L^2(\mathbb{R}^3) \oplus L^2(\mathbb{R}^3)$ that satisfies $T_t T_s = T_{s+t}$, $\forall s, t \in \mathbb{R}$.

c) Define the bilinear form $\sigma : L^2(\mathbb{R}^3) \oplus L^2(\mathbb{R}^3) \times L^2(\mathbb{R}^3) \oplus L^2(\mathbb{R}^3) \rightarrow \mathbb{R}$ by

$$\sigma(\Phi_1, \Phi_2) = \langle \Phi_1, \mathbf{J} \Phi_2 \rangle = \langle \varphi_1, \pi_2 \rangle - \langle \pi_1, \varphi_2 \rangle$$

for $\Phi_1 = (\varphi_1, \pi_1), \Phi_2 = (\varphi_2, \pi_2)$. Prove that σ is invariant under the dynamics (3.7), that is $\sigma(T_t \Phi_1, T_t \Phi_2) = \sigma(\Phi_1, \Phi_2)$ for every $t \in \mathbb{R}$.

d) Prove that $T_t \mathbf{J}H = \mathbf{J}HT_t$ for every $t \in \mathbb{R}$.

The Hamiltonian formulation (3.7), Problem 3.3 and the canonical quantization method suggest to find the quantized free field by constructing operator-valued distributions $(\varphi(f))_{f \in \mathcal{S}(\mathbb{R}^3)}, (\pi(f))_{f \in \mathcal{S}(\mathbb{R}^3)}$ on some Hilbert space $\mathcal{H}$ that satisfy on a suitable, dense subspace the equal time commutation relations

$$[\varphi(f), \varphi(g)] = [\pi(f), \pi(g)] = 0, \quad [\varphi(f), \pi(g)] = i\langle f, g \rangle, \quad \forall f, g \in \mathcal{S}(\mathbb{R}^3) \quad (3.8)$$

and then to define for $\Phi = (\varphi, \pi)$ (in a suitable distributional sense)

$$\Phi_t(f) = (\varphi_t(f), \pi_t(f)) = T_t \Phi(f), \quad \forall f \in \mathcal{S}(\mathbb{R}^3).$$

This yields in particular an operator-valued distribution $(\phi(f))_{f \in \mathcal{S}(\mathbb{R}^4)}$, interpreted as the quantized free field, for which the Wightman axioms can be verified.

We make this construction rigorous in the setting of $\mathcal{H} = \mathcal{F}_s(L^2(\mathbb{R}^3, d\mathbf{p}))$. In the remainder, we consider all operator identities as identities on the dense domain $\mathcal{D}_{\mathcal{S}}$, introduced in Section 2.4. We need the following two preparations.

Lemma 3.3. *Let $(\varphi(f))_{f \in \mathcal{S}(\mathbb{R}^3)}, (\pi(f))_{f \in \mathcal{S}(\mathbb{R}^3)}$ be operator-valued distributions. Define*

$$\sigma(\Phi, F) = \varphi(g) - \pi(f), \quad \forall F = (f, g) \in \mathcal{S}(\mathbb{R}^3) \times \mathcal{S}(\mathbb{R}^3), \Phi = (\varphi, \pi).$$

Then, $(\varphi(f) = \sigma(\Phi, (0, f)))_{f \in \mathcal{S}(\mathbb{R}^3)}, (\pi(f) = -\sigma(\Phi, (f, 0)))_{f \in \mathcal{S}(\mathbb{R}^3)}$ satisfy (3.8) if and only if

$$[\sigma(\Phi, F_1), \sigma(\Phi, F_2)] = i\sigma(F_1, F_2), \quad \forall F_1, F_2 \in \mathcal{S}(\mathbb{R}^3) \times \mathcal{S}(\mathbb{R}^3).$$

Proof. On the one hand, we compute

$$[\sigma(\Phi, F_1), \sigma(\Phi, F_2)] = [\varphi(g_1) - \pi(f_1), \varphi(g_2) - \pi(f_2)] = -i\langle g_1, f_2 \rangle + i\langle f_1, g_2 \rangle = i\sigma(F_1, F_2).$$

On the other hand, given $(\sigma(\Phi, F))_{F \in \mathcal{S}(\mathbb{R}^3) \times \mathcal{S}(\mathbb{R}^3)}$ as above and defining $\varphi(f) = \sigma(\Phi, (0, f))$, $\pi(f) = -\sigma(\Phi, (f, 0))$ for every $f \in \mathcal{S}(\mathbb{R}^3)$, we find e.g. that

$$[\varphi(f), \pi(g)] = [\sigma(\Phi, (0, f)), -\sigma(\Phi, (g, 0))] = i\sigma((0, f), (g, 0)) = i\langle f, g \rangle, \quad \forall f, g \in \mathcal{S}(\mathbb{R}^3).$$

The verification of the remaining commutation relations in (3.8) is left as an *exercise*. $\square$

Lemma 3.4. *Define $\kappa : \mathcal{S}(\mathbb{R}^3) \times \mathcal{S}(\mathbb{R}^3) \rightarrow L^2(\mathbb{R}^3, d\mathbf{p})$ by*

$$\kappa(f, g) = \frac{1}{\sqrt{2}}(\sqrt{\omega} \widehat{f} + i\sqrt{\omega}^{-1} \widehat{g}),$$

where $\omega(\mathbf{p}) = \sqrt{|\mathbf{p}|^2 + m^2}$ for $\mathbf{p} \in \mathbb{R}^3$. Then, we have that

$$2\text{Im}\langle \kappa F_1, \kappa F_2 \rangle = \sigma(F_1, F_2), \quad \forall F_1, F_2 \in \mathcal{S}(\mathbb{R}^3) \times \mathcal{S}(\mathbb{R}^3).$$

Furthermore, recalling $(T_t)_{t \in \mathbb{R}}$ from Problem 3.3, we have that $\kappa T_t = e^{-i\omega t} \kappa$, $\forall t \in \mathbb{R}$.

Proof. For $F_1 = (f_1, g_2), F_2 = (f_2, g_2)$, we use Plancherel and verify that

$$\begin{aligned} 2\text{Im}\langle \kappa F_1, \kappa F_2 \rangle &= \text{Im}\langle \sqrt{\omega}\hat{f}_1 + i\sqrt{\omega}^{-1}\hat{g}_1, \sqrt{\omega}\hat{f}_2 + i\sqrt{\omega}^{-1}\hat{g}_2 \rangle \\ &= \text{Im} i\langle \sqrt{\omega}\hat{f}_1, \sqrt{\omega}^{-1}\hat{g}_2 \rangle + \text{Im}(-i)\langle \sqrt{\omega}^{-1}\hat{g}_1, \sqrt{\omega}\hat{f}_2 \rangle \\ &= \sigma(F_1, F_2). \end{aligned}$$

Note here that we used the assumption that F_1, F_2 are real-valued.

For the second statement, we observe that

$$\begin{aligned} \kappa T_t F &= \kappa \begin{pmatrix} \cos(\omega(\nabla)t)f + \omega(\nabla)^{-1} \sin(\omega(\nabla)t)g \\ -\omega(\nabla) \sin(\omega(\nabla)t)f + \cos(\omega(\nabla)t)g \end{pmatrix} \\ &= \frac{1}{\sqrt{2}}(\sqrt{\omega}(\cos(\omega t) - i \sin(\omega t))\hat{f} + i\sqrt{\omega}^{-1}(\cos(\omega t) - i \sin(\omega t))\hat{g}) = e^{-i\omega t} \kappa F \end{aligned}$$

for every $F = (f, g)$. □

We are now ready to define the free, scalar field. Denoting by $a(f), a^*(g)$ the standard creation and annihilation operators on $\mathcal{H}$, we first define

$$\sigma(\Phi, F) = i(a(\kappa F) - a^*(\kappa F)) = \frac{i}{\sqrt{2}}(a(\sqrt{\omega}\hat{f} + i\sqrt{\omega}^{-1}\hat{g}) - \text{h.c.}),$$

for every $F = (f, g) \in \mathcal{S}(\mathbb{R}^3) \times \mathcal{S}(\mathbb{R}^3)$. These are defined so that

$$[\sigma(\Phi, F_1), \sigma(\Phi, F_2)] = 2\text{Im}\langle \kappa F_1, \kappa F_2 \rangle = \sigma(F_1, F_2)$$

for every $F_1, F_2 \in \mathcal{S}(\mathbb{R}^3) \times \mathcal{S}(\mathbb{R}^3)$, by Lemma 3.4. In particular, the operators

$$\varphi(f) = \sigma(\Phi, (0, f)), \quad \varphi(f) = \sigma(\Phi, (f, 0)), \quad \forall f \in \mathcal{S}(\mathbb{R}^3)$$

define a family of canonical field operators satisfying (3.8), by Lemma 3.3. We then define a time-dependent family $(\sigma(\Phi_t, F))_{t \in \mathbb{R}, F \in \mathcal{S}(\mathbb{R}^3) \times \mathcal{S}(\mathbb{R}^3)}$ of operator-valued distributions by

$$\sigma(\Phi_t, F) = \sigma(\Phi, T_{-t}F), \quad \forall F \in \mathcal{S}(\mathbb{R}^3) \times \mathcal{S}(\mathbb{R}^3).$$

This can be understood as an operator-distributional solution of (3.7), because

$$\frac{d}{dt}\sigma(\Phi_t, F) = -\sigma(\Phi, \mathbf{J}HT_{-t}F) = -\sigma(\Phi_t, \mathbf{J}HF),$$

so that with $\varphi_t(f) = \sigma(\Phi_t, (0, f)), \pi_t = -\sigma(\Phi_t, (f, 0))$, we have symbolically that

$$\int_{\mathbb{R}^3} dx \left(\frac{d}{dt}\varphi_t(x) - \pi_t(x) \right) f(x) = 0, \quad \int_{\mathbb{R}^3} dx \left(\frac{d}{dt}\pi_t(x) + (-\Delta + m^2)\varphi_t(x) \right) f(x) = 0, \quad \forall f \in \mathcal{S}(\mathbb{R}^3).$$

We thus define the free, scalar quantum field $(\phi(f))_{f \in \mathcal{S}(\mathbb{R}^4)}$ of mass $m > 0$ by

$$\begin{aligned}\phi(f) &= \int_{\mathbb{R} \times \mathbb{R}^3} dt d\mathbf{x} \varphi_t(\mathbf{x}) f(t, \mathbf{x}) \\ &= \int_{\mathbb{R}} dt \sigma(\Phi_t, (0, f(t, \cdot))) \\ &= \int_{\mathbb{R}} \frac{dt}{\sqrt{2}} \left(a(e^{i\omega t} \sqrt{\omega}^{-1} \hat{f}(t, \cdot)) + a^*(e^{i\omega t} \sqrt{\omega}^{-1} \hat{f}(t, \cdot)) \right).\end{aligned}\tag{3.9}$$

Identifying heuristically $\phi(f) = \int_{\mathbb{R}^4} dx \phi_x f(x)$, this means that

$$\phi_x = \int_{\mathbb{R}^3} \frac{d\mathbf{p}}{\sqrt{2\pi}^3} \left(\frac{e^{-i(\omega(\mathbf{p})t - \mathbf{p}\mathbf{x})}}{\sqrt{2\omega(\mathbf{p})}} a_{\mathbf{p}} + \frac{e^{i(\omega(\mathbf{p})t - \mathbf{p}\mathbf{x})}}{\sqrt{2\omega(\mathbf{p})}} a_{\mathbf{p}}^* \right), \quad \forall x = (t, \mathbf{x}) \in \mathbb{R}^4. \tag{3.10}$$

3.2 Representations of the Poincaré Group

3.2.1 Finite Dimensional Unitary Representations of $\mathrm{SU}(2)$

3.2.2 Finite Dimensional Representations of $\mathrm{SL}(2, \mathbb{C})$

3.3 The Free Massive Vector Field

3.4 The Free Massive Dirac Field

4 Mathematical Interlude

In this chapter, we continue our discussion on the theory of distributions. Many of the general properties of quantum field theories rely on certain analyticity properties of the Wightman functions. The latter correspond to tempered distributions, equal to the boundary values of certain holomorphic functions. The connection to analytic functions enables us to use complex function theory to study their properties and the tool that links the Wightman functions to holomorphic functions is the Laplace transform. We therefore collect some general properties of the Laplace transform and holomorphic functions in the next Section 4.1. Afterwards, we discuss some refined properties, more specific to the applications in quantum field theory, in Section 4.2.

4.1 Laplace Transforms and Holomorphic Functions

A motivating example one may keep in mind for the following considerations is the map

$$x \mapsto g(x) = \int_0^\infty dx e^{2\pi i p x} f(p) \in \mathcal{S}(\mathbb{R}),$$

for some $f \in \mathcal{S}(\mathbb{R})$. Then, $g \in \mathcal{S}(\mathbb{R})$ coincides with the boundary value of the function

$$\mathbb{C}_+ = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\} \ni z \mapsto G(z) = \int_0^\infty dx e^{2\pi i p z} f(p)$$

That is, $g(x) = \lim_{y \rightarrow 0} G(x + iy)$, $\forall x \in \mathbb{R}$, and G is holomorphic in the upper half plane $\mathbb{C}_+$. In this section, we generalize this example to the setting of several complex variables and to boundary values that equal tempered distributions instead of smooth functions.

Let's recall a few basic facts from complex function theory. First, as a consequence of Cauchy's integral formula, $f : U \rightarrow \mathbb{C}$ is holomorphic in an open subset $U \subset \mathbb{C}$ if and only if it is analytic in U . That is, for all $z_0 \in U$, we can expand f into a power series

$$f(z) = \sum_{k=0}^{\infty} a_k (z - z_0)^k$$

which converges locally uniformly at $z_0 \in U$ (e.g. in a small disc $B_r(z_0)$ centered at $z_0 \in U$). The coefficients are uniquely determined by f and can be computed by

$$a_k = \frac{f^{(k)}(z_0)}{k!} = \frac{1}{2\pi i} \int_{\gamma} d\zeta \frac{f(\zeta)}{(\zeta - z_0)^{k+1}}$$

for a suitable positively oriented, closed loop γ centered at z_0 (e.g. $\gamma = \partial B_r(z_0)$).

Second, $f : U \rightarrow \mathbb{C}$ is holomorphic in U if and only if it is real differentiable in $U \subset \mathbb{R}^2$ and satisfies the Cauchy-Riemann equations, that is

$$\partial_{\bar{z}} f = \frac{1}{2}(\partial_x f + i\partial_y f) = 0 \iff \partial_x f_1 = \partial_y f_2, \partial_y f_1 = -\partial_x f_2,$$

with the conventions that $z = x + iy \in \mathbb{C}$, $x, y \in \mathbb{R}$ and $\operatorname{Re} f = f_1, \operatorname{Im} f = f_2$. In this case, the derivative of f is equal to $\partial_z f = \frac{1}{2}(\partial_x f - i\partial_y f)$.

In the sequel, we study functions that are holomorphic in several complex variables. In these notes, we employ only a few basic properties of functions holomorphic in several variables which allows us to use a hands on approach (we identify holomorphy with analyticity); for an introduction to the theory of functions of several complex variables, see e.g. [29, Chapter 7] and the references therein. We say that $f : U \rightarrow \mathbb{C}$ is holomorphic in an open subset $U \subset \mathbb{C}^n$ if it is complex analytic in U , i.e. for every $w \in U$, we have

$$f(z) = \sum_{k_1=0}^{\infty} \dots \sum_{k_n=0}^{\infty} a_{k_1 \dots k_n} (z_1 - w_1)^{k_1} \dots (z_n - w_n)^{k_n}$$

for all $z \in O$ in an open neighborhood $O \subset U$ around $w \in U$. By standard properties of power series, recall that this implies locally absolute and uniform convergence of the power series to f in a small, closed polydisc $P_r(w) = \{z \in \mathbb{C}^n : |z_j - w_j| \leq r, \forall j = 1, \dots, n\}$. Moreover, one can differentiate the power series locally term by term so that

$$a_{k_1 \dots k_n} = \frac{\partial_{z_1}^{k_1} \dots \partial_{z_n}^{k_n} f}{k_1! k_2! \dots k_n!} \Big|_{z=w}.$$

This implies that two holomorphic functions $f, g : U \rightarrow \mathbb{C}$ in an open, connected set $U \subset \mathbb{C}^n$ that agree on an open subset $D \subset \mathbb{R}^n$ with $D \subset U \cap \mathbb{R}^n$, agree in fact everywhere in U . Moreover, it implies that a function that is real analytic in an open subset $O \subset \mathbb{R}^n$ admits a unique, holomorphic extension on an open set $U \subset \mathbb{C}^n$ that contains $O \subset U$.

The following holomorphy criterion is used repeatedly below.

Lemma 4.1. *Let $U \subset \mathbb{C}^n$ be open. Then, $f : U \rightarrow \mathbb{C}$ is holomorphic in U if and only if it is continuous in U and holomorphic in each variable separately.*

Remark 4.1. *Remarkably, the equivalence remains true if the continuity of f is dropped, by Hartog's theorem (see e.g. [29, Chapter 7, Problem 1*] and the references therein).*

Proof. It is clear that analyticity (hence, continuity) of f implies analyticity in each variable separately. So, assume that f is continuous and holomorphic in each variable. Pick $w = (w_1, \dots, w_n) \in U$. By Cauchy's integral formula for holomorphic functions in one variable, we have locally in a polydisc $P_r(w) \subset U$ for small $r > 0$ that

$$f(z_1, \dots, z_n) = \frac{1}{2\pi i} \int_{\partial B_r(w_1)} \frac{d\zeta_1}{\zeta_1 - z_1} f(\zeta_1, z_2, \dots, z_n), \quad \forall z = (z_1, \dots, z_n) \in P_r(w).$$

Iterating the integral formula n times, we arrive at

$$\begin{aligned} f(z) &= \frac{1}{2\pi i} \int_{\partial B_r(w_1)} \frac{d\zeta_1}{\zeta_1 - z_1} \cdots \left(\frac{1}{2\pi i} \int_{\partial B_r(w_n)} \frac{d\zeta_n}{\zeta_n - z_n} f(\zeta_1, \zeta_2, \dots, \zeta_n) \right) \cdots \\ &= \frac{1}{(2\pi i)^n} \int_{\partial P_r(w)} \frac{d\zeta_1}{\zeta_1 - z_1} \cdots \frac{d\zeta_n}{\zeta_n - z_n} f(\zeta_1, \zeta_2, \dots, \zeta_n), \quad \forall z = (z_1, \dots, z_n) \in P_r(w), \end{aligned}$$

where the second step follows from integrability of the integrand (by continuity of f) and by Fubini's theorem. Now, expand

$$\prod_{j=1}^n \frac{1}{\zeta_j - z_j} = \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} \prod_{j=1}^n \frac{(z_j - w_j)^{k_j}}{(\zeta_j - w_j)^{k_j+1}},$$

then the power series on the r.h.s. converges absolutely and uniformly in $\zeta \in \partial P_r(w)$, for every $z \in (P_r(w))^\circ$. Inserting this into the integral above and exchanging integration with summation by uniform convergence of the integrand on $\partial P_r(w)$, we conclude that

$$f(z) = \sum_{k_1=0}^{\infty} \cdots \sum_{k_n=0}^{\infty} a_{k_1 \dots k_n} (z_1 - w_1)^{k_1} \cdots (z_n - w_n)^{k_n}, \quad \forall z \in (P_r(w))^\circ$$

for suitable coefficients $(a_{k_1 \dots k_n})_{k_1, \dots, k_n \in \mathbb{N}_0}$ in $\mathbb{C}$. Hence, $f : U \rightarrow \mathbb{C}$ is holomorphic. $\square$

A holomorphic function $f : U \rightarrow \mathbb{C}$ is clearly continuous as a $\mathbb{C}$ -valued map on $U \subset \mathbb{R}^{2n}$, viewed as a subset of $\mathbb{R}^{2n}$. In these notes, there are essentially two ways how we interpret f in a distributional sense. First, we can view it as a distribution $\Lambda_f \in \mathcal{D}'(U)$ on $U \subset \mathbb{R}^{2n}$ in the usual sense that

$$\begin{aligned} \Lambda_f(\varphi) &= \int_U dx_1 dy_1 \cdots dy_n dx_n f(z_1, \dots, z_n) \varphi(x_1, y_1, \dots, x_n, y_n) \\ &= \int_U dx_1 dy_1 \cdots dx_n dy_n f(x_1 + iy_1, \dots, x_n + iy_n) \varphi(x_1, y_1, \dots, x_n, y_n), \quad \forall \varphi \in \mathcal{D}(U). \end{aligned}$$

Second, when we discuss distributional boundary values of holomorphic functions, we consider instead (suitable subsets of) open subsets of the form $U = \mathbb{R}^n + iC \subset \mathbb{C}^n$ for some open, convex cone $C \subset \mathbb{R}^n$. In this situation, we interpret f as a family $(\Lambda_{f_y} = \Lambda_{f(\cdot + iy)})_{y \in C}$ of distributions in $\mathcal{D}'(\mathbb{R}^n)$, defined by

$$\Lambda_{f_y}(\varphi) = \int_{\mathbb{R}^n} dx_1 \dots dx_n f(x_1 + iy_1, \dots, x_n + iy_n) \varphi(x_1, \dots, x_n), \quad \forall \varphi \in \mathcal{D}(\mathbb{R}^n), y \in C.$$

In the sequel, we make these identifications (open subsets of $\mathbb{C}^n$ as open subsets of $\mathbb{R}^{2n}$, holomorphic functions $f : U \rightarrow \mathbb{C}$ as distributions) without further notice. It should always be clear from context in which distributional sense we interpret f .

Lemma 4.2. *Let $(f_k)_{k \in \mathbb{N}}$ be a sequence of holomorphic functions on an open subset $U \subset \mathbb{C}^n$. Then, $(f_k)_{k \in \mathbb{N}}$ converges in $\mathcal{D}'(U)$ if and only if $(f_k)_{k \in \mathbb{N}}$ converges compactly (as a sequence of holomorphic functions). In this case, $(f_k)_{k \in \mathbb{N}}$ converges to a holomorphic limiting function $f : U \rightarrow \mathbb{C}$.*

Proof. We recall from complex analysis that the Weierstrass convergence theorem (see e.g. [25]) states that if a sequence of holomorphic functions $(g_k)_{k \in \mathbb{N}}$ in $D \subset \mathbb{C}$ converges compactly to a function $g : D \rightarrow \mathbb{C}$ (that is, it converges uniformly on compact subsets of D to g), then g is holomorphic in D and the l -th derivatives $(g_k^{(l)})_{k \in \mathbb{N}}$ converge compactly to $g^{(l)}$ as well, for every $l \in \mathbb{N}$. Now, given a sequence of holomorphic functions $(f_k)_{k \in \mathbb{N}}$ that converges uniformly on compact subsets of $U \subset \mathbb{C}^n$ to some $f : U \rightarrow \mathbb{C}$, the locally uniform convergence implies that $f : U \rightarrow \mathbb{C}$ is continuous. Applying the Weierstrass convergence theorem in each variable separately, we conclude with Lemma 4.1 that $f : U \rightarrow \mathbb{C}$ is holomorphic in U , too. In other words, Weierstrass' theorem remains valid for holomorphic functions in several complex variables. Given such a sequence, we clearly have that $\lim_{k \rightarrow \infty} \Lambda_{f_k}(\varphi) = \Lambda_f(\varphi), \forall \varphi \in \mathcal{D}(U)$, i.e. $(\Lambda_{f_k})_{k \in \mathbb{N}}$ converges in $\mathcal{D}'(U)$.

On the other hand, suppose $(\Lambda_{f_k})_{k \in \mathbb{N}}$ converges in $\mathcal{D}'(U)$. Then, by Theorem 2.11, $\lim_{k \rightarrow \infty} \Lambda_{f_k} = \Lambda$ in $\mathcal{D}'(U)$ for some $\Lambda \in \mathcal{D}'(U)$. For $w \in U$ and $r = r_w > 0$ sufficiently small, we apply Cauchy's integral formula to write

$$f_k(z) = \int_0^1 dt_1 \dots \int_0^1 dt_n f_k(z_1 + re^{2\pi i t_1}, \dots, z_n + re^{2\pi i t_n}), \quad \forall z \in P_r(w) \subset U.$$

Now, pick some $g \in C_c^\infty((0, \epsilon))$ for $0 < \epsilon < r$, such that $\int_0^\infty dr r g(r) = 1$ and define

$$\varphi_w(x_1, y_1, \dots, x_n, y_n) = \prod_{j=1}^n g(|z_j - w_j|), \quad \forall (z_1, \dots, z_n) = (x_1, y_1, \dots, x_n, y_n) \in U \subset \mathbb{R}^{2n}.$$

Then $\varphi_w \in \mathcal{D}((P_\epsilon(w))^\circ) \subset \mathcal{D}(U)$ and a basic change of variables implies that

$$\begin{aligned} f_k(w) &= \int_0^1 dt_1 \dots \int_0^1 dt_n f_k(w_1 + re^{2\pi i t_1}, \dots, w_n + re^{2\pi i t_n}) \\ &\quad \times \int_0^\infty dr_1 \dots \int_0^\infty dr_n r_1 \dots r_n \varphi(r_1, \dots, r_n) \\ &= \int_U dx_1 dy_1 \dots dx_n dy_n f_k(x_1, y_1, \dots, x_n, y_n) \varphi_w(x_1, y_1, \dots, x_n, y_n) = \Lambda_{f_k}(\varphi_w). \end{aligned}$$

As $k \rightarrow \infty$, we conclude that $\lim_{k \rightarrow \infty} f_k(w) = \Lambda(\varphi_w) = f(w)$, for every $w \in U$. If $z \in P_\delta(w)$ varies in a sufficiently small, compact neighborhood $P_\delta(w)$, the set $E = \{\varphi_z \in \mathcal{D}(U) : z \in P_\delta(w)\} \subset \mathcal{D}(U)$ is bounded. Indeed, we find constants $C_j > 0$ such that

$$\sup_{\varphi \in E} |\varphi|_j \leq C_j, \quad \forall j \in \mathbb{N}_0$$

so that Theorem 2.9 *iii*) implies the boundedness of E . Applying once more Theorem 2.11, this shows that $(f_k)_{k \in \mathbb{N}}$ converges locally uniformly and hence compactly to f . In particular, by the first part of the proof, $f : U \rightarrow \mathbb{C}$ is holomorphic. $\square$

Now, let's define the Laplace transform of a distribution $\Lambda \in \mathcal{D}'(\mathbb{R}^n)$. For $\Lambda \in \mathcal{D}'(\mathbb{R}^n)$, we certainly have that $e^{-2\pi(\cdot)y}\Lambda \in \mathcal{D}'(\mathbb{R}^n)$, for every $y \in \mathbb{R}^n$, by the local boundedness of $e^{-2\pi(\cdot)y} \in C^\infty(\mathbb{R}^n)$. For some $y \in \mathbb{R}^n$, we might even have the property that $e^{-2\pi(\cdot)y}\Lambda \in \mathcal{S}'(\mathbb{R}^n)$. For such $y \in \mathbb{R}^n$, we define the Laplace transform $\mathcal{L}(\Lambda)_y \in \mathcal{S}'(\mathbb{R}^n)$ at $y \in \mathbb{R}^n$ by

$$\mathcal{L}(\Lambda)_y = (\widehat{e^{-2\pi(\cdot)y}\Lambda}) \iff \mathcal{L}(\Lambda)_y(\varphi) = \Lambda(e^{-2\pi(\cdot)y}\widehat{\varphi}), \quad \forall \varphi \in \mathcal{S}(\mathbb{R}^n).$$

If $\Lambda = \Lambda_f$ for some $f \in \mathcal{S}(\mathbb{R}^n)$, this means that $\mathcal{L}(\Lambda_f)_y$ satisfies

$$\mathcal{L}(\Lambda_f)_y(\varphi) = \int_{\mathbb{R}^n} dp f(p) e^{-2\pi p y} \widehat{\varphi}(p) = \int_{\mathbb{R}^n} dx \left(\int_{\mathbb{R}^n} dp e^{-2\pi i p(x-iy)} f(p) \right) \varphi(x)$$

for all $\varphi \in \mathcal{S}(\mathbb{R}^n)$. In other words, $\mathcal{L}(\Lambda_f)_y = \Lambda_{\mathcal{L}(f)(\cdot, y)}$, where $\mathcal{L}(f)$ is the usual (two-sided) Laplace transform with complex argument $z = x - iy \in \mathbb{C}^n$, $x, y \in \mathbb{R}^n$ (the minus sign in front of the imaginary part is conventional physics notation), i.e.

$$(x, y) \mapsto \mathcal{L}(f)(x, y) = \int_{\mathbb{R}^n} dp e^{-2\pi i p(x-iy)} f(p).$$

This suggests to say that $\mathcal{L}(\Lambda) = (\mathcal{L}(\Lambda)_y)_{y \in \Gamma}$ corresponds to a holomorphic function $f : U \rightarrow \mathbb{C}$ in the variable $z = x - iy \in \mathbb{R}^n - i\Gamma \subset \mathbb{C}^n$, for some open $\Gamma \subset \mathbb{R}^n$, if

$$\mathcal{L}(\Lambda)_y(\varphi) = \int_{\mathbb{R}^n} dx f(x - iy) \varphi(x), \quad \forall y \in \Gamma, \varphi \in \mathcal{S}(\mathbb{R}^n).$$

In the sequel, whenever $\mathcal{L}(\Lambda)$ corresponds to a holomorphic function in the above sense, we denote the function for simplicity again by $(x, y) \mapsto \mathcal{L}(\Lambda)(x, y) = \mathcal{L}(x - iy)$. It is useful to keep in mind that with our sign convention, $z = x - iy \mapsto f(z)$ is holomorphic in the j -th variable $z_j = x_j - iy_j$ if and only if $\partial_{x_j} f = i \partial_{y_j} f$.

In the rest of this section, we study the holomorphy properties of the Laplace transform of distributions in $\mathcal{D}'(\mathbb{R}^n)$ as defined above and we characterize those holomorphic functions which arise as such Laplace transforms. Afterwards, we collect refined results on tempered distributions with suitable support assumptions that are particularly relevant in the study of the Wightman functions of quantum fields. Before stating the first main result, we record two lemmas.

Lemma 4.3. *Let $\Lambda \in \mathcal{D}'(\mathbb{R}^n)$. Then, $\Gamma = \{y \in \mathbb{R}^n : e^{-2\pi(\cdot)y}\Lambda \in \mathcal{S}'(\mathbb{R}^n)\}$ is convex.*

Proof. Let $y_1, y_2 \in \Gamma$, $t \in [0, 1]$ and define $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ by

$$\varphi(p) = \frac{e^{-2\pi p(ty_1 + (1-t)y_2)}}{e^{-2\pi p y_1} + e^{-2\pi p y_2}}.$$

Clearly, $\varphi > 0$ on $\mathbb{R}^n$ and, by convexity of $x \mapsto e^{-2\pi p x}$, we have that $\varphi \leq 1$. Furthermore, it is clear that $\varphi \in C_b^\infty(\mathbb{R}^n)$ and that its derivatives are bounded, too (*exercise*). By Example 2.6, this shows $\varphi(\cdot)e^{-2\pi(\cdot)y_1}\Lambda \in \mathcal{S}'(\mathbb{R}^n)$, $\varphi(\cdot)e^{-2\pi(\cdot)y_2}\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ so that also

$$e^{-2\pi(\cdot)(ty_1 + (1-t)y_2)}\Lambda = \varphi(\cdot)e^{-2\pi(\cdot)y_1}\Lambda + \varphi(\cdot)e^{-2\pi(\cdot)y_2}\Lambda \in \mathcal{S}'(\mathbb{R}^n),$$

as a linear combination of tempered distributions. $\square$

We call a tempered distribution to be of fast decrease if it is equal to the product $\varphi\Lambda$ of some $\varphi \in \mathcal{S}(\mathbb{R}^n)$ and some $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$.

Lemma 4.4. *Let $\varphi\Lambda \in \mathcal{S}'(\mathbb{R}^n)$, for some $\varphi \in \mathcal{S}(\mathbb{R}^n)$ and $\Lambda \in \mathcal{S}'(\mathbb{R}^n)$, be a tempered distribution of fast decrease. Then, its Fourier transform $\widehat{\varphi\Lambda} \in \mathcal{S}'(\mathbb{R}^n)$ is equal to a polynomially bounded function $f \in C^\infty(\mathbb{R}^n)$, given by $f(x) = \Lambda(e^{-2\pi i x(\cdot)}\varphi)$, $\forall x \in \mathbb{R}^n$.*

Proof. Consider $\varphi\Lambda$ for some $\varphi \in \mathcal{S}(\mathbb{R}^n)$ and $\Lambda \in \mathcal{S}(\mathbb{R}^n)$. Then $\psi = \check{\varphi}(-\cdot) \in \mathcal{S}(\mathbb{R}^n)$, $\Sigma = \widehat{\Lambda} \in \mathcal{S}'(\mathbb{R}^n)$ and we have for every $\zeta \in \mathcal{S}(\mathbb{R}^n)$ that

$$(\varphi\Lambda)(\zeta) = \Lambda(\varphi\zeta) = \Sigma(\widetilde{\varphi\zeta}) = \Sigma(\psi(-\cdot) * \check{\zeta}) = (\psi * \Sigma)(\check{\zeta}) = \widetilde{\psi * \Sigma}(\zeta).$$

In other words, $\widehat{\varphi\Lambda} = \psi * \Sigma$. Now, recall that for every $\zeta \in \mathcal{S}(\mathbb{R}^n)$, it holds true that

$$(\psi * \Sigma)(\zeta) = \Sigma\left(\int_{\mathbb{R}^n} dx \psi(x - \cdot)\zeta(x)\right) = \int_{\mathbb{R}^n} dx \Sigma(\psi(x - \cdot))\zeta(x).$$

The second equality can be justified using an approximation argument that allows to consider w.l.o.g. to $\zeta, \psi \in \mathcal{D}(\mathbb{R}^n)$, a Riemann sum approximation applied to the integral on the r.h.s. and the smoothness of the function $\mathbb{R}^n \ni x \mapsto f(x) = \Sigma(\psi(x - \cdot)) \in C^\infty(\mathbb{R}^n)$. We leave the detailed proof as an *exercise* (a careful exposition of the key arguments can be found in e.g. [15, Chapter 6]). Hence, $\widehat{\varphi\Lambda} = \psi * \Sigma = \Lambda_f$ corresponds to a smooth function $f \in C^\infty(\mathbb{R}^n)$. Note that in terms of φ and Λ , we have explicitly that

$$f(x) = \Sigma(\psi(x - \cdot)) = \widehat{\Lambda}(\check{\varphi}(\cdot - x)) = \Lambda(e^{-2\pi i x(\cdot)}\varphi) \left(= \widehat{\Lambda}(\widehat{\varphi}(x - \cdot)) = (\widehat{\varphi} * \widehat{\Lambda})(x) \right).$$

By Theorem 2.15, $f \in C^\infty(\mathbb{R}^n)$, which is uniquely determined by $\widehat{\varphi\Lambda} \in \mathcal{S}'(\mathbb{R}^n)$, is necessarily polynomially bounded. By Prop. 2.5, this also follows directly from

$$|f(x)| = |\Lambda(e^{-2\pi i x(\cdot)}\varphi)| \leq C \sum_{\alpha, \beta \in \mathbb{N}_0^n : |\alpha|, |\beta| \leq m} |e^{-2\pi i x(\cdot)}\varphi|_{\alpha, \beta} \leq C(1 + |x|^m)$$

for some $m = m_\Lambda$ and, analogously, $|\partial^\alpha f| \leq C_\alpha(1 + |x|^m)$, for each $\alpha \in \mathbb{N}_0^n$. $\square$

The following result characterizes precisely those holomorphic functions that arise as Laplace transforms of tempered distributions.

Theorem 4.1. *Let $\Gamma \subset \mathbb{R}^n$ be a non-empty, convex, open subset of $\mathbb{R}^n$ and $\Lambda \in \mathcal{D}'(\mathbb{R}^n)$ be such that $e^{-2\pi(\cdot)y}\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ for every $y \in \Gamma$. Then $\mathcal{L}(\Lambda)$ corresponds to a holomorphic function in the tube $\mathbb{R}^n - i\Gamma$. This function has the property that for every compact $K \subset \Gamma$, there exists a polynomial $P_K : \mathbb{R}^n \rightarrow \mathbb{C}$ such that*

$$\sup_{y \in K} |\mathcal{L}(\Gamma)(x - iy)| \leq |P_K(x)|, \quad \forall x \in \mathbb{R}^n. \quad (4.1)$$

Conversely, if $f : \mathbb{R}^n - i\Gamma \rightarrow \mathbb{C}$ is holomorphic in the tube $\mathbb{R}^n - i\Gamma$ and satisfies the bound (4.1) as above, then it is the Laplace transform of a distribution $\Lambda \in \mathcal{D}'(\mathbb{R}^n)$ with the property that $e^{-2\pi(\cdot)y}\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ for every $y \in \Gamma$.

Remark 4.2. *In general, the polynomial boundedness (4.1) breaks down at the boundary of Γ . For $n = 1$ and $\Gamma = (0, \infty)$, consider for example the Laplace transform*

$$\mathbb{R} - i\Gamma \ni x - iy \mapsto \int_{\mathbb{R}} dp e^{2\pi ip(x-iy)} e^{-\frac{1}{2}p^2} = e^{-\frac{1}{2}(x-iy)^2},$$

which grows like $e^{\frac{1}{2}y^2}$ as $y \rightarrow \infty$ and thus faster than any polynomial.

Proof. Let $\Lambda \in \mathcal{D}'(\mathbb{R}^n)$ be such that $e^{-2\pi(\cdot)y}\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ for every $y \in \Gamma$. To identify $\mathcal{L}(\Lambda)$ with a holomorphic function, we first prove that $e^{-2\pi(\cdot)y}\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ is a linear combination of tempered distributions of fast decrease. By Lemma 4.4, this implies that we can identify $\mathcal{L}(\Lambda)$ with a smooth function in $\mathbb{R}^n \times \Gamma \subset \mathbb{R}^{2n}$. Afterwards, we verify that the latter is in fact holomorphic in $\mathbb{R}^n - i\Gamma$, by verifying the Cauchy-Riemann equations in each variable $z_j = x_j - iy_j$, $j \in \{1, \dots, n\}$, separately.

Let $y^{(1)}, \dots, y^{(k)} \in \Gamma$ be a set of vectors so that their convex hull $H = H_{y^{(1)}, \dots, y^{(k)}} \subset \Gamma$ has non-empty interior, $H^\circ \neq \emptyset$. Then, we define $\phi = \phi_{y^{(1)}, \dots, y^{(k)}} : \mathbb{R}^n \times H$ by

$$\phi(p, y) = \frac{e^{-2\pi py}}{\sum_{j=1}^k e^{-2\pi py^{(j)}}}.$$

Similarly as in the proof of Lemma 4.4, convexity implies (*exercise*) that

$$\sup_{(p, y) \in \mathbb{R}^n \times H} |\phi(p, y)| \leq 1.$$

More generally, it is straightforward to see that for every $(\alpha, \beta) \in \mathbb{N}_0^n \times \mathbb{N}_0^n$, we have that

$$\sup_{y \in H} |\partial^{(\alpha, \beta)} \phi(p, y)| \leq C(1 + |p|^{|\beta|})$$

for some C (that depends on H). Note that for every $y \in H$, we have that

$$e^{-2\pi(\cdot)y}\Lambda = \sum_{j=1}^k \phi(\cdot, y) e^{-2\pi(\cdot)y^{(j)}} \Lambda.$$

This representation of $e^{-2\pi(\cdot)y}\Lambda$ is not, yet, sufficient to apply Lemma 4.4, because $\phi(\cdot, y)$ is not a Schwartz function. However, for $y \in H^\circ$, it turns out that $\phi(\cdot, y)$ decays exponentially (as a motivation, consider e.g. the case $n = 1$, two points $0 < a < b$ such that $H_{a,b} = (a, b)$ and $y = a + t(b - a)$ for some $0 < t < 1$). To make this more precise, fix a compact set $K \subset \Gamma$. Then, we claim that there exists some $\epsilon > 0$ such that

$$\exp(\epsilon\sqrt{1 + |\cdot|^2})(e^{-2\pi(\cdot)y}\Lambda) \in \mathcal{S}'(\mathbb{R}^n), \forall y \in K. \quad (4.2)$$

Since $\zeta_{-\epsilon} = \exp(-\epsilon\sqrt{1 + |\cdot|^2}) \in \mathcal{S}(\mathbb{R}^n)$, this implies that $e^{-2\pi(\cdot)y}\Lambda$ is a tempered distribution of fast decrease which is used below to identify $\mathcal{L}(\Lambda)$ with a smooth function.

So, let's prove (4.2). To this end, fix some $y \in K$. Then for some $\epsilon > 0$ sufficiently small, we have that $\overline{B}_{2\epsilon}(y) \subset H_{y^{(1)}, \dots, y^{(k)}}^\circ = H^\circ$ for a suitable set of vectors in Γ . Choosing ϕ as above and defining

$$\psi_\epsilon(p, u) = \zeta_\epsilon(p)\phi(p, u) = \exp(\epsilon\sqrt{1 + |p|^2})\phi(p, u),$$

this implies

$$\begin{aligned} \sup_{(p,u) \in \mathbb{R}^n \times B_\epsilon(y)} |\psi_\epsilon(p, u)| &\leq \sup_{(p,u) \in \mathbb{R}^n \times B_\epsilon(y)} e^{\epsilon(1+|p|)} \phi(p, u) \\ &= e^\epsilon \sup_{(p,u) \in \mathbb{R}^n \times B_\epsilon(y)} \sup_{x \in \mathbb{R}^n: |2\pi x| \leq \epsilon} e^{-2\pi p x} \phi(p, u) \\ &\leq e^\epsilon \sup_{(p,u) \in \mathbb{R}^n \times B_{2\epsilon}(y)} \phi(p, u) \leq e^\epsilon. \end{aligned}$$

Similarly, using the bounds on ϕ discussed above, we conclude that the derivatives of ψ_ϵ are polynomially bounded in $\mathbb{R}^n$, uniformly in $u \in B_\epsilon(y)$, and with upper bounds that depend continuously on $y \in K$ (*exercise*). This uses the observation that the derivatives of ζ_ϵ are bounded by some polynomial times $\zeta_\epsilon \in \mathcal{S}(\mathbb{R}^n)$ itself. Since multiplication of tempered distributions by polynomially bounded functions yields tempered distributions, we conclude that for every $y \in K$, we find some open neighborhood $B_\epsilon(y) \subset \Gamma$ such that

$$\exp(\epsilon\sqrt{1 + |\cdot|^2})(e^{-2\pi(\cdot)u}\Lambda) = \sum_{j=1}^k \psi_\epsilon(\cdot, u) e^{-2\pi(\cdot)y^{(j)}} \Lambda \in \mathcal{S}'(\mathbb{R}^n), \forall u \in B_\epsilon(y).$$

Observe that this implies $\exp(\epsilon'\sqrt{1 + |\cdot|^2})(e^{-2\pi(\cdot)u}\Lambda) \in \mathcal{S}'(\mathbb{R}^n)$ for every $\epsilon' < \epsilon$ as well. As a consequence, we may cover $K \subset \Gamma$ by finitely many such open sets $(B_{\epsilon_j}(y_j))_{j=1}^l$, choose $\epsilon = \min\{\epsilon_1, \dots, \epsilon_l\}$ and conclude (4.2).

Based on (4.2) and Lemma 4.4, we obtain that for every compact $K \subset \Gamma$ and every $y \in K$, there exists a polynomially bounded function $f_y \in C^\infty(\mathbb{R}^n)$ such that $\mathcal{L}(\Lambda)_y = \Lambda_{f_y}$. The preceding analysis and Lemma 4.4 imply furthermore that

$$f_u(x) = f(x, u) = \sum_{j=1}^k \Lambda(\psi_\epsilon(\cdot, u) e^{-2\pi i(\cdot)(x - iy^{(j)})} \zeta_\epsilon) \quad (4.3)$$

locally in $u \in B_\epsilon(y)$, for suitable points $y^{(1)}, \dots, y^{(k)} \in \Gamma$ such that $\overline{B_{2\epsilon}}(y) \subset H_{y^{(1)}, \dots, y^{(k)}}^\circ$. The local identity (4.3) clearly implies the smoothness of f , viewed as a function in $\mathbb{R}^n \times B_\epsilon(y)$. Since $K \subset \Gamma$ and $y \in K$ were arbitrary, this justifies that $\mathcal{L}(\Lambda) \in C^\infty(\mathbb{R}^n \times \Gamma)$. Moreover, covering a given compact set $K \subset \Gamma$ by finitely many balls $(B_{\epsilon_j}(y_j))_{j=1}^l$ and using the continuous dependence of the polynomial upper bounds on the derivatives of ψ_ϵ on $y \in K$, mentioned above, it follows similarly as in Lemma 4.4 that

$$\sup_{y \in K} |\mathcal{L}(\Gamma)(x, y)| \leq |P_K(x)|, \quad \forall x \in \mathbb{R}^n.$$

Finally, in order to interpret $\mathcal{L}(\Gamma)(x, y) = \mathcal{L}(\Lambda)(x - iy)$ as a holomorphic function in $\mathbb{R}^n - i\Gamma$ with respect to $z = x - iy$, note that (4.3) and the definition of ψ_ϵ and ϕ imply

$$\partial_{x_j} \mathcal{L}(\Gamma)(x, y) = -2\pi i \sum_{j=1}^k \Lambda(\psi_\epsilon(\cdot, y)(\cdot)_j e^{-2\pi i(\cdot)(x - iy^{(j)})} \zeta_\epsilon) = i \partial_{y_j} \mathcal{L}(\Gamma)(x, y)$$

for every $j \in \{1, \dots, n\}$. That is, $\mathcal{L}(\Lambda)$ satisfies the Cauchy-Riemann equations in z_j for each $j \in \{1, \dots, n\}$ and, by Lemma 4.1, this concludes the proof of the first direction.

Conversely, suppose that $f : \mathbb{R}^n - i\Gamma \rightarrow \mathbb{C}$ is holomorphic and satisfies (4.1), for every compact $K \subset \Gamma$. We first recall that if f satisfies (4.1), analogous bounds hold true as well for its derivatives $\partial_{y_j} f$ with respect to $y \in \Gamma$. For simplicity of notation, let's focus on the case $n = 1$. The general case follows along the same lines and we leave the details as an *exercise*. In our application below, we only need the claimed polynomial boundedness on compact intervals $[a, b] \subset \Gamma$. By (4.1), we find some $k \in \mathbb{N}$ such that

$$\sup_{z \in \mathbb{R} - i[a, b]} |f(z)(z + i(a - \epsilon))^{-k}| \leq \sup_{x - iy \in \mathbb{R} - i[a, b]} |P_{[a, b]}(x)| |x - i(y - \epsilon - a)|^{-k} \leq C,$$

where $\epsilon > 0$ is chosen sufficiently small. Taking as a contour γ_L in $\mathbb{C}$ the concatenation of the straight paths from $L - ia$ to $-L - ia$, from $-L - ia$ to $-L - ib$, from $-L - ib$ to $L - ib$ and from $L - ib$ back to $L - ia$, and sending $L \rightarrow \infty$, the integrability of $\mathbb{R} \ni x \mapsto (1 + x)^{-2} \in \mathbb{R}$ and Cauchy's integral formula imply that

$$\begin{aligned} \frac{f(z)}{(z + i(a - \epsilon))^{k+2}} &= \lim_{L \rightarrow \infty} \frac{1}{2\pi i} \int_{\gamma_L} \frac{d\zeta}{(\zeta - z)} \frac{f(\zeta)}{(\zeta + i(a - \epsilon))^{k+2}} \\ &= \frac{1}{2\pi i} \int_{\gamma} \frac{d\zeta}{(\zeta - z)} \frac{f(\zeta)}{(\zeta - z + i(a - \epsilon))^{k+2}} \end{aligned}$$

for all $z \in \mathbb{R} - i(a, b)$, where the trace of γ in $\mathbb{C}$ is equal to $(\mathbb{R} - ia) \cup (\mathbb{R} - ib)$. As a consequence, the complex derivative $f' = \partial_x f + i \partial_y f$ satisfies

$$\frac{f'(z)}{(z + i(a - \epsilon))^{k+2}} = \frac{(k+2)f(z)}{(z + i(a - \epsilon))^{k+3}} + \frac{1}{2\pi i} \int_{\gamma} \frac{d\zeta}{(\zeta - z)^2} \frac{f(\zeta)}{(\zeta + i(a - \epsilon))^{k+2}}$$

so that, by the uniform bound (4.1), we get for every sufficiently small $\delta > 0$ that

$$\sup_{z \in \mathbb{R} - i[a+\delta, b-\delta]} |f'(z)(z + i(a - \epsilon))^{-k-2}| \leq C + \frac{C}{\delta^2} \int_{\gamma} \frac{d\zeta}{|\zeta + i(a - \epsilon)|^2} \leq C.$$

Since $a, b \in \mathbb{R}, \delta > 0$ were arbitrary, this proves that $\sup_{y \in [a, b]} |f'(\cdot - iy)|$ is polynomially bounded for each compact interval $[a, b] \subset \Gamma$ and, by the Cauchy-Riemann equations, the same statement applies to $\partial_y f$ (and $\partial_x f$).

Now, we seek $\Lambda \in \mathcal{D}'(\mathbb{R}^n)$ such that $e^{-2\pi(\cdot)y}\Lambda \in \mathcal{S}'(\mathbb{R}^n)$ and with the property that

$$\mathcal{L}(\Lambda)_y = (\widehat{e^{-2\pi(\cdot)y}\Lambda}) = \Lambda_{f(\cdot - iy)}, \quad \forall y \in \Gamma.$$

So, consider $\widehat{\Lambda_{f(\cdot - iy)}}$ and note that $\widehat{\Lambda_{f(\cdot - iy)}} \in \mathcal{S}'(\mathbb{R}^n)$ by the polynomial boundedness of $f(\cdot - iy)$. As a consequence, we have for every $y \in \Gamma$ that

$$\Lambda_y = e^{2\pi(\cdot)y} \widehat{\Lambda_{f(\cdot - iy)}} \in \mathcal{D}'(\mathbb{R}^n)$$

and our goal is to show that Λ_y does in fact not depend on $y \in \Gamma$. To see this, note that for every $\varphi \in \mathcal{D}(\mathbb{R}^n)$ the polynomial boundedness of the derivatives of $f(\cdot - iy)$ in all compact intervals in Γ , dominated convergence and integration by parts imply that

$$\begin{aligned} \partial_{y_j} \left(e^{2\pi(\cdot)y} \widehat{\Lambda_{f(\cdot - iy)}}(\varphi) \right) &= \partial_{y_j} \int_{\mathbb{R}^n} dx f(x - iy) (\widehat{e^{2\pi(\cdot)y}\varphi})(x) \\ &= \int_{\mathbb{R}^n} dx f(x - iy) (2\pi(\cdot)_j \widehat{e^{2\pi(\cdot)y}\varphi})(x) \\ &\quad + \int_{\mathbb{R}^n} dx (\partial_{y_j} f)(x - iy) (\widehat{e^{2\pi(\cdot)y}\varphi})(x) \\ &= \int_{\mathbb{R}^n} dx f(x - iy) (-i\partial_{x_j} (\widehat{e^{2\pi(\cdot)y}\varphi}))(x) \\ &\quad + \int_{\mathbb{R}^n} dx (\partial_{y_j} f)(x - iy) (\widehat{e^{2\pi(\cdot)y}\varphi})(x) \\ &= \int_{\mathbb{R}^n} dx (i\partial_{x_j} f + \partial_{y_j} f)(x - iy) (\widehat{e^{2\pi(\cdot)y}\varphi})(x) = 0, \end{aligned}$$

for all $j \in \{1, \dots, n\}$. Hence, $\Lambda_y(\varphi) = \Lambda_{y_0}(\varphi)$ for a fixed $y_0 \in \Gamma$ (recall that Γ is assumed to be open and convex) and all $y \in \Gamma, \varphi \in \mathcal{D}(\mathbb{R}^n)$. In other words, $\Lambda_y \in \mathcal{D}'(\mathbb{R}^n)$ does actually not depend on $y \in \Gamma$ so that $\Lambda = \Lambda_{y_0} \in \mathcal{D}'(\mathbb{R}^n)$ has the desired properties. $\square$

4.2 Analytic Continuation and Edge of the Wedge Theorem

5 General Properties of Quantum Field Theories

5.1 Wightman Functions and Reconstruction Theorem

5.2 Locality and Irreducibility

5.3 Spin and Statistics

5.4 The Interaction Picture and Haag's Theorem

6 Mathematical Interlude

- 6.1 Gaussian Measures on Locally Convex Vector Spaces
- 6.2 The Euclidean Massive Klein-Gordon Field
- 7 The Euclidean Approach to Quantum Field Theory
 - 7.1 Path Integral Formulation of Quantum Field Theories
 - 7.2 Osterwalder-Schrader Reconstruction Theorem
 - 7.3 Construction of the Euclidean ϕ_2^4 Field Theory
- 8 Basic Results on Lattice Yang-Mills Theories
 - 8.1 Gauge Theories in Physics
 - 8.2 Leading Order Partition Function of U(N) Lattice Yang-Mills

References

- [1] V. Bargmann. On Unitary Ray Representations of Continuous Groups. *Ann. of Math.* **59**, No. 2 (1954).
- [2] V. Bargmann. Note on Wigner's Theorem on Symmetry Operations. *J. Math. Phys.* **5**, pp. 862-868 (1964).
- [3] P. Billingsley. Convergence of Probability Measures. 2nd Edition. *Wiley Series in Probability and Statistics*, Wiley (1990).s
- [4] D. L. Cohn. Measure Theory (2nd Edition). *Birkhäuser Advanced Texts*, Birkhäuser (2013).
- [5] J. Dimock. Quantum Mechanics and Quantum Field Theory. A Mathematical Primer. *Cambridge University Press* (2011).
- [6] I. M. Gel'fand, N. Ya. Vilenkin. Generalized Functions Vol. 4: Applications of Harmonic Analysis. *Academic Press* (1964).
- [7] B. C. Hall. Quantum Theory for Mathematicians. Graduate Texts in Mathematics **267**, Springer (2013).
- [8] B. C. Hall. Lie Groups, Lie Algebras, and Representations. 2nd Edition. Graduate Texts in Mathematics **222**, Springer (2015).
- [9] J. Hilgert, K.-H. Neeb. Structure and Geometry of Lie Groups. *Springer New York* (2012).
- [10] J. Horváth. Topological Vector Spaces and Distributions. *Dover Publications, Inc.* (2012).
- [11] L. D. Landau, E. M. Lifschitz. Mechanics. Course of Theoretical Physics, Volume 1 (2nd Edition). *Pergamon Press* (1969).
- [12] L. D. Landau, E. M. Lifschitz. The Classical Theory of Fields. Course of Theoretical Physics, Volume 2 (2nd Edition). *Pergamon Press* (1959).
- [13] L. D. Landau, E. M. Lifschitz. Quantum Mechanics: Non-Relativistic Theory. Course of Theoretical Physics, Volume 3 (2nd Edition). *Pergamon Press* (1965).
- [14] J. M. Lee. Introduction to Smooth Manifolds. *Graduate Texts in Mathematics*, Springer (2000).
- [15] E. H. Lieb, M. Loss. Analysis. *Graduate Studies in Mathematics* **14**, American Mathematical Society (2001).
- [16] W. Miller, Jr. Symmetry Groups and Their Applications. *Academic Press, Inc.* (1972).

- [17] W. Moretti. Spectral Theory and Quantum Mechanics: Mathematical Foundations of Quantum Theories, Symmetries and Introduction to the Algebraic Formulation. 2nd Edition. *Springer Cham* (2018).
- [18] M. Nakahara. Geometry, Topology and Physics. 2nd Edition. Graduate Student Series in Physics. *Taylor & Francis* (2003).
- [19] B. O'Neill. Semi-Riemannian Geometry. With Applications to Relativity. *Academic Press* (1983).
- [20] E. Nelson. Analytic vectors. *Ann. Math.* **70** (1959).
- [21] M. Reed, B. Simon. Methods of Modern Mathematical Physics I: Functional Analysis. *Academic Press, Inc.* (1972).
- [22] M. Reed, B. Simon. Methods of Modern Mathematical Physics II: Fourier Analysis, Self-Adjointness. *Academic Press, Inc.* (1975).
- [23] M. Reed, B. Simon. Methods of Modern Mathematical Physics III: Scattering Theory. *Academic Press, Inc.* (1979).
- [24] M. Reed, B. Simon. Methods of Modern Mathematical Physics IV: Analysis of Operators. *Academic Press, Inc.* (1978).
- [25] R. Remmert. Theory of Complex Functions. *Graduate Texts in Mathematics* **122**, Springer (1991).
- [26] W. Rudin. Functional Analysis. 2nd Edition. *International Series in Pure and Applied Mathematics*, McGraw Hill Education (India) (2006).
- [27] R. U. Sexl, H. K. Urbantke. Relativität, Gruppen, Teilchen. 3rd Edition. *Springer* (1992).
- [28] C. D. Sogge. Lectures on Non-Linear Wave Equations. 2nd Edition. *International Press* (2008).
- [29] E. M. Stein, R. Shakarchi. Princeton Lectures in Analysis IV. Functional Analysis. Introduction to Further Topics in Analysis. *Princeton University Press* (2011).
- [30] R. F. Streater, A. S. Wightman. PCT, Spin and Statistics, and All That. Princeton Landmarks in Mathematics and Physics, *Princeton University Press* (2000).
- [31] M. Talagrand. What is a Quantum Field Theory? A First Introduction for Mathematicians. *Cambridge University Press* (2022).
- [32] R. M. Wald. General Relativity. *The University of Chicago Press* (1984).
- [33] S. Weinberg. The Quantum Theory of Fields, Volume 1: Foundations. *Cambridge University Press* (2005).

- [34] E. Wigner. Gruppentheorie und ihre Anwendungen auf die Quantenmechanik der Atomspektren. *Springer* (1931).
- [35] E. Wigner. On Unitary Representations of the Inhomogeneous Lorentz Group. *Ann. of Math.* **40**, No. 1 (1939).