Klaas Landsman

Introductory Mathematical Physics

June 18, 2018

Springer

Contents

Mathematical Physics: historical introduction

The aim of mathematical physics lies in the cross-fertilization of mathematics and physics, to the mutual benefit of both. Ideally, the frontier of mathematics thereby intersects cutting edge research in physics, but also, perhaps more conservatively, known mathematics may be used in order to provide a rigorous formulation and foundation of the laws of physics. Typically, however, physics turns out to be a little ahead of mathematics, so that the latter has to move forward to keep track of the former, enriching both fields. The fact that so far this has always been possible is quite remarkable, suggesting that the Universe indeed has an underlying mathematical structure, as probably first argued (at least in Western Civilization) by ancient Greek thinkers such as Pythagoras (-572– -500)and Plato (-427– -347).

Perhaps Johannes Kepler (1571-1630) should be labeled as the first mathematical physicist in history, since he was certainly exceptionally talented in both fielExrds (not to speak of his epoch-making genius and persistence in astronomy, for which he is mainly remembered), and he did make an effort towards the aims just described. However, this effort was slightly off (as in his attempt to explain the known planetary orbits in terms of the five Platonic solids), so that Christiaan Huygens (1629–1695) was arguably the first really *successful* mathematical physicist (the mathematics of Galilei (1564–1642) was too primitive to be mentioned in this context, although he famously was among the first to recognize and promote the supporting role of mathematics for physics). Huygens brilliantly combined novel constructions in Euclidean geometry (as well as some pre-calculus) with inventions in optics, mechanics, and astronomy, whence it is highly appropriate that the Mathematical Physics Department op IMAPP is housed in the Huygens Building of the Faculty of Science.

However, both were eclipsed by the towering figure of Isaac Newton (1642–1727), whose status as the greatest scientist of all times somewhat masks the fact that he was first and foremost a *mathematical physicist*. Indeed, as the inventor of the Calculus he paved the way for modern mathematics, whilst as the author of *Principia* he virtually single-handedly created classical physics, including the derivation of Kepler's Laws from his own laws of gravity and mechanics (an achievement whose importance and brilliance cannot be overestimated, both conceptually and technically). Moreover, his mathematics and his physics were clearly inseparable.

Also after Newton's death, breakthroughs typically took place simultaneously in both areas, and throughout the 18th and 19th centuries progress at both fronts often resulted from the work of a single scientist, such as, Euler, Lagrange, Laplace, Fourier, Gauss, Poisson, Cauchy, Jacobi, Hamilton, Riemann, and Poincaré. In particular, the theory of partial differential equations, which at the time comprised at least half of mathematics and practically all of mathematical physics, largely evolved from physical applications to vibrating strings, fluid mechanics, heat, etc.

As beautifully described in the book *Plato's Ghost* by Jeremy Gray (Princeton University Press, 2008), in the late 19th century, mathematics underwent a 'modernist' turn and became independent from its application, with its own, internal criteria of relevance and truth. This has led to the creation of modern versions of e.g. logic, algebraic geometry, and algebraic topology, as well as to closely related new fields like category theory. The leading figure in this modernist transformation was undoubtedly David Hilbert (1862–1943), one of the greatest mathematicians of all times, but, seemingly paradoxically given his central role in this move towards complete abstraction, also one of the founders of modern mathematical physics.

First, Hilbert played a key role in providing the mathematical underpinning of Einstein's Theory of General Relativity from 1915, both directly and through his pupil Hermann Weyl (1885–1955), whose book *Raum - Zeit - Materie: Vorlesungen über allgemeine Relativitätstheorie* from 1917 is still a classic. Note that General Relativity was mathematically based on Riemann's revolutionary theory of (metric) differential geometry developed 65 years earlier, and could not possibly have been found without it (in developing his theory, Einstein received substantial help from the mathematician Marcel Grossmann). Apart from giving us a physical theory of breathtaking beauty, which continues to describe the observational Universe to the present day with astonishing accuracy, this also launched a period of intense interaction between geometry and physics that has continued unabated ever since.

Second, immediately after the successful development of quantum mechanics as a physical theory by Planck, Einstein, Bohr, Heisenberg, Schrödinger, Born, Jordan, Dirac, and others during the period 1900–1927, Hilbert initiated the study of its mathematical structure, which, as he saw at once, turned out to be grounded in functional analysis (especially in the theory of what we now call Hilbert spaces). His pioneering work in this direction was brilliantly completed by Weyl, notably in his book *Gruppentheorie und Quantenmechanik* from 1928, and, even more impressively, by Hilbert's assistant John von Neumann (1903–1957) at the time, whose book *Mathematische Grundlagen der Quantenmechanik* from 1932 describes the mathematical formalism of quantum mechanics as we still use it today.

Thus modern mathematical physics (which has continued to study the mathematical structure of quantum theory and general relativity) heavily relies on, and in turn continues to influence, areas of mathematics such as differential geometry, functional analysis, Lie theory, representation theory, and stochastics. For example, the theory of operator algebras (founded by von Neumann in the 1930s and Gelfand in the 1940s, and most recently culminating in the 'noncommutative geometry' of Alain Connes) has been developed in close connection with attempts to establish a mathematically and conceptually rigorous foundation for quantum field theory.

Chapter 1
Poisson brackets and Hamilton formalism

We do not assume familiarity with differential geometry and analysis on manifolds, so in what follows one may assume that $M = \mathbb{R}^k$ for some k. However, whenever possible we will phrase definitions and results in such a way that their more general meaning should be clear to those who *are* familiar with differential geometry etc.

Throughout this chapter, all vector spaces are defined over the real numbers.

1.1 Vector fields and their flows

An old-fashioned ***vector field*** on $M = \mathbb{R}^k$ is a map $X : \mathbb{R}^k \to \mathbb{R}^k$, given by components $X(x) = (X^1(x), \ldots, X^k(x))$ that describe something like a hyper-arrow at x. We therefore define the set of vector fields on M as[1]

$$\mathrm{Vec}(M) = C^\infty(M, \mathbb{R}^k). \tag{1.1}$$

This is not just a set: it is also:

- A real vector space under pointwise operations, viz. $(X+Y)(x) = X(x) + Y(x)$, and scalar multiplication $(\lambda X)(x) = \lambda X(x)$, $\lambda \in \mathbb{R}$, so that inverse of X is $-X$.
- A Lie algebra[2] under the bracket

$$[X,Y]^i = \sum_j \left(X^j \frac{\partial Y^i}{\partial x^j} - Y^j \frac{\partial X^i}{\partial x^j} \right); \tag{1.2}$$

- A $C^\infty(M)$ module[3] under the action $C^\infty(M) \times \mathrm{Vec}(M) \to \mathrm{Vec}(M)$ given by

[1] For an arbitrary manifold one should define $\mathrm{Vec}(M) = \Gamma(TM)$, the smooth sections of the tangent bundle TM to M, often also called $\mathfrak{X}(M)$. In general this is no longer of the form $C^\infty(M, \mathbb{R}^k)$.

[2] A ***Lie algebra*** over $\mathbb{R}$ is a vector space over $\mathbb{R}$ equipped with a bilinear map $[\cdot,\cdot] : A \times A \to A$ that satisfies $[a,b] = -[b,a]$ as well as $[a,[b,c]] + [c,[a,b]] + [b,[c,a]] = 0$ for all $a,b,c \in A$.

[3] Recall that an ***algebra*** A over $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$) is simply a vector space over $\mathbb{K}$ equipped with a bilinear map $\cdot : A \times A \to A$ (i.e., $\cdot$ is linear in each entry); we write $ab \equiv a \cdot b \equiv \cdot(a,b)$. We say that

$$(f \cdot X)(x) = f(x)X(x). \tag{1.3}$$

Here we regard $C^\infty(M) \equiv C^\infty(M,\mathbb{R})$ as a commutative algebra under pointwise operations, that is, $(\lambda f + \mu g)(x) = \lambda f(x) + \mu g(x)$ and $(fg)(x) = f(x)g(x)$. A vector field X defines a first-order differential operator $\delta_X : C^\infty(M) \to C^\infty(M)$ by

$$\delta_X f(x) = \sum_{j=1}^{k} X^j(x) \frac{\partial f(x)}{\partial x^j}. \tag{1.4}$$

To make the idea precise that a vector field on M is really the same as a first-order differential operator on $C^\infty(M)$, we note that it easily follows from (1.4) that

$$\delta_X(fg) = \delta_X(f)g + f\delta_X(g), \text{ for all } g \in C^\infty(M). \tag{1.5}$$

More generally, a ***derivation*** of an algebra A is a linear map $\delta : A \to A$ satisfying

$$\delta(ab) = \delta(a)b + a\delta(b), \text{ for all } a \in A. \tag{1.6}$$

Thus any vector field X on M gives a derivation δ_X of the algebra $C^\infty(M)$ by (1.4). Let $\mathrm{Der}(A)$ be the set of all derivations on an algebra A. Then $\mathrm{Der}(C^\infty(M))$ is:

- A real vector space under pointwise operations: $(\delta_1 + \delta_2)(f) = \delta_1(f) + \delta_2(f)$), scalar multiplication is $(\lambda\delta)(f) = \lambda \cdot \delta(f)$, $\lambda \in \mathbb{R}$, hence the inverse of δ is $-\delta$.
- A Lie algebra under the bracket

$$[\delta_1, \delta_2] = \delta_1 \circ \delta_2 - \delta_2 \circ \delta_1; \tag{1.7}$$

- A $C^\infty(M)$ module under the natural map $C^\infty(M) \times \mathrm{Der}(C^\infty(M)) \to \mathrm{Der}(C^\infty(M))$ given by $(g\delta)(f) = gf$.

Theorem 1.1. *For any manifold M, each derivation of $C^\infty(M)$ takes the form* (1.4)*, at least locally (for $M = \mathbb{R}^k$ also globally). This gives an isomorphism*

$$\mathrm{Vec}(M) \xrightarrow{\cong} \mathrm{Der}(C^\infty(M)); \tag{1.8}$$
$$X \mapsto \delta_X \tag{1.9}$$

of vector spaces, Lie algebras, and $C^\infty(M)$ modules, that is (for all $X, Y \in \mathrm{Vec}(M)$):

$$\delta_{\lambda X + \mu Y} = \lambda\delta_X + \mu\delta_Y \ \ (\lambda, \mu \in \mathbb{R}); \tag{1.10}$$
$$[\delta_X, \delta_Y] = \delta_{[X,Y]}; \tag{1.11}$$
$$\delta_{fX} = f\delta_X, \ (f \in C^\infty(M)). \tag{1.12}$$

The proof for $M = \mathbb{R}^k$ is an exercise. We often identify vector fields with derivations.

A is ***commutative*** if $ab = ba$ for all $a, b \in A$, and ***associative*** if $(ab)c = a(bc)$ for all $a, b, c \in A$. All algebras we look at are associative by definition, *except Lie algebras*. A ***module*** over an algebra A is a vector space V with a bilinear map $A \times V \to V$, written $(a, v) \mapsto av$, such that $a(bv) = (ab)v$.

Vector fields (or, equivalently, derivations) may be 'integrated', at least *locally*, in the following sense. First, a ***curve*** through $x_0 \in M$ is a smooth map $c : I \to M$, where $I \subset \mathbb{R}$ is open and $c(t_0) = x_0$ for some $x_0 \in I$. We usually assume that $0 \in I$ with $t_0 = 0$ and hence $c(0) = x_0$. We then say that c ***integrates*** X near x_0 if

$$\dot{c}(t) = X(c(t)), \tag{1.13}$$

a symbolic equality that can be interpreted in two equivalent ways:

- Describing $c : I \to \mathbb{R}^k$ by k functions $c^j : I \to \mathbb{R}$ ($j = 1, \ldots, k$), eq. (1.13) stands for
$$\frac{dc^j(t)}{dt} = X^j(c^1(t), \ldots, c^k(t)), \; j = 1, \ldots, k. \tag{1.14}$$
- More abstractly, eq. (1.13) means that for any $f \in C^\infty(M)$ we have
$$\delta_X f(c(t)) = \frac{d}{dt} f(c(t)). \tag{1.15}$$

To pass from (1.15) to (1.14), we just have to recall (1.4), and note that

$$\frac{d}{dt} f(c(t)) = \frac{d}{dt} f(c^1(t), \ldots, c^k(t)) = \sum_{j=1} \frac{dc^j(t)}{dt} \frac{\partial f(c(t))}{\partial x^j}.$$

The theory of ordinary differential equations shows that such local integral curves exist near any point $x_0 \in M$, and that they are unique in the sense that if two curves $c_1 : I_1 \to M$ and $c_2 : I_2 \to M$ both satisfy (1.13) with $c_1(0) = c_2(0) = x_0$, then $c_1 = c_2$ on $I_1 \cap I_2$. However, curves that integrate X near some point may not be defined for all t, i.e., for $I = \mathbb{R}$. This makes the important concept of a ***flow*** of a vector field X, which is meant to encapsulate all integral curves of X, a bit complicated.

We start with the simplest case. We say that a vector field X is ***complete*** if for any $x_0 \in M$ there is a curve $c : \mathbb{R} \to M$ satisfying (1.13) with $c(0) = x_0$. An example of a complete vector field is $M = \mathbb{R}$ and $\delta_X = d/dx$, so that $\varphi_t(x) = x + t$. A *sufficient* condition for a (smooth) vector field X to be complete is that it has *compact support*, in that the functions $x \mapsto X^i$ vanish outside a compact (i.e. closed and bounded) subset of M (so if M is compact, then every vector field is complete). See exercises.

For an incomplete example, take $M = (0, 1)$ with the same X, or see the exercises.

Definition 1.1. *Let M be a manifold and let $X \in \mathrm{Vec}(M)$ be a complete vector field. A* **flow** *of X is a smooth map $\varphi : \mathbb{R} \times M \to M$, written $\varphi_t(x) \equiv \varphi(t, x)$, that satisfies*

$$\varphi_0(x) = x; \tag{1.16}$$
$$\delta_X f(\varphi_t(x)) = \frac{d}{dt} f(\varphi_t(x)), \; \forall t \in \mathbb{R}, x \in M, f \in C^\infty(M). \tag{1.17}$$

In local coordinates, where $\varphi_t(x) = (\varphi_t^1(x), \ldots, \varphi_t^k(x))$, eq. (1.17) becomes

$$\frac{d\varphi_t^j(x)}{dt} = X^j(\varphi_t(x)), \; j = 1, \ldots, k. \tag{1.18}$$

Therefore, analogously to (1.13), eq. (1.17) may symbolically be written as

$$X(\varphi_t(x)) = \frac{d}{dt}\varphi_t(x), \;\; \forall t \in \mathbb{R}, x \in M. \tag{1.19}$$

Thus the flow φ of X gives the integral curve c of X through x_0 by $c(t) = \varphi_t(x_0)$.

According to the theory of ordinary differential equations, notably the Picard–Lindelöf Theorem, any complete vector field has a unique flow. Moreover, the uniqueness part of this theorem implies the composition rule

$$\varphi_s \circ \varphi_t = \varphi_{s+t}. \tag{1.20}$$

From a group-theoretic point of view, conditions (1.16) - (1.20) therefore state that a flow is a smooth action of $\mathbb{R}$ (as an additive group) on M *that in addition integrates* X. In particular, (1.20) implies $\varphi_{-t} = \varphi_t^{-1}$, so that each $\varphi_t : M \to M$ is an invertible smooth map with smooth inverse. In other words, each φ_t is a *diffeomorphism* of M.

In particular, M is a disjoint union of the integral curves of X, which can never cross each other because of the uniqueness of the solution of (1.13) with $c(0) = x_0$.

If X is not complete, we do the best we can by defining the set

$$D_X = \{(t,x) \in \mathbb{R} \times M \mid \exists c : I \to M, c(0) = x, \dot{c}(t) = X(c(t)), t \in I\}. \tag{1.21}$$

cf. (1.13) and subsequent text. Obviously $\{0\} \times M \subset D_X$, and (less trivially) it turns out that D_X is open. Then a flow of X is a map $\varphi : D_X \to M$ that satisfies (1.16) for all x and eq. (1.17) for $(t,x) \in D_X$. Eq. (1.20) then holds whenever it is defined. The idea is to start with each $x \in M$, find some maximal I, and hence include all $t \in I$.

Exercises for week 1 (inleveropgaven: 1, 5)

1. Prove Theorem 1.1 for $M = \mathbb{R}^k$. Hint: given δ, define $X^i = \delta(x^i)$.
2. Show that $\mathrm{Vec}(M)$ and $\mathrm{Der}(C^\infty(M))$ are Lie algebras under the natural brackets (1.2) and (1.7), respectively. Why is $\delta_1 \circ \delta_2$ not, in general, a derivation?
3. For any real algebra A (or just for $A = M_n(\mathbb{R})$, the real $n \times n$ matrices), show that:
   a. The formula $[a,b] = ab - ba$ defines a Lie bracket on A;
   b. For fixed $b \in A$, the formula $\delta(a) = [a,b]$ defines a derivation on A.
4. Take $M = \mathbb{R}$ and show that the vector field $X(x) = x^2 d/dx$ is incomplete (in other words, show that for suitable $x_0 \in \mathbb{R}$ the differential equation $\dot{x} = x^2$ with initial condition $x(0) = x_0$ fails to have a solution for all $t \in \mathbb{R}$). If you find it easier to integrate, take $\dot{x} = x^2 + 1$ (such equations describe population growth in a situation where reproduction is proportional to the number of pairs of individuals).
5. Show that a vector field X on $M = \mathbb{R}^k$ with compact support (in the sense that the set $\{x \in M \mid X(x) \neq 0\}$ is bounded) is complete (i.e. for any $x_0 \in M$, the equations $\dot{x}^j = X^j x)$ with initial condition $x^j(0) = x_0^j$ have a solution for all $t \in \mathbb{R}$.

Hint: Start from the following fact about ordinary differential equations:
For any $x_0 \in M$ *there exist* $\varepsilon > 0$ *and a neighbourhood* U *of* x_0 *such that the flow* $\varphi : (-\varepsilon, \varepsilon) \times U \to M$ *of* X *is defined (i.e., for any* $x \in U$ *and* $t \in (-\varepsilon, \varepsilon)$ *the point* $\varphi_t(x)$ *is defined and* (1.16) - (1.20) *and* (1.17) *hold, the second one whenever* $|t+s| < \varepsilon$*).* This gives a cover of M by open sets. Choose a finite subcover (U_i) of the compact support of X, and take $\varepsilon_0 = \min\{\varepsilon_i\}$. Now represent $t \in \mathbb{R}$ (uniquely) as $t = \frac{1}{2}n\varepsilon_0 + r$, where $n \in \mathbb{Z}$ and $0 \leq r \leq \varepsilon_0/2$, and define $\varphi_t = (\varphi_{\varepsilon_0/2})^n \varphi_r$. Show that (1.16) - (1.20) and (1.17) are satisfied for all $x \in M$ and $t \in \mathbb{R}$.

1.2 Poisson manifolds and Hamiltonian Mechanics

Classical mechanics started with the *Principia* (1687) by Newton and the equations named after him (first written down by Euler). In the 19th century these were generalized and brought into what we now call *Hamiltonian form* by Hamilton, whose formalism was further generalized in the 20th century into *Poisson geometry*. The older field of symplectic geometry (which until the 1980's was seen as the correct mathematical setting for classical mechanics) is a special case of Poisson geometry, whose mathematical structure is actually simpler than that of symplectic geometry.

Definition 1.2. *A* **Poisson bracket** *on a manifold M is a Lie bracket* $\{-,-\}$ *on (the real vector space)* $C^\infty(M)$*, such that for each* $h \in C^\infty(M)$ *the map*

$$\delta_h : f \mapsto \{h, f\} \tag{1.22}$$

is a vector field on M (or, equivalently, a derivation of $C^\infty(M,\mathbb{R})$ *with respect to its structure of a commutative algebra under pointwise multiplication). A manifold M equipped with a Poisson bracket is called a* **Poisson manifold***,* $(C^\infty(M),\{\,,\,\})$ *is called a* **Poisson algebra***, and* δ_h *is called the* **Hamiltonian vector field** *of h.*

Unfolding, we have a bilinear map $\{-,-\} : C^\infty(M)\times C^\infty(M) \to C^\infty(M)$ that satisfies

$$\{g, f\} = -\{f, g\}; \tag{1.23}$$

$$\{f,\{g,h\}\}+\{h,\{f,g\}\}+\{g,\{h,f\}\} = 0; \tag{1.24}$$

$$\{f, gh\} = \{f,g\}h + g\{f,h\}. \tag{1.25}$$

The simplest and oldest example is $M = \mathbb{R}^2$, with coordinates $x = (p,q)$ and bracket

$$\{f,g\} = \frac{\partial f}{\partial p}\frac{\partial g}{\partial q} - \frac{\partial f}{\partial q}\frac{\partial g}{\partial p}, \tag{1.26}$$

so that, using (1.4). the Hamiltonian vector field of $h \in C^\infty(\mathbb{R}^2)$ is given by

$$\delta_h = \frac{\partial h}{\partial p}\frac{\partial}{\partial q} - \frac{\partial h}{\partial q}\frac{\partial}{\partial p}. \tag{1.27}$$

This immediately generalizes to $M = \mathbb{R}^{2n}$, where we define

$$\{f,g\} = \sum_{j=1}^{n}\left(\frac{\partial f}{\partial p_j}\frac{\partial g}{\partial q^j} - \frac{\partial f}{\partial q^j}\frac{\partial g}{\partial p_j}\right). \tag{1.28}$$

In that case, the Hamiltonian vector field of h is obviously given by

$$\delta_h = \sum_{j=1}^{n}\left(\frac{\partial h}{\partial p_j}\frac{\partial}{\partial q^j} - \frac{\partial h}{\partial q^j}\frac{\partial}{\partial p_j}\right). \tag{1.29}$$

The flow of δ_h gives the motion of a system with Hamiltonian h. Writing

$$\varphi_t(p,q) = (p(t), q(t)), \tag{1.30}$$

we see from (1.18) that this flow is given by ***Hamilton's equations***

$$\frac{dp_j(t)}{dt} = -\frac{\partial h(p(t),q(t))}{\partial q^j}; \tag{1.31}$$

$$\frac{dq^j(t)}{dt} = \frac{\partial h(p(t),q(t))}{\partial p_j}. \tag{1.32}$$

Hamiltonians of the special form

$$h(p,q) = \frac{\sum_j p_j^2}{2m} + V(q) \tag{1.33}$$

give ***Newton's equation***

$$m\frac{d^2q^j(t)}{dt^2} = F_j(q(t)) \tag{1.34}$$

where $F_j = -\partial V/\partial q^j$, better know as $F = ma$ (more precisely, $ma = F$). Finding this equation, including the underlying concepts of dynamics, force, mass, and acceleration, has been one of the greatest intellectual achievements of mankind.[4]

A different kind of example is $M = \mathbb{R}^3$, which is *odd-dimensional*, where we put

$$\begin{aligned}\{f,g\}(x,y,z) = {} & x\left(\frac{\partial f}{\partial y}\frac{\partial g}{\partial z} - \frac{\partial f}{\partial z}\frac{\partial g}{\partial y}\right)\\ & + y\left(\frac{\partial f}{\partial z}\frac{\partial g}{\partial x} - \frac{\partial f}{\partial x}\frac{\partial g}{\partial z}\right)\\ & + z\left(\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}\right).\end{aligned} \tag{1.35}$$

The following result is conceptually nice; the proof is a straightforward exercise.

Proposition 1.1. *A Poisson bracket on M defines a Lie algebra homomorphism*

$$C^\infty(M) \to \mathrm{Der}(C^\infty(M)); \tag{1.36}$$

$$h \mapsto \delta_h. \tag{1.37}$$

In other words, for any $f,g \in C^\infty(M)$ and $s,t \in \mathbb{R}$ we have (see (1.7)*)*

$$\delta_{sf+tg} = s\delta_f + t\delta_g; \tag{1.38}$$

$$[\delta_f, \delta_g] = \delta_{\{f,g\}}. \tag{1.39}$$

For any derivation δ on a manifold M, we say that a function $f \in C^\infty(M)$ is ***conserved*** if f is constant along the flow of δ. Eqs. (1.17) and (1.22) imply that if M is a Poisson manifold and $\delta = \delta_h$ is Hamiltonian, then f is conserved iff $\{h, f\} = 0$.

[4] See *Requiem voor Newton* (Contact, Amsterdam, 2005) by the author of these notes.

For computations, it is often convenient to define a Poisson bracket on X in terms of a ***Poisson tensor***. In coordinates, this is just an anti-symmetric matrix $B^{ij}(x)$ that satisfies, for each (i,j,k),

$$\sum_l \left(B^{li}\frac{\partial B^{jk}}{\partial x_l} + B^{lj}\frac{\partial B^{ki}}{\partial x_l} + B^{lk}\frac{\partial B^{ij}}{\partial x_l} \right) = 0. \tag{1.40}$$

In terms of B, the Poisson bracket is defined by

$$\{f,g\}(x) = \sum_{i,j} B^{ij}(x)\frac{\partial f(x)}{\partial x^i}\frac{\partial g(x)}{\partial x^j}, \tag{1.41}$$

or, more abstractly, by

$$\{f,g\} = B(df,dg). \tag{1.42}$$

For example, the Poisson bracket (1.26) comes from the constant Poisson tensor

$$B(x) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{1.43}$$

It is not difficult to see that any Poisson bracket must come from a Poisson tensor, given that by Theorem 1.1, for any derivation δ on $C^\infty(X)$, the function $\delta(g)$ depends linearly on dg (i.e. on the derivatives $\partial g(x)/\partial x^i$). Hence $\{f,g\} = \delta_f(g) = -\delta_g(f)$ depends linearly on both df and dg. This enforces (1.41) or (1.42), upon which (1.41) and the Jacobi identity (1.24) imply (1.40).

Finally, we say that a Poisson manifold is ***symplectic*** if the corresponding Poisson tensor $B(x)$ is given by an *invertible* matrix, for each $x \in M$. For example, $\mathbb{R}^{2n}$ with Poisson bracket (1.28) is symplectic. This requires M to be *even-dimensional*. In that case, the inverse of the Poisson tensor B is called the ***symplectic form*** on M, usually denoted by σ, and M is called a ***symplectic manifold***. This gives rise to the field of ***symplectic geometry***. One of the main results of symplectic geometry (due to A. Weinstein) is that one may always choose local coordinates (p,q) such that B is locally constant and hence takes the form (1.43). Another is that any Poisson manifold is foliated by symplectic manifolds (called its symplectic leaves).

1.3 Lie group actions on manifolds

Symmetries are extremely important in physics.We describe symmetries through ***Lie groups***, which in classical mechanics act on Poisson manifolds. Hardly any Lie group theory is needed for what follows: in physics we only need *linear* Lie groups, which by definition are closed subgroups of $GL_n(\mathbb{R})$ or $GL_n(\mathbb{C})$, with group multiplication simply given by matrix multiplication. For example, $SO(3)$ is the subgroup of $GL_3(\mathbb{R})$ consisting of matrices R that satisfy $R^T R = 1_3$ and $\det(R) = 1$. Also abelian Lie groups like the additive groups $\mathbb{R}^n$ fall under this scope, since one may identify $a \in \mathbb{R}^n$ with the $2n \times 2n$-matrix

$$a \equiv \begin{pmatrix} 1_n & \mathrm{diag}(a) \\ 0 & 1_n \end{pmatrix}, \tag{1.44}$$

where diag(a) is the diaginal $n \times n$ matrix with entries $(a_1, \dots, a_n)$ on the diagonal. Indeed, matrix multiplication indeed reproduces addition. However, we put the n-torus $G = \mathbb{T}^n = U(1)^n$ (where $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$, with multiplication as the group operation), which is the compact sister of $\mathbb{R}^n$, into the diagonal of $GL_n(\mathbb{C})$.

The ***Lie algebra*** of a linear Lie group $G \subset GL_n(\mathbb{K})$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, is

$$\mathfrak{g} = \{A \in M_n(\mathbb{K}) \mid e^{tA} \in G \forall t \in \mathbb{R}\}, \tag{1.45}$$

where the exponential map $\exp : \mathfrak{g} \to G$ is just given by its usual (norm-convergent) power series. Even if $\mathbb{K} = \mathbb{C}$, we always regard $\mathfrak{g}$ as a *real* vector space. A crucial property of $\mathfrak{g}$ is that it is closed under the ***commutator*** (or ***Lie bracket***) defined by

$$[A, B] = AB - BA. \tag{1.46}$$

For example, the Lie algebra $\mathfrak{so}(3)$ of $SO(3)$ consists of all real 3×3 matrices X that satisfy $X^T = -X$. As a vector space have $\mathfrak{so}(3) \cong \mathbb{R}^3$, which follows by choosing a basis

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, J_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, J_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{1.47}$$

of the 3×3 real antisymmeric matrices. The commutators of these elements are

$$[J_1, J_2] = J_3; \ [J_3, J_1] = J_2; \ [J_2, J_3] = J_1. \tag{1.48}$$

The abelian Lie group $G = \mathbb{R}^n$ has $\mathfrak{g} = \mathbb{R}^n$ as its Lie algebra, realized as

$$A \equiv \begin{pmatrix} 0_n & \mathrm{diag}(A) \\ 0 & 0_n \end{pmatrix}. \tag{1.49}$$

All Lie brackets $[A, B]$ then vanish, and the exponential map is the identity, in that $\exp(a)$, computed from the right-hand side of (1.44), equals $A = a$, realized as in (1.49). The Lie algebra $\mathfrak{t}$ of the torus $\mathbb{T}$ is $i\mathbb{R}$, again with vanishing Lie bracket.

We now turn to Lie group actions on manifolds, leaving out the Poisson bracket for the moment. Let G be a Lie group, and let M be a manifold. The following concepts are almost trivially equivalent:

- Group homomorphisms $\varphi : G \to \mathrm{Diff}(M)$, where the ***diffeomorphism group*** $\mathrm{Diff}(M)$ consists of al smooth invertible maps $M \to M$ with smooth inverse;
- Smooth group actions $\varphi : G \times M \to M$, which we write as $\varphi(\gamma, x) \equiv \gamma \cdot x \equiv \varphi_\gamma(x)$.

For each $A \in \mathfrak{g}$ we then define a map $\delta_A : C^\infty(M) \to C^\infty(M)$ by

$$\delta_A f(x) = \frac{d}{dt} f(e^{-tA} \cdot x)_{|t=0}. \tag{1.50}$$

This map is obviously linear. Moreover, it can be shown that δ is well behaved:

Proposition 1.2. *The map $\delta : \mathfrak{g} \to \mathrm{Der}(C^\infty(M))$, $A \mapsto \delta_A$ is a homomorphisms of Lie algebra, i.e., each δ_A is a derivation, δ is linear, and, for each $A, B \in \mathfrak{g}$,*

$$[\delta_A, \delta_B] = \delta_{[A,B]}. \tag{1.51}$$

The proof is an exercise in differential geometry (which alas is too difficult now), but some examples might help.

1. Let $G = \mathbb{R}^d$ act on $M = \mathbb{R}^d$ by $\varphi(a,x) = x + a$. Let $(e_1, \ldots, e_d)$ be the usual basis of $\mathbb{R}^d$, i.e., $e_1 = (1, 0, \ldots, 0)$, etc. Then

$$\delta_{e_i} f(x) = \frac{d}{dt} f(x_1, \ldots, x_i - t, \ldots, x_d)_{|t=0} = -\frac{\partial f}{\partial x^i}(x), \tag{1.52}$$

so that $\delta_{e_i} = -\partial/\partial x^i$. More generally, for $A \in \mathfrak{g}$, expressed in the same basis,

$$\delta_A = -\sum_{i=1}^{d} A_i \frac{\partial}{\partial x^i}. \tag{1.53}$$

2. Let $G = SO(3)$ act on $M = \mathbb{R}^3$ in the usual way. Using (1.47), we obtain

$$\begin{aligned} \delta_{J_1} f(x) &= \frac{d}{dt} f\left(e^{-tJ_1}(x,y,z)\right)_{|t=0} = \frac{d}{dt} f((1_3 - tJ_1) \cdot (x,y,x))_{|t=0} \\ &= \frac{d}{dt} f(x, y + tz, z - ty)_{|t=0} = \left(z \frac{\partial f}{\partial y} - y \frac{\partial f}{\partial z}\right)(x,y,z). \end{aligned} \tag{1.54}$$

In this way we obtain

$$\delta_{J_1} = z\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}; \tag{1.55}$$

$$\delta_{J_2} = x\frac{\partial}{\partial z} - z\frac{\partial}{\partial x}; \tag{1.56}$$

$$\delta_{J_1} = y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}, \tag{1.57}$$

from which one may verify (1.51), for example, $[\delta_{J_1}, \delta_{J_2}] = \delta_{J_3}$.

3. Let G act on $M = G$ by left multiplication, i.e., $\varphi(x,y) = xy$. The map $A \mapsto \delta_A$ is injective and realizes the Lie algebra of G as an algebra of first-order differential operators on $C^\infty(G)$ (which is the way the Lie algebra is sometimes *defined*).

Exercises for week 2 (Inleveropgaven: 2 and 3, bonus: 5 and 6)

1. Take $M = \mathbb{R}^2$, with the unusual bracket

$$\{f,g\}(p,q) = p\left(\frac{\partial f}{\partial p}\frac{\partial g}{\partial q} - \frac{\partial f}{\partial q}\frac{\partial g}{\partial p}\right). \tag{1.58}$$

   Show that this is a Poisson bracket. Is it symplectic?
2. Prove Proposition 1.1. Furthermore, check (1.39) in examples (1.26) and (1.35).
3. Verify that (1.35) defines a Poisson bracket on $M = \mathbb{R}^3$ and write down the corresponding Poisson tensor B. As an application, take a rigid body, which is described by the angular velocity $\Omega = (\Omega_1, \Omega_2, \Omega_3)$ and the moment of inertia $I = (I_1, I_2, I_3)$, with associated angular momentum $\Pi = (\Pi_1, \Pi_2, \Pi_3)$ with $\Pi_i = I_i \Omega_i$. The latter are the coordinates $(x,y,z) = (\Pi_1, \Pi_2, \Pi_3)$ on phase space $M = \mathbb{R}^3$, with Poisson bracket (1.35). The Hamiltonian for a rigid body is then given by

$$h(\Pi) = \tfrac{1}{2}(I_1\Omega_1^2 + I_2\Omega_2^2 + I_3\Omega_3^2) = \tfrac{1}{2}\left(\frac{\Pi_1^2}{I_1} + \frac{\Pi_2^2}{I_2} + \frac{\Pi_3^2}{I_3}\right); \tag{1.59}$$

   the corresponding equations of motion for the Ω_i are the ***Euler equations***. Derive the Euler equations.
4. Show that (1.41) defines a Poisson bracket, provided that B satisfies (1.40).
5. Show that if $A, B \in \mathfrak{g}$ as defined by (1.45), then $[A,B] \in \mathfrak{g}$, cf. (1.46).
6. Prove Proposition 1.2 for $M = \mathbb{R}^k$,

1.4 Symmetries of Poisson manifolds and the momentum map

We now introduce the Poisson bracket into the game.

Definition 1.3. *Let G be a Lie group acting on a Poisson manifold M. We say that G acts by* **Poisson symmetries** *if for each $\gamma \in G$ and $f, g \in C^\infty(M)$ we have*

$$\varphi_\gamma^*(\{f,g\}) = \{\varphi_\gamma^*(f), \varphi_\gamma^*(g)\}. \tag{1.60}$$

In terms of the Poisson tensor B, see (1.41), and the explicit form

$$\varphi_\gamma(x) = (\varphi^1(x_1, \ldots, x^d), \ldots, \varphi^d(x_1, \ldots, x^d)), \tag{1.61}$$

the defining condition (1.60) for a Poisson symmetry is equivalent to (exercise)

$$B^{ij}(\varphi_\gamma(x)) = \sum_{k,l} \frac{\partial \varphi_\gamma^i}{\partial x^k}(x) \frac{\partial \varphi_\gamma^j}{\partial x^l}(x) B^{kl}(x), \tag{1.62}$$

which in the language of differential geometry would be written succinctly as

$$(\varphi_\gamma)_* B = B, \tag{1.63}$$

where the ***push-forward*** $\varphi_* B$ of B under any smooth map $\varphi : M \to M$ is defined by

$$\varphi_* B(df, dg) = B(d\varphi^* f, d\varphi^* g), \tag{1.64}$$

where $\varphi^* f = f \circ \varphi$, as usual, and we may also writ $\varphi^* df$ instead of $d\varphi^* f$, etc.
Here are some examples of Poisson symmetries (with details in the exercises).

1. Let $G = SO(3)$ act on $M = \mathbb{R}^3$ in the natural way. This action is a symmetry of the Poisson bracket (1.35)
2. Take $M = \mathbb{R}^6 = \mathbb{R}^3 \times \mathbb{R}^3$, with coordinates $x = (\mathbf{p}, \mathbf{q})$, where $\mathbf{p} = (p_1, p_2, p_2)$ and $\mathbf{q} = (q^1, q^2, q^3)$, equipped with the 'canonical' Poisson bracket (1.28), i.e.,

$$\{f, g\} = \sum_{j=1}^{3} \left(\frac{\partial f}{\partial p_j} \frac{\partial g}{\partial q^j} - \frac{\partial f}{\partial q_j} \frac{\partial g}{\partial p^j} \right). \tag{1.65}$$

   a. Let $G = \mathbb{R}^6$ act on $M = \mathbb{R}^6$ by

$$(\mathbf{a}, \mathbf{b}) \cdot (\mathbf{p}, \mathbf{q}) = (\mathbf{p} + \mathbf{a}, \mathbf{q} + \mathbf{b}). \tag{1.66}$$

   b. Let $G = SO(3)$ act on $\mathbb{R}^6$ by

$$R \cdot (\mathbf{p}, \mathbf{q}) = (R\mathbf{p}, R\mathbf{q}). \tag{1.67}$$

3. Define the ***symplectic group*** $\mathrm{Sp}_{2n}(\mathbb{R})$ in dimension $2n$ as the subgroup of $GL_{2n}(\mathbb{R})$ consisting of invertible matrices Γ that satisfy

$$\Gamma^T B \Gamma = B, \tag{1.68}$$

where

$$B = \begin{pmatrix} 0 & 1_n \\ -1_n & 0 \end{pmatrix}, \tag{1.69}$$

In other words, if, given the standard inner product $\langle \cdot, \cdot \rangle$ on $\mathbb{R}^{2n}$, we define a new bilinear form ω on $\mathbb{R}^{2n}$ by

$$\omega(x,y) = \langle x, By \rangle, \tag{1.70}$$

where we should really write $S = -B^{-1} = B$, since it is (minus) the *inverse* of B that defines a symplectic form, then $\Gamma \in \mathrm{Sp}_{2n}(\mathbb{R})$ iff

$$\omega(\Gamma x, \Gamma y) = \omega(x,y) \tag{1.71}$$

for all $x, y \in \mathbb{R}^{2n}$. Note that ω is *bilinear, anti-symmetric*, and *nondegenerate*; a form on a real vector space V with these properties is by definition ***symplectic***. This easily implies that the dimension of V is *even* (provided it is finite), so that $V \cong \mathbb{R}^{2n}$, and it can be shown that one can always choose a basis so that a symplectic form ω is given by (1.70), with (1.69). Since ω is nondegenerate, eq. (1.71) implies that Γ is invertible, i.e. $\Gamma \in GL_{2n}(\mathbb{R})$ automatically.
The group $\mathrm{Sp}_{2n}(\mathbb{R})$*, then, naturally acts on* $\mathbb{R}^{2n}$ *by Poisson symmetries.*
For later use, we record the fact that the Lie algebra of the symplectic group $\mathrm{Sp}_{2n}(\mathbb{R})$ is given by those real $2n \times 2n$ matrices A that satisfy

$$A^T B + BA = 0; \tag{1.72}$$

to see this, take $\Gamma = \exp(tA)$ in (1.71) and compute $d/dt(\cdots)_{|t=0}$.

4. Let G be a Lie group with Lie algebra $\mathfrak{g}$. Choose a basis (T_a) of $\mathfrak{g}$, with associated ***structure constants*** C^c_{ab} defined by the Lie bracket on $\mathfrak{g}$ as

$$[T_a, T_b] = \sum_c C^c_{ab} T_c. \tag{1.73}$$

We write θ in the dual vector space $\mathfrak{g}^*$ as $\theta = \sum_a \theta_a \omega^a$, where (ω_a) is the dual basis to a chosen basis (T_a) of $\mathfrak{g}$, i.e., $\omega_a(T_b) = \delta_{ab}$. In terms of these coordinates, the ***Lie–Poisson bracket*** on $C^\infty(\mathfrak{g}^*)$ is defined by

$$\{f, g\}(\theta) = C^c_{ab} \theta_c \frac{\partial f(\theta)}{\partial \theta_a} \frac{\partial g(\theta)}{\partial \theta_b}. \tag{1.74}$$

Equivalently, the Poisson bracket (1.74) may be defined by the condition

$$\{\hat{A}, \hat{B}\} = \widehat{[A,B]}, \tag{1.75}$$

where $A, B \in \mathfrak{g}$ and $\hat{A} \in C^\infty(\mathfrak{g}^*)$ is the evaluation map $\hat{A}(\theta) = \theta(A)$.
Now G canonically acts on $\mathfrak{g}^*$ through the ***coadjoint representation***, defined by

$$(x \cdot \theta)(A) = \theta(x^{-1}Ax). \tag{1.76}$$

This action is Hamiltonian with respect to the Lie–Poisson bracket (1.74)

Writing $\gamma = \exp(-tA)$ for $t \in \mathbb{R}$ and $A \in \mathfrak{g}$, and applying (1.50) to (1.60), we find

$$\delta_A(\{f,g\}) = \{\delta_A(f),g\} + \{f,\delta_A(g)\}. \tag{1.77}$$

For *connected* Lie groups this 'infinitesimal' property is equivalent to (1.60); this relies on some Lie theory to the effect that G is generated by the image of the exponential map, in the sense that if G is connected, then every $\gamma \in G$ takes the form

$$\gamma = \exp(A_1)\cdots\exp(A_n) \ \ (A_1,\ldots,A_n \in \mathfrak{g}). \tag{1.78}$$

Compare (1.77) with the following property δ_A already has since it is a derivation:

$$\delta_A(fg) = \delta_A(f)g + f\delta_A(g). \tag{1.79}$$

We may call a derivation $\delta : C^\infty(M) \to C^\infty(M)$ satisfying the like of (1.77), i.e.,

$$\delta(\{f,g\}) = \{\delta(f),g\} + \{f,\delta(g)\}, \tag{1.80}$$

a ***Poisson derivation***. We are already familiar with a large class of Poisson derivations: for each $h \in C^\infty(M)$, the corresponding map δ_h defined by (1.22) is a Poisson derivation (this follows from the Jacobi identity). Let us call a Poisson derivation of the kind δ_h ***inner***. This raises the question if our derivations δ_A are inner.

Definition 1.4. *A* **momentum map** *for a Lie group G acting on a Poisson manifold M is a map*

$$J : M \to \mathfrak{g}^* \tag{1.81}$$

such that for each $A \in \mathfrak{g}$,

$$\delta_A = \delta_{J_A}, \tag{1.82}$$

where the function $J_A \in C^\infty(M)$ *is defined by by*

$$J_A(x) = \langle J(x), A\rangle \equiv J(x)(A), \tag{1.83}$$

i.e.,

$$\delta_A(f) = \{J_A, f\}, \ \ f \in C^\infty(M). \tag{1.84}$$

We return to the previous list of examples, which all turn out to be Hamiltonian.

1. With $\mathfrak{so}(3) \cong \mathbb{R}^3$ under the choice of basis (J_1,J_2,J_3), and hence also $\mathfrak{g}^* \cong \mathbb{R}^3$ under the usual inner product on $\mathbb{R}^3$, the momentum map $J : M \to \mathfrak{g}^*$ for the natural $SO(3)$-action on $\mathbb{R}^3$ should be a map $J : \mathbb{R}^3 \to \mathbb{R}^3$. Indeed, we have

$$J(\mathbf{x}) = \mathbf{x}. \tag{1.85}$$

2. More generally, the momentum map for the coadjoint action of G on $\mathfrak{g}^*$ is simply the identity map $\mathfrak{g}^* \to \mathfrak{g}^*$, i.e., $J(\theta) = \theta$, or, equivalently, $J_A = \hat{A}$. Indeed,

$$\begin{aligned}\delta_A \tilde{B}(\theta) &= \frac{d}{dt}\tilde{B}(e^{-tA}\cdot\theta)_{|t=0} = \frac{d}{dt}\theta(e^{tA}Be^{-tA})_{|t=0}\\ &= \theta([A,B]) = \widehat{[A,B]}(\theta) = \{\hat{A},\hat{B}\}(\theta)\\ &= \{J_A,\hat{B}\}(\theta).\end{aligned}$$

3. The $M=\mathbb{R}^6$ examples are all Hamiltonian as well:
   a. For $G=\mathbb{R}^6$ one finds
   $$J(\mathbf{p},\mathbf{q}) = (\mathbf{q},-\mathbf{p}). \tag{1.86}$$
   In particular, if $\mathbb{R}^3$ acts on M by $\mathbf{b}:(\mathbf{p},\mathbf{q})\mapsto(\mathbf{p},\mathbf{q}+\mathbf{b})$, we simply have
   $$J(\mathbf{p},\mathbf{q}) = -\mathbf{p}. \tag{1.87}$$
   b. For $G=SO(3)$ we find
   $$J(\mathbf{p},\mathbf{q}) = -\mathbf{q}\times\mathbf{p}. \tag{1.88}$$

This explains why the momentum map generalizes the idea of (angular) momentum.

In Definition 1.4 we did not require the G-action to consist of Poisson symmetries, because this is often a consequence of the existence of a momentum map:

Proposition 1.3. *Let G be a connected Lie group that acts on a Poisson manifold M. If this action is Hamiltonian (i.e., if it has a momentum map), then G acts on (M,B) by Poisson symmetries (in the sense that* (1.60) *holds).*

Proof. An easy computation shows that (1.77) holds (see exercise). The comment following this equation finishes the proof.

The converse is not true: if G acts by Poisson symmetries, the action is not necessarily Hamiltonian. See exercises.

Thus the momentum map is a generalization of (minus) the momentum, whence its name; note that the quantity in (1.88) is (minus) the angular momentum. Of course, these annoying minus signs could be removed by putting a minus sign in (1.82), but that would have other negative (*sic*) consequences.

It would be natural to expect that J, if it exists, satisfies

$$\{J_A,J_B\} = J_{[A,B]} \quad (A,B\in\mathfrak{g}), \tag{1.89}$$

which expresses the Lie algebra homomorphism property of the map $A\mapsto J_A$ from $\mathfrak{g}$ to $C^\infty(M)$, where the Lie algebra structure of the former is given by the commutator or matrices, and the Lie bracket on the latter is the Poisson bracket. This property sometimes holds, and sometimes doesn't, the latter cases not even necessarily being pathological at all. For example, the examples with $G=SO(3)$ and $G=\mathrm{Sp}_{2n}(\mathbb{R})$ satisfy (1.89), as does the $\mathbb{R}^3$-action on $\mathbb{R}^6$, but the $\mathbb{R}^{2n}$-action on $\mathbb{R}^{2n}$ by translation, where $\mathbb{R}^{2n}$ has the canonical Poisson bracket (1.28), does not (why?).

One advantage of the formalism so far is that the well-known link between symmetries and conserved quantities has a clean formulation, called ***Noether's Theorem*** (in Hamiltonian form).

Theorem 1.2. *Let M be a Poisson manifold equipped with a Hamiltonian action of some Lie group G (so that there is a momentum map $J : M \to \mathfrak{g}^*$). Suppose $h \in C^\infty(M)$ is G-invariant, in that $h(\gamma \cdot x) = h(x)$ for each $\gamma \in G$ and $x \in X$. Then for each $A \in g$, the function J_A is constant along the flow of δ_h. In other words,*

$$J_A(\varphi_t(x)) = J_A(x) \tag{1.90}$$

for any $x \in M$ and any $t \in \mathbb{R}$ for which the flow $\varphi_t(x)$ of δ_h is defined.

Proof. Using all assumptions as well as the definition of a flow, we compute:

$$\begin{aligned}\frac{d}{dt} J_A(\varphi_t(x)) = \delta_h(J_A)(\varphi_t(x)) = \{h, J_A\}(\varphi_t(x)) = -\{J_A, h\}(\varphi_t(x)) \\ = -\delta_A(h)(\varphi_t(x)) = \frac{d}{ds} h(e^{sA}\varphi_t(x))_{|s=0} \\ = \frac{d}{ds} h(\varphi_t(x))_{|s=0} = 0.\end{aligned}$$

□

Here are some examples.

1. Take $M = \mathbb{R}^3$ with the Poisson bracket (1.35) and take the defining action of $G = SO(3)$ on $\mathbb{R}^3$. This action has a momentum map (1.85). If $h \in C^\infty(\mathbb{R}^3)$ is invariant under rotations, then it is a function of $x^2 + y^2 + z^2$. But this implies that $\{h, f\} = 0$ for any $f \in C^\infty(\mathbb{R}^3)$. Hence $\delta_h = 0$ and $\varphi_t(x) = x$ for any t, x, where φ is the flow of δ_h. This makes (1.90) a tautology, so that Noether's Theorem is trivial in this case.
2. Let $M = \mathbb{R}^6$ and let $G = SO(3)$ act on M by (1.67). Then a Hamiltonian of the kind (1.33) is G-invariant if the potential V is rotation-invariant. In that case angular momentum $\mathbf{q} \times \mathbf{p}$ is conserved.
3. Still taking $M = \mathbb{R}^6$, let $G = \mathbb{R}^3$ act on M by spatial translations, i.e.,

$$\mathbf{b} \cdot (\mathbf{p}, \mathbf{q}) = (\mathbf{p}, \mathbf{q} + \mathbf{b}). \tag{1.91}$$

   This action has a momentum map

$$J(\mathbf{p}, \mathbf{q}) = -\mathbf{p}. \tag{1.92}$$

   A Hamiltonian of the kind (1.33) is G-invariant if the potential V is translation-invariant, in which case momentum $\mathbf{p}$ is conserved.

Exercises for week 3 (inleveropgaven: nos. 5 and 6)

1. Show that (1.60) and (1.62) are equivalent.
2. Show that the Lie group actions (1.66) and (1.67) define Poisson symmetries.
3. Prove (1.77).
4. Prove (1.85), (1.86) and (1.88).
5. Find (and verify) the momentum map for the defining action of the symplectic group $G = \mathrm{Sp}_{2n}(\mathbb{R})$ on $\mathbb{R}^{2n}$.
6. Let $G = \mathbb{R}$ act on $M = \mathbb{R}^2$ by $b \cdot (p,q) = (p, q+b)$. Equip $\mathbb{R}^2$ with the unusual Poisson bracket (1.58).

   a. Show that this action satisfies (1.60).
   b. Compute the derivation δ_T for the single generator $T = 1$ of $\mathfrak{g} = \mathbb{R}$.
   c. Show that δ_T is not inner.

1.5 The probabilistic structure of classical physics

As a goal in itself, and also in preparation for quantum mechanics (which is a probabilistic theory by its very nature), in this section we discuss the probabilistic structure of classical physics. In Newtoninan physics, one assumes that the state of the system is exactly known; if M is the phase space of the system, henceforth called X (which is the usual notation in probability theory, whereas M is the usual notation for manifolds), then an 'exact' state is just a point $x \in X$. We will soon formalize this, but the idea is that if x is known, then everything can be predicted with certainty, at least in principle (some would say: everything is determined by x), like the values of all observables $f : X \to \mathbb{R}$ (which are given by $f(x)$) and the future (or even past) time evolution of the system. However, if $X = \mathbb{R}^{6N}$ with $N \sim 10^{23}$, then it is not feasible to assume that $x \in X$ is known exactly. This led to (classical) *statistical mechanics*, which was developed in the 19th century mainly by Maxwell, Boltzmann, and Gibbs. Furthermore, even if X is relatively small, the existence of *deterministic chaos* (first studied by Poincaré and others in connection with the stability of the solar system, in which context it was probably familiar already to Newton) indicated that in many dynamical systems x should really be known with almost infinite precision in order to make accurate predictions, which is unfeasible in practice.

Both cases suggested replacing 'exact' or 'pure' states $x \in X$ by *probability measures* on X, which incorporate the confession that although we do not know the exact state, we have at least some insight into the likelihood what it is.

As a warm-up we first assume that X is a ***finite set***, playing the role of the configuration space of some physical system, or, equivalently (as we shall see), of its pure state space. In general, X should be thought of as the phase space, but as long as X is discrete, the phase space coincides with the (intuitively more appealing) configuration space. Finite sets are not at all boring. For example, if X is supposed to describe the possible configurations of N bits (numbered $0, \ldots, N-1$), then $X = \underline{2}^{\underline{N}}$. Here $\underline{N} = \{0, 1, \ldots, N-1\}$ (so that, in particular, $\underline{2} = \{0, 1\}$), and, for arbitrary sets Y and Z, the set Y^Z consists of all functions $x : Z \to Y$. Of course, by binary coding the set $\underline{2}^{\underline{N}}$ may be identified with the set $\{0, 1, \ldots, 2^N - 1\}$, i.e., for $x \in \underline{2}^{\underline{N}}$ the binary number $x(N-1) \cdots x(0)$ is identified with its decimal counterpart $\sum_{k=0}^{N-1} x(k)2^k$. S

In any case, as we will see, X defines both the states and the observables, from the totality of each of which X may in turn be reconstructed. Furthermore, the states and the observables define each other even without relying on the underlying set X.

Definition 1.5. *Recall that the power set $\mathscr{P}(X)$ of X is the set of all subsets of X.*

1. *An* **event** *is a subset $U \subseteq X$, i.e., $U \in \mathscr{P}(X)$.*
2. *A* **probability distribution** *on X is a function $p : X \to [0, 1]$ such that $\sum_x p(x) = 1$.*
3. *A* **probability measure** *on X is a function $\mu : \mathscr{P}(X) \to [0, 1]$ such that $\mu(X) = 1$ and $\mu(U \cup V) = \mu(U) + \mu(V)$ whenever $U \cap V = \emptyset$.*
4. *A* **random variable** *on X is a function $f : X \to \mathbb{R}$.*
5. *The* **spectrum** *of a random variable f is the subset $\sigma(f) = \{f(x) \mid x \in X\}$ of $\mathbb{R}$.*

Probability distributions p and probability measures μ on finite sets come down to the same thing, as the former gives rise to the latter by

$$\mu(U) = \sum_{x \in U} p(x), \tag{1.93}$$

whilst *vice versa* one has

$$p(x) = \mu(\{x\}). \tag{1.94}$$

It is a matter of convenience which one is used.

A special class of probability measures stands out: each $y \in X$ defines a probability distribution p_y by $p_y(x) = \delta_{xy}$, or explicitly,

$$p_y(x) = 1 \text{ if } x = y; \tag{1.95}$$
$$p_y(x) = 0 \text{ if } x \neq y. \tag{1.96}$$

Equivalently, for the corresponding probability measure one has

$$\mu_y(U) = 1 \text{ if } y \in U; \tag{1.97}$$
$$\mu_y(U) = 0 \text{ if } y \notin U. \tag{1.98}$$

The single most important construction in probability theory, then, is as follows.

Theorem 1.3. *A probability measure μ on X and a random variable $f : X \to \mathbb{R}$ jointly yield a probability measure μ_f on the spectrum $\sigma(f)$ by means of*

$$\mu_f(\Delta) = \mu(f \in \Delta), \tag{1.99}$$

where $\Delta \subseteq \sigma(f)$ and $f \in \Delta$ denotes the event $\{x \in X \mid f(x) \in \Delta\}$ in X.

In terms of the corresponding probability distribution p on X, one has

$$p_f(\lambda) = \sum_{x \in X \mid f(x) = \lambda} p(x). \tag{1.100}$$

Instead of $f \in \Delta$, the notation $f^{-1}(\Delta)$ is often used. The proof is an exercise.

Given a finite set X, we may form the set $C(X)$ of all complex-valued functions on X, enriched with the structure of a complex vector space under pointwise operations:

$$(\lambda \cdot f)(x) = \lambda f(x) \ (\lambda \in \mathbb{C}); \tag{1.101}$$
$$(f+g)(x) = f(x) + g(x). \tag{1.102}$$

We use the notation $C(X)$ with some foresight, anticipating the case where X is no longer finite, but in any case, since for the moment it is, every function is continuous. Moreover, the vector space structure on $C(X)$ may be extended to that of a commutative algebra (where, by convention, all our algebras are associative and are defined over the complex scalars) by defining multiplication pointwisely, too:

$$(f \cdot g)(x) = f(x)g(x). \tag{1.103}$$

Note that this algebra has a unit 1_X, i.e., the function identically equal to 1. We also equip $C(X)$ with an ***involution***, which on an arbitrary (not necessarily commutative) algebra A is defined as an anti-linear anti-homomorphism, i.e., a real-linear map $*: A \to A$ (written $a \mapsto a^*$) that satisfies $(\lambda a)^* = \overline{\lambda} a^*$ and $(ab)^* = b^* a^*$. In our case $A = C(X)$, which is commutative, the latter property simply becomes $(fg)^* = f^* g^*$. In any case, we define this involution by pointwise complex conjugation, i.e.,

$$f^*(x) = \overline{f(x)}. \tag{1.104}$$

Finally, $C(X)$ has a natural ***norm***

$$\|f\|_\infty = \sup_{x \in X}\{|f(x)|\}. \tag{1.105}$$

These structures turn $C(X)$ into a ***commutative C*-algebra***.

Definition 1.6. *A C*-algebra is an associative algebra (over $\mathbb{C}$) equipped with an involution as well as a norm in which A is complete (i.e., a Banach space), such that algebra, involution, and norm are related by the axioms*

$$\|ab\| \leq \|a\|\,\|b\|; \tag{1.106}$$
$$\|a^* a\| = \|a\|^2. \tag{1.107}$$

If X is compact (but not necessarily finite), then $C(X)$ is a commutative C*-algebra. if X is locally compact but not compact, one should take the space $C_0(X)$ of all continuous functions $f : X \to \mathbb{C}$ that vanish at infinity (i.e., for any $\varepsilon > 0$ the set $\{x \in X \mid |f(x)| \geq \varepsilon\}$ is compact). It is of fundamental importance that $C(X)$ and $C_0(X)$ are *commutative*. The elements of $C_0(X)$ are called ***observables***. We already noted that $C(X)$ has a unit (as an algebra), namely the function 1_X; this is still the case if X is compact, but $C_0(X)$ has no unit.

Definition 1.7. *A* state *on a C*-algebra A with unit is a linear map $\omega : A \to \mathbb{C}$ that is* positive*, i.e.,*

$$\omega(a^* a) \geq 0 \tag{1.108}$$

for each $a \in A$, as well as normalized *in that*

$$\omega(1_A) = 1. \tag{1.109}$$

If A has no unit, like $C_0(X)$, then (1.109) should be replaced the condition

$$\|\omega\| = 1, \tag{1.110}$$

where $\|\cdot\|$ is the usual norm on the Banach dual A^*, i.e.,

$$\|\omega\| = \sup\{|\omega(a)|, a \in A, \|a\| \leq 1\|. \tag{1.111}$$

In fact, it can be shown that if (1.108) holds, then ω is bounded, and if in addition A has a unit 1_A, then ω satisfies $\|\omega\| = \omega(1_A)$. Therefore, in the presence of a unit and condition (1.108), the normalization conditions (1.110) and (1.111) are equivalent. A special case of this which is within our reach is:

Proposition 1.4. *Let X be a compact Hausdorff space. If a linear map $\varphi : C(X) \to \mathbb{C}$ is positive, then it is bounded, with norm $\|\varphi\| = \varphi(1_X)$.*

The proof is an exercise.

If we specialize Definition 1.7 to the case $A = C(X)$, where X is finite or compact, we note that if $f = a^*a$, then $f(x) = |a(x)|^2$, so that $f(x) \geq 0$ for each x. Conversely, if $f(x) \geq 0$ for each x, then we have $f = a^*a$ for $a = \sqrt{f}$. Hence we have:

Lemma 1.1. *A state on $C(X)$ is a complex-linear map $\omega : C(X) \to \mathbb{C}$ that satisfies:*

1. *$\omega(f) \geq 0$ for each $f \geq 0$ (i.e., $f(x) \geq 0$ for each x);*
2. *$\omega(1_X) = 1$.*

Proposition 1.5. *For finite X there is a bijective correspondence between states ω on $C(X)$ and probability measures μ on X, given by*

$$\omega(f) = \sum_{x\in X} p(x)f(x) = \sum_{\lambda\in\sigma(f)} \mu(f = \lambda)\cdot\lambda; \tag{1.112}$$

$$\mu(U) = \omega(1_U), \tag{1.113}$$

where $p(x)$ is given by (1.94) *and 1_U is the characteristic function of U (defined by $1_U(x) = 1$ if $x \in U$ and $1_U(x) = 0$ if $x \notin U$).*

The proof is an exercise. The ***state space*** $S(C(X))$ of the algebra $C(X)$ is the set of all states on $C(X)$; by the theorem, this is essentially the same as the set $Pr(X)$ of all probability distributions p on X. These are examples of ***compact convex sets***.

Definition 1.8. *A subset K of a real vector space V is called* **convex** *if whenever $v, w \in K$ and $t \in (0,1)$, one has $tv + (1-t)w \in K$. This is equivalent with the following property: if $t_1, \ldots, t_n$ are numbers in $[0,1]$ such that $\sum_i t_i = 1$ and if $v_1, \ldots, v_n$ are in K, then $\sum_{i=1}^n t_i \cdot v_i$ is in K.*

Thus any linear subspace of V is trivially convex, but the interesting convex sets are *compact* in the topology inherited from V (provided it has one, always assumed Hausdorff). We will usually have $V = W^*$, the space of linear functionals from some other (finite-dimensional) real vector space W to $\mathbb{R}$, and the topology on V is the so-called w^*-topology, defined by saying that $v_n \to v$ iff $v_n(w) \to v(w)$ for each $w \in W$. Then K is compact in V iff each infinite sequence in K has a convergent subsequence. If V is finite-dimensional, as is the case in this section, then compact just means closed and bounded.

Definition 1.9. *The* **(extreme) boundary** *$\partial_e K$ of a convex set K consists of all $v \in K$ satisfying the condition:*

$$\text{if } v = tw + (1-t)x \text{ for certain } w, x \in K \text{ and } t \in (0,1), \text{ then } v = w = x.$$

Elements $v \in \partial_e K$ of the boundary are called **extremal points** *of K.*

The main example of interest to us is:

Theorem 1.4. *For any locally compact Hausdorff space X, the boundary $\partial_e Pr(X)$ of the convex set $Pr(X)$ of all probability measures on X is isomorphic to X through $\delta_x \leftrightarrow x$, where $\delta_x : C_0(X) \to \mathbb{C}$ is the evaluation map*

$$\delta_x(f) = f(x). \tag{1.114}$$

The proof is (once again) an exercise. Of course, the probability distribution p corresponding to the state δ_x is p_x, see (1.96)

The situation is almost the same for compact Hausdorff spaces X, taking $A = C(X)$, and even for locally compact Hausdorff spaces, provided we take $A = C_0(X)$. Instead of Proposition 1.5 we now have the ***Riesz–Radon Representation Theorem***:

Theorem 1.5. *Let X be a locally compact Hausdorff space. There is a bijective correspondence between* states $\omega : C_0(X) \to \mathbb{C}$ *and* complete regular probability measures μ *on X, given as follows:*

- *The measure μ defines the state ω through*

$$\omega(f) = \int_X d\mu\, f,\ f \in C_0(X). \tag{1.115}$$

- *The state ω defines the measure μ on open sets U and on compact sets K by*

$$\mu(U) = \sup\{\omega(f) \mid f \in C_c(U), 0 \leq f \leq 1_X\}; \tag{1.116}$$
$$\mu(K) = \inf\{\omega(f) \mid f \in C_c(X), 0 \leq f \leq 1_X, f_{|K} = 1_K\}, \tag{1.117}$$

respectively; this eventually gives $\mu(A)$ for all (Borel) measurable sets $A \subset X$.

See appendix for an explanation of the terminology. Theorem 1.4 remains valid:

Theorem 1.6. *For any locally compact Hausdorff space X, the correspondence $\delta_x \leftrightarrow x$ gives a bijection $\partial_e S(C_0(X)) \cong X$.*

Exercises for week 4 (inleveropgaven: 5 and 6)[5]

1. Prove Proposition 1.4. Hint: use the pre-inner product $\langle f, g\rangle = \varphi(f^* g)$ on $C(X)$.
2. Prove the second equality in (1.112).
3. Prove Proposition 1.5.
4. Show directly from their definitions that $S(C(X))$ and $Pr(X)$ are convex sets.
5. Prove Theorem 1.4.
6. Define the ***variance*** $\Delta_\omega(f)$ of $f \in C_0(X)$ with respect to some probability measure μ on X or the corresponding state ω by

$$\Delta_\omega(f) = \omega(f^2) - \omega(f)^2. \tag{1.118}$$

   Show that $\mu = \mu_y$ for some $y \in X$ iff $\Delta_\omega(f) = 0$ for all $f \in C_0(X, \mathbb{R})$.

[5] Physics students without Introductory Functional Analysis only need to do these for finite X; others should do the Full Monty!

In finite dimension, there are two kinds of compact convex sets: smooth ones, like the (closed) unit disc in $\mathbb{R}^2$ or the (closed) unit ball in $\mathbb{R}^3$ (which we will encounter in quantum mechanics), and ***convex polytopes***, which by definition are convex hulls of finitely many points (that is, the smallest convex sets containing these points). Examples of convex polytopes are ***regular polyhedra***. These were classified (up to affine isomorphism, i.e., bijections preserving convex sums) by Schläfli in 1852, who showed that the only possibilities are:

- The *simplices* $\Delta_n = \{x \in \mathbb{R}^{n+1} \mid x_i \geq 0, \sum_i x_i = 1\}$, $n \geq 1$;
- The *cubes* $Q_n = \{x \in \mathbb{R}^n \mid -1 \leq x_i \leq 1\}$, $n > 1$;
- The *cross-polytopes* $O_n = \{x \in \mathbb{R}^n \mid \sum_i |x_i| \leq 1\}$, $n > 1$;
- The countably many *regular polygons* in $\mathbb{R}^2$ (which include Q_2, O_2, Δ_2);
- The five *platonic solids* in $\mathbb{R}^3$ (which include Q_3, O_3, Δ_3);
- The six *regular polychora* in $\mathbb{R}^4$ (which include Q_4, O_4, Δ_4).

Here Δ_n is affinely homeomorphic to the convex hull of $n+1$ *linearly independent* points, and this property uniquely defines it (up to affine isomorphism). It is almost tautological that the simplex Δ_n is the set $Pr(X_{n+1})$ of all probability distributions on a set X_{n+1} of cardinality $n+1$.

Theorem 1.4 may be rewritten as follows:

Theorem 1.7. *The boundary of the n-dimensional simplex Δ_n is given by*

$$\partial_e \Delta_n = \{\mathbf{e}_1, \ldots . \mathbf{e}_{n+1}\}, \tag{1.119}$$

where $(\mathbf{e}_1, \ldots . \mathbf{e}_{n+1})$ is the standard basis of $\mathbb{R}^{n+1}$ (i.e., $\mathbf{e}_1 = (1, 0, \ldots, 0)$, etc.).

It follows that $|\partial_e \Delta_n| = n+1$, i.e., the boundary of Δ_n has $n+1$ points. This is another way to single out the simplices among all regular polyhedra. The simplest example is $\Delta_1 \cong [0,1]$, so that $\partial_e \Delta_1 \cong \{0,1\}$. Note that $\partial_e(0,1) = \emptyset$, so that the boundary of a convex set may well be empty (another example is the *open* disc). This cannot happen if K is compact (in finite dimension this was proved by Caratheodory, whereas in general it follows from the Krein–Milman Theorem of functional analysis).

1.6 Appendix: some measure theory

In what follows, X will typically be a locally compact Hausdorff space (we will often omit the predicate 'Hausdorff', but it always applies). Measure theory studies ***measure spaces*** (X, Σ, μ), where X is a set, and,

- $\Sigma \subseteq \mathscr{P}(X)$ is a so-called ***σ-algebra*** of subsets of X, which means that:

  1. $X \in \Sigma$;
  2. If $A \in \Sigma$, then $A^c \in \Sigma$ (where $A^c \equiv X \backslash A$ is the complement of A);
  3. If $A_n \in \Sigma$ for $n \in \mathbb{N}$, then $\cup_n A_n \in \Sigma$ (i.e., Σ is closed under countable *unions*).

  It follows that $\emptyset \in \Sigma$, and that Σ is closed under countable *intersections*, too.
- $\mu : \Sigma \to [0, \infty]$, called a (positive) ***measure***, is ***countably additive***, i.e.,

$$\mu(\cup_n A_n) = \sum_n \mu(A_n), \tag{1.120}$$

  whenever $A_n \in \Sigma$, $n \in \mathbb{N}$, $A_i \cap A_j = \emptyset$ for all $i \neq j$ (this seemingly technical condition is indispensable in almost every limit argument in measure theory). The obvious convention here is that $t + \infty = \infty$ for any $t \in \mathbb{R}^+$, as well as $\infty + \infty = \infty$.

A ***probability space*** is a measure space (X,Σ,μ) for which $\mu(X)=1$, in which case μ is called a ***probability measure***. More generally, a measure space is called ***finite*** if $\mu(X)<\infty$, which evidently implies $\mu(A)<\infty$ for any $A\in\Sigma$, and ***σ-finite*** if X is a countable union $X=\cup_n A_n$ with $\mu(A_n)<\infty$ for each n. For example, with regard to Lebesgue measure $X=\mathbb{R}$ is σ-finite whilst $X=[0,1]$ is finite (the non-σ-finite case is pathological). In what follows, we only need to deal with probability measures.

This definition of a σ-algebra marks a difference with a topology on X, which is a collection $\mathscr{O}(X)$ of 'open' subsets (containing X and the empty set $\emptyset$) that is closed under *arbitrary* unions and *finite* intersections (but *not* under complementation!).

Nonetheless, topology and measure theory are closely related:

1. Any topological space X admits a canonical σ-algebra $\mathscr{B}(X)$, viz. the smallest σ-algebra in $\mathscr{P}(X)$ that contains $\mathscr{O}(X)$ (this exists and equals the intersection of all σ-algebra that contain $\mathscr{O}(X)$); elements of $\mathscr{B}(X)$ are called ***Borel sets*** in X.
2. The definition of a ***continuous*** function $f:X\to Y$ between topological spaces X and Y as a function for which $f^{-1}(V)\in\mathscr{O}(X)$ for each $V\in\mathscr{O}(Y)$, is copied by saying that $f:X\to Y$ is ***measurable*** with respect to given σ-algebras Σ_X (on X) and Σ_Y (on Y) if $f^{-1}(B)\in\Sigma_X$ for any $B\in\Sigma_Y$.
3. If X and Y are topological spaces and $\Sigma_X=\mathscr{B}(X)$, $\Sigma_Y=\mathscr{B}(Y)$, then it it easy to show that f is (Borel) measurable iff $f^{-1}(B)\in\Sigma_X$ merely for any $B\in\mathscr{O}(Y)$, from which it follows that each continuous function is measurable. For $f:X\to\mathbb{R}$ to be measurable it is even sufficient that $f^{-1}((t,\infty))\in\Sigma_X$ for each $t\in\mathbb{R}$.
4. The above condition of σ-finiteness is often used just in case the A_i are *compact*.

An important goal of measure theory is to provide a rigorous theory of ***integration***; here the key idea (due to Lebesgue) is that in defining the integral of some *measurable* function $f:X\to\mathbb{R}$, one should partition the *range* $\mathbb{R}$ rather than the *domain* X, as had been done in the Calculus since Newton (where typically $X\subseteq\mathbb{R}^n$). This, in turn, suggests that f should first be approximated by ***simple*** functions.
These are *measurable* functions $s:X\to\mathbb{R}^+$ with finite range, or, equivalently,

$$s=\sum_i \lambda_i 1_{A_i}, \tag{1.121}$$

where $\lambda_i\geq 0$, $A_i\in\Sigma$, and $n<\infty$. Such a representation is unique if we require that the sets A_i are mutually disjoint and the coefficients λ_i are distinct; namely, if $\{x_1,\ldots,x_n\}$ are the distinct vales of s, one takes $A_i=s^{-1}(x_i)$ and $\lambda_i=x_i$. Given some measure μ, we further restrict the class of simple functions to those for which $\mu(A_i)<\infty$. One then first defines the integral of a simple function s, as in (1.121), by

$$\int_X d\mu\, s=\sum_i \lambda_i\mu(A_i); \tag{1.122}$$

a nontrivial argument shows that the right-hand side is independent of the particular representation (1.121) of s used on the left. Granting this, linearity of the integral on simple functions is immediate. Subsequently, for *positive* measurable functions $f\geq 0$, writing $s\leq f$ iff $s(x)\leq f(x)$ for each $x\in X$, one defines the integral by

$$\int_X d\mu\, f = \sup\left\{\int_X d\mu\, s \mid 0 \leq s \leq f, s \text{ simple}\right\}. \tag{1.123}$$

For measurable functions $f : X \to \mathbb{C}$, one first decomposes f as

$$f = \sum_{k=0}^{3} i^k f_k,\ f_k \geq 0. \tag{1.124}$$

On this basis, one then defines the integral by linear extension of (1.123), that is,

$$\int_X d\mu\, f = \sum_{k=0}^{3} i^k \int_X d\mu\, f_k. \tag{1.125}$$

We call f ***integrable*** with respect to μ, writing $f \in \mathscr{L}^1(X, \Sigma, \mu)$, if

$$\int_X d\mu\, |f| < \infty; \tag{1.126}$$

this implies that each positive part f_k, and hence also f itself, is integrable, i.e.,

$$\int_X d\mu\, f < \infty. \tag{1.127}$$

One then has the useful estimates

$$\left|\int_X d\mu\, f\right| \leq \int_X d\mu\, |f| \leq \|f\|_\infty^{\text{ess}} \mu(X), \tag{1.128}$$

where the ***essential supremum*** of f (with respect to μ) is defined by

$$\|f\|_\infty^{\text{ess}} = \inf\{t \in [0,\infty] \mid |f| \leq t \text{ a.e.}\}, \tag{1.129}$$

in which one or both terms in the third expression may well be infinite (in which case the second estimate still holds, of course!). However, if X is a compact space (see the next section), $f \in C(X)$, and μ is finite, then all of (1.128) is useful.

Linearity of the integral is far from trivial: the proof relies on both Lebesgue Theorems of integration.

For us it suffices to deal with ***locally compact Hausdorff spaces*** X (think of $X = \mathbb{R}^n$, or of compact spaces like $X = [0,1]$). As before, we say that a map $\varphi : C(X) \to \mathbb{C}$ is ***positive*** if $\varphi(f) \geq 0$ whenever $f \geq 0$ (pointwise). We also write $\mathscr{O}(X)$ for the set of *open* subsets of X, whilst $\mathscr{K}(X)$ denotes the set of all *compact* subsets of X.

We start with the compact case. The idea is quite straightforward: any finite measure $\mu : \mathscr{B}(X) \to [0,\infty)$ gives rise to a positive linear map $\varphi : C(X) \to \mathbb{C}$ via

$$\varphi(f) = \int_X d\mu\, f,\ f \in C(X). \tag{1.130}$$

Conversely, any such map canonically defines a finite measure μ at least on opens $U \in \mathscr{O}(X)$ and on compacta $K \in \mathscr{K}(X)$ (which are key examples of Borel sets) by

$$\mu(U) = \sup\{\varphi(f) \mid f \in C_c(U), 0 \le f \le 1_X\}; \tag{1.131}$$
$$\mu(K) = \inf\{\varphi(f) \mid f \in C_c(X), 0 \le f \le 1_X, f_{|K} = 1_K\}. \tag{1.132}$$

Subsequently, this preliminary measure is (hopefully!) to be extended to at least all of $\mathscr{B}(X)$, i.e., to all Borel sets, in such a way that μ recovers φ via (1.130).

This works, and one even obtains a bijective correspondence between finite measure spaces (X, Σ, μ) and positive linear maps $\varphi : C(X) \to \mathbb{C}$ if the former are subjected to two additional conditions, predicated on having $\mathscr{B}(X) \subset \Sigma$, namely:

- ***completeness***, in that $\mu(B) = 0$ and $A \subset B$ for $A \in \mathscr{P}(X)$, $B \in \Sigma$ imply $A \in \Sigma$;
- ***regularity***, i.e., for a given measure $\mu : \Sigma \to [0, \infty]$, for any $A \in \Sigma$, one has

$$\mu^*(A) = \mu_*(A) = \mu(A), \tag{1.133}$$

  where the ***outer measure*** μ^* and ***inner measure*** μ_* are defined by

$$\mu^*(A) = \inf\{\mu(U) \mid U \supseteq A, U \in \mathscr{O}(X)\}; \tag{1.134}$$
$$\mu_*(A) = \sup\{\mu(K) \mid K \subseteq A, K \in \mathscr{K}(X)\}, \tag{1.135}$$

  respectively. These expressions apparently make sense for *all* subsets $A \subset X$, but lovers of the Banach–Tarski Paradox may be reassured that μ^* and μ_* typically fail to be countable additive if they are seen as maps from $\mathscr{P}(X)$ to $[0, \infty]$.

Theorem 1.8. *Let X be a compact Hausdorff space. There is a bijective correspondence between* complete regular finite measure spaces (X, Σ, μ) *and* positive linear maps $\varphi : C(X) \to \mathbb{C}$*, explicitly given as follows:*

- *The measure space* (X, Σ, μ) *defines* φ *through* (1.130)*, assuming* (1.123) - (1.125)*;*
- *The map* φ *defines the pair* (Σ, μ) *in three steps:*
  1. μ *is given on opens U and on compacta K by* (1.131) *and* (1.132)*, respectively;*
  2. Σ *is defined as the collection of all sets* $A \in \mathscr{P}(X)$ *where* $\mu^*(A) = \mu_*(A)$*;*
  3. μ *is given on all of* Σ *by* $\mu(A) = \mu^*(A)$*, using* (1.134)*, or, equivalently (given the previous point), by* $\mu(A) = \mu_*(A)$*, based on* (1.135)*.*

We omit the proof of this ***Riesz Representation Theorem*** (often attributed also to ***Radon***), which is highly technical (especially in proving countable additivity), and just note that expressions like (1.131) and (1.132) are really desparate attempts to define "$\mu(A) = \varphi(1_A)$", which is OK for finite X, but in general is ill defined because even for Borel sets A, the characteristic function 1_A is rarely continuous on X.

We note that μ has to be finite, since obviously $\mu(X) = \varphi(1_X)$. One can say a little more about this. A linear map $\varphi : C(X) \to \mathbb{C}$ is ***bounded*** if, for some $0 < C < \infty$,

$$|\varphi(f)| \leq C\|f\|_\infty. \tag{1.136}$$

In that case, the following expression, called the ***norm*** of φ, is $\leq C$, hence finite:

$$\|\varphi\| = \sup\{|\varphi(f)|, f \in C(X), \|f\|_\infty = 1\}. \tag{1.137}$$

The proof is an exercise. As in the finite case, a ***state*** on $C(X)$ remains a positive linear functional $\omega : C(X) \to \mathbb{C}$ with $\omega(1_X) = 1$. Theorem 1.8 then yields:

Corollary 1.1. *If X is a compact Hausdorff space, there is a bijective correspondence between states on $C(X)$ and complete regular probability measures on X.*

In order to generalize Corollary 1.1 to the locally compact case, we involve the space $C_0(X)$ of all continuous functions $f : X \to \mathbb{C}$ that ***vanish at infinity***, i.e., for any $\varepsilon > 0$ the set $\{x \in X \mid |f(x)| \geq \varepsilon\}$ is compact, or, equivalently, for any $\varepsilon > 0$ there is a compact set $K \subset X$ such that $|f(x)| < \varepsilon$ for all $x \notin K$. (This choice may sound contrived, but the seemingly more natural choice $C_b(X)$ of all bounded continuous functions turns out to fail on Proposition 1.4; in fact, the pure state space of $C_0(X)$ will turn out to be X, whereas $C_b(X)$ has far more pure states that together form βX, the terrible Ceh–Stone compactification of X, which is much, much larger than X).

Lemma 1.2. *A positive linear map $\varphi : C_0(X) \to \mathbb{C}$ is bounded.*

The proof is an exercise.

We now define a ***state*** on $C_0(X)$ as a positive (and hence bounded) linear functional $\omega : C_0(X) \to \mathbb{C}$ with $\|\omega\| = 1$; this is consistent with the terminology for the compact case because of Proposition 1.4. The following key theorem, which is the central result of this section, generalizes Proposition 1.5. As we see, this generalization requires the full power of measure theory on locally compact spaces.

Theorem 1.9. *Let X be a locally compact Hausdorff space. There is a bijective correspondence between states on $C_0(X)$ and complete regular probability measures on X, explicitly given by* (1.130) *and by the bullet points of Theorem 1.5. In particular, the state space $S(C_0(X))$ may be identified with the space $Pr(X)$ of all complete regular probability measures on X.*

It is an exercise to show that under (1.130), one has

$$\|\varphi\| = \mu(X). \tag{1.138}$$

In particular, if φ is a state, then $\mu(X) = 1$, i.e., μ is a probability measure. It turns out that also Theorem 1.4 still holds, with a slightly more involved proof.

Chapter 2
Quantum theory on Hilbert space

Notation: For any H, we use the following notation: two arbitrary vectors $\psi, \varphi \in H$ define an operator $|\psi\rangle\langle\varphi|$ through Dirac's 'bra-ket' notation

$$|\psi\rangle\langle\varphi|\chi = \langle\varphi, \chi\rangle\psi. \tag{2.1}$$

A basis of a Hilbert space by convention always means an *orthonormal* basis.

2.1 Finite-dimensional Hilbert space

The quantum analogue of a finite set X (in its role as a phase space in classical mechanics) is the finite-dimensional Hilbert space $\ell^2(X)$, by which we mean the vector space of functions $\psi : X \to \mathbb{C}$, equipped with the inner product

$$\langle\psi, \varphi\rangle = \sum_{x\in X} \overline{\psi(x)}\phi(x). \tag{2.2}$$

For finite X we have $X \cong \underline{n} = \{1, 2, \ldots, n\}$ as sets, where n is the cardinality of X, inducing the unitary isomorphism $\ell^2(\underline{n}) \cong \mathbb{C}^n$ of Hilbert space through the map $\psi \mapsto (\psi(1), \ldots, \psi(n))$, where $\mathbb{C}^n$ has the standard inner product

$$\langle w, z\rangle = \sum_i \overline{w}_i z_i. \tag{2.3}$$

In particular, $\delta_k \in \ell^2(\underline{n})$ is mapped to the k'th standard basis vector $\upsilon_k \equiv |k\rangle$ of $\mathbb{C}^n$.

If H is finite-dimensional, we may therefore assume that $H = \mathbb{C}^n$ and that $B(H)$, i.e., the algebra of all bounded linear maps $a : H \to H$, is just the algebra $M_n(\mathbb{C})$ of $n \times n$ matrices (if $\dim(H) < \infty$, linear maps are automatically continuous, and we often refrain from making the subtle difference between linear maps $a : \mathbb{C}^n \to \mathbb{C}^n$ and the matrices representing such maps, once a basis of $\mathbb{C}^n$ has been chosen). In any case, $B(H)$ or $M_n(\mathbb{C})$ is the quantum analogue of the algebra $C(X)$ in the previous section. Like $C(X)$, it is a C*-algebra (with unit): the involution on $M_n(\mathbb{C})$ is given by hermitian conjugation, i.e.,

$$(a^*)_{ij} = \overline{a_{ji}}, \tag{2.4}$$

and, more abstractly, the involution on $B(H)$ is the map $a \mapsto a^*$, where a^* is the unique operator such that

$$\langle a^*\varphi, \psi\rangle = \langle \varphi, a\psi\rangle, \tag{2.5}$$

for each $\varphi, \psi \in H$. The unit is simply the unit operator $1 \equiv 1_H$. i.e., $1_H(\psi) = \psi$. Finally, it goes without saying that the algebraic structure on $M_n(\mathbb{C})$ (or $B(H)$) is given by matrix (or operator) multiplication and addition, that is,

$$(\lambda \cdot a)\psi = \lambda(a\psi); \tag{2.6}$$

$$(a+b)\psi = a\psi + b\psi; \tag{2.7}$$

$$(ab)\psi = a(b\psi), \tag{2.8}$$

The ***spectrum*** $\sigma(a)$ of $a \in B(H)$ consists of all eigenvalues of a, i.e., $\lambda \in \sigma(a)$ iff there exists $\psi \neq 0$ such that $a\psi = \lambda\psi$.

2.2 States

The key to the probabilistic setting of quantum mechanics is given by the following quantum counterpart of of a classical probability measure.

Definition 2.1. *Let H be a finite-dimensional Hilbert space. A* **density operator** *is a positive operator ρ on H such that*

$$\mathrm{Tr}\,(\rho) = 1. \tag{2.9}$$

The set of all density operators on H is called $\mathscr{D}(H)$. It is easily seen to be convex.

We recall the definition of the trace, and, afterwards, of positivity.

Lemma 2.1. *If (υ_i) and (υ_i') are bases of H, then for any operator $a : H \to H$,*

$$\sum_i \langle \upsilon_i, a\upsilon_i\rangle = \sum_i \langle \upsilon_i', a\upsilon_i'\rangle.$$

This lemma (which you can prove for yourself) allows us to define the ***trace*** of a by

$$\mathrm{Tr}\,(a) = \sum_i \langle \upsilon_i, a\upsilon_i\rangle, \tag{2.10}$$

where (υ_i) is any basis of H. We obtain

$$\mathrm{Tr}\,(ab) = \sum_{i,j} \langle \upsilon_i, a\upsilon_j\rangle\langle \upsilon_j, b\upsilon_i\rangle = \sum_{i,j} \langle \upsilon_i, b\upsilon_j\rangle\langle \upsilon_j, a\upsilon_i\rangle = \mathrm{Tr}\,(ba). \tag{2.11}$$

If u is ***unitary*** (in that $uu^* = u^*u = 1$), then from either Lemma 2.1 or (2.11),

$$\mathrm{Tr}\,(uau^*) = \mathrm{Tr}\,(a). \tag{2.12}$$

Finally, if $a^* = a$, then we may use the spectral theorem

$$a = \sum_{\lambda \in \sigma(a)} \lambda \cdot e_\lambda; \tag{2.13}$$

$$1_H = \sum_{\lambda \in \sigma(a)} e_\lambda, \tag{2.14}$$

where

$$H_\lambda = \{\psi \in H \mid a\psi = \lambda\psi\}, \tag{2.15}$$

is the eigenspace for $\lambda \in \sigma(a)$ and e_λ is the unique projection $H \to H$ with image H_λ. We may also write the spectral decomposition of $a^* = a$ as

$$a = \sum_{i=1}^{\dim(H)} \lambda_i |\upsilon_i\rangle\langle\upsilon_i|; \tag{2.16}$$

$$1_H = \sum_{i=1}^{\dim(H)} |\upsilon_i\rangle\langle\upsilon_i|, \tag{2.17}$$

where λ_i is the eigenvalue corresponding to the eigenvector υ_i (i.e., $a\upsilon_i = \lambda_i\upsilon_i$), and the υ_i form a basis of H consisting of eigenvectors of a. Taking the trace over the basis in (2.16) then yields

$$\mathrm{Tr}\,(a) = \sum_{i=1}^{\dim(H)} \lambda_i = \sum_{\lambda \in \sigma(a)} m_\lambda \cdot \lambda, \tag{2.18}$$

where $m_\lambda = \dim(H_\lambda)$ is the multiplicity of λ.

We say that an operator $a : H \to H$ is ***positive*** if $\langle\psi, a\psi\rangle \geq 0$ for arbitrary $\psi \in H$, in which case we write $a \geq 0$. Without proof we mention some facts about positivity:

Proposition 2.1. *The following condition on an operator $a : H \to H$ are equivalent:*

1. $\langle\psi, a\psi\rangle \geq 0$ *for arbitrary* $\psi \in H$*;*
2. $a^* = a$ *and* $\sigma(a) \subset \mathbb{R}^+$*;*
3. $a = c^2$ *for some hermitian operator* c*;*
4. $a = b^*b$ *for some operator* b*.*

Being positive, a density operator ρ is hermitian, so by (2.17), we have

$$\rho = \sum_i p_i |\upsilon_i\rangle\langle\upsilon_i|, \; p_i \geq 0, \; \sum_i p_i = 1, \tag{2.19}$$

where the (υ_i) form an orthonormal set in H and $|\upsilon_i\rangle\langle\upsilon_i|$ is the (orthogonal) projection on the one-dimensional subspace $\mathbb{C} \cdot \upsilon_i$. Conversely, an operator of the form (2.19) is a density operator. A special class of density operators stands out:

- Each *unit vector* $\psi \in H$ defines a density operator

$$e_\psi = |\psi\rangle\langle\psi|, \tag{2.20}$$

i.e., the (orthogonal) projection on the one-dimensional subspace $\mathbb{C}\cdot\psi$. A basis of eigenvectors of e_ψ consists of $\upsilon_0 = \psi$ itself, supplemented by any basis $(\upsilon_1, \ldots, \upsilon_{\dim(H)-1})$ of the orthogonal complement of $\mathbb{C}\cdot\psi$. The corresponding probabilities are evidently $p_0 = 1$ and $p_i = 0$ for all $i > 0$.

It makes good sense to copy Definition 1.1, *mutatis mutandis*:

Definition 2.2. *A* **state** *on $B(H)$ is a complex-linear map $\omega : B(H) \to \mathbb{C}$ satisfying:*

1. *$\omega(a) \geq 0$ for each positive $a \in B(H)$, i.e., for each $a \geq 0$ (***positivity***);*
2. *$\omega(1) = 1$ (***normalization***).*

Despite its easy proof, the following result is of fundamental importance.

Theorem 2.1. *If H is finite-dimensional, there is a bijective correspondence between states ω on $B(H)$ and density operators ρ on H, given by*

$$\omega(a) = \mathrm{Tr}\,(\rho a). \tag{2.21}$$

The proof is an exercise; for finite-dimensional H the identification (2.21) even works for any linear map $\omega : B(H) \to \mathbb{C}$, matching it with some $\rho \in B(H)$.

Definition 2.3. *The* **state space** *$S(B(H))$ is the set of all states $\omega : B(H) \to \mathbb{C}$, seen as a subspace of $B(H)^*$ (in the w^*-topology).*

The quantum analogue of Proposition 1.5 is as follows.

Corollary 2.1. *Let H be a finite-dimensional Hilbert space. The state space $S(B(H))$ is isomorphic as a compact convex set to the set $\mathscr{D}(H)$ of density matrices on H.*

The case $H = \mathbb{C}^2$ provides a beautiful illustration of this theorem (see exercise).

Proposition 2.2. *The state space $S(M_2(\mathbb{C}))$ of the 2×2 matrices is isomorphic (as a compact convex set) to the closed unit ball $B^3 = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 \leq 1\}$.*

On this isomorphism, the (extreme) boundary (cf. Definition 1.9)

$$\partial B^3 = S^2 = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 = 1\} \tag{2.22}$$

corresponds to the set of all density matrices $\rho = e_\psi$, where $\psi \in \mathbb{C}^2$ with $\|\psi\| = 1$ (these are exactly the one-dimensional projections on $\mathbb{C}^2$).

The proof is an exercise; you may use the fact that any hermitian 2×2 matrix may be parametrized by $(t,x,y,z) \in \mathbb{R}^4$ as

$$\rho(t,x,y,z) = \tfrac{1}{2}\begin{pmatrix} t+z & x-iy \\ x+iy & t-z \end{pmatrix}. \tag{2.23}$$

In this example, the pure state space of $B(H)$ is by no means empty, and we will now see that also in general, the special density operators e_ψ in (2.20) to some extent play the role of the points $x \in X$. Let $\mathscr{P}(H)$ be the set of all projections on H, that is,

$$\mathscr{P}(H) = \{e \in B(H) \mid e^2 = e^* = e\}, \tag{2.24}$$

and let $\mathscr{P}_1(H) \subset \mathscr{P}(H)$ the the subset of all *one-dimensional* projections, so we have $e \in \mathscr{P}_1(H)$ iff $e = e_\psi$, see (2.20), where $\psi \in H$ is a unit vector. This also means that $e \in \mathscr{P}_1(H)$ iff $e \in \mathscr{P}(H)$ and $\dim(eH) = \mathrm{Tr}\,(e) = 1$.

Proposition 2.3. *A density operator ρ is an extremal point of the convex set $\mathscr{D}(H)$ of all density operators on H iff $\rho = e_\psi$ for some unit vector $\psi \in H$.*

The proof is an exercise. Consequently,

$$P(B(H)) = \mathscr{P}_1(H). \tag{2.25}$$

Denoting the state ω defined by the density operator $\rho = e_\psi$ via (2.21) by ω_ψ,

$$\omega_\psi = \langle \psi, a\psi \rangle; \tag{2.26}$$

to see this, take the trace over some basis that contains ψ.

2.3 Density operators on infinite-dimensional Hilbert spaces

Our first aim is to generalize Theorem 2.1 to infinite-dimensional Hilbert spaces. To this effect, we first need density operators, which in turn hinge on the *trace*, which for finite-dimensional H was defined by (2.10). There are (at least) two difficulties. Firstly, not every operator has a trace in the sense of a naive generalization of (2.10); for example, take $a = 1_H$, so that $\mathrm{Tr}\,(1_H) = \dim(H)$, which may be infinite. Secondly, Lemma 2.1 is no longer valid in general. Let

$$B(H)_+ = \{a \in B(H) \mid a \geq 0\}. \tag{2.27}$$

In view of Exercises 5 and 6, we have a well-defined map

$$\mathrm{Tr} : B(H)_+ \to [0, \infty], \tag{2.28}$$

defined as usual by

$$\mathrm{Tr}\,(a) = \sum_i \langle \upsilon_i, a\upsilon_i \rangle, \tag{2.29}$$

where (υ_i) is an arbitrary basis of H. Thus we may define

$$B_1(H)_+ = \{a \in B(H)_+ \mid \mathrm{Tr}\,(a) < \infty\}, \tag{2.30}$$

where $B(H)_+$ consists of all positive bounded operators on H. We then define the trace-class $B_1(H)$ as the complex linear span of $B_1(H)_+$, i.e. $a \in B_1(H)$ iff $a =$

$\sum_i c_i a_i$ with $c_i \in \mathbb{C}$ and $a_i \in B(H)_+$ (finite sum). It can be shown that the trace is well defined (and linear) by $\mathrm{Tr}\,(\sum_i c_i a_i) = \sum_i c_i \mathrm{Tr}\,(a_i)$, and that it may computed as in the finite-dimensional case, i.e.,

$$\mathrm{Tr} : B_1(H) \to \mathbb{C}; \tag{2.31}$$

$$b \mapsto \sum_i \langle \upsilon_i, b\upsilon_i \rangle, \tag{2.32}$$

independently of the choice of basis. It also continues to satisfy the rules (2.11) and (2.12) from the finite-dimensional case.

It is extremely important for quantum theory that $B_1(H)$ is a two-sided ideal in $B(H)$, i.e., $b \in B_1(H)$ and $a \in B(H)$, then $ba \in B_1(H)$ and $ab \in B_1(H)$. In particular, $\mathrm{Tr}\,(ba)$ is well defined.

We now give a nice illustration of the trace class (see Functional Analysis).

Proposition 2.4. *Let $H = \ell^2 \equiv \ell^2(\mathbb{N})$ and for $f \in \ell^\infty$ define the multiplication operator m_f as usual by $m_f \psi = f\psi$. You should know that m_f is bounded, with*

$$\|m_f\| = \|f\|_\infty. \tag{2.33}$$

Then $m_f \in B_1(H)$ iff $f \in \ell^1(\mathbb{N})$, with $\|m_f\| = \|f\|_1$.

Returning to our basic question to what extent Theorem 2.1 generalizes to infinite-dimensional Hilbert space, we we may first literally copy Definition 2.1: as in the finite-dimensional case, a ***density operator*** is a positive operator ρ on H such that $\mathrm{Tr}\,(\rho) = 1$. Since ρ is positive, this implies $\rho \in B_1(H)$. We define the state space as usual: for any C*-algebra A with unit, we put

$$S(A) = \{\omega \in A^* \mid \omega(a^*a) \geq 0 \,\forall a \in A, \omega(1_A) = 1\}. \tag{2.34}$$

It is an exercise to show that the state space $S(B(H))$ contains the set $\mathscr{D}(H)$ of all density operators on H through (2.21), but if $\dim(H) = \infty$, then $S(B(H))$ *is strictly larger than* $\mathscr{D}(H)$. In other words, each density operator ρ still defines a state ω on $B(H)$ by (2.21), but if $\dim(H) = \infty$, there are states that are not of that form. We denote the set of all states of the form (2.21), where $\rho \in \mathscr{D}(H)$, by $S_n(B(H))$, called the *normal state space of* $B(H)$.

Corollary 2.2. *With $\mathscr{D}(H)$ the convex set of all density operators on H, i.e., of all positive operators $\rho \in B_1(H)$ with* $\mathrm{Tr}\,(\rho) = 1$*, eq.* (2.21) *induces an isomorphisms*

$$S_n(B(H)) \cong \mathscr{D}(H) \tag{2.35}$$

of convex sets (i.e., $\omega \leftrightarrow \rho$). Furthermore, for the corresponding pure states we have

$$P_n(B(H)) \equiv \partial_e S_n(B(H)) \cong \mathscr{P}_1(H), \tag{2.36}$$

so that any normal pure state on $B(H)$ is given by (2.26)*.*

Exercises for week 5 (Inleveropgaven: 2, 3, 4)

1. Prove Theorem 2.1. *Hint:* if $\dim(H) < \infty$, as we assume in all exercises for this week, then $B(H)$ is a Hilbert space in the ***Hilbert–Schmidt*** inner product
$$\langle a, b\rangle = \mathrm{Tr}\,(a^* b). \tag{2.37}$$
2. Prove Proposition 2.2.
3. Prove Proposition 2.3.
4. The variance of a state ω may be defined as in the classical case, viz.
$$\Delta_\omega(a) = \omega(a^2) - \omega(a)^2, \tag{2.38}$$
or, using (2.21), with slight abuse of notation we may write
$$\Delta_\rho(a) = \mathrm{Tr}\,(\rho a^2) - (\mathrm{Tr}\,(\rho a))^2. \tag{2.39}$$
As a major contrast between classical and quantum mechanics, prove that:
   a. For each $\omega \in S(B(H))$ there exists an operator $a = a^*$ such that $\Delta_\omega(a) > 0$.
   b. For each hermitian operator a, some density operator ρ satisfies $\Delta_\rho(a) = 0$ iff the image of ρ lies in some fixed eigenspace of a (i.e., $\rho = \sum_i p_i e_{\upsilon_i}$, where all $p_i > 0$, $\sum_i p_i = 1$, and $a\upsilon_i = \lambda\upsilon_i$ where λ is independent of i).
5. Find an operator $a \in B(H)$ and bases (υ_i) and (υ'_i) of H for which
$$\sum_i \langle \upsilon_i, a\upsilon_i\rangle \neq \sum_i \langle \upsilon'_i, a\upsilon'_i\rangle.$$
it is even possible that one of these expressions is finite, whereas the other is infinite!
6. Let $a \in B(H)$ be *positive*, i.e., $a \geq 0$ in the usual sense that $\langle \psi, a\psi\rangle \geq 0$ for each $\psi \in H$. Show that for any two bases (υ_i) and (υ'_i) of H,
$$\sum_i \langle \upsilon_i, a\upsilon_i\rangle = \sum_i \langle \upsilon'_i, a\upsilon'_i\rangle,$$
where both sides may be infinite. *Hint:* for a given basis (υ_i) and operator $a \geq 0$, show that for each unitary operator u,
$$\sum_i \langle \upsilon_i, u^* a u\upsilon_i\rangle = \sum_i \langle \upsilon_i, a\upsilon_i\rangle.$$

2.4 Functional analysis of the trace class

We now analyze the trace-class operators more deeply in case that H is an *separable* Hilbert space. We first define square roots of *positive* operators $c \geq 0$ (we avoid our generic notation a here, since below we will take $c = a^*a$). This can be done through the continuous functional calculus $C(\sigma(c)) \to C^*(c)$, since if $c \geq 0$, then $\sigma(a) \subset \mathbb{R}^+$, and hence $\sqrt{\cdot}$ is defined on $\sigma(c)$. Alternatively, one may use the following construction (which is actually is specialization of the previous approach).

Lemma 2.2. *For any positive operator $c \in B(H)$ there is a (unique) positive operator $\sqrt{c} \in C^*(c)$ that satisfies $\sqrt{c}^2 = c$ (called the* **square root** *of c).*

Proof. If necessary, first rescale c so that $\|c\| \leq 1$, take the power series for

$$\sqrt{1-x} = \sum_{k\geq 0} t_k x^k$$

(in which $t_0 = 1$), which converges absolutely for $|x| \leq 1$, and put

$$\sqrt{c} = \sum_{k\geq 0} t_k (1_H - c)^k.$$

As in the numerical case, squaring the series and rearranging terms yields $\sqrt{c}^2 = c$ (since uniqueness will not be needed, we omit the proof). □

We are now in a position to define the ***absolute value*** $|a|$ of any $a \in B(H)$ by

$$|a| = \sqrt{a^*a}, \tag{2.40}$$

since $a^*a \geq 0$. By construction (of the square root), $|a| \geq$ and if $a \geq 0$, then $|a| = a$. This allows us to redefine the trace-class operators as

$$B_1(H) = \{a \in B(H) \mid \mathrm{Tr}\,(|a|) < \infty\}. \tag{2.41}$$

Without proof we state some pretty deep analytic results about $B_1(H)$.

Theorem 2.2. *1. Any trace-class operator is compact, i.e.,*

$$B_1(H) \subset B_0(H) \subset B(H), \tag{2.42}$$

where $B_0(H)$ is the C-algebra of* compact operators *on H (see below).*

2. *The trace-class operators $B_1(H)$ form a Banach space in the* new *norm*

$$\|a\|_1 = \mathrm{Tr}\,(|a|). \tag{2.43}$$

3. *The trace-class operators $B_1(H)$ form a two-sided ideal in $B(H)$; in other words, if $b \in B_1(H)$ and $a \in B(H)$, then $ba \in B_1(H)$ and $ab \in B_1(H)$.*
4. *With the same notation we have*

$$|\mathrm{Tr}\,(ab)| \leq \|a\|\|b\|_1. \tag{2.44}$$

5. The map

$$\mathrm{Tr} : B_1(H) \to \mathbb{C}; \tag{2.45}$$
$$b \mapsto \sum_i \langle \upsilon_i, b\upsilon_i \rangle, \tag{2.46}$$

where (υ_i) is some basis of H, is well defined, linear, and independent of the choice of basis.

6. We then have

$$\mathrm{Tr}\,(ab) = \mathrm{Tr}\,(ba), \tag{2.47}$$

and hence

$$\mathrm{Tr}\,(ubu^*) = \mathrm{Tr}\,(b), \tag{2.48}$$

for any $b \in B_1(H)$ and any unitary u.

7. We have the dualities

$$B_0(H)^* \cong B_1(H); \tag{2.49}$$
$$B_1(H)^* \cong B(H), \tag{2.50}$$

where the symbol $\cong$ stands for isometric isomorphism*, given as follows:*

- *Any norm-continuous linear map $\omega : B_0(H) \to \mathbb{C}$ takes the form* (2.21) *for some $b \in B_1(H)$, and* vice versa*, giving a bijective correspondence between $\omega \in B_0(H)^*$ and $b \in B_1(H)$ satisfying*

$$\|\omega\| = \|b\|_1; \tag{2.51}$$

- *Any norm-continuous linear map $\chi : B_1(H) \to \mathbb{C}$ takes the form*

$$\chi(b) = \mathrm{Tr}\,(ba), \tag{2.52}$$

for some $a \in B(H)$, and vice versa*, giving a bijective correspondence between $\chi \in B_1(H)^*$ and $a \in B(H)$ satisfying*

$$\|\chi\| = \|a\|. \tag{2.53}$$

8. It follows from (2.49) - (2.50) *that*

$$B_0(H)^{**} \cong B(H); \tag{2.54}$$
$$B(H)^* \cong B_1(H)^{**}. \tag{2.55}$$

Note that (2.49) - (2.50) reflect the (hopefully) familiar function space dualities

$$\ell_0(\mathbb{N})^* \cong \ell^1(\mathbb{N}); \tag{2.56}$$
$$\ell^1(\mathbb{N})^* \cong \ell^\infty. \tag{2.57}$$

Recall that $a \in B_0(H)$ iff a is a norm-limit of finite-rank operators (these are bounded operators whose range is finite-dimensional, such as finite-dimensional projections). We note that $B_0(H)$ is a C*-algebra in $B(H)$ under the usual operator norm. Moreover, $B_0(H)$ is a two-sided ideal in $B(H)$, like $B_1(H)$. There is a subtlety in (2.42): as a normed space, $B_0(H)$ simply inherits the norm of $B(H)$, in which it is complete. Clearly, $B_1(H)$ also inherits the norm of $B(H)$, but that is the wrong one: firstly, $B_1(H)$ is not complete in the operator norm (indeed, its completion is $B_0(H)$) and secondly, the operator norm is the wrong one for the fundamental dualities below. The only regrettable thing about $B_1(H)$ is that it fails to be a C*-algebra: in the usual operator norm it is not complete, whereas in the trace-norm (2.43), in which it is complete, the fundamental C^*-axiom $\|a^*a\| = \|a\|^2$ does not hold. Note that finiteness of the trace is immediate from (2.44), in which the special case $a = 1_H$ gives $\mathrm{Tr}\,(b) \leq \|b\|_1$. Independence of the basis follows by decomposing b as a sum of positive operators, using linearity of the trace, and the second exercise.

Since the C*-algebra $B_0(H)$ for infinite-dimensional H does not have a unit (like $A = C_0(X)$ for non-compact X), the state space is defined as

$$S(A) = \{\omega \in A^* \mid \omega(a^*a) \geq 0 \forall a \in A, \|\omega\| = 1\}. \tag{2.58}$$

Corollary 2.3. *With $\mathscr{D}(H)$ the convex set of all density operators on H, i.e., of all positive operators $\rho \in B_1(H)$ with* $\mathrm{Tr}\,(\rho) = 1$*, eq.* (2.52) *induces isomorphisms*

$$S(B_0(H)) = S_n(B(H)) \cong \mathscr{D}(H) \tag{2.59}$$

of convex sets (i.e., $\omega \leftrightarrow \rho$). Furthermore, for the corresponding pure states we have

$$P(B_0(H)) = P_n(B(H)) \cong \mathscr{P}_1(H), \tag{2.60}$$

where $P(B_0(H)) = \partial S(B_0(H))$ and $P_n(B(H)) = \partial S_n(B_0(H))$, cf. (2.25)*, so that any pure state on $B_0(H)$, as well as any normal pure state on $B(H)$, is given by* (2.26).

More precisely, the equality in (2.59) should be read as saying that any state ω on $B_0(H)$ has a unique *normal* extension to $B(H)$, and similarly in (2.60).

The adjective 'normal' for states is often used in a different way (which, as we shall see, is compatible with our previous use). First, we say that a family (e_i) of projections is *orthogonal* $e_i^* = e_i$ and $e_i e_j = \delta_{ij} e_i$. For example, if (υ_i) is a basis (= orthonormal basis) of H, then $e_i = |\upsilon_i\rangle\langle\upsilon_i|$ is an orthogonal family of projections.

Definition 2.4. *A state $\omega : B(H) \to \mathbb{C}$ is called* **normal** *if for each orthogonal family of projections one has*

$$\omega\left(\sum_i e_i\right) = \sum_i \omega(e_i). \tag{2.61}$$

Here $\sum_i e_i$ may be defined either as the projection on the smallest closed subspace K of H that contains each $e_i H$, or as the strong operator sum (i.e. $\sum_i e_i \psi$ converges in H for each fixed $\psi \in H$). If $e_i = |\upsilon_i\rangle\langle\upsilon_i|$, as above, then the additivity condition (2.61) obviously implies $\sum_i \omega(e_i) = 1$. The main result on normal states is as follows.

Theorem 2.3. *A state ω on $B(H)$ takes the form $\omega(a) = \mathrm{Tr}(\rho a)$ for some (unique) density operator $\rho \in \mathscr{D}(H)$ iff it is normal (in the sense of the preceding definition).*

We now construct pure states on $B(H)$ that fail to be normal. We first recall some facts from functional analysis.

Definition 2.5. *Let H be a Hilbert space, and let $a \in B(H)$ be a bounded operator.*

1. *The* **resolvent** *$\rho(a)$ consists of all $z \in \mathbb{C}$ for which $a - z$ has a bounded inverse.*
2. *The* **spectrum** *$\sigma(a) = \mathbb{C}\backslash\rho(a)$ is the complement of the resolvent.*

If $\dim(H) < \infty$, then $\sigma(a)$ is the set of eigenvalues of a. If H is infinite-dimensional, then even self-adjoint operators $a = a^* \in B(H)$ may not have eigenvalues at all. If they do, we denote the set of eigenvalues of a by $\sigma_d(a)$; the rest is called the *continuous spectrum* $\sigma_c(a) = \sigma(a)\backslash\sigma_d(a)$. For example, if $H = L^2(0,1)$ and $a\psi(x) = x\psi(x)$ (this is a bounded version of the position operator of quantum mechanics; the unbounded version on $H = L^2(\mathbb{R})$ will be discussed later in these notes). In that case, $\sigma(a) = \sigma_c(a) = [0,1]$ and a has no eigenvalues, i.e. $\sigma_d(a) = \emptyset$. We also recall the spectral theorem for $a = a^* \in B(H)$. Let $C^*(a)$ be the C*-algebra generated by a and 1_H (i.e., the norm-closure of the algebra of all polynomials in a).

Theorem 2.4. *If $a^* = a \in B(H)$, then $C^*(a)$ is commutative, $\sigma(a) \subset \mathbb{R}$ is compact, and there is an isomorphism of (commutative) C*-algebras*

$$C(\sigma(a)) \cong C^*(a), \tag{2.62}$$

written $f \mapsto f(a)$, which is unique if it is subject to the following conditions:

- *the unit function $1_{\sigma(a)} : \lambda \mapsto 1$ corresponds to the unit operator 1_H;*
- *the identity function $\mathrm{id}_{\sigma(a)} : \lambda \mapsto \lambda$ is mapped to the given operator a.*

Furthermore, this **continuous functional calculus** *satisfies the rule*

$$(tf + g)(a) = tf(a) + g(a); \tag{2.63}$$

$$(fg)(a) = f(a)g(a); \tag{2.64}$$

$$f(a)^* = \overline{f}(a). \tag{2.65}$$

For any $\lambda \in \sigma(a)$, we now define a linear map

$$\omega_\lambda : C^*(a) \to \mathbb{C}; \tag{2.66}$$

$$\omega_\lambda(f(a)) = f(\lambda). \tag{2.67}$$

It follows from (2.63) that ω_λ is linear, while (2.64) - (2.65) that it is also positive. Furthermore, $\omega_\lambda(1_H) = \omega_\lambda(1_{\sigma(a)}(a)) = 1_{\sigma(a)}(a))(\lambda) = 1$, where $1_X : X \to \mathbb{C}$ is the function that equals 1 for each $x \in X$, for any set X, and we used the property $1_{\sigma(a)}(a)) = 1_H$ of the continuous functional calculus. Hence ω_λ is a state on $C^*(a)$, which is *multiplicative* in the sense that $\omega(bc) = \omega(b)\omega(c)$ for all $b, c \in C^*(a)$.

A fundamental result, which we mention without proof, is that a state on a commutative C*-algebras is multiplicative iff it is pure, so ω_λ is a *pure state* on $C^*(a)$.

Proposition 2.5. *Let $a = a^* \in B(H)$ have non-empty continuous spectrum $\sigma_c(a)$, so that there is some $\lambda \in \sigma_c(a) \subseteq \sigma(a)$ that is not an eigenvalue of a. Then there exists an extension of the functional ω_λ from $C^*(a)$ to $B(H)$ that defines a pure state on $B(H)$, and no such extension is normal.*

Proof. Since ω_λ is a state, its satisfies $\omega_\lambda(1_H) = 1$ and hence $\|\omega_\lambda\| = 1$ (note that 1_H is the unit of both $B(H)$ and $C^*(a)$). The Hahn–Banach Theorem yields a linear extension $\omega'_\lambda : B(H) \to \mathbb{C}$ of ω_λ with $\|\omega'_\lambda\| = 1$. Trivially (see previous comment in brackets), $\omega'_\lambda(1_H) = 1$. We now use a lemma from C*-algebras: if $\omega : B(H) \to \mathbb{C}$ is linear and bounded and satisfies $\|\omega\| = \omega(1_H)$, then ω is positive, and hence defines a state on $B(H)$. it follows that ω'_λ is a state on $B(H)$. A technical argument based on the Krein–Milman Theorem of functional analysis shows that ω'_λ mat be chosen so as to be a *pure* state on $B(H)$. All this is true for any $\lambda \in \sigma(a)$. The final step of the proof is to show that if $\lambda \in \sigma_c(a)$, then ω'_λ cannot be normal (exercise). □

2.5 The Born measure and the Born rule of quantum mechanics

The ***Born rule*** provides a link between the mathematical formalism of quantum theory and experiment, and as such is almost single-handedly responsible for practically all predictions of quantum physics. On a par with the Heisenberg uncertainty relations the Born rule is often seen as a turning point where indeterminism entered fundamental physics. For these two reasons, its importance cannot be overestimated.

A simple version of the Born rule was first stated by Max Born (1882-1970) in the context of scattering theory, following a slightly earlier paper in which he famously omitted the absolute value squared signs (though he corrected this is a footnote added in proof). The modern formulation below is due to von Neumann.

We first construct the Born measure, which is a successful attempt to adapt Theorem 1.3 to quantum mechanics. We initially assume that H is finite-dimensional. For given $a = a^* \in B(H)$ and $\Delta \subset \sigma(a)$, we use the notation

$$e_\Delta = \sum_{\lambda \in \Delta} e_\lambda, \tag{2.68}$$

where e_λ is the projection onto the eigenspace $H_\lambda = \{\psi \in H \mid a\psi = \lambda\psi\}$, see (2.15).

Theorem 2.5. *A density operator ρ on H and a hermitian operator $a : H \to H$ jointly yield a probability distribution p_a on the spectrum $\sigma(a)$ by*

$$p_a(\lambda) = \mathrm{Tr}(\rho e_\lambda). \tag{2.69}$$

The associated probability measure μ_a, called the **Born measure** *associated to the observable a and the density operator ρ, is given at $\Delta \subseteq \sigma(a)$ by (cf.* (2.68)*)*

$$\mu_a(\Delta) = \mathrm{Tr}(\rho e_\Delta). \tag{2.70}$$

In terms of the state ω associated to ρ, cf. (2.21)*, we simply have*

$$p_a(\lambda) = \omega(e_\lambda); \tag{2.71}$$

$$\mu_a(\Delta) = \omega(e_\Delta). \tag{2.72}$$

Note that this relationship between p_a and μ_a is the same as the one in (1.93) - (1.94): from the point of view of Theorem (1.5), we have

$$\mu_a(\Delta) = \omega(e_\Delta) = \omega\left(\sum_{\lambda\in\Delta} e_\lambda\right) = \sum_{\lambda\in\Delta} \omega(e_\lambda) = \sum_{\lambda\in\Delta} p_a(\lambda). \tag{2.73}$$

It is an exercise to verify that p_a is indeed a probability distribution on $\sigma(a)$, or, equivalently, that μ_a is a probability measure on $\sigma(a)$. Expectation values may then be rewritten as follows:

$$\omega(a) = \sum_{\lambda\in\sigma(a)} \lambda \cdot p_a(\lambda). \tag{2.74}$$

If $\rho = e_\psi$, writing p_a^ψ for the associated probability, (2.69) yields

$$p_a^\psi(\lambda) = \langle \psi, e_\lambda \psi\rangle = \|e_\lambda \psi\|^2. \tag{2.75}$$

If in addition $\lambda \in \sigma(a)$ is non-degenerate, so that $e_\lambda = |\upsilon_\lambda\rangle\langle\upsilon_\lambda|$ for some unit vector υ_λ with $a\upsilon_\lambda = \lambda\upsilon_\lambda$, then the Born rule (2.70) assumes its original form

$$p_a^\psi(\lambda) = |\langle\psi, \upsilon_\lambda\rangle|^2. \tag{2.76}$$

What does the Born measure mean physically? First, quantum mechanics stipulates that measurements of some observable a always have outcomes in its spectrum $\sigma(a)$. Following up on this, the ***Born rule states that:***

If an observable a is measured in a state ω, then the probability that the outcome lies in $\Delta \subset \sigma(a)$ equals $\mu_a(\Delta)$, where μ_a is the Born measure defined by a and ω. In particular, the probability that the outcome is $\lambda \in \sigma(a)$ is equal to $p_a(\lambda)$.

We now generalize the Born measure to infinite-dimensional Hilbert spaces. In that case the underlying probability distribution may not exist, so we will only talk about measures. Combining Theorem 2.4 with Theorem 1.1 gives the Born measure:

Corollary 2.4. *Let H be a Hilbert space, let $a^* = a \in B(H)$, and let ω be a state on $B(H)$. There exists a unique probability measure μ_a on the spectrum $\sigma(a)$ such that*

$$\omega(f(a)) = \int_{\sigma(a)} d\mu_a f, \ f \in C(\sigma(a)). \tag{2.77}$$

This measure μ_a on $\sigma(a)$ is called the **Born measure** *(defined by ω and a).*

Wha happens conceptually is that the state ω is restricted from $B(H)$ to $C^*(a)$ and subsequently transferred to a state on $C(\sigma(a))$, where it is given by a probability measure on $\sigma(a)$. Thus it is the restriction of states to commutative algebras that gives rise to the probability structure of quantum mechanics, at least according to the Copenhagen Interpretation. Heisenberg wrote the following about this procedure:

> 'One may call these uncertainties objective, in that they are simply a consequence of the fact that we describe the experiment in terms of classical physics; they do not depend in detail on the observer. One may call them subjective, in that they reflect our incomplete knowledge of the world.'

Taking $f(\lambda)=\lambda$ and hence $f(a)=a$, we obtain the expectation value

$$\omega(a)=\int_{\sigma(a)} d\mu_a(\lambda)\,\lambda, \tag{2.78}$$

As an important special case, suppose that

$$\omega(a)=\langle\psi,a\psi\rangle, \tag{2.79}$$

for some unit vector $\psi\in H$; it is easily verified that ω is a state. If $\sigma(a)=\sigma_d(a)$,

$$\mu_a(\{\lambda\}) = \|e_\lambda\psi\|^2; \tag{2.80}$$
$$\mu_a(\Delta) = \|e_\Delta\psi\|^2, \tag{2.81}$$

where e_λ is the projection onto the eigenspace H_λ, where, as before,

$$H_\lambda = \{\psi\in H \mid a\psi=\lambda\psi\}; \tag{2.82}$$
$$e_\Delta = \sum_{\lambda\in\Delta} e_\lambda. \tag{2.83}$$

In order to generalize (2.81) to general spectra we need the following extension of Theorem 2.4, in which $\mathscr{B}(\sigma(a))$ is the space of bounded (Borel) measurable functions $f:\sigma(a)\to\mathbb{C}$ (this is a C*-algebra under the same operations as $C(\sigma(a))$).

Theorem 2.6. *Let $a^*=a\in B(H)$. The isomorphism $C(\sigma(a))\to C^*(a)$ of Theorem 2.4 has a unique extension to a homomorphism*

$$\mathscr{B}(\sigma(a))\to B(H),\ f\mapsto f(a), \tag{2.84}$$

i.e. (2.63) - (2.65) *hold, but instead of isometry we now have*

$$\|f(a)\|\leq\|f\|_\infty, \tag{2.85}$$

with equality for $F\in C(\sigma(a))$. Finally, (2.77) *remains valid for each $f\in\mathscr{B}(\sigma(a))$.*

For any measurable subset $\Delta\subseteq\sigma(a)$ we then *define* an operator $e_\Delta\in B(H)$ by

$$e_\Delta=1_\Delta(a). \tag{2.86}$$

It is an exercise to show that e_Δ is a projection, which in finite dimension coincides with the one defined in (2.68). Eq. (2.77) then gives

$$\mu_a(\Delta)=\omega(e_\Delta), \tag{2.87}$$

and in the special case (2.79) we recover (2.81).

Exercises for week 6 (inleveropgave for students with functional analysis: nos. 1, 2, 5; for those without: 4, 5, 6)

1. Prove Theorem 2.3.
2. Finish the proof of Proposition 2.5 by showing that if $\lambda \in \sigma_c(a)$, then ω'_λ cannot be normal.
3. For the projection $a = e_L$ defined by a linear subspace $L \subset H$, so that $\sigma(a) = \{0, 1\}$, compute the Born probabilities $p_a(0)$ and $p_a(1)$ for a state (2.79).
4. In the previous exercise, specialize your answer to the case where $L = \mathbb{C} \cdot \varphi$ is one-dimensional, so that $e_L = |\varphi\rangle\langle\varphi|$ for some unit vector $\varphi \in H$.
5. Prove Corollary 2.4. In particular, show that the functional $f \mapsto \langle\psi, f(a)\psi\rangle$ is positive and linear, and that it satisfies $1_{\sigma(a)} \mapsto 1$.
6. Prove (2.74).

Appendix: proofs

This section contains proofs of the claims in the preceding two sections. Some technical arguments rely on the *polar decomposition* of operators, which in turn uses the concept of a *partial isometry*, of which a *unitary (operator)* is a special case,

Definition 2.6. *Let H be a Hilbert space.*

1. *An* **isometry** *is an operator $u \in B(H)$ such that $u^*u = 1_H$.*
2. *A* **partial isometry** *is an operator $u \in B(H)$ for which $u^*u = p$ is a projection.*
3. *A* **unitary** *s an operator $u \in B(H)$ such that $u^*u = uu^* = 1_H$.*

Exercise 2.1. *Let $u \in B(H)$. Show that:*

- *u is an isometry iff $\|u\psi\| = \|\psi\|$ for each $\psi \in H$.*
- *u is unitary iff it is an invertible partial isometry, or, equivalently, a surjective isometry (in particular, an isometry is automatically injective).*
- *If $\dim(H) < \infty$, then an isometry is automatically unitary (and give a counterexample to this if $\dim(H) = \infty$).*

Lemma 2.3. *Let H be a Hilbert space with a partial isometry $u \in B(H)$.*

- *Also u^* is a partial isometry, or, equivalently, $uu^* = q$ is a projection.*
- *The kernel of u is $(pH)^\perp$, and its range is qH, and u is unitary from pH to qH.*
- *Conversely, an operator u on H for which there is a closed subspace $L \subset H$ on which u is isometric (and hence unitary from L to $\mathrm{ran}(u) = uH$), whilst it is identically zero on $L^\perp$, is a partial isometry (note that $\mathrm{ran}(u)$ is automatically closed if u is an isometry).*
- *If $u \neq 0$, then $\|u\| = 1$.*
- *An partial isometry u is unitary iff $u^*u = uu^* = 1_H$ (i.e., $p = q = 1_H$).*

The following proposition (due to von Neumann) gives the ***polar decomposition*** of an operator, which mimics the polar decomposition $z = r\exp(i\varphi)$ of $z \in \mathbb{C}$.

Proposition 2.6. *For $a \in B(H)$, assumed nonzero, the operator u given by*

$$u|a|\psi = a\psi, \ |a|\psi \in \mathrm{ran}\,|a|; \tag{2.88}$$
$$u\psi = 0, \ \psi \in (\mathrm{ran}\,|a|)^{\perp} = \ker|a|, \tag{2.89}$$

1. is well defined;
2. is a partial isometry (and hence has norm $\|u\| = 1$);
3. is unitary from $(\mathrm{ran}\,|a|)^{-}$ to $(\mathrm{ran}\,a)^{-}$;
4. satisfies

$$\||a|\psi\| = \|a\psi\|; \tag{2.90}$$
$$u^*u|a| = |a| = |a|u^*u. \tag{2.91}$$

Given that u is a partial isometry, it is characterized by the two properties

$$\ker u = \ker a; \tag{2.92}$$
$$a = u|a|. \tag{2.93}$$

Furthermore, if $a \neq 0$, then a is invertible iff u is unitary.

Lemma 2.4. *1. For any $b \in B_1(H)$ we have*

$$\|b\| \leq \|b\|_1. \tag{2.94}$$

2. Any trace-class operator is compact, i.e., $B_1(H) \subset B_0(H)$.
3. For $a \in B(H)$ and $b \in B_1(H)$ one has

$$|\mathrm{Tr}\,(ba)| \leq \|b\|_1\|a\|. \tag{2.95}$$

Part 4 will shortly be improved to $B_1(H)$ actually being a Banach space.

Proof. 1. For every $\varepsilon > 0$ there is a unit vector $\psi \in H$ such that $\|b\|^2 \leq \|b\psi\|^2 + \varepsilon$. Put $b' = (b^*b)^{1/4}$, and note that $\|(b^*b)^{1/4}\|^2 = \||b|\| = \|b\|$ by (2.90) Completing ψ to a basis (υ_i), and noting that

$$\sum_i \|(b^*b)^{1/4}\upsilon_i\|^2 = \sum_i \langle (b^*b)^{1/4}\upsilon_i, (b^*b)^{1/4}\upsilon_i\rangle = \sum_i \langle \upsilon_i, |b|\upsilon_i\rangle = \|b\|_1,$$

we have

$$\|b\| = \|(b^*b)^{1/4}\|^2 \leq \|(b^*b)^{1/4}\psi\|^2 + \varepsilon \leq \sum_i \|(b^*b)^{1/4}\upsilon_i\|^2 + \varepsilon = \|b\|_1 + \varepsilon.$$

Since this holds for all $\varepsilon \geq 0$, one has (2.94).

2. Let $b \in B_1(H)$. Since $\sum_i \langle \upsilon_i, |b|\upsilon_i\rangle < \infty$, for each $\varepsilon > 0$ we can find n such that $\sum_{i>n} \langle \upsilon_i, |b|\upsilon_i\rangle < \varepsilon$. Let e_n be the projection onto the linear span of $\{\upsilon_i\}_{i=1,\dots,n}$. Using $\|a\|^2 = \|aa^*\|$ (which is valid by (2.94)), we have

$$\|e_n^\perp |b|^{1/2}\|^2 = \|e_n^\perp |b| e_n^\perp\| \leq \|e_n^\perp |b| e_n^\perp\|_1 = \sum_i \langle \upsilon_i, e_n^\perp |b| e_n^\perp \upsilon_i\rangle = \sum_{i>n} \langle \upsilon_i, |b|\upsilon_i\rangle < \varepsilon,$$

for $|(e_n^\perp |b| e_n^\perp)| = e_n^\perp |b| e_n^\perp$ (if $a \geq 0$ then $b^*ab \geq 0$ for any $a, b \in B(H)$). Since $e_n^\perp = 1 - e_n$, it follows that $e_n|b|^{1/2} \to |b|^{1/2}$ in the norm topology. Since each operator $e_n|b|^{1/2}$ obviously has finite rank, $|b|^{1/2}$ and hence $|b|$ is compact. But b has polar decomposition $b = u|b|$ and $B_0(H)$ is a two-sided ideal in $B(H)$.
3. We just showed that b is compact. By Theorem 2.7 below, also b^*b is compact, and since it is self-adjoint, Theorem 2.7 applies. This gives an expansion

$$|b| = \sum_{i=1}^{m \leq n} p_i |\upsilon_i\rangle\langle\upsilon_i|; \tag{2.96}$$

although the sum may be infinite, this is no problem, as it is norm-convergent. It follows from Lemma 2.3 and Proposition 2.6 that $(\upsilon_i' = u\upsilon_i)$ also forms an orthonormal set, like the υ_i themselves, since the closed linear space spanned by the unit vectors υ_i is just $(\text{ran}|b|)^-$ and u is unitary from this space onto its image $(\text{ran} b)^-$. Taking the trace over any basis that contains the vectors υ_i', we compute

$$\begin{aligned} |\text{Tr}\,(ba)| &= |\text{Tr}\,(u|b|u^*ua)| = |\sum_i p_i \langle \upsilon_i', ua\upsilon_i'\rangle| \\ &\leq \sum_i p_i |\langle \upsilon_i', ua\upsilon_i'\rangle| \leq \sum_i p_i \|a\|\|u\|\|\upsilon_i\| = \|b\|_1 \|a\|, \end{aligned} \tag{2.97}$$

where we used $\|b\|_1 = \sum_i p_i$, given (2.96).

We leave it to the reader to prove that the trace-class operators $B_1(H)$ form a vector space with norm (2.43).

We recall the spectral theorem for compact operators:

Theorem 2.7. *If $a = a^*$ is a compact operator on a separable Hilbert space H:*

1. H has an orthonormal basis consisting of eigenvectors of a.
2. The eigenspaces H_λ corresponding to different eigenvalues λ are orthogonal.
3. One has the spectral decompositions

$$\sum_{\lambda \in \sigma_d(a)} e_\lambda = 1; \tag{2.98}$$

$$\sum_{\lambda \in \sigma_d(a)} \lambda \cdot e_\lambda = a, \tag{2.99}$$

where the first sum converges strongly, that is, for each $\psi \in H$ one has

$$\sum_{\lambda \in \sigma_d(a)} e_\lambda \psi = \psi, \tag{2.100}$$

whereas the second sum converges in operator norm (and hence strongly, too).
4. *The spectrum of a is real.*
5. *If $\lambda \in \sigma(a)$ and $\lambda \neq 0$, then $\lambda \in \sigma_d(a)$ and* $\dim(H_\lambda) < \infty$.
6. *If* $\dim(H) = \infty$, *then $\sigma_d(a)$ has 0 as its only accumulation point.*

Conversely, if properties 1–6 hold, then a is compact and self-adjoint.

Proposition 2.7. *The map*

$$\mathrm{Tr} : B_1(H) \to \mathbb{C}; \tag{2.101}$$
$$b \mapsto \sum_i \langle \upsilon_i, b\upsilon_i \rangle, \tag{2.102}$$

where (υ_i) is some basis of H, is well defined, (obviously) linear, and independent of the choice of basis. Furthermore, we have

$$\mathrm{Tr}\,(ab) = \mathrm{Tr}\,(ba). \tag{2.103}$$

Proof. Taking $a = 1_H$ in (2.95), we have $|\mathrm{Tr}\,(b)| \leq \|b\|_1 < \infty$ for $b \in B_1(H)$. Independence of the choice of basis follows by first decomposing $b = b' + ib''$, with $b' = \frac{1}{2}(b + b^*)$ and $b'' = -\frac{1}{2}i(b - b^*)$ self-adjoint, as usual, and subsequently using (2.99) to write $b' = b'_+ - b'_-$, with

$$b'_\pm = \pm \sum_{\lambda \in \sigma_d(b') \cap \mathbb{R}_\pm} \lambda \cdot e_\lambda, \tag{2.104}$$

and likewise for b''. This makes b is a linear combination of four positive operators, whence the claim follows from Exercise 5 and the obvious linearity of (2.102).

To establish (2.103), we first note that $\mathrm{Tr}\,(au) = \mathrm{Tr}\,(ua)$ for any unitary u; this is the same as

$$\mathrm{Tr}\,(uau^*) = \mathrm{Tr}\,(a), \tag{2.105}$$

The claim then follows from the following (generally useful) lemma. □

Lemma 2.5. *Any $a \in B(H)$ is a linear combination of at most four unitaries.*

This can be proved from the continuous functional calculus.

Theorem 2.8. *We have dualities*

$$B_0(H)^* \cong B_1(H); \tag{2.106}$$
$$B_1(H)^* \cong B(H), \tag{2.107}$$

where the symbol $\cong$ stands for isometric isomorphism. *Explicitly:*

- *Any norm-continuous linear map $\omega : B_0(H) \to \mathbb{C}$ takes the form*

$$\omega(a) = \mathrm{Tr}\,(\rho a), \tag{2.108}$$

for some $\rho \in B_1(H)$*, and* vice versa*, giving a bijective correspondence between* $\omega \in B_0(H)^*$ *and* $\rho \in B_1(H)$ *satisfying*

$$\|\omega\| = \|\rho\|_1; \tag{2.109}$$

- *Any norm-continuous linear map* $\chi : B_1(H) \to \mathbb{C}$ *takes the form*

$$\chi(\rho) = \mathrm{Tr}\,(\rho a), \tag{2.110}$$

for some $a \in B(H)$*, and* vice versa*, giving a bijective correspondence between* $\chi \in B_1(H)^*$ *and* $a \in B(H)$ *satisfying*

$$\|\chi\| = \|a\|. \tag{2.111}$$

Proof. It is clear from (2.95) that $B_1(H) \subseteq B_0(H)^*$, with $\|\omega\| \leq \|\rho\|_1$. For the opposite direction, we return to the projections e_n in the proof of part 2 of Lemma 2.4. Taking the trace over the basis (υ_i), we have

$$\|\rho\|_1 = \mathrm{Tr}\,(|\rho|) = \lim_n \mathrm{Tr}\,(e_n|\rho|e_n) = \lim_n \mathrm{Tr}\,(e_n|\rho|) = \lim_n \mathrm{Tr}\,(e_n u^*\rho) = \lim_n \omega(e_n u^*);$$

since $\omega(e_n u^*) \geq 0$ we have $\omega(e_n u^*) \leq \|\omega\|\|e_n u^*\| \leq \|\omega\|$, whence $\|\rho\|_1 \leq \|\omega\|$ (note that the limiting procedure is necessary here, since $\omega(u^*)$ would not be defined because typically u^* is not compact). This proves (2.109).

To prove (2.106), it remains to be shown that every $\omega \in B_0(H)^*$ can be represented as (2.108). Noting that $B_0(H)$ is the norm-closure of the linear span of all operators of the sort $a = |\psi\rangle\langle\varphi|$, where $\psi, \varphi \in H$ are unit vectors, the functional ω is determined by its values on those operators. Given ω, we define ρ by its matrix elements

$$\langle\varphi, \rho\psi\rangle = \omega(|\psi\rangle\langle\varphi|). \tag{2.112}$$

It is easy to show that

$$\mathrm{Tr}\,(\rho|\psi\rangle\langle\varphi|) = \langle\varphi, \rho\psi\rangle,$$

which gives (2.108) on operators a of the said form, upon which the general case follows by continuity.

We now prove (2.50). As in the previous case, the inclusion $B(H) \subset B_1(H)^*$ is clear from (2.95), as is the inequality $\|\chi\| \leq \|a\|$. This time, the proof of the opposite inequality uses $\rho = |\psi\rangle\langle\varphi|$, in which case one easily obtains

$$\||\psi\rangle\langle\varphi|\|_1 = \|\psi\|\|\varphi\|, \tag{2.113}$$

which in the case of unit vectors equals unity. Assuming (2.110), this gives

$$|\chi(\rho)| = |\chi(|\psi\rangle\langle\varphi|)| = |\mathrm{Tr}\,(|\psi\rangle\langle\varphi|a)| = |\langle\varphi, a\psi\rangle| \leq \|\chi\|\||\psi\rangle\langle\varphi|\|_1 = \|\chi\|.$$

Combined with the following formula for the norm,

$$\|a\| = \sup\{|\langle\varphi, a\psi\rangle|, \psi, \varphi \in H, \|\psi\| = \|\varphi\| = 1\}, \tag{2.114}$$

this gives $\|a\| \leq \|\chi\|$, and hence (2.111).

Finally, as in the previous case, given χ we find a though its matrix elements

$$\langle\varphi, a\psi\rangle = \chi(|\psi\rangle\langle\varphi|), \tag{2.115}$$

which gives (2.110) on the special trace-class operators $\rho = |\psi\rangle\langle\varphi|$. Noting that the linear span of such operators in dense (in the trace-norm) in $B_1(H)$, once again this gives the general case by continuity. □

Corollary 2.5. *The vector space $B_1(H)$ is complete in the norm* (2.43) *and is a two-sided ideal in $B(H)$: if $a \in B(H)$ and $b \in B_1(H)$, then ab and ba lie in $B_1(H)$.*

Proof. The first claim follows from (2.106) and the completeness of $B_0(H)^*$ (the dual V^* of any normed space V is a Banach space). The second follows from (2.95) and (2.103). □

As a normed space, $B_0(H)$ simply inherits the norm of $B(H)$, in which it is complete. $B_1(H)$ also inherits the norm of $B(H)$, but that is the wrong one: firstly, $B_1(H)$ is not complete in the operator norm (its completion is $B_0(H)$) and secondly, the operator norm is the wrong one for the fundamental dualities stated in Theorem 2.8.

The proof of Theorem 2.6 is based on two lemma's we state without proof:

Lemma 2.6. *Any* positive *function $f \in \mathscr{B}(\sigma(a))$ is a* pointwise *limit of some monotone increasing bounded sequence (f_n) in $C(\sigma(a))$, written $f_n \nearrow f$. That is,*

$$0 \leq f_1(x) \leq \cdots \leq f_n(x) \leq f_{n+1}(x) \leq \cdots \leq c \cdot 1_{\sigma(a)}; \tag{2.116}$$

$$f(x) = \lim_{n\to\infty} f_n(x),\ x \in \sigma(a). \tag{2.117}$$

Lemma 2.7. *If (a_n) is a sequence of positive operators on H for which*

$$0 \leq a_1 \leq \cdots \leq a_n \leq a_{n+1} \leq \cdots \leq c1_H, \tag{2.118}$$

where $a_i \leq a_j$ means that $\langle\psi, a_i\psi\rangle \leq \langle\psi, a_j\psi\rangle$ for each $\psi \in H$, then there exists a unique positive operator a such that $a_n \nearrow a$ strongly, i.e., for each $\psi \in H$,

$$a\psi = \lim_{n\to\infty} a_n\psi. \tag{2.119}$$

Furthermore, $a = \sup_n a_n$ with respect to the partial ordering $\leq$ on the set of positive bounded operators (that is, $a_n \leq a$ for each n, and if $a_n \leq b$ for each n, then $a \leq b$).

Given these lemma's, the proof of Theorem 2.6 is an exercise.

2.6 Tensor products

Our aim is the tensor products of Hilbert spaces, which in the context of quantum mechanics replaces the cartesian product of phase spaces. Although this is not sufficient, we start with the *algebraic* tensor product of two vector spaces A and B.

Proposition 2.8. *Let A and B be (complex) vector spaces. There is a vector space called $A \otimes B$, in words the* **algebraic tensor product** *of A and B (over $\mathbb{C}$), and a map $p : A \times B \to A \otimes B$, such that for any vector space C and any bilinear map $\beta : A \times B \to C$, there is a unique linear map $\beta' : A \otimes B \to C$ such that $\beta = \beta' \circ p$.*

In other words, the following diagram commutes:

$$\begin{array}{ccc} A \times B & \xrightarrow{p} & A \otimes B \\ & \searrow^{\beta} & \downarrow{\scriptstyle \exists! \beta'} \\ & & C \end{array} \tag{2.120}$$

This universal property implies that $A \otimes B$ is unique up to isomorphism (exercise).

Proof. In preparation for an explicit construction of $A \otimes B$, define the (complex) ***free vector space*** on any non-empty set X as $C_c(X)$, where X has the discrete topology (i.e., $C_c(X)$ consists of all functions $f : X \to \mathbb{C}$ with finite support), and pointwise operations. For each $y \in X$, the delta-function $\delta_y \in C_c(X)$ is defined by $\delta_y(x) = \delta_{xy}$, so that each element f of $C_c(X)$ is a finite sum $f = \sum_i \lambda_i \delta_{x_i}$, where $\lambda_i \in \mathbb{C}$ and $x_i \in X$.

If A and B are (complex) vector spaces, $A \otimes B$ is the quotient of the free vector space $C_c(A \times B)$ on $X = A \times B$ by the equivalence relation generated by the relations:

$$\delta_{(a_1+a_2,b)} \sim \delta_{(a_1,b)} + \delta_{(a_2,b)}; \tag{2.121}$$
$$\delta_{(a,b_1+b_2)} \sim \delta_{(a,b_1)} + \delta_{(a,b_2)}; \tag{2.122}$$
$$\lambda \delta_{(a,b)} \sim \delta_{(\lambda a,b)}; \tag{2.123}$$
$$\lambda \delta_{(a,b)} \sim \delta_{(a,\lambda b)}. \tag{2.124}$$

For $a \in A, b \in B$, the image of $\delta_{(a,b)}$ in $A \otimes B$ is called $a \otimes b$, so that by construction,

$$(a_1 + a_2) \otimes b = a_1 \otimes b + a_2 \otimes b; \tag{2.125}$$
$$a \otimes (b_1 + b_2) = a \otimes b_1 + a \otimes b_2; \tag{2.126}$$
$$\lambda(a \otimes b) = (\lambda a) \otimes b = a \otimes (\lambda b). \tag{2.127}$$

Elements of the algebraic tensor product $A \otimes B$ may therefore be written as finite sums $c = \sum_i a_i \otimes b_i$, with $a_i \in A$, $b_i \in B$, subject to the above relations.

Now consider some bilinear map $\beta : A \times B \to C$. We extend β to a map

$$\tilde{\beta} : C_c(A \times B) \to C; \tag{2.128}$$
$$\tilde{\beta}\left(\sum_i \lambda_i \delta_{(a_i,b_i)}\right) = \sum_i \lambda_i \beta(a_i, b_i). \tag{2.129}$$

Since β is bilinear, it respects the above equivalence relation, so that it duly quotients to $\beta' : A \otimes B \to C$, upon which the property $\beta = \beta' \circ p$ holds by construction. Finally, since p is surjective the latter property uniquely determines β'. □

We give two alternative constructions of $A \otimes B$:

1. Let $\mathrm{Hom}(A^* \times B^*, \mathbb{C})$ be the space of bilinear maps from $A^* \times B^*$ to $\mathbb{C}$, seen as a vector space in the obvious way, and for $a \in A$ and $b \in B$ define
$$a \otimes b : A^* \times B^* \to \mathbb{C}; \tag{2.130}$$
$$(a \otimes b)(\theta, \rho) = \theta(a)\rho(b), \tag{2.131}$$
where $\theta \in A^*$ and $\rho \in B^*$. Then $a \otimes b$ is clearly bilinear and hence
$$a \otimes b \in \mathrm{Hom}(A^* \times B^*, \mathbb{C}). \tag{2.132}$$
Since $\mathrm{Hom}(A^* \times B^*, \mathbb{C})$ is a vector space, we can define $A \otimes B$ as the linear span of all $a \otimes b$ within $\mathrm{Hom}(A^* \times B^*, \mathbb{C})$, where $a \in A$ and $b \in B$, so that elements of $A \otimes B$ are finite linear combinations $\sum_i \lambda_1 a_i \otimes b_i$, where $\lambda_i \in \mathbb{C}$, $a_i \in A$, and $b_i \in B$.
2. Similarly, we may regard $A \otimes B$ as a subspace of the vector space $L(A^*, B)$ of linear maps from A^* to B, where this time the elementary tensors are defined by
$$a \otimes b : A^* \to B; \tag{2.133}$$
$$(a \otimes b)(\theta) = \theta(a)b. \tag{2.134}$$
Once again, $A \otimes B$ is the (finite) linear span of all $a \otimes b$ (absorb the λ's in a or b).

If A and B are finite-dimensional, then $A \otimes B$ *equals* $\mathrm{Hom}(A^* \times B^*, \mathbb{C})$ or $L(A^*, B)$.

If A and B are Hilbert spaces, we call them H_1 and H_2, denote their elements by α and β, respectively, and attempt to define a sesquilinear form on $H_1 \otimes H_2$ by
$$\langle \sum_j \alpha'_j \otimes \beta'_j, \sum_i \alpha_i \otimes \beta_i \rangle = \sum_{i,j} \langle \alpha'_j, \alpha_i \rangle_{H_1} \langle \beta'_j, \beta_i \rangle_{H_2}. \tag{2.135}$$
It is a non-trivial fact that this form is well defined, because representations $\sum_i \alpha_i \otimes \beta_i$ of vectors in $H_1 \otimes H_2$ may not be unique. For example, if $H_1 = H_2 = H = \mathbb{C}^n$, and (α_i) and (α'_i) are two bases of H, then $\sum_i \alpha_i \otimes \alpha_i = \sum_i \alpha'_i \otimes \alpha'_i$ (to see this, take inner products with an arbitrary elementary tensor $\psi \otimes \varphi$, yielding the same result). Thus the above inner product is indeed well defined iff the equality
$$\sum_i \alpha_i \otimes \beta_i = \sum_i \alpha''_i \otimes \beta''_i \tag{2.136}$$
implies
$$\sum_{i,j} \langle \alpha'_j, \alpha_i \rangle_{H_1} \langle \beta'_j, \beta_i \rangle_{H_2} = \sum_{i,j} \langle \alpha'_j, \alpha''_i \rangle_{H_1} \langle \beta'_j, \beta''_i \rangle_{H_2}, \tag{2.137}$$
for all vectors $\sum_j \alpha'_j \otimes \beta'_j$. But this is true, which we see if we use the injection
$$H_1 \otimes H_2 \hookrightarrow L(H_1^* \times H_2^*, \mathbb{C}) \tag{2.138}$$

just discussed, combined with the well-known isomorphism $H^* \cong \overline{H}$ (where $\overline{H}$ is the same as H as a real vector space, but having complex scalar mulitplication defined by $\overline{\lambda}$ instead of $\lambda \in \mathbb{C}$, and hence the map $\psi \mapsto f_\psi$ from $\overline{H}$ to H^* defined by

$$f_\psi(\varphi) = \langle \psi, \varphi \rangle \tag{2.139}$$

is a *linear* isomorphism). This turns (2.138) into an injection

$$H_1 \otimes H_2 \hookrightarrow \mathrm{Hom}(\overline{H}_1 \times \overline{H}_2, \mathbb{C}), \tag{2.140}$$

i.e., the space of *bi-anti-linear* maps from $H_1 \times H_2$ to $\mathbb{C}$, where $\alpha \otimes \beta$ is the map

$$\alpha \otimes \beta : \overline{H}_1 \times \overline{H}_2 \to \mathbb{C}; \tag{2.141}$$

$$\alpha \otimes \beta(\alpha' \otimes \beta') = \langle \alpha', \alpha \rangle_{H_1} \langle \beta', \beta \rangle_{H_2}. \tag{2.142}$$

Consequently, the equality (2.137), which we need to prove, is the same as

$$\sum_i \alpha_i \otimes \beta_i \left(\sum_j \alpha'_j \otimes \beta'_j \right) = \sum_i \alpha''_i \otimes \beta''_i \left(\sum_j \alpha'_j \otimes \beta'_j \right), \tag{2.143}$$

which is the same as

$$\left(\sum_i \alpha_i \otimes \beta_i - \sum_i \alpha''_i \otimes \beta''_i \right) \left(\sum_j \alpha'_j \otimes \beta'_j \right) = 0. \tag{2.144}$$

And this is true for all $\sum_j \alpha'_j \otimes \beta'_j$, since (2.136) holds by assumption, and hence the left-hand side of (2.144) is zero as an element of $\mathrm{Hom}(\overline{H}_1 \times \overline{H}_2, \mathbb{C})$.

It remains to show that (2.135) is an inner product, i.e., that it is positive definite; this is an exercise. Finally, we complete $H_1 \otimes H_2$ in the norm defined by the inner product (2.135); with abuse of notation the ensuing Hilbert space is often just called $H_1 \otimes H_2$, but it would be more precise to denote it by $H_1 \overline{\otimes} H_2$, as we will do.

It is easy to show that if (e_i) and (f_j) are bases for H_1 and H_2, respectively, then $(e_i \otimes f_j)$ is a basis of $H_1 \overline{\otimes} H_2$ (exercise). Furthermore, we have (exercise!)

$$L^2(\mathbb{R}^n) \overline{\otimes} L^2(\mathbb{R}^m) \cong L^2(\mathbb{R}^{n+m}), \tag{2.145}$$

where the appropriate unitary isomorphism (from left to right) maps $f \otimes g$ to fg.[1]

Another interesting application is to Hilbert–Schmidt operators.

[1] More generally, if (X, Σ, μ) and (X', Σ', μ') are σ-finite measure spaces with X and X' well behaved (e.g., Polish), so that the L^2-spaces are separable, one has a natural isomorphism

$$L^2(X, \Sigma, \mu) \hat{\otimes} L^2(X', \Sigma', \mu') \cong L^2(X \times X', \Sigma \times \Sigma', \mu \times \mu'),$$

obtained as the closure of the isometric (and hence bounded) map that sends the vector $\sum_i \psi_i \otimes \psi'_i$ into the function $(x, x') \mapsto \sum_i \psi_i(x) \psi'_i(x')$ on $X \times X'$. Here $\Sigma \times \Sigma'$ is the smallest σ-algebra on $X \times X'$ that contains all sets $A \times A'$, $A \in \Sigma$, $A' \in \Sigma'$, and $\mu \times \mu'$ is the product measure defined on elementary measurable sets by $\mu \times \mu'(A \times A') = \mu(A)\mu'(A')$.

Definition 2.7. *Let H be a separable Hilbert space. An operator $a \in B(H)$ is called a* **Hilbert–Schmidt operator** *if for some (and hence any) basis (e_i) of H,*

$$\sum_i \|ae_i\|^2 < \infty, \tag{2.146}$$

We write $B_2(H)$ for the set of all Hilbert–Schmidt operators on H.

The argument that the sum in (2.146) is independent of the basis is analogous to the trace, this time even without the complication of the absolute value $|a|$. For $a \in B_2(H)$, with foresight we define the expression (where (e_i) is any basis of H):

$$\|a\|_2 = \sqrt{\mathrm{Tr}\,(a^*a)} = \left(\sum_i \|ae_i\|^2\right)^{1/2}. \tag{2.147}$$

Theorem 2.9. *The Hilbert–Schmidt operators $B_2(H)$ on a separable Hilbert space form a Hilbert space themselves with inner product*

$$\langle a,b\rangle_2 = \mathrm{Tr}\,(a^*b), \tag{2.148}$$

and a Banach space in the ensuing norm, which equals (2.147). Moreover, we have

$$H\overline{\otimes}\overline{H} \cong B_2(H); \tag{2.149}$$

$$\psi\otimes\varphi \mapsto |\psi\rangle\langle\varphi|, \tag{2.150}$$

in that (2.150) *extends by linearity and continuity to a unitary isomorphism* (2.149).

The proof is an exercise, in which one should use the fact that the linear span of the operators $|\psi\rangle\langle\varphi|$ is dense in $B_2(H)$ in the $\|\cdot\|_1$ norm (and similarly for $B_1(H)$).

We now turn to operators on tensor product Hilbert spaces. For $a \in B(H_1)$ and $b \in B(H_2)$ we form the operator $a\otimes b$ on $H_1\overline{\otimes}H_2$, initially defined on the linear span $H_1\otimes H_2$ (i.e. the *algebraic* tensor product of H_1 and H_2) of all vectors $\alpha\otimes\beta$ by

$$a\otimes b\left(\sum_i \alpha_i\otimes\beta_i\right) = \sum_i (a\alpha_i)\otimes(b\beta_i), \tag{2.151}$$

from which it is easy to show that $a\otimes b$ is bounded on $H_1\otimes H_2$, i.e.,

$$\sup\{\|a\otimes b(\psi)\|, \psi\in H_1\otimes H_2, \|\psi\| = 1\} \leq \|a\|\|b\| < \infty, \tag{2.152}$$

which (by a standard Banach space argument) also implies that $a\otimes b$ is bounded on the completion $H_1\overline{\otimes}H_2$, with the same norm as on $H_1\otimes H_2$. One way to prove (2.152) is to factor the operator $a\otimes b$ into two parts by writing

$$a\otimes b = (a\otimes 1_{H_2})\cdot(1_{H_1}\otimes b), \tag{2.153}$$

where 1_{H_i} is the unit of $B(H_i)$, $i = 1,2$, and prove the easier bounds

$$\|a \otimes 1_{H_2}\| = \|a\|; \tag{2.154}$$
$$\|1_{H_1} \otimes b\| = \|b\|, \tag{2.155}$$

which imply

$$\|a \otimes b\| = \|(a \otimes 1_{H_2})(1_{H_1} \otimes b)\| \leq \|a \otimes 1_{H_2}\| \|1_{H_1} \otimes b\| = \|a\| \|b\|. \tag{2.156}$$

For example,

$$\left\|(a \otimes 1_{H_2})\left(\sum_k \alpha_k \otimes \beta_k\right)\right\|^2 = \left\|\sum_k (a\alpha_k) \otimes \beta_k\right\|^2 \leq \sum_k \|(a\alpha_k) \otimes \beta_k\|^2$$
$$\leq \|a\|^2 \left\|\sum_k \alpha_k \otimes \beta_k)\right\|^2. \tag{2.157}$$

This strategy also gives us linear isometric injections

$$B(H_1) \hookrightarrow B(H_1 \overline{\otimes} H_2); \tag{2.158}$$
$$a \mapsto a \otimes 1_{H_2}; \tag{2.159}$$
$$B(H_2) \hookrightarrow B(H_1 \overline{\otimes} H_2); \tag{2.160}$$
$$b \mapsto 1_{H_1} \otimes b, \tag{2.161}$$

which play a major role in our next topic of interest about Hilbert space tensor products, namely the ***partial trace*** of the physicsts. Here is all there is to say:

Theorem 2.10. *Let H_1 and H_2 be Hilbert spaces. Then there is a linear map*

$$B_1(H_1 \overline{\otimes} H_2) \to B_1(H_1); \tag{2.162}$$
$$\rho \mapsto \rho_1, \tag{2.163}$$

which restricts to a convex map between the pertinent density matrices, i.e.,

$$\mathscr{D}(H_1 \overline{\otimes} H_2) \to \mathscr{D}(H_1), \tag{2.164}$$

and which is completely characterized by any one the equivalent properties

$$(a \otimes b)_1 = \mathrm{Tr}_{H_2}(b) \cdot a \ \ (a \in B_1(H_1), b \in B_1(H_2)); \tag{2.165}$$
$$\langle \alpha_1, \rho_1 \alpha_2 \rangle_{H_1} = \sum_j \langle \alpha_1 \otimes f_j, \rho(\alpha_2 \otimes f_j) \rangle_{H_1 \overline{\otimes} H_2} \ \ (\alpha_1, \alpha_2 \in H_1); \tag{2.166}$$
$$\mathrm{Tr}_{H_1}(\rho_1 a) = \mathrm{Tr}_{H_1 \overline{\otimes} H_2}(\rho(a \otimes 1_{H_2})) \ \ (a \in B_1(H), \tag{2.167}$$

where (f_i) is an arbitrary orthonormal basis of H_2. Analogously, we have a map

$$B_1(H_1 \overline{\otimes} H_2) \to B_1(H_2); \tag{2.168}$$
$$\rho \mapsto \rho_2, \tag{2.169}$$

restricting to $\mathscr{D}(H_1\overline{\otimes}H_2) \to \mathscr{D}(H_2)$, that is characterized by any one the properties

$$(a \otimes b)_2 = \mathrm{Tr}_{H_1}(a) \cdot b; \tag{2.170}$$

$$\langle \beta_1, \rho_2 \beta_2 \rangle_{H_2} = \sum_i \langle e_i \otimes \beta_1, \rho(e_i \otimes \beta_2) \rangle_{H_1 \overline{\otimes} H_2}; \tag{2.171}$$

$$\mathrm{Tr}_{H_2}(\rho_2 b) = \mathrm{Tr}_{H_1 \overline{\otimes} H_2}(\rho(1_{H_1} \otimes b)). \tag{2.172}$$

where (e_i) is an arbitrary orthonormal basis of H_1.

Apart from functional-analytic details (which are settled by the remark that the linear span of all $a \otimes b$, with $a \in B_1(H)$ and $b \in B(H_2)$ is dense in $B_1(H_1\overline{\otimes}H_2)$ in the trace-norm), this is just a computation (exercise). The significance of (2.167) and (2.172) comes out particularly well if we use (2.21). If we denote the (convex) set of *normal* states on $B(H)$ by $S_n(B(H))$, so that $\omega \in S_n(B(H))$ iff it takes the form $\omega(a) = \mathrm{Tr}(\rho a)$ for some $\rho \in \mathscr{D}(H)$, then Theorem 2.10 produces maps $S_n(B(H_1\overline{\otimes}H_2)) \to S_n(B(H_1))$, written $\omega \mapsto \omega_1$, and $S_n(B(H_1\overline{\otimes}H_2)) \to S_n(B(H_2))$, written $\omega \mapsto \omega_2$, that are defined by the counterparts of (2.167) and (2.172), namely

$$\omega_1(a) = \omega(a \otimes 1_{H_2}); \tag{2.173}$$

$$\omega_2(b) = \omega(1_{H_1} \otimes b). \tag{2.174}$$

Thus ω_1 is simply the restriction of ω to $B(H_1) \subset (B(H_1\overline{\otimes}H_2)$, cf. (2.158) - (2.159), and similarly, ω_2 is the restriction of ω to $B(H_2) \subset (B(H_1\overline{\otimes}H_2)$, cf. (2.160) - (2.161).

Let $\psi \in H_1\overline{\otimes}H_2$ be a unit vector with associated pure state $\rho = |\psi\rangle\langle\psi|$ on $B(H_1\overline{\otimes}H_2)$. If $\psi = \alpha \otimes \beta$ for unit vectors $\alpha \in H_1$ and $\beta \in H_2$, then $\rho_1 = |\alpha\rangle\langle\alpha|$ and $\rho_s = |\beta\rangle\langle\beta|$ are both pure, too. However, the simple example of the ***Bell state***

$$\psi = \frac{1}{\sqrt{2}}(e_1 \otimes e_2 - e_2 \otimes e_1) \in \mathbb{C}^2 \otimes \mathbb{C}^2, \tag{2.175}$$

where (e_1, e_2) is any orthonormal basis of $\mathbb{C}^2$, shows that ρ_1 and ρ_2 may be mixed:

$$\rho_1 = \tfrac{1}{2} \cdot 1_{H_1} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}; \tag{2.176}$$

$$\rho_2 = \tfrac{1}{2} \cdot 1_{H_2} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}. \tag{2.177}$$

More generally, if $\psi = \sum_{i,j} c_{ij} e_i \otimes f_j \in H_1\overline{\otimes}H_2$, with $\sum_{i.j} |c_{ij}|^2 = 1$, then

$$\rho_1 = \sum_j p_j |\alpha_j\rangle\langle\alpha_j|, \tag{2.178}$$

where $p_j = \|\varphi_j\|^2$ for $\varphi_j = \sum_i c_{ij} e_i$ (so that $\sum_j p_j = 1$), and $\alpha_j = \varphi_j / \|\varphi_j\|$.

Exercises for week 7 (inleveropgaven: 4, 6, 7)

1. Show that $A \otimes B$, defined as in Proposition 2.8, is unique up to isomorphisms.
2. Show that (2.135) is an inner product.
3. Show that if (e_i) and (f_j) are bases for H_1 and H_2, respectively, then $(e_i \otimes f_j)$ is a basis of $H_1 \overline{\otimes} H_2$.
4. Prove (2.145) by giving a unitary map $u : L^2(\mathbb{R}^n)\overline{\otimes}L^2(\mathbb{R}^m) \to L^2(\mathbb{R}^{n+m})$.
5. Prove Theorem 2.9 (physicists: assume H is finite-dimensional).
6. Prove the equivalence between the three conditions (2.165), (2.166), and (2.167).
7. One may iterate the tensor product construction, forming

$$H_1\overline{\otimes}H_2\overline{\otimes}H_3 = (H_1\overline{\otimes}H_2)\overline{\otimes}H_3 \cong H_1\overline{\otimes}(H_2\overline{\otimes}H_3), \tag{2.179}$$

et cetera. Hence for any $N \in \mathbb{N}$ one has the Hilbert space $H_N = (\mathbb{C}^2)^{\otimes N}$ for N qubits, i.e. the N-fold tensor product of $\mathbb{C}^2$. The underlying system consists of N bits (often called c-bits in this context), with configuration space $X_N = \underline{2}^{\underline{N}}$, i.e. the set of all maps $s : \underline{N} \to \underline{2}$, where $\underline{N} = \{0, 1, \ldots, N-1\}$ and $\underline{2} = \{0, 1\}$ as usual. Prove that

$$(\mathbb{C}^2)^{\otimes N} \cong \ell^2(\underline{2}^{\underline{N}}), \tag{2.180}$$

by constructing a unitary operator $(\mathbb{C}^2)^{\otimes N} \to \ell^2(\underline{2}^{\underline{N}})$.

Chapter 3
Symmetry in quantum physics

Roughly speaking, a symmetry of some mathematical object is an invertible transformation that leaves all relevant structure as it is. Thus a symmetry of a set is just a bijection, a symmetry of a vector space is an invertible linear map, a symmetry of a topological space is a homeomorphism, a symmetry of a smooth manifold is a diffeomorphism, a symmetry of a Banach space is a linear isometric isomorphism, and a symmetry of a Hilbert space H is a unitary operator, i.e., a bounded map $u : H \to H$ satisfying $uu^* = u^*u = 1_H$, or, equivalently, an invertible linear map $u : H \to H$ satisfying $\langle u\varphi, u\psi\rangle = \langle \varphi, \psi\rangle$ for each $\varphi, \psi \in H$. Note that if $\dim(H) < \infty$, then the last condition implies that u is invertible, since it makes u isometric, hence (being linear) also injective, hence (H being finite-dimensional) also surjective.

The discussion of symmetries in quantum physics is based on this idea, but we will see that the mathematically obvious choice (namely unitary operators) is not the physically correct one. Given a Hilbert space H, we have encountered two mathematical structures relevant to quantum mechanics:

1. The *(normal) state space* of all density matrices on H, i.e.,

$$\mathscr{D}(H) = \{\rho \in B_1(H) \mid \rho \geq 0, \mathrm{Tr}(\rho) = 1\}. \tag{3.1}$$

2. The *(normal) pure state space* of all one-dimensional projections on H:

$$\mathscr{P}_1(H) = \{\rho \in \mathscr{D}(H) \mid \rho^2 = \rho\} \tag{3.2}$$
$$= \{e \in \mathscr{P}(H) \mid \mathrm{Tr}(e) = 1\}. \tag{3.3}$$

What are the relevant structures on these sets that give rise to a good notion of symmetry? The following answers were proposed by von Neumann, Jordan, and Wigner in the 1930s. They are not at all obvious, but turn out to be mutually coherent:

1. $\mathscr{D}(H)$ is a convex set (in the real vector space $B(H)_{\mathrm{sa}}$).
2. $\mathscr{P}_1(H)$ is a set with a ***transition probability***

$$\tau : \mathscr{P}_1(H) \times \mathscr{P}_1(H) \to [0,1]; \tag{3.4}$$
$$\tau(e, f) = \mathrm{Tr}(ef). \tag{3.5}$$

To see what this means, assume $e = |\psi\rangle\langle\psi|$ and $f = |\varphi\rangle\langle\varphi|$ for unit vectors $\psi, \varphi \in H$. Then (see exercises)

$$\mathrm{Tr}\,(ef) = |\langle\psi,\varphi\rangle|^2, \tag{3.6}$$

which is the usual expression for the transition probability $\psi \to \varphi$ in physics.

Each of these structures comes with its own notion of symmetry:

Definition 3.1. *Let H be a Hilbert space.*

1. *A* **Kadison symmetry** *is an affine bijection $\mathsf{K} : \mathscr{D}(H) \to \mathscr{D}(H)$, where being* **affine** *means that K preserves convex sums, i.e., for $t \in (0,1)$ and $\rho_1, \rho_2 \in \mathscr{D}(H)$,*

$$\mathsf{K}(t\rho_1 + (1-t)\rho_2) = t\mathsf{K}\rho_1 + (1-t)\mathsf{K}\rho_2. \tag{3.7}$$

2. *A***Wigner symmetry** *is a bijection $\mathsf{W} : \mathscr{P}_1(H) \to \mathscr{P}_1(H)$ that satisfies*

$$\mathrm{Tr}\,(\mathsf{W}(e)\mathsf{W}(f)) = \mathrm{Tr}\,(ef),\ e, f \in \mathscr{P}_1(H). \tag{3.8}$$

Each unitary operator u defines the relevant symmetry in the obvious way:

$$\mathsf{K}(\rho) = u\rho u^*\ (\rho \in \mathscr{D}(H)); \tag{3.9}$$

$$\mathsf{W}(e) = ueu^*\ (e \in \mathscr{P}_1(H)); \tag{3.10}$$

However, further possibilities arise if we change $\mathbb{C}$-linearity to $\mathbb{C}$-*anti-linearity*. An ***anti-linear operator*** $u : H \to H$ is a real-linear map that satisfies $u(z\psi) = \overline{z}\psi$, $z \in \mathbb{C}$. An ***anti-unitary operator*** on H is an invertible anti-linear operator satisfying

$$\langle u\varphi, u\psi\rangle = \overline{\langle\varphi,\psi\rangle}\ (\varphi,\psi \in H). \tag{3.11}$$

Equivalently, $uu^* = u^*u = 1_H$. A simple example is the map

$$J : \mathbb{C}^n \to \mathbb{C}^n; \tag{3.12}$$

$$Jz = \overline{z}, \tag{3.13}$$

i.e., if $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$, then $(Jz)_i = \overline{z}_i$. Similarly, one may define $J : \ell^2 \to \ell^2$ or $J : L^2 \to L^2$ by $J\psi = \overline{\psi}$, where complex conjugation is defined pointwise, that is, $(J\psi)(x) = \overline{\psi(x)}$. For any Hilbert space one may pick a basis (υ_i) and define J relative to this basis by

$$J\left(\sum_i c_i\upsilon_i\right) = \sum_i \overline{c}_i\upsilon_i. \tag{3.14}$$

It follows that (3.9) - (3.10) also define symmetries if u is anti-unitary (exercise). In the next few sections we will see that there are no other possibilities.

3.1 Wigner's Theorem for $H = \mathbb{C}^2$

The key theorem about symmetries in quantum mechanics, first proved by Wigner for pure states and hence called ***Wigner's Theorem***, states that (3.9) - (3.10) is it:

Theorem 3.1. *Let H be a Hilbert space, with* $\dim(H) > 1$.

1. *Each Kadison symmetry takes the form* (3.9)*;*
2. *Each Wigner symmetry takes the form* (3.10)*;*

where in all cases the operator u is either unitary or anti-unitary, and u is uniquely determined by the symmetry in question up to a phase (that is, u u′ implement the same symmetry by conjugation iff $u' = zu$*, where* $z \in \mathbb{T}$*).*

The aim of this section is to prove part 2 of this theorem for $H = \mathbb{C}^2$. This special case gives key insights by itself and also serves as a lemma for the general proof.

We start with some background, part of which already played a role in Proposition 2.2. Any complex 2×2 matrix a can be written as

$$a = a(x_0, x_1, x_2, x_3) = \tfrac{1}{2} \sum_{\mu=0}^{3} x_\mu \sigma_\mu, \; x_\mu \in \mathbb{C}; \tag{3.15}$$

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.16}$$

i.e., the ***Pauli matrices***. Writing $\mathbf{x} = (x_1, x_2, x_3)$, some interesting special cases are:

- $x_\mu \in \mathbb{R}$ for each $\mu = 0,1,2,3$, which is the case iff $a^* = a$.
- $x_0 = 1$, $\mathbf{x} \in \mathbb{R}^3$, and $\|\mathbf{x}\| = 1$, which holds iff a is a one-dimensional projection.
- $x_0 = 1$, $\mathbf{x} \in \mathbb{R}^3$, and $\|\mathbf{x}\| \leq 1$, which holds iff a is density matrix.

Assume the second case, so that $a = e$ with $e^2 = e^* = e$ and $\mathrm{Tr}(e) = 1$. If a linear map $u : \mathbb{C}^2 \to \mathbb{C}^2$ is unitary, then simple computations show that $e' = ueu^*$ is a one-dimensional projection, too, given by $e' = \frac{1}{2}\sum_{\mu=0}^{3} x'_\mu \sigma_\mu$ with $x'_0 = 1$, $\mathbf{x}' \in \mathbb{R}^3$, and $\|\mathbf{x}'\| = 1$. Writing $\mathbf{x}' = R\mathbf{x}$ for some map $R : S^2 \to S^2$, we have

$$u(\mathbf{x} \cdot \sigma)u^* = (R\mathbf{x}) \cdot \sigma, \tag{3.17}$$

where $\mathbf{x} \cdot \sigma = \sum_{j=1}^{3} x_j \sigma_j$. This also shows that R extends to a linear isometry $R : \mathbb{R}^3 \to \mathbb{R}^3$. Using the formula

$$\mathrm{Tr}(\sigma_i \sigma_j) = 2\delta_{ij}, \tag{3.18}$$

the matrix-form of R follows as

$$R_{ij} = \tfrac{1}{2}\mathrm{Tr}(u\sigma_i u^* \sigma_j). \tag{3.19}$$

Define $U(2)$ as the group of all unitary 2×2 matrices. The topological space $U(2)$ is connected. Also, recall that $O(3)$ is the group of all real orthogonal 3×3 matrices M, a condition that may be expressed in (at least) four equivalent ways:

- $MM^T = M^M M = 1_3$;
- M invertible and $M^T = M^{-1}$;
- M is an isometry (and hence it is injective and therefore invertible);
- M preserves the inner product: $\langle M\mathbf{x}, M\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$.

This implies $\det(M) = \pm 1$ (as can be seen by diagonalizing M; being a real linear isometry, its eigenvalues can only be ± 1, and $\det(M)$ is their product). Thus $O(3)$ breaks up into two parts $O_\pm(3) = \{R \in O(3) \mid \det(R) \pm 1\}$, of which $O_+ \equiv SO(3)$ consists of rotations. Using an explicit parametrization of $SO(3)$, e.g., through Euler angles, or, using surjectivity of the exponential map (from the Lie algebra of $SO(3)$, which consist of anti-symmetric real matrices), it follows that $O_\pm(3)$ are precisely the two connected components of $O(3)$, the identity of course lying in $O_+(3)$.

Proposition 3.1. *The map $u \mapsto R$ defined by* (3.17) *is a homomorphism from $U(2)$ onto $SO(3)$, with kernel $U(1)$, seen as the diagonal matrices $z \cdot 1_2$, with $z \in \mathbb{T}$.*

Proof. As a finite-dimensional linear isometry, R is invertible (this also follows from unitarity and hence invertibility of u), hence $R \in O(3)$. It is obvious from (3.17) that $u \mapsto R$ is a continuous homomorphism (of groups). Since $U(2)$ is connected and $u \mapsto R$ is continuous, R must lie in the connected component of $O(3)$ containing the identity, whence $R \in SO(3)$. To show surjectivity of $u \mapsto R$, take some unit vector $\mathbf{u} \in \mathbb{R}^3$ and define $u = \cos(\frac{1}{2}\theta) + i\sin(\frac{1}{2}\theta)\mathbf{u}\cdot\sigma$. The corresponding rotation $R_\theta(\mathbf{u})$ is the one around $\mathbf{u}$ by an angle θ, and such rotations generate $SO(3)$. □

To incorporate $O_-(3)$, let $U_a(2)$ be the set of all anti-unitary 2×2 matrices. These do not form a group, as the product of two anti-unitaries is unitary, but the union $U(2) \cup U_a(2)$ is a disconnected Lie group with identity component $U(2)$.

Proposition 3.2. *The map $u \mapsto R$ defined by* (3.17) *is a surjective homomorphism*

$$\tilde{\pi}' : U(2) \cup U_a(2) \to O(3), \tag{3.20}$$

with Moreover, $U(2)$ maps onto $SO(3)$ and $U_a(2)$ maps onto $O_-(3)$.

Proof. The map $u \mapsto R$ in (3.17) sends the anti-unitary operator $u = J$ on $\mathbb{C}^2$ to $R = \mathrm{diag}(1, -1, 1) \in O_-(3)$. Since $U_a(2) = J \cdot U(2)$ and similarly $O_-(3) = R \cdot SO(3)$, the last claim follows. The computation of the kernel may now be restricted to $U(2)$, and then follows as in the last step op the proof of the previous proposition. □

For $H = \mathbb{C}^2$, part 3 of Theorem 3.1 (Wigner's Theorem) explicitly reads:

Theorem 3.2. *Each bijection $\mathsf{W} : \mathscr{P}_1(\mathbb{C}^2) \to \mathscr{P}_1(\mathbb{C}^2)$ that satisfies*

$$\mathrm{Tr}\,(\mathsf{W}(e)\mathsf{W}(f)) = \mathrm{Tr}\,(ef) \tag{3.21}$$

for each $e, f \in \mathscr{P}_1(\mathbb{C}^2)$ takes the form $\mathsf{W}(e) = ueu^*$, where u is either unitary or anti-unitary, and is uniquely determined by W up to a phase.

To prove this we extend the pure state part of Proposition 2.2:

Proposition 3.3. *The pure state space $\mathscr{P}_1(\mathbb{C}^2)$ corresponds bijectively to the sphere*

$$S^2 = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 = 1\},$$

in that each one-dimensional projection $e \in \mathscr{P}_1(\mathbb{C}^2)$ may be expressed uniquely as

$$e(x,y,z) = \tfrac{1}{2}\begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix}, \tag{3.22}$$

where $(x,y,z) \in \mathbb{R}^3$ and $x^2+y^2+z^2 = 1$. Under this bijection $\mathscr{P}_1(\mathbb{C}^2) \cong S^2$, Wigner symmetries W of $\mathbb{C}^2$ correspond to bijections $\mathsf{W}' : S^2 \to S^2$ that preserve the inner product of any two unit vectors in $\mathbb{R}^3$.

Proof. The first claim restates Proposition 2.2. If ψ and ψ' are unit vectors in $\mathbb{C}^2$ with corresponding one-dimensional projections $e_\psi(x,y,z)$ and $e_{\psi'}(x',y',z')$ then, as one easily verifies, the corresponding transition probability takes the form

$$\mathrm{Tr}\,(e_\psi e_{\psi'}) = \tfrac{1}{2}(1 + \langle \mathbf{x}, \mathbf{x}' \rangle) = \cos^2(\tfrac{1}{2}\theta(\mathbf{x},\mathbf{x}')), \tag{3.23}$$

where $\theta(\mathbf{x},\mathbf{y})$ is the arc (i.e., geodesic) distance between $\mathbf{x}$ and $\mathbf{y}$. Consequently, $\mathsf{W} : \mathscr{P}_1(\mathbb{C}^2) \to \mathscr{P}_1(\mathbb{C}^2)$ satisfies (3.8), iff the map $\mathsf{W}' : S^2 \to S^2$ satisfies

$$\langle \mathsf{W}'(\mathbf{x}), \mathsf{W}'(\mathbf{x}') \rangle = \langle \mathbf{x}, \mathbf{x}' \rangle \;\; (\mathbf{x}, \mathbf{x}' \in S^2). \tag{3.24}$$

Lemma 3.1. *If some bijection $\mathsf{W}' : S^2 \to S^2$ satisfies* (3.24)*, then W' extends (uniquely) to an orthogonal linear map $R : \mathbb{R}^3 \to \mathbb{R}^3$.*

Wigner's Theorem then follows by combining Propositions 3.2 and 3.3: given the linear map R just constructed, read (3.17) from right to left, where u exists by surjectivity of the map (3.20), and the precise lack of uniqueness of u as claimed in Theorem 3.1 is just a restatement of the fact that (3.20) has $U(1)$ as its kernel. □

Exercises for week 8 (Inleveropgave: 5, 6, 7)

1. Prove (3.6).
2. Show that the product of two anti-unitary operators is unitary.
3. Show that any anti-unitary operator $u : H \to H$ takes the form $u = Jv$, where v is unitary and J is an anti-unitary operator on H of the kind constructed above.
4. Show that the maps defined by (3.9) - (3.10) are symmetries according to Definition 3.1, both when u is unitary and when it is anti-unitary.
5. Prove the last claim in the proof of Proposition 3.1.
6. Prove (3.23).
7. Prove Lemma 3.1.

3.2 Proof of Wigner's Theorem

The problem is to lift a given map $\mathsf{W} : \mathscr{P}_1(H) \to \mathscr{P}_1(H)$ that satisfies (3.8) to either a unitary or an anti-unitary map $u : H \to H$ such that $e_\psi = e_{u\psi} = ue_\psi u^*$. Suppose $e_\psi = e_{\psi'}$. Since $e_{z\psi} = e_\psi$ for any $z \in \mathbb{T}$, and likewise for $e_{\psi'}$, this means that $u\psi = z\psi'$ for some $z \in \mathbb{T}$; the problem is to choose the z's coherently all over the unit sphere of H. There are many proofs in the literature, of which the following one—partly based on an earlier proof by Bargmann (1964)—has the advantage of making at least the construction of u explicit (at the cost of opaque proofs of some crucial lemma's, partly packed into exercises). We assume $\dim(H) > 2$.

Lemma 3.2. *If $V \subset H$ is a k-dimensional subspace (where $k < \infty$), then there is a unique k-dimensional subspace $V' \subset H$ with the following property: for any unit vector $\psi \in H$ we have $\psi \in V$ iff $\psi' \in V'$ for any (unit) vector $\psi' \in \mathsf{W}(e_\psi)H$.*

Proof. Pick a basis $(\upsilon_1, \ldots, \upsilon_k)$ of V and find unit vectors $\upsilon_i' \in H$ such that $\upsilon_i' \in \mathsf{W}(e_{\upsilon_i})H$, $i = 1, \ldots, k$. Then

$$|\langle \upsilon_i', \upsilon_j' \rangle|^2 = \mathrm{Tr}\,(e_{\upsilon_i'} e_{\upsilon_j'}) = \mathrm{Tr}\,(\mathsf{W}(e_{\upsilon_i})\mathsf{W}(e_{\upsilon_j})) = \mathrm{Tr}\,(e_{\upsilon_i} e_{\upsilon_j}) = |\langle \upsilon_i, \upsilon_j \rangle|^2 = \delta_{ij},$$

so that the vectors $(\upsilon_1', \ldots, \upsilon_k')$ form an orthonormal set and hence form a basis of their linear span V'. Now, as mentioned below (**??**), we have $\psi \in H$ iff $\sum_{i=1}^k |\langle \upsilon_i, \psi \rangle|^2 = 1$ and similarly $\psi' \in H$ iff $\sum_{i=1}^k |\langle \upsilon_i', \psi' \rangle|^2 = 1$. Since W preserves transition probabilities, a similar computation gives $\sum_{i=1}^k |\langle \upsilon_i, \psi \rangle|^2 = \sum_{i=1}^k |\langle \upsilon_i', \psi' \rangle|^2$, which proves the claim. □

We have already proved Wigner's Theorem for $H = \mathbb{C}^2$. This implies (exercise):

Lemma 3.3. *If V and V' are related as in Lemma 3.2, and $\dim(V) = \dim(V') = 2$, then there is a unitary or anti-unitary operator $u_V : V \to V'$ such that $\mathsf{W}(e) = u_V e u_V^*$ for any $e \in \mathscr{P}_1(V)$, where $\mathscr{P}_1(V) \subset \mathscr{P}_1(H)$ consists of all $e \in \mathscr{P}_1(H)$ with $eH \subset V$.*

So far this has been relatively "canonical". The tricky steps of the proof are:

Lemma 3.4. *Either u_V is unitary for all $V \subset H$ (with $\dim(V) = 2$) or u_V is anti-unitary for all $V \subset H$ (with $\dim(V) = 2$).*

Lemma 3.5. *Wigner's Theorem holds for $H = \mathbb{C}^3$. Therefore (cf. Lemma 3.2), if $\dim(V) = \dim(V') = 3$, then there is a unitary or anti-unitary operator $u_V : V \to V'$ such that $\mathsf{W}(e) = u_V e u_V^*$ for any $e \in \mathscr{P}_1(V)$,*

For completeness' sake we give the proofs at the end of this section in small print.

We are now in a position to finish the proof of Wigner's Theorem. Fix unit vectors $\psi \in H$ and $\psi' \in \mathsf{W}(e_\omega)H$; clearly, ψ' is unique up to multiplication by $z \in \mathbb{T}$, and its choice turns out to completely determine u (i.e., the ambiguity in ψ' is the only one in the entire construction). For a modest start, we put

$$u\psi = \psi'. \tag{3.25}$$

We assume that the outcome of Lemma 3.4 is that each u_V is unitary; the anti-unitary case requires obvious modifications of the argument below. The first step is, of course, to define $u(\lambda\psi) = \lambda u\psi$, $\lambda \in \mathbb{C}$ (so this would have been $\overline{\lambda} u\psi$ in the anti-unitary case). Let $\varphi \in H$ be linearly independent of ψ and consider the two-dimensional space V spanned by ψ and φ. Define $u(\varphi) = u_V\varphi$. With (3.25), this defines u on all of H. To prove that u is linear, take φ_1 and φ_2 linearly independent of each other and of ψ, so that the linear span V_3 of ψ, φ_1, and φ_2 is three-dimensional. Let V_i be the two-dimensional linear span of ψ and φ_i, $i = 1, 2$. Then $u\varphi_i = u_{V_i}\varphi_i$, where the phase of u_{V_i} is fixed by (3.25). Let $w : V_3 \to V_3'$ be the unitary that implements W according to Lemma 3.5, with phase determined by (3.25). Since u_{V_1} and u_{V_2} and w are unique up to a phase and this phase has been fixed for each in the same way, we must have $u_{V_1} = w_{|V_1}$ and $u_{V_2} = w_{|V_2}$. Now w is unitary and hence linear, so

$$\begin{aligned} u(\varphi_1 + \varphi_2) &= w(\varphi_1 + \varphi_2) = w(\varphi_1) + w(\varphi_2) \\ &= u_{V_1}(\varphi_1) + u_{V_2}(\varphi_2) = u(\varphi_1) + u(\varphi_2), \end{aligned}$$

since this is how u was defined. Since each u_V is unitary, so is u and similarly it is easy to verify that u implements W, because each u_V does so. □

Proof of Lemma 3.4. We first design a "unitarity test" for W. Define a function

$$T : \mathscr{P}_1(H) \times \mathscr{P}_1(H) \times \mathscr{P}_1(H) \to \mathbb{C}; \tag{3.26}$$

$$T(e, f, g) = \mathrm{Tr}\,(efg), \tag{3.27}$$

$$T(e_{\psi_1}, e_{\psi_2}, e_{\psi_3}) = \langle \psi_1, \psi_2 \rangle \langle \psi_2, \psi_3 \rangle \langle \psi_3, \psi_1 \rangle. \tag{3.28}$$

Let $V \subset H$ be two-dimensional and pick an orthonormal basis (υ_1, υ_2). Define

$$\chi_1 = \upsilon_1, \;\; \chi_2 = (\upsilon_1 - \upsilon_2)/\sqrt{2}, \;\; \chi_3 = (\upsilon_1 - i\upsilon_2)/\sqrt{2}. \tag{3.29}$$

A simple computation then shows that

$$T(e_{\chi_1}, e_{\chi_2}, e_{\chi_3}) = \tfrac{1}{4}(1 + i). \tag{3.30}$$

It follows from (3.28) that for u unitary and v anti-unitary, we have

$$T(e_{u\psi_1}, e_{u\psi_2}, e_{u\psi_3}) = T(e_{\psi_1}, e_{\psi_2}, e_{\psi_3}); \tag{3.31}$$

$$T(e_{v\psi_1}, e_{v\psi_2}, e_{v\psi_3}) = \overline{T(e_{\psi_1}, e_{\psi_2}, e_{\psi_3})}. \tag{3.32}$$

Eq. (3.30) implies that if $\mathsf{W} : V \to V'$ is (anti-) unitarily implemented, we have

$$T(\mathsf{W}(e_{\chi_1}), \mathsf{W}(e_{\chi_2}), \mathsf{W}(e_{\chi_3})) = T(e_{u\chi_1}, e_{u\chi_2}, e_{u\chi_3}) = \tfrac{1}{4}(1 \pm i), \tag{3.33}$$

with a plus sign if u is unitary and a minus sign if u is anti-unitary. Now take a second pair $(\tilde{V}, \tilde{V}')$ as above, and pick a basis $(\tilde{\upsilon}_1, \tilde{\upsilon}_2)$ of $\tilde{V}$, with associated vectors $(\tilde{\chi}_1, \tilde{\chi}_2, \tilde{\chi}_3)$, as in (3.29). Suppose $u : V \to V'$ implementing W is unitary, whereas $\tilde{u} : \tilde{V} \to \tilde{V}'$ implementing W is anti-unitary. It then follows from (3.33) that

$$T(\mathsf{W}(e_{\chi_1}), \mathsf{W}(e_{\chi_2}), \mathsf{W}(e_{\chi_3})) = T(e_{u\chi_1}, e_{u\chi_2}, e_{u\chi_3}) = \tfrac{1}{4}(1 + i); \tag{3.34}$$

$$T(\mathsf{W}(e_{\tilde{\chi}_1}), \mathsf{W}(e_{\tilde{\chi}_2}), \mathsf{W}(e_{\tilde{\chi}_3})) = T(e_{\tilde{u}\tilde{\chi}_1}, e_{\tilde{u}\tilde{\chi}_2}, e_{\tilde{u}\tilde{\chi}_3}) = \tfrac{1}{4}(1 - i). \tag{3.35}$$

The following expression defies a metric d on $\mathscr{P}_1(H)$:

$$d(e_\psi, e_\varphi) = \|\omega_\psi - \omega_\varphi\| = \|e_\psi - e_\varphi\|_1 = 2\sqrt{1 - |\langle \varphi, \psi \rangle|^2}, \tag{3.36}$$

with respect to which both W and T are continuous (the latter with respect to the product metric on $\mathscr{P}_1(H)^3$, of course). Let $t \mapsto (\upsilon_1(t), \upsilon_2(t))$ be a continuous path of orthonormal vectors (i.e., in $H \times H$), with associated vectors $(\chi_1(t), \chi_2(t), \chi_3(t))$, as in (3.29). Then the function $f(t) = T(\mathsf{W}(\chi_1(t)), \mathsf{W}(\chi_2(t)), \mathsf{W}(\chi_3(t)))$ is continuous, and by (3.33) it can only take the values $\frac{1}{4}(1 \pm i)$. Hence $f(t)$ must be constant. However, taking a path such that $(\upsilon_1(0), \upsilon_2(0)) = (\upsilon_1, \upsilon_2)$ and $(\upsilon_1(1), \upsilon_2(1)) = (\tilde{\upsilon}_1, \tilde{\upsilon}_2)$, gives $f(0) = \frac{1}{4}(1 + i)$ and $f(1) = \frac{1}{4}(1 - i)$, which is a contradiction. □

Proof of Lemma 3.5. Let $(\upsilon_1, \upsilon_2, \upsilon_3)$ be some basis of of H (like the usual basis of $H = \mathbb{C}^3$). We first show that if W is the identity if restricted to both $\mathrm{span}(\upsilon_1, \upsilon_2)$ and $\mathrm{span}(\upsilon_1, \upsilon_3)$, then W is the identity on H altogether. To this end, take $\psi = \sum_i c_i \upsilon_i$, initially with $c_1 \in \mathbb{R}\backslash\{0\}$. Take a unit vector $\psi' \in \mathsf{W}(e_\psi)$, with $\psi = \sum_i c_i' \upsilon_i$. By the first assumption on W we have $|\langle \upsilon, \psi' \rangle| = |\langle \upsilon, \psi \rangle|$ for any unit vector $\upsilon \in \mathrm{span}(\upsilon_1, \upsilon_2)$. Taking

$$\upsilon = \upsilon_1, \qquad \upsilon = \upsilon_2, \qquad \upsilon = (\upsilon_1 + \upsilon_2)/\sqrt{2}, \qquad \upsilon = (\upsilon_1 + i\upsilon_2)/\sqrt{2}, \tag{3.37}$$

gives the equations

$$|c_1'| = |c_1|, \qquad |c_2'| = |c_2|, \qquad |c_1' + c_2'| = |c_1 + c_2|, \qquad |c_1' - ic_2'| = |c_1 - ic_2|, \tag{3.38}$$

respectively. By a choice of phase we may and will assume $c_1' = c_1$, in which case the only solution is $c_2 = c_2'$ (geometrically, the solution c_2' lies in the intersection of three different circles in the complex plane, which is either empty or consists of a single point). Similarly, the second assumption on W gives $c_3 = c_3'$, whence $\psi' = \psi$. The case $c_1 = 0$ may be settled by a straightforward limit argument, since inner products (and hence their absolute values) are continuous on $H \times H$.

Given a Wigner symmetry $\mathsf{W} : \mathscr{P}_1(H) \to \mathscr{P}_1(H)$, we now construct u as follows.

1. Fix a basis $(\upsilon_1, \upsilon_2, \upsilon_3)$ with "image" $(\upsilon_1', \upsilon_2', \upsilon_3')$ under W, i.e, $\mathsf{W}(e_{\upsilon_i}) = e_{\upsilon_i'}$.
2. The unitarity test in the proof of Lemma 3.4 settles if the operators should be chosen to be unitary or anti-unitary; for simplicity we assume the unitary case.
3. Define a unitary $u_1 : H \to H$ by $u_1 \upsilon_i' = \upsilon_i$ for $i = 1,2,3$, and subsequently define $\mathsf{W}_1 = \alpha_{u_1} \circ \mathsf{W}$, which (being the composition of two Wigner symmetries) is a Wigner symmetry. Clearly, $\mathsf{W}_1(e_{\upsilon_i}) = e_{\upsilon_i}$ $(i = 1,2,3)$, so that W_1 maps $\mathscr{P}_1(H_{(12)})$ to itself, where $H_{(12)} \equiv \mathrm{span}(\upsilon_1, \upsilon_2)$. Hence Lemma 3.3 gives a unitary map $\tilde{u}_1 : H_{(12)} \to H_{(12)}$ such that the restriction of W_1 to $H_{(12)}$ is $\alpha_{\tilde{u}_1}$.
4. Define a unitary $u_2 : H \to H$ by $u_2 = \tilde{u}_1^{-1}$ on $H_{(12)}$ and $u_2 \upsilon_3 = \upsilon_3$, followed by the Wigner symmetry $\mathsf{W}_2 = \alpha_{u_2} \circ \mathsf{W}_1$. By construction, $\mathsf{W}_2(e_{\upsilon_i}) = e_{\upsilon_i}$ for $i = 1,2,3$) (W_2 is even the identity on $\mathscr{P}_1(H_{(12)})$), so that W_2 maps $\mathscr{P}_1(H_{(13)})$ to itself, where $H_{(13)} \equiv \mathrm{span}(\upsilon_1, \upsilon_3)$. Hence the restriction of W_2 to $H_{(13)}$ is implemented by a unitary $\tilde{u}_2 : H_{(13)} \to H_{(13)}$, whose phase may be fixed by requiring $\tilde{u}_2 \upsilon_1 = \upsilon_1$.
5. Similarly to u_2, we define $u_3 : H \to H$ by $u_3 = \tilde{u}_2^{-1}$ on $H_{(13)}$ and $u_3 \upsilon_2 = \upsilon_2$, so that u_3 is the identity on $H_{(12)}$. Of course, we now define a Wigner symmetry

$$\mathsf{W}_3 = \alpha_{u_3} \circ \mathsf{W}_2 = \alpha_{u_3} \circ \alpha_{u_2} \circ \alpha_{u_1} \circ \mathsf{W}, \tag{3.39}$$

which by construction is the identity on both $\mathscr{P}_1(H_{(12)})$ and $\mathscr{P}_1(H_{(13)})$, and so by the first part of the proof it must be the identity on all of $\mathscr{P}_1(H)$. Hence

$$\mathsf{W} = \alpha_{u_1^{-1}} \circ \alpha_{u_2^{-1}} \circ \alpha_{u_3^{-1}} = \alpha_u \qquad (u = u_1^{-1} u_2^{-1} u_3^{-1}). \qquad \square$$

3.3 Equivalence between Kadison and Wigner symmetries

The equivalence between Wigner's Theorem (i.e., part 2 of Theorem 3.1) and Kadison's (i.e., part 1) is based on the following identification of their assumptions.

Proposition 3.4. *There is a bijective correspondence between:*

- *affine bijections* $\mathsf{K} : \mathscr{D}(H) \to \mathscr{D}(H)$*;*
- *bijections* $\mathsf{W} : \mathscr{P}_1(H) \to \mathscr{P}_1(H)$ *that satisfy* (3.8)*, viz.*

$$\mathsf{W} = \mathsf{K}_{|\mathscr{P}_1(H)}; \tag{3.40}$$

$$\mathsf{K}\left(\sum_i \lambda_i e_{\upsilon_i}\right) = \sum_i \lambda_i \mathsf{W}(\upsilon_{\upsilon_i}), \tag{3.41}$$

where $\rho = \sum_i \lambda_i e_{\upsilon_i}$ *is some (not necessarily unique) expansion of* $\rho \in \mathscr{D}(H)$ *in terms of a basis of eigenvector* υ_i *with eigenvalues* λ_i*, where* $\lambda_i \geq 0$ *and* $\sum_i \lambda_i = 1$*. In particular,* (3.40) *and* (3.41) *are well defined.*

Proof. It is important to distinguish between $B(H)_{\text{sa}}$ as a Banach space in the usual operator norm $\|\cdot\|$, and $B_1(H)_{\text{sa}}$, the Banach space of trace-class operators in its intrinsic norm $\|\cdot\|_1$. Of course, if $\dim(H) < \infty$, then $B(H)_{\text{sa}} = B_1(H)_{\text{sa}}$ as vector spaces, but even in that case the two norms do not coincide. We start with (3.40).

1. Since $\mathscr{P}_1(H) = \partial_e \mathscr{D}(H)$, any affine bijection of the convex set $\mathscr{P}_1(H)$ must preserve its boundary (why?), so that K bijectively (why?) maps $P_1(H)$ into itself. The goal of the next two steps is to prove that (3.40) satisfies (3.8).
2. An affine bijection $\mathsf{K} : \mathscr{D}(H) \to \mathscr{D}(H)$ extends to an isometric isomorphism $\mathsf{K}_1 : B_1(H)_{\text{sa}} \to B_1(H)_{\text{sa}}$ with respect to the trace-norm $\|\cdot\|_1$, as follows:

   a. Put $\mathsf{K}_1(0) = 0$ and for $b \geq 0$, $b \in B_1(H)$, i.e. $b \in B_1(H)_+$, and $b \neq 0$, define

   $$\mathsf{K}_1(b) = \|b\|_1 \mathsf{K}(b/\|b\|_1). \tag{3.42}$$

   By construction, this is isometric and preserves positivity. Note that for $b \in B_1(H)_+$ we have $\text{Tr}\,(b) = \|b\|_1$, hence $b/\|b\|_1 \in \mathscr{D}(H)$, where K is defined. Linearity of K_1 with positive coefficients follows from the affine property of K (exercise). Note that if $a, b \in B_1(H)_+$, then $a + b \in B_1(H)_+$.

   b. For $b \in B_1(H)_{\text{sa}}$, decompose $b = b_+ - b_-$, where $b_\pm \geq 0$; the decomposition is unique subject to $b_+ b_- = 0$. For general $b = b^* \in B(H)$ this follows from the spectral theorem in the form $C^*(b) \cong C(\sigma(b))$, see exercises, but in the case at hand b is compact (since $B_1(H) \subset B_0(H)$ and so it has a spectral expansion

   $$b = \sum_{\lambda \in \sigma(b)} \lambda \cdot e_\lambda, \tag{3.43}$$

   where $\sigma(b) \subset \mathbb{R}$ and hence each $\lambda \in \mathbb{R}$; hence we may take

$$b_\pm = \pm \sum_{\lambda \in \sigma(a) \cap \mathbb{R}^\pm} \lambda \cdot e_\lambda, \tag{3.44}$$

We then define

$$\mathsf{K}_1(b) = \mathsf{K}_1(b_+) - \mathsf{K}_1(b_-). \tag{3.45}$$

It is an exercise to show that this makes K_1 linear on all of $B_1(H)_{\mathrm{sa}}$.

The key point in verifying isometry of K_1 is the property

$$|b| = b_+ + b_-, \tag{3.46}$$

which follows from either (3.44) or the continuous functional calculus. Using (3.46), we have

$$\begin{aligned} \|\mathsf{K}_1(b)\|_1 &= \mathrm{Tr}\,(|\mathsf{K}_1 b|) = \mathrm{Tr}\,(|\mathsf{K}_1(b_+) - \mathsf{K}_1(b_-)|) = \mathrm{Tr}\,(\mathsf{K}_1(b_+) + \mathsf{K}_1(b_-)) \\ &= \mathrm{Tr}\,(b_+ + b_-) = \mathrm{Tr}\,(|b_+ - b_-|) = \mathrm{Tr}\,(|b|) = \|b\|_1. \end{aligned}$$

3. For any $e, f \in P_1(H)$ we have the formula

$$\|e - f\|_1 = 2\sqrt{1 - \mathrm{Tr}\,(ef)}, \tag{3.47}$$

which can be proved by a calculation with 2×2 matrices: just take $e = e_\psi$ and $f = e_\varphi$ and work inside the two-dimensional subspace spanned by ψ and φ, except when $\varphi = z\psi$, $z \in \mathbb{T}$, in which case (3.47) reads $0 = 0$. Since K_1 is linear as well as isometric with respect to the trace-norm, we have

$$\|\mathsf{K}_1(e) - \mathsf{K}_1(f)\|_1 = \|\mathsf{K}_1(e - f)\|_1 = \|e - f\|_1,$$

and hence, by (3.47), $\mathrm{Tr}\,(\mathsf{K}_1(e)\mathsf{K}_1(f)) = \mathrm{Tr}\,(ef)$. Eq. (3.40) gives (3.8).

We move on to (3.41). The main concern is that this expression be well defined, since in case some eigenvalue $\lambda > 0$ of ρ is degenerate (necessarily with finite multiplicity, even in infinite dimension, since ρ is compact), the basis of the eigenspace H_λ that takes part in the sum $\sum_i \lambda_i e_{\upsilon_i}$ is far from unique. This is settled as follows:

Lemma 3.6. *Let $\mathsf{W} : \mathscr{P}_1(H) \to \mathscr{P}_1(H)$ be a bijection that satisfies* (3.8)*, let $L \subset H$ be a (finite-dimensional) subspace, and let (υ_i) and (υ_i') be two bases of L, with associated one-dimensional projections $e_i \equiv e_{\upsilon_i}$ and $e_i' \equiv e_{\upsilon_i'}$. Then*

$$\sum_i \mathsf{W}(e_i) = \sum_i \mathsf{W}(e_i'). \tag{3.48}$$

Proof. For projections e and f on H we write $e \leqslant f$ iff $eH \subseteq fH$, which is true iff $ef = e$ (which implies that e and f commute). Here we go:

1. From elementary Hilbert space theory (Bessel's inequality/Parseval's formula), for any unit vector $\psi \in H$ we have $\sum_i |\langle \upsilon_i, \psi \rangle|^2 = 1$ iff $\psi \in L$ and so for any $e \in P_1(H)$ we have $e \leqslant e_L$ iff $\sum_i \mathrm{Tr}\,(e_i e) = 1$.

2. If $e, f \in P_1(H)$ with $ef = 0$ (i.e. $eH \perp fH$), then $e + f \in P(H)$ and $e \leqslant e_L$, and $f \leqslant e_L$ imply $e + f \leqslant e_L$. More generally, if $e_i e_i = \delta_{ij} e_i$ (in other words, the e_i are mutually orthogonal), for some $e_i \in P_1(H)$, then $\sum_i e_i$ is a projection and $e_i \leqslant e_L$ for each i implies $\sum_i e_i \leqslant e_L$. We use this for the e_i defined in the lemma.
3. Although $\mathrm{Tr}\,(a) = 0$ does not imply $a = 0$, in the case at hand we do have

$$e_i \perp e_j \Leftrightarrow e_i e_j = 0 \Leftrightarrow \mathrm{Tr}\,(e_i e_j) = 0,$$

since the nontrivial implication $\Leftarrow$ in the second $\Leftrightarrow$ arises from $e_i = e_{\upsilon_i}$ and hence $\mathrm{Tr}\,(e_i e_j) = |\langle \upsilon_i, \upsilon_j \rangle|^2$, which vanishes iff $\langle \upsilon_i, \upsilon_j \rangle = 0$ and hence $e_i e_j = 0$.
4. Therefore, by (3.8) the $\mathsf{W}(e_i)$ are orthogonal: if $e_i \perp e_j$, then $W(e_i) \perp W(e_j)$, and hence $\hat{L} = \sum_i \mathsf{W}(e_i)$ is a projection, like $L = \sum_i e_i$.
5. Since $\upsilon'_j \in L$ we have $e'_j \leqslant e_L$ and hence $\sum_i \mathrm{Tr}\,(e_i e'_j) = 1$, see no. 1 above, and hence, by (3.8), $\sum_i \mathrm{Tr}\,(W(e_i) W(e'_j)) = 1$. This is equivalent to

$$W(e'_j) \leqslant \sum_i \mathsf{W}(e_i) = \hat{L}, \tag{3.49}$$

and hence by no. 2 we also have $\sum_j W(e'_j) \leqslant \sum_i \mathsf{W}(e_i)$. Interchanging the roles of the two bases gives the converse inequality, yielding (3.48). □

Finally, to prove bijectivity of the correspondence $\mathsf{K} \leftrightarrow \mathsf{W}$, we need the property

$$\mathsf{K}\left(\sum_i \lambda_i e_{\upsilon_i}\right) = \sum_i \lambda_i \mathsf{K}(e_{\upsilon_i}), \tag{3.50}$$

since this implies that K is determined by its action on $\mathscr{P}_1(H) \subset \mathscr{D}(H)$. In finite dimension this follows from convexity of K, and we are done. In infinite dimension, we in addition need continuity of K, as well as convergence of the sum $\sum_i \lambda_i e_{\upsilon_i}$ not only in the operator norm (as follows from the spectral theorem for self-adjoint compact operators), but also in the trace norm: for finite n, m,

$$\|\sum_{i=n}^{m} \lambda_i e_{\upsilon_i}\|_1 \leq \sum_{i=n}^{m} |\lambda_i| \|e_{\upsilon_i}\|_1 = \sum_{i=n}^{m} \lambda_i,$$

since $\|e_{\upsilon_i}\|_1 = 1$. Because $\sum_i \lambda_i = 1$, the above expression vanishes as $n, m \to \infty$, whence $\rho_n = \sum_{i=1}^{n} \lambda_i e_{\upsilon_i}$ is a Cauchy sequence in $B_1(H)$, which by completeness of the latter converges (to an element of $\mathscr{D}(H)$, as one easily verifies). The proof of continuity is completed by noting that K is continuous with respect to the trace norm, for it is isometric and hence bounded (see step 2 above). □

Exercises for week 9 (Inleveropgaven: 3, 5, 6)

1. Give the details for step 1 in the proof of Proposition 3.4.
2. Show that the map K_1 constructed in the text is linear (assuming K is affine). First prove this on $B_1(H)_+$ (i.e. with positive coefficients), then on $B_1(H)_{\mathrm{sa}}$.
3. Prove (3.47).
4. Prove Lemma 3.6 directly, without using the partial order on projections.
5. Prove that the map K constructed from W is indeed an affine bijection of $\mathscr{D}(H)$.
6. Prove Proposition 3.5 in the next section.

3.4 Symmetry groups and projective representations

Since Wigner symmetries W are implemented by either unitary or anti-unitary operators that are determined by W only up to a phase $z \in \mathbb{T}$, we may conclude that the quantum-mechanical symmetry group $\mathscr{G}^H$ of a Hilbert space H is given by

$$\mathscr{G}^H = (U(H) \cup U_a(H))/\mathbb{T}, \tag{3.51}$$

where $U(H)$ is the group of unitary operators on H, and $U_a(H)$ is the set of anti-unitary operators on H; the latter is not a group (since the product of two anti-unitaries is unitary) but their union is. Furthermore, $\mathbb{T}$ is identified with the normal subgroup $\mathbb{T} \equiv \mathbb{T} \cdot 1_H = \{z \cdot 1_H \mid z \in \mathbb{T}\}$ of $U(H) \cup U_a(H)$ (and also of $U(H)$) consisting of multiples of the unit operators by a phase; thus the quotient $\mathscr{G}^H$ is a group.

The fact that $\mathscr{G}^H$ rather than $U(H)$ is the symmetry group of quantum mechanics has profound consequences. Namely, let G be a group. In mathematics, the natural kind of action of G on a Hilbert space H is a ***unitary representation***, i.e., a homomorphism $u : G \to U(H)$, $x \mapsto u(x)$ (so that $u(x)^{-1} = u(x^{-1}) = u(x)^*$ and $u(x)u(y) = u(xy)$, which imply $u(e) = 1_H$). As to the possible continuity properties of unitary representations in case that G is a *topological* group,[1] it turns out to be optimal to equip $U(H)$ with the *strong* operator topology (in which $a_\lambda \to a$ iff $a_\lambda \psi \to a\psi$ for each $\psi \in H$).[2] One reason for this choice of topology (as opposed to the norm topology) is pragmatic: in most examples (where H is infinite-dimensional), norm-continuity of $x \mapsto u(x)$ is simply not the case. Fortunately, the strong topology on $U(H)$ is just what is needed:

Proposition 3.5. *If $u : x \mapsto u(x)$ is a unitary representation of some topological group G on a Hilbert space H, then the following conditions are equivalent:*

1. *The map $G \times H \to H$, $(x, \psi) \mapsto u(x)\psi$, is continuous;*
2. *The map $G \to U(H)$, $x \mapsto u(x)$, is continuous in the strong topology on $U(H)$.*

However, we have just seen that in quantum mechanics one should look at homomorphisms $h : G \mapsto \mathscr{G}^H$. To simplify the discussion, we now drop $U_a(H)$ from consideration and just deal with the connected component $\mathscr{G}^H_0 = U(H)/\mathbb{T}$ of the identity. This restriction may be justified by noting that in what follows we will only deal with symmetries given by *connected Lie groups*.[3] Thus in what follows we are going to study continuous homomorphisms

---

[1] A ***topological group*** is, a group G that is also a topological space, such that group multiplication $G \times G \mapsto G$ and inverse $G \to G$ are continuous; typical groups that appear in mathematical physics, such as $\mathbb{R}^n$ or $SO(3)$ are naturally endowed with a compatible topology (in the latter case, it comes from the embedding in $M_3(\mathbb{C}) \cong \mathbb{R}^9$).

[2] This topology may also be defined on $U_a(H)$, of course; if J is a conjugation on H, as used before, the map $u \mapsto Ju$ from $U(H)$ to $U_a(H)$ is a homeomorphism.

[3] These are generated by some neighborhood of the identity, in which each element is a square, $x = y^2$. In that case, $h(x) = h(y)^2$ is always a square and hence it cannot lie in the component $U_a(H)/\mathbb{T}$. All elements are products of such squares and hence they must also lie in $U(H)/\mathbb{T}$.

$$h: G \to U(H)/\mathbb{T}. \tag{3.52}$$

Since it is inconvenient to deal with such a quotient, and one would like to use the mathematical literature on unitary representations, we try to lift h to some map

$$u: G \to U(H), \tag{3.53}$$

where, in terms of the canonical projection (which is a group homomorphism)

$$\pi: U(H) \to U(H)/\mathbb{T}, \tag{3.54}$$

we have

$$\pi \circ u = h. \tag{3.55}$$

As we will see, such maps u always exist, but u often loses the homomorphism property, *though in a controlled way*. To see how this comes about, in the worst case use the axiom of choice to choose a cross-section s of π in (3.54), i.e., a map

$$s: U(H)/\mathbb{T} \to U(H) \tag{3.56}$$

such that

$$\pi \circ s = \mathrm{id}. \tag{3.57}$$

This can be done in a measurable way, but generally not in a continuous way, and also s usually fails to be a group homomorphism.[4] Anyway, from h and s we obtain

$$u: G \to U(H); \tag{3.58}$$

$$u = s \circ h, \tag{3.59}$$

written $x \mapsto u(x)$. In particular, the cross-section property (3.57) yields (3.55). Since s typically fails to be a group homomorphism, also u is not a homomorphism. Now different choices of s must differ by a phase, and h is a homomorphism of groups, hence the operator $u(x)u(y)u(xy)^*$ must be a multiple of the identity, i.e.,

$$u(x)u(y)u(xy)^* = c(x,y) \cdot 1_H, \tag{3.60}$$

for some function $c: G \times G \to \mathbb{T}$ (exercise). Equivalently, we have

$$u(x)u(y) = c(x,y)u(xy) \ (x,y \in G). \tag{3.61}$$

Associativity of multiplication in G and the homomorphism property of h yield

$$c(x,y)c(xy,z) = c(x,yz)c(y,z), \tag{3.62}$$

and if we impose the natural requirement

[4] This is true much more generally: if N is a normal subgroup of G, then G/N is a group and the canonical projection $\pi: G \to G/N$ is a homomorphism, but usually there is no cross-section $s: G/N \to G$ of π that is a homomorphism.

$$s(e_{U(H)/\mathbb{T}}) = e_{U(H)} = 1_H, \tag{3.63}$$

or, equivalently,

$$u(e) = 1_H, \tag{3.64}$$

we also have

$$c(e,x) = c(x,e) = 1. \tag{3.65}$$

A function $c : G \times G \to \mathbb{T}$ satisfying (3.62) and (3.65) is called a ***multiplier*** or ***2-cocycle*** on G, and the set $Z^2(G,\mathbb{T})$ of such multipliers is an abelian group under (pointwise) operations in $\mathbb{T}$. In the topological case one requires c to be Borel measurable (and in the Lie group case it should in addition be smooth *near the identity*). A map $x \mapsto u(x)$ from G to $U(H)$ for which (3.61) holds is called a ***projective*** unitary representation of G (on H), with multiplier c. Suppose we change s in (3.56) to

$$s :' U(H)/\mathbb{T} \to U(H), \tag{3.66}$$

such that

$$\pi \circ s' = \mathrm{id} \tag{3.67}$$

as well as

$$s'(e) = 1_H. \tag{3.68}$$

It follows that

$$s'(\pi(v)) = \beta(\pi(v))s(\pi(v)), \tag{3.69}$$

where $v \in U(H)$ and hence $\pi(v) \in U(H)/\mathbb{T}$ (why?). Changing s to s' changes $u = s \circ h$ to $u' = s' \circ h$, where, putting $\pi(v) = h(x)$ in (3.69), and defining

$$b = \beta \circ h, \tag{3.70}$$

we obtain

$$u(x)' = b(x)u(x). \tag{3.71}$$

In fact, one could also change u to u' via (3.71) straight away, without returning to the cross-sections s and s'. Either way,

$$b : G \to \mathbb{T} \tag{3.72}$$

is a measurable function satisfying

$$b(e) = 1. \tag{3.73}$$

This condition follows from $s'(e) = 1_H$, or otherwise is imposed so that

$$u'(e) = 1_H, \tag{3.74}$$

just like u. Then the associated multiplier c', defined by

$$u(x)'u(y)' = c'(x,y)u'(xy), \tag{3.75}$$

as in (3.60), changes to

$$c'(x,y) = \frac{b(x)b(y)}{b(xy)} c(x,y). \tag{3.76}$$

This identifies a special class of multipliers, called ***2-coboundaries***, namely functions c of the form

$$c(x,y) \equiv \partial_1 b(x,y) = \frac{b(xy)}{b(x)b(y)}. \tag{3.77}$$

A multiplier of the form (3.77) may be removed by changing to (3.71), so that

$$u(x)'u(y)' = u'(xy), \tag{3.78}$$

and hence, since we already had (3.74), the map $u' : G \to U(H)$ is an ordinary unitary representation of G. The set of 2-coboundaries forms a subgroup $B^2(G,\mathbb{T})$ of $Z^2(G,\mathbb{T})$, and even a normal subgroup, since $Z^2(G,\mathbb{T})$ is abelian. We define $B^1(G,\mathbb{T})$ as the set of all $b : G \to \mathbb{T}$ that satisfy $b(e) = 1$ (and are smooth near the identity, in case G is a Lie group), and define $B^2(G,\mathbb{T})$ as the image of the map

$$\partial_1 : B^1(G,\mathbb{T}) \to Z^2(G,\mathbb{T}); \tag{3.79}$$

$$b \mapsto c : (x,y) \mapsto \frac{b(xy)}{b(x)b(y)}. \tag{3.80}$$

Hence the quotient

$$H^2(G,\mathbb{T}) = \frac{Z^2(G,\mathbb{T})}{B^2(G,\mathbb{T})} \tag{3.81}$$

is called the ***second cohomology group*** of G with coefficients in $\mathbb{T}$.

Theorem 3.3. *if $H^2(G,\mathbb{T})$ is trivial, then any multiplier can be removed by modifying the lift u of h, and the ensuing map $u' : G \to U(H)$ is a homomorphism and hence a unitary representation of G on H.*

This is true by construction. Conversely, if $H^2(G,\mathbb{T})$ is non-trivial, then G will have projective representations that cannot be turned into ordinary ones by a change of phase (for it can be shown that any multiplier $c \in Z^2(G,\mathbb{T})$ is realized by some projective representation). Thus it is important to compute $H^2(G,\mathbb{T})$ for any group G, and see what can be done if it is non-trivial. Three important results are:[5]

$$H^2(\mathbb{R},\mathbb{T}) = 1; \tag{3.82}$$

$$H^2(SU(2),\mathbb{T}) = 1; \tag{3.83}$$

$$H^2(SO(3),\mathbb{T}) = \mathbb{Z}_2. \tag{3.84}$$

Corollary 3.1. *For $G = \mathbb{R}$ or $G = SU(2)$, any homomorphism $h : G \to U(H)/\mathbb{T}$ as in* (3.52) *comes from a continuous unitary representation $u : G \to U(H)$ by* (3.55).

[5] The first has nothing to do with commutativity or topological triviality of $\mathbb{R}$, e.g. $H^2(\mathbb{R}^2,\mathbb{T}) = \mathbb{R}$.

Exercises for week 10 Part 1 (inleveropgave: no. 3)

1. Prove (3.61) from (3.60), i.e. show that the operator $u(x)u(y)u(xy)^*$ must be a multiple of the identity.
2. Verify (3.62) and (3.65).
3. Here is another perspective on multipliers. For some given multiplier c on G, define a group G_c by putting $G_c = G \times \mathbb{T}$ as a set, with group operations

$$(x,z)\cdot(y,w) = (xy, c(x,y)zw); \tag{3.85}$$
$$(x,z)^{-1} = (x^{-1}, \overline{c(x,x^{-1})}\overline{z}), \tag{3.86}$$

where $x,y \in G$ and $z,w \in \mathbb{T}$.

a. Prove that G_c is a group and that $G_c/\mathbb{T} \cong G$ (where $\mathbb{T} \subset G_c$ via $z \mapsto (e,z)$).
b. Show that G_c is isomorphic *as a group* to $G \times \mathbb{T}$ iff $c \in B^2(G,\mathbb{T})$.
c. Show that projective representations u of G with multiplier c bijectively correspond to (ordinary) unitary representations u_c of G_c that satisfy $u_x(e,z) \in \mathbb{T}\cdot 1_H$ for each $z \in \mathbb{T}$, via $u(x) = u_c(x,1)$.
d. Given a homomorphism $h : G \to U(H)/\mathbb{T}$, define

$$G_h = \{(x,u) \in G \times U(H) \mid h(x) = \pi(u)\}. \tag{3.87}$$

Show that this is a subgroup of $G \times U(H)$ (and hence a group). Show that any cross-section $s : U(H)/\mathbb{T} \to U(H)$ of π with associated multiplier c gives an isomorphism $G_h \cong G_c$.

3.5 Self-adjoint operators and Stone's Theorem

Corollary 3.1 shows that for $G = \mathbb{R}$ we may forget projective representations and focus on (strongly continuous) unitary representations of $\mathbb{R}$ on a Hilbert space H, i.e. a map $t \mapsto u_t$, where $t \in \mathbb{R}$ and each $u_t \in B(H)$ is unitary, such that:

$$u_s u_t = u_{s+t},\ s,t \in \mathbb{R}; \tag{3.88}$$

$$u_0 = 1_H; \tag{3.89}$$

$$\lim_{t\to 0} u_t \psi = \psi,\ t \in \mathbb{R},\ \psi \in H. \tag{3.90}$$

These conditions imply, for each $s \in \mathbb{R}$ and $\psi \in H$,

$$\lim_{t\to s} u_t \psi = u_s \psi. \tag{3.91}$$

In the context of quantum mechanics, physicists formally write

$$u_t = e^{-ita}, \tag{3.92}$$

where $a = a^*$ is typically the Hamiltonian of the system (but we avoid the notation h instead of a here, partly in order to rightly suggest far greater generality of the construction; if h *is* the Hamiltonian, one would have $a = h/\hbar$ in (3.92)). ***Stone's Theorem*** makes this rigorous, and even turns the passage from the generator a to the unitary group $t \mapsto u_t$ (and back) into a bijective correspondence.

Theorem 3.4 (Stone). *Let H be a Hilbert space.*

1. *If a is a self-adjoint operator on H, the map $t \mapsto u_t$ defined by* (3.92) *defines a strongly continuous unitary representations of the additive group $\mathbb{R}$ on H.*
2. *Conversely, given such a group of unitary operators, the operator a defined by*

$$a\psi = i \lim_{s\to 0} \frac{u_s\psi - \psi}{s}\ \ (\psi \in D(a)) \tag{3.93}$$

   exists and is self-adjoint.
3. *These constructions are mutually inverse.*

However, our phrasing of this theorem, especially the words "self-adjoint", hide the true complexity of the theorem. This complexity only enters the case where H is infinite-dimensional, and if $\dim(H) < \infty$ we take the theorem at face value and "self-adjoint" is the same as "hermitian". The easiest case is $H = \mathbb{C}$, where the claim is still nontrivial, implying that each map $t \mapsto u_t$ must take the form

$$u_t = e^{-iat}, \tag{3.94}$$

where $a \in \mathbb{R}$ (here identified with a self-adjoint operator on $\mathbb{C}$). This is an exercise. Also, the finite-dimensional case is an exercise: the direction from a to u_t follows by defining the exponential using the continuous functional calculus, with all its associated good properties, see Theorem 2.4. The direction from u_t to a exploits the properties of the map $t \mapsto u_t$ in order to show that the limit in (3.93) exists and gives a self-adjoint operator.

However, if H is infinite-dimensional, one need the theory of *unbounded* self-adjoint operators (created by J. von Neumann). This theory was directly inspired by quantum mechanics, but also formalized many examples from Hilbert's school.

Definition 3.2. *An* **unbounded operator** *on a Hilbert space H is a linear map*

$$a : D(a) \to H,$$

where $D(a)$, called the **domain** *of a. is a* dense *linear subspace of H, for which*

$$\sup\{\|a\psi\|, \psi \in D(a), \|\psi\| = 1\} = \infty. \tag{3.95}$$

If $a \in B(H)$, we may take any dense linear subspace such subspace $D(a) \subset H$, and restrict a to $D(a)$. Then a is, of course, still bounded on $D(a)$, i.e., the supremum in (3.95) is finite, and we may extend it to H by continuity. Nothing has been gained against defining a on all of H in the first place. The cases of interest are therefore those where a satisfies (3.95); since boundedness is the same as continuity (at least for linear operators), one may also say that unbounded operators are discontinuous.

Definition 3.3. *1. The adjoint a^* of an unbounded operator $a : D(a) \to H$ has domain $D(a^*) \subset H$ consisting of all $\psi \in H$ for which the functional*

$$f^a_\psi : D(a) \to \mathbb{C}; \tag{3.96}$$

$$f^a_\psi(\varphi) \; = \; \langle \psi, a\varphi \rangle, \tag{3.97}$$

is bounded (i.e., for which there is $C > 0$ such that $|f^a_\psi(\varphi)| \leq C\|\varphi\|$ for all $\varphi \in H$).

2. For $\psi \in D(a^*)$, the functional f^a_ψ has a unique bounded extension $f^a_\psi : H \to \mathbb{C}$, so by the Riesz–Fréchet Theorem there is a unique vector $\chi \in H$ such that

$$f^a_\psi(\varphi) = \langle \chi, \varphi \rangle. \tag{3.98}$$

3. The adjoint $a^* : D(a^*) \subset H$, then, is defined by $a^*\psi = \chi$, or, equivalently, by

$$\langle a^*\psi, \varphi \rangle = \langle \psi, a\varphi \rangle, \;\; \psi \in D(a^*), \varphi \in D(a). \tag{3.99}$$

Note that, given our assumption that $D(a)$ be dense in H, i.e., $D(a)^- = H$, eq. (3.99) *indeed uniquely specifies $a^*\psi$ (since the inner products $\langle \chi, \varphi \rangle$ of some vector $\chi \in H$ with vectors φ in a dense subset of H uniquely determine χ).*

4. An unbounded operator $a : D(a) \to H$ is called **self-adjoint***, written $a^* = a$, when*

$$D(a^*) = D(a); \tag{3.100}$$

$$a^*\psi = a\psi \;\; \forall \psi \in D(a). \tag{3.101}$$

This definition also makes sense for bounded operators: if $a : D(a) \to H$ is bounded, then as already mentioned a has a unique extension to a bounded operator $a : H \to H$, whose adjoint a^* may be either defined through Definition 3.3 as the adjoint of $a : D(a) \to H$, or, equivalently, as the adjoint of the extension $a : H \to H$. If $D(a) = H$ to begin with, and a is bounded, then also $D(a^*) = H$ (exercise). Thus for $a \in B(H)$, Definition 3.3 reduces to the usual definition of the adjoint.

With this definition, Stone's Theorem holds as stated, adding the claim that the collection of vectors $\psi \in H$ for which the limit in (3.93) exists is dense in H; this collections is the domain $D(a)$, and a is self-adjoint on this specific domain. Moreover, if a is self-adjoint, then (3.92) is rigorously defined by the spectral theorem for possibly unbounded self-adjoint operators (which we will not discuss). The relevant part of this theorem states that if a is a self-adjoint operator on H, and $f : \mathbb{R} \to \mathbb{C}$ is a *bounded* continuous measurable function, then $f(a)$ is a well-defined bounded operator. If f takes values in $\mathbb{R}$, then $f(a)$ is self-adjoint, and if f takes values in $\mathbb{T}$, then $f(a)$ is unitary. Furthermore, one has the same rules (2.63) - (2.65) as for the bounded case. In particular, the operator (3.92) is defined, through this spectral theorem, by $f_t(x) = \exp(-itx)$.

Simple examples of unbounded operators are:

$$H = \ell^2(\mathbb{N}),\ D(a_1) = \ell_c(\mathbb{N}),\ a_1\psi(x) = x\psi(x); \tag{3.102}$$

$$H = L^2(\mathbb{R}),\ D(a_2) = C_c(\mathbb{R}),\ a_2\psi(x) = x\psi(x); \tag{3.103}$$

the latter is (almost) the position operator of quantum mechanics (see below). It is easy to show that (3.95) holds in both cases (exercise). Neither the operator defined by (3.102) nor the one in (3.103) is self-adjoint; it is an exercise to show that

$$D(a_1^*) = \{\psi \in \ell^2 \mid x\psi \in \ell^2\} \equiv D_{max}(q_c); \tag{3.104}$$

$$D(a_2^*) = \{\psi \in L^2 \mid x\psi \in L^2\} \equiv D_{max}(q_d), \tag{3.105}$$

where in both cases we somewhat sloppily write $x\psi$ for the function $x \mapsto x\psi(x)$. Define $q_c : D_{max}(q_c) \to L^2(\mathbb{R})$ and $q_d : D_{max}(q_d) \to \ell^2$ by $q_c\psi(x) = x\psi(x)$ etc. Then

$$D(q_c^*) = D_{max}(q_c); \tag{3.106}$$

$$D(q_d^*) = D_{max}(q_d), \tag{3.107}$$

and hence q_c is self-adjoint on $D(q_c) = D_{max}(q_c)$, as is q_d on $D(q_d) = D_{max}(q_d)$.

Corollary 3.2. *With $t \mapsto u_t$ and a defined and related as in Theorem 3.4, if $\psi \in D(a)$, for each $t \in \mathbb{R}$ the vector $\psi_t = u_t\psi$ lies in $D(a)$ and satisfies*

$$a\psi_t = i\frac{d\psi_t}{dt}, \tag{3.108}$$

whence $t \mapsto \psi_t$ is the unique solution of (3.108) *with initial value $\psi_0 = \psi$.*

With $a = h/\hbar$ (as above), this is the ***time-dependent Schrödinger equation***

$$h\psi_t = i\hbar\frac{d\psi_t}{dt}. \tag{3.109}$$

Exercises for week 10 Part 2 (inleveropgaven: nos. 2, 4).

1. Prove Stone's Theorem for $H = \mathbb{C}$, i.e. show that each continuous unitary representation of $\mathbb{R}$ on $\mathbb{C}$ takes the form (3.94), for some $a \in \mathbb{R}$.
2. Prove Theorem 3.4 if $\dim(H) < \infty$ (see hints in main text).
3. Show that the operators defined in (3.102) - (3.103) satisfy (3.95).
4. Prove (3.106) and (3.107).

Exercise1 for week 11 ((inleveropgave)

1. Prove Corollary 3.2 from Theorem 3.4. Without proof, you may use the fact that the solution of (3.108) with fixed initial condition is unique (and so you need to show that $t \mapsto u_t \psi$ solves (3.108)).

3.6 Position, momentum, and free Hamiltonian

The three basic operators of non-relativistic quantum mechanics are position, denoted q, momentum, p, and the free Hamiltonian h_0. Assuming for simplicity that the particle moves in one dimension, these are informally given on $H = L^2(\mathbb{R})$ by

$$q\psi(x) = x\psi(x); \tag{3.110}$$

$$p\psi(x) = -i\hbar \frac{d}{dx}\psi(x); \tag{3.111}$$

$$h_0\psi(x) = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi(x), \tag{3.112}$$

where m is the mass of the particle under consideration. We put $\hbar = 1$ and $m = 1/2$.

The problem is that these operators are unbounded, like (in general) the abstract Hamiltonians arising from Stone's Theorem in the previous sections. We would like them to be self-adjoint, but the domain on which they are is not obvious. The case of q has been dealt with in the previous section, see (3.105), which gives:

Theorem 3.5. *The operator $q : D(q) \to L^2(\mathbb{R})$, which is formally defined by*

$$q\psi(x) = x\psi(x), \tag{3.113}$$

is self-adjoint on the domain

$$D(q) = \{\psi \in L^2(\mathbb{R}) \mid q\psi \in L^2(\mathbb{R})\}. \tag{3.114}$$

The momentum operator p can be analyzed in two ways. The first relies on the Fourier transform. It is easier to use the Fourier transform. Without proof, we state:

Theorem 3.6. *Let $f \in L^1(\mathbb{R})$. Then*

$$\hat{f}(k) = \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} e^{-ikx} f(x); \tag{3.115}$$

$$\check{f}(x) = \int_{-\infty}^{\infty} \frac{dk}{\sqrt{2\pi}} e^{ikx} f(k), \tag{3.116}$$

are well defined. If $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then $\hat{f} \in L^2(\mathbb{R})$ and $\check{f} \in L^2(\mathbb{R})$, and

$$\|\hat{f}\|_2 = \|\check{f}\|_2 = \|f\|_2. \tag{3.117}$$

Thus the map $f \mapsto \hat{f} \equiv \mathscr{F}(f)$ extends by continuity to a unitary isomorphism

$$\mathscr{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R}), \tag{3.118}$$

with inverse $f \mapsto \check{f} \equiv \mathscr{F}^{-1}(f)$, i.e., for $f \in L^1(\mathbb{R})$ such that $\hat{f} \in L^1(\mathbb{R})$ we have

$$\check{\hat{f}} = \hat{\check{f}} = f. \tag{3.119}$$

A formal computation (which can be made rigorous using distribution theory) gives

$$\mathscr{F}(x^n f^{(m)})(k) = (id/dk)^n (ik)^m \mathscr{F}(f)(k), \tag{3.120}$$

of which we single out the special case $n = 0$, $m = 1$, i.e.,

$$\mathscr{F}(pf)(k) = k\mathscr{F}(f)(k), \tag{3.121}$$

where $p = -id/dx$ is the momentum operator. We may also write this equation as

$$p = \mathscr{F}^{-1} q \mathscr{F}, \tag{3.122}$$

where q is the position operator. This would make perfectly good sense if all variables were called x, but it is in line with the above notation to let q act on the k-variable (or in 'momentum space', as the physicists say), i.e., $q\hat{\psi}(k) = k\hat{\psi}(k)$.

Lemma 3.7. *Let $u : H_1 \to H_2$ be a unitary operator between Hilbert spaces H_1 and H_2, and suppose $a : D(a) \to H_1$ is self-adjoint (where $D(a) \subset H_1$). Then*

$$\tilde{a} = uau^* : D(\tilde{a}) \to H_2; \tag{3.123}$$

$$D(\tilde{a}) = \{\psi_2 \in H_2 \mid u^*\psi_2 \in D(a)\} \tag{3.124}$$

is self-adjoint.

Theorem 3.7. *The momentum operator is self-adjoint on the domain*

$$\begin{aligned} D(p) &= \{\psi \in L^2(\mathbb{R}) \mid \hat{\psi} \in D(q)\} \\ &= \{\hat{\psi} \in L^2(\mathbb{R}) \mid k \mapsto k\hat{\psi}(k) \in L^2(\mathbb{R})\}. \end{aligned} \tag{3.125}$$

The proof of Lemma 3.7 is an exercise; Theorem 3.7 follows immediately as a special case, where u is the Fourier transform. Similarly, for the free Hamiltonian,

$$D(h_0) = D(d^2/dx^2) = \{\psi \in L^2(\mathbb{R}) \mid k \mapsto k^2\hat{\psi}(k) \in L^2(\mathbb{R})\}, \tag{3.126}$$

and $h_0^* = h_0$ on this domain. Similarly for any power of p, $(d/dx)^n$ is self-adjoint on

$$D((d/dx)^n) = \{\psi \in L^2(\mathbb{R}) \mid k \mapsto k^n\hat{\psi}(k) \in L^2(\mathbb{R})\}. \tag{3.127}$$

However, the domain of p^n is expressed in terms of the Fourier transform $\hat{\psi}$ rather than ψ itself, which is unsatisfactory. This could be resolved by writing:

$$D((d/dx)^n) = \{\psi \in L^2(\mathbb{R}) \mid \psi^{(n)} \in L^2(\mathbb{R})\}, \tag{3.128}$$

but what does the n'th derivative $\psi^{(n)}$ mean for L^2-functions? Let $\mathscr{D}(\mathbb{R})$ be $C_c^\infty(\mathbb{R})$ as a set, equipped with the topology in which $\varphi_\lambda \to \varphi$ iff there is a compact $K \subset \mathbb{R}$ such that $\text{supp}(\varphi_\lambda) \subseteq K$ for all λ, and for all $m \in \mathbb{N}$ (including zero) one has

$$\|D^m(\varphi_\lambda - \varphi)\|_\infty \to 0, \tag{3.129}$$

where $D = d/dx$. Elements of $\mathscr{D}(\mathbb{R})$ are called ***test functions***. A linear map $u : \mathscr{D}(\mathbb{R}) \to \mathbb{C}$ is continuous iff for each compact $K \subset \mathbb{R}$ there is $m \in \mathbb{N}$ and $C > 0$ with

$$|\langle u, \varphi \rangle| \equiv |u(\varphi)| \leq C \|D^m \varphi\|_\infty. \tag{3.130}$$

Distributions are elements of the space $\mathscr{D}'(\mathbb{R})$ of all continuous maps $u : \mathscr{D}(\mathbb{R}) \to \mathbb{C}$. This space carries the weak topology, in which $u_\lambda \to u$ iff $\langle u_\lambda, \varphi \rangle \to \langle u, \varphi \rangle$ for each $\varphi \in \mathscr{D}(\mathbb{R})$. In this topology, $\mathscr{D}(\mathbb{R})$ is dense in $\mathscr{D}'(\mathbb{R})$, where $u \in \mathscr{D}(\mathbb{R})$ defines $u \in \mathscr{D}'(\mathbb{R})$ through the L^2 inner product, i.e., $\langle u, \varphi \rangle = \langle \overline{u}, \varphi \rangle_{L^2(\mathbb{R})}$. We have

$$\mathscr{D}(\mathbb{R}) \subset L^2(\mathbb{R}) \subset \mathscr{D}'(\mathbb{R}), \tag{3.131}$$

in which each embedding is continuous and dense. This is a ***Gelfand triple***.

For each $m \in \mathbb{N}$, the ***weak derivative*** $D^n u$ of $u \in \mathscr{D}'(\mathbb{R})$ is defined by

$$\langle D^m u, \varphi \rangle = (-1)^m \langle u, D^m \varphi \rangle. \tag{3.132}$$

This definition may be motivated by faking the formula $\langle u, \varphi \rangle = \int_{\mathbb{R}} u(x)\varphi(x)$, which on repeated partial integration gives (3.132). In view of (3.131), for $\psi \in L^2(\mathbb{R})$ the (weak) derivative $D^m \psi$ is defined, and hence (3.128) is well defined. The fact that D^n is self-adjoint on the domain (3.128) follows from the earlier results using the Fourier transform, but it may also be shown directly, just from the above definitions and from the definition of self-adjointness. This is not trivial, and as an illustration we just prove that the momentum operator $p = -id/dx$ is self-adjoint on the domain

$$D(p) = \{\psi \in L^2(\mathbb{R}) \mid \psi' \in L^2(\mathbb{R})\}, \tag{3.133}$$

where ψ' denotes the weak derivative of $\psi \in L^2(\mathbb{R}) \subset \mathscr{D}'(\mathbb{R})$. Let $\psi \in D(p^*)$, so that by definition for each $\varphi \in D(p)$ we have

$$\langle p^* \psi, \varphi \rangle = \langle \psi, p\varphi \rangle = -i\langle \psi, \varphi' \rangle. \tag{3.134}$$

Since $\mathscr{D}(\mathbb{R}) \subset D(p)$, this is true in particular for each $\varphi \in \mathscr{D}(\mathbb{R})$, in which case the right-hand side equals $-i\psi'(\varphi)$, where the derivative is weak. But this equals $\langle p^* \psi, \varphi \rangle$ and so the distribution $-i\psi'$ is given by taking the inner product with $p^* \psi \in L^2(\mathbb{R})$. Hence $-i\psi' = p^* \psi \in L^2(\mathbb{R})$, and in particular $\psi' \in L^2(\mathbb{R})$, so that $\psi \in D(p)$. This proves that $D(p^*) \subseteq D(p)$. Conversely, it is immediate from the definition of the adjoint that $\mathscr{D}(\mathbb{R}) \subset D(p^*)$, but how to proceed?

Definition 3.4. *1. The* **graph** *of an operator* $a : D(a) \to H$ *is the vector space*

$$G(a) = \{(\psi, a\psi), \psi \in D(a)\} \subset H \oplus H, \tag{3.135}$$

where $H \oplus H$ *is* $H \times H$ *as a set, seen as a vector space in the obvious way, and seen as a Hilbert space in the inner product*

$$\langle (v, w), (v', w') \rangle_{H \oplus H} = \langle v, v' \rangle_H + \langle w, w' \rangle_H. \tag{3.136}$$

Thus $H \oplus H$ has a norm $\|(v,w)\| = \sqrt{\|v\|^2 + \|w\|^2}$.

2. *An operator $a : D(a) \to H$ is* **closed** *if its graph is closed.*
3. *An operator $a : D(a) \to H$ is* **closable** *if the closure of its graph is the graph of some (uniquely defined) operator, called the* **closure** *a^- of a.*
4. *An operator $a : D(a) \to H$ is* **essentially self-adjoint** *if its closure is self-adjoint, i.e., if $(a^-)^* = a^-$, or, equivalently, if $a^- = a^*$, or if $a^{**} = a^*$.*
5. *If $a \subset a^*$ i.e., if $\langle a\varphi, \psi\rangle = \langle \varphi, a\psi\rangle$, $\varphi, \psi \in D(a)$, then a is called* **symmetric**.[6]

If a is closable, the domain $D(a^-)$ of its closure consists of all $\psi \in H$ for which there exists a sequence (ψ_n) in $D(a)$ such that $\psi_n \to \psi$ *and* $a\psi_n$ converges, on which we define a^- by $a^-\psi = \lim_n a\psi_n$. It is easy to show that a^* is always closed (provided $D(a)$ is dense), and hence a self-adjoint operator is closed. In practice, self-adjoint operators often arise as closures of essentially self-adjoint operators a.

As a case in point, we return to to the momentum operator. Define $\tilde{p}$ as the restriction of p to $\mathscr{D}(\mathbb{R})$. Then $D(\tilde{p}) \subset D(p^*)$, and hence $D(\tilde{p}^-) \subset D(p^*)$, since p^* is closed. Finally, $\tilde{p}^- = p$, so that $D(p) \subset D(p^*)$. Since we already had the opposite inclusion, it follows that $D(p) = D(p^*)$, and since the actions of p and p^* on this domain coincide, we have $p^* = p$. In other words, p is self-adjoint (i.e. on $D(p)$).

We finally discuss the nature of Heisenberg's ***canonical commutation relations***

$$[q, p] = i\hbar \cdot 1_H. \tag{3.137}$$

The problem is that the naive verification of these relations on the domains were p and q are defined and self-adjoint does not work, because if $\psi \in D(p)$ then $p\psi$ may not lie in $D(q)$, and similarly, if $\psi \in D(q)$, then $q\psi$ may not lie in $D(p)$, so that neither $qp\psi$ nor $pq\psi$ may be defined. This is somewhat resolved by finding a common ***core*** for the operators in question, i.e. a domain on which each is essentially self-adjoint and (3.137) holds. An example of such a core is $\mathscr{D}(\mathbb{R})$ (exercise).

Exercises for week 11 Part 2 (inleveropgave: 3,4)

1. Prove Lemma 3.7.
2. Prove that a^* is always closed (provided $D(a)$ is dense).
3. We have already seen that p is essentially self-adjoint on $\mathscr{D}(\mathbb{R})$. Show that q is also essentially self-adjoint on $\mathscr{D}(\mathbb{R})$.
4. Using the closed graph theorem, prove the ***Hellinger–Toeplitz Theorem***:

   Theorem 3.8. *If a linear map $a : H \to H$ satisfies $a^* = a$, then it is bounded.*

---

[6] We write $a \subset b$ if $D(a) \subset D(b)$ and $a\psi = b\psi$ for all $\psi \in D(a)$.

3.7 Some abstract representation theory

Having dealt with the unitary representation theory of the important non-compact Lie group $\mathbb{R}$, we now move towards compact groups, like $SU(2)$ and $SO(3)$. Unlike those of non-compact groups, unitary representations of compact groups are built from elementary bricks called *irreducible* representations. Moreover, for compact groups these are finite-dimensional, which makes it a lot easier to classify them.

Definition 3.5. *Let u be a (projective) unitary representation of a group G on a Hilbert space H and let $K \subset H$ be a closed linear subspace of H.*[7] *We call K* **stable** *under $u(G)$ if $u(x)\psi \in K$ for all $x \in G$ and $\psi \in K$.*[8] *We call H* **irreducible** *under u (equivalently: we call u irreducible, given H) if the only closed linear subspaces of H that are stable under u are either $K = H$ or $K = \{0\}$. If not, (H,u) is* **reducible**.

For example, any one-dimensional representation of any group is trivially irreducible. Furthermore, the defining representations of $SU(2)$ on $\mathbb{C}^2$ and of $SO(3)$ on $\mathbb{C}^3$ (here seen as the complexification of $\mathbb{R}^3$) are irreducible. One the other hand, suppose we have unitary representations $(u_i(G))$ of a given group G, where u_i acts on a Hilbert space H_i. We can the form the ***direct sum*** $H = \bigoplus_i H_i$, defined as the set of all sequences $\psi = (\psi_i)$ with $\psi_i \in H_i$ and

$$\sum_i \|\psi_i\|^2_{H_i} < \infty, \tag{3.138}$$

with pointwise vector space operations, and inner product

$$\langle \psi, \varphi \rangle = \sum_i \langle \psi_i, \varphi_i \rangle_{H_i}. \tag{3.139}$$

It is easy to show that $\bigoplus_i H_i$ is complete, and carries a unitary representation of G:

$$\bigoplus_i u_i(x)\psi = (u_i(x)\psi_i). \tag{3.140}$$

Clearly, this representation (call it u) is reducible, since any of the $H_i \subset \bigoplus_j H_j$ is stable under u and hence can act like the K in Definition 3.5 (on top of which H_i may not be irreducible itself, given rise to further stable Hilbert subspaces $K \subset H_i$).

Definition 3.6. *Let u be a (projective) unitary representation of a group G on a Hilbert space H. We call H* **completely reducible** *under u if H has irreducible subspaces H_i such that $H = \bigoplus_i H_i$ and $u = \bigoplus_i u_i$.*

On might expect that any reducible representation is completely reducible, but this is not the case for general groups G. Consider $G = \mathbb{R}$ and $H = L^2(\mathbb{R})$, where

$$u(y)\psi(x) = \psi(x - y). \tag{3.141}$$

[7] If $\dim(H) < \infty$, then any linear subspace is closed.

[8] Stable subspaces are often called *invariant*, but we reserve this terminology for subspaces $L \subset H$ such that $u(x)\psi = \psi$ for each $x \in G$ and $\psi \in L$. Of course an invariant subspace is also stable.

This is easier to analyze in momentum space, i.e., we compute the Fourier transform

$$\hat{u}(y)\hat{\psi}(k) \equiv \mathscr{F} u(y)\mathscr{F}^{-1}\hat{\psi}(k) = \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}}\, e^{-ikx} u(y)\psi(x) = \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}}\, e^{-ikx}\psi(x-y)$$
$$= e^{-iky}\hat{\psi}(k). \tag{3.142}$$

From this, we see that for any measurable subset $\Omega \subset \mathbb{R}$ the subset $L^2(\Omega) \subset L^2(\mathbb{R})$ defined by $\psi \in L^2(\Omega)$ iff $\psi(x) = 0$ for all $x \notin \Omega$ (or, equivalently, $L^2(\Omega) = 1_\Omega L^2(\mathbb{R})$, where 1_Ω is the multiplication operator by the characteristic function of Ω) is stable under each $\hat{u}(y)$. Consequently, $L^2(\mathbb{R})$ is not irreducible under $\hat{u}$ (and, since the property of (ir)reducibility is preserved under unitary equivalence, it is not irreducible under u either). However, any subspace $L^2(\Omega)$ contains subspaces $L^2(\tilde{\Omega})$, where $\tilde{\Omega} \subset \Omega$, so no $L^2(\Omega)$ is irreducible: it should be clear from (3.142) that the only stable subspaces would correspond to $\Omega = \{k_0\}$ for some $k_0 \in \mathbb{R}$, but since points have (Lebesgue) measure zero in $\mathbb{R}$, we have $L^2(\{k_0\}) = 0$. Consequently, although $L^2(\mathbb{R})$ is reducible under $u(\mathbb{R})$, it is not completely reducible.

Without proof we quote two important results about irreducibility. The first is ***Schur's Lemma***, in which the commutant S' of a subset $S \subset B(H)$ is defined by

$$S' = \{a \in B(H) \mid ab = ba \forall b' \in S\}. \tag{3.143}$$

For example, if $S = \mathbb{C}\cdot 1_H$ then $S' = B(H)$, and conversely, $B(H)' = \mathbb{C}\cdot 1_H$. If $H = \mathbb{C}^n$ and D_n are the diagonal matrices, then $D_n' = D_n$ (exercise).

Lemma 3.8 (Schur). *A unitary representation u of a group G is irreducible iff*

$$u(G)' = \mathbb{C}\cdot 1, \tag{3.144}$$

i.e., if $au(x) = u(x)a$ for each $x \in G$, then $a = \lambda \cdot 1_H$ for some $\lambda \in \mathbb{C}$.

Theorem 3.9 (Peter–Weyl). *Let G be a compact group.*

1. *Every irreducible representation of G is finite-dimensional.*
2. *Every unitary representation of G is a direct sum of irreducible representations.*

The first claim also holds for some non-compact groups, notably abelian groups, for which Schur's Lemma implies that any irreducible representation must be one-dimensional (exercise). But there are many (necessarily non-compact and non-abelian) groups, all of whose nontrivial unitary irreducible representations are infinite-dimensional (such as the Lorentz group or the Poincaré group or $SL(2,\mathbb{R})$).

Let u be a finite-dimensional unitary representation of some Lie group G (not necessarily compact, although we well use the following construction in the compact case only). Since H is finite-dimensional, the following operation is unproblematic: for $X \in \mathfrak{g}$ we define an operator

$$du(X)\ :\ H \to H; \tag{3.145}$$

$$du(X)\psi = \frac{d}{dt} u\left(e^{tX}\right) \psi_{|t=0}. \tag{3.146}$$

This gives a linear map $du : \mathfrak{g} \to B(H)$, which satisfies (exercise)

$$[du(X), du(Y)] = du([X,Y]); \tag{3.147}$$
$$du(X)^* = -du(X). \tag{3.148}$$
$$e^{du(X)} = u\left(e^X\right). \tag{3.149}$$

Note that physicists use Planck's constant $\hbar > 0$ and like to write

$$\rho(X) = i\hbar du(X), \tag{3.150}$$

so that $\rho(X)^* = \rho(X)$ and

$$[\rho(X), \rho(Y)] = i\hbar\rho([X,Y]). \tag{3.151}$$

More generally, we call a linear map $\pi : \mathfrak{g} \to B(H)$ (where $H \cong \mathbb{C}^n$ is finite-dimensional, so that $\pi : \mathfrak{g} \to M_n(\mathbb{C})$), a ***skew-adjoint*** representation of $\mathfrak{g}$ on H if

$$[\pi(X), \pi(Y)] = \pi([X,Y]); \tag{3.152}$$
$$\pi(X)^* = -\pi(X). \tag{3.153}$$

If the representation u is projectve, see (3.60), instead of (3.147) we obtain

$$[du(X), du(Y)] = du([X,Y]) + i\varphi(X,Y) \cdot 1_H, \tag{3.154}$$

where φ is given by

$$\varphi(X,Y) = \frac{d}{ds}\frac{d}{dt}\left[\Gamma\left(e^{tX}, e^{sY}\right)\right]_{|s=t=0} - (X \leftrightarrow Y), \tag{3.155}$$

assuming u satisfies (3.60) with $c(x,y) = \exp(i\Gamma(x,y))$, where $\Gamma : G \times G \to \mathbb{R}$.

More generally, suppose $\pi : \mathfrak{g} \to B(H)$ satisfies

$$[\pi(X), \pi(Y)] = \pi([X,Y]) + i\varphi(X,Y) \cdot 1_H, \tag{3.156}$$

for some function $\varphi : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$, which necessarily satisfies (exercise)

$$\varphi(X,Y) = -\varphi(Y,X); \tag{3.157}$$
$$\varphi(X,[Y,Z]) + \varphi(Z,[X,Y]) + \varphi(Y,[Z,X]) = 0. \tag{3.158}$$

Taking the trace (which is finite dimension is unproblematic) yields

$$\varphi(X,Y) = \frac{i}{n}\mathrm{Tr}\left(\pi([X,Y])\right), \tag{3.159}$$

where $n = \dim(H) < \infty$. We may define a linear function $\theta : \mathfrak{g} \to \mathbb{R}$ by

$$\theta(X) = \frac{i}{n}\mathrm{Tr}\left(\pi(X)\right) \tag{3.160}$$

so that $\varphi(X,Y) = \theta([X,Y])$, and we may now remove the spoiler φ by redefining

$$\tilde{\pi}(X) = \pi(X) + i\theta(X) \cdot 1_H, \tag{3.161}$$

since we then have (3.152) - (3.153) with $\pi \rightsquigarrow \tilde{\pi}$ (exercise). Consequently, any *finite-dimensional* skew-adjoint projective representation of a Lie algebra $\mathfrak{g}$ (i.e., a linear map $\pi : \mathfrak{g} \to B(H)$ satisfying (3.153) and (3.156)) may be redefined so as to obtain a true representation of the Lie algebra in question. We now need:

Theorem 3.10. *Let G_1 and G_2 be Lie groups, with Lie algebras $\mathfrak{g}_1$ and $\mathfrak{g}_2$, respectively, and suppose that G_1 is connected and simply connected. Then every Lie algebra homomorphism $\varphi : \mathfrak{g}_1 \to \mathfrak{g}_2$ comes from a unique Lie group homomorphism $\Phi : G_1 \to G_2$ through $\varphi = d\Phi$, where (realizing G_1 and G_2 as matrices)*

$$d\Phi(X) = \frac{d}{dt} u\left(e^{tX}\right)_{|t=0}. \tag{3.162}$$

Let H be a finite-dimensional Hilbert space, so that $B(H) \cong M_n(\mathbb{C})$, where $n = \dim(H)$, and take $U(H) \cong U_n(\mathbb{C})$ to be the group of all unitary matrices on $\mathbb{C}^n$. The Lie algebra $\mathfrak{u}_n(\mathbb{C})$ of $U_n(\mathbb{C})$ consists of all skew-adjoint $n \times n$ complex matrices. Since irreducibility is preserved under the correspondence $u(G) \leftrightarrow du(\mathfrak{g})$, we infer:

Corollary 3.3. *Let $\tilde{G}$ be a connected and simply connected Lie group with Lie algebra $\mathfrak{g}$. Any finite-dimensional skew-adjoint representation $\pi : \mathfrak{g} \to \mathfrak{u}_n(\mathbb{C})$ of $\mathfrak{g}$ comes from a unique unitary representation $u(\tilde{G})$ through* (3.146)*, i.e.,*

$$\pi(X)\psi = \frac{d}{dt} u\left(e^{tX}\right) \psi_{|t=0}; \tag{3.163}$$

$$e^{\pi(X)} = u\left(e^{X}\right). \tag{3.164}$$

Thus there is a bijective correspondence between finite-dimensional unitary representations of $\tilde{G}$ and finite-dimensional skew-adjoint representations of $\mathfrak{g}$. In particular, if $\tilde{G}$ is compact, this specializes to a bijective correspondence between unitary irreducible representations of $\tilde{G}$ and skew-adjoint irreducible representations of $\mathfrak{g}$.

What about a group like $SO(3)$, which is not simply connected? Here another deep result from Lie theory is needed (called ***Lie's Third Theorem***, proved by É. Cartan):

Theorem 3.11. *Let G be a connected Lie group G with Lie algebra $\mathfrak{g}$. There exists a simply connected Lie group $\tilde{G}$, unique up to isomorphism, such that:*

- *The Lie algebra of $\tilde{G}$ is $\mathfrak{g}$.*
- *$G \cong \tilde{G}/D$, where D is a discrete normal subgroup of the center of $\tilde{G}$.*
- *$D \cong \pi_1(G)$, i.e., the fundamental group of G, which is therefore abelian.*

Proving this would be challenging even in a course on Lie groups! If we now start from a (possibly) projective finite-dimensional representation $u : G \to U(H)$, pass to $du(\mathfrak{g})$, and move to a true representation $\widetilde{du}(\mathfrak{g})$. By Corollary 3.3 we may exponentiate $\widetilde{du}$ to a unitary representation $\tilde{u}(\tilde{G})$, so that

$$\widetilde{du} = d\tilde{u}. \tag{3.165}$$

Subsequently, in the setting of Theorem 3.11, we pick a (Borel measurable) cross-section

$$\tilde{s} : G \to \tilde{G} \tag{3.166}$$

of the canonical projection

$$\tilde{\pi} : \tilde{G} \to G = \tilde{G}/D. \tag{3.167}$$

As always, this means that $\tilde{\pi} \circ \tilde{s} = \mathrm{id}_G$, and $\tilde{s}$ is supposed to be smooth near the identity, and chosen such that $\tilde{s}(e_G) = e_{\tilde{G}}$, where e_G and $e_{\tilde{G}}$ are the unit elements of G and $\tilde{G}$, respectively. We then obtain a map $u' : G \to U(H)$ by $u' = \tilde{u} \circ \tilde{s}$, which is easily shown to satisfy $\pi \circ u' = \pi \circ u$, where $\pi : U(H) \to U(H)/\mathbb{T}$ is the canonical projection, as before. This means that any homomorphism $h : G \to U(H)/\mathbb{T}$ that gave rise to the projective representation $u : G \to U(H)$ through (3.59), always comes from some unitary representation of $\tilde{G}$. In particular, all homomorphisms $h : SO(3) \to U(H)/\mathbb{T}$ come from unitary representations of $SU(2)$ on H in the way just described. These will be classified next week.

Exercises for week 12 (inleveropgaven: nos. 1, 2, 4)

1. Prove that $D'_n = D_n$.
2. Infer from Schur's Lemma that any irreducible representation of an abelian group is one-dimensional. Now prove that (assuming representations are continuous):
   - Any irreducible representation of $\mathbb{R}$ is given by $u_p(x) = e^{ipx}$, for some $p \in \mathbb{R}$;
   - Any irreducible representation of $\mathbb{Z}$ is given by $u_z(n) = z^n$, for some $z \in \mathbb{T}$;
   - Any irreducible representation of $\mathbb{T}$ is given by $u_n(z) = z^n$, for some $n \in \mathbb{Z}$.
3. Prove (3.147) - (3.149). *Hint:* regarding G as a matrix group, we have
$$[X,Y] = \frac{d}{ds}\frac{d}{dt}\left(e^{tX}e^{sY}e^{-tX}\right)_{|s=t=0}. \tag{3.168}$$
4. Prove (3.154) - (3.155).
5. Show that (3.161) satisfies (3.152) - (3.153) with $\pi \rightsquigarrow \tilde{\pi}$.
6. Verify the claims after (3.167).

3.8 Irreducible representations of $SU(2)$ and $SO(3)$

In (mathematical) physics the most important compact Lie groups are $SU(2)$ and $SO(3)$. The latter describes spatial rotation symmetries, and as we shall see its representation theory is closely related to that of $SU2)$. The latter possess an additional relevance to physics, too, as an internal symmetry group of elementary particles.

Recall that $\mathfrak{su}(2) \cong \mathbb{R}^3$ as a vector space, with basis (S_1, S_2, S_3), where $S_k = -\frac{1}{2}i\sigma_k$ $(k = 1,2,3)$, where $(\sigma_1, \sigma_2, \sigma_3)$ are the Pauli matrices (3.16), and hence

$$[S_i, S_j] = \varepsilon_{ijk} S_k, \tag{3.169}$$

as the commutation relations.[9] Since $SU(2)$ is compact, all its unitary irreducible representations are finite-dimensional. Since $G = SU(2)$ is also connected and simply connected, its irreducible (unitary) representations u bijectively correspond to irreducible (skew-adjoint) representations du of its Lie algebra $\mathfrak{g}$. This correspondence is given by (3.163) - (3.164). By Schur's Lemma, some representation $\pi(\mathfrak{g})$ is irreducible iff the only operators that commute with all $\pi(X)$ are multiples of the unit operator. We will find "all" skew-adjoint irreducible representations π of $\mathfrak{su}(2)$ on (necessarily finite-dimensional) Hilbert spaces, up to (unitary) equivalence.[10]

Assume $H \cong \mathbb{C}^n$; this we may do, since we classify up to equivalence. By linearity, finding a Lie algebra homomorphism $\pi : \mathfrak{su}(2) \to \mathfrak{u}_n(\mathbb{C})$ (where $\mathfrak{u}_n(\mathbb{C})$, consisting of all skew-adjoint $n \times n$ matrices, is the Lie algebra of the unitary group $U_n(\mathbb{C})$ of all unitary $n \times n$ matrices),[11] is the same as finding $n \times n$ matrices

$$L_k = i\pi(S_k) \tag{3.170}$$

that satisfy

$$[L_i, L_j] = i\varepsilon_{ijk} L_k, \tag{3.171}$$

i.e., $[L_1, L_2] = iL_3$, etc., and

$$L_k^* = L_k. \tag{3.172}$$

It turns out to be convenient to introduce the ***ladder operators***

$$L_\pm = L_1 \pm iL_2, \tag{3.173}$$

with ensuing commutation relations

$$[L_3, L_\pm] = \pm L_\pm; \tag{3.174}$$

$$[L_+, L_-] = 2L_3. \tag{3.175}$$

---

[9] Here ε_{ijk} is the totally anti-symmetric symbol with $\varepsilon_{123} = 1$ etc., so that (3.169) comes down to $[S_1, S_2] = S_3$, $[S_3, S_1] = S_2$, and $[S_2, S_3] = S_1$.

[10] Let G be a Lie group with Lie algebra $\mathfrak{g}$. Two unitary representations $u_i : G \to U(H_i)$, $i = 1,2$, are ***equivalent*** if there is a unitary operator $v : H_1 \to H_2$ such that $u_2(x) = vu_1(x)v^*$ for each $x \in G$. Similarly, if the Hilbert space H_i are finite-dimensional, two skew-adjoint representations $\pi_i : \mathfrak{g} \to U(H_i)$, $i = 1,2$, are ***equivalent*** if there is a unitary $v : H_1 \to H_2$ such that $\pi_2(X) = v\pi_1(X)v^*$, $X \in \mathfrak{g}$.

[11] The reader may prefer to simply write $\pi : \mathfrak{su}(2) \to M_n(\mathbb{C})$, as in the lectures.

The simple idea is now to diagonalize L_3, which is possible as $L_3^* = L_3$. Hence

$$H = \bigoplus_{\lambda \in \sigma(L_3)} H_\lambda, \tag{3.176}$$

where $\sigma(L_3)$ is the spectrum of L_3 (which in this finite-dimensional case consists of its eigenvalues), and H_λ is the eigenspace of L_3 for eigenvalue λ (i.e., if $\upsilon \in H_\lambda$, then $L_3\upsilon = \lambda\upsilon$). The structure of this decomposition is as follows.

1. For any $\lambda \in \sigma(L_3)$ and nonzero $\upsilon_\lambda \in H_\lambda$, eq. (3.174) implies:
   - either $\lambda + 1 \in \sigma(L_3)$ and $L_+\upsilon_\lambda \in H_{\lambda+1}$ (as a nonzero vector);
   - or $L_+\upsilon_\lambda = 0$.

   Similarly, either $\lambda - 1 \in \sigma(L_3)$ and $L_-\upsilon_\lambda \in H_{\lambda-1}$, or $L_-\upsilon_\lambda = 0$.
2. Let $\lambda_0 = \min\sigma(L_3)$ be the smallest eigenvalue of L_3. Pick some $0 \neq \upsilon_{\lambda_0} \in H_{\lambda_0}$. Then there must be some $k \in \mathbb{N}$ such that $L_+^{k+1}\upsilon_{\lambda_0} = 0$, whereas all vectors $L_+^l\upsilon_{\lambda_0}$ for $l = 0, \ldots, k$ are nonzero (and lie in H_{λ_0+l}). If not, H would be infinite-dimensional. Likewise, for any $\lambda \in \sigma(L_3)$ there must be $l \in \mathbb{N}$ such that $L_-^l v_\lambda = 0$ (for otherwise H would be infinite-dimensional).

It is an exercise to show that $\lambda_0 = -k/2$, so that, repeatedly applying L_+, we obtain

$$\{-k/2, -k/2+1, \ldots, k/2-1, k/2\} \subseteq \sigma(L_3). \tag{3.177}$$

Now consider $H' = \mathbb{C}\cdot\upsilon_{\lambda_0} \oplus \mathbb{C}\cdot L_+\upsilon_{\lambda_0} \oplus \cdots \oplus L_+^{k-1}\upsilon_{\lambda_0} \oplus L_+^k\upsilon_{\lambda_0} \subseteq H$; this is the subspace of H with basis $(\upsilon_{\lambda_0}, L_+\upsilon_{\lambda_0}, \ldots, L_+^{k-1}\upsilon_{\lambda_0}, L_+^k\upsilon_{\lambda_0})$. As a further exercise, it follows that irreducibility (assumed!) implies $H' = H$. To finish, we need to define an inner product (or normalize the v_λ) so that $L_3^* = L_3$ and $L_+^* = L_-$ (which of course implies $L_-^* = L^+$), so that $\pi(A)^* = -\pi(A)$ for all $A \in \mathfrak{su}(2)$. If (u_l) is the standard orthonormal basis of $\mathbb{C}^n$ with $n = 2j+1$, labeled by $l = 0, 1, \ldots, 2j$, we put

$$L_3 u_l = (l-j)u_l; \tag{3.178}$$

$$L_+ u_l = \sqrt{(l+1)(n-l-1)}\,u_{l+1}; \tag{3.179}$$

$$L_- u_l = \sqrt{l(n-l)}\,u_{l-1}. \tag{3.180}$$

Here (3.179) is even formally correct for $l = 2j$, since in that case $n - 2j - 1 = 0$, and similarly, (3.180) formally holds even for $l = 0$. Clearly, $\lambda = l - j$, i.e., $v_\lambda = u_{\lambda+j}$.

This leads to the following theorem stating the representation theory of $SU(2)$.

Theorem 3.12. *Let π be an irreducible representation of the Lie algebra $\mathfrak{g}$ of $SU(2)$ on some finite-dimensional Hilbert space H. Then $H \cong \mathbb{C}^{2j+1}$ for some positive half-integer or integer j, and any such j is possible (i.e., $j = 0, 1/2, 1, 3/2, \ldots$).*

Furthermore, the spectrum $\sigma(L_3)$ of the self-adjoint operator $L_3 = i\pi(S_3)$ is

$$\sigma(L_3) = \{-j, -j+1, \cdots, j-1, j\}, \tag{3.181}$$

and if (3.176) *is the spectral decomposition of H relative to L_3, then:*

1. Each H_λ is one-dimensional, $\lambda \in \sigma(L_3)$;
2. For $\lambda < j$ the operator L_+ maps H_λ to $H_{\lambda+1}$, whereas $L_+ = 0$ on H_j;
3. For $\lambda > -j$ the operator L_- maps H_λ to $H_{\lambda-1}$, whereas $L_- = 0$ on H_{-j}.

Conversely, these properties determine π up to unitary equivalence, so this gives a complete list of all irreducible representations of $\mathfrak{g}$, and hence also of $SU(2)$ (up to unitary equivalence). In particular, irreducible representations of $SU(2)$ are determined by their dimension $n = \dim(H)$, and each dimension $n = 2j+1 \in \mathbb{N}$ occurs.

In principle, we should find the unitary representation of $SU(2)$ corresponding to some representation of its Lie algebra by exponentiation. However, it is easier to proceed in the oppositie direction. For each $j \in \mathbb{N}/2$, define H_j as the complex vector space of all homogeneous polynomials p in two variables $z = (z_1, z_2)$ of degree $2j$. A basis of H_j is given by $(z_1^{2j}, z_1^{2j-1}z_2, \ldots, z_1 z_2^{2j-1}, z_2^{2j})$, which has $2j+1$ elements. So $\dim(H_j) = 2j+1$, which is a good start. We then consider the map

$$D_j : SU(2) \to B(H_j); \tag{3.182}$$
$$D_j(u)f(z) = f(zu). \tag{3.183}$$

Clearly $D_j(e)f(z) = f(z \cdot 1_2 = f(z))$, so $D_j(e) = 1$, and

$$D_j(u)D_j(v)f(z) = D_j(v)f(zu) = f(zuv) = D_j(uv)f(z),$$

so $D_j(u)D_j(v) = D_j(uv)$. Hence D_j is a representation of $SU(2)$.

We now compute $L_3 = -\frac{1}{2}iS_3$ on this space. From (3.163) with $\tilde{u} \rightsquigarrow D_j$, we have

$$L_3 = -\tfrac{1}{2}idD_j \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = -\tfrac{1}{2}i\frac{d}{dt}D_j \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}_{t=0}, \tag{3.184}$$

so that

$$L_3 f(z) = -\tfrac{1}{2}i\frac{d}{dt}f(e^{it}z_1, e^{-it}z_2)_{t=0} = \tfrac{1}{2}\left(z_1\frac{\partial_e f(z)}{\partial z_1} - z_2\frac{\partial f(z)}{\partial z_2}\right). \tag{3.185}$$

Similarly, we obtain

$$L_+ f(z) = z_1\frac{\partial f(z)}{\partial z_2}; \tag{3.186}$$
$$L_- f(z) = z_2\frac{\partial f(z)}{\partial z_1}. \tag{3.187}$$

Hence $f_{2j}(z) = z_1^{2j}$ gives $L_3 f_{2j} = jf_{2j}$, and $f_0(z) = z_2^{2j}$ gives $L_3 f_0 = -jf_0$. In general, $f_l(z) = z_1^l z_2^{2j-l}$ spans the eigenspace H_λ of L_3 with eigenvalue $\lambda = l - j$, so that, as in (3.178), we may identify f_l with the standard basis vector u_l of $H_j \cong C^{2j+1}$, where $l = 0, 1, \ldots, 2j$. This confirms (3.181), as well as the fact that the corresponding eigenspaces are all one-dimensional. The rest is easily checked, too, except for the unitarity of the representation (which follows from the skew-adjointness of the Lie algebra elements but should better be proved directly).

We now pass to $SO(3)$, The Lie algebra $\mathfrak{so}(3)$ is isomorphic to $\mathfrak{su}(2)$ via the map

$$\rho : \mathfrak{so}(3) \to \mathfrak{su}(2) \tag{3.188}$$

defined by linear extension of

$$\rho(J_k) = -\tfrac{1}{2} i\sigma_k, \tag{3.189}$$

where the (J_1, J_2, J_3) form a basis of $\mathfrak{so}(3) \cong \mathbb{R}^3$, see (1.47). The only difference between $SU(2)$ and $SO(3)$ therefore lies in their global properties, which were already clarified in Proposition 3.1 in §3.1: in conclusion, $SU(2)$ is a double covering of $SO(3)$, and since $SU(2)$ is connected and simply connected, it is in fact the universal covering space of $SO(3)$. Proposition 3.1 shows that if $u : SU(2) \to U(H)$ is a unitary representation of $SU(2)$ for which

$$u(\mathrm{diag}(-1,-1)) = 1_H, \tag{3.190}$$

and $\tilde{s} : SO(3) \to SU(2)$ is any cross-section of the projection $\tilde{\pi} : SU(2) \to SO(3)$

$$\tilde{\pi} \circ \tilde{s} = \mathrm{id}_{SO(3)}, \tag{3.191}$$

then $\tilde{u} = u \circ \tilde{s}$ is a unitary representation of $SO(3)$. Indeed, the desired property $\tilde{u}(R_1)\tilde{u}(R_2) = \tilde{u}(R_1R_2)$ is equivalent to the property

$$u(\tilde{s}(R_1)\tilde{s}(R_2)\tilde{s}(R_1R_2)^{-1}) = 1_H; \tag{3.192}$$

applying the homomorphism $\tilde{\pi} : SU(2) \to SO(3)$ to the argument of u and recalling (3.191) gives $\tilde{\pi}(\tilde{s}(R_1)\tilde{s}(R_2)\tilde{s}(R_1R_2)^{-1}) = e$, i.e., the unit of $SO(3)$, so that $\tilde{s}(R_1)\tilde{s}(R_2)\tilde{s}(R_1R_2)^{-1} \in \ker(\tilde{\pi})$, which is the center of $SU(2)$ consisting of $\pm 1_2$. Consequently, (3.190) with $u(\mathrm{diag}(1,1)) = 1_H$ implies (3.192).

In fact, all unitary representations of $SO(3)$ arise in this way: given some such representation $\tilde{u}$, pass to the corresponding representation $d\tilde{u}$ of $\mathfrak{so}(3) \cong \mathfrak{su}(2)$, integrate the latter to a unitary representation of $SU(2)$, which by (3.190) quotients to a representation of $SO(3)$, which must be the given one.

Moreover, up to unitary equivalence, one also obtains all projective unitary representations of $SO(3)$ in a similar way: given such a representation $u : SO(3) \to U(H)$, follow the procedure described at the end of the preceding section, i.e., pass to the corresponding representation du' of $\mathfrak{su}(2)$, and integrate the latter to a unitary representation u' of $SU(2)$. Then

$$\tilde{u} = u' \circ \tilde{s} : SO(3) \to U(H) \tag{3.193}$$

is equivalent to the given projective representation u, where the (correct) notion of equivalence of two projective unitary representations $\tilde{u}$ used here is as follows:

$$\tilde{u}_1 \sim \tilde{u}_2 \text{ iff } \pi \circ \tilde{u}_1 = \pi \circ \tilde{u}_2, \tag{3.194}$$

that is, $\tilde{u}_1 \sim \tilde{u}_2$ iff $\tilde{u}_1$ and $\tilde{u}_2$ induce the same map $h : G \to U(H)/\mathbb{T}$, cf. (3.55).

Finally, the representation D_j defined by (3.183) satisfies (3.190) ff $j \in \mathbb{N}$. Therefore, we now have a complete list of all unitary irreducible representations of $SO(3)$: these are labeled by $j \in \mathbb{N}$ and correspond to the $D_j(SU(2))$ in the said way.

Exercises for week 13 (inleveropgaven: nos. 2 and 3)

1. Verify that the operators (3.178) - (3.180) satisfy (3.174) - (3.175).
2. Finish the proof of Theorem 3.12. Hints: The ***Casimir operator***

$$C = L_1^2 + L_2^2 + L_3^2, \tag{3.195}$$

commutes with each L_k. By Schur's lemma, in any irreducible representation we therefore must have

$$C = c \cdot 1_H, \tag{3.196}$$

where $c \in \mathbb{R}$ (in fact, $c \geq 0$). You may also use the additional algebraic relations

$$L_+L_- = C - L_3(L_3 - 1_H); \tag{3.197}$$
$$L_-L_+ = C - L_3(L_3 + 1_H). \tag{3.198}$$

a. With c defined as in (3.196), use the above relations to show that

$$c - (\lambda_0 + k)(\lambda_0 + k + 1) = 0; \tag{3.199}$$
$$c - \lambda_0(\lambda_0 - 1) = 0, \tag{3.200}$$

and infer that these relations imply $\lambda_0 = -k/2$

b. Consider $H' \subseteq H$ as in the main text and show that π is irreducible iff $H' = H$.

c. Show that the operators in (3.178) - (3.180) satisfy $L_3^* = L_3$ and $L_\pm^* = L_\mp$.

3. Diagonalize the harmonic oscillator Hamiltonian $h = \frac{1}{2}(-d^2/dx^2 + x^2)$ on $H = L^2(\mathbb{R})$ in a similar way (not worrying about unbounded operators):

a. Define $L_\pm = x \mp d/dx$, and show that $[h, L_\pm] = \pm L_\pm$ and $h = \frac{1}{2}(L_+L_- + 1_H)$.

b. Show that h is positive (i.e. $\langle \psi, h\psi \rangle \geq 0$ for all $\psi \in D(h)$), and show from this that any eigenvalue of h must be positive (or zero).

c. Prove that there is a unique $\psi_0 \in L^2(\mathbb{R})$ (up to normalization and a phase) such that $L_-\psi_0 = 0$ (and give it!).

d. Show that the eigenvalues of h are $\{n + \frac{1}{2}, n \in \mathbb{N}\}$, each with multiplicity one.[12]

[12] It can be shown that the spectrum of h is discrete, so you have found the entire spectrum and the corresponding eigenvectors form a basis of $L^2(\mathbb{R})$.

3.9 Unitary irreducible representations of semi-direct products

Another case of great importance to quantum mechanics is the representation theory of regular ***semi-direct products*** $L \ltimes V$, such as the Poincaré-group, which plays a central role in relativistic quantum theory (including quantum field theory).

Let some group L act linearly on a vector space V, seen as an abelian group under addition, or more generally, let L act on some abelian group V by automorphisms (i.e., $\lambda \cdot (v+w) = \lambda \cdot v + \lambda \cdot w$), where we write the group operation in V as +). Then

$$(\lambda, v) \cdot (\lambda', v') = (\lambda\lambda', v + \lambda \cdot v'); \tag{3.201}$$
$$(\lambda, v)^{-1} = (\lambda^{-1}, -\lambda^{-1} \cdot v), \tag{3.202}$$

turn $L \ltimes V$ into a group, called the ***semi-direct product*** of L and V. In what follows, the *dual* L-action on $\hat{V}$ plays a role, where the ***dual group***

$$\hat{V} = \mathrm{Hom}(V, \mathbb{T}), \tag{3.203}$$

of V is the set of (continuous) homomorphisms $\chi : V \to \mathbb{T}$, seen as a (topological) group under (the compact-open topology and) pointwise multiplication

$$(\chi_1 \cdot \chi_2)(v) = \chi_1(v)\chi_2(v). \tag{3.204}$$

For example, previous exercises give

$$\hat{\mathbb{R}}^n \cong \mathbb{R}^n, \quad \chi_p(x) = e^{ipx}; \tag{3.205}$$
$$\hat{\mathbb{Z}} \cong \mathbb{T}, \quad \chi_z(n) = z^n; \tag{3.206}$$
$$\hat{\mathbb{T}} \cong \mathbb{Z}, \quad \chi_n(z) = z^n, \tag{3.207}$$

where in the first line px is the inner product $\langle p, x \rangle$ in $\mathbb{R}^n$. For (real) vector spaces V we therefore have $\hat{V} \cong V^*$, where $V^* = \mathrm{Hom}(V, \mathbb{R})$ is the linear dual of V, in that $\chi = \chi_p \in \hat{V}$, defined in (3.205), is identified with $p \in V^*$. Since we also have $V^* \cong V$ upon the choice of an inner product, we end up with an isomorphism $\hat{V} \cong V$ for real vector spaces (but one should be careful and diffident with this identification).

The ***dual action*** λ^* of $\lambda \in L$ on $\hat{V}$ is then given, for $\chi \in \hat{V}$ and $v \in V$, by

$$(\lambda^* \cdot \chi)(v) = \chi(\lambda^{-1} \cdot v). \tag{3.208}$$

If $V = \mathbb{R}^n$ and $L \subset GL_n(\mathbb{R})$, upon identifying $\hat{V} \cong V$ as above, this comes down to $\lambda^* p = (\lambda^{-1})^T p$, which in general differs from λ itself. This difference does matter for the Poincaré-group, where $V = \mathbb{R}^4$ and $L = SO(3,1)$. However, if $L = SO(n)$ and $V = \mathbb{R}^n$, then, exceptionally, $\lambda^* = (\lambda^{-1})^T = \lambda$, since $\lambda^{-1} = \lambda^T$ for all $\lambda \in SO(n)$.

We say that the semi-direct product $L \ltimes V$ is ***regular*** if the L-action on V^* is regular, in a sense that may be stated in a number of equivalent ways (whose equivalence is not at all obvious, see exercises). Perhaps the simplest condition is:[13]

[13] A space is T_0 if for any two distinct points there is an open set that contains exactly one of them.

Definition 3.7. *A continuous action of a topological group G on a space X is called* **regular** *if the quotient space X/G of G-orbits in X is T_0.*

Using the definition of the quotient topology on X/G (which is the finest topology for which the projection $\pi : X \to X/G$ is continuous), this condition is equivalent to: for each $x \in X$ and $y \in X$ such that $y \notin Gx$ (so that $[x] \neq [y]$ in X/G) there is an open $U \subset X$ such that either $Gx \subseteq GU$ but $Gy \nsubseteq GU$, or $Gy \subset GU$ but $Gx \nsubseteq GU$. Most of our examples, like $SO(n)$ acting on $\mathbb{R}^n$, and also $SO(3,1)$ acting on $\mathbb{R}^4$, are regular. The simplest example of a *non-regular* action is the action $\mathbb{Z} \circlearrowleft \mathbb{T}$ given by

$$n : z \mapsto e^{2\pi i n\theta} z \quad (\theta \in \mathbb{R}\backslash\mathbb{Q}). \tag{3.209}$$

Indeed, each $\mathbb{Z}$-orbit in $\mathbb{T}$ is dense, so that for any open $U \subset \mathbb{T}$ we have $GU = \mathbb{T}$. However, the $\mathbb{Z}$-action does not give automorphisms of $\mathbb{T}$ (as a group), so for a non-regular semi-direct product we need a slightly different example (see exercises).

For any subgroup $H \subset G$, we denote the equivalence class of x in G/H by $[x]$. Furthermore, each $x \in X$ defines a G-orbit through x denoted by $G \cdot x$, as well as a stabilizer (or "little group")

$$G_x = \{g \in G \mid g \cdot x = x\}. \tag{3.210}$$

We then have $G \cdot x \cong G/G_x$ under the map $g \cdot x \mapsto [g]$ (which is a homeomorphism if the G-action is regular, see exercises). We apply these concepts to $G = L$ acting on $X = \hat{V}$, just in the case where $\hat{V} \cong V^*$. In what follows, any L-orbit $\mathscr{O} \subset V^*$ may be equipped with a ***quasi-invariant measure*** μ, i.e., $\mu(A) = 0$ implies $\mu(\lambda \cdot A) = 0$ for any $\lambda \in L$ (for any measurable $A \subset \mathscr{O}$), as well as $\mu(U) > 0$ for any open $U \subset \mathscr{O}$, and any such measure may be used to define $L^2(\mathscr{O})$. For simplicity (and because this is enough for our applications), we state the formulae only in the more special case that μ is ***invariant***, i.e., $\mu(\lambda \cdot A) = \mu(A)$ for all $\lambda \in L$. Furthermore, we pick a (measurable) cross-section $s : \mathscr{O} \to L$, i.e., a *right* inverse to the canonical projection $\pi : L \to \mathscr{O}$. Realizing $\mathscr{O}$ as $\mathscr{O} = L \cdot \theta_0$ for some $\theta_0 \in \mathscr{O}$ with stabilizer $L_{\theta_0} \equiv L_0$, we have $\pi(\lambda) = \lambda\theta_0$, whereas using the isomorphism $\mathscr{O} \cong L/L_0$ under $\lambda\theta_0 \mapsto [\lambda]$, as above, we have $\pi : L \to L/L_0$ defined as $\pi(\lambda) = [\lambda]$. Then s has to satisfy

$$\pi \circ s = \mathrm{id}_{\mathscr{O}}. \tag{3.211}$$

It may not be possible to make s continuous, and also, s is not a *left* inverse to π; instead, there exists a unique function $h_s : L \to L_0$ such that $s \circ \pi(\lambda) = \lambda h_s(\lambda)$, i.e.,

$$h_s(\lambda) = \lambda^{-1} s([\lambda]) = \lambda^{-1} s \circ \pi(\lambda). \tag{3.212}$$

Theorem 3.13. *Up to unitary equivalence, the irreducible unitary representations of a regular semi-direct product $L \ltimes V$ are classified by pairs $(\mathscr{O}, \chi)$, where $\mathscr{O}$ is an L-orbit in V^* and χ labels the unitary irreducible representations of the stabilizer $L_0 \subset L$ of an arbitrary point $\theta_0 \in \mathscr{O}$, up to unitary equivalence. In what follows, u_χ is some unitary irreducible representation labeled by χ, defined on a Hilbert space H_χ. The representation $\tilde{u}^{(\mathscr{O},\chi)}(L \ltimes V)$ may then be realized on the Hilbert space*

$$\tilde{H}^{(\mathscr{O},\chi)} = L^2(\mathscr{O}) \otimes H_\chi, \tag{3.213}$$

and is explicitly given, for $(\lambda, v) \in L \ltimes V$*, by*

$$\tilde{u}^{(\mathscr{O},\chi)}(\lambda, v)\tilde{\psi}(\theta) = e^{i\theta(v)} u_\chi(s(\theta)^{-1}\lambda s(\lambda^{-1}\theta))\tilde{\psi}(\lambda^{-1}\theta). \tag{3.214}$$

Here we regard $\tilde{\psi} \in L^2(\mathscr{O}) \otimes H_\chi$ as a function $\tilde{\psi} : \mathscr{O} \to H_\chi$, so that in (3.214) we have $\tilde{\psi}(\theta) \in H_\chi$, on which the operator $u_\chi(s(\theta)^{-1}\lambda s(\lambda^{-1}\theta))$ on H_χ acts. Note that the ***Wigner cocycle*** $s(\theta)^{-1}\lambda s(\lambda^{-1}\theta)$ lies in L_0 (exercise). It is a nontrivial matter to show that each operator $\tilde{u}^{(\mathscr{O},\chi)}(\lambda, v)$ is unitary on $\tilde{H}^{(\mathscr{O},\chi)}$ and that one indeed has a representation. To the latter effect, it is useful to know the following: if one has (unitary) representations $u_1(V)$ and $u_2(L)$ on the same Hilbert space H that satisfy

$$u_2(\lambda)u_1(v)u_2(\lambda^{-1}) = u_1(\lambda \cdot v), \tag{3.215}$$

then one obtains a (unitary) representation u of $L \ltimes V$ by means of

$$u(\lambda, v) = u_1(v)u_2(\lambda). \tag{3.216}$$

Conversely, a (unitary) representation u of $L \ltimes V$ on H gives rise to (unitary) representations $u_1(V)$ and $u_2(L)$ by restriction, i.e., $u_1(v) = u(1, v)$ and $u_2(\lambda) = u(\lambda, 0)$, and this pair satisfies (3.216).

We briefly discuss two basic examples, each of which is easily seen to be regular. We write $p \in V^*$ (instead of θ), which is seen as a momentum variable.

- $G = E(2) = SO(2) \ltimes \mathbb{R}^2$, defined like $E(3)$, i.e., with respect to the usual action of $SO(2)$ on $\mathbb{R}^2$ (this group will also play a surprising role in the representation theory of the Poincaré-group). We find the same action of $SO(2)$ on $(\mathbb{R}^2)^* = \mathbb{R}^2$, so that the orbits are $\mathscr{O}_0 = \{0\}$ with $L_0 = SO(2)$ and

$$\mathscr{O}_r = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = r^2\} \tag{3.217}$$

for $r > 0$, with $L_0 = \{e\}$. Thus the Hilbert spaces and representations are

$$\tilde{H}^{(0,n)} = \mathbb{C}; \tag{3.218}$$
$$\tilde{u}^{(0,n)}(\lambda, v) = e^{2\pi i n\lambda}; \tag{3.219}$$
$$\tilde{H}^r = L^2(0,1); \tag{3.220}$$
$$\tilde{u}^r(\lambda, v)\tilde{\psi}(p) = e^{ir(v_1 \cos p' + v_2 \sin p')}\psi(p - \lambda|\text{mod}\, 1), \tag{3.221}$$

where $n \in \mathbb{Z}$, $\lambda \in [0, 1)$, $p \in (0, 1)$, and $p' = 2\pi p$. In the first case $\mathbb{R}^2 \subset E(2)$ is represented trivially, whereas in the second the r-dependence of the representation lies entirely in $\mathbb{R}^2$ (since $\tilde{H}^r$ and $\tilde{u}^r(\lambda, 0)$ are evidently independent of r).
- $G = E(3) = SO(3) \ltimes \mathbb{R}^3$, as before with the defining action of $SO(3)$. The $SO(3)$-orbits in $(\mathbb{R}^3)^* = \mathbb{R}^3$ are spheres $S^2_r \cong SO(3)/SO(2)$ with radius $r > 0$, as well as the origin ($r = 0$) with stabilizer $SO(3)$, so that for the Hilbert spaces we obtain

$$\tilde{H}^{(0,j)} = \mathbb{C}^{2j+1}; \tag{3.222}$$
$$\tilde{H}^{(r,n)} = L^2(S^2); \tag{3.223}$$

where $j = 0, 1, \ldots$ labels the unitary irreducible representations of $SO(3)$ on $H_j = \mathbb{C}^{2j+1}$, whereas $n \in \mathbb{Z}$ labels the irreducible representations of $SO(2)$ on $\mathbb{C}$ (we write $S^2 \equiv S^2_1$). In the second case, the representation $u^{(r,n)}$ of $SO(3) \subset E(3)$ depends explicitly on n through the Wigner cocycle; for $n = 0$ we simply obtain

$$\tilde{u}^{(r,0)}(R, v)\tilde{\psi}(p) = e^{irp\cdot v}\tilde{\psi}(R^{-1}p). \tag{3.224}$$

Mathematicians prefer a different realization of the above representation $\tilde{u}^{(\mathscr{O},\chi)}$. Consider (measurable) functions $\psi : L \to H_\chi$ that satisfy the constraint

$$\psi(\lambda s) = u_\chi(s^{-1})\psi(\lambda), \tag{3.225}$$

for (almost) every $\lambda \in L$ and $s \in L_0$. Now if ψ and φ both satisfy (3.225), then by unitarity of u_χ their inner product $\langle\varphi(\lambda), \psi(\lambda)\rangle_{H_\chi}$ in H_χ is L_0-invariant, in that

$$\langle\varphi(\lambda s), \psi(\lambda s)\rangle_{H_\chi} = \langle\varphi(\lambda), \psi(\lambda)\rangle_{H_\chi} \ \ (s \in L_0). \tag{3.226}$$

Hence the function $\lambda \mapsto \langle\varphi(\lambda), \psi(\lambda)\rangle_{H_\chi}$, *a priori* defined from G to $\mathbb{C}$, induces a function $[\lambda] \mapsto \langle\varphi(\lambda), \psi(\lambda)\rangle_{H_\chi}$ from L/L_0 to $\mathbb{C}$. We write the latter function as $\langle\varphi, \psi\rangle_{H_\chi}[\lambda]$; in particular, taking $\varphi = \psi$, we write $\|\psi\|^2_{H_\chi}[\lambda] = \langle\psi(\lambda), \psi(\lambda)\rangle_{H_\chi}$. We may then define a new Hilbert space $H^{(\mathscr{O},\chi)}$ that consists of all measurable functions $\psi : L \to H_\chi$ that for each $s \in L_0$ satisfy (3.225) and

$$\int_{L/L_0} d\mu([\lambda])\, \|\psi\|^2_{H_\chi}[\lambda] < \infty. \tag{3.227}$$

This space turns out to be complete in the natural inner product

$$\langle\varphi, \psi\rangle = \int_{L/L_0} d\mu([\lambda])\, \langle\varphi, \psi\rangle_{H_\chi}[\lambda] \tag{3.228}$$

A cross-section $s : L/L_0 \to L$ as above then gives rise to a unitary isomorphism

$$w_s : H^{(\mathscr{O},\chi)} \to \tilde{H}^{(\mathscr{O},\chi)}; \tag{3.229}$$
$$w_s\psi(\theta) = \psi(s(\theta)); \tag{3.230}$$
$$w_s^{-1}\tilde{\psi}(x) = u_\chi(h_s(x))\tilde{\psi}([x]), \tag{3.231}$$

from which the corresponding representation $u^{(\mathscr{O},\chi)}(\lambda, v) = w_s^{-1}\tilde{u}^{(\mathscr{O},\chi)}(\lambda, v)w_s$ on our mathematicians' Hilbert space $H^{(\mathscr{O},\chi)}$ can be computed (exercise).

Exercises for week 14 (inleveropgaven: nos. 2, 5)

1. *Only for fanatic topologists. Not compulsory even as oefenopgave.* Show that the following conditions (on a continuous G-action on a space X are equivalent:[14]
   a. The the quotient space X/G of G-orbits in X is T_0;
   b. Each G-orbit in X is relatively open in its closure;
   c. Each G-orbit in X is locally closed;
   d. Each map $[g] \mapsto gx$ is a homeomorphism from G/G_x to the orbit $G \cdot x$ $(x \in X)$.
2. Let $L = \mathbb{Z}$ and $V = \mathbb{C}$ as an abelian group under addition, where the $\mathbb{Z}$-action on $\mathbb{C}$ is given by (3.209). Compute the corresponding L-action on $\hat{V}$ and show that the semi-direct product $\mathbb{Z} \ltimes \mathbb{C}$ is not regular.
3. Show that $s(\theta)^{-1}\lambda s(\lambda^{-1}\theta) \in L_0$.
4. Show that (3.214) defines a unitary representation of $L \ltimes V$.
5. Show that w_s and w_s^{-1} are unitary and compute $u^{(\mathscr{O},\chi)}(\lambda, v) = w_s^{-1}\tilde{u}^{(\mathscr{O},\chi)}(\lambda, v)w_s$.

[14] Let X be a space. One calls $Y \subset Y' \subseteq X$ ***relatively open*** in Y' if there is an open set $U \subset X$ such that $Y = Y' \cap U$. A subset $Y \subset X$ is ***locally closed*** if each $y \in Y$ has an open neighbourhood U in X such that $U \cap Y$ is closed.

3.10 Unitary irreducible representations of the Poincaré group

The Poincaré group is the symmetry group of relativistic physics. We first define

$$\langle x, y\rangle_M = x^0 y^0 - \mathbf{x}\cdot\mathbf{y}; \tag{3.232}$$
$$O(3,1) = \{\lambda \in GL_4(\mathbb{R}) \mid \langle \lambda x, \lambda y\rangle_M = \langle x, y\rangle_M \, \forall x, y \in \mathbb{R}^4\}, \tag{3.233}$$

where we write $x = (x^0, x^1, x^2, x^3) = (x^0, \mathbf{x})$. Note that the ***Minkowski inner product*** $\langle\cdot,\cdot,\rangle_M$ defined by (3.232) is not really an inner product (since it fails to be positive definite), but an indefinite non-degenerate bilinear form. Unlike $O(4)$, which consists of matrices leaving the usual inner product on $\mathbb{R}^4$ invariant, $O(3,1)$ is not compact. It has four connected components, which may be identified by the (independent) conditions $\det(\lambda) = \pm 1$ and $\pm\lambda^0{}_0 > 0$. For simplicity we restrict ourselves to the connected component L of the identity, in which $\det(\lambda) = 1$ and $\lambda^0{}_0 > 0$. See exercises. This group is called the ***(proper orthochronous) Lorentz group***.

For our later study of projective representations, it is interesting to note that we have a surjective homomorphism

$$\tilde{\pi} : SL(2,\mathbb{C}) \to L, \tag{3.234}$$

with kernel $\mathbb{Z}_2 = \{\pm 1_2\}$, which, noting that $SU(2) \subset SL(2,\mathbb{C})$ and $SO(3) \subset L$ (as matrices with unit entries on the left and the upper outer edges), restricts to

$$\tilde{\pi} : SU(2) \to SO(3). \tag{3.235}$$

Both are double covering maps, and in fact $SL(2,\mathbb{C})$ and $SU(2)$ are the universal covering spaces of L and $SO(3)$, respectively. See also Proposition 3.1, in which context (3.20) may be obtained from (3.234) by restriction to $U(2) \subset SL(2,\mathbb{C})$, so that in turn (3.235) equivalently follows from (3.20) by restriction to $SU(2) \subset U(2)$. To see this, we again start from the four matrices $(\sigma_0, \sigma_1, \sigma_2, \sigma_3)$ in (3.16), and note:

- These form a basis for the (real) vector space of all self-adjoint 2×2 matrices;
- For any $x \in \mathbb{R}^4$ we have $\det(\sum_{\mu=0}^3 x_\mu \sigma_\mu) = \langle x, a\rangle_M = (x^0)^2 - \|\mathbf{x}\|^2$;
- For any $\tilde{\lambda} \in SL(2,\mathbb{C})$ and $a \in M_2(\mathbb{C})$ we have $\det(\tilde{\lambda} a \tilde{\lambda}^*) = \det(a)$;
- For any $\tilde{\lambda} \in SL(2,\mathbb{C})$ and self-adjoint $a \in M_2(\mathbb{C})$, $\tilde{\lambda} a \tilde{\lambda}^*$ is again self-adjoint.

Taking $a = \sum_\mu x_\mu \sigma_\mu$, it follows that for $\tilde{\lambda} \in SL(2,\mathbb{C})$ and $x \in \mathbb{R}^4$ there must be $\lambda \in O(3,1)$ such that $\tilde{\lambda} \sum_\mu x_\mu \sigma_\mu \tilde{\lambda}^* = \sum_\mu (\lambda\cdot x)_\mu \sigma_\mu$. By continuity and the fact that $SL(2,\mathbb{C})$ is connected it follows that in fact $\lambda \in L$, so we put $\tilde{\pi}(\tilde{\lambda}) = \lambda$.

We now introduce the ***(proper orthochronous) Poincaré group***

$$P = L \ltimes \mathbb{R}^4, \tag{3.236}$$

defined with respect to the natural action of $L \subset O(3,1)$ on $\mathbb{R}^4$, as well as its cover

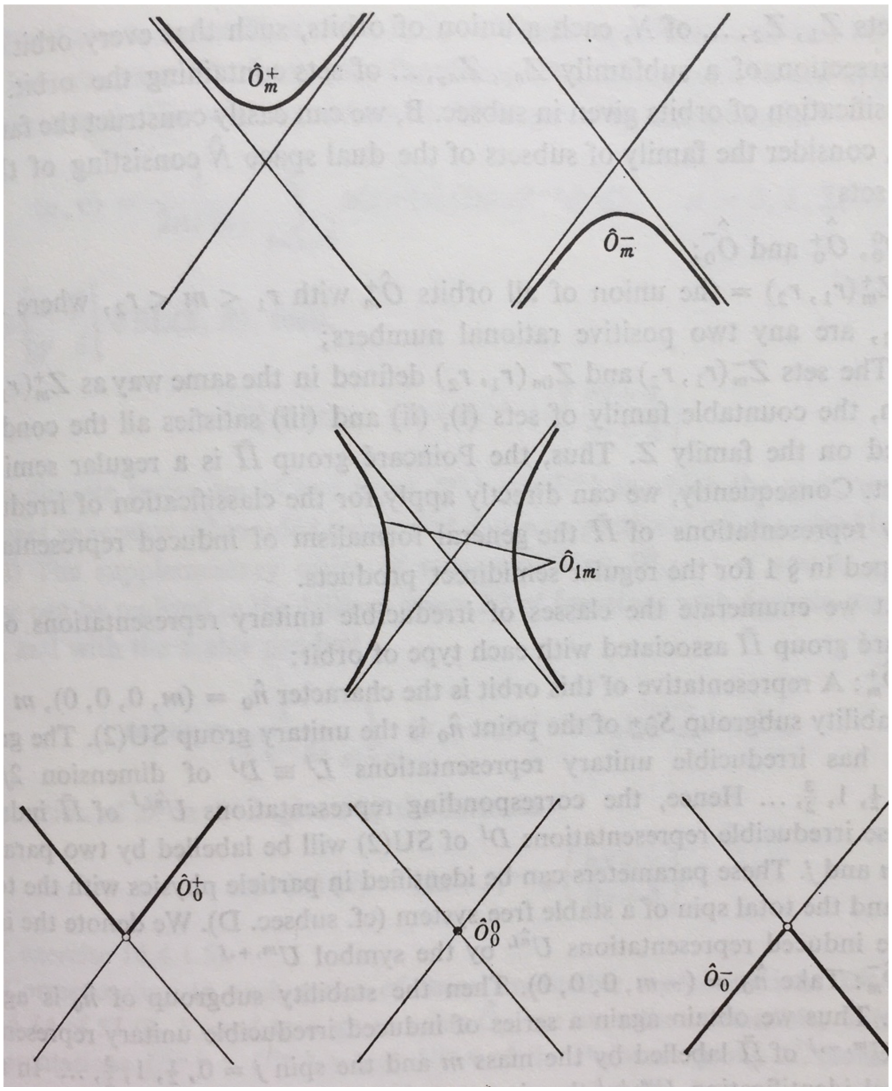
$$\tilde{P} = SL(2,\mathbb{C}) \ltimes \mathbb{R}^4, \tag{3.237}$$

where $SL(2,\mathbb{C}) \equiv \tilde{L}$ acts on $\mathbb{R}^4$ via the map $\tilde{\pi}$ in (3.234).

Writing $p^2 = p_0^2 - p_1^2 - p_2^2 - p_3^2$, the L-orbits $\mathscr{O}$ in $(\mathbb{R}^4)^* = \mathbb{R}^4$ are seen to be:

1. $\mathscr{O}_0^0 = \{(0,0,0,0)\}$, with stabilizer $L_0 = L$;
2. $\mathscr{O}_m^\pm = \{p \in \mathbb{R}^4 \mid p^2 = m^2, \pm p_0 > 0\}$, $m > 0$, with $L_0 = SO(3)$;
3. $\mathscr{O}_0^\pm = \{p \in \mathbb{R}^4 \mid p^2 = 0, \pm p_0 > 0\}$, with $L_0 \cong E(2) = SO(2) \ltimes \mathbb{R}^2$, see exercises;
4. $\mathscr{O}_{im} = \{p \in \mathbb{R}^4 \mid p^2 = -m^2, \pm p_0 > 0\}$, $m > 0$, with $L_0 = SO(2,1)$.

Here the stabilizers L_0 are found by taking the reference points $(\pm m, 0,0,0)$ in case 2, $(\pm 1, 0, 0, -1)$ in case 3, and $(0,0,0,m)$ in case 4. The physically relevant cases are (probably) $\mathscr{O}_m^+$ and $\mathscr{O}_0^+$, since $\mathscr{O}_{im}$ describes tachyons (which do not exist).



It can be shown that all projective irreducible representations of P come from irreducible representations of the covering $\tilde{P}$. This enlarges the stabilizers, as follows:

1. For $\mathscr{O}^+_{m^2}$ we now obtain $\tilde{L}_0 = SU(2)$, leading to a family of unitary irreducible representations $u^{m,j}$ labeled by ***mass*** $m > 0$ and ***spin*** $j = 0, \frac{1}{2}, 1, \ldots$.
2. For $\mathscr{O}^+_0$ the stabilizer $\tilde{L}_0$ of $(1,0,0,1)$ is a double cover $E(2)'$ of $E(2)$, whose unitary irreducible representations are labeled by either $(0,n)$ with $n \in \mathbb{Z}/2$ (called ***helicity***) or by $r > 0$. The latter case does not seem to occur in nature.

   On the one hand, this classification is a triumph of mathematical physics, but on the other hand, it fails to single out which cases actually occur in nature: as far as we know, these are spin $j = 0$ and $j = \frac{1}{2}$ and helicity $n = \pm 1$ and $n = \pm 2$ (so even small half-integral helicities, which in principle are allowed, do not occur).

   As the simplest example of a realization, we consider the massive case. Here $\mathscr{O}^+_m \cong \mathbb{R}^3$ under $(\omega_{\mathbf{p}}, \mathbf{p}) \leftrightarrow \mathbf{p}$, where $\omega_{\mathbf{p}} = \sqrt{\|\mathbf{p}\|^2 + m^2}$, and with this parametrization the Lorentz-invariant measure on $\mathscr{O}^+_m$ is $d^3\mathbf{p}/\omega_{\mathbf{p}}$ (rather than $d^3\mathbf{p}$, as might have been expected). For each $\mathbf{p} \in \mathbb{R}^3$ there is a unique boost $b_{\mathbf{p}} \in L$ that maps $(m,0,0,0)$ to $(\omega_{\mathbf{p}}, \mathbf{p})$, so we take $s(\mathbf{p}) = b_{\mathbf{p}}$. The Hilbert space (3.213) is

   $$\tilde{H}^{m,j} = L^2(\mathbb{R}^3) \otimes H_j, \tag{3.238}$$

   and for the corresponding unitary representation (3.214) of P we duly obtain

   $$\tilde{u}^{m,j}(\lambda), a^0\mathbf{a})\tilde{\psi}(\mathbf{p}) = e^{i(a^0\omega_{\mathbf{p}} - \langle \mathbf{a}, \mathbf{p} \rangle)} D_j(b^{-1}_{\mathbf{p}} \lambda b_{\lambda^{-1}\mathbf{p}}) \tilde{\psi}(\lambda^{-1}\mathbf{p}). \tag{3.239}$$

The ***Galilei group*** is the non-relativistic analogue of the Poincaré group. With $E(3) = SO(3) \ltimes \mathbb{R}^3$, acting on $\mathbb{R}^4$ by

$$(R, \mathbf{v}) : (a^0, \mathbf{a}) \mapsto (a^0, R\mathbf{a} + a^0\mathbf{v}), \tag{3.240}$$

the Galilei group is defined by

$$G = E(3) \ltimes \mathbb{R}^4. \tag{3.241}$$

Note that $\mathbf{v}$ is physically interpreted as a velocity, whereas earlier $\mathbf{a} \in \mathbb{R}^3 \subset E(3)$ was a position variable. This is clear from the natural G-action on $\mathbb{R}^4$, given by

$$(R, \mathbf{v}, a^0, \mathbf{a}) : (t, \mathbf{x}) \mapsto (t + a^0, R\mathbf{x} + \mathbf{a} + t\mathbf{v}), \tag{3.242}$$

which in fact determines the action (3.240). Either way, we obtain the group law

$$(R, \mathbf{v}, a^0, \mathbf{a}) \cdot (R', \mathbf{v}', (a^0)', \mathbf{a}') = (RR', \mathbf{v} + R\mathbf{v}', a^0 + (a^0)', \mathbf{a} + R\mathbf{a}' + (a^0)'\mathbf{v}). \tag{3.243}$$

We therefore see that the role of the Lorentz group L in the Poincaré group is now played by the Euclidean group $E(3)$. Since from (3.243) the inverse is found to be

$$(R, \mathbf{v}, a^0, \mathbf{a})^{-1} = (R^{-1}, -R^{-1}\mathbf{v}, -a^0, -R^{-1}(\mathbf{a} - a^0\mathbf{v})), \tag{3.244}$$

the dual $E(3)$-action on $(\mathbb{R}^4)^* \cong \mathbb{R}^4$ is given (in non-relativistic notation) by

$$(R,\mathbf{v}) : (E,\mathbf{p}) \mapsto (E - \langle \mathbf{v}, R\mathbf{p}\rangle, R\mathbf{p}). \tag{3.245}$$

In an exercise you determine the dual $E(3)$ orbits in $\mathbb{R}^4$, to find that none of the representations of G constructed from these orbits and their stabilizers reproduce some recognizable version of non-relativistic quantum mechanics.

Indeed we need to pass to projective representations of G; this is far more important for the Galilei group than it is for the Poincaré group. It can be shown that projective representations of G correspond to ordinary representations of the group

$$\check{G} = \widetilde{E(3)} \ltimes \mathbb{R}^5, \tag{3.246}$$

where $\widetilde{E(3)} = SU(2) \ltimes \mathbb{R}^3$, in which $SU(2)$ acts on $\mathbb{R}^3$ through the projection (3.235), that is, writing $\tilde{\pi}(u) \equiv R(u)$, we have $u \cdot \mathbf{x} = R(u)\mathbf{x}$. Furthermore, in order to define the semi-direct product (3.246) we need an action of $\widetilde{E(3)}$ on $\mathbb{R}^5$, namely

$$(u,\mathbf{v}) : (a^0,\mathbf{a},c) \mapsto (a^0, R(u)\mathbf{a} + a^0\mathbf{v}, c + \tfrac{1}{2}a^0\|\mathbf{v}\|^2 + \langle \mathbf{v}, R(u)\mathbf{a}\rangle). \tag{3.247}$$

Consequently, writing $\tilde{x} = (R,\mathbf{v},a^0,\mathbf{a})$, for the group law in $\check{G}$ we obtain

$$(\tilde{x},c)\cdot(\tilde{x}',c') = (\tilde{x}\cdot\tilde{x}', c + c' + \langle \mathbf{v}, R(u)\mathbf{a}'\rangle + \tfrac{1}{2}(a^0)'\|\mathbf{v}\|^2). \tag{3.248}$$

Eq. (3.247) implies the following dual $\widetilde{E(3)}$-action on $(\mathbb{R}^5)^* = \mathbb{R}^5$:

$$(u,\mathbf{v}) : (E,\mathbf{p},m) \mapsto (E - \langle \mathbf{v}, R(u)\mathbf{p}\rangle + \tfrac{1}{2}m\|\mathbf{v}\|^2, R(u)\mathbf{p} - m\mathbf{v}, m). \tag{3.249}$$

It is a nice exercise to find the corresponding orbits and stabilizers.

Exercises for week 15 (inleveren: 2 and 4)

1. Show that $O(3,1)$ has the four connected components explained in the main text. *Hint*: show that any matrix $\lambda \in O(3,1)$, where $(\lambda x)^\mu = \sum_{\nu=0}^{4} \lambda^\mu{}_\nu x^\nu$, satisfies

$$(\lambda^0{}_0)^2 - \sum_{k=1}^{3} (\lambda^k{}_0)^2 = 1, \tag{3.250}$$

so that $|\lambda^0{}_0| \geq 1$. Then note that both $\mathrm{sgn}(\lambda^0{}_0)$ and $\det(\lambda)$ are continuous functions on $O(3,1)$.
2. For $\mathscr{O}_0^{\pm} = \{p \in \mathbb{R}^4 \mid p^2 = 0, \pm p_0 > 0\}$, show that $L_0 \cong E(2)$.
3. Derive (3.245) and find the dual $E(3)$-orbits in $\mathbb{R}^4$ with their stabilizers.
4. Derive (3.249) and find the dual $\widetilde{E(3)}$-orbits in $\mathbb{R}^5$, as well as their stabilizers. Try to find out which orbits and unitary irreducible representations of the stabilizers correspond to known non-relativistic particles.